

THEORETICAL GEOPHYSICAL FLUID DYNAMICS

ENVIRONMENTAL FLUID MECHANICS

Volume 6

Managing Editor:

G. T. CSANADY, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

Editorial Board:

A. J. DAVENPORT, University of Western Ontario, London, Ontario

B. B. HICKS, Atmospheric Turbulence and Diffusion Laboratory, Oak Ridge, Tennessee

G. R. HILST, Electric Power Research Institute, Palo Alto, California

R. E. MUNN, IIASA, A-2361 Laxenburg, Austria

J. D. SMITH, University of Washington, Seattle, Washington

The titles in this series are listed at the end of this volume.

A. S. MONIN

*P. P. Shirshov Institute of Oceanology, Academy of Sciences of the U.S.S.R.,
Moscow*

Theoretical Geophysical Fluid Dynamics

Translated by Ron Hardin



KLUWER ACADEMIC PUBLISHERS
DORDRECHT / BOSTON / LONDON

Library of Congress Cataloging-in-Publication Data

CIP

Monin, A. S. (Andreĭ Sergeevich), 1921-
[Teoreticheskie osnovy geofizicheskoi gidrodinamiki. English]
Fundamentals of geophysical fluid dynamics / A.S. Monin ;
translated by Ron Hardin.
p. cm. -- (Environmental fluid mechanics)
Translation of: Teoreticheskie osnovy geofizicheskoi
gidrodinamiki.
Includes bibliographical references.
ISBN 0-7923-0426-8
1. Fluid dynamics. 2. Geophysics. I. Title. II. Series.
QC809.F5M6613 1990
550'.1'53205--dc20 89-36640
ISBN-13:978-94-010-7339-4 e-ISBN-13:978-94-009-1880-1
DOI: 10.1007/978-94-009-1880-1

*Originally published by Gidrometeoizdat, © 1988
under the title 'Teoreticheskie Osnovy Geofizicheskoi Gidrodinamiki'*

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Kluwer Academic Publishers incorporates
the publishing programmes of
D. Reidel, Martinus Nijhoff, Dr W. Junk and MTP Press.

Sold and distributed in the U.S.A. and Canada
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers Group,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

printed on acid free paper

All Rights Reserved

© 1990 by Kluwer Academic Publishers

No part of the material protected by this copyright notice may be reproduced or
utilized in any form or by any means, electronic or mechanical,
including photocopying, recording, or by any information storage and
retrieval system, without written permission from the copyright owner.

Table of Contents

AUTHOR'S PREFACE	ix
INTRODUCTION	xi
PART I: GENERAL CONCEPTS	
CHAPTER 1. EQUATIONS OF GEOPHYSICAL FLUID DYNAMICS	3
1. Thermodynamics	3
2. Stratification	7
3. Fluid Dynamics	9
4. Potential Vorticity	13
5. Energy	18
6. Hydrodynamic Systems	22
CHAPTER 2. SMALL OSCILLATIONS	33
7. Types of Adiabatic Motions	33
8. Small Oscillations of an Isothermal Atmosphere	43
9. Adaptation	51
10. Laplace Tidal Equation	55
CHAPTER 3. HYDRODYNAMIC INSTABILITY	65
11. Instability of an Ideal Fluid	65
12. Baroclinic Instability	73
13. Bifurcations	81
14. Instability of a Viscous Fluid	85
15. Stochasticity	93
16. Stochastization Scenarios	100
17. Taylor and Lorenz Flows	108
PART II: PROCESSES	
CHAPTER 4. SURFACE WAVES	119
18. Linear Theory	119
19. Ray Theory	123
20. Stationary Waves	128
21. Stationary Potential Waves	130

22. Stationary Gravity Waves	138
23. Stationary Capillary Waves	148
24. Hamiltonian Formalism	155
25. Wind Waves	160
CHAPTER 5. INTERNAL WAVES	165
26. Linear Theory	165
27. Ray Theory	172
28. Stationary Waves	180
29. Generation of Microstructure	185
30. Hamiltonian Formalism	188
31. Spectra	194
32. Inertial Oscillations	199
CHAPTER 6. GEOPHYSICAL TURBULENCE	202
33. Statistical Fluid Mechanics	202
34. Turbulence in Stratified Media	209
35. Vertical Thinly Laminated Microstructure	218
36. Boundary Layers	222
37. Geostrophic Turbulence	228
CHAPTER 7. ROSSBY WAVES	237
38. Linear Theory	237
39. Topographic Waves	242
40. Stationary Waves	246
41. Hamiltonian Formalism	261
42. Generation and Dissipation	269
PART III: GLOBAL PROBLEMS	
CHAPTER 8. GENERAL CIRCULATION OF THE ATMOSPHERE AND OCEAN	279
43. General Circulation of the Atmosphere	279
44. Similarity Theory for the Circulation of Planetary Atmospheres	285
45. Numerical Models of the General Circulation of the Atmosphere	290
46. Circulation of the World Ocean	297
47. Eddy-Resolving Models	306
CHAPTER 9. THEORY OF CLIMATE	312
48. Concept of Climate	312
49. Simplified Climate Models	315
50. Three-Dimensional Numerical Models	322
51. Modeling of Glacial Periods	334
CHAPTER 10. FLUID DYNAMICS OF PLANETARY INTERIORS	342
52. Fluid Dynamics of the Earth's Core	342

53. Hydromagnetic Dynamos	348
54. Evolution of Planetary Interiors	361
55. Mantle Convection	380
SELECTED REFERENCES	391
INDEX	397

Author's Preface

This book grew out of lectures on geophysical fluid dynamics delivered over many years at the Moscow Institute of Physics and Technology by the author (and, with regard to some parts of the book, by his colleagues). During these lectures the students were advised to read many books, and sometimes individual articles, in order to acquaint themselves with the necessary material, since there was no single book available which provided a sufficiently complete and systematic account (except, perhaps, the volumes on *Hydrophysics of the Ocean*, *Hydrodynamics of the Ocean*, and *Geodynamics* in the ten-volume *Oceanology* series published by Nauka Press in 1978–1979; these refer, however, specifically to the ocean, and anyway they are much too massive to be convenient for study by students). As far as we know, no text corresponding to our understanding of geophysical fluid dynamics has as yet been published outside the Soviet Union. The present book is designed to fill this gap.

Since it is customary to write the preface after the entire book has been completed, the author has an opportunity there to raise some points of possible criticism by the reviewers and readers. First of all, note that this work presents the *theoretical fundamentals* of geophysical fluid dynamics, and that observational and experimental data (which in the natural sciences are always very copious) are referred to only rarely and briefly. Any complete presentation of these data would have, according to my reckoning, doubled or even tripled the size of the book and the time needed to write it. Therefore, it had to be assumed that the readers either were somewhat acquainted with the empirical data or could be referred to the appropriate available sources.

However, even in just a theoretical account, a review of the pertinent problems and subjects is inevitably somewhat subjective; the authors will invariably describe topics that they know best and are more interested in. Usually these are the areas in which they have themselves worked. Accordingly, this book does not contain sections on the theory of tides (except for an explanation in Section 7 of the possibility of tide resonance), mesoscale phenomena (tropical hurricanes, tornados, oceanic interthermocline eddies, etc.), Rossby waves in hydrostatically unstable media, or coherent structures in turbulent flows. Moreover, only a general description of the hydrodynamic forecasting of weather is given, without any detailed discussion of its procedure, no applications of Lagrangian methods are given, and so forth. Some of these subjects (such as the theory of tides) are treated adequately in the literature, while for others (such as coherent structures) no theory has as yet been formulated.

In order to keep up, as much as possible, with the state of the art, the author has not shied away from new mathematical techniques, making use not only of differential equations, vector and tensor symbolism, and variational methods, but also, when necessary, of probability theory, ergodic theory, etc. Trivial calculations are not given, only their results being presented, together with any information needed to enable the readers to carry out all the proofs for themselves. Since it was not the author's intention to provide material for light reading, the book is replete with problems for students.

For practical reasons, only some of the many references cited are given in the reference list. The author has endeavored to explain everything necessary in the text, confining himself to priority indications in those ideas and results which seem to him (perhaps subjectively) the most important and interesting.

The notation could not, unfortunately, be kept uniform even in individual chapters. So many concepts were used (the newest of these are denoted by italics, to distinguish them) that no alphabets would have been sufficient for the notation. Therefore, in different places the same symbol may well stand for different physical quantities. In most places the notation is carefully explained (although theoretical physicists have become adept at guessing the meaning of a symbol without any specific explanation), and the author hopes that this duplication will not hamper understanding.

The author would like to express his gratitude to his old friends and ideological coauthors A. M. Obukhov and A. M. Yaglom, many of whose own works are reflected in this book, for their interest and their moral support. In addition, he would like to thank V. M. Kamenkovich, D. V. Chalikov, V. E. Zakharov, G. S. Golitsyn, V. P. Trubitsyn, and M. Ya. Marov for their help in selecting the scientific material, E. G. Agafonova for her great assistance in preparing the manuscript, G. Yu. Aleksandrova for doing the illustrations, and especially N. I. Solntseva for organizing the entire project of preparing the book for publication.

Introduction

Preferring to leave history to the historians of science, we will not present here a historical survey, but rather will only cite some examples of priorities during the course of the book. We will just note that separating geophysical fluid dynamics from general fluid dynamics was a risky step to take, since it could well have cut the field off from general science by creating a specialized jargon that was incomprehensible to the uninitiated. Pseudosciences can thrive and flourish on such soil over an area which is vast but which fails to reach any heights. This has occurred, for instance, in meteorology (here we do not include atmospheric physics), in which interest in astrological techniques for weather forecasting periodically rears its head (even now unverifiable forecasting methods ‘for the season’ are still used), in the agricultural sciences, in pedagogy, etc.

However, things are quite different if, in an important sphere of knowledge, there is a solid nugget of seriousness, requiring a special study of the material and the creation of new concepts (the creation of new concepts goes deeper than just the recognition of patterns; it is the very basis of scientific creativity, distinguishing the investigator from the computers of the foreseeable future), and if, basing itself on these concepts, the new science contacts and interacts with other branches of knowledge and becomes generally used.

However, the scientific community is not yet capable of consciously carrying out the separation or nonseparation of new sciences, as decided by authoritative groups of experts. For the present this is happening spontaneously, although, of course, the initiative and activity of major scientists (or, in the worst of cases, of major pseudoscientists) play an important role. In this spontaneous manner the field of geophysical fluid dynamics appeared on the scene as well, although the way was prepared for its separation from the other sciences in the early 1970s, when Joseph Smagorinsky organized the Geophysical Fluid Dynamics Laboratory at Princeton.

The characteristic properties of the objects of geophysical fluid dynamics (natural flows of liquids and gases) are rotation and stratification. In engineering fluid dynamics, on the other hand, these properties are in most cases unimportant, so that in courses on general fluid dynamics rotation and stratification are largely ignored.

The rotation affecting a natural flow is the rotation ω of the celestial body supporting the flow (planet, star, nebula, etc.). In a reference system moving with the supporting body, its rotation causes a Coriolis force $2\rho\omega \times \mathbf{u}$ acting on the

relative flows $\mathbf{u}$, and this force gives rise to a number of specific effects. For instance, in the Northern Hemisphere of the Earth, where the Coriolis force is directed toward the right of the flow, rivers scour their right banks. The trade winds near the Earth's surface, which are directed toward the thermal equator, acquire an easterly component, while the high antitrades, directed toward the pole, acquire a westerly component and produce a westerly transfer of air masses in the middle latitudes which overtakes the rotation of the Earth. By balancing the pressure differentials, the Coriolis force imparts geostrophicity to the flows. It sets up specific local oscillations in the ocean and atmosphere with an inertial frequency $f = 2\omega_z$, while variations of f with latitude (the β effect) produce Rossby waves, at the troughs and crests of which are formed the main generators of weather: cyclones and anticyclones. The gravitational effects of external bodies create tidal oscillations at rotating celestial bodies. Finally, in the presence of electrically conducting layers of liquid or gas, rotation generates the magnetic fields of celestial bodies.

The existence of an intrinsic gravitational field of a celestial body (with a free-fall acceleration g), together with the compressibility of all real gases and liquids, produces in them a stratification, primarily a quasihydrostatic increase in the fluid density ρ with depth. Moreover, the stratification is complicated by the fact that real gases and liquids are baroclinic, that is, their density depends not only on the pressure p (as in a barotropic medium) but also on the temperature T and the concentration s of thermodynamically active admixtures.

Baroclinicity has a dynamic significance, since it gives rise to an eddy-formation factor: a term $\text{curl}(-\rho^{-1} \nabla p) = \rho^{-2} \nabla \rho \times \nabla p$ in the expression for the vorticity of acceleration (in a barotropic fluid it is equal to zero). Even more important is the fact that fluctuations p' , T' , s' cause fluctuations of the density ρ' , and thus of the buoyancy force $\rho'g$ as well. For unstable stratification (in the absence of s , for an entropy increase with depth), they lead to the development of convection, at the expense of gravitational potential energy. For stable stratification (in the absence of s , for an entropy decrease with depth), on the other hand, kinetic energy is consumed as work against the buoyant forces, and instead of convection conditions are suitable for the development of internal gravity waves.

Consequently, geophysical fluid dynamics can be defined as *the fluid dynamics of natural flows of rotating baroclinic stratified fluids and gases, or, in short,*

$$\text{GFD} = \text{FD RBSF}.$$

Geophysical fluid dynamics poses problems for, and finds applications in, meteorology, oceanography, climatology, Earth physics, planetology, and astrophysics. It interrelates with magnetohydrodynamics, plasma physics, nonlinear mechanics (especially the theory of nonlinear waves), and, of course, in accordance with modern methodology, with computational mathematics.

PART I

GENERAL CONCEPTS

1. EQUATIONS OF GEOPHYSICAL FLUID DYNAMICS
2. SMALL OSCILLATIONS
3. HYDRODYNAMIC INSTABILITY

Equations of Geophysical Fluid Dynamics

1. THERMODYNAMICS

For greater generality let us assume that the medium being considered (atmospheric air, ocean water, a melt in the liquid outer layer of the Earth's core, etc.) is described by three independent thermodynamic variables which, in this section, will be taken to be the temperature T , the pressure p , and the mass concentration s of a *thermodynamically active admixture*. Concentration s may refer to the water vapor in unsaturated moist air (then s is the specific humidity) or to the salt content of ocean water (then s is the salinity). Just as in the case of s , in the following we will use only the *mass* thermodynamic characteristics of a medium (that is, the values per unit mass). As the fundamental function typifying the thermodynamic state of the medium, we take its *thermodynamic potential* (*free enthalpy*) $\varphi(T, p, s)$, the differential of which is defined by the formula

$$d\varphi = -\eta dT + \frac{1}{\rho} dp + \mu ds, \quad (1.1)$$

where the coefficients are the *entropy* $\eta = -\partial\varphi/\partial T$, the *specific volume* $1/\rho = \partial\varphi/\partial p$ (where ρ is the density of the medium), and the *chemical potential* $\mu = \partial\varphi/\partial s$. If these are known (as functions of T , p , and s), then we can also find the *free energy* $f = \varphi - p/\rho$, the *internal energy* $\varepsilon = \varphi - p/\rho + T\eta$, and the *enthalpy* $w = \varphi + T\eta$, so that any of the functions f , ε , or w can be selected instead of φ (by the way, later in the book these will as a rule not be used, and the notations f , ε , and w will be employed for completely different quantities). The differentials of the coefficients in (1.1) can be written as follows:

$$d\eta = \frac{c_p}{T} dT - \frac{\alpha}{\rho} dp + \delta ds, \quad (1.2)$$

$$d\rho = -\alpha\rho dT + \beta\rho dp + \gamma\rho ds, \quad (1.3)$$

$$d\mu = -\delta dT - \frac{\gamma}{\rho} dp + \lambda ds. \quad (1.4)$$

Here $c_p = T(\partial\eta/\partial T)_{p,s}$ is the *specific heat* at constant p and s , $\alpha = -(1/\rho)\partial\rho/\partial T$ is the coefficient of *thermal expansion*, $\beta = (1/\rho)\partial\rho/\partial p$ and $\gamma = (1/\rho)\partial\rho/\partial s$ are the coefficients of compressibility at the expense of p and at the expense of s , $\delta = \partial\eta/\partial s = -\partial\mu/\partial T$, and $\lambda = \partial\mu/\partial s$, where in (1.2)–(1.4) the

permutability of the second derivatives of the function $\varphi(T, p, s)$ is taken into account. Equation (1.3) is the differential form of the *equation of state* of the medium $\rho = \rho(T, p, s)$ (if it has the simplified form $\rho = \rho(p)$, that is, if the density of the medium is a function only of the pressure, then the medium is *barotropic*; otherwise it is *baroclinic*). Another important thermodynamic characteristic of a medium is the *velocity of sound*

$$c(T, p, s) = \left[\left(\frac{\partial p}{\partial \rho} \right)_{\eta, s} \right]^{1/2}, \quad (1.5)$$

where indexes η and s show that the derivatives are taken for constant η and s . If functions $\rho(T, p, s)$ and $c(T, p, s)$ are known, then we can also define the quantities

$$c_v = T \left(\frac{\partial \eta}{\partial T} \right)_{\rho, s} = \frac{1}{\kappa - 1} \frac{\alpha^2 T}{\beta \rho}; \quad \kappa = \frac{c_p}{c_v} = \beta \rho c^2, \quad (1.6)$$

where c_v is the specific heat at constant ρ and s (these formulas are readily derived from (1.2), (1.3)). Then all the coefficients in (1.2)–(1.4) are known except δ and λ (moreover, if we write differentials $d\delta$ and $d\lambda$ in a form similar to (1.2)–(1.4), taking into account the permutability of the second derivatives of functions η, ρ, μ , it is clear that for the determination of coefficients δ and λ , albeit accurate to the constant components, we have only to know derivatives $\partial\delta/\partial s$ and $\partial\lambda/\partial s$, and that as long as these are unknown the incompleteness of our knowledge of the thermodynamic characteristics of the medium is expressed by two arbitrary functions of s).

As an example, let us consider moist (unsaturated) atmospheric air. With an accuracy sufficient for geophysical fluid dynamics, it can be considered to be a mixture of ideal gases, water vapor being the thermodynamically active admixture, so that s is here the specific humidity. The equation of state for moist air has the form

$$p = R\rho T; \quad R = (1 - s)R_a + sR_v, \quad (1.7)$$

where $R_a = R_0/\mu_a \approx 287 \text{ J/kg K}$ and $R_v = R_0/\mu_v \approx 461 \text{ J/kg K}$ are the gas constants of dry air and water vapor ($R_0 \approx 8.314 \text{ J/(mole K)}$ is the universal gas constant, and $\mu_a \approx 29$ and $\mu_v \approx 18$ are the relative molecular masses of dry air and water vapor). From (1.7) it ensues that coefficient β in (1.3) equals $1/p$ and thus that, according to the second formula of (1.6), the velocity of sound is given by the relation

$$c^2 = \frac{\kappa p}{\rho} = \kappa R T, \quad (1.8)$$

where κ , generally speaking, depends on s (this dependence is readily derived with the aid of the following formulas (1.9) and (1.10)). The partial pressure of water vapor

$$p_v = s[s + \alpha_v(1 - s)]^{-1}p; \quad \alpha_v = \mu_v/\mu_a \approx 0.622. \quad (1.9)$$

The partial pressure of dry air is here $p_a = p - p_v$.

The entropy of moist air is given by the formulas:

$$\eta = (1 - s)\eta_a + s\eta_v; \quad (1.10)$$

$$\eta_a = c_{p_a} \ln T - R_a \ln p_a + \text{const}; \quad \eta_v = c_{p_v} \ln T - R_v \ln p_v + \text{const},$$

where $c_{p_a} \approx 1003 \text{ J/(kg K)}$ and $c_{p_v} \approx 1810 \text{ J/(kg K)}$ are the specific heats of dry air and water vapor at constant p and s . Generally s does not exceed 3 or 4% (in clouds and fogs the content of liquid water and ice is considerably lower than the vapor content, so that in many problems the same formulas (1.7) and (1.10) can be used approximately, but with s determined in them from (1.9) according to the pressure of saturated water vapor $p_{vm}(T)$). However, despite the smallness of s , it is important to take the moisture content into account, not only directly for calculating clouds, precipitation, and the very substantial contribution of the moisture content to the radiation transfer, but also, because of the sizable thermal effects of evaporation and condensation (and sometimes the sublimation of ice), as an effective heat-transfer agent. For instance, if we write the turbulent fluctuations of density and entropy (denoted by primes) approximately as

$$\frac{\rho'}{\rho} \approx -\frac{T'}{T} - (R_v/R_a - 1)s'; \quad \eta' \approx \frac{c_{p_a} T'}{T} + (\eta_v - \eta_a)s', \quad (1.11)$$

then the mean specific work of the buoyancy forces (equal to $g\rho$) and the turbulent entropy flux can be expressed by the formulas:

$$-\frac{g}{\rho} \overline{\rho' w'} \approx \frac{g}{c_{p_a} \rho T} (q_{\text{turb}} + b q_{\text{lat}}); \quad (1.12)$$

$$b = \left(\frac{R_v}{R_a} - 1 \right) \frac{c_{p_a} T}{\mathcal{L}} \approx 0.07,$$

$$\overline{\rho \eta' \mathbf{u}'} \approx \frac{1}{T} \left(\mathbf{q}_{\text{turb}} + \frac{\eta_v - \eta_a}{c_{p_a}} \frac{c_{p_a} T}{\mathcal{L}} \mathbf{q}_{\text{lat}} \right), \quad (1.13)$$

(Monin, 1967), where $\mathcal{L} \approx 2,500,000 \text{ J/kg}$ is the latent heat of vaporization; $\mathbf{u}'$ is a vector denoting the fluctuation of the wind velocity; w' is its vertical component; $\mathbf{q}_{\text{turb}} = c_{p_a} \rho \overline{T' \mathbf{u}'}$ is the turbulent heat flux; $\mathbf{q}_{\text{lat}} = \mathcal{L} \rho \overline{s' \mathbf{u}'}$ is the turbulent flux of latent heat. The contribution of $\mathbf{q}_{\text{lat}}$ to (1.12) and (1.13) may well be significant.

A second example is ocean water, in which the salinity s is the thermodynamically active admixture. It should be noted that an analytical calculation of the thermodynamic characteristics of a liquid is in general much more difficult, in comparison with a gas, since liquids are complicated by the presence in them of remote interactions between molecules (which have to be disrupted during evaporation). Moreover, water is a liquid which is in many respects anomalous, primarily because it possesses a *temperature of maximum density* $T = T_1(p, s) \approx 4^\circ\text{C}$ at $p < p_1 \approx 27 \times 10^4 \text{ mbar}$ and $s < s_1 \approx 24\text{‰}$, which is higher than the freezing point $T = T_2(p, s)$ (with a rise in p or s both T_1 and T_2 decrease, but T_1

drops more rapidly than T_2). Thus, the density of water, which always increases with a rise in p or s , for $T > T_1$ normally increases with a drop in T but in the interval $T_1 > T > T_2$ behaves anomalously and decreases, while at freezing this density decreases even more (as a result of which ice turns out to be lighter than liquid water and does not sink, but rather floats on the water surface, an especially important factor for the climate of the Earth). Finally, the salts dissolved in ocean water exist in the form of ions, between which additional Coulomb forces act as well as the ordinary molecular forces. In view of all these complications, the equation of state for ocean water can be obtained only by empirical means. It is customarily written as

$$\rho(T, p, s) = \rho_0(1 + 10^{-3}\sigma_t)[1 - \mu(p - p_a)]^{-1}, \quad (1.14)$$

where $p_a = 1013.25$ mbar is standard atmospheric pressure; $\rho_0 = \rho(4^\circ\text{C}, p_a, 0)$; $\sigma_t = 10^3[\rho(T, p_a, s)/\rho_0 - 1]$ is the so-called *relative density*; $\mu = [K_0 + A(p - p_a) + B(p - p_a)^2]^{-1} \sim 4 \times 10^{-8}$ to 5×10^{-8} mbar $^{-1}$ is the isothermal compressibility of water, where functions σ_t , K_0 , A , and B are described by polynomials in low powers of T and $(s - 35\text{‰})^{1/2}$. The velocity of sound (1.5) in ocean water is also studied only empirically. It increases with increasing temperature, pressure, and salinity, being described by the formula of Fry and Page (1971), in which to the standard value $c = 1449.30$ m/s is added a temperature correction in the form of a cubic polynomial in $T^\circ\text{C}$, as well as a pressure correction of fourth degree in p , a salinity correction of third degree in $(s - 35\text{‰})$, and a mixed correction of three terms in T^2p , Tp^2 , and $T(s - 35\text{‰})$. Coefficients δ and λ in (1.2) and (1.4) for ocean water have not yet been measured.

Considering a liquid or a gas as a continuous medium, we can nevertheless represent the fluid as an aggregate of *infinitesimal* (but, all the same, *macroscopic*, that is, containing a sufficiently large number of molecules to allow the use of the concepts of continuum mechanics: temperature, pressure, density, velocity, etc.) *individual* (retaining their individuality for some time, that is, not exchanging molecules with the surroundings much) *fluid particles*. These particles can be *labelled* (for instance, by their initial spatial coordinates), and their trajectories can be traced out (*Lagrangian* description of motion). The characteristics A of the fluid particles which do not vary with the time t (that is, which satisfy the equation $dA/dt = 0$, where d/dt is the individual derivative) are called the *Lagrangian invariants* of motion.

Now let us introduce the concept of a *thermodynamic process*, defined here as a variation with time of certain thermodynamic characteristics, caused by factors such as the molecular exchange of heat and thermodynamically active admixtures, phase transformations of substances in the medium, and so forth. Let us introduce, in particular, the following notation for the rates of change of quantities η and s in the fluid particles:

$$\frac{d\eta}{dt} = \varepsilon_\eta; \quad (1.15)$$

$$\frac{ds}{dt} = \varepsilon_s \quad (1.16)$$

Quantity ε_η is the rate of influx of heat to a unit mass of fluid, divided by the temperature. Processes for which $\varepsilon_\eta \equiv 0$ (so that the entropy is a Lagrangian invariant) are naturally referred to as *isentropic processes*. If $\varepsilon_s \equiv 0$ as well (so that the concentration of thermodynamically active admixtures is also a Lagrangian invariant), then the process is an *adiabatic process*. According to this definition, for adiabatic motions the trajectories of the fluid particles (and for stationary flows, the streamlines as well) will be isentropes on $s = \text{const}$ surfaces. From (1.5) it follows that for an adiabatic process

$$\frac{dp}{dt} = c^2 \frac{d\rho}{dt}. \quad (1.17)$$

This equation can sometimes conveniently be used instead of the condition $d\eta/dt = 0$.

2. STRATIFICATION

The stratification of a medium refers to the distribution of its density ρ *along the vertical*; that is, in the direction of gravity. We could also consider the stratification of the other thermodynamic characteristics, such as the temperature T , the concentration s of thermodynamically active admixtures, the sound velocity c , and the entropy η , but it is the density stratification which is of the greatest dynamic significance in geophysical fluid dynamics, since it directly determines the effect of the buoyancy forces specific to natural flows. In order to estimate the effect of these forces on fluid particles moving *adiabatically* along the vertical, let us for the present use η , s , and p as our independent thermodynamic variables and let us introduce the vertical coordinate z , reckoned, for instance, upward into the atmosphere from the surface of the equilibrium figure of the *geoid* $r = r_0(\theta, \lambda)$ (where r is the distance from the center of the Earth; $\theta = \pi/2 - \varphi$, the colatitude, is the complement of the latitude φ , and λ is the longitude), so that z is the height above sea level.

In a stratified medium parameters η , s , and p vary in general with the height z , so that the density stratification is described by the function $\rho(z) = \rho[\eta(z), s(z), p(z)]$. For adiabatic displacement of a fluid particle from a height z_1 , where $p = p_1$, its density at a height z will be $\rho_1(z) = \rho[\eta(z_1), s(z_1), p(z_1)]$. For a small displacement $z - z_1$ this density will differ from that of the surrounding medium by an amount $\rho_1(z) - \rho(z) \approx (\partial\rho_*/\partial z)_{z=z_1}(z_1 - z)$, where $\rho_*(\eta, s) = \rho(\eta, s, p_1)$ is the *potential density*, that is, the density reduced adiabatically to the pressure p_1 .

Obviously, for $\partial\rho_*/\partial z < 0$ the buoyancy force $g(\rho - \rho_1)$ will act on an adiabatically displaced fluid particle in a restoring manner, that is, the stratification will be *stable*. The vertical displacement of the fluid particle here requires the consumption of energy to work against the buoyancy force. If such a displacement $\xi = z - z_1$ is carried out, then, as long as no dissipative forces act, the particle will oscillate along the vertical relative to its equilibrium level z_1 with a frequency N . This frequency can be found by equating the acceleration of the particle $\partial^2\xi/\partial t^2$ to

the buoyancy acceleration $(g/\rho)(\partial\rho_*/\partial z)\zeta$, which gives

$$N^2 = - \left(\frac{g}{\rho} \right) \left(\frac{\partial\rho_*}{\partial z} \right). \quad (2.1)$$

This quantity is known as the *Brunt–Väisälä frequency*.

For $\partial\rho_*/\partial z > 0$, on the other hand, the buoyancy force will act on a moving fluid particle so as to accelerate it, that is, the stratification will be *unstable*. Moreover, if dissipative factors do not prevent it, convection will develop in the medium, and it will via the work of the buoyancy forces pick up kinetic energy at the expense of the potential energy of the unstable stratification.

Finally, for $\partial\rho_*/\partial z = 0$, the stratification will be *neutral*. If we set $dp = 0$ in formulas (1.2) and (1.3) and then eliminate dT from them, the differential of the potential density can be put into the form

$$d\rho_* = \frac{\rho_* \alpha_* T_*}{c_{p*}} \left[-d\eta + \left(\delta_* + \frac{\gamma_* c_{p*}}{\alpha_* T_*} \right) ds \right], \quad (2.2)$$

where asterisks below signify the values of the functions of η , s , and p at $p = p_1$. This shows that neutral stratification does not have to be adiabatic (that is, it does not have to satisfy the conditions $\partial\eta/\partial z = \partial s/\partial z = 0$); it can correspond to all possible stratifications $\eta(z)$ and $s(z)$ satisfying the condition

$$\frac{\partial\eta}{\partial z} = \left(\delta_* + \frac{\gamma_* c_{p*}}{\alpha_* T_*} \right) \frac{\partial s}{\partial z}.$$

This multiplicity of neutral stratifications manifests itself clearly in the World Ocean, where some very diverse combinations of profiles $T(z)$ and $s(z)$ are observed: the so-called *T-S curves* $z(s, T)$, examples of which are given in Figure 2.1. Formula (2.2) can be simplified considerably, if we substitute into it the entropy in the form $\eta(\rho, p, s)$; that is, if we determine $d\eta$ using formula (1.2), after eliminating dT from it with the aid of (1.3). Then, for $z = z_1$, using relations (1.6) we get

$$\frac{\partial\rho_*}{\partial z} = \frac{\partial\rho}{\partial z} - \frac{1}{c^2} \frac{\partial p}{\partial z}. \quad (2.3)$$

In most instances the variations of pressure with height are described accurately enough by the *hydrostatic equation* $\partial p/\partial z = -g\rho$, and then from (2.1) and (2.3) we arrive at the familiar formula

$$N^2 = - \left(\frac{g}{\rho} \right) \left(\frac{\partial\rho}{\partial z} + \frac{g\rho}{c^2} \right). \quad (2.4)$$

This relation implies that the condition for hydrostatic stability of the stratification $N^2 > 0$ leads to the inequality $\partial\rho/\partial z < -g\rho/c^2$. Consequently, just a decrease in density with height is not enough to ensure stability: the density must decrease with height *sufficiently rapidly*.

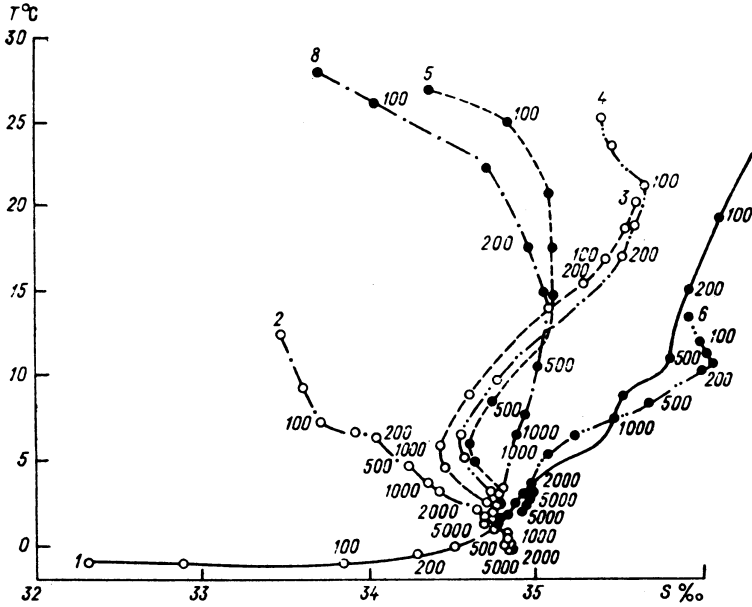


Fig. 2.1. Types of T-S curves in the World Ocean, according to Stepanov (1965): (1) polar, (2) subpolar, (3) middle tropical, (4) tropical, (5) equatorial, (6) North Atlantic, (7) Mediterranean, (8) Indo-Malayan.

Finally, let us give another expression for N^2 , obtained from (2.4) using formula (1.3), the hydrostatic equation, and relations (1.6):

$$N^2 = \alpha g \left(\frac{\partial T}{\partial z} + \gamma_a - \frac{\gamma}{\alpha} \frac{\partial s}{\partial z} \right); \quad \gamma_a = \frac{\kappa - 1}{\alpha} \frac{g}{c^2}, \quad (2.5)$$

where γ_a is the adiabatic temperature gradient (adiabatic lapse rate). For an atmosphere with equation of state (1.7) it takes the form $\gamma_a = (\kappa - 1)/\kappa \times g/R = g/c_p$, which gives for $c_p \approx 1003 \text{ J/(kg K)}$ and $g = 9.8 \text{ m/s}^2$ a value $\gamma_a \approx 9.8 \text{ K/km}$. In a *homogeneous* ($\partial T/\partial z = \partial s/\partial z = 0$) hydrostatic atmosphere we have $N = (\alpha g \gamma_a)^{1/2} = (\kappa - 1)^{1/2} g/c \approx 2 \times 10^{-2} \text{ s}^{-1}$; note, by the way, that in such an atmosphere $\partial \ln p/\partial z = \partial \ln \rho/\partial z = -1/H$, where

$$H = \frac{RT}{g} = \frac{p}{\rho g} = \frac{c^2}{\kappa g} \quad (2.6)$$

is the so-called *height of the homogeneous atmosphere* (for $T \approx 0^\circ \text{C}$ H is approximately 8 km).

3. FLUID DYNAMICS

A Eulerian description of a moving fluid makes use of the fields of density $\rho(\mathbf{x}, t)$

and velocity $\mathbf{u}(\mathbf{x}, t)$; we will employ space coordinates $\mathbf{x}$ and the *relative* velocity $\mathbf{u}$ in a reference system rotating at a constant angular velocity $\boldsymbol{\omega}$ (usually the mean angular velocity of rotation of the planet). Let us write the *evolution* equations of fluid dynamics (that is, the equations containing the derivatives with respect to time $\partial/\partial t$ of the unknown fields), namely the continuity equation and the equation of motion, as

$$\frac{d\rho}{dt} + \rho D = 0; \quad D = \text{div } \mathbf{u}; \quad (3.1)$$

$$\mathbf{w} = \frac{d\mathbf{u}}{dt} + 2\boldsymbol{\omega} \times \mathbf{u} = -\frac{\nabla p}{\rho} + \mathbf{f}, \quad (3.2)$$

where $d/dt = \partial/\partial t + \mathbf{u} \cdot \nabla$ is the individual derivative with respect to time; $\mathbf{w}$ is the acceleration in the absolute reference system ($2\boldsymbol{\omega} \times \mathbf{u}$ is the *Coriolis acceleration*); $\mathbf{f}$ is the sum of the accelerations due to viscosity and the external mass forces, of which in geophysical fluid dynamics just the gravity force usually needs to be taken into account (in the study of tides the gravity force and the gravitational *tide-producing* force are added together). If the moving fluid is electrically conducting and is in a magnetic field (of intensity $\mathbf{H}$), then in general the *ponderomotive* Lorentz force also has to be taken into account; in the approximation of magnetic fluid dynamics (neglecting the effects of the finiteness of the velocity c_0 of electromagnetic disturbances, that is, neglecting the displacement current $(1/4\pi)(\partial\mathbf{E}/\partial t)$ in comparison with the currents of conduction and induction

$$\mathbf{j} = \sigma \left(\mathbf{E} + \frac{1}{c_0} \mathbf{u} \times \mathbf{H} \right),$$

where σ is the electrical conductivity of the medium, and $\mathbf{E}$ is the electric-field strength; in this approximation $\text{curl } \mathbf{H} = (4\pi/c_0)\mathbf{j}$, the Lorentz acceleration can be written as

$$\mathbf{f}_L = \frac{1}{c_0\rho} \mathbf{j} \times \mathbf{H} = \frac{1}{4\pi\rho} \text{curl } \mathbf{H} \times \mathbf{H}. \quad (3.3)$$

It is often convenient to consider separately on the right-hand side of equations of motion (3.2) the acceleration of the buoyancy forces. This is easy to do approximately, if we specify the *main quasistatic* ('equilibrium') stratification of the medium, for instance, in terms of functions $T = T_0(z)$ and $s = s_0(z)$, whereby functions $p = p_0(z)$ and $\rho = \rho_0(z)$ are found from the system of equations of state $\rho_0 = \rho(T_0, p_0, s_0)$ and hydrostatics $\nabla p_0 = \rho_0 \mathbf{g}$ (and then any of the other thermodynamic characteristics can be found as well), and assuming the deviations of the actual thermodynamic fields from these equilibrium values $\rho' = \rho - \rho_0$, $p' = p - p_0$ to be small (in comparison with their equilibrium values). Then, accurate to first-order terms, taking into account the relation $\nabla p_0 = \rho_0 \mathbf{g}$, we get

$$-\frac{\nabla p}{\rho} + \mathbf{g} = -\frac{\nabla p_0 + \nabla p'}{\rho_0 + \rho'} + \mathbf{g} \approx -\frac{\nabla p'}{\rho_0} + \frac{\rho'}{\rho_0} \mathbf{g}. \quad (3.4)$$

The use of the right-hand side of this equation in (3.2) instead of the left-hand side, and the simultaneous substitution of the condition $D = 0$ of nondivergence (solenoidality) of the velocity field in place of continuity equation (3.1), is known as the *Boussinesq approximation*. In this approximation the buoyancy acceleration is equal to $\rho'g/\rho_0$. If ρ_0 varies greatly with height within the layer being considered (for instance, in the convective layer of the Sun, by a factor of 10^6), then it may be convenient to introduce the function $\rho_{00}(z)$ from the relation $(g/\rho_{00})(\partial\rho_{00}/\partial z) = -N_0^2$ (where ρ_{00} is analogous to the potential density), as well as the reduced pressure $p'' = (\rho_{00}/\rho_0)p'$. Now the right-hand side of (3.4) can, with the aid of (2.4), be written as

$$-\frac{\nabla p'}{\rho_0} + \frac{\rho'}{\rho_0} \mathbf{g} = -\frac{\nabla p''}{\rho_{00}} - \eta'' \mathbf{g}; \quad \eta'' = \frac{p' - c_0^2 \rho'}{\rho_0 c_0^2}, \quad (3.5)$$

where η'' denotes the dimensionless perturbations of the specific entropy. For small $N_0^2 h/g$ (where h is the thickness of the layer) we can assume that $\rho_{00} \approx \text{const}$, and the first term on the right-hand side will be approximately potential, while the second term is proportional to the linearized entropy η' . If $\mathbf{f}_L$ is not taken into account, then the four scalar equations (3.1)–(3.2) contain five unknown scalar functions ρ , $\mathbf{u}$, and p , and the system of equations must be supplemented by the equation of state (which does not contain the derivatives with respect to time; that is, it is a *synchronous relation* between the thermodynamic fields). If the medium is to be *barotropic*, that is, $\rho = \rho(p)$, then the system of equations obtained should be closed. If, on the other hand, the medium is *baroclinic*, and $\rho = \rho(T, p)$ (or $\rho = \rho(T, p, s)$), then still another synchronous relation has to be added: the formula for the entropy $\eta = \eta(T, p)$ (or $\eta = \eta(T, p, s)$), and evolution equation (1.15) for η (or the two evolution equations (1.15) and (1.16) for η and s). Finally, if $\mathbf{f}_L$ is also taken into account, then the system of equations must also include the *induction equation*

$$\frac{\partial \mathbf{H}}{\partial t} - \text{curl}(\mathbf{u} \times \mathbf{H}) = -\text{curl}(v_m \text{curl} \mathbf{H}); \quad v_m = \frac{c_0^2}{4\pi\sigma}, \quad (3.6)$$

where v_m is the *magnetic viscosity* (here $\text{div} \mathbf{H} = 0$).

Now let us consider equations of motion (3.2) and let us note that the acceleration field $\mathbf{w}$, and in general any vector field, is completely determined by its curl, divergence, and normal component at the boundary S of the given region of space; such a description of the motion turns out to be very convenient in a number of problems. When calculating $\text{curl} \mathbf{w}$ we will use a tensor (transforming vectors into vectors) hydrodynamic (with coefficients depending on the velocity field $\mathbf{u}$) linear differential *Helmholtzian* operator helm . This operator, which was introduced by A. A. Fridman, is defined for each vector field $\mathbf{A}$ by the formula

$$\text{helm} \mathbf{A} = \frac{d\mathbf{A}}{dt} - (\mathbf{A} \cdot \nabla) \mathbf{u} + D\mathbf{A}. \quad (3.7)$$

Its meaning in fluid dynamics is as follows: $\text{helm} \mathbf{A} = 0$ signifies ‘freezing’ of the vector field $\mathbf{A}$ into the moving fluid. If field $\mathbf{A}$ is ‘frozen in’, then each vector line of it (that is, a line whose tangent at any point has the direction of field $\mathbf{A}$ at this point) is ‘conserved’, that is, it always consists of the same fluid particles, while the

intensities of the vector tubes (the walls of which consist of vector lines), that is, the fluxes $\int_S \mathbf{A} \cdot d\mathbf{S}$ of vector $\mathbf{A}$ through any cross sections S of these tubes, do not vary with time.

Thus, for instance, the left-hand side of induction equation (3.6), allowing for the condition $\text{div } \mathbf{H} = 0$, can be put into the form $\text{helm } \mathbf{H}$, and this equation becomes

$$\text{helm } \mathbf{H} = -\text{curl}(\nu_m \text{curl } \mathbf{H}), \quad (3.6')$$

from which we have at the limit as $\nu_m \rightarrow 0$ (that is, as $\sigma \rightarrow \infty$) $\text{helm } \mathbf{H} = 0$: in a superconducting moving medium the magnetic field turns out to be 'frozen in.' For $\nu_m \approx \text{const}$ the right-hand side of Equation (3.6') can be rewritten as $\nu_m \Delta \mathbf{H}$, and then it is evident that the freezing in of the magnetic field will in general be disrupted because of its diffusion as a result of the magnetic viscosity. For instance, in the liquid layer of the Earth's core for an electrical conductivity $\sigma \sim 3 \times 10^9 \Omega^{-1} \text{cm}^{-1}$ the magnetic viscosity $\nu_m \sim 2.5 \times 10^4 \text{cm}^2/\text{s}$, the *magnetic Reynolds number* $\text{Re}_m = \nu_m^{-1} UL$ for motions on a global scale $L \sim 10^4 \text{km}$ with velocities of the order of $U \sim 10^{-2} \text{cm/s}$ is of the order of 10^3 , and the diffusion time of the global magnetic field $\nu_m^{-1} L^2 \sim 10^6$ years is comparable to the fundamental periods of reversals of the polarity of the geomagnetic field (as observed by paleomagnetologists).

Calculating the curl of the acceleration (3.2) and using notation (3.7), we obtain Fridman's equation for the vorticity:

$$\text{curl } \mathbf{w} = \text{helm}(\boldsymbol{\Omega} + 2\boldsymbol{\omega}) = \rho^{-2} \nabla \rho \times \nabla p + \text{curl } \mathbf{f}; \quad \boldsymbol{\Omega} = \text{curl } \mathbf{u}. \quad (3.8)$$

Here $\boldsymbol{\Omega}$ is the vorticity of the relative motion, while $2\boldsymbol{\omega}$ is the vorticity of the translational motion (rotation of planet), so that $\boldsymbol{\Omega} + 2\boldsymbol{\omega}$ is the absolute vorticity, and thus the curl of the absolute acceleration equals the Helmholtzian of the absolute vorticity. The right-hand side of equation (3.8) describes the internal (that is, not at the boundary) eddy sources. The first term $\rho^{-2} \nabla \rho \times \nabla p$ differs from zero only if the medium is baroclinic; then the families of surfaces $p = \text{const}$ and $\rho = \text{const}$ do not coincide, and their intersections form tubes (known as *isobaro-isosteric solenoids*), in the cross sections of which circulations arise. According to equation of state (1.3),

$$\rho^{-2} \nabla \rho \times \nabla p = -\rho^{-1} \alpha \nabla T \times \nabla p + \rho^{-1} \gamma \nabla s \times \nabla p, \quad (3.9)$$

where the first term describes the eddy-producing effect of *isobaro-isothermal solenoids* and the second term describes the effect of *isobaro-isoconcentration solenoids*. Since the free-fall acceleration $\mathbf{g}$ is the gradient of the gravitational potential, it is not a vortical acceleration, and its contribution to $\text{curl } \mathbf{f}$ will be nil. However, the curl of the buoyancy acceleration entering into (3.9) is not zero; also vortical, in general, are the acceleration of the viscosity force and the acceleration of the Lorentz force (3.3). If we also calculate the divergence of acceleration (3.2)

and its normal component w_N at a rigid stationary boundary S , then we get

$$\begin{aligned}\operatorname{div} \mathbf{w} &= \frac{dD}{dt} + \frac{1}{2} \Delta(u^2) - \mathbf{u} \cdot \Delta \mathbf{u} - \boldsymbol{\Omega}(\boldsymbol{\Omega} + 2\boldsymbol{\omega}) \\ &= -\operatorname{div} \frac{\nabla p}{\rho} + \operatorname{div} \mathbf{f};\end{aligned}\quad (3.10)$$

$$w_N = -\frac{1}{\rho} \frac{\partial p}{\partial N} + f_N = 0 \quad \text{at } S. \quad (3.11)$$

As already noted, the system of equations (3.8), (3.10), (3.11) is equivalent to the equations of motion (with the boundary condition $u_N = 0$ at S).

4. POTENTIAL VORTICITY

Many natural thermohydrodynamic processes are approximately adiabatic (at least during quite short time intervals). However, if the disturbances of the adiabaticity are considerable, then they are able to shape the specific properties of the process. Therefore, the field of the entropy η (and, if a thermodynamically active admixture is present, the field of its concentration s as well) deserves special consideration in geophysical fluid dynamics, as we saw, for instance, in Section 2, where the vertical stability of the medium was seen to be determined by the stratification of the potential density $\rho_*(\eta, s)$.

In particular, it will be interesting to consider the direction $\mathbf{e} = \nabla \eta |\nabla \eta|^{-1}$ of the largest entropy variation at a given point (called the *thermodynamic vertical*, since, at least on a global scale, the $\eta = \text{const}$ surfaces are approximately horizontal, so that $\mathbf{e}$ is, as a rule, approximately vertical — see the example below, in Figure 4.1), and the components $A_e = \mathbf{e} \cdot \mathbf{A}$ of the vector fields $\mathbf{A}$ in this direction. The latter can be described using the following hydrodynamic theorem (Monin, 1971), easily proven with the aid of definition (3.7): *for any vector field $\mathbf{A}$ and for any scalar field η satisfying equation (1.15), the following relation is valid:*

$$\rho \frac{d}{dt} (\rho^{-1} \mathbf{A} \cdot \nabla \eta) = \nabla \eta \cdot \operatorname{helm} \mathbf{A} + \mathbf{A} \cdot \nabla \varepsilon_\eta. \quad (4.1)$$

Let us choose as $\mathbf{A}$ the absolute vorticity $\boldsymbol{\Omega} + 2\boldsymbol{\omega}$, satisfying Equation (3.8), and let us consider for the time being that thermodynamically active admixtures are absent (as in the case of dry air or fresh water), so that $\eta = \eta(\rho, p)$ and thus $\nabla \eta (\nabla \rho \times \nabla p) = 0$. Then, from (4.1) and (3.8) we get

$$\begin{aligned}\rho \frac{d\Omega_*}{dt} &= \nabla \eta \cdot \operatorname{curl} \mathbf{f} + (\boldsymbol{\Omega} + 2\boldsymbol{\omega}) \cdot \nabla \varepsilon_\eta = \operatorname{div} [\eta \operatorname{curl} \mathbf{f} + \varepsilon_\eta (\boldsymbol{\Omega} + 2\boldsymbol{\omega})]; \\ \Omega_* &= \rho^{-1} (\boldsymbol{\Omega} + 2\boldsymbol{\omega}) \cdot \nabla \eta\end{aligned}\quad (4.2)$$

(Obukhov, 1962). The quantity Ω_* is called the *potential vorticity* (Rossby, 1940); combining the effects of rotation $\boldsymbol{\Omega} + 2\boldsymbol{\omega}$ and stratification $\nabla \eta$, the function $\Omega_*(\mathbf{x}, t)$ is the field most typical of geophysical fluid dynamics. Its evolution

equation (4.2) is essentially the projection of the equation for the vorticity (3.8) onto the thermodynamic vertical. If the Lorentz force is not acting, then in (4.2) $\mathbf{f}$ is just the acceleration of the viscosity forces, and for adiabatic processes, when $\mathbf{f} = \varepsilon_\eta = 0$, we have $d\Omega_*/dt = 0$ (Ertel, 1942), that is, Ω_* is an *adiabatic Lagrangian invariant*, and η is as well, according to the definition of an adiabatic process (Obukhov's second equation (4.2) shows that $d\Omega_*/dt = 0$, that is, Ω_* is a Lagrangian invariant not only for adiabatic processes in the absence of a Lorentz force; for this it is sufficient that the vector field $\eta \text{ curl } \mathbf{f} + \varepsilon_\eta(\boldsymbol{\Omega} + 2\boldsymbol{\omega})$ be solenoidal).

The general procedure for constructing the adiabatic Lagrangian invariants (Gavrilin and Zaslavskii, 1970) consists in using for the field of the *absolute* velocity the generalized Clebsch representation $\mathbf{v} = \beta \nabla \eta + \lambda \nabla \mu + \nabla W$, where $I_1 = \eta$, $I_2 = \lambda$, and $I_3 = \mu$ are the primary invariants, $\beta = \int_0^t T dt$ is the thermodynamic action, and $W = \int_0^t \Lambda dt$ is the action with a Lagrangian $\Lambda = (v^2/2) - \varphi - w$ (here φ is the gravitational potential and w is the enthalpy). Each trio of invariants can be used to construct a new invariant $\rho^{-1} \nabla I_1 (\nabla I_2 \times \nabla I_3)$. For instance, from invariants η , λ , μ we get $I_4 = \Omega_*$, and then the invariants of Holman (1964) $I_5 = \rho^{-1} \lambda \nabla \mu (\nabla \eta \times \nabla \Omega_*)$, $I_6 = \rho^{-1} \lambda \nabla \mu (\nabla \eta \times \nabla I_5)$, and $I_7 = \rho^{-1} \nabla I_5 \times (\nabla \eta \times \nabla \Omega_*)$, etc. (the invariance of I_1 , I_4 , I_5 , I_6 , I_7 does not yet yield equations for determining W , since it only provides five equations in the four unknowns η , λ , μ , and ρ).

Moiseev *et al.* (1982) cited another wide class of Lagrangian invariants: fields $\rho^{-1} \mathbf{A} \cdot \nabla I$, where I is a Lagrangian invariant and $\mathbf{A}$ is the 'freezing-in' integral, that is, a vector field with a Helmholtzian equal to zero. They also indicated methods of constructing the freezing-in integrals. In one such method the fields $\mathbf{A} = \nabla I_1 \times \nabla I_2$ are constructed from two invariants, while in the general case the fields

$$\mathbf{A}'_\alpha = \varepsilon_{\alpha\beta\gamma} \varepsilon_{ikl} I_i \left(\frac{\partial I_k}{\partial x_\beta} \right) \left(\frac{\partial I_l}{\partial x_\gamma} \right),$$

are constructed from three invariants I_1, I_2, I_3 (even more general is the field,

$$\mathbf{A}''_\alpha = \varepsilon_{\alpha\beta\gamma} I (\partial I' / \partial x_\beta \times \partial I'' / \partial x_\gamma),$$

where I, I', I'' are three different vector fields, whose components are Lagrangian invariants).

For simplicity, let us confine ourselves here to solenoidal frozen-in fields $\mathbf{A} = \text{curl } \mathbf{B}$, where $\mathbf{B}$ is the vector potential (this limitation is easily lifted). Then the helicity of field $\mathbf{B}$, that is, the integral $\int \mathbf{B} \cdot \text{curl } \mathbf{B} dV$, does not vary with time. It characterizes the topological entanglement of the force lines of field $\mathbf{A}$. For the field $\mathbf{A} = \nabla I \times \nabla I_2$ the helicity is zero, so that the force lines are not entangled and do not have knots. For fields $\mathbf{A}'_\alpha$, on the other hand, the helicity need not be zero, so that their force lines may become entangled and knotted. With the aid of velocity fields of this type, Moiseev *et al.* constructed some examples of topological solitons.

Clearly, any function of η and Ω_* will itself also be an adiabatic Lagrangian invariant. It is convenient to take two such independent functions as the Lagrangian coordinates of the fluid particles. The corresponding coordinate

surfaces intersect the medium at tubes, through the walls of which fluid cannot pass, so that the adiabatic evolution of the medium will consist in deformations of these tubes. As an example, we consider, following Obukhov (1964), an atmosphere of dry air with equation of state (1.7) for $s = 0$, in which the entropy is given by the formula

$$d\eta = c_p d \ln \Theta; \quad \Theta = T \left(\frac{p_0}{p} \right)^{(\kappa-1)/\kappa}, \quad (4.3)$$

where Θ is the so-called *potential temperature*, and p_0 is the standard pressure, say, at sea level; according to (2.5), $\partial\Theta/\partial z = (\Theta/T)(\gamma_a - \gamma)$, where $\gamma = -\partial T/\partial z$ is the lapse rate. According to the statistical calculations of Oort and Rasmusson (1971), in the real terrestrial atmosphere on the average for 1958 to 1963 the mean annual values of Θ everywhere increased with height (that is, the atmosphere was hydrostatically stable; quite rare exceptions to this were the convective clouds which sometimes appeared here and there, especially in the tropics), while the mean pressure at the isentropic surfaces $p_*(\Theta)$ decreased from 920 mbar at $\Theta = 260$ K to 100 mbar at $\Theta = 400$ K. From the equator to the pole at the 1000-mbar level Θ decreased (from 299.5 K at the equator to 261.2 K at 75°N). At a level of about 200 mbar, Θ varied little (staying about 345 K), while at the 50-mbar level it increased (from 489.2 K to 522.7 K). Therefore, on a global scale surfaces $\Theta = \text{const}$ are quasihorizontal (so that Θ can serve as a vertical coordinate), and approximately $\Omega_* \approx \rho^{-1}(\Omega_z + 2\omega_z)(c_p/T)(\gamma_a - \gamma)$. Consequently, Obukhov took as the Lagrangian coordinates of the air particles the quantities Θ and

$$\tilde{\Omega} = (\Omega_z + 2\omega_z) \frac{p_*(\Theta)}{p} \frac{\gamma_a - \gamma}{\gamma_a - \gamma_*(\Theta)}, \quad (4.4)$$

where $\gamma_*(\Theta)$ is the mean lapse rate at the isentropic surfaces. On a large horizontal scale outside the tropical zone $\Omega_z + 2\omega_z \approx 2\omega_z \approx 2\omega \cos \Theta$ (this is the *Coriolis parameter*, usually designated as f), so that coordinate Ω can replace the latitude. Accordingly, a good representation of the invariant $(\Theta, \tilde{\Omega})$ tubes is given by their meridional sections (an example of such a section is given in Figure 4.1).

Now let $m(\Theta, \tilde{\Omega}) d\Theta d\tilde{\Omega}$ be the fraction of the mass of the atmosphere contained in an infinitely thin $(\Theta, \tilde{\Omega})$ tube. Then $m(\Theta, \tilde{\Omega})$ represents the probability density for coordinates $\Theta, \tilde{\Omega}$ of a randomly chosen air particle. According to empirical data, this probability is concentrated mainly in the region $280 \text{ K} < \Theta < 305 \text{ K}$; $0.5 \text{ s}^{-1} < 10^4 \tilde{\Omega} < 1.5 \text{ s}^{-1}$ with a maximum near the point $\Theta = 290 \text{ K}$, $10^4 \tilde{\Omega} = 1 \text{ s}^{-1}$ and a minimum in the region $297 \text{ K} < \Theta < 305 \text{ K}$, $2.2 \text{ s}^{-1} < 10^4 \tilde{\Omega} < 3.2 \text{ s}^{-1}$ (see example in Figure 4.2).

Obukhov introduced the concept of an *equivalent zonal model of the atmosphere* (EZMA), having the same probability density $m(\Theta, \tilde{\Omega})$ as the actual atmosphere. Kurganskii (1976) constructed approximate two-level EZMA in which the linearized dimensionless values of the potential vorticity $\xi_{1,2} = (2\omega)^{-1}[\Delta\psi_{1,2} + f \pm L_R^{-2}(\psi_1 - \psi_2)]$ at the $p_1 = 700$ mbar and $p_2 = 300$ mbar levels were used as the Lagrangian invariants. Here $\psi_{1,2}$ are the horizontal stream functions, and $L_R \sim HN/f$ is the horizontal scale of the synoptic processes (known as the *Rossby deformation radius*); the probability densities $m(\xi_{1,2})$ turn out to be

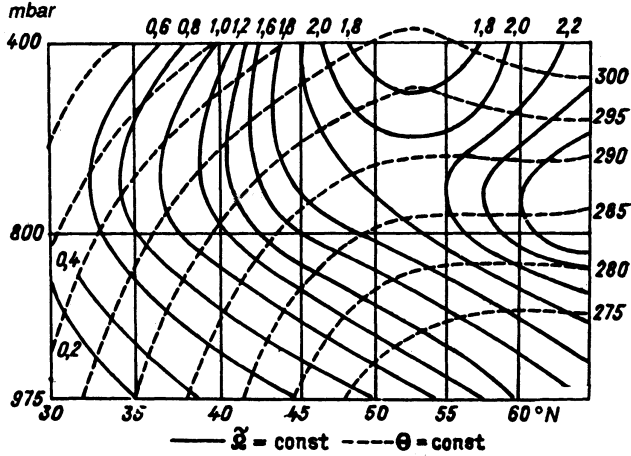


Fig. 4.1. Meridional section of $(\Theta, \bar{\Omega})$ tubes at 100°W on 1 April 1962, according to Obukhov (1964); $\bar{\Omega}$ in units of 10^{-4} s^{-1} .

considerably more diffuse at the upper level than at the lower level, in accordance with Figure 4.2.

Next we consider more complicated media, in which the role of thermodynamically active admixtures in function $\eta(\rho, p, s)$ is considerable and the term $\rho^{-2} \nabla \eta (\nabla \rho \times \nabla p)$ on the right-hand side of the equation does not equal zero; in this case the potential vorticity is no longer a Lagrangian invariant, even for adiabatic processes. However, this difficulty can be mitigated (Monin, 1974) if instead of the entropy we use the *pseudoentropy* $\eta^* = \eta(\rho, p, s^*)$, that is, the entropy reduced isopycnically and isobarically to some standard concentration of the admixture $s = s^*$ (the values for $s = s^*$ of the functions of ρ, p, s will be denoted by superscript asterisks). From (1.2), (1.3), (1.6) we then get the following relation for the pseudoentropy:

$$d\eta^* = \frac{c_p^*}{\rho \alpha^* T^* c^{*2}} (dp - c^{*2} d\rho), \quad (4.5)$$

which gives, taking (1.15), (1.16), and (3.1) into account,

$$\begin{aligned} \varepsilon^* = \frac{d\eta^*}{dt} = & \frac{c_p^*}{\alpha^* T^*} \left(1 - \frac{c^2}{c^{*2}} \right) D + \\ & + \frac{\alpha T c^2}{\alpha^* T^* c^{*2}} \frac{c_p^*}{c_p} \left[\varepsilon_\eta - \left(\delta + \frac{c_p \gamma}{\alpha T} \right) \varepsilon_s \right]. \end{aligned} \quad (4.6)$$

Therefore, if η, ε_η are replaced by η^*, ε^* , equation (4.2) remains valid; the quantity $\Omega^* = \rho^{-1}(\bar{\Omega} + 2\omega) \cdot \nabla \eta^*$ can be called the *pseudopotential vorticity*. The

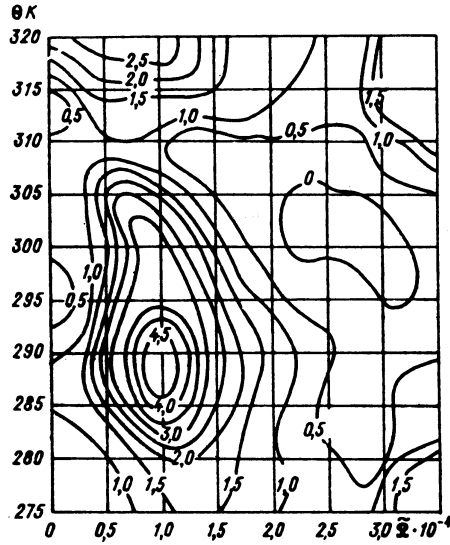


Fig. 4.2. Probability density $m(\Theta, \hat{\Omega})$ for period from 1 to 10 April 1962 (according to Karunin).

quantities η^* and Ω^* are, strictly speaking, not adiabatic Lagrangian invariants, since for $\varepsilon_\eta = \varepsilon_s = 0$ the quantity ε^* does not go to zero, but rather is equal to the first term on the right-hand side of Equation (4.6). On the other hand, this term is usually very small (to equate it with the real values of ε_η , for instance, is sensible), both because $c \approx c^*$ (for instance, in the case of (1.8), $c^2/c^{*2} = \kappa/\kappa^*$) and in the case of quasisolenoidal motions, when $D \approx 0$ (which is, as a rule, true for large-scale motions), so that η^* and Ω^* can be assumed approximately to be adiabatic Lagrangian invariants.

It is clear from (4.2) that taking the Lorentz force (3.3) into account in magnetohydrodynamics disrupts the adiabatic invariance of the potential vorticity. However, in magnetohydrodynamics there will be another adiabatic Lagrangian invariant. For instance, if we set $\mathbf{A} = \mathbf{H}$ in (4.1) and make use of induction equation (3.6'), then for any function $\eta(\rho, p, s)$ we get

$$\rho \frac{dH_*}{dt} = -\nabla \eta \cdot \text{curl}(\nu_m \text{curl} \mathbf{H}) + \mathbf{H} \cdot \nabla \varepsilon_\eta;$$

$$H_* = \rho^{-1} \mathbf{H} \cdot \nabla \eta \quad (4.7)$$

(Monin, 1971). Hence we see that in the adiabatic approximation, when it is assumed that $\nu_m = \varepsilon_\eta = 0$ (where $\nu_m = 0$ means that we neglect the ohmic dissipation of the magnetic-field energy, that is, the generation of Joule heat, which takes place at a rate $\nu_m/4\pi \times (\text{curl} \mathbf{H})^2$), we obtain $dH_*/dt = 0$, so that the quantity H_* , which we will call the *charge of the magnetic field*, turns out to be a Lagrangian invariant.

5. ENERGY

Let us write equations of motion (3.2) in a somewhat different form:

$$\frac{\partial \rho \mathbf{v}}{\partial t} = -\text{div } \Pi + \rho \nabla \Phi, \quad (5.1)$$

where $\mathbf{v}$ is the *absolute* velocity of motion, Φ is the gravitational potential, and Π is the tensor of hydromagnetic stresses, having the Cartesian components:

$$\begin{aligned} \Pi_{ik} &= \rho v_i v_k + p \delta_{ik} - \sigma'_{ik} - \Pi'_{ik}; \\ \sigma'_{ik} &= \rho \nu \left(\frac{\partial v_i}{\partial x_k} + \frac{\partial v_k}{\partial x_i} \right) + \rho \left(\nu_1 - \frac{2}{3} \nu \right) D \delta_{ik}; \\ \Pi'_{ik} &= (4\pi)^{-1} \left(H_i H_k - \frac{1}{2} H^2 \delta_{ik} \right), \end{aligned} \quad (5.2)$$

where x_i , v_i , H_i are the Cartesian coordinates and the components of the velocity and the magnetic-field strength; δ_{ik} is the unit tensor; σ'_{ik} are the viscous stresses (ν and ν_1 being the kinematic viscosities); and Π'_{ik} are the magnetic Maxwell stresses. Now, if we take the scalar product times $\mathbf{v}$, then after some identity transformations we get the following *local* equation for the density of the kinetic energy $\rho v^2/2$:

$$\begin{aligned} \partial \left(\frac{\rho v^2}{2} \right) / \partial t + \text{div} \left[\rho \mathbf{v} \left(\frac{v^2}{2} + \frac{p}{\rho} \right) - \mathbf{v} \cdot \boldsymbol{\sigma}' - \mathbf{v} \cdot \boldsymbol{\Pi}' \right] \\ = pD - \sigma'_{ik} \frac{\partial v_i}{\partial x_k} - \Pi'_{ik} \frac{\partial v_i}{\partial x_k} + \rho \mathbf{v} \cdot \nabla \Phi. \end{aligned} \quad (5.3)$$

The expression in square brackets on the left-hand side of this equation is the flux density of kinetic energy. Similarly, if we take the scalar product of induction equation (3.6) times $\mathbf{H}$, we arrive at the following equation for the energy density of the magnetic field $H^2/8\pi$:

$$\begin{aligned} \frac{\partial (H^2/8\pi)}{\partial t} + \text{div} \left[\frac{1}{4\pi} \mathbf{H} \times (\mathbf{v} \times \mathbf{H}) - \frac{\nu_m}{4\pi} \mathbf{H} \times \text{curl } \mathbf{H} + \mathbf{v} \cdot \boldsymbol{\Pi}' \right] \\ = -\frac{\nu_m}{4\pi} (\text{curl } \mathbf{H})^2 + \Pi'_{ik} \frac{\partial v_i}{\partial x_k}. \end{aligned} \quad (5.4)$$

In the flux density of magnetic energy (given inside the square brackets on the left-hand side of this equation) the first two terms represent the Poynting vector in the magnetohydrodynamic approximation. Next, using the definition $\varepsilon = \phi - p/\rho + T\eta$ and Equation (1.1) and interpreting the right-hand sides of (1.15) and

(1.16), we obtain the following equation for the density of internal energy $\rho\varepsilon$:

$$\begin{aligned} \frac{\partial \rho\varepsilon}{\partial t} + \operatorname{div}(\rho \mathbf{v} \varepsilon + \mathbf{q}) \\ = -pD + \sigma'_{ik} \frac{\partial v_i}{\partial x_k} + \frac{\nu_m}{4\pi} (\operatorname{curl} \mathbf{H})^2 + Q, \end{aligned} \quad (5.5)$$

where $\mathbf{q}$ is the flux density of internal energy, produced by molecular processes, and Q is the density of outside sources of internal energy. A comparison of the right-hand sides of Equations (5.3), (5.4), and (5.5) shows that term pD describes the exchange between kinetic and internal energies due to adiabatic heating or cooling of the medium as it contracts or expands; term $\sigma'_{ik} \partial v_i / \partial x_k$ describes the viscous dissipation of kinetic energy as heat; $\Pi'_{ik} \partial v_i / \partial x_k$ describes the mutual transformations of kinetic and magnetic energies due to the work of ponderomotive forces — the Maxwell stresses (where this term is positive, it represents the generation of a magnetic field; where it is negative, it represents the dynamo effect, that is, the motion of a current-carrying conductor in a magnetic field); $\rho \mathbf{v} \cdot \nabla \Phi$ describes the exchange between kinetic energy and potential energy in the gravitational field ($-\rho\Phi$); and, finally, term $(\nu_m/4\pi)(\operatorname{curl} \mathbf{H})^2$ describes the ohmic dissipation of magnetic energy as Joule heat. For the total energy we now obtain the local equation

$$\begin{aligned} \frac{\partial}{\partial t} \left[\rho \left(\frac{v^2}{2} + \varepsilon - \Phi \right) + \frac{H^2}{4\pi} \right] + \operatorname{div} \left[\rho \mathbf{v} \left(\frac{v^2}{2} + \frac{p}{\rho} + \varepsilon - \Phi \right) + \right. \\ \left. + \frac{1}{4\pi} \mathbf{H} \times (\mathbf{v} \times \mathbf{H}) - \mathbf{v} \cdot \boldsymbol{\sigma}' - \frac{\nu_m}{4\pi} \mathbf{H} \times \operatorname{curl} \mathbf{H} + \mathbf{q} \right] = Q. \end{aligned} \quad (5.6)$$

From (5.6) we will derive an integral equation for the total energy, limited here to the nonmagnetic case (so that all the terms containing $\mathbf{H}$ drop out), but then, having in mind, in particular, the World Ocean, we consider the case in which the volume V occupied by the medium is bounded at the bottom by a stationary rigid surface, the 'floor', with a coordinate $z = -h(\mathbf{x}_h)$ (where $\mathbf{x}_h$ are the horizontal coordinates), and at the top by a fluid boundary surface S with a coordinate $z = \xi(\mathbf{x}_h, t)$. Let us integrate local equation (5.6) over volume V , taking into account the equations:

$$\int \frac{\partial A}{\partial t} dV = \frac{\partial}{\partial t} \int A dV - \int \hat{A} \frac{\partial \xi}{\partial t} d\Sigma; \quad (5.7)$$

$$\int \operatorname{div} \mathbf{A} dV = \int \hat{\mathbf{A}} \cdot \mathbf{n} dS = \int \hat{\mathbf{A}} \cdot \mathbf{n} [1 + (\nabla \xi)^2]^{1/2} d\Sigma, \quad (5.8)$$

where $\mathbf{n}$ is the unit vector of an outward normal to surface S , $dS = [1 + (\nabla \xi)^2]^{1/2} d\Sigma$ is an area element of this surface, and $d\Sigma$ is an element of the horizontal area; a

circumflex indicates values of the functions on surface S ; the first of Equations (5.8) represents the Gauss theorem. We will also make use of the familiar *kinematic boundary condition at a fluid surface*:

$$\frac{\partial \xi}{\partial t} = -\hat{\mathbf{v}} \cdot \nabla \xi + \hat{v}_z = \hat{\mathbf{v}} \cdot \mathbf{n} [1 + (\nabla \xi)^2]^{1/2}. \quad (5.9)$$

With the aid of (5.7)–(5.9), we can now write the integral of (5.6) over V as

$$\begin{aligned} \frac{d\mathcal{E}}{dt} = & - \int \hat{p} \frac{\partial \xi}{\partial t} d\Sigma + \int (\hat{\mathbf{v}} \cdot \hat{\sigma}' - \hat{\mathbf{q}}) \cdot \mathbf{n} dS + \int Q dV; \\ \mathcal{E} = & \int \left(\frac{v^2}{2} + \varepsilon - \Phi \right) dm, \end{aligned} \quad (5.10)$$

where $dm = \rho dV$ is a mass element. When calculating the first two terms on the right-hand side, we have to take into account, in addition to the kinematic boundary condition (5.9), the requirement of continuity of the velocities, forces, temperatures, and heat fluxes upon transition through the fluid boundary surface from one medium to the other. For the forces this requirement takes the form

$$\begin{aligned} \mathbf{n}(\hat{p} - \hat{p}_a) = & -\mathbf{n}\gamma\hat{p} \operatorname{div}\{[1 + |\nabla \xi|^2]^{-1/2} \nabla \xi\} + \\ & + \nabla(\gamma\hat{p}) + \mathbf{n} \cdot (\hat{\sigma}' - \hat{\sigma}'_a). \end{aligned} \quad (5.11)$$

Here subscript a designates a characteristic of the upper medium (of the atmosphere, in the case of the ocean surface). The first term on the right-hand side is the *surface pressure*, a quantity proportional to the Gaussian curvature of surface S (the multiplier with the divergence); the proportionality coefficient $\gamma\hat{p}$ is called the *coefficient of surface tension* (at the water-air interface $\gamma \approx 72.5 \text{ cm}^3/\text{s}^2$ at 20°C). The second term describes the effect of horizontal inhomogeneities of the surface tension, and the third reflects the difference between the forces of viscous friction in the two media at surface S . If $\gamma\hat{p}$ is taken to be quasiconstant, then taking into account the first term on the right-hand side of (5.11) reduces to adding the following term to the potential energy $\mathcal{P} = -\int \Phi dm$:

$$\mathcal{P}_\gamma = \int \gamma\hat{p} \{ [1 + (\nabla \xi)^2]^{1/2} - 1 \} d\Sigma. \quad (5.12)$$

For adiabatic processes we have to set $\sigma' = \mathbf{q} = Q = 0$ in (5.10), and in (5.11) we have to drop the last term and the term $\nabla(\gamma\hat{p})$ (since it is directed along the tangent, while the other terms lie along the normal to S). Thus (5.10) simplifies to

$$\begin{aligned} \left(\frac{\partial}{\partial t} \right) (\mathcal{K} + \mathcal{L}) = & - \int \hat{p}_a \frac{\partial \xi}{\partial t} d\Sigma; \\ \mathcal{K} = & \int \frac{v^2}{2} dm; \mathcal{L} = \int (\varepsilon - \Phi) dm + \mathcal{P}_\gamma. \end{aligned} \quad (5.13)$$

where $\mathcal{K}$ is the kinetic energy, and $\mathcal{L}$ is the so-called *labile energy*, consisting of the internal energy plus the potential energy. In cases where the surface S is a *free surface* (that is, free of external influences), so that the right-hand side of (5.13) goes to zero (this occurs, in particular, for $\hat{p}_a = \text{const}$), the total energy $\mathcal{K} + \mathcal{L}$ does not vary with time. Therefore, for adiabatic processes only exchange between the kinetic and labile energies is possible. Obviously, during such exchange the labile energy cannot drop below some minimum $\mathcal{L}_*$, known as the *unavailable labile energy*. The difference $\Pi = \mathcal{L} - \mathcal{L}_*$, called the *available potential energy* by Lorenz (1956), is the maximum amount of kinetic energy obtainable at the expense of the labile energy with the aid of adiabatic processes. The quantities $\mathcal{K} + \mathcal{L}$ and $\mathcal{L}_*$ (and thus their difference $\mathcal{K} + \Pi$ as well) are adiabatic invariants.

The unavailable labile energy $\mathcal{L}_*$ has to be found from the variational problem of a conditional minimum of $\mathcal{L}$, as a functional of the spatial distributions of the independent thermodynamic variables, on condition that their variations satisfy the requirements of continuity and adiabaticity, for fixed values of the total mass, energy, entropy, and amount of thermodynamically active admixture. Let us assume (along with Lorenz, who considered dry air with equation of state (1.7) for $s = 0$) that in a state with $\mathcal{L} = \mathcal{L}_*$ all the thermodynamic parameters are constant along a streamline of the adiabatic motion, that is, they can be functions only of η and s . Designating these functions by subscript asterisks, we get

$$\Pi = \int [\rho \varepsilon - \rho_* \varepsilon_* - \Phi(\rho - \rho_*)] dV. \quad (5.14)$$

Let us take as an example a terrestrial atmosphere of dry air with equation of state (1.7) for $s = 0$. Its internal energy $\mathcal{I} = \int c_v T dm$ is estimated to be 8.6×10^{23} J (or for a column of unit cross section 1.7×10^9 J/m²); the potential energy $\mathcal{P} = \int gz dm$ in the quasistatic approximation is $\int RT dm$ (because then

$$\int gz \rho dz = - \int z dp = \int p dz = \int R \rho T dz)$$

and is estimated to be 3.6×10^{23} J (or 7.2×10^8 J/m²); the labile energy

$$\begin{aligned} \mathcal{L} &= \mathcal{I} + \mathcal{P} = \int (c_v + R) T dm \\ &= \int c_p T dm = \int c_p \Theta \left(\frac{p}{p_0} \right)^{(\kappa-1)/\kappa} dm. \end{aligned} \quad (5.15)$$

is thus 1.22×10^{24} J (or 2.42×10^9 J/m²). According to the rule given above, the value of $\mathcal{L}_*$ is obtained from (5.15) by replacing p by $p_*(\Theta)$, a pressure which is constant over an isentropic surface (introduced in formula (4.4) above). Then we

have

$$\Pi = \int c_p \Theta \left\{ \left(\frac{p}{p_0} \right)^{(\kappa-1)/\kappa} - \left[\frac{p_*(\Theta)}{p_0} \right]^{(\kappa-1)/\kappa} \right\} dm. \quad (5.16)$$

Oort (1964) estimated this quantity to be $5.5 \times 10^6 \text{ J/m}^2$, that is, only about 0.2% of $\mathcal{L}$ (while $\mathcal{K} = 1.5 \times 10^6 \text{ J/m}^2$, that is, about $\frac{1}{4}$ of Π). If the atmosphere is brought adiabatically to a state with neutral stratification (in which $\partial\Theta/\partial z = 0$), then the value $\mathcal{L}_1$ of the labile energy in this state is found from the right-hand side of (5.15) by replacing the factor $p^{(\kappa-1)/\kappa}$ in the integrand function by its average over the mass in the vertical air column $[(\kappa/(2\kappa-1))p_s^{(\kappa-1)/\kappa}]$, where p_s is the surface pressure. The quantity $\mathcal{L}_1$ is also an adiabatic invariant. The difference $\mathcal{L}_1 - \mathcal{L}$, known as the *macrostability parameter*, determines the amount of kinetic energy liberated or consumed during the adiabatic transition from the actual stratification of the atmosphere to the neutral stratification (for three-parameter media containing thermodynamically active admixtures, there are many neutral stratifications, and thus many macrostability parameters).

To conclude this section, we note that when studying the hydrodynamic stability of the motion of a medium it is useful to consider the adiabatic invariant $\mathcal{E} + \mathcal{F}$, where $\mathcal{F} = \int \varphi(\eta, \Omega_*) dm$ for function φ (which is quadratic in Ω_*) is some measure of the nonuniformity of the field of Ω_* . For adiabatic small perturbations of the motion, the principal part of functional $\mathcal{E} + \mathcal{F}$ remains invariant in time, and if it is positive-definite then the motion is stable with regard to small perturbations.

6. HYDRODYNAMIC SYSTEMS

In most cases practical solutions of specific fluid-dynamical problems are obtainable only if the *continuous* fields describing the state of the medium are replaced approximately by *discrete* (and finite) sets of parameters. For example, a continuously stratified medium can be approximated by a finite number of horizontal layers, with consideration of the thermohydrodynamic fields just at their boundaries. In general, for an Eulerian description of the motion, the space occupied by the fluid is divided into cells. The state of the medium is then characterized by the values of the thermohydrodynamic fields at the cell centers, these values at different points of a selected space grid being related with the aid of difference equations, derivable from the equations of fluid dynamics. Similarly, the space of the Lagrangian coordinates can also be divided into cells, attributing to the centers of the cells their total masses (method of *finite elements*) and deriving from the Lagrange equations of fluid dynamics expressions for the forces acting on the finite elements.

Finally, using *Galerkin's method* the thermohydrodynamic fields are represented approximately as linear combinations of a finite number of certain basis functions, while the coefficients of these linear combinations also form a finite set of parameters, describing the state of the medium (Galerkin, 1915). We will demonstrate this method with the aid of a boundary problem with initial values for

the Navier-Stokes equations, describing the motion of an incompressible homogeneous viscous fluid in a region of space V , bounded by a rigid surface S . In this case $\rho = \text{const}$, continuity equation (3.1) becomes $\text{div } \mathbf{v} = 0$, and equations of motion (3.2) are written as

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = - \frac{\nabla p}{\rho} + \nu \Delta \mathbf{v} + \mathbf{f}, \quad (6.1)$$

where, as in (5.2), ν is the kinematic viscosity and $\mathbf{f}$ is the acceleration of the external mass forces; the boundary conditions here have the form $\mathbf{v} = 0$ on S , and the initial values $\mathbf{v}(\mathbf{x}, 0) = \mathbf{v}_0(\mathbf{x})$. The solution $\mathbf{v}(\mathbf{x}, t)$ is sought in the functional space R of the solenoidal vector fields in a region V with zero values at surface S , in which we will use the 'scalar product' $[\mathbf{v}_1 \cdot \mathbf{v}_2] = \int \mathbf{v}_1 \cdot \mathbf{v}_2 \, dV$. Following Brushlinskaya (1964) and Ladyzhenskaya (1970), let us select in R as a basis the complete system of orthonormalized eigenfunctions $\psi_k(\mathbf{x})$ of the equation

$$\nu \Delta \psi_k = -\mu_k \psi_k; \quad k = 1, 2, \dots \quad (6.2)$$

and let us seek an approximate solution of equation (6.1) (n th Galerkin approximation) in the form

$$\mathbf{v}_n(\mathbf{x}, t) = u^k(t) \psi_k(\mathbf{x}); \quad k = \overline{1, n}, \quad (6.3)$$

where the summation from 1 to n is over the recurring indexes. Next we substitute (6.3) into (6.1), take the scalar product of both sides of the obtained equation times $\psi_i(\mathbf{x})$, and integrate over the region V . The term with the pressure gradient vanishes in view of the zero values on S of the functions of R , since $[\psi_i \times \nabla p] = \int p \mathbf{n} \cdot \psi_i \, dS = 0$ (so that the Galerkin procedure, like the curl operator, excludes from the equations of motion the pressure field), and we have

$$\begin{aligned} \dot{u}^i &= \frac{1}{2} \Gamma_{jk}^i u^j u^k - \nu \mu_i u^i + f^i; \quad i = \overline{1, n}; \\ \Gamma_{jk}^i &= -2[\psi_i \cdot (\psi_j \cdot \nabla) \psi_k]; \quad f^i = [\psi_i \cdot \mathbf{f}] \end{aligned} \quad (6.4)$$

(as an exception to the general rule, here the summation over i is *not* carried out). Consequently, 'parameters' $u^i(t)$ are found from a system of n ordinary differential equations (6.4) for the initial conditions $u^i(0) = [\psi_i \cdot \mathbf{v}_0]$. Below we will pay special attention to the case of zero viscosity and no external forces, in which the total kinetic energy $E = \frac{1}{2} \rho [\mathbf{v}_n \cdot \mathbf{v}_n] = \frac{1}{2} \rho \sum_i (u^i)^2$ should not vary with time, that is, $\partial E / \partial t = 0$; according to Equation (6.4), for $\nu = f^i = 0$ this condition becomes

$$\Gamma_{jk}^i u^i u^j u^k \equiv 0. \quad (6.5)$$

Moreover, in this case, as first shown by Lee (1952), for a reliable finite-dimensional approximation of the fluid-dynamical equations, the *regularity* condition is satisfied, that is, the phase volume $\partial u^k / \partial u^k = 0$ is conserved, just as for Hamilton systems (Liouville theorem).

Following Obukhov (1969) and his later works with coauthors, in order to generalize the above-mentioned properties of the Galerkin approximations of the fluid-dynamical equations, let us use the concept of *systems of the hydrodynamic type*. These are regular systems with n degrees of freedom $u^1, \dots, u^n$, with

quadratic-nonlinear equations of motion

$$\dot{u}^i = \frac{1}{2} \Gamma_{jk}^i u^j u^k \quad (6.6)$$

(where the *dynamic tensor* Γ_{jk}^i can be assumed symmetric with respect to its subscripts, so that it has $n^2(n+1)/2$ components), and with at least one quadratic positive-definite integral of motion (*energy*)

$$E = \frac{1}{2} g_{ik} u^i u^k \quad (6.7)$$

(where g_{ik} is a positive-definite *metric tensor*, which can be used to lower the indexes, for example, $g_{ik} \Gamma_{jk}^\lambda = \Gamma_{i,jk}$; the reciprocal tensor g^{ik} , satisfying the condition $g^{ik} g_{kl} = \delta_k^i$, where δ_k^i is the unit tensor, can be used to raise the indexes; the representation in which $g_{ik} = \delta_{ik}$, so that $E = \frac{1}{2} \sum_1 (u^i)^2$, called the *energy representation*, corresponds to some set of orthogonal coordinates $u^1, \dots, u^n$). The requirement of conservation of energy $\dot{E} \equiv 0$ because of the equations of motion imposes on the *dynamic tensor* $n(n+1)(n+2)/6$ conditions

$$\Gamma_{i,jk} + \Gamma_{j,ki} + \Gamma_{k,ij} = 0. \quad (6.8)$$

The requirement of regularity $\partial \dot{u}^i / \partial u^i = \Gamma_{ik}^i u^k \equiv 0$ gives n conditions

$$\gamma_k = \Gamma_{ik}^i = 0. \quad (6.9)$$

It is evident that the condition of regularity is necessary and sufficient for the invariance in time of the *canonical distribution* $f(\mathbf{u}) = C \exp(-E/\Theta)$ (where Θ is an analog of the temperature), describing the ‘white noise’. Thus regular systems (and these alone) are incapable of converting white noise into any other distribution. This property of systems of the hydrodynamic type is similar to the second law of thermodynamics.

For conditions (6.8) and (6.9) the dynamic tensor $\Gamma_{i,jk}$ has $N = n^2(n+1)/2 - n(n+1)(n+2)/6 - n = n(n^2-4)/3$ independent components. This number is positive only when $n \geq 3$, so that the least possible number of degrees of freedom for systems of the hydrodynamic type is three (for $n = 2$, putting aside regularity requirement (6.9) but making use of energy conservation condition (6.8), equations of motion (6.6) can with the aid of linear transformations be put into the form $\dot{u}^1 = (u^2)^2$, $\dot{u}^2 = -u^1 u^2$; however, here we have $\partial \dot{u}^i / \partial u^i = -u^1 \neq 0$, as a result of which the motion is nonperiodic, with u^1 increasing monotonically with time). As an example of a system of the hydrodynamic type with three degrees of freedom, let us cite the turbulent motion of an incompressible homogeneous ideal fluid inside an ellipsoid (with a surface S defined by the equation $S \equiv a_1^{-2} x_1^2 + a_2^{-2} x_2^2 + a_3^{-2} x_3^2 - 1 = 0$). In accordance with (3.8), for $\rho \equiv \text{const}$ and $\mathbf{f} = 0$, this motion is described by the *Helmholtz equation*

$$\begin{aligned} \text{helm } \mathbf{\Omega} &= 0; & \mathbf{\Omega} &= \text{curl } \mathbf{v}; \\ \text{div } \mathbf{v} &= 0; & (\hat{\mathbf{v}} \cdot \nabla) S &= 0. \end{aligned} \quad (6.10)$$

Let us introduce *uniform* (that is, with $\mathbf{\Omega} \equiv \text{const}$) ‘elliptical rotations’ about the principal axes of the ellipsoid, with the following velocity fields, depending linearly

on the space coordinates x_i :

$$\psi_i(\mathbf{x}) = -\frac{a_j}{a_k} x_k \mathbf{e}_j + \frac{a_k}{a_j} x_j \mathbf{e}_k, \quad (6.11)$$

where (i, j, k) are the cyclical permutations of indexes $(1, 2, 3)$, summation over repeating indexes not being carried out here, and $\mathbf{e}_i$ are the unit vectors, directed along the principal semiaxes of the ellipsoid. Let us also introduce here the so-called *Poincaré parameters* $\omega_i = a_j a_k (a_j^2 + a_k^2)^{-1} \Omega_i$, so that for $\omega_i = \text{const}$ the velocity fields $\mathbf{v} = \omega_i \psi_i(\mathbf{x})$ will be steady-state solutions of equations (6.10), the unsteady-state solutions being sought in the form

$$\mathbf{v}(\mathbf{x}, t) = \sum_{i=1}^3 \omega_i(t) \psi_i(\mathbf{x}). \quad (6.12)$$

After substituting this expression into the Helmholtz equation, for variables $\omega_i(t)$ we obtain a system of the hydrodynamic type of third order (a so-called *triplet*) with the equations:

$$I_i \dot{\omega}_i = (I_k - I_j) \omega_j \omega_k; \quad I_i = a_j^2 + a_k^2. \quad (6.13)$$

If we introduce vectors $\boldsymbol{\omega} = \omega_1 \mathbf{e}_1 + \omega_2 \mathbf{e}_2 + \omega_3 \mathbf{e}_3$ and $\mathbf{M} = I\boldsymbol{\omega}$, where I is a diagonal matrix with diagonal elements I_1, I_2, I_3 , then the equations of the triplet can be put into vector form:

$$\dot{\mathbf{M}} = \boldsymbol{\omega} \times \mathbf{M}. \quad (6.13')$$

Now, accurate to the replacement of $\boldsymbol{\omega}$ by $-\boldsymbol{\omega}$, this system of the hydrodynamic type is identical to the Euler equations for the motion of a gyroscope, with a tensor of the moments of inertia I , an angular velocity of rotation $\boldsymbol{\omega}$, and an angular momentum $\mathbf{M}$, in a reference system moving with the gyroscope. The Euler equations have two quadratic integrals of motion: the kinetic energy $E = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{M}$ and the square of the angular momentum $J = M^2$ (so that $\dot{E} = \dot{J} = 0$). For a 'fluid gyroscope' the kinetic energy $K = \frac{1}{2} m E$, where $m = \frac{4}{3} \pi a_1 a_2 a_3 \rho$ is the fluid mass inside the ellipsoid, and $J = \Pi^{-2} (\Gamma_1^2 + \Gamma_2^2 + \Gamma_3^2)$ is the sum of the squares of the circulations $\Gamma_i = \pi a_i^{-1} a_1 a_2 a_3 \Omega_i$ around the contours of the principal sections of the ellipsoid.

The triplet is not just an example, it is the general form of the third-order system of the hydrodynamic type (in the sense that each such system of the hydrodynamic type can be put into the form of a triplet by means of orthogonal transformations of its phase coordinates). Actually, for $n = 3$ the dynamic tensor of the system has $N = n(n^2 - 4)/3 = 5$ independent components $\Gamma_{i,jk}$, as well as a symmetric tensor of second rank A_k^i with a zero trace A_i^i . Having established a correspondence between these via the formula $A_k^i = \varepsilon^{ijl} \Gamma_{j,ik}$, where ε^{ijl} is a completely asymmetric tensor, and using an orthogonal transformation to put matrix A_k^i into diagonal form A'_{ij} , we see that $\Gamma'_{i,jk} = \frac{1}{3} (A'_{kk} - A'_{jj})$ (the indexes being defined as in (6.11)), that is, the system of hydrodynamic type is indeed a triplet. This general role played by the triplet accords it a special importance. We

will describe the *free* motion of triplet (6.13'), assuming to be more definite that $I_1 > I_2 > I_3$ and writing its invariants as

$$I_1^{-1}M_1^2 + I_2^{-1}M_2^2 + I_3^{-1}M_3^2 = 2E; \quad M_1^2 + M_2^2 + M_3^2 = J. \quad (6.14)$$

In axes M_1, M_2, M_3 these are the equations for the surfaces of an ellipsoid with semi-axes $\alpha_i = (2EI_i)^{1/2}$ and a sphere of radius $J^{1/2}$, so that the end of vector $\mathbf{M}$ moves along the line of intersection of these surfaces. If the sphere and ellipsoid touch at points $M = \alpha_1$ or $M = \alpha_3$, then at these points there will be reached, respectively, a maximum or minimum of energy, corresponding to steady *stable* rotations about axis M_1 or axis M_3 (they are stable, because for slight deviations from them trajectories $\mathbf{M}(t)$ are closed curves around the undisturbed points). For $M = \alpha_2$ surfaces (6.14) intersect along the *separatrix*, two two-dimensional curves (circles) intersecting at *saddle* points on axis M_2 , and this corresponds to unstable rotation around axis M_2 (since slight deviations from it cause the trajectories around the saddle point to move far away from it).

In the energy variables used just above (which we will denote as x_i), when the matrix with zero trace A'_{ij} is diagonal, let us write $p_i = \Gamma'_{i,jk}$, so that we will have $p_1 + p_2 + p_3 = 0$, with p_1 and p_3 positive and p_2 negative, that is, x_2 is an unstable mode. We introduce three integrals of motion $S_i = p_i^{-1}x_j^2 - p_k^{-1}x_k^2$, related by the expression $S_1 + S_2 + S_3 = 0$, or the integrals $\varepsilon_i = S_k - S_j$ equivalent to these, which we normalize using the formula $e_i = -(\frac{2}{3})^{1/3}p_1p_2p_3\varepsilon_i$. If we introduce $\gamma = 6^{-1/3}(e_1 - e_3)^{1/2}$, then we can solve the equations of the triplet in Jacobi elliptic functions:

$$x_1 = A_1 \operatorname{dn} \gamma t; \quad x_2 = A_2 \operatorname{cn} \gamma t; \quad x_3 = A_3 \operatorname{sn} \gamma t, \quad (6.15)$$

where the origin of time is so chosen that at $t = 0$ we have $x_3 = 0$, the amplitudes are defined by the formulas $A_i = 6^{-1/3}[(e_k - e_j)/p_jp_k]^{1/2}$, and the modulus k of the elliptic functions is defined by the formula

$$k^2 = \frac{e_2 - e_3}{e_1 - e_2} = \frac{(I_2 - I_3)(2EI_1 - J)}{(I_1 - I_2)(J - 2EI_3)}. \quad (6.16)$$

Functions (6.15) are periodic, with a period $4\gamma^{-1}K(k)$, where $K(k)$ is a complete elliptic integral of the first kind. For small k^2 (when $I_3 \approx I_2$) they degenerate into trigonometric functions ($\operatorname{sn} \tau \approx \sin \tau$; $\operatorname{cn} \tau \approx \cos \tau$; $\operatorname{dn} \tau \approx 1$), and for $k^2 \approx 1$ the trajectories of the system are close to the separatrix. Figure 6.1 shows plots of functions (6.15) for $k^2 = 0.4$ and $k^2 = 0.999$.

Forced motion of the triplet under laboratory conditions, whereby the fluid inside the ellipsoid uncoils under the influence of some field of the moment of the external forces (for instance, in the case of an electrical conductor like mercury, the Lorentz force (3.3)), is difficult to achieve. This is because ensuring a linear dependence (6.11) of the velocity field on the space coordinates hinders the adhesion of the fluid to the walls and the creation of a boundary layer at the walls. However, if the field of the moment of the external forces is a linear function of the space coordinates, then the boundary layer only slows down a fluid core of ellipsoidal shape with an approximately linear internal velocity distribution.

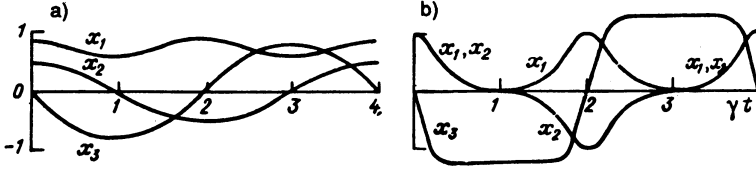


Fig. 6.1. Plots of functions (6.15) for $k^2 = 0.4$ (a) and 0.999 (b).

Assuming this friction to be proportional to the angular velocity of rotation of the core, we can describe this rotation with the aid of the equation of forced motion of a triplet

$$\dot{\mathbf{M}} = \boldsymbol{\omega} \times \mathbf{M} - \lambda \mathbf{M} + \mathbf{F} \quad (6.17)$$

with $\lambda = \text{const}$ and $\mathbf{F} = \text{const}$. To be more definite let us set $a_1 > a_2 > a_3$ and let us confine ourselves to the most interesting case: when the moment of the external forces acts in the direction of unstable rotation, that is, $\mathbf{F} = (0, F, 0)$ (and, when $I_2 = (I_1 + I_3)/2$, it can be shown that the energy of the unstable mode will be imparted most rapidly to the other modes). If the following new variables and parameters are introduced into (6.17):

$$\begin{aligned} u_1 &= I_2^{1/2} \omega_2; & u_2 &= \left(\frac{I_1}{2}\right)^{1/2} \omega_1 + \left(\frac{I_3}{2}\right)^{1/2} \omega_3; \\ u_3 &= \left(\frac{I_3}{2}\right)^{1/2} \omega_3 - \left(\frac{I_1}{2}\right)^{1/2} \omega_1; & f &= I_2^{-1/2} F; \end{aligned} \quad (6.18)$$

$$L = p^{-1} = 2(a_1^2 + a_2^2)^{1/2}(a_2^2 + a_3^2)^{1/2}(a_3^2 + a_1^2)^{1/2}(a_1^2 - a_3^2)^{-1},$$

then Equations (6.17) become:

$$\begin{aligned} \dot{u}_1 &= p(u_3^2 - u_2^2) - \lambda u_1 + f; & \dot{u}_2 &= p u_1 u_2 - \lambda u_2; \\ \dot{u}_3 &= -p u_1 u_3 - \lambda u_3. \end{aligned} \quad (6.19)$$

These have three steady-state solutions. One of these, let us call it solution *A*, has the form $u_1 = f/\lambda$, $u_2 = u_3 = 0$. The other two, which we will call *B* and *B'*, for $f > 0$ have the form $u_1 = \lambda/p$, $u_2 = \pm (f/p - \lambda^2/p^2)^{1/2}$, $u_3 = 0$ and for $f < 0$ are $u_1 = -\lambda/p$, $u_2 = 0$, $u_3 = \pm (|f|/p - \lambda^2/p^2)^{1/2}$. A study indicates that their stability depends on the effective Reynolds number $\text{Re} = |fp\lambda^{-2}|$. For $\text{Re} < 1$ there is only one steady-state solution *A*, and it is stable. For $\text{Re} = 1$ the three solutions are identical. For $\text{Re} > 1$ solution *A* is unstable, and any disturbance will 'flip' the triplet from state *A* to one of the two stable states *B* or *B'* (depending on the sign of the initial disturbance).

Let us consider the case $f > 0$ (the opposite case can be analyzed similarly). Then for an initial value $u_3 = 0$ we get $u_3 \equiv 0$, and after renormalization $pu_1/\lambda \rightarrow u_1$, $pu_2/\lambda \rightarrow u_2$, $\lambda t \rightarrow t$ Equations (6.19) become:

$$\dot{u}_1 = \text{Re} - u_1 - u_2^2; \quad \dot{u}_2 = -u_2 + u_1 u_2. \quad (6.20)$$

In phase space u_1, u_2 (Figure 6.2) their steady-state solutions are depicted by points A ($u_1 = \text{Re}, u_2 = 0$, at the vertex of the parabola $u_1 = \text{Re} - u_2^2$), B , and B' ($u_1 = 1, u_2 = \pm(\text{Re} - 1)^{1/2}$, at the intersections of the parabola with the line $u_1 = 1$). For $\text{Re} < 1$ point A is a node, and all the phase trajectories are directed toward it. With increasing Re , after bifurcation for $1 < \text{Re} < 9/8$ the phase trajectories are pushed away from A and are attracted to nodes B or B' (depending on the initial sign of u_2). After a new bifurcation for $\text{Re} > 9/8$ nodes B and B' change into *foci*, so that instead of monotonic motion toward points B or B' the system performs damped nonlinear oscillations about these points.

Now let us look at the case where, in addition to the regular external force $\mathbf{F} =$

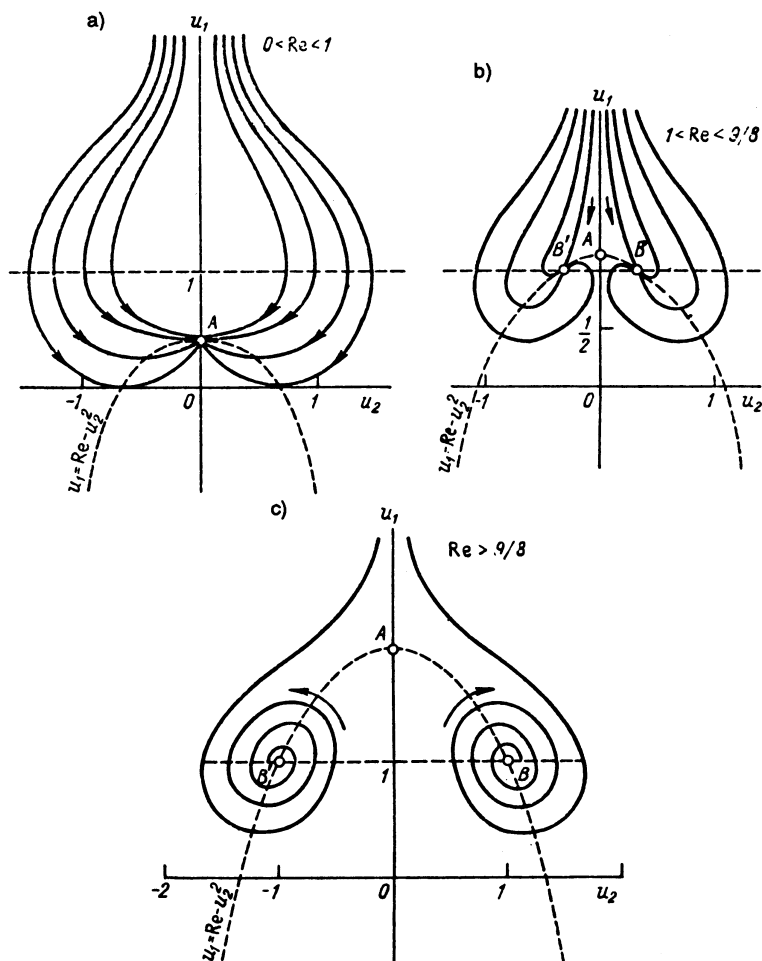


Fig. 6.2. Phase portraits of fluid gyroscope (6.20): (a) for $0 < \text{Re} < 1$, (b) for $1 < \text{Re} < 9/8$, (c) for $\text{Re} > 9/8$.

$(0, F, 0)$, random forces $\mathbf{F}'(t)$ also act on the triplet. If these take the form of 'white noise', then Equations (6.19), after the renormalization used in (6.20), become:

$$\begin{aligned}\dot{u}_1 &= \text{Re} - u_1 - u_2^2 + u_3^2 + f'_1(t); \\ \dot{u}_2 &= -u_2 + u_1 u_2 + f'_2(t); \quad \dot{u}_3 = -u_3 - u_1 u_3 + f'_3(t),\end{aligned}\tag{6.21}$$

where the mathematical expectation of the random forces $\langle f'_i(t) \rangle$ is zero, and their correlation function is $\langle f'_j(t + \tau) f'_j(t) \rangle = 2\sigma^2 \delta_{ij} \delta(\tau)$.

If $\text{Re} > 1$, then after averaging over the ensemble of random-force realizations for component u_2 we obtain a steady-state probability distribution with maxima in the vicinities of the points $u_2 = \pm(\text{Re} - 1)^{1/2}$ and a minimum at the point $u_2 = 0$, corresponding to an unstable state. If for some realization $\mathbf{f}'(t)$ the system has arrived at one of the most probable states, then after some time T during which quite high values of $|\mathbf{f}'|$ have appeared the system may cross over into another most probable state (Obukhov has called this the *crossover phenomenon*; it may play a major role in geophysics). The period T should be longer, the lower the value of σ^2 (that is, the more infrequently large values of $|\mathbf{f}'|$ are observed) and the higher the value of Re (that is, the greater the distance between the most probable values of u_2). In the examples of numerical integration of Equations (6.21) considered by Obukhov and his coauthors, with $\text{Re} = 6$ for $\sigma = 0.01$ no crossover was observed, whereas for $\sigma = 0.1$ it took place after a time $T \sim 10$.

Therefore, under the influence of weak noise (thermal noise, for instance) crossover in the triplet is virtually impossible. As Obukhov and his coauthors have shown, this conclusion holds true if the linear friction $\lambda \mathbf{M}$ in (6.17) is replaced by the square-law friction. At the same time, their laboratory experiments with a fluid gyroscope demonstrated that crossover takes place after a proportioned (that is, for some specific interval T_1) deviation of an external force $\mathbf{F}$, produced by an agitator rotating about the central axis of the ellipsoid; T_1 was found to be shorter, the more rapid the rotation (for an analytical description of the crossover, a system of the hydrodynamic type more complicated than a triplet had to be used).

It should be noted that, if the external force is switched on and off with a period $T_2 \lesssim T_1$, then the sequence of crossovers will have a period $2T_2$. This *doubling of the period* is typical of the phenomenon of *parametric resonance*, appearing in some oscillating systems for a periodic variation of their parameters as a result of external factors. A typical example of this is a pendulum with an oscillating (at a period $T = 2\pi/\omega$) suspension, the motion of which is described by the Mathieu equation

$$\ddot{x} + a^2(1 + \varepsilon \cos \omega t)x = 0.\tag{6.22}$$

For $\varepsilon \ll 1$ and $a \approx \omega/2$ this equation has periodic solutions with a period $2T$. Obukhov proposes using a similar mechanism to explain the quasibiennial cyclicity of the circulation of the atmosphere, which manifests itself most clearly in the reversal of the direction of the zonal (latitudinal) wind in the equatorial stratosphere. The 'crossover' to the opposite phase of the biennial cycle (or, on the other

hand, the omission of the next annual ‘crossover’) occasionally occurring in nature, leads to a slight increase in the *mean* period of this cyclicity (up to 26 months).

A system of the hydrodynamic type which is more complicated than a triplet can be constructed by superposing triplets. In this case we define the dynamic tensor:

$$\Gamma_{i,jk} = \sum_{\alpha=1}^m c_{\alpha} \Gamma_{i,jk}^{(\alpha)}, \quad (6.23)$$

where $\Gamma_{i,jk}^{(\alpha)}$ are the dynamic tensors of the triplets, the components of which are nonzero only for values of indexes i, j, k corresponding to the numbers of coordinates u^i of the appropriate triplet. As an example, let us consider the magnetohydrodynamic generalization of Helmholtz equation (6.10) when there is a superconducting fluid inside the ellipse:

$$\text{helm } \mathbf{\Omega} = \text{curl } \mathbf{f}_L; \quad \text{helm } \mathbf{H} = 0, \quad (6.24)$$

where $\mathbf{f}_L$ is the Lorentz acceleration (3.3). If we use for the magnetic-field strength $\mathbf{H}$ an expression analogous to (6.12), with Poincaré parameters $h_i(t)$, then instead of the triplet equation (6.13') we get

$$\dot{\mathbf{M}} = \boldsymbol{\omega} \times \mathbf{M} + \mathbf{j} \times \mathbf{h}; \quad \dot{\mathbf{h}} = \boldsymbol{\omega} \times \mathbf{h}, \quad (6.25)$$

where $\mathbf{j} = I\mathbf{h}$, and I being the same diagonal matrix as in the definition $\mathbf{M} = I\boldsymbol{\omega}$. The sixth-order system of the hydrodynamic type (6.25) can be represented as a superposition of four triplets $(\omega_1, \omega_2, \omega_3)$, (ω_1, h_2, h_3) , (ω_2, h_1, h_3) , and (ω_3, h_1, h_2) :

$I_1 \dot{\omega}_1 =$	$(I_3 - I_2)\omega_2\omega_3$	$+ (I_3 - I_2)h_2h_3$		
$I_2 \dot{\omega}_2 =$	$(I_1 - I_3)\omega_3\omega_1$		$+ (I_1 - I_3)h_3h_1$	
$I_3 \dot{\omega}_3 =$	$(I_2 - I_1)\omega_1\omega_2$			$+ (I_2 - I_1)h_1h_2$
$\dot{h}_1 =$				$+ h_2\omega_3$
$\dot{h}_2 =$		$h_3\omega_1$	$- h_3\omega_2$	$- h_1\omega_3$
$\dot{h}_3 =$		$- h_2\omega_1$	$+ h_1\omega_2$	

Here the second, third, and fourth triplets are linked to the first triplet and also to one another. Simpler, nonmultiple linkages turn out to be sufficient for modeling a cascade process of transfer of kinetic energy from large-scale motions to small-scale motions.

Let us assume that an external force f supplies energy to the unstable mode u_{11} of the first triplet (the first index of the mode indicates the number of the ‘stage’ in the cascade, and the second indicates the number of the mode at this stage). This

energy can be imparted to either one or both stable modes of the first triplet (u_{21} and u_{22}). These become unstable modes of the next two triplets and may impart energy to the stable modes of the latter, that is, $u_{21} \rightarrow (u_{31}, u_{32})$, $u_{22} \rightarrow (u_{33}, u_{34})$, etc. Therefore, we obtain a chain of equations:

$$\begin{aligned}\dot{u}_{11} &= f - \lambda_1 u_{11} + p_1(u_{21}^2 - u_{22}^2); \\ \dot{u}_{i,2k-1} &= -\lambda_i u_{i,2k-1} - p_{i-1} u_{i-1,k} u_{i,2k-1} + \\ &\quad + p_i(u_{i+1,4k-3}^2 - u_{i+1,4k-2}^2); \\ \dot{u}_{i,2k} &= -\lambda_i u_{i,2k} + p_{i-1} u_{i-1,k} u_{i,2k} + p_i(u_{i+1,4k-1}^2 - u_{i+1,4k}^2),\end{aligned}\tag{6.26}$$

where $i = 2, 3, \dots$, and $1 \leq k \leq 2^{i-1}$, while the wave numbers p_i and the friction factors λ_i increase with an increase in their order number (for example, $\lambda_i \sim \nu p_i^2$). In such a branching n -stage system it is possible to excite 2^n equally probable chains, but between them at each stage crossovers are possible, and if the crossover occurs at an early stage, then all the subsequent chains disappear (this leads to the phenomenon of intermittence at late stages). If $\varepsilon = f u_{11}$ is the supplied power and $p_{i+1}/p_i = q = \text{const}$, then in a fixed excited chain, neglecting friction at earlier stages, we arrive at the steady-state solution

$$u_i = q^{2/3} \varepsilon^{1/3} p_i^{-1/3},\tag{6.27}$$

corresponding to the Kolmogorov—Obukhov spectrum for the kinetic energy of a developed turbulent flow ($p_i^{-1} u_i^2 \sim \varepsilon^{2/3} p_i^{-5/3}$). By analogy with (6.26), we can construct a multistage system of the hydrodynamic type with double linkage, in which each i th stage is linked to the two preceding stages, $(i-1)$ and $(i-2)$, and the two subsequent stages, $(i+1)$ and $(i+2)$, and in which neglecting friction there exist two quadratic integrals of motion: the total energy and the mean square of the two-dimensional vorticity, the so-called *enstrophy* (Gledzer, 1973). In such a system of the hydrodynamic type, for excitation of any intermediate stage enstrophy is transferred to the later stages, forming a spectrum of energy $p_i^{-1} u_i^2 \sim \varepsilon_\Omega^{2/3} p_i^{-3}$, while energy is transferred to the earlier stages, forming a spectrum $\varepsilon_\Omega^{2/3} p_i^{-5/3}$. We will return to such processes in subsequent chapters.

Developing Obukhov's theory of systems of the hydrodynamic type constituted an important stage in the systematic formulation of the general fundamentals of nonlinear mechanics (or the theory of the nonlinear equations of mathematical physics). The next stage could be the study of so-called *symmetrized* systems (Obukhov, 1977), generalizing the concept of a *Hamilton* system, the equations of motion of which can be written as

$$\dot{u}^i = S^{ij} \frac{\partial F}{\partial u^j},\tag{6.28}$$

where F is the *generating function* (a generalization of the Hamiltonian), and the *symmetrizer* $S = (S^{ij})$ is a nonsingular square matrix.

If $A = (\partial \dot{u}^i / \partial u^j)$ is the *stability matrix*, then symmetrization of a dynamic system is equivalent to the existence of a nonsingular real matrix S such that the matrix $S^{-1}A$ is symmetric. Symmetrized systems, for which we can introduce a

symmetric (or skew-symmetric) symmetrizer, are called θ systems (or H systems). H systems exist only in even-numbered phase spaces R_{2n} , and in such spaces there will always exist linear coordinates $p_1, \dots, p_n; q_1, \dots, q_n$, in which the H system reduces to canonical Hamilton form $p_i = -\partial H/\partial q_i, \dot{q}_i = \partial H/\partial p_i$.

For quadratic-nonlinear systems we can define the *characteristic quadratic form* $B = (\frac{1}{2})(\partial \dot{u}^i/\partial u^j)(\partial \dot{u}^j/\partial u^i)$, the matrix of which is denoted by the same letter $B = (B_{ij})$; a quadratic-nonlinear regular θ system with a symmetrizer $S \equiv B^{-1}$ is called a B system. The simplest example of a symmetrized system (θ system or B system) is the canonical triplet (6.13); it has a generating function $F \sim \omega_1 \omega_2 \omega_3$. Another example of a θ system or B system is a multistage system of the hydrodynamic type with simple linkage (6.26). For a θ system with the specified symmetrizer we can introduce an operator $\mathcal{L} = S^j \partial^2/\partial u^i \partial u^j$, similar to the Laplacian, so that we will have $\partial \dot{u}^i/\partial u^i = \mathcal{L}F$, and thus for regular θ systems $\mathcal{L}F = 0$, that is, the generating function F is harmonic for operator $\mathcal{L}$. On the other hand, if we assume that $s(\mathbf{u}, \mathbf{v})$ is a bilinear form with a matrix S^{-1} and that $\Theta(\mathbf{u}) = \frac{1}{2}s(\mathbf{u}, \mathbf{v})$ is the corresponding quadratic form, then it is easy to see that $\Theta = u^j \partial F/\partial u^j$, so that if $F(\mathbf{u})$ is a homogeneous function of degree m then it will equal Θ/m . Let us assume further that for the generating functions we can define a bilinear operation

$$[F, G](\mathbf{u}) = S^j \left(\frac{\partial F}{\partial u^i} \right) \left(\frac{\partial G}{\partial u^j} \right) = [G, F](\mathbf{u}), \quad (6.29)$$

analogous to the Poisson bracket for the Hamiltonians. Thus, if G is the first integral of a θ system with a generating function F (that is, $\dot{G}_F \equiv 0$), then F is the first integral of the θ system with a generating function G (and with the same symmetrizer). Finally, let us consider θ systems with a quadratic positive-definite integral of motion E , which we will specify in an energy representation (the halfsum of the squares of coordinates u^j) such that the quadratic form $\Theta(\mathbf{u})$ is referred to the principal axes (the diagonal symmetrizer, the diagonal elements of which let us call $\theta_1, \dots, \theta_n$). Invariance of E here places constraints on F ; for instance, for quadratically nonlinear systems we have

$$F = \sum_{2\theta_i + \theta_j = 0} c_{ij}(u^i)^2 u^j + \sum_{\theta_i + \theta_j + \theta_k = 0} c_{ijk} u^i u^j u^k \quad (6.30)$$

(if, moreover, $\sum_j c_{ij} \equiv 0$, then the system is regular, that is, it is a θ system of the hydrodynamic type). In this case the quadratic form $B(\mathbf{u})$ can also be referred to the principal axes (for which in general matrices B and Θ must be commutative, which places constraints on the symmetrizer).

Small Oscillations

7. TYPES OF ADIABATIC MOTIONS

All the motions of a fluid during an adiabatic process are wave motions. In order to classify these, it is sufficient to consider waves of small amplitude, that is, small oscillations of the medium relative to some *equilibrium* state. As such a state, let us take the state of *relative* rest, in which the medium as a whole rotates at some constant angular velocity $\boldsymbol{\omega}$, while its thermodynamic characteristics (denoted by subscript zero) depend only on the vertical coordinate $z = r - r_0$ (where r is the distance from the center of gravity) and are related by the hydrostatic equation $\nabla p_0 = \rho_0 \mathbf{g}$ (where $\mathbf{g} = -g \nabla r$ is the free-fall acceleration, and ∇r is the unit vector in the vertical direction). The square of the Brunt–Väisälä frequency (2.4) in the equilibrium state will be called N_0^2 , and the equilibrium gradient of the admixture concentration $\partial s_0 / \partial z = \Gamma_0$. The relative velocity $\mathbf{u}$ and the deviations of the thermodynamic characteristics from their equilibrium values (denoted by primes) will be assumed to be small. We will linearize (relative to these quantities) the hydrothermodynamical equations of the adiabatic processes (obtained from the equations of Sections 1 and 3 neglecting in them all nonadiabatic factors). The linearized equations of continuity and motion (3.1), (3.2) then become:

$$\frac{\partial \rho'}{\partial t} + u_r \frac{\partial \rho_0}{\partial z} = -\rho_0 \operatorname{div} \mathbf{u}; \quad (7.1)$$

$$\frac{\partial \mathbf{u}}{\partial t} + 2\boldsymbol{\omega} \times \mathbf{u} = -\frac{\nabla p'}{\rho_0} - g \frac{\rho'}{\rho_0} \nabla r = -\frac{\nabla p''}{\rho_{00}} + g\eta'' \nabla r; \quad (7.2)$$

where the second equation in (7.2) is the same as (3.5). Linearization of the adiabaticity equation (1.17) and the condition of admixture conservation (1.16) for $\varepsilon_s = 0$ gives the following equations for η'' and s' :

$$\frac{\partial \eta''}{\partial t} + \frac{N_0^2}{g} u_r = 0; \quad (7.3)$$

$$\frac{\partial s'}{\partial t} + \Gamma_0 u_r = 0. \quad (7.4)$$

At the rigid walls confining the fluid the normal component of the velocity goes to zero. If the fluid has a free surface $r = r_0 + \zeta$, where ζ is the perturbation of its

level, then at $r = r_0$ the linearized kinematic and dynamic boundary conditions (5.9) and (5.11) are satisfied. In this case they become

$$\frac{\partial \zeta}{\partial t} = \hat{u}_r; \quad \frac{\partial \hat{p}'}{\partial t} - g\hat{\rho}_0 \hat{u}_r = -\gamma_0 \rho_0 \frac{\partial \Delta \zeta}{\partial t} \quad (7.5)$$

(the second of these equations is obtained by individual differentiation with respect to t and linearization of the boundary condition $\hat{p} - p_a \approx -\gamma\hat{\rho}\Delta\zeta$ for $p_a \approx \text{const}$). Note that if, in addition to η and s , some adiabatic Lagrangian invariant A exists (for instance, the potential vorticity), that is, for adiabatic processes $dA/dt = 0$, then after linearization this equation becomes $\partial A'/\partial t + u_r \partial A_0/\partial z = 0$. Now, after finding the value of u_r from (7.3), (7.4), we see that the partial time derivatives of $gN_0^{-2}\eta''$, $\Gamma_0^{-1}s'$, and $(\partial A_0/\partial z)^{-1}A'$ are equal to one another. Thus the differences of these quantities are *linear invariants* (that is, they are time-independent linear combinations of the unknown functions). If any of the quantities N_0 , Γ_0 , or $\partial A_0/\partial z$ is equal to zero (making the equilibrium state relative to this characteristic *uniform*), then the invariant becomes η'' , s' , or A' . For $r = r_0$ there will also exist the boundary invariants $\hat{p}' - g\hat{\rho}_0\zeta + \gamma_0\rho_0 \Delta\zeta$ and differences between ζ and the boundary values $gN_0^{-2}\eta''$, $\Gamma_0^{-1}s'$, and $(\partial A_0/\partial z)^{-1}A'$ (the boundary values of the last three quantities at *horizontal* rigid walls are also invariants).

Finally, let us integrate over the volume V (bounded at the top by the equilibrium free surface $r = r_0$) the local energy equation, following from (7.1)–(7.3):

$$\frac{1}{2} \frac{\partial}{\partial t} \left(\rho_0 u^2 + \frac{p'^2}{\rho_0 c_0^2} + \rho_0 \frac{g^2 \eta'^2}{N_0^2} \right) + \text{div } p' \mathbf{u} = 0 \quad (7.6)$$

to obtain a quadratic integral invariant, the total energy of the perturbations

$$\begin{aligned} \mathcal{E} = \frac{1}{2} \iint \left[u^2 + \left(\frac{p'}{\rho_0 c_0} \right)^2 + \left(\frac{g\eta''}{N_0} \right)^2 \right] \rho_0 dV + \\ + \frac{1}{2} \int [g\zeta^2 + \gamma_0(\nabla\zeta)^2] \hat{\rho}_0 d\Sigma, \end{aligned} \quad (7.7)$$

where the first term in the volume integral is the *kinetic energy*, the second term is the *elastic energy*, and the third term is the so-called *thermobaric energy* (that is, the portion of the potential energy associated with the buoyancy forces acting on the fluid particles, which are moved along the vertical from their equilibrium position); the integral over the surface (of $\hat{p}'$, $\hat{u}_r$, where $\hat{p}' = g\hat{\rho}_0\zeta - \gamma_0\rho_0 \Delta\zeta$, and $\hat{u}_r = \partial\zeta/\partial c$) corresponds to the potential energy of its perturbations ζ , the first term being the gravitational energy and the second being due to the surface tension (it is obtained from (5.12) for small $(\Delta\zeta)^2$).

Let us consider the possible steady-state solutions of equations (7.1)–(7.4), which describe the so-called *motions of the first kind*. These solutions possess the

following four properties: (G₁) the distribution of density perturbations ρ' is described by the law

$$\rho' = -(g\omega_r)^{-1} \boldsymbol{\omega} \cdot \nabla \rho' \quad (7.8)$$

(obtained from (7.2) via scalar multiplication by $\boldsymbol{\omega}$), so that it is nonhydrostatic; it becomes hydrostatic only in the so-called *traditional approximation*, whereby $\boldsymbol{\omega}$ is replaced by $\omega_r \nabla r$. The hydrostatic approximation, on the other hand, in which the vertical projection of equation (7.2) is written as $\nabla p' = -g\rho' \nabla r$, must be supplemented by the traditional approximation in order for the work of the Coriolis force $\rho_0 \mathbf{u}(2\boldsymbol{\omega} \times \mathbf{u})$ to remain zero and for the law of energy conservation not to be violated. In this case, however, the equations will contain an artificial singularity at the equator. If one of the quantities N_0 and Γ_0 is not zero, then the motions are *horizontal* (G₂), that is, $u_r = 0$; *geostrophic* (G₃), that is, an equilibrium exists between the pressure gradient and the Coriolis force for horizontal motions $2\boldsymbol{\omega} \times \mathbf{u}_k = -\rho_0^{-1} \nabla p'$, so that

$$\mathbf{u} = \mathbf{u}_G = (2\omega_r \rho_0)^{-1} \nabla r \times \nabla p', \quad (7.9)$$

and *nondivergent* (G₄), that is, $\text{div } \mathbf{u} = 0$.

The combined effect of these properties is to make the pressure perturbations p' independent of the longitude λ . Thus it turns out that they may be an arbitrary function $p'_s(r, \theta)$ and, according to (7.9), $u_\theta = 0$, that is, the flows are directed along latitude circles, with a geostrophic velocity $u_\lambda = (2r\omega_r \rho_0)^{-1} \partial p'_s / \partial \theta$; such flows will be possible only in models with rigid walls that do not vary along the longitude, and thus are impossible for actual land-surface and ocean-floor reliefs (where, accordingly, steady-state perturbations must be nonadiabatic). If $N_0 = \Gamma_0 = 0$ (thermodynamically uniform equilibrium state), then the requirement $u_r = 0$ is not obligatory, and steady-state motions with a velocity $\mathbf{u} = \mathbf{u}_G + u_r \boldsymbol{\omega} / \omega_r$ are also possible; these motions satisfy the conditions $u_n = 0$ at the rigid walls and $u_r = 0$ at $r = r_0$, while p' and u_r are related by the expression:

$$\text{div } \frac{u_r \boldsymbol{\omega}}{\omega_r} - \frac{g u_r}{c_0^2} = (2r^2 \rho_0 \cos^2 \theta)^{-1} \frac{\partial p'}{\partial \lambda}. \quad (7.10)$$

With regard to unsteady-state solutions, we note that if the perturbations of the fields $\mathbf{u}$, p , ρ at the initial moment $t = 0$ are so *adapted* to one another that they possess all the above-mentioned properties of steady-state solutions (for a thermodynamically nonuniform equilibrium state these are properties (G₁)–(G₄)), then they will not vary with time. In the opposite case the solution of the Cauchy problem for Equations (7.1)–(7.4) will be the sum of the steady-state solution, defined as the adapted part of the initial perturbations (completely characterized by the linear invariants found on the basis of the initial data, which for a thermodynamically nonuniform equilibrium state is the mean zonal part of the initial pressure field), plus the unadapted unsteady-state perturbations (with the linear invariants equal to zero). Let us call these *motions of the second kind* or *proper motions*. These, as already noted, have the nature of waves, and if the initial perturbations were concentrated in a limited region of space V , then with the passage of time these waves will leave region V , and inside V their energy density

will diminish. Ultimately, the only waves remaining there will be adapted hydrodynamic waves, possessing properties (G_1) – (G_4) ; this process can be called *adaptation* of the hydrodynamic fields (the adaptation is carried out completely only in certain regions of *unbounded space*, while, for instance, in the spherical layers a reverberation background remains from the waves; however, in reality this background becomes damped because of the effects of certain dissipative factors not allowed for here).

In order to describe the proper motions, we will use operational calculus, with $D = (2\omega^{-1}) \partial/\partial t$ as the dimensionless operator of differentiation with respect to time. We also define $\mathbf{k} = \omega^{-1}\boldsymbol{\omega}$ as a unit vector along the rotation axis of the reference system, for instance, the Earth (in the traditional approximation $k \rightarrow k_r$, ∇r , $k^2 \rightarrow k_r^2$).

Let us introduce a linear (tensor) operator $\mathbf{L}$, inherent in a homogeneous rotating medium, and let us assume that for each vector field $\mathbf{A}$

$$\mathbf{L}(\mathbf{A}) = (D^2 + k^2)^{-1} [D^2\mathbf{A} - D\mathbf{k} \times \mathbf{A} + (\mathbf{k} \cdot \mathbf{A})\mathbf{k}]. \quad (7.11)$$

In a thermodynamically inhomogeneous medium (the inhomogeneity of which will be characterized by the dimensionless Brunt–Väisälä frequency $N = N_0/2\omega$) this operator has to be replaced by a more complicated operator $\mathbf{M}$:

$$\mathbf{M}(\mathbf{A}) = \mathbf{L}(\mathbf{A}) + Q^{-1}N^2\mathbf{L}(\nabla r)L_r(\mathbf{A}); \quad Q = -[D^2 + N^2L_r(\nabla r)] \quad (7.12)$$

(note that in the traditional approximation $L_r(\nabla r) \rightarrow 1$). Using these notations, we obtain from Equations (7.1)–(7.3) the following equations in $\mathbf{u}$ and $P = p''/2\omega$:

$$\begin{aligned} \mathbf{u} + \frac{\mathbf{M}(\nabla P)}{\rho_{00}D} &= \frac{\mathbf{M}(\mathbf{u}_m)}{D} + \frac{g\eta''_m\mathbf{M}(\nabla r)}{2\omega D^2}; \\ \text{div } \mathbf{u} - \frac{gu_r}{c_0^2} + \frac{4\omega^2 DP}{\rho_{00}c_0^2} &= \frac{2\omega p'_m}{\rho_0 c_0^2}, \end{aligned} \quad (7.13)$$

where the subscript m denotes the nonadapted part of the initial values. If this part is nonzero only in some confined region V , which is nowhere contiguous with the boundaries of the fluid, then the fluid motions will not be related by the boundary conditions, unless they reach these boundaries. For example, we consider for a uniform equilibrium state ($N = 0$, $\mathbf{M} = \mathbf{L}$, and η'' is an invariant, which we will set equal to zero, so that $p' = c_0^2 \rho'$) proper motions, in which pressure perturbations are absent (such motions can be called *inertial motions*). According to (7.13), inertial motions appear if the initial perturbations are related by the condition $\text{div } \mathbf{L}(\mathbf{u}_m) = gc_0^{-2} L_r(\mathbf{u}_m)$, which according to definition (7.11) can, since $\mathbf{u}_m$ is independent of D , be divided into three conditions: fields $\mathbf{u}_m$, $\mathbf{k} \times \mathbf{u}_m$, and $(\mathbf{k} \cdot \mathbf{u}_m)\mathbf{k}$ must satisfy the equation $\text{div } \mathbf{A} = gc_0^2 A_r$. Nonadapted fields $\mathbf{u}_m$ with these properties must be orthogonal to $\mathbf{k}$ and thus cannot satisfy the boundary conditions at arbitrary rigid walls. Consequently, inertial motions are possible only locally. This accounts for the observed intermittence and rapid damping of the spatial coherence of the inertial oscillations in the ocean (in the atmosphere, on the other hand, where except for large mountain massifs no horizontal boundaries exist, the intermittence of the inertial oscillations should be weaker).

Periodic proper motions constitute *eigen oscillations* of the entire medium as a whole, in which case fields $\mathbf{u}$, p' , ρ' , s' depend on time according to an $\exp(2i\omega ft)$ law and must satisfy all the boundary conditions (here f are the eigenfrequencies, measured in units of the semidiurnal frequency 2ω ; due to the invariance of the total energy (7.7), all these are real). The equations for the complex amplitudes of the eigen oscillations $\mathbf{u}$ and P are obtained from Equations (7.13) by replacing their right-hand sides by zeros and the operator D by its eigenvalue if , to yield:

$$\mathbf{u} = \frac{i}{f\rho_{00}} \mathbf{M}(\nabla P); \quad (7.14)$$

$$\operatorname{div} \frac{\mathbf{M}(\nabla P)}{\rho_{00}} - \frac{gM_r(\nabla P)}{\rho_{00}c_0^2} + \frac{4\omega^2 f^2 P}{\rho_{00}c_0^2} = 0. \quad (7.15)$$

The boundary condition for Equation (7.15) when $r = r_0$ is

$$4\omega^2 f^2 \hat{P} = g\hat{M}_r(\nabla P) - \gamma_0 \Delta \hat{M}_r(\nabla P). \quad (7.16)$$

It follows from (7.14) and (7.15) that inertial eigen oscillations (without pressure perturbations) do not exist. The space variables r and θ in Equation (7.15) are not separated, due to the presence of multiplier Q in operator $\mathbf{M}$. An exception is the case of a uniform equilibrium state ($N = 0$, $\mathbf{M} = \mathbf{L}$), in which, assuming for simplicity the Boussinesq approximation $\operatorname{div} \mathbf{u} = 0$ and filtering out acoustic waves, instead of (7.15) for P we obtain an equation of *gyroscopic waves*

$$[f^2 \nabla^2 - (\mathbf{k} \cdot \nabla)^2]P = 0, \quad (7.17)$$

having particular solutions with separated variables

$$P = r^n e^{im\lambda} (f^2 - \sin^2 \theta)^{n/2} F_n^m(\xi), \quad \text{where} \quad \xi = f \cos \theta (f^2 - \sin^2 \theta)^{-1/2},$$

and $F_n^m(\xi)$ are Legendre functions.

In the traditional approximation, Equation (7.15) becomes the *Laplace tidal equation*, in which variables r and θ are separated (since in this approximation Q can depend only on r); this property of the equation is widely used when studying the eigen oscillations of rotating fluids in models with *horizontal* rigid walls. However, for an arbitrary wall relief solutions with separated variables are unsuitable, and the analytical advantages of the traditional approximation are lost. In the general case, for the eigen frequencies f we can derive an algebraic equation if we multiply (7.15) by $Q_1^2 P^*$ and integrate the result over the volume V ; here the asterisk indicates the complex conjugate, and the factor $Q_1^2 = (f^2 - k^2)^2 Q^2$ was introduced to ensure that the integrands would not have numerators containing f :

$$\begin{aligned} & \int [\mathbf{M}(\nabla P) \cdot \nabla Q_1^2 P^* + g c_0^{-2} M_r(\nabla P) Q_1^2 P^* - 4\omega^2 f^2 c_0^{-2} Q_1^2 P P^*] \times \\ & \times \rho_{00}^{-1} dV = \hat{\rho}_{00}^{-1} \int \hat{M}_r(\nabla P) \hat{Q}_1^2 \hat{P}^* d\Sigma. \end{aligned} \quad (7.18)$$

This is an algebraic equation in f , the coefficients of which are quadratic functionals of the complex amplitude of the eigen oscillations of the pressure P . After extracting the branches of the roots f of the equation, corresponding to acoustic, surface-capillary, gravity, internal-gravity, and gyroscopic waves, we can determine the eigen frequencies using a variational method, by selecting extremals of P minimizing successive roots of the necessary branch. Let us demonstrate some different types of eigen frequencies, beginning with acoustic waves, whose physical origins can be traced to the compressibility of the medium (the characteristic sound velocity c_0), so that when describing these we can neglect the other factors: rotation (set $\mathbf{k} = 0$ everywhere), density stratification (set $N = 0$), and oscillations of the free surface (set $\dot{P} = 0$). Then Equation (7.18) becomes radically simplified (for instance, without stratification $\mathbf{M}(\mathbf{A}) = \mathbf{L}(\mathbf{A})$, and without rotation $\mathbf{L}(\mathbf{A}) = \mathbf{A}$), and for the dimensional frequencies $\sigma = 2\omega f$ of the acoustical waves we obtain from it the formula

$$\sigma_{ac}^2 = \frac{c_0^2}{H^2}; \quad H^2 = \left(\int |\nabla P|^2 dV \right)^{-1} \int |P|^2 dV, \quad (7.19)$$

where H is the height scale of the homogeneous atmosphere (or the depth scale of the homogeneous ocean). When describing all other types of waves, however, the compressibility of the medium can be neglected as a first approximation, assuming the Boussinesq approximation $\text{div } \mathbf{u} = 0$, so that on the left-hand sides of Equations (7.15) and (7.18) only the first term is retained.

To describe surface waves, neglecting rotation and stratification, we use boundary condition (7.16). After dropping the first term on its right-hand side, we get for the capillary-wave frequencies

$$\sigma_{cap}^2 = \gamma_0 K_{cap}^3; \quad K_{cap}^3 = \left(\int |\nabla P|^2 dV \right)^{-1} \int \left| \nabla_h \frac{\partial P}{\partial z} \right|^2 d\Sigma. \quad (7.20)$$

If, on the other hand, we neglect the second term on the right-hand side of (7.16), we obtain for the frequencies of surface-gravity waves

$$\sigma_{grav}^2 = gK_{grav}; \quad K_{grav} = \left(\int |\dot{P}|^2 d\Sigma \right)^{-1} \int |\nabla P|^2 dV. \quad (7.21)$$

To describe internal-gravity waves, we neglect rotation (although this is allowable only for the high-frequency components of these waves) and oscillations of the free surface (so that in Equation (7.18) only the first term on the left-hand side remains), and then Equation (7.18) becomes

$$\begin{aligned} \sigma^2 \int \left[(\sigma^2 - N_0^2) |\nabla P|^2 - 2 \frac{\partial N_0^2}{\partial z} P^* \frac{\partial P}{\partial z} \right] \rho_{00}^{-1} dV = \\ = \int N_0^2 (\sigma^2 - N_0^2) |\nabla_h P|^2 \rho_{00}^{-1} dV. \end{aligned} \quad (7.22)$$

This is a biquadratic equation in the frequencies σ_{int}^2 of the internal waves. For quasiconstant N_0 from (7.22) we obtain

$$\sigma_{\text{int}}^2 \approx N_0^2 \left(\int |\nabla P|^2 \rho_{00}^{-1} dV \right)^{-1} \int |\nabla_h P|^2 \rho_{00}^{-1} dV, \quad (7.23)$$

from which it is evident that $|\sigma_{\text{int}}| < \max N_0(z)$ (in the traditional approximation it turns out that $|\sigma_{\text{int}}| > 2|\omega_r|$). Finally, to describe gyroscopic waves, we neglect oscillations of the free surface and, for the time being, stratification (which leads to a loss of *baroclinic* gyroscopic waves); without using the traditional approximation (and thus setting $k = 1$), we obtain from (7.18)

$$f^2 \int |\nabla P|^2 dV + if \mathbf{k} \cdot \int \nabla P \times \nabla P^* dV - \int |\mathbf{k} \cdot \nabla P|^2 dV = 0, \quad (7.24)$$

(in the traditional approximation, due to the fictitious dependence of $k^2 \rightarrow k_r^2$ on θ , instead of (7.24) for f_{gyro} we get an equation of fourth order with two nonphysical roots). For quasisolenoidal fields P the second term here is small, so that

$$\sigma_{\text{gyro}}^2 \approx 4\omega^2 \left(\int |\nabla P|^2 dV \right)^{-1} \int |\mathbf{k} \cdot \nabla P|^2 dV, \quad (7.25)$$

which yields the inequality $|\sigma_{\text{gyro}}| < 2\omega$.

As we have already mentioned, if at time $t = 0$ in some region of space V the conditions for stationarity of the perturbations $(G_1) - (G_4)$ are disrupted, then in this region in addition to motions of the first kind (stationary perturbations) there will appear motions of the second kind. The latter take the form of fast waves, and when they leave region V only motions of the first kind remain in it, these being determined by the fields of the linear invariants (which can be constructed according to the initial data). This restoration of the conditions of compatibility $(G_1) - (G_4)$ is known as adaptation of the fluid-dynamical fields. Adaptation to state (G_1) is accomplished by the generation and dispersal of acoustic waves, which takes an amount of time approximately equal to the time for the acoustic-wave front to traverse the main thickness of the medium (in the atmosphere this takes some tens of seconds). After this the adaptation process continues, to the state of geostrophic equilibrium $(G_2) - (G_4)$, via the generation and dispersal of gravity waves. Behind the fronts of these waves there is a continuous 'wake' in which the oscillations are damped.

Dikii (1969) has shown that, as a result, a stationary state $(G_1) - (G_4)$ is restored such that the difference between the initial and final states (let us call this the *departing wave*) has the lowest energy; by varying the energy of the departing wave, we can obtain the equation of conservation of potential vorticity. Since in the atmosphere the kinetic energy of the perturbations in (7.7) is, as a rule, an order of magnitude higher than the potential energy (for instance, for $u \sim 10$ m/s, $p'/p_0 \sim 10^{-2}$, and $p_0 \sim \rho_0 c_0^2 \sim \rho_0 gH$, we obtain $u^2 \sim 10^2$ m²/s² and

$(p'/\rho_0 c_0)^2 \sim (p'/p_0)^2 gH \approx 10 \text{ m}^2/\text{s}^2$), during the adaptation process the rearrangement of the pressure field involves a lower energy expenditure for producing the departing wave than does the rearrangement of the wind field. Consequently, the pressure field usually has to accommodate itself to the wind field, rather than vice versa.

The acoustic and gravity waves which disperse during the adaptation period almost always have very small amplitudes, and thus they can be considered to be 'hydrodynamic noise', of little importance with regard to the 'weather', which is created by motions of the first kind (and by gyroscopic waves on a synoptic scale). Motions of the first kind have, generally speaking, fairly large amplitudes and must be described by nonlinear equations. Nonlinearity inevitably leads to unsteadiness and in general to a continual disruption of the conditions for compatibility of the fluid-dynamical fields $(G_1) - (G_4)$. Thus there is constantly a concurrence between the disruption of the compatibility and the adaptation, which serves to weaken the overall effect of these conflicting processes. Because the adaptation takes place more rapidly, the disruptions of the compatibility conditions $(G_1) - (G_4)$ are usually minor.

To conclude this section, let us consider briefly the *forced* oscillations produced when a periodic external force acts on a linear oscillatory system, described by Equations (7.1)–(7.5). For instance, this may be a tide-generating force, giving rise to an acceleration $g\nabla\tilde{\xi} = -\nabla U$, where U is the sum of the tidal potentials of the external bodies:

$$U_i = Gm_i(r_i^{-1} - |\mathbf{r}_i - \mathbf{r}|^{-1} + r_i^{-3} \mathbf{r}_i \cdot \mathbf{r}), \quad (7.26)$$

and G is the gravitational constant, m_i is the mass of the tide-producing body, and $\mathbf{r}_i = \mathbf{r}_i(t)$ is its radius vector (found astronomically; on the Earth the geocentric radius vector of the Moon or Sun). The correction for the pull of the tidal bulges of water is given by the factor $(1 - 3\rho/5\rho_1)^{-1} \approx 1.12$ (where ρ_1 is the mean density of the land and water and ρ is the density of water, so that $\rho/\rho_1 \approx 0.18$). On the Earth, potential U contains harmonics with all possible frequencies $n_1\sigma_1 + \dots + n_6\sigma_6$, $n_1, \dots, n_6 = 0, \pm 1, \pm 2, \dots$, where $\sigma_i = 2\pi/\tau_i$ are the six fundamental frequencies in the Earth–Moon–Sun system ($\tau_1 = 24 \text{ h } 50.47 \text{ min}$ is a lunar day, $\tau_2 = 27.321582 \text{ days}$ is a tropical month, $\tau_3 = 365.242199 \text{ days}$ is a year, $\tau_4 = 8.847 \text{ years}$ is the period of rotation of the line of apsides of the lunar orbit, $\tau_5 = 18.613 \text{ years}$ is the period of rotation of the line of nodes of the lunar orbit, $\tau_6 = 20,940 \text{ years}$ is the period of longitude variation of the solar perigee). To describe the forced oscillations, it is convenient to introduce columns of functions $\mathbf{B} = \{\mathbf{u}, P, \eta'', \xi\}$ with a 'scalar product'

$$\begin{aligned} (\mathbf{B}_1 \cdot \mathbf{B}_2) = & \frac{1}{2} \iint \left[\mathbf{u}_1 \cdot \mathbf{u}_2^* + \frac{4\omega^2 P_1 P_2^*}{\rho_{00}^2 c_0^2} + \frac{g^2 \eta_1'' \eta_2''^*}{N_0^2} \right] \rho_0 dV + \\ & + \frac{1}{2} \int [g\xi_1 \xi_2^* + \gamma_0 \nabla \xi_1 \cdot \nabla \xi_2^*] \rho_0 d\mathbf{S} \end{aligned} \quad (7.27)$$

(so that $(\mathbf{B} \cdot \mathbf{B})$ is the total energy (7.7)). Let $\{\mathbf{B}_n\}$; $n = 1, 2, \dots$ be a complete orthonormalized (in the sense of (7.27)) system of complex amplitudes of the eigen oscillations with all possible eigen frequencies f_n , and let $\Sigma c_n \mathbf{B}_n$ be the expansion in this system of the functions of the complex amplitude of the acceleration of a periodic external force with a fixed dimensionless frequency f_0 . Then the equation describing the forced oscillations and its solution can be put into matrix form:

$$D\mathbf{B} = \mathcal{L}\mathbf{B} + e^{2if_0\omega t} \Sigma c_n \mathbf{B}_n; \quad (7.28)$$

$$\mathbf{B} = -ie^{2if_0\omega t} \Sigma \frac{c_n \mathbf{B}_n}{f_0 - f_n}, \quad (7.29)$$

where $\mathcal{L}$ is the matrix of the space operators entering into Equations (7.1)–(7.3), (7.5). Thus we see that, if the frequency of the effect f_0 is close to any of the eigen frequencies f_n , then the reaction (7.29) to the effect will be a resonant reaction. For tides on the Earth in a World Ocean with real configurations of its shores and bottom, the natural frequencies f_n were calculated numerically by Kagan and Gotlib (1980) using the ‘traditional approximation’ in a barotropic ocean and the so-called ‘shallow-water approximation’ or ‘long-wave approximation’, when the wavelengths are considerably longer than the thickness of the fluid layer (the depth of the ocean) and we need only consider the equations on the average over the thickness of this layer.

Using these approximations, we arrive at the Laplace tidal equations with unknown functions $\mathbf{B} = (\bar{u}_\theta, \bar{u}_\lambda, \zeta)$, where the bar signifies averaging over the layer thickness, and the eigen oscillations are surface gravity waves and barotropic Rossby waves. The natural frequencies were calculated not using the variational method indicated above, but rather by solving the problem with the initial data (in the form of two spherical harmonics) and finding the Fourier transforms of the oscillations obtained. The eigen periods were found to be between 8 and 144 h, with a maximum amplitude $B_{\max}$ corresponding to a eigen oscillation with a period of 39.78 h; a total of 34 eigen oscillations were detected, with squared amplitudes $B^2 \geq 0.2 B_{\max}^2$ (see their spectrum $B^2(\sigma)$ in Figure 7.1).

The eigen periods found (12.50, 12.68, and 12.83 h) are very close to the periods of the semidiurnal tides N_2 and M_2 , while the eigen periods (22.50, 23.87, and 25.90 h) are close to the period of the diurnal tide K_1 . These results imply that the semidiurnal and diurnal tides in the World Ocean probably have a resonance character; empirical data reveal (Munk and Cartwright, 1966) that some nonlinear effects even appear in them.

Without the ‘traditional approximation’ (but for $\text{div } \mathbf{u} = 0$, that is, in the Boussinesq approximation) the spectrum of the eigen frequencies of a spherical layer was investigated by Kivman and Maslova (1984), who established that

$$|\sigma| \leq \frac{1}{2} [1 + (1 + 4N_{\max}^2)^{1/2}].$$

If the spectrum is discrete, then it will be everywhere dense on the segment $\pm(1 +$

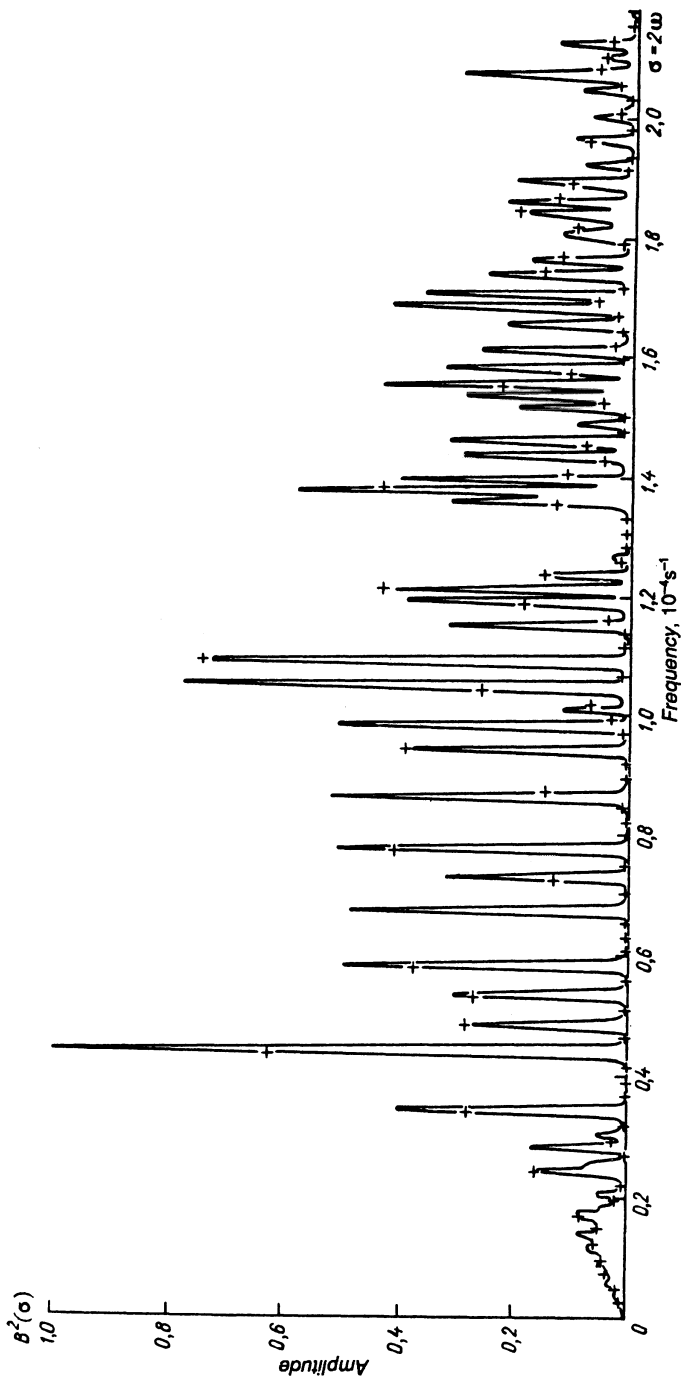


Fig. 7.1. Spectrum $B^2(\sigma)$ of eigen frequencies of Laplace tidal equations.

$N_{\max}^2)^{1/2}$, which is associated mainly with the presence of short waves. At low frequencies these are localized in certain latitude zones, depending on the stratification of the medium. In the traditional approximation these zones change; in particular, for weak stratification (low N) the waves in the equatorial zone are filtered out.

8. SMALL OSCILLATIONS OF AN ISOTHERMAL ATMOSPHERE

In this section we will illustrate the main subjects of Section 7, following Monin and Obukhov (1958), who considered the simplest example of an atmosphere (without a fluid upper surface S). We assume equation of state (1.7) (for $s = 0$) and we use the traditional approximation with a *constant* Coriolis parameter $f = 2\omega_z$ (the latter, that is, ignoring the dependence of f on the latitude, 'filters out' from the solutions of the equations the Rossby waves, which will be considered in Section 10); in this example all the calculations are carried out analytically. Let us employ the local Cartesian coordinates x, y, z and the corresponding velocity components u, v, w and let us set

$$\rho_0 u = \frac{\partial \varphi}{\partial x} - \frac{\partial \psi}{\partial y}; \quad \rho_0 v = \frac{\partial \varphi}{\partial y} + \frac{\partial \psi}{\partial x}; \quad \rho_0 w = \chi. \quad (8.1)$$

Then, in the above-indicated approximation, Equations (7.1)–(7.3) can be rewritten as

$$\frac{\partial \rho'}{\partial t} = -\Delta_h \varphi - \frac{\partial \chi}{\partial z}; \quad (8.2)$$

$$\frac{\partial \psi}{\partial t} + f\varphi = 0; \quad \frac{\partial \varphi}{\partial t} - f\psi = -p'; \quad \lambda \frac{\partial \chi}{\partial t} = -\frac{\partial p'}{\partial z} - g\rho'; \quad (8.3)$$

$$\frac{\partial p'}{\partial t} = -c_0^2 \Delta_h \varphi - g_1 \chi - c_0^2 \frac{\partial \chi}{\partial z}; \quad (8.4)$$

$$g_1 = \frac{c_0^2 N_0^2}{g} = (\kappa - 1)g + \frac{\partial c_0^2}{\partial z},$$

where Δ_h is the horizontal Laplace operator, and factor λ , equal to unity here, will be zero in the hydrostatic approximation. Reasonable boundary conditions for Equations (8.2)–(8.4) are imposed on the vertical mass flow: $\chi \rightarrow 0$ as $z \rightarrow 0$ and as $z \rightarrow \infty$ (the second of these conditions may prove to be insufficient, in which case it must be replaced by the more stringent condition that the kinetic energy of the vertical oscillations $\rho_0 w^2 \rightarrow 0$).

Let us introduce, as in (7.27), a column of unknown functions $\mathbf{B} = \{\psi, \varphi, p', \chi, \rho'\}$ and let us write Equations (8.2)–(8.4) in a matrix form similar to (7.28): $\partial \mathbf{B} / \partial t = \mathcal{L} \mathbf{B}$, where $\mathcal{L} = (\mathcal{L}_{ij})$ is the matrix of the space differential operators entering into these equations. We will seek a *linear invariant* of the equations in the form $J_1 = \gamma^k B_k$ (summation over k from 1 to 5), where γ^k are time-independent linear operators. Since we know that $\partial J_1 / \partial t = \gamma^k \partial B_k / \partial t = \gamma^k \mathcal{L}_k^i B_i$, operators γ^k must satisfy the system of equations $\gamma^k \mathcal{L}_k^i \equiv 0$ (conjugate to the system $\mathcal{L}_k^i B_i = 0$,

defining the steady-state solution **B**). From the specific form of matrix $\mathcal{L}$, determined according to (8.2)–(8.4), it follows that since in its first column only the second matrix element $\mathcal{L}_1^2 = f$ is nonzero, while in the fifth column only the fourth element $\mathcal{L}_5^4 = -g/\lambda$ is nonzero, we should set $\gamma^2 = \gamma^4 = 0$; then the equation derived with the aid of the third column $\mathcal{L}_3^i$ is satisfied identically. Using the second and fourth columns $\mathcal{L}_2^i$ and $\mathcal{L}_4^i$, we obtain the equations

$$\gamma^1 f + \gamma^3 c_0^2 \Delta_h + \gamma^5 \Delta_h = 0 \quad \text{and} \quad \gamma^3 \left(g_1 + c_0^2 \frac{\partial}{\partial z} \right) + \gamma^5 \frac{\partial}{\partial z},$$

from these, setting $\gamma^1 = \Delta_h$ and seeking γ^3 and γ^5 in the form of linear functions of $\partial/\partial z$, we obtain

$$\gamma^3 = f \frac{\partial}{\partial z} \frac{1}{g_1}, \quad \text{and} \quad \gamma^5 = -f \frac{\partial}{\partial z} \frac{c_0^2}{g_1} - f.$$

Next, with the aid of these results, we obtain the linear invariant

$$J_1 = \Delta_h \psi + f \frac{\partial}{\partial z} \frac{p' - c_0^2 \rho'}{g_1} - f \rho', \quad (8.5)$$

which is a linearized form of the potential vorticity. From (8.2) and (8.4) we get the equation $(\partial/\partial t) (p' - c_0^2 \rho') = -g_1 \chi$, which for $z = 0$ yields the additional invariant

$$J_2 = \hat{p}' - \hat{c}_0^2 \hat{\rho}'. \quad (8.6)$$

It should be noted that invariants J_1 and J_2 do not contain λ , so that they will not vary upon transition to the hydrostatic approximation.

Equations (8.2)–(8.4) possess some family of steady-state [stationary] solutions, which all satisfy conditions (G₁)–(G₄) of the preceding section, expressible here in the following form:

$$\begin{aligned} (G_1) \quad \rho' &= -\frac{f}{g} \frac{\partial \psi}{\partial z}; & (G_2) \quad \chi &= 0; \\ (G_3) \quad p' &= f\psi; & (G_4) \quad \varphi &= 0. \end{aligned} \quad (8.7)$$

Therefore, the stationary motions are horizontal and nondivergent, and for them the geostrophic-wind formulas and the hydrostatic equation are valid:

$$u = -\frac{1}{f\rho_0} \frac{\partial p'}{\partial y}; \quad v = \frac{1}{f\rho_0} \frac{\partial p'}{\partial x}; \quad \frac{\partial p'}{\partial z} = -g\rho'. \quad (8.8)$$

It follows from (8.7) that each steady-state solution is completely determined if we specify a single function $\psi = \psi_s(\mathbf{x})$. On the other hand, each solution with properties (8.7) describes some stationary horizontal eddy motion. Moreover, any solution for which $\varphi = 0$ constitutes a steady-state solution: in this case the first two equations (8.3) imply that ψ and p' do not depend on t (being related by the formula $p' = f\psi$); Equation (8.4) becomes $c_0^2 \partial \chi / \partial z + g_1 \chi = 0$ and for the zero boundary condition at $z = 0$ it has only a zero solution, and then from (8.2) the time independence of ρ' ensues as well. The solution of Equations (8.2)–(8.4) for

arbitrary initial data $\psi_0, \varphi_0, p'_0, \chi'_0, \rho'_0$ will be steady-state if, and only if, it satisfies conditions (8.7). The necessity in this case is obvious, and to prove sufficiency let us eliminate from (8.2)–(8.4) the unknowns $\psi, p',$ and $\rho',$ which yields for φ and χ the following equations and initial data:

$$\begin{aligned} \left(\frac{\partial^2}{\partial t^2} + f^2 \right) \varphi &= g_1 \chi + c_0^2 \frac{\partial \chi}{\partial z} + c_0^2 \Delta_h \varphi; \\ \lambda \frac{\partial^2 \chi}{\partial t^2} &= \frac{\partial}{\partial z} \left(g_1 \chi + c_0^2 \frac{\partial \chi}{\partial z} + c_0^2 \Delta_h \varphi \right) + g \left(\frac{\partial \chi}{\partial z} + \Delta_h \varphi \right); \\ t = 0; \quad \varphi &= \varphi_0; \quad \frac{\partial \varphi}{\partial t} = f\varphi_0 - p'_0; \\ \chi &= \chi_0; \quad \lambda \frac{\partial \chi}{\partial t} = - \left(\frac{\partial p'_0}{\partial z} + g\rho' \right). \end{aligned} \quad (8.9)$$

If conditions (8.7) are satisfied, then all these initial data go to zero, and the equations have only the trivial solution $\varphi = \chi = 0$, from which follows the steady-state nature of the solution of Equations (8.2)–(8.4). If, on the other hand, the initial data for these equations do not satisfy conditions (8.7), then the solution can be represented as the sum of a steady-state solution, determined by some function $\psi_s(\mathbf{x})$ (for which invariants J_1 and J_2 are found from the initial data), plus an unsteady-state, so-called *wave*, solution (for which invariants J_1 and J_2 are zero), satisfying the following initial conditions at $t = 0$:

$$\begin{aligned} \psi &= \psi_0 - \psi_s; \quad \varphi = \varphi_0; \quad p' = p'_0 - f\psi_s; \quad \chi = \chi_0; \quad \rho' = \rho'_0 + \\ &+ \frac{f}{g} \frac{\partial \psi_s}{\partial z}. \end{aligned} \quad (8.10)$$

Now let us find, in terms of invariants J_1 and J_2 , the function ψ_s determining the steady-state solution. With the aid of formulas (8.5)–(8.7) this function can be related as follows to J_1 and J_2 :

$$\Delta_h \psi_s + f^2 \frac{\partial}{\partial z} \left(\frac{\psi_s}{g} + \frac{\psi_s}{g_1} + \frac{c_0^2}{gg_1} \frac{\partial \psi_s}{\partial z} \right) = J_1; \quad (8.11)$$

$$f\hat{\psi}_s + \frac{fc_0^2}{g} \frac{\partial \psi_s}{\partial z} = J_2. \quad (8.12)$$

Therefore, function ψ_s can be found as the solution of the nonhomogeneous Equation (8.11) for the nonhomogeneous boundary condition (8.12); the second boundary condition can in this case be the requirement that ψ_s be bounded as $z \rightarrow \infty$. Note, by the way, that terms in (8.11) with higher derivatives of ψ_s , taking into account the relation $gg_1 = c_0^2 N_0^2$ given in (8.4), can be expressed in the form $(\Delta_h + f^2/N_0^2)(\partial^2/\partial z^2)\psi_s$. For $N_0^2 > 0$ (that is, for stable stratification) this operator is elliptic; in it the ratio of the second term to the first is $(L/L_R)^2$, where L is the

horizontal scale, and $L_R = HN_0 f^{-1}$ is the Rossby deformation radius, introduced back in Section 4. For $N_0^2 < 0$ (that is, for unstable stratification) the operator turns out to be hyperbolic and describes the waves distributed over the vertical. Note, too, that Equations (8.11), (8.12) for ψ_s can be replaced by a single integral equation, if we introduce the combined invariant

$$\begin{aligned} \frac{f^0 \hat{\rho}_0}{g} \tilde{\Omega} &= \int_0^\infty J_1 dz + \frac{\hat{\rho}_0}{\kappa} \int_0^z \frac{g_1 dz}{p_0} \int_z^\infty J_1 dz + f \tilde{g}_1^{-1} J_2 = \\ &= \int_0^\infty (\Delta_h \psi - f \rho') dz + \frac{\hat{\rho}_0}{\kappa} \int_0^z \frac{g_1 dz}{\rho_0} \int_z^\infty (\Delta_h \psi - f \rho') dz - \\ &\quad - \frac{f \hat{\rho}_0}{\kappa} \int_0^z (p' - c_0^2 \rho') p_0^{-1} dz, \end{aligned} \quad (8.13)$$

which is an integral form of the potential vorticity. Next let us express all the unknown functions on the right-hand side of (8.13) in terms of ψ_s using formulas (8.7), and let us introduce a new unknown $\tilde{\psi} = \psi_s / \rho_0$, a horizontal scale $L_0 = f^{-1}(\hat{p}_0 / \hat{\rho}_0)^{1/2}$, a dimensionless horizontal Laplacian $\Delta = L_0^2 \Delta_h$, and a new variable of integration $\xi = p_0 / \hat{p}_0$, to obtain

$$\begin{aligned} \Delta \left(\int_0^1 \tilde{\psi} d\xi + \int_\xi^1 \frac{\alpha^2 d\xi}{\xi^2} \int_0^\xi \tilde{\psi} d\xi \right) - \tilde{\psi} &= \tilde{\Omega}; \\ \alpha^2 &= \frac{\kappa - 1}{\kappa} \frac{\gamma_a - \gamma}{\gamma_a} \frac{T_0}{\hat{T}_0} = \frac{H^2 N_0^2}{c_0^2}. \end{aligned} \quad (8.14)$$

The equation for ψ_s was obtained in this form by Monin (1958); in the case of a quasiconstant *positive* stratification parameter α^2 (strictly speaking, it is constant — and positive — in an isothermal atmosphere; viz. title of present section), a solution of this equation was also obtained, having the form

$$\Psi(x, y, \xi) = -\frac{1}{2\pi} \int_0^1 d\xi_1 \int_{-\infty}^\infty \int G(r, \xi, \xi_1) \tilde{\Omega}(x_1, y_1, \xi_1) dx_1 dy_1, \quad (8.15)$$

where Ψ is the integral expression in parentheses in (8.14); x, y, ξ are the horizontal coordinates and the undisturbed pressure p_0 , measured by the scales L_0 and $\hat{p}_0$; $r = [(x - x_1)^2 + (y - y_1)^2]^{1/2}$, and $G(r, \xi, \xi_1)$ is the solution for $\tilde{\Omega} = -2\pi \delta(x - x_1) \delta(y - y_1) \delta(\xi - \xi_1)$. In order to find function G , let us change variables $(x, y, r) = \alpha(\xi, \eta, \rho)$, $\xi = \exp(-\theta)$, $G = 2\pi G_1 \exp[(\theta + \theta_1)/2]$, and then after differentiating Equation (8.14) for function $G_1(\rho, \theta, \theta_1)$ we get

$$\begin{aligned} \frac{\partial^2 G_1}{\partial \xi^2} + \frac{\partial^2 G_1}{\partial \eta^2} + \frac{\partial^2 G_1}{\partial \theta^2} - \frac{1}{4} G_1 &= -\delta(\xi - \xi_1) \delta(\eta - \eta_1) \delta(\theta - \theta_1); \\ \left[\frac{\partial G_1}{\partial \theta} + \left(\frac{1}{2} - \alpha^2 \right) G_1 \right]_{\theta=0} &= 0; \quad e^{-\theta/2} \left(\frac{\partial G_1}{\partial \theta} + \frac{1}{2} G_1 \right)_{\theta \rightarrow \infty} \rightarrow 0. \end{aligned} \quad (8.16)$$

For this equation the action function of a point source in unbounded space is

$$G_0(\rho, \theta, \theta_1) = (4\pi A)^{-1} e^{-A/2}; \quad A = [\rho^2 + (\theta - \theta_1)^2]^{1/2}. \quad (8.17)$$

In the halfspace $\theta \geq 0$ for boundary conditions (8.16) the action function of a point source is obtained by the addition of fictitious sources on the negative θ semiaxis:

$$\begin{aligned} G_1 = G_0(\rho, \theta, \theta_1) + G_0(\rho, \theta, -\theta_1) + (1 - 2\alpha^2) \exp\left[-\left(\frac{1}{2} - \alpha^2\right) \theta_1\right] \times \\ \times \int_{-\infty}^{-\theta_1} G_0(\rho, \theta, \theta') \exp\left[-\left(\frac{1}{2} - \alpha^2\right) \theta'\right] d\theta'. \end{aligned} \quad (8.18)$$

The integral appearing here is the difference between the integrals over the regions $(-\infty, +\infty)$ and $(-\theta_1, \infty)$; the first of these, via a change of variables

$$\theta' = \theta + \frac{1}{2} \frac{\tau}{1 - \alpha^2} - \rho^2 \frac{1 - \alpha^2}{2\tau}$$

becomes

$$\begin{aligned} \frac{1}{4\pi} \exp\left[-\left(\frac{1}{2} - \alpha^2\right) \theta\right] \int_0^\infty \exp\left[-\frac{\tau}{2} - \alpha^2 \rho^2 \frac{1 - \alpha^2}{2\tau}\right] \frac{d\tau}{\tau} = \\ = \frac{1}{2\pi} \exp\left[-\left(\frac{1}{2} - \alpha^2\right) \theta\right] K_0(\alpha \rho \sqrt{1 - \alpha^2}); \end{aligned} \quad (8.19)$$

where K_0 is a cylindrical MacDonald function. Now, collecting the results obtained and returning to the original variables, we obtain

$$\begin{aligned} G(r, \xi, \xi_1) = \frac{1 - 2\alpha^2}{(\xi \xi_1)^{\alpha^2}} K_0(r \sqrt{1 - \alpha^2}) - \frac{1}{2} \frac{1 - 2\alpha^2}{(\xi \xi_1)^{\alpha^2}} \int_{\ln \xi \xi_1}^\infty \times \\ \times B^{-1} \exp\left[-\frac{B}{2} - \left(\frac{1}{2} - \alpha^2\right) \theta\right] d\theta + \\ + \frac{1}{2\sqrt{\xi \xi_1}} \left(\frac{e^{-C/2}}{C} + \frac{e^{-C'/2}}{C'} \right); \end{aligned} \quad (8.20)$$

$$B = \alpha^{-1}(r^2 + \alpha^2 \theta^2)^{1/2}; \quad C = \alpha^{-1} \left(r^2 + \alpha^2 \ln^2 \frac{\xi}{\xi_1} \right)^{1/2};$$

$$C' = \alpha^{-1}(r^2 + \alpha^2 \ln^2 \xi \xi_1)^{1/2}.$$

Note that function G has a singularity of the form C^{-1} at the point $r = 0$, $\xi = \xi_1$. If in function (8.20) we consider the first term, which has at all points of the

vertical line $r = 0$ a logarithmic singularity, we can readily apply the limit transition $\alpha \rightarrow 0$ to the formulas for a barotropic atmosphere. This limit transition is nonuniform with regard to x, y, ζ , but for a sufficiently smooth function $\tilde{\Omega}$ it can also be carried out directly in formula (8.15). Consequently, this formula takes the form proposed much earlier by Obukhov (1949):

$$\Psi(x, y) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_0(r) \tilde{\Omega}(x_1, y_1) dx_1 dy_1, \quad (8.21)$$

where $\tilde{\Omega}$ here refers to the mean potential vorticity over the thickness of the atmosphere.

After finding the function ψ_s which determines the steady-state solution, from the initial data we can distinguish the part of (8.10) determining the wave solution. While finding the latter, we can replace one of Equations (8.2)–(8.4) by the condition of zero potential vorticity, so that the equations describing the wave solution are fourth-order equations with respect to time. From these we can in principle express ψ' , p' , and ρ' in terms of φ and χ without integrating over time, while functions φ and χ are determined from Equation (8.9), not containing ψ_s . Since the coefficients in Equation (8.9) do not depend on x, y, t , it follows that these equations have particular solutions in the form of harmonic waves with amplitudes depending on z :

$$\{\varphi, \chi\}(\mathbf{x}, z, t) = \{\Phi(z), X(z)\} \exp[i(\mathbf{k} \cdot \mathbf{x} - \sigma t)], \quad (8.22)$$

where $\mathbf{k} = (k_x, k_y)$ are the horizontal wave numbers, which can be arbitrary, while $\sigma = \sigma(\mathbf{k})$ are the frequencies to be determined. Since in dynamical equations neither the influxes nor the dissipation of energy are taken into account, the total energy of the waves (8.22) (that is, the simplified analog of quantity (7.7)) cannot vary with time. Consequently, all the eigen frequencies σ are real. The solution of Equations (8.9) for arbitrary initial data can be represented as a superposition of elementary wave solutions (8.22) with all possible values of $\mathbf{k}$. Therefore, if we determine the nature of the elementary waves (8.22) we are able to understand the structure of any wave solution. To study the elementary waves, we have to ascertain the height dependences of their amplitudes $\Phi(z)$ and $X(z)$ and the spectrum of their frequencies $\sigma(\mathbf{k})$. By substituting (8.22) and (8.9), we get

$$\begin{aligned} (f^2 + k^2 c_0^2 - \sigma^2) \Phi &= g_1 X + c_0^2 \frac{\partial X}{\partial z}; \\ (f^2 - \sigma^2) \left(c_0^2 \frac{\partial \Phi}{\partial z} + g \Phi \right) &= (gg_1 - \lambda c_0^2 \sigma^2) X. \end{aligned} \quad (8.23)$$

It should be noted, firstly, that these equations have the following nontrivial particular solution (satisfying the boundary conditions):

$$X = 0; \quad \Phi = \hat{\Phi} \left(\frac{p_0}{\hat{p}_0} \right)^{1/\kappa}; \quad \sigma^2 = f^2 + k^2 c_0^2. \quad (8.24)$$

The corresponding *two-dimensional* waves encompass the entire thickness of the atmosphere, propagate only horizontally (with a maximum group velocity c_0), and do not cause vertical oscillations of air particles; the hydrostatic approximation ($\lambda = 0$) does not change them. Waves with $X \neq 0$ will be considered only for quasiconstant c_0 and g_1 (which is precisely true in an isothermal atmosphere). Then, from (8.23) it follows that both Φ and X satisfy the equation

$$(f^2 - \sigma^2) \left(\frac{\partial^2 F}{\partial z^2} + \frac{g + g_1}{c_0^2} \frac{\partial F}{\partial z} + \frac{\lambda \sigma^2}{c_0^2} F \right) - k^2 \left(\frac{gg_1}{c_0^2} - \lambda \sigma^2 \right) F = 0. \quad (8.25)$$

Let us seek particular solutions of this equation in the form $F = \hat{F} \exp[-M + im]z]$, where M and m are real numbers. By substituting this expression into (8.25) and equating to zero the imaginary part of the resulting equation, we obtain $M = (g + g_1)/2c_0^2$. Then, by setting the real part of the equation equal to zero, we arrive at the relation

$$(\sigma^2 - f^2)(\lambda \sigma^2 - c_0^2 m_1^2) \left(\sigma^2 - \frac{gg_1}{c_0^2} \right)^{-1} = k^2 c_0^2; \quad m_1^2 = m^2 + \left(\frac{g + g_1}{2c_0^2} \right)^2. \quad (8.26)$$

These particular solutions correspond to so-called *internal* plane harmonic waves with a horizontal wave vector $\mathbf{k}$ and a vertical wave number m ; the amplitudes of these waves decrease exponentially with height, while $m \neq 0$ (since for $m = 0$ no combination of such waves can satisfy the boundary condition $\chi \rightarrow 0$ as $z \rightarrow 0$). Consequently, the propagation velocity of these waves always has a vertical component, that is, the waves are essentially three-dimensional. Because internal waves with any $\mathbf{k}$ and m may enter into the solution of the Cauchy problem (8.9) for arbitrary initial data, therefore in (8.26) they should be assumed arbitrary, and this becomes an equation in σ . Note that the left-hand side of (8.26) is nonnegative, so that, since

$$f^2 < \frac{gg_1}{c_0^2} < \frac{(g + g_1)^2}{4c_0^2},$$

it is easy to see that the roots σ^2 can be found only in two intervals:

$$\sigma^2 > \frac{(g + g_1)^2}{4\lambda c_0^2}; \quad f^2 \leq \sigma^2 < \frac{gg_1}{c_0^2}, \quad (8.27)$$

and, in view of the continuous dependence of the roots σ^2 on $\mathbf{k}$ and m , all points within these intervals are possible frequencies of the internal waves. Let us write the solutions σ^2 of Equation (8.26) as

$$\sigma_{a,g}^2 = \frac{1}{2} c_0^2 K^2 \left[1 \pm \left(1 - \frac{4gg_1}{\lambda c_0^2} K_1^2 K^{-4} \right)^{1/2} \right]; \quad (8.28)$$

$$K^2 = k^2 + L^{-2} + m_1^2 \lambda^{-1}; \quad K_1^2 = k^2 + m_1^2 c_0^4 (gg_1 L^2)^{-1},$$

where $L = c_0/f$ is a typical scale of the horizontal motions of a compressible fluid

in a Coriolis force field (Obukhov, 1949). From this we see that the effect of the Coriolis force is substantial only for small $k \geq L^{-1}$. In order to ascertain the behavior of waves with frequencies σ_a and σ_g , we ignore the effect of the Coriolis force, that is, we set $L^{-1} = 0$. Let us also assume that $\lambda = 1$ and let us consider an isothermal atmosphere, in which $c_0^2 = \kappa g H$ and $g_1 = (\kappa - 1)g$. Then formulas (8.28) become

$$\sigma_{a,g}^2 = \frac{1}{2} \kappa g H K^2 \left[1 \pm \left(1 - 4 \frac{\kappa - 1}{\kappa^2} \frac{k^2}{H^2} K^{-4} \right)^{1/2} \right]; \quad (8.28')$$

$$K^2 = k^2 + m^2 + \frac{1}{4H^2}.$$

Recall that we are considering adiabatic processes, for which according to (4.3) and (1.7) the quantity $\rho p^{-1/\kappa}$ is conserved in fluid particles, that is, these are essentially polytropic processes with a polytropic index κ . For $\kappa \rightarrow \infty$ the quantity which remains constant is ρ , so that this case corresponds to an incompressible fluid. Here all the frequencies σ_a go to ∞ , that is, waves with these frequencies disappear; then only oscillations with frequencies σ_g remain possible, where σ_g is found from the formula $\sigma_g^2 = g H^{-1} k^2 K^{-2}$. These are obviously internal gravity waves.

For $\kappa \rightarrow 1$ the conserved quantity is T , so that we are dealing with isothermal processes, for which the isothermal stratification is neutral. Here all the frequencies σ_g go to zero, that is, oscillations with these frequencies disappear (are transformed into stationary eddy motions), and the only possible oscillations are those with frequencies σ_a , given by the formula $\sigma_a^2 = g H K^2$. These are obviously acoustic waves. A similar discussion shows that the two-dimensional waves (8.24) are acoustic waves.

With regard to adiabatic processes with $\kappa \approx 1.4$, we can conclude that in (8.28') frequencies σ_a correspond to acoustic waves, distorted because the isothermal stratification differs from neutral, while σ_g corresponds to internal gravity waves, distorted because of the compressibility of the fluid. Waves corresponding to frequencies (8.28) and containing an additional distortion due to the effect of the Coriolis force can be interpreted similarly. For $\lambda = 1$ and sufficiently high k and m , formulas (8.28) become

$$\sigma_a^2 \approx c_0^2 K^2 > \frac{(g + g_1)^2}{4c_0^2}; \quad \sigma_g^2 \approx \frac{g g_1}{c_0^2} \frac{K_1^2}{K^2} < \frac{g g_1}{c_0^2}, \quad (8.29)$$

where the inequalities shown here are strictly valid for any k and m . In an isothermal atmosphere for $\kappa \approx 1.4$ and $H \approx 8$ km we get $\tau_a = 2\pi/\sigma_a < 300$ s and $\tau_g = 2\pi/\sigma_g > 330$ s. This result shows a good fit with spectra of atmospheric microbarograms $\sigma S_{pp}(\sigma)$, having a minimum for periods $\tau \sim 300$ s.

In the hydrostatic approximation the frequencies of the possible small oscillations are obtained from (8.28) at the limit as $\lambda \rightarrow 0$. Here all the frequencies σ_a go to infinity, that is, the internal acoustic waves are completely filtered out; of the

internal waves only the gravity waves remain, with frequencies

$$\sigma_g^2 \approx \frac{gg_1}{c_0^2} \frac{K_1^2}{m_1^2}. \quad (8.30)$$

These frequencies are too high, in comparison with (8.29), and they are higher for higher wave numbers k ; however, for $k \ll m$ (long waves) this error is negligible.

9. ADAPTATION

If at the initial time in some region of space V the hydrostatic equilibrium (G_1) has been disturbed, then internal acoustic waves form in V , and when they are scattered it can be assumed that the fluid has essentially already accommodated itself to state (G_1). After this the adaptation of the fluid-dynamical fields to the state of geostrophic equilibrium (G_2)–(G_4) continues, and on the average over the thickness of the layer filled with fluid this state is established after the scattering of the two-dimensional waves. Somewhat later, after the slower internal gravity waves have been scattered, it becomes established at all heights. Let us now describe this process of adaptation, following Monin (1958).

Monin took advantage of a conversion for *hydrostatic equilibrium* from Cartesian coordinates (x, y, z) to curvilinear nonsteady-state *isobaric* coordinates (x, y, p). In the latter the coordinate surfaces are vertical $x = \text{const}$ and $y = \text{const}$ planes, and instead of the pressure p (now an independent variable) a new unknown function appears in the fluid-dynamical equations, namely the height of the isobaric surfaces $z = z(x, y, p, t)$. Then the relations $\partial^*/\partial s = \partial/\partial s + \partial p/\partial s \partial/\partial p$ are valid, where $s = x, y$ or t , and $\partial^*/\partial s$ is the local derivative in Cartesian coordinates, as well as the relations

$$\frac{\partial}{\partial z} = \frac{\partial p}{\partial z} \frac{\partial}{\partial p} = -g\rho \frac{\partial}{\partial p}; \quad \frac{d^*}{dt} = \frac{d_h}{dt} + w^* \frac{\partial}{\partial p},$$

where d_h/dt is the individual derivative with respect to horizontal motion on the isobaric surfaces, and $w^* = dp/dt$ replaces the vertical velocity;

$$\frac{\partial p}{\partial s} = - \frac{\partial p}{\partial z} \frac{\partial z}{\partial s} = g\rho \frac{\partial z}{\partial s}.$$

With the aid of these relations the continuity equation (3.1) and the horizontal components of motion (3.2) in the traditional approximation (when replacing the vertical component of the equations of motion by the hydrostatic equation, we must keep the work of the Coriolis force equal to zero) become

$$D_h + \frac{\partial w^*}{\partial p} = 0; \quad D_h = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}; \quad (9.1)$$

$$\frac{d_h u}{dt} + w^* \frac{\partial u}{\partial p} - fv = -g \frac{\partial z}{\partial x}; \quad \frac{d_h v}{dt} + w^* \frac{\partial v}{\partial p} + fu = -g \frac{\partial z}{\partial y} \quad (9.2)$$

(recall that here u and v are the velocity components in the old Cartesian coordinates, rather than in the new curvilinear coordinates). Finally, if we define $\rho =$

$-(g(\partial z/\partial p)^{-1})$ from the hydrostatic equation, then the adiabaticity Equation (1.17) becomes

$$gp^2 \frac{d_h}{dt} \frac{\partial z}{\partial p} + \alpha^2 c^2 w^* = 0, \quad (9.3)$$

where α^2 is the stratification parameter, defined in (8.14). If the velocity of sound c is here assumed to be specified (as a function of ρ , p , and $T = p/R\rho$, say), then (9.1)–(9.3) is a system of four equations in the unknowns u , v , w^* , and z . The advantages of using isobaric coordinates include linearity of the continuity Equation (9.1) and linearity and potentiality of the terms with the pressure gradient $g\nabla_h z$ on the right-hand sides of the equations of motion. A disadvantage, however, is a more complicated boundary condition $\hat{w} = 0$ at the Earth's surface, obtained from the definition

$$w^* = \frac{dp}{dt} = g\rho \left(\frac{d_h z}{dt} - w \right);$$

if this condition is approximated for the mean surface pressure $p = \hat{p}$, while w^* is expressed in terms of D_h by integrating over p with the condition $w^* \rightarrow 0$ ($p \rightarrow 0$), then we get

$$g\hat{p} \frac{d_h \hat{z}}{dt} = - \int_0^{\hat{p}} D_h dp. \quad (9.4)$$

By linearizing Equations (9.1)–(9.4) relative to the state of rest, introducing the stream function and the potential of the horizontal-flow velocity as in (8.1) (but without the factor ρ_0), and using dimensionless variables similar to those in (8.14), we can describe the wave solutions of the linearized equations of the velocity potential φ with the aid of the equation

$$\frac{\partial^2 \varphi}{\partial t^2} + \varphi = \Delta \left(\int_0^1 \varphi d\xi + \int_\xi^1 \frac{\alpha^2 d\xi}{\xi^2} \int_0^\xi \varphi d\xi \right) \quad (9.5)$$

for

$$t = 0; \quad \varphi = \varphi_0; \quad \frac{\partial \varphi}{\partial t} = \psi_0 - z_0.$$

To solve this equation, let us use a Laplace transform with respect to t . We denote operator $\partial/\partial t$ by the letter s , the integral expression in parentheses on the right-hand side of (9.5) by the letter Φ , and its Laplace transform by $\tilde{\Phi}$. Then for $\tilde{\Phi}$ we obtain an equation completely analogous to the equation for function Ψ of formula (8.15), the solution of which for quasiconstant α^2 is

$$\begin{aligned} \tilde{\Phi}(x, y, \xi, s) = & \frac{1}{2\pi} \int_0^1 d\xi_1 \int_{-\infty}^{\infty} G(r\sqrt{1+s^2}, \xi, \xi_1) \times \\ & \times (s\varphi_0 + \psi_0 - z_0)_{x_1, y_1, \xi_1} dx_1 dy_1, \end{aligned} \quad (9.6)$$

where $G(r, \zeta, \xi_1)$ is given by formula (8.20). After performing an inverse Laplace transformation, we obtain

$$\begin{aligned} \Phi(x, y, \zeta, t) = & \frac{1}{2\pi} \int_0^1 d\xi_1 \int_{-\infty}^{\infty} \Gamma(r, \zeta, \xi_1, t) (\psi_0 - z_0)_{x_1, y_1, \xi_1} dx_1 dy_1 + \\ & + \frac{1}{2\pi} \frac{\partial}{\partial t} \int_0^1 d\xi_1 \int_{-\infty}^{\infty} \Gamma(r, \zeta, \xi_1, t) \varphi_0(x_1, y_1, \xi_1) dx_1 dy_1, \end{aligned} \quad (9.7)$$

where $\Gamma(r, \zeta, \xi_1, t)$ is the inverse transform of $G(r\sqrt{1+s^2}, \zeta, \xi_1)$, which can be found from formula (8.20) with the aid of tables of Laplace transforms. The inverse Laplace transformation can be performed in (8.20) under the integral sign as well, since because of the exponentially damping (for $\alpha^2 < \frac{1}{2}$) factor $\exp[-(\frac{1}{2} - \alpha^2)\xi]$ this integral converges uniformly with respect to r . As a result we get

$$\begin{aligned} \Gamma(r, \zeta, \xi_1, t) = & \frac{1 - 2\alpha^2}{(\xi\xi_1)^{\alpha^2}} \frac{\cos[t^2 - (1 - \alpha^2)r^2]^{1/2}}{[t^2 - (1 - \alpha^2)r^2]^{1/2}} - \frac{\alpha}{2r} \frac{1 - 2\alpha^2}{(\xi\xi_1)^{\alpha^2}} \times \\ & \times \int_{\ln \xi\xi_1}^{\infty} \exp\left[-\left(\frac{1}{2} - \alpha^2\right)\xi\right] J_0\left[\left(1 + \frac{\alpha^2\xi^2}{r^2}\right)^{1/2} \left(t^2 - \frac{r^2}{4\alpha^2}\right)^{1/2}\right] d\xi + \\ & + \frac{\alpha}{2r(\xi\xi_1)^{1/2}} J_0\left[\left(1 + \frac{\alpha^2}{r^2} \ln^2 \frac{\xi}{\xi_1}\right)^{1/2} \left(t^2 - \frac{r^2}{4\alpha^2}\right)^{1/2}\right] + \\ & + \frac{\alpha}{2r(\xi\xi_1)^{1/2}} J_0\left[\left(1 + \frac{\alpha^2}{r^2} \ln^2 \xi\xi_1\right)^{1/2} \left(t^2 - \frac{r^2}{4\alpha^2}\right)^{1/2}\right], \end{aligned} \quad (9.8)$$

where the first term is nonzero only for $r \leq t(1 - \alpha^2)^{-1/2}$, and the other terms are nonzero only for $r \leq 2\alpha t$. This expression implies that waves whose dispersal leads to adaptation of the fluid-dynamical fields can be put into one of two classes: (1) *barotropic waves*, propagating with a (dimensional) velocity $c_0(1 - \alpha^2)^{-1/2}$ and corresponding to the waves arising in a barotropic atmosphere due to perturbations of the geostrophic equilibrium between the velocity and pressure fields averaged over the height (Obukhov, 1949); (2) slower *baroclinic waves*, propagating with a velocity $2\alpha c_0$ (these do not appear in a barotropic atmosphere).

The limit of the formulas for a barotropic atmosphere is obtained when $\alpha \rightarrow 0$. Then functions Ψ and Φ become the stream function and the velocity potential of the horizontal flow, averaged over the thickness of the atmosphere. The action function (8.20) at the limit takes the form $G = K_0(r)$, so that for small α (in the actual terrestrial atmosphere $\alpha \sim 0.1$) and for values of ξ_1 not too close to ξ the first term in (8.20) will be the main term. Formula (8.15) is transformed into Obukhov's formula (8.21). After integration over ξ equation (9.5) becomes the following equation obtained by Obukhov:

$$\frac{\partial^2 \Phi}{\partial t^2} + \Phi = \Delta \Phi, \quad (9.5')$$

which differs from the usual wave equation by the presence of the second term on the left-hand side, characterizing the *gyroscopic rigidity* of a rotating fluid film. Function (9.8) at the limit is equal to $(t^2 - r^2)^{-1/2} \cos(t^2 - r^2)^{1/2}$; for small α the first term in (9.8) is the main term.

As mentioned in Section 7, during the adaptation process the velocity field usually varies relatively little, while the pressure distribution, on the other hand, varies sharply, approaching the distribution related by the formulas of the geostrophic wind to the limit velocity field (differing little from the initial field); therefore, it makes sense to refer to an adaptation of the pressure field to the wind field. As an example, we consider in a baroclinic atmosphere the initial conditions

$$\varphi_0 = z_0 = 0; \quad \psi_0 = A \left[\frac{1 + \alpha^2(1 - \xi)}{R^2} \left(\frac{x^2}{R^2} - 1 \right) - 2\xi \right] e^{-x^2/2R^2}, \quad (9.9)$$

for which the velocity field corresponds to plane-parallel flow along the y axis, antisymmetric relative to this axis with a cyclonic rotation of the wind at the transition through this axis (idealized model of flow at frontal interface), while the pressure field is undisturbed. As a result of the adaptation, here we have

$$\varphi = 0; \quad \psi_s = z_s = A \frac{1 + \alpha^2(1 - \xi)}{R^2} \left(\frac{x^2}{R^2} - 1 \right) e^{-x^2/2R^2}, \quad (9.10)$$

that is, the velocity field varies little, while in the pressure field a distinct frontal baric trough is formed. For $L_0 = 2500$ km, $R = 0.2$, and $A = 0.0002$ the halfwidth of the trough RL_0 is 500 km, while the depth is 4 m; the kinetic energy is reduced by 3% during adaptation, some of the energy being used to form the baric trough and some being carried away by the waves.

The action function $G(r, \xi, \zeta_1)$ of the linear operator from the left-hand side of Equation (8.14), which enters into formulas (8.15) and (9.6), finds another useful application, associated with the *quasigeostrophic approximation*. The latter is obtained by expanding the unknown functions in Equations (9.1)–(9.4) in asymptotic series, in powers of the Kibel' number $Ki = U/fL$, which yields in terms of the dimensionless variables used above the relations

$$\Omega = \Delta\psi = \Delta z + O(Ki); \quad D_h = \Delta\varphi = O(Ki). \quad (9.11)$$

Accurate to $O(Ki)$, from (9.1)–(9.4) we can derive the equation

$$\begin{aligned} & \Delta \left(\int_0^1 \frac{\partial z}{\partial t} d\xi + \int_\xi^1 \frac{\alpha^2 d\xi}{\xi^2} \int_0^\xi \frac{\partial z}{\partial t} d\xi \right) - \beta^2 \frac{\partial z}{\partial t} = \\ & = - \left(\int_0^1 [z, \Delta z] d\xi + \int_\xi^1 \frac{\alpha^2 d\xi}{\xi^2} \int_0^\xi [z, \Delta z] d\xi \right) - \beta^2 \int_0^1 \left[z, \frac{\partial z}{\partial \xi} \right] d\xi, \end{aligned} \quad (9.12)$$

where $\beta = L/L_0$, and square brackets $[A, B]$ signify the Jacobian $\partial(A, B)/\partial(x, y)$, so that the quantity $[z, \Delta z]$ describes in the quasigeostrophic approximation the

vorticity transfer, while $[z, \partial z / \partial \xi]$ describes the temperature transfer. In the linear theory the right-hand side of this equation goes to zero, and from this it follows that $\partial z / \partial t = 0$, that is, the field $z(x, y, \xi, t)$ is stationary. Then, integrating this equation over t and taking geostrophicity into account ($z \sim \psi$), we return to Equation (8.14) considered above. However, if the nonlinear right-hand side of Equation (9.12) is taken into account, then field $z(x, y, \xi, t)$ stops being stationary and starts to vary gradually with time (gradually in comparison with the fields of the fast waves participating in the adaptation processes). Equation (9.12) differs from (8.14) only in the use of another horizontal scale (L instead of L_0) and in the specific form of the right-hand side. Thus, its solution $\partial z / \partial t$ can be expressed in terms of the right-hand side in the form of an integral of type (8.15) with an action function expressed in terms of function $G(r, \xi, \xi_1)$ of formula (8.20):

$$\begin{aligned} \frac{\partial z(x, y, \xi, t)}{\partial t} = & \frac{1}{2\pi} \int_0^1 d\xi_1 \iint_{-\infty}^{\infty} \left\{ G(\beta r, \xi, \xi_1) [z, \Delta z]_{x_1, y_1, \xi_1, t} - \right. \\ & \left. - \frac{\beta^2 \xi_1^2}{\alpha^2} \frac{\partial G(\beta r, \xi, \xi_1)}{\partial \xi_1} \left[z, \frac{\partial z}{\partial \xi} \right]_{x_1, y_1, \xi_1, t} \right\} dx_1 dy_1. \end{aligned} \quad (9.13)$$

Hence we see that for a *baric tendency*, that is, for a rate of pressure variations $\partial z / \partial t$, the action function of the field of vorticity transfer $[z, \Delta z]$ will be G , while that of the field of temperature transfer $[z, \partial z / \partial \xi]$ will be the field $(\beta^2 \xi_1^2 / \alpha^2) \partial G / \partial \xi_1$ (these action functions were published in 1958 in coordinates k_x, k_y, z by Obukhov and Chaplygina, in coordinates k_x, k_y, ξ by Buleev and Marchuk, and in the above form by Monin). In the limit case of a barotropic atmosphere ($\alpha \rightarrow 0$) the derivatives $\partial z / \partial s$ stop depending on ξ , so that Equation (9.12) and its solution (9.13) become:

$$\Delta \frac{\partial z}{\partial t} - \beta^2 \frac{\partial z}{\partial t} = -[z, \Delta z]; \quad (9.12')$$

$$\frac{\partial z(x, y, t)}{\partial t} = \frac{1}{2\pi} \iint_{-\infty}^{\infty} K_0(\beta r) [z, \Delta z]_{x_1, y_1, t} dx_1 dy_1 \quad (9.13')$$

(Obukhov, 1949).

10. LAPLACE TIDAL EQUATION

Let us return to Equation (7.15) for the complex amplitude P of the eigen oscillations of the reduced pressure. We will consider this equation in the traditional approximation $\mathbf{k} \rightarrow k_r \nabla r$, $k_r = \cos \theta$ (actually, instead of replacing the constant field on a sphere $\mathbf{k} = \text{const}$ by a variable radial field $\cos \theta \cdot \nabla r$, which would not be justified at all, we replace the Coriolis acceleration $2\omega \mathbf{k} \times \mathbf{u}$ by the approximate expression $2\omega \cos \theta \cdot \nabla r \times \mathbf{u}$, leading only to a loss of vertical acceleration $2\omega u_\lambda \sin \theta$, usually very small in comparison with the quantity $\rho_0^{-1}(\partial p' / \partial z + g\phi')$, and to a loss in the λ projection of the acceleration of the term $2\omega u_r \sin \theta$, which

$2\omega u_r \sin \theta$, which is very small in comparison with the remaining term $2\omega u_\theta \cos \theta$ everywhere except in a very narrow equatorial zone). In this approximation we obtain

$$\mathbf{L}(\mathbf{A}) \rightarrow (f^2 - \cos^2 \theta)^{-1} [f^2 \mathbf{A} - if \cos \theta \mathbf{A} \times \nabla r - \cos^2 \theta A_r \nabla r], \quad (10.1)$$

so that now $L_r(\mathbf{A}) = A_r$, $\mathbf{L}(\nabla r) = \nabla r$, and $L_r(\nabla r) = 1$. Hence from formulas (7.12) it follows that $Q = f^2 - N_0^2$ and also

$$\mathbf{M}(\mathbf{A}) = \mathbf{L}(\mathbf{A}) + \frac{N_0^2 A_r \nabla r}{f^2 - N_0^2}; \quad M_r(\mathbf{A}) = \frac{f^2 A_r}{f^2 - N_0^2}. \quad (10.2)$$

In view of definitions (10.1) and (10.2), Equation (7.15) takes the form of the Laplace tidal equation, in which all three of the space variables are separated. In a fluid layer with *horizontal* lower and upper boundaries $r = r_0 - h$ and $r = r_0$, it is sufficient to consider solutions with separable variables, depending on the longitude periodically according to an $e^{im\lambda}$ law with integral values of m , that is, solutions of the form $P = e^{im\lambda} P_1(r) P_2(\theta)$. Then, after separating the variables, from the Laplace tidal equation for $P_1(r)$ and $P_2(\theta)$ we obtain the equations:

$$\frac{\partial}{\partial r} \frac{r^2}{\rho_{00}(f^2 - N_0^2)} \frac{\partial P_1}{\partial r} - \frac{r^2}{\rho_{00} c_0^2} \left(\frac{g}{f^2 - N_0^2} \frac{\partial P_1}{\partial r} - 4\omega^2 P_1 \right) - \frac{\varepsilon}{\rho_{00}} P_1 = 0; \quad (10.3)$$

$$\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \frac{\sin \theta}{f^2 - \cos^2 \theta} \left(\frac{\partial}{\partial \theta} + \frac{m}{f} \operatorname{ctg} \theta \right) + \right. \\ \left. + \frac{1}{f^2 - \cos^2 \theta} \left(-\frac{m}{f} \operatorname{ctg} \theta \frac{\partial}{\partial \theta} - \frac{m^2}{\sin^2 \theta} \right) \right] P_2 + \varepsilon P_2 = 0, \quad (10.4)$$

where $\varepsilon = 4r_0^2 \omega^2 / gh_0$ is an arbitrary constant, appearing during the separation of variables (in the case of a homogeneous ocean the arbitrary constant h_0 introduced here is equal to the ocean depth; in general it is known as the *dynamically equivalent depth* for the eigen oscillation being considered). As our boundary condition for Equation (10.3), at the *rigid* lower boundary of the layer (the *bottom* $r = r_0 - h$) we set the vertical velocity $u_r \sim M_r(\nabla P) \sim \partial P / \partial r$ equal to zero, while at the *free* upper surface we invoke condition (7.16) (in which here to simplify the calculations we set $\gamma_0 = 0$, that is, we neglect capillary effects; in an atmosphere, which has no distinct upper surface, instead of this condition it is sufficient to assume boundedness of the solution).

Equation (10.3) with these boundary conditions will have a nontrivial solution $P_1 \neq 0$ only for certain special values of frequencies f , which for fixed physical parameters will be functions $f = f_1(\varepsilon)$ of the constant of the separation of variables (describing the so-called *eigencurves* of the 'vertical' equation (10.3) in the ε, f plane). Similarly, the 'horizontal' equation (10.4) has a nontrivial solution $P_2 \neq 0$, bounded for $0 \leq \theta \leq \pi$, only for certain special values $f = f_2(\varepsilon)$. Now, the eigenfrequencies of the Laplace tidal equation will correspond to the points of intersection between the 'vertical' and 'horizontal' eigencurves $f = f_1(\varepsilon)$ and $f =$

$f_2(\varepsilon)$. It should be noted that all the effects of compressibility, gravity, stratification, and the conditions at the upper boundary of the layer show up only in equation (10.3), whereas the effects of the rotation and sphericity of the planet show up only in Equation (10.4).

Following Kamenkovich (1977), let us rewrite Equation (10.3), after a simple renormalization of the density $\rho_{00} \rightarrow \rho_1$, as a system of two equations of first order:

$$\begin{aligned} \frac{\partial P_1}{\partial z} + (\sigma^2 - N^2)\rho_1 W &= 0; & \frac{\partial W}{\partial z} + \left(\varepsilon_1 - \frac{1}{c_0^2} \right) \frac{P_1}{\rho_1} &= 0; \\ \frac{\partial P_1}{\partial z} &= 0 \quad \text{for } z = -h; & P_1 + g\rho_1 W &= 0 \quad \text{for } z = 0, \end{aligned} \quad (10.3')$$

where $\varepsilon_1 = \varepsilon/4r_0^2\omega^2 = (gh_0)^{-1}$, and $W(z)$ is the normalized vertical profile of the vertical velocity. Note that the eigenfrequencies are real, due to the adiabaticity; the realness of ε_1 will be proven below. Therefore, eigenfunctions P_1 and W can also be assumed to be real. Now, after differentiating the first equation (10.3') with respect to ε_1 , multiplying the result by W , eliminating from it the quantity $(\sigma^2 - N^2)\rho_1 W$ with the aid of the initial equation, and then integrating over z , we get

$$\begin{aligned} \frac{\partial \sigma^2}{\partial \varepsilon_1} \int \rho_1 W^2 dz &= \int \left(\frac{\partial P_1}{\partial \varepsilon_1} \frac{\partial W}{\partial z} - W \frac{\partial^2 P_1}{\partial z \partial \varepsilon_1} \right) dz = \\ &= \int \left(\frac{\partial P_1}{\partial \varepsilon_1} \frac{\partial W}{\partial z} - P_1 \frac{\partial^2 W}{\partial \varepsilon_1 \partial z} \right) dz. \end{aligned} \quad (10.5)$$

The right-hand side of (10.5) was obtained by integration by parts using boundary conditions (10.3'). If we eliminate $\partial^2 W / \partial \varepsilon_1 \partial z$ from it with the aid of the second Equation (10.3'), this side becomes $\int P_1^2 \rho_1^{-1} dz$, so that from (10.5) it follows that $\partial \sigma^2 / \partial \varepsilon_1 > 0$. Consequently, the eigencurves of the 'vertical' equation in the ε_1 σ plane are, firstly, symmetrical about the axis $\sigma = 0$ and, secondly, rise monotonically with increasing ε_1 for $\sigma > 0$ (for $\sigma < 0$, on the other hand, they drop monotonically). Similarly to the derivation of Equation (10.5), we now calculate the quantity $W(\partial P_1^* / \partial z) + P_1^*(\partial W / \partial z)$ with the aid of equation (10.3') and integrate it over z to obtain

$$\int (N^2 - \sigma^2)\rho_1 |W|^2 dz + g\rho_1 |W|_{z=0}^2 = \int (\varepsilon_1 - c_0^{-2})\rho_1^{-1} |P_1|^2 dz, \quad (10.6)$$

from which we see that ε_1 is real and that for $\sigma^2 < N_{\min}^2$ it must be true that $\varepsilon_1 \geq c_{0\max}^{-2}$.

The complete calculation of all the 'vertical' eigencurves is carried out without difficulty in the simplest case of an incompressible fluid ($c_0 \rightarrow \infty$), when in

Equations (10.3') ρ_1 and N can be taken to be quasicontant and they yield

$$\frac{\partial^2 W}{\partial z^2} + \varepsilon_1(N^2 - \sigma^2) W = 0 \quad (10.7)$$

with boundary conditions $W = 0$ for $z = -h$ and $\partial W / \partial z = g\varepsilon_1 W$ for $z = 0$. For $\varepsilon_1 > 0$ and $\sigma^2 > N^2$ we obtain

$$W = \text{sh}[\varepsilon_1^{1/2}(\sigma^2 - N^2)^{1/2}(z + h)]; \quad (10.8a)$$

$$\text{th}(b\varepsilon_1^{1/2}) = a\varepsilon_1^{-1/2}; \quad a = g^{-1}(\sigma^2 - N^2)^{1/2}; \quad b = h(\sigma^2 - N^2)^{1/2}.$$

This equation in ε_1 has a single root, which we will call ε_0 . The eigencurve $\varepsilon_1 = \varepsilon_0(\sigma)$ will be called the 'zero' curve. On it $\varepsilon_0 \rightarrow \infty$ ($\sigma \rightarrow \infty$) and asymptotically $\varepsilon_0 \sim \sigma^2 g^{-2}$, while for $\sigma \rightarrow 0$ we get $\varepsilon_0(0) \approx (gh)^{-1}$. For $\varepsilon_1 > 0$ and $\sigma^2 < N^2$ instead of (10.8a) we have

$$W = \sin[\varepsilon_1^{1/2}(N^2 - \sigma^2)^{1/2}(z + h)]; \quad (10.8b)$$

$$\text{tg}(b\varepsilon_1^{1/2}) = a\varepsilon_1^{-1/2}; \quad a = g^{-1}(N^2 - \sigma^2)^{1/2}; \quad b = h(N^2 - \sigma^2)^{1/2}.$$

This equation has a denumerable set of roots, determined approximately by the formula $\varepsilon_n \approx (\pi n)^2 h^{-2}(N^2 - \sigma^2)^{-1}$; $n = 1, 2, 3, \dots$. Let us call the corresponding eigencurves 'positive'. All these possess an asymptote $\sigma^2 \sim N^2$ ($\varepsilon_1 \rightarrow \infty$). For $\varepsilon_1 < 0$ and $\sigma^2 < N^2$ function W is defined by the same formula as in (10.8a), while the eigencurves are given by the equation

$$\text{th}[b(-\varepsilon_1)^{1/2}] = -a(-\varepsilon_1)^{-1/2}; \quad a = g^{-1}(N^2 - \sigma^2)^{1/2}; \quad (10.8c)$$

$$b = h(N^2 - \sigma^2)^{1/2}.$$

which has no solutions, so that in this quadrant of the $\varepsilon_1\sigma$ plane no eigencurves exist.

Finally, for $\varepsilon_1 < 0$ and $\sigma^2 > N^2$ function W is defined by formula (10.8b), and the eigencurves are given by the equation

$$\text{tg}[b(-\varepsilon_1)^{1/2}] = -a(-\varepsilon_1)^{-1/2}; \quad a = g^{-1}(\sigma^2 - N^2)^{1/2}; \quad (10.8d)$$

$$b = h(\sigma^2 - N^2)^{1/2},$$

having a denumerable set of roots $\varepsilon_n \approx -(\pi n)^2 h^{-2}(\sigma^2 - N^2)^{-1}$, $n = -1, -2, -3, \dots$. Let us call these eigencurves 'negative'. They all have asymptotes $\varepsilon_1 \sim 0$ ($\sigma \rightarrow \infty$) and $\sigma^2 \sim N^2$ ($\varepsilon_1 \rightarrow -\infty$).

Returning from (10.7) to the general case of Equations (10.3') with variable coefficients, we discover that the overall character of the eigencurves in the $\varepsilon_1\sigma$ plane remains essentially the same (although the asymptotes are shifted), there being once again a 'zero' eigencurve (with the same asymptote $\varepsilon_0 = \sigma^2/g^2$ as above), 'positive' eigencurves (with asymptotes $\varepsilon_1 \sim N_{\max}^2$ as $\varepsilon_1 \rightarrow \infty$), and 'negative' eigencurves (with asymptotes $\varepsilon_1 \sim \sigma_{0\min}^{-2}$ for $\sigma^2 \rightarrow \infty$ and $\sigma^2 \sim N_{\min}^2$ for $\varepsilon_1 \rightarrow -\infty$); see Figure 10.1.

Let us demonstrate the means used to prove these statements, using as an

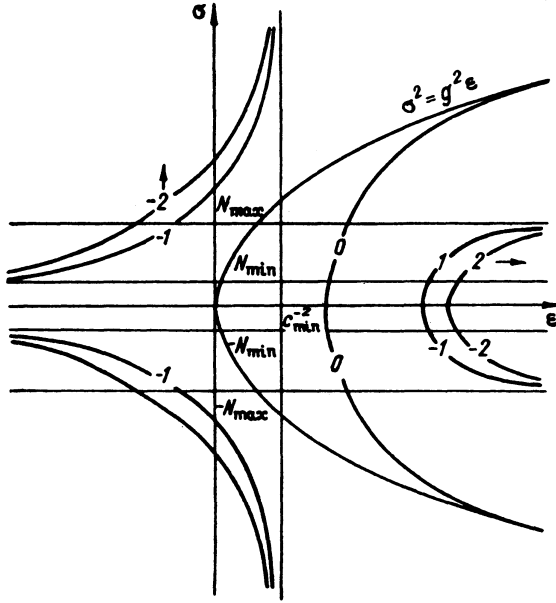


Fig. 10.1. 'Vertical' eigencurves.

example the region $\sigma^2 < N_{\min}^2$, $\varepsilon_1 > c_{0\max}^{-2}$ of the $\varepsilon_1\sigma$ plane. After eliminating function W from Equations (10.3'), we arrive at the following equation and boundary conditions for P_1 :

$$\begin{aligned} \frac{\partial}{\partial z} K \frac{\partial P_1}{\partial z} - GP_1 &= 0; & K &= \rho_1^{-1}(N^2 - \sigma^2)^{-1}; & G &= \rho_1^{-1}(c_0^{-2} - \varepsilon_1); \\ \frac{\partial P_1}{\partial z} &= 0 \quad \text{for } z = -h; & P_1 + g(N^2 - \sigma^2)^{-1} \frac{\partial P_1}{\partial z} &= 0 \quad \text{for } z = 0. \end{aligned} \quad (10.9)$$

Let $P(z)$ be the solution of this equation for the boundary conditions $P = 1$, $\partial P/\partial z = 0$ at $z = -h$, and let us consider the function $F(\varepsilon_1) = K(0)P^{-1}(0) \partial P(0)/\partial z$. As ε_1 increases from $-\infty$ to $+\infty$, function K does not change, G decreases, and according to Sturm's first theorem, the number of zeros of function $P(z)$ in the interval $-h \leq z \leq 0$ can only increase; each new zero originates at the point $z = 0$, and then it shifts to the point $z = -h$ (beyond which the zeros cannot pass). Since $G \rightarrow -\infty$ ($\varepsilon_1 \rightarrow \infty$), the values $\bar{\varepsilon}_n$ of parameter ε_1 for which $P(0) = 0$ will be infinitely many. According to Sturm's second comparison theorem, in the intervals $(-\infty, \bar{\varepsilon}_1)$, $(\bar{\varepsilon}_1, \bar{\varepsilon}_2)$, $\dots$, function $F(\varepsilon_1)$ will be a monotonically decreasing function. If in (10.9) we replace K by $K_{\min} = \rho_1^{-1} \max(N_{\max}^2 - \sigma^2)^{-1}$ and G by $G_{\min} = \rho_1^{-1} \max(c_{0\max}^{-2} - \varepsilon_1)$, then with the aid of the second comparison theorem we can readily show that $F \rightarrow \infty$ ($\varepsilon_1 \rightarrow -\infty$). The intersections of the $F = F(\varepsilon_1)$ curves with the line $F = -[g\rho_1(0)]^{-1}$ give the eigenvalues $\varepsilon_0, \varepsilon_1, \varepsilon_2, \dots$.

For $\varepsilon_1 > c_{0\max}^{-2}$ in the interval $-h \leq z \leq 0$ there will always be a subinterval in which $G < 0$. Then as $\sigma^2 \rightarrow -\infty$ function K decreases, and according to the first

comparison theorem $P(z)$ will oscillate in the specified subinterval, and thus over the whole interval. Therefore, for any fixed $\varepsilon_1 > c_{0\max}^{-2}$ there will exist an infinite number of eigenvalues σ^2 . Since the eigencurves do not intersect, for $\sigma^2 \rightarrow -\infty$ they will all approach the vertical asymptote $\varepsilon_1 = c_{0\max}^{-2}$.

If we replace K and G by $K_{\min}$ and $G_{\min}$ and by $K_{\max}$ and $G_{\max}$, and if we find in an elementary manner for the transformed equations the eigenvalues ε_1 , then it is easy to derive the inequalities:

$$\begin{aligned} \frac{\rho_{1\min}}{\rho_{1\max}} \frac{(\pi n)^2}{(N_{\max}^2 - \sigma^2)h^2} + c_{0\max}^{-2} < \varepsilon_n(\sigma^2) < c_{0\min}^{-2} + \\ + \frac{\rho_{1\max}}{\rho_{1\min}} \frac{[\pi(n+1)]^2}{(N_{\min}^2 - \sigma^2)h^2}, \end{aligned} \quad (10.10)$$

which give some idea of the configuration of the eigencurves $\varepsilon_n(\sigma^2)$. Note, too, that the eigencurves cannot intersect the line $\sigma^2 = g^2\varepsilon_1$, so that in view of their monotonicity they will not possess a vertical asymptote for $\varepsilon_1 > c_{0\max}^{-2}$ and they must enter the region $\varepsilon_1 > c_{0\min}^{-2}$. Their behavior in this region, as well as in other regions of the $\varepsilon_1\sigma^2$ plane, can be considered similarly.

Now let us examine the 'horizontal' Equation (10.4). Note that for this equation the poles $\theta = 0, \pi$ constitute singularities (in the vicinity of each of these poles there exists both a solution in the form of the product of $(\sin \theta)^m$ times an analytic function and a solution, inapplicable for us, with logarithmic branching) and 'critical latitudes', in which $\cos \theta = \pm f$ (but in the vicinities of which the solutions are continuous and even analytical). If instead of the latitude we introduce the variable $\mu = \cos \theta$, then we can write (10.4) as a system of two first-order equations:

$$\left[(1 - \mu^2) \frac{\partial}{\partial \mu} - \frac{m}{f} \mu \right] \psi = (f^2 - \mu^2) \xi; \quad (10.11)$$

$$\left[(1 - \mu^2) \frac{\partial}{\partial \mu} + \frac{m}{f} \mu \right] \xi = \left[\frac{m^2}{f^2} - (1 - \mu^2) \varepsilon \right] \psi, \quad (10.12)$$

where $\psi = P_2$ is the old unknown, and the first of these equations is used to determine the additional unknown ξ . Next we multiply the first of these equations by the complex conjugate ξ^* and the equation complex-conjugate to the second one by ψ ; then, by adding up the results, dividing the sum by $1 - \mu^2$, and integrating over μ from -1 to $+1$, we obtain

$$\begin{aligned} \varepsilon^* = & \left(\int_{-1}^1 |\psi|^2 d\mu \right)^{-1} \left[\int_{-1}^1 \frac{f^2 - \mu^2}{1 - \mu^2} |\xi|^2 d\mu + \right. \\ & \left. + \int_{-1}^1 \frac{m^2 |\psi|^2 d\mu}{f^2(1 - \mu^2)} \right], \end{aligned} \quad (10.13)$$

from which it is evident that, firstly, ε is real (and may be either positive or negative) and, secondly, for $f^2 \geq 1$ (that is, for eigenperiods shorter than a half

day) ε is positive. Similarly, we can demonstrate the orthogonality $\int_{-1}^1 \psi_1 \psi_2^* d\mu = 0$ of eigenfunctions with different ε (for the same f). It is also easy to show that different eigencurves $f = f(\varepsilon)$ do not intersect.

The behavior of the eigencurves is easy to explain in the region of high $|f|$, where Equations (10.11) and (10.12) have an asymptotic form:

$$(1 - \mu^2) \frac{\partial \psi}{\partial \mu} = f^2 \xi; \quad (1 - \mu^2) \frac{\partial \xi}{\partial \mu} = \left[\frac{m^2}{f^2} - (1 - \mu^2) \varepsilon \right] \psi, \quad (10.14)$$

from which after eliminating ξ we get the Legendre equation

$$(1 - \mu^2) \frac{\partial^2 \psi}{\partial \mu^2} - 2\mu \frac{\partial \psi}{\partial \mu} - \frac{m^2}{1 - \mu^2} \psi + \varepsilon f^2 \psi = 0, \quad (10.15)$$

having regular solutions: the Legendre functions $\psi = P_n^m(\mu)$ for $\varepsilon f^2 = n(n+1)$, where $n = m, m+1, \dots$. Asymptotes $|f| \sim [n(n+1)/\varepsilon]^{1/2}$ for $\varepsilon \rightarrow +0$ are shown in Figure 10.2. They correspond to the so-called first type of eigencurves. However, as $\varepsilon \rightarrow 0$ not all of the eigencurves approach the ordinate axis: for low ε and finite f in Equations (10.11), (10.12) only the term $-(1 - \mu^2)\varepsilon\psi$ drops out asymptotically and after eliminating ψ from these we again obtain for ξ the Legendre equation

$$(1 - \mu^2) \frac{\partial^2 \xi}{\partial \mu^2} - 2\mu \frac{\partial \xi}{\partial \mu} - \frac{m^2}{1 - \mu^2} \xi + \frac{m}{f} \xi = 0 \quad (10.16)$$

with regular solutions for $m/f = n(n+1)$ (here ψ is a linear combination of P_{n+1}^m

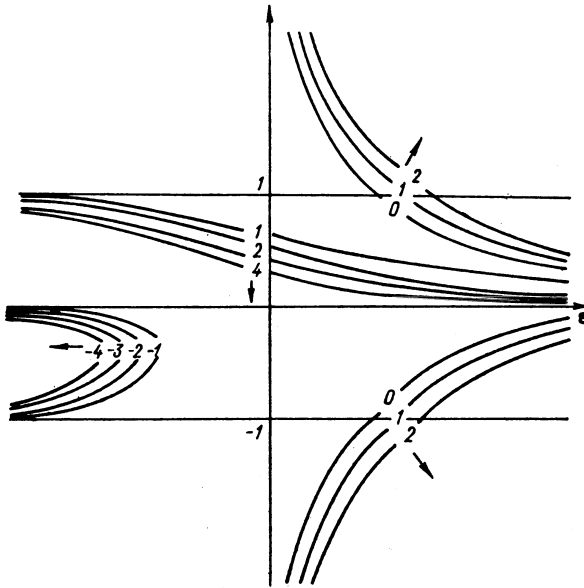


Fig. 10.2. 'Horizontal' eigencurves. The ordinate is f .

and P_{n-1}^m). These asymptotes (corresponding to the so-called second type of eigencurves) apply for both positive and negative values of ε . Note, too, that positive ratios m/f are obtained for waves propagating from east to west.

Now let us consider the asymptote of the eigencurves for $\varepsilon \rightarrow \infty$ (or $h \rightarrow 0$) and $f \rightarrow 0$, using the following equation implied by (10.11), (10.12):

$$\left(\mathcal{L} + \frac{m}{f}\right) \xi = \varepsilon \left[\left(\frac{m}{f} + 2\right) \mu - (1 - \mu^2) \frac{\partial}{\partial \mu} \right] \psi;$$

$$\mathcal{L} = \frac{\partial}{\partial \mu} (1 - \mu^2) \frac{\partial}{\partial \mu} - \frac{m^2}{1 - \mu^2}. \quad (10.17)$$

Since $f \rightarrow 0$, on the right-hand side of (10.17) we can set $(m/f) + 2 \approx m/f$ and, using Equation (10.11), we can replace it with $(\mu^2 - f^2)\xi$, which yields

$$-\varepsilon^{-1} \mathcal{L} \xi + \mu^2 \xi = E \xi; \quad E = f^2 + \frac{m}{f\varepsilon}, \quad (10.17')$$

and we have to study the asymptotic behavior of the eigenvalues E of the differential operator with a small parameter ε^{-1} for the higher derivative. This problem is analogous to the *quasiclassical approximation* in quantum mechanics, whereby the Schrödinger equation $-\hbar \partial^2 \psi / \partial x^2 + U(x) \psi = E \psi$ yields an asymptote for the total energy E for small Planck constants $\hbar \rightarrow 0$; in (10.17') μ^2 plays the role of the potential energy $U(x)$, and instead of $\partial^2 / \partial x^2$ we have operator $\mathcal{L}$. It is evident from (10.17') that the finite function $\xi(\mu)$ must have large derivatives, so that we can assume it to be 'localized', that is, it will be considerably different from zero only on a small segment of μ , for instance, for $\mu^2 \ll 1$. Then Equation (10.17') will there have the approximate form

$$\left(\frac{\partial^2}{\partial \theta^2} + E\varepsilon^{1/2} - \theta^2 \right) \xi = 0; \quad \theta = \mu\varepsilon^{1/4}. \quad (10.17'')$$

The solutions of this equation, determined for all θ and bounded for $\varepsilon \rightarrow \infty$, have the form $\varepsilon = \exp(-\theta^2/2) H_\nu(\theta)$, where H_ν are Hermite polynomials with numbers $\nu = 0, 1, 2, \dots$ satisfying the equation $E\varepsilon^{1/2} = 2\nu + 1$. Recalling the definition of E and considering this relation to be a cubic equation in f , we can write its roots in series form:

$$f = \pm (2\nu + 1)^{1/2} \varepsilon^{-1/4} - \frac{m\varepsilon^{-1/2}}{4\nu + 2} + \dots; \quad \nu = 0, 1, 2, \dots; \quad (10.18)$$

$$f = \frac{m\varepsilon^{-1/2}}{2\nu + 1} + \dots; \quad \nu = 1, 2, \dots \quad (10.19)$$

Substitution of the obtained series into the initial equation for ξ verifies that they are valid for all ν in (10.19) except $\nu = 0$ (which was, therefore, excluded), corresponding to the case $m^2 = \varepsilon f^2$, requiring separate study. For $f = -m\varepsilon^{-1/2}$ from (10.11) and (10.12) we get $\psi \approx \exp(-\theta^2/2)$, $\xi \approx 0$, while for $f = +m\varepsilon^{-1/2}$ no solution exists.

We still have to examine the asymptotes of the eigencurves for $\varepsilon \rightarrow -\infty$ (with $f^2 < 1$). First let us consider eigencurves of the second type. Numerical calculations show that $f \rightarrow 1$ as $\varepsilon \rightarrow -\infty$, and from (10.11) and (10.12) we see that $1 - \mu^2$ has to be small, meaning that function $\varepsilon(\mu)$ is ‘localized’ near the poles. For example, near the north pole, introducing the independent variable $x = (-\varepsilon)^{1/2} \theta^2$, where θ is the colatitude, and seeking the frequencies in the form $f = 1 - q(-\varepsilon)^{-1/2} + O(\varepsilon^{-1})$, where coefficient q is to be determined, we obtain the following approximate equation for ξ :

$$\frac{\partial^2 \xi}{\partial x^2} + \left(-\frac{1}{4} + \frac{q}{2x} - \frac{m^2 - 2m}{4x^2} \right) \xi = 0. \quad (10.20)$$

A bounded (for $0 \leq x < \infty$) solution here exists only for $q = m + 2\nu$; $\nu = 0, 1, 2, \dots$, and takes the form

$$\xi = e^{-x/2} x^{m/2} \frac{d^{m-1}}{dx^{m-1}} L_{m-1+\nu}(x),$$

where $L_n(x)$ are Laguerre polynomials. A similar asymptote is obtained near the south pole.

Finally, let us discuss eigencurves of the third type, with an asymptote of one of branches $f = -1 + q(-\varepsilon)^{-1/2} + O(\varepsilon^{-1})$; here, similarly to (10.20), we obtain $q = m + 2\nu + 2$; $\nu = 0, 1, 2, \dots$ and

$$\xi = e^{-x/2} x^{(m+2)/2} \frac{d^{m+1}}{dx^{m+1}} L_{m+1+\nu}(x).$$

The second branch of these eigencurves for $\varepsilon \rightarrow -\infty$ has an asymptote $f = 0$, so that we can set

$$f = -m(-\varepsilon)^{-1} - 2qm(-\varepsilon)^{-3/2} + O(\varepsilon^{-2}),$$

and, similarly to (10.20), we obtain $q = m + 2\nu + 1$; $\nu = 0, 1, 2, \dots$ and

$$\xi = e^{-x/2} x^{m/2} \frac{d^m}{dx^m} L_{m+\nu}(x).$$

All the above asymptotes of the ‘horizontal’ eigencurves were plotted in Figure 10.2. The case $m = 0$ (zonal motion) has to be considered separately; curves of the second type do not appear in this case.

Intersections between the ‘vertical’ and ‘horizontal’ eigencurves (for $N_{\min} < 2\omega < N_{\max}$) are shown in Figure 10.3. Points $1^\pm$ of intersection between the negative ‘vertical’ curves and the ‘horizontal’ curves of the first type correspond to acoustic waves: in an incompressible fluid (for $c_0 \rightarrow \infty$) they vanish, but they remain for $\rho_0 = \text{const}$, $N_0 = 0$, and in a nonrotating plane. For these points the following inequality is valid:

$$\sigma^2 > N_{\min}^2 + \frac{\rho_{1 \min}}{\rho_{1 \max}} \left(\frac{\pi c_{0 \min}}{2h} \right)^2. \quad (10.21)$$

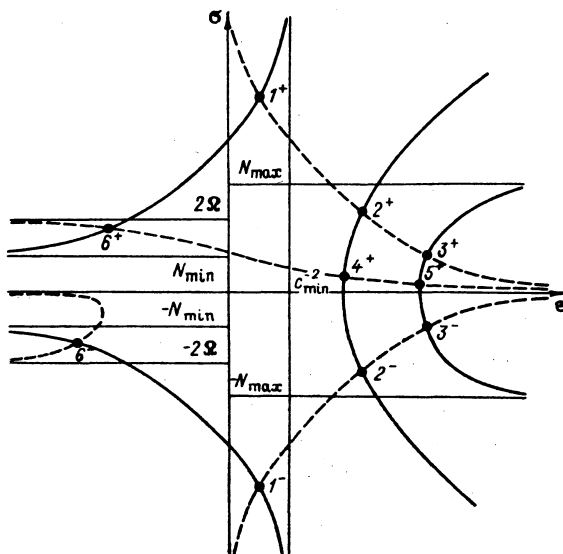


Fig. 10.3. Intersections of 'vertical' (solid) and 'horizontal' (dashed) eigencurves for $N_{\min} < 2\omega < N_{\max}$. Only one pair of eigencurve intersections is shown for each family. Intersections $1^\pm$ correspond to acoustic waves; $2^\pm$ to surface waves; $3^\pm$ to internal gravity waves; $4^\pm$ to barotropic waves; $5^\pm$ to baroclinic Rossby waves; $6^\pm$ to gyroscopic waves.

On curves of the first type there also exist points $2^\pm$ of intersection with the zero 'vertical' curve, corresponding to surface gravity waves (the zero curve vanishes when the free surface is replaced by a rigid plane), and points $3^\pm$ of intersection with the positive 'vertical' curves correspond to internal gravity waves, for which from (10.10) we obtain

$$0 < |\sigma| < N_{\max}; \quad \varepsilon > c_{0\max}^{-2} + \frac{\rho_{1\min}}{\rho_{1\max}} \left(\frac{\pi}{N_{\max}h} \right)^2. \quad (10.22)$$

Points $4^\pm$ of intersection of the 'horizontal' curves of the second type with the zero 'vertical' curve correspond to *barotropic* Rossby waves, produced by the combined effect of the rotation and sphericity of the planet (the condition for existence of curves of the second type); for these points $0 < \sigma < \omega$. Point $5^\pm$ of intersection with the positive 'vertical' curves correspond to *baroclinic* Rossby waves (these disappear for $N_{\max} \rightarrow 0$); for these points $0 < \sigma < \min(\omega, N_{\max})$. The Rossby waves propagate only toward the west.

Finally, points $6^\pm$ of intersection of the negative 'vertical' curves with 'horizontal' curves of the second and third types correspond to *gyroscopic* waves, produced by sufficiently rapid rotation ($2\omega > N_{\min}$); for these points

$$N_{\min} < |\sigma| < 2\omega; \quad \varepsilon < c_{0\min}^{-2} - \frac{\rho_{1\min}}{\rho_{1\max}} \frac{\pi^2}{4h^2(4\omega^2 - N_{\min}^2)}. \quad (10.23)$$

Hydrodynamic Instability

11. INSTABILITY OF AN IDEAL FLUID

In the preceding chapter we considered adiabatic small oscillations of a fluid relative to a state of rest (or relative to steady motion, described by one of the solutions of the linearized fluid-dynamical equations). From the invariability in time of the total energy of the eigen oscillations (normal modes) (7.7), we deduced their frequencies σ to be real (and then verified this deduction by calculating all the possible such frequencies), so that small initial perturbations did not grow and did not decay with time. Thus in Chapter 2 it was invariably assumed that the stratification of the density is stable, that is $N_0^2 > 0$.

Now let us examine the opposite case of unstable stratification of the density, when $N_0^2 = -M^2 < 0$. We limit ourselves here (and throughout the chapter) to the Boussinesq approximation, replacing continuity Equation (7.1) by the condition of solenoidality of the velocity field (whereby acoustic waves are filtered out) $\text{div } \mathbf{u} = 0$. Moreover, we neglect for the time being the rotation of the reference system ($\omega \rightarrow 0$), so that Equations (7.1)–(7.3) become

$$\text{div } \mathbf{u} = 0; \quad \frac{\partial \mathbf{u}}{\partial t} = -\frac{\nabla p''}{\rho_{00}} + g\eta'' \nabla r; \quad \frac{\partial \eta''}{\partial t} = \frac{M^2}{g} u_r. \quad (11.1)$$

Instead of the local energy Equation (7.6) we now have

$$\frac{1}{2} \frac{\partial}{\partial t} \left(\rho_{00} u^2 - \rho_{00} \frac{g^2 \eta'^2}{M^2} \right) + \text{div } p'' \mathbf{u} = 0, \quad (11.2)$$

and, for example, for a fluid layer $0 \leq z \leq h$ between horizontal rigid walls we obtain

$$\frac{\partial}{\partial t} \left(\int \frac{v^2}{2} \rho_{00} dV - \int \frac{g^2 \eta'^2}{2M^2} \rho_{00} dV \right) = 0, \quad (11.3)$$

that is, the *difference* between the total kinetic energy and the total thermobaric energy of the perturbations does not vary with time. Regarding the individual behavior of each of these two kinds of energy, however, Equation (11.3) tells us nothing. Soon we will see, though, that both of these energies increase exponentially with time.

We consider only two-dimensional perturbations of the motion in the xz plane, and we introduce, using the nondivergence of the velocity, the stream function ψ , where $u = -\partial\psi/\partial z$ and $w = \partial\psi/\partial x$. Now, taking the curl of the equations of motion (assuming quasiconstancy of the potential density ρ_{00}) and eliminating η'' with the aid of the third Equation of (11.1), we arrive at the following equation for ψ :

$$\frac{\partial^2 \Delta \psi}{\partial t^2} - M^2 \frac{\partial^2 \psi}{\partial x^2} = 0 \quad (11.4)$$

with boundary conditions $w = 0$, that is, $\psi = 0$ for $z = 0, h$. Since M^2 is independent of x and t , elementary solutions of this equation can be sought in the form $\psi = \psi(z) \exp[i(kx - \sigma t)]$. For the complex amplitude $\psi(z)$ we then get from (11.4) the equation

$$\sigma^2 \left(\frac{\partial^2 \psi}{\partial z^2} - k^2 \psi \right) - k^2 M^2 \psi = 0. \quad (11.5)$$

For quasiconstant M the solution, going to zero at $z = 0$, has the form $\psi \sim \sin mz$, where $m = -ik(1 + M^2/\sigma^2)^{1/2}$. If ψ also goes to zero at $z = h$, then we can set $mh \equiv \pi n$; $n = 0, \pm 1, \pm 2, \dots$, so that

$$\sigma = \pm iM \left(1 + \frac{\pi^2 n^2}{k^2 h^2} \right)^{-1/2}, \quad (11.6)$$

meaning that for every fixed k there exists a denumerable set of elementary solutions, for all of which the values of σ are purely imaginary. Accordingly, here there will always exist elementary solutions which increase with time. This phenomenon is known as static (Taylor) instability. Its mechanism was explained in Section 2 as an accelerating effect of the buoyancy forces on fluid particles moving along the vertical from their equilibrium position. In this case, according to the third Equation (11.1), the interaction of the vertical velocity u_r with the unstable potential-density gradient M^2/g causes an increase in the entropy perturbations η'' . Likewise, according to the second Equation (11.1), their contribution $g\eta''$ to the buoyancy acceleration leads to an increase in u_r .

Small perturbations can be augmented by certain forms of the velocity distribution in the equilibrium state, as well as by an unstable density distribution. Let us consider, for example, the instability of a plane-parallel steady flow of an incompressible ideal fluid, directed along the x axis and having a velocity $u_0 = \{U(z), 0, 0\}$. Then, instead of (7.1), (7.2), the linearized fluid-dynamical equations will take the following form:

$$\operatorname{div} \mathbf{u} = 0; \quad \frac{\partial \mathbf{u}}{\partial t} + U \frac{\partial \mathbf{u}}{\partial x} + w \frac{\partial \mathbf{u}_0}{\partial z} = - \frac{\nabla p'}{\rho_0}. \quad (11.7)$$

The local energy equation accordingly becomes

$$\frac{1}{2} \frac{\partial}{\partial t} \rho_0 |\mathbf{u}|^2 + \frac{1}{2} \frac{\partial}{\partial x} \rho_0 U |\mathbf{u}|^2 + \rho_0 u w \frac{\partial U}{\partial z} + \operatorname{div} p' \mathbf{u} = 0, \quad (11.8)$$

from which, for instance for the layer $0 \leq z \leq h$ between horizontal rigid walls, we obtain

$$\frac{\partial}{\partial t} \int \frac{|\mathbf{u}|^2}{2} \rho_0 dV = - \int \rho_0 u w \frac{\partial U}{\partial z} dV, \quad (11.9)$$

indicating that the kinetic energy of the perturbations can change with time because of the stress work $-\rho_0 u w$ on the gradient of the equilibrium velocity $\partial U / \partial z$. As above, we consider only two-dimensional perturbations $u = -\partial \psi / \partial z$, $w = \partial \psi / \partial x$. Then, by taking the curl of Equation (11.7), we arrive at an equation for ψ analogous to (11.4):

$$\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \Delta \psi - \frac{\partial^2 U}{\partial z^2} \frac{\partial \psi}{\partial x} = 0. \quad (11.10)$$

However, the behavior of the solutions ψ of this equation turns out to be much more complicated than in the case of (11.4). Let us once again seek elementary wave solutions in the form $\psi = \psi(z) \exp[i(kx - \sigma t)]$. Then for the complex amplitude $\psi(z)$ we get the *Rayleigh equation*

$$(U - c) \left(\frac{\partial^2 \psi}{\partial z^2} - k^2 \psi \right) - \frac{\partial^2 U}{\partial z^2} \psi = 0, \quad (11.11)$$

where $c = \sigma/k$ is the phase velocity. Firstly, this equation possesses a singularity z_0 , at which for neutral perturbations (real c) the coefficient $U - c \approx U'(z_0)(z - z_0)$ of the higher derivative $\partial^2 \psi / \partial z^2$ can go to zero, so that there the derivative $\partial \psi / \partial z \sim (\psi U' / U)_0 \ln(z - z_0)$ is multivalued, and the problem arises of choosing the required branch of the solution. Secondly, the set of discrete eigenvalues c is, as a rule, only finite, so that by no means can each solution of equation (11.10) be represented as a superposition of the corresponding elementary wave solutions, in order to evaluate the stability. Here, it is true, there is also a continuum of eigenvalues c , as is evident if we rewrite Equation (11.11) as

$$(U - U'' \mathcal{L}^{-1}) \psi_1 = c \psi_1; \quad \mathcal{L} = \frac{\partial^2}{\partial z^2} - k^2; \quad \psi_1 = \mathcal{L} \psi \quad (11.11')$$

(here and below, a prime signifies the derivative with respect to z), which implies that c is an eigenvalue of the sum of the operator of multiplication by function $U(z)$ (having a real continuous spectrum, filling the interval $U_{\min} \leq c \leq U_{\max}$) and the entirely continuous operator $-U'' \mathcal{L}^{-1}$ (the addition of which does not alter the continuous spectrum). However, operator $U'' \mathcal{L}^{-1}$ is not self-adjoint, and the theorem on the resolution of an arbitrary function into a discrete and continuous spectrum does not apply here. Thirdly, because the indicated operator is not self-adjoint, even the real eigenvalues c of, say, the discrete spectrum, if they are degenerate, may lead to instability in the guise of 'secular' perturbations, which increase linearly with time.

The example of the Rayleigh equation shows that it is advisable to start from a

more general definition of hydrodynamic instability than that ascribing eigenvalues with negative imaginary parts to the linearized equations. In order to arrive at such a general definition, let us introduce the concept of the *phase space* of the fluid, the 'points' M of which constitute complete sets of independent (that is, not connected by synchronous relations) thermohydrodynamic fields, characterizing the instantaneous states of the moving fluid. In the case of an incompressible fluid this is the velocity field $\mathbf{u}(\mathbf{x})$ in the region of space occupied by the fluid, satisfying the necessary boundary conditions; in the general case we must add to the velocity field the density field $\rho(\mathbf{x})$, the entropy field $\eta(\mathbf{x})$, and the admixture-concentration field $s(\mathbf{x})$; in the Galerkin approximations (6.3) these are the sets of coefficients $u^1, \dots, u^n$ (playing the role of generalized coordinates or degrees of freedom of the flow; there are thus, generally speaking, an infinite number of these). The evolution of the fluid flow in time is represented in phase space by some line $M = M(t)$, the phase trajectory of the flow; for a steady flow it consists of a single point, while for a periodic flow it is a closed line (cycle). The set $M(t) = F^t M(0)$ of phase trajectories passing through all the points of the phase space $M = M(0)$ and extended onto the entire time axis defines the group of mappings of F^t onto the phase space, known as the *phase flow*. It describes the evolution of all the fluid flows in the given geometry for all possible initial data.

Now let us introduce in the phase space the norm of its elements $\|M\|$. Then the general definition of stability of the phase trajectory $M = M_0(t)$, according to Lyapunov, is as follows: for each arbitrarily small positive number ε there exists some positive number $\delta = \delta(\varepsilon)$ such that for any trajectory $M = M(t)$ with an initial value $M(0)$ satisfying the condition $\|M(0) - M_0(0)\| < \delta$ the inequality $\|M(t) - M_0(t)\| < \varepsilon$ will be satisfied for all $t > 0$. It is easy to see that, if only a single unstable infinitesimal wave perturbation $M'(t) = M(t) - M_0(t)$ is present (with a negative imaginary part $\gamma = \text{Im } \sigma < 0$ of the eigenfrequency σ), Lyapunov instability of the trajectory $M_0(t)$ ensues. Actually, while the perturbation $M'(t)$ is small, it grows according to the $e^{|\gamma|t}$ law of linear theory, but then the nonlinear terms of the equations hinder this growth and, as a rule, some final limit is reached. Lowering the amplitude of the initial perturbation only serves to lengthen this process, without changing the final limit; this nonuniformity in time of the convergence to zero also signifies that Lyapunov stability is absent. Since in reality small perturbations are always present, linear instability of the flow $M_0(t)$ means that the flow is unrealizable in practice (for instance, a state of rest for unstable stratification); this is actually the main significance of the concept of hydrodynamic instability.

In a number of cases the reverse can also be shown to be true: that linear stability guarantees Lyapunov stability. For instance, for Equation (11.10) the following theorem is valid (Dikii, 1976): a two-dimensional plane-parallel flow with a monotonic velocity profile $U(z)$, $0 \leq z \leq h$, in which $U(0)$ and $U(h)$ are not eigenvalues, can be unstable only if nonreal or multiple real eigenvalues are present in the discrete spectrum. The proof of this is based on solution of the Cauchy problem for Equation (11.10) (with ψ depending on x according to an e^{ikx} law) for an arbitrary initial value $\psi(z, 0) = \psi_0(z)$, with the aid of a Laplace

transform with respect to time:

$$\psi^*(z, s) = \int_0^\infty e^{ikst} \psi(z, t) dt; \quad (11.12)$$

$$(U - s) \left(\frac{\partial^2}{\partial z^2} - k^2 \right) \psi^* - U'' \psi^* = f \equiv -\frac{i}{k} (\psi_0'' - k^2 \psi_0), \quad (11.13)$$

which yields

$$\psi^*(z, s) = \int_0^h G^*(z, \xi, s) f(\xi) d\xi; \quad (11.14)$$

$$G^*(z, \xi, s) = \begin{cases} \psi_1(z) \psi_2(\xi) W^{-1}[U(\xi) - s]^{-1}; & z < \xi, \\ \psi_1(\xi) \psi_2(z) W^{-1}[U(\xi) - s]^{-1}; & z > \xi, \end{cases} \quad (11.15)$$

where ψ_1, ψ_2 are the solutions of the homogeneous Equation (11.13) for the conditions $\psi_1(0) = 0, \psi_1'(0) = 1$, and $\psi_2(h) = 0, \psi_2'(h) = 1$, while $W = \psi_1(h) = -\psi_2(0)$ is their Wronskian. Then

$$\psi(z, t) = \int_0^h G(z, \xi, t) f(\xi) d\xi; \quad (11.16)$$

$$G(z, \xi, t) = \frac{k}{2\pi} \int_{i\gamma - \infty}^{i\gamma + \infty} e^{-ikst} G^*(z, \xi, s) ds, \quad (11.17)$$

where $\gamma > 0$ is so large that above the integration contour the function G^* is analytic in s . In order to prove the boundedness of $\psi(z, t)$ as $t \rightarrow \infty$, it is necessary to study the behavior of integral (11.17). It is easy to see that function $W(s)$ continues analytically from the upper halfplane through the real axis, everywhere except at the points $s = U(0)$ and $U(h)$, and that it goes through zero wherever s is equal to the eigenvalues of the homogeneous equation. Then function $G^*(s)$ continues analytically from the upper halfplane through the real axis, everywhere except at the points $s = U(0), U(h), U(z)$, and $U(\xi)$, and it goes through poles where (and with the same multiplicity) s equals the eigenvalues of the homogeneous equation. Function $\psi_1(z, s)$ at points $s = U(0), U(z)$ has a branching, but it can be shown to be continuous there; consequently, the Wronskian $W(s)$ is continuous everywhere in the region $\text{Im } s > -\varepsilon$ for sufficiently small $\varepsilon > 0$. Moreover, if $s = U(0)$ and $U(h)$ are not eigenvalues, then at the upper halfplane and in the vicinity of segment $\{U(0), U(h)\}$, on which $W(s)$ continues analytically for any cross sections from its ends, there will be a finite number of eigenvalues (actually, if $s = U(0)$ is not an eigenvalue, then $W = \psi_1(h, s) \neq 0$); due to continuity $\psi_1(z, s) \neq 0$ in the vicinity of point $s = U(0)$ either;

in addition, this function is analytic and cannot go through zero for large $|s|$. Therefore, eigenvalues with arbitrarily large moduli cannot exist either, and according to Hobard's theorem they all lie within a circle with its diameter on the segment $[U_{\min}, U_{\max}]$.

In the following the paths of integration over ξ in (11.14) and (11.16) will be deformed in such a way that they pass below the point $\xi = \xi_s$ at which $U(\xi_s)$ is equal to the real eigenvalues s , while in (11.17) the contour of integration over s will be lowered a certain distance below the real axis, and will rise from there along the edges of the vertical sections to the points $s = U(0)$, $U(z)$, $U(\xi)$, $U(h)$ and then pass around these points on circles with radii of the order of c/t , nowhere coming closer to the eigenvalues than at a distance of half the radius. Then, as $t \rightarrow \infty$ the integrals over the horizontal parts of the contour decrease exponentially, while on the circles and the vertical segments they remain bounded, and it remains only to add the sum of the residues at the poles between the old and new integration contours, at those points s which are eigenvalues. At a simple pole the residue is proportional to $G^*(z, \xi, s) e^{-ikst}$, and at a k th pole to $\partial^{k-1} \{G^*(z, \xi, s) e^{-ikst}\} / \partial s^{k-1}$. These residues distinguish from (11.16) the wave solutions, which are the only ones capable of causing an unrestricted growth of $\psi(z, t)$, which verifies Dikii's theorem.

We mention without proof that for stability of a flow with a velocity profile of 'type A', that is, with monotonic $U'(z) > 0$ and with one inflection point $z = \bar{z}$ (so that $\bar{U}'' = 0$), at which $U'(z)$ is a maximum (so that $\bar{U}''' < 0$), it is necessary and sufficient that there be no nonreal eigenvalues c at $k = 0$ or that the following condition of Rosenblut and Simon (1964) be satisfied:

$$[U'(U - \bar{U})]^{-1} \Big|_0^h + \int_0^h U''(U')^{-2} (U - \bar{U})^{-1} dz > 0. \quad (11.18)$$

In particular, it is sufficient that there exists a constant K_0 such that $(U - K_0) U'' \geq 0$ (Fjørtoft, 1950), or simply that there be no inflection points (Rayleigh, 1880). On the other hand, a flow with a 'type A' velocity profile for $0 \leq z \leq h/2$, symmetric about the point $z = \frac{1}{2}$, is always unstable (Tollmien, 1935). A tangential velocity discontinuity $U(z) = U_0$ for $z > 0$ and $(-U_0)$ for $z < 0$ is also unstable (this is known as Helmholtz instability).

The theorems of Rayleigh and Fjørtoft can be verified for arbitrary initial data, without considering the elementary wave solutions. For example, from Equation (11.10), for its solution ψ depending on x according to the e^{ikx} law, we obtain the integral relation

$$\frac{\partial}{\partial t} \int (|\psi'|^2 + k^2 |\psi|^2) dz - 2 \operatorname{Re} ik \int U(\psi'' - k^2 \psi) \psi^* dz = 0. \quad (11.19)$$

Assume there exists some constant K_0 such that function $(U - K_0) (U'')^{-1}$ is everywhere continuous. Then, multiplying (11.10) by $(\psi^{*''} - k^2 \psi^*) (U - K_0)$

$(U'')^{-1}$, integrating over z , taking the real part, and adding it to (11.19), we obtain

$$\frac{\partial}{\partial t} \int [|\psi'|^2 + k^2|\psi|^2 + (U - K_0)|\psi'' - k^2\psi|^2 (U'')^{-1}] dz = 0, \quad (11.20)$$

so that for Fjørtoft's condition smallness of this integral implies smallness of the mean square values of ψ , ψ' , and ψ'' at any moment of time, that is, stability of the flow. The square integral invariant in (11.20) can be interpreted, if we consider, following Arnol'd (1965), the Lyapunov instability of a nonlinear *two-dimensional* flow of an incompressible ideal fluid, the stream function $\psi(x, z, t)$ of which satisfies the vorticity equation

$$\frac{\partial \Delta \psi}{\partial t} + \frac{\partial(\psi, \Delta \psi)}{\partial(x, z)} = 0 \quad (11.21)$$

(its linearization relative to $\psi_1 = \psi + \int^z U(z) dz$ brings us back to Equation (11.10)), with boundary conditions specifying an impenetrable boundary $\psi(\Gamma) = \text{const}$ and conservation there of the circulation $\oint (\partial \psi / \partial n) d\Gamma = \text{const}$. Here not only is the kinetic energy (per unit mass) $E = \frac{1}{2} \int |\nabla \psi|^2 dx dz$ invariable with time, but also any functional of the vorticity $F = \int \Phi(\Delta \psi) dx dz$. In order to investigate the stability of a steady flow with stream function ψ_0 (which we will assume to be a monotonic function of the vorticity $\psi_0 = \Psi(\Delta \psi_0)$), let us choose a conserved functional $F(\psi) = E + F$ such that for $\psi = \psi_0$ it has an extremum (see the end of Section 5). Clearly, when the first variation δF goes to zero, we should set $\Phi' = \Psi$. For a plane-parallel flow with a velocity $u_0 = U(z)$ we will have $\Psi'(\Delta \psi_0) = U(z)/U''(z)$, and in a frame of reference moving relative to the initial frame at a constant velocity K_0 , we get

$$F(\psi) - F(\psi_0) = \frac{1}{2} \int [(\nabla \delta \psi)^2 + (U - K_0)(U'')^{-1}(\Delta \delta \psi)^2] dx dz + \dots \quad (11.22)$$

For the Fjørtoft condition this quadratic form relative to $\delta \psi$ is positive and can be taken for the square of the norm $\|\delta \psi\|_{(1)}^2$. For linearized equations it is, according to (11.20), an exact invariant. For nonlinear equations, on the other hand, it is the square of the norm $\|\delta \psi\|_{(2)}^2$ specified by the left-hand side of (11.22). However, these norms are equivalent, in the sense that there are positive constants $C_1 \leq C_2$ satisfying the inequality $C_1 \|\delta \psi\|_{(1)} \leq \|\delta \psi\|_{(2)} \leq C_2 \|\delta \psi\|_{(1)}$. If the initial norm $\|\delta \psi\|_{(1)}$ is small, then the invariant norm $\|\delta \psi\|_{(2)}$ will also be small, so that $\|\delta \psi\|_{(1)}$ always stays small, and here linear stability implies nonlinear Lyapunov stability.

In natural flows it often occurs that the stabilizing effect of a stable stratification $N_0^2(z) > 0$ is accompanied by the destabilizing effect of an unstable velocity profile $U(z)$. In such cases instead of Equations (11.4) or (11.10) we have the

following combination of these:

$$\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \left[\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \Delta \psi - U'' \frac{\partial \psi}{\partial x} \right] + N_0^2 \frac{\partial^2 \psi}{\partial x^2} = 0. \quad (11.23)$$

For elementary wave solutions of form $\psi(z) \exp[ik(x - ct)]$ we have

$$(U - c)[(U - c)(\psi'' - k^2 \psi) - U'' \psi] + N_0^2 \psi = 0. \quad (11.24)$$

Now, if we denote $U - c = W$ and introduce a new unknown function $\Psi = \psi W^{-1/2}$, then Equation (11.24) becomes

$$(W\Psi)' - \left[\frac{1}{4} U'' + k^2 W + U'^2 W^{-1} \left(\frac{1}{4} - \text{Ri} \right) \right] \Psi = 0, \quad (11.24')$$

where

$$\text{Ri}(z) = N_0^2 (U')^{-2} \quad (11.25)$$

is the *Richardson number*. Next we multiply Equation (11.24') by the complex conjugate Ψ^* , integrate the result over the layer $0 \leq z \leq h$ for the boundary conditions $\Psi = 0$ (that is, $w = 0$) at both ends of this segment, and write out the imaginary part of the equation obtained:

$$(\text{Im } c) \int_0^h \left[|\Psi'|^2 + k^2 |\Psi|^2 - \left(\frac{1}{4} - \text{Ri} \right) |U' \Psi W^{-1}|^2 \right] dz = 0. \quad (11.26)$$

If $\text{Ri}(z) \geq \frac{1}{4}$ everywhere, then this equality holds true only on condition that $\text{Im } c = 0$, that is, in the absence of unstable wave solutions. Therefore, the condition $\text{Ri} \geq \frac{1}{4}$ is sufficient for stability of a stratified flow (Miles and Howard, 1961).

So far we have considered only two-dimensional wave perturbations $u, w(x, z)$ of a plane-parallel flow $\mathbf{u}_0 = \{U(z), 0, 0\}$, since their stability is sufficient for the stability of three-dimensional wave perturbations as well (only the projection $\mathbf{u}_0 \cdot \mathbf{k} k^{-1}$ of the basic flow influences the $e^{i(k_1 x + k_2 y - \sigma t)}$ waves). With the aid of invariant (11.20) this can also be shown for perturbations depending arbitrarily on the time: if the two-dimensional perturbations of the velocity and vorticity are bounded, then the three-dimensional perturbations of the velocity and the z component of the vorticity, multiplied by U'' , will also be bounded (and the other vorticity components can increase linearly with time).

Now let us consider briefly the stability of steady axisymmetric flows, directed along the x axis and having a velocity profile $U(r)$, $r = (y^2 + z^2)^{1/2}$. Using cylindrical coordinates (x, r, ϕ) , we will consider small wave perturbations of the form $\mathbf{u}(r) \exp[ik(x - ct) + i n \phi]$. Then for the complex amplitude $f = u_r(r)$ from (11.7) we obtain the following analog of Rayleigh's Equation (11.11):

$$(U - c) \left[\frac{\partial}{\partial r} \frac{r}{n^2 + k^2 r^2} \frac{\partial r f}{\partial r} - f \right] - r \frac{\partial Q}{\partial r} f = 0;$$

$$Q = \frac{r}{n^2 + k^2 r^2} \frac{\partial U}{\partial r} \quad (11.27)$$

with boundary conditions $f \rightarrow 0$ as $r \rightarrow \infty$ (for an axisymmetric jet in unbounded space) or $f(R) = 0$ (for a round tube of radius R). Multiplication of (11.27) by $rf^*(U - c)^{-1}$, integration over r , and selection of the imaginary part then yields

$$\operatorname{Im} c \int r^2 |f|^2 |U - c|^{-2} \frac{\partial Q}{\partial r} dr = 0, \quad (11.28)$$

from which it is evident that for instability (satisfaction of the condition $\operatorname{Im} c \neq 0$) it is sufficient that the derivative $\partial Q / \partial r$ at any point change sign (Rayleigh, 1892).

For instance, in Poiseuille flow in a round tube $U(r) = U_{\max} (1 - r^2/R^2)$, and growing nonaxisymmetric wave perturbations ($n \neq 0$) are impossible (for $n = 0$ here $\partial Q / \partial r \equiv 0$ and Rayleigh's equation has no eigenvalues at all). If $\partial Q / \partial r$ changes sign at a point r_c and $U(r_c) = U_c$, then for instability it is necessary that somewhere in the flow a condition of the Fjørtoft type ($U - U_c$) $\partial Q / \partial r < 0$ be satisfied. The stability problem here apparently, as above, reduces just to a study of the discrete spectrum of the Rayleigh equation. We note that Batchelor and Gill (1962) for axisymmetric jets and Sato and Okada (1966) for axisymmetric wakes detected the possibility of instability only for $n = 1$.

12. BAROCLINIC INSTABILITY

In the preceding section we saw that in an incompressible fluid adiabatic perturbations of the velocity field can grow only at the expense of the kinetic energy of the basic flow (see energy Equation (11.9)). Such instability is called *barotropic*, because it is in general typical of barotropic fluids (since in them the barotropic potential energy appearing due to the two-dimensional compressibility is very low). In baroclinic fluids, on the other hand, so-called *baroclinic instability* is also possible: the growth of perturbations at the expense of the available potential energy of the basic state (5.14). This type of instability plays a major role, in particular, in the formation of synoptic processes in the atmosphere and in the World Ocean.

Let us take as an example quasihydrostatic currents in the atmosphere, which will be described using the fluid-dynamical equations in isobaric coordinates (9.1)–(9.4) (with the x axis pointing to the east and the y axis to the north). We consider these processes also to be quasigeostrophic in the sense of (9.11), so that the unknown functions are represented as asymptotic series in powers of the Kibel' number, having the form $u_0 + u_1 \text{Ki} + \dots$, $v_0 + v_1 \text{Ki} + \dots$, $z_0 + z_1 \text{Ki} + \dots$, and $w_1^* \text{Ki} + \dots$. Here we will also allow for a possible gradual variation of the Coriolis parameter with latitude, setting $f = f_0 + f_1 \text{Ki}$, where f_0 is quasiconstant. In the zeroth approximation from (9.2) we obtain the formulas of the geostrophic wind:

$$u_0 = -\frac{g}{f_0} \frac{\partial z_0}{\partial y}, \quad v_0 = \frac{g}{f_0} \frac{\partial z_0}{\partial x}, \quad (12.1)$$

so that $gf_0^{-1}z_0$ is the stream function of the geostrophic current; here $D_{h0} = 0$. In

the first approximation from (9.2) we get

$$\begin{aligned}\frac{d_{h0}u_0}{dt} - f_0v_1 - f_1v_0 &= -g \frac{\partial z_1}{\partial x}; \\ \frac{d_{h0}v_0}{dt} + f_0u_1 + f_1u_0 &= -g \frac{\partial z_1}{\partial y},\end{aligned}\tag{12.2}$$

where

$$\frac{d_{h0}}{dt} = \frac{\partial}{\partial t} + u_0 \frac{\partial}{\partial x} + v_0 \frac{\partial}{\partial y}$$

is the individual time derivative with respect to the geostrophic current. Thus, for the vorticity of the geostrophic current

$$\frac{\partial v_0}{\partial x} - \frac{\partial u_0}{\partial y} = \frac{g}{f_0} \Delta z_0$$

we get

$$\frac{d_{h0}}{dt} \left(\frac{g}{f_0} \Delta z_0 + f \right) + f_0 \left(\frac{\partial u_1}{\partial x} + \frac{\partial v_1}{\partial y} \right) = 0.\tag{12.3}$$

At the same time, from Equation (9.3), taken in the first approximation, divided by $\alpha^2 c_0^2$, and then differentiated with respect to p (for quasiconstant c_0^2), we obtain

$$\frac{\partial w_1^*}{\partial p} = - \frac{g}{c_0^2} \frac{d_{h0}}{dt} \frac{\partial}{\partial p} \frac{p^2}{\alpha^2} \frac{\partial z_0}{\partial p}.\tag{12.4}$$

Next, taking into account that $\partial u_1/\partial x + \partial v_1/\partial y + \partial w_1^*/\partial p = 0$, we add together (12.3) and (12.4) to get

$$\frac{d_{h0}\tilde{\Omega}}{dt} = 0; \quad \tilde{\Omega} = \frac{g}{f_0} \Delta z_0 + f + \frac{g f_0}{c_0^2} \frac{\partial}{\partial p} \frac{p^2}{\alpha^2} \frac{\partial z_0}{\partial p}.\tag{12.5}$$

The quantity $\tilde{\Omega}$ represents in the quasigeostrophic approximation the potential vorticity (4.4), and (12.5) implies that it is an invariant of the geostrophic current. Dropping the zero of z_0 everywhere in the following, let us rewrite (12.5) in a form analogous to the usual two-dimensional equation for the vorticity (11.21):

$$\begin{aligned}\frac{\partial \mathcal{L}\psi}{\partial t} + \frac{\partial(\psi, \mathcal{L}\psi + f)}{\partial(x, y)} &= 0; \\ \mathcal{L}\psi &= \Delta\psi + \frac{f_0^2}{c_0^2} \frac{\partial}{\partial p} \frac{p^2}{\alpha^2} \frac{\partial \psi}{\partial p}; \quad \psi = g f_0^{-1} z.\end{aligned}\tag{12.6}$$

The differences from (11.21) consist in the replacement of the two-dimensional Laplace operator Δ by a three-dimensional (for stable stratification, elliptic) operator $\mathcal{L}$ and in the addition of the term $(\partial f/\partial y) (\partial \psi/\partial x)$, related to the Earth's rotation and sphericity. Above we presented integral forms (9.12) and (9.13) of Equation (12.6), derived for the boundary condition (9.4), which can also be

written as

$$\frac{d_{h0}}{dt} \left(p \frac{\partial z}{\partial p} + \alpha^2 z \right) = 0 \text{ for } p = p_0. \quad (12.7)$$

Let us show that Equations (12.6) and (12.7) have a square integral invariant, the total energy

$$\begin{aligned} E = & \frac{1}{2} \int \left[(\nabla_h z)^2 + \frac{f_0^2}{\alpha^2 c_0^2} p \left(\frac{\partial z}{\partial p} \right)^2 \right] dx dy dp + \\ & + \frac{1}{2} \int \frac{f_0^2 p_0}{c_0^2} z^2 dx dy. \end{aligned} \quad (12.8)$$

Actually, if we differentiate this expression with respect to time, integrate the quantities of type $(\partial z / \partial x) (\partial / \partial x) (\partial z / \partial t)$ by parts, and apply the Gauss theorem, we get

$$\begin{aligned} \frac{\partial E}{\partial t} = & - \frac{f_0}{g} \int z \frac{\partial \tilde{\Omega}}{\partial t} dx dy dp + \int z \mathbf{n} \cdot \nabla \frac{\partial z}{\partial t} dp d\Sigma + \\ & + \int \frac{f_0^2}{\alpha^2 c_0^2} \frac{\partial}{\partial t} \left[p \frac{\partial z}{\partial p} + \alpha^2 z \right] pz dS, \end{aligned} \quad (12.9)$$

where Σ is the lateral cylindrical surface, $\mathbf{n}$ is the unit vector normal to it, and S is the lower surface $p = p_0$. In view of (12.6) the first term is here proportional to the integral of $\partial(z^2, \tilde{\Omega}) / \partial(x, y)$, which can be integrated by parts to obtain in integral of $(z^2 \tilde{\Omega}_x)_y - (z^2 \tilde{\Omega}_y)_x$, and then the integral $\int dp \oint z^2 (\tilde{\Omega}_x dx + \tilde{\Omega}_y dy)$ or, hence, $\int dp \oint z^2 d\tilde{\Omega} = 0$, since at the lateral boundary for $p = \text{const}$ we know that $z = \text{const}$ (nonpassage of flow condition). The second term in (12.9) goes to zero, due to the constancy of the circulation at the boundary $\oint \mathbf{n} \cdot \nabla z dl = \text{const}$. The third term can be shown to be zero, as in the case of the first term, but using (12.7) instead of (12.6). Therefore, $\partial E / \partial t = 0$.

Note that $z - F(p)$ for arbitrary $F(p)$ is once again a solution of Equation (12.6) for boundary condition (12.7). Therefore, in (12.8) from $\partial z / \partial p$ we can subtract an arbitrary function of p , and from $\hat{z}$ an arbitrary constant C . From this follows the constancy in time of the mean (over x, y) value $\partial \bar{z} / \partial p$ of $\partial z / \partial p$ for each p , as well as of the mean value $\bar{z}$ for $p = p_0$. If we set $F(p) = \bar{z}(p_0)$ and $C = \bar{z}(p_0)$, then we get the minimum value of the energy (12.8):

$$\begin{aligned} E = & \frac{1}{2} \int \left[(\nabla_h z)^2 + \frac{f_0^2 p^2}{\alpha^2 c_0^2} \left(\frac{\partial z}{\partial p} - \frac{\partial \bar{z}}{\partial p} \right)^2 \right] dx dy dp + \\ & + \frac{1}{2} \int \frac{f_0^2 p_0}{c_0^2} (z - \bar{z})^2 dx dy. \end{aligned} \quad (12.10)$$

The first term here corresponds to the kinetic energy, the second to the baroclinic energy, and the third to the barotropic available potential energy. The ratios of these terms are as follows: $1: (L/L_R)^2: (L/L_0)^2$, where L is the horizontal scale of the perturbations, $L_R = HN_0 f_0^{-1}$ is the Rossby deformation radius, introduced back in Section 4, and $L_0 = c_0 f_0^{-1}$ is the scale of the barotropic perturbations, introduced by Obukhov (1949).

We will begin our study of the stability of quasigeostrophic currents (12.1), (12.6), and (12.7) with the simpler barotropic current. The method for distinguishing a barotropic current, which was formulated above, in Section 9 (following Monin (1958)), consists in taking the limit $\alpha \rightarrow 0$ in integral Equation (9.12), whereby the derivatives $\partial z/\partial s$ stop depending on p . As a result, Equation (12.6) becomes

$$\frac{\partial \mathcal{L}_1 \psi}{\partial t} + \frac{\partial(\psi, \mathcal{L}_1 \psi + f)}{\partial(x, y)} = 0; \quad (12.11)$$

$$\mathcal{L}_1 \psi = \Delta \psi - \frac{f_0^2}{c_0^2} \psi; \quad \psi = g f_0^{-1} z,$$

the potential vorticity (mean value over the thickness of the atmosphere) $\tilde{\Omega} = \mathcal{L}_1 \psi + f$, and the second term (the baroclinic available potential energy) in the total energy (12.10) drops out. Now let us use the method of Arnold to analyze the stability of a steady zonal barotropic current with a potential vorticity $\tilde{\Omega}_0$, depending monotonically on the latitude, so that the stream function can be represented as $\psi_0 = \Psi(\tilde{\Omega}_0)$. Then, by analogy with (11.20), for the functional $F(\psi) = E + \int \Phi(\tilde{\Omega}) dx dy$ the first variation goes to zero for $\Phi' = \Psi$, while the second is

$$\delta^2 F = \int \left[(\nabla \delta \psi)^2 + \frac{f_0^2}{c_0^2} (\delta \psi)^2 + \Psi'(\tilde{\Omega}_0) (\delta \tilde{\Omega})^2 \right] dx dy, \quad (12.12)$$

which implies that for stability of the current (minimum F at $\psi = \psi_0$, that is, the condition $\delta^2 F > 0$ is satisfied) it is sufficient that $\Psi'(\tilde{\Omega}_0) > 0$. In the general circulation of the atmosphere the opposite condition $\Psi'(\tilde{\Omega}_0) < 0$ is usually satisfied, so that the zonal current is from west to east, that is, ψ_0 increases from north to south, while the behavior of $\tilde{\Omega}_0$ is determined mainly by the component f , increasing from south to north: ψ_0 turns out to be a decreasing function of $\tilde{\Omega}_0$. However, this behavior may become the opposite, if we pass to a frame of reference rotating about the initial frame at a constant angular velocity and *overtaking the zonal current*. Thus, the following analog of Rayleigh's theorem holds true: for stability of a zonal barotropic current it is sufficient that the potential vorticity depend monotonically on the latitude (Go, 1949).

The sufficient conditions for stability of a zonal baroclinic current are completely analogous. In this case, according to (12.7), there still exists a boundary invariant $((\partial \psi / \partial p) + (\alpha^2 / p) \psi)_p = p_0$, and stability of a zonal current with a stream

function ψ_0 yields a minimum for $\psi = \psi_0$ of the functional

$$F(\psi) = E + \int \Phi(p, \tilde{\Omega}) dx dy dp + \int_{p=p_0} \Gamma \left(\frac{\partial \psi}{\partial p} + \frac{\alpha^2}{p} \psi \right) dx dy, \quad (12.13)$$

where $\Phi(p, \tilde{\Omega})$ is an arbitrary function of two variables, while Γ is an arbitrary function of one variable. If the first variation δF goes to zero for $\psi = \psi_0$, we can assume that $\tilde{\Omega}_0$ for each p is a monotonic function of the latitude, so that we can set $\psi_0 = \Psi(p, \tilde{\Omega}_0)$, and that $((\partial \psi_0 / \partial p) + (\alpha^2/p)\gamma_0)_{p=p_0}$ is likewise a monotonic function of the latitude, so that for $p = p_0$ we can set

$$\psi_0 = X \left(\frac{\partial \psi_0}{\partial p} + \frac{\alpha^2}{p} \psi_0 \right).$$

Then it is easy to show that for satisfaction of the requirement $\delta F = 0$ we should set $\partial \Phi / \partial \tilde{\Omega} = \Psi$ and

$$\Gamma' = - \left(\frac{f_0^2}{c_0^2} \frac{p^2}{\alpha^2} \right)_{p=p_0} X$$

and also that the second variation $\delta^2 F$ then becomes

$$\begin{aligned} \delta^2 F = & \int \left[(\nabla_h \delta \psi)^2 + \frac{f_0^2}{c_0^2} \frac{p^2}{\alpha^2} \left(\frac{\partial \delta \psi}{\partial p} \right)^2 + \frac{\partial \psi_0}{\partial \tilde{\Omega}_0} (\delta \tilde{\Omega})^2 \right] dx dy dp + \\ & + \left(\frac{f_0^2}{c_0^2} \frac{p^2}{\alpha^2} \right)_{p=p_0} \int_{p=p_0} \left[\frac{\alpha^2}{p} (\delta \psi)^2 - \right. \\ & \left. - X' \left(\frac{\partial \delta \psi}{\partial p} + \frac{\alpha^2}{p} \delta \psi \right)^2 \right] dx dy. \end{aligned} \quad (12.14)$$

For stability of a zonal current, that is, $\delta^2 F > 0$, it would be sufficient for the two conditions $\partial \psi_0 / \partial \tilde{\Omega}_0 > 0$ for each p and $X' < 0$ for $p = p_0$ to be satisfied. As in (12.12), the first of these conditions can be ensured if we change to a rotating frame of reference leading the zonal current. However, the second condition reduces to, roughly speaking, an unnatural requirement that the surface density increase from the poles to the equator. Let us replace this by the condition

$$\left(\frac{\partial \psi}{\partial p} + \frac{\alpha^2}{p} \psi \right)_{p=p_0} = \text{const}$$

(which means, more or less, that the surface density is constant); then the variations of this expression will be equal to zero, and for stability of a zonal baroclinic current it is sufficient that for each p the potential vorticity decreases from north to south (Charney and Stern, 1962).

However, these sufficient conditions for baroclinic stability are apparently satisfied in the atmosphere and in the World Ocean only very rarely. In order to study the unstable perturbations, let us linearize Equations (12.6) and (12.7) relative to a steady plane-parallel current along the x axis with a velocity $-\partial\psi_0/\partial y = U(y, p)$:

$$\begin{aligned} & \left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \left(\Delta\psi + \frac{f_0^2}{c_0^2} \frac{\partial}{\partial p} \frac{p^2}{\alpha^2} \frac{\partial\psi}{\partial p} \right) + \\ & + \left(-\frac{\partial^2 U}{\partial y^2} + \frac{\partial f}{\partial y} - \frac{f_0^2}{c_0^2} \frac{\partial}{\partial p} \frac{p^2}{\alpha^2} \frac{\partial U}{\partial p} \right) \frac{\partial\psi}{\partial x} = 0; \\ & \left(\frac{\partial}{\partial t} + \hat{U} \frac{\partial}{\partial x} \right) \left(\frac{\partial\psi}{\partial p} + \frac{\alpha^2}{p} \psi \right) - \\ & - \left(\frac{\partial U}{\partial p} + \frac{\alpha^2}{p} U \right) \frac{\partial\psi}{\partial x} = 0 \quad \text{for } p = p_0. \end{aligned} \quad (12.15)$$

The energy equation for the perturbations, which is analogous to (11.9), now becomes

$$\begin{aligned} & \frac{1}{2} \frac{\partial}{\partial t} \left\{ \iint \left[\overline{(\nabla_h \psi)^2} + \frac{f_0^2 p^2}{\alpha^2 c_0^2} \overline{\left(\frac{\partial\psi}{\partial p} \right)^2} \right] dy dp + \int_{p=p_0} \frac{f_0^2 p_0}{c_0^2} \overline{z^2} dy \right\} = \\ & = \iint \left(\overline{\frac{\partial\psi}{\partial x} \frac{\partial\psi}{\partial y} \frac{\partial U}{\partial y}} + \frac{f_0^2 p^2}{\alpha^2 c_0^2} \overline{\frac{\partial\psi}{\partial x} \frac{\partial\psi}{\partial p} \frac{\partial U}{\partial p}} \right) dy dp, \end{aligned} \quad (12.16)$$

where the overbar indicates a mean value over x . The second term on the right-hand side, which is specific to baroclinic instability, is proportional to the quantity $-\overline{v'\rho'(\partial\rho_0/\partial y)}$ and has the same sign as it does. Therefore, for instance for $\partial\rho_0/\partial y > 0$, for instability we must have $\overline{v'\rho'} < 0$, so that if the perturbations increase then the heavier particles ($\rho' > 0$) must descend and move southward ($v' < 0$), while the lighter particles ($\rho' < 0$) must, conversely, rise and move northward ($v' > 0$). Considering this effect further, we will assume that U depends only on p and we will seek wave solutions ψ of Equations (12.15) in the form $\psi(p) \exp[ik_1(x - ct) + ik_2y]$. Then, setting $k_1^2 + k_2^2 = k^2$, we obtain for the complex

amplitude $\psi(p)$ the relations

$$(U - c) \left(\frac{f_0^2}{c_0^2} \frac{\partial}{\partial p} \frac{p^2}{\alpha^2} \frac{\partial \psi}{\partial p} - k^2 \psi \right) + B\psi = 0;$$

$$B = \frac{\partial f}{\partial y} - \frac{f_0^2}{c_0^2} \frac{\partial}{\partial p} \frac{p^2}{\alpha^2} \frac{\partial U}{\partial p}; \quad (12.17)$$

for $p = p_0$

$$(U - c) \left(\frac{\partial c}{\partial p} + \frac{\alpha^2}{p} \psi \right) - F\psi = 0; \quad F = \frac{\partial U}{\partial p} + \frac{\alpha^2}{p} U. \quad (12.18)$$

Equation (12.17) differs from Rayleigh's Equation (11.11) mainly just by the term $\partial f/\partial y$. But boundary condition (12.18) is now more complicated, since it contains the eigenvalue c . Replacing it by the condition $\partial \psi/\partial p + \alpha^2 \psi/p = 0$, we can demonstrate stability for $B > 0$ (that is, Rayleigh's theorem or a particular case of the Charney—Stern theorem) or, if B changes sign once, stability for $(U - K)B < 0$, where K is the value of U at the point where B changes sign (analog of Fj\o rtoft's theorem). However, condition (12.18) can cause instability only provided that for each K there does not exist more than one increasing wave solution. This can be proven for a finite-difference approximation of Equations (12.17), (12.18), for which the interval $0 \leq p \leq p_0$ is divided at points $p_1, \dots, p_{N-1}$ into N different parts of length $\delta = p_0/N$, the equations being written in difference form as

$$(U_n - c) \left(r_{n-1/2}^2 \frac{\psi_{n-1} - \psi_n}{\delta^2} - r_{n+1/2}^2 \frac{\psi_n - \psi_{n+1}}{\delta^2} - k^2 \psi_n \right) +$$

$$+ B_n \psi_n = 0; \quad (12.17')$$

$$(U_0 - c) \left(\frac{\psi_0 - \psi_1}{\delta} + s_0^2 \psi_0 \right) - F_0 \psi_0 = 0, \quad (12.18')$$

where $r_{n-1/2}^2$ is some mean value of $f_0^2 p^2/\alpha^2 c_0^2$ between the points p_n and p_{n-1} , while $s_0^2 = (\alpha^2/p)_{p=p_0}$. Then the following theorem (Dikii, 1973) will be true: *if all $B_n > 0$ or all $B_n < 0$ (Rayleigh condition) or if the sequence B_n changes sign once and there exists a constant K_0 such that $(U_n - K_0) B_n < 0$ (Fj\o rtoft condition), then Equations (12.17') with boundary conditions*

$$\psi_N = \psi_{N-1}; \quad (12.19)$$

$$\psi_1 - \psi_0 = K(c)\psi_0; \quad K(c) = a + \frac{K}{b - c}; \quad a \geq 0, \quad (12.20)$$

possess no more than one pair of nonreal complex-conjugate eigenvalues c .

For the proof let us specify a real c and, beginning with (12.19), let us find successively from Equations (12.17'): ψ_{N-2} in the form of the ratio of a first-degree polynomial in c to $U_{N-1} - c$; then ψ_{N-3} in the form of the ratio of a second-degree polynomial to $(U_{N-1} - c)(U_{N-2} - c)$; etc., up to ψ_0 , the ratio of an $(N-1)$ th-degree polynomial to $(U_{N-1} - c) \dots (U_1 - c)$. In this way the eigenvalues of Equations (12.17') with boundary condition (12.20) are obtained, as intersections of the plots of function $M(c) = (\psi_1 - \psi_0)/\psi_0$ with the hyperbola $K(c)$. In order to calculate $\partial M/\partial c$, we differentiate Equation (12.17') with respect to c and divide by $U_n - c$, to obtain

$$\begin{aligned} & -k^2 \varphi_n + r_{n-1/2}^2 \frac{\varphi_{n-1} - \varphi_n}{\delta^2} - r_{n+1/2}^2 \frac{\varphi_n - \varphi_{n+1}}{\delta^2} + \\ & + \frac{B_n \varphi_n}{U_n - c} + \frac{B_n \psi_n}{(U_n - c)^2} = 0, \end{aligned} \quad (12.21)$$

where $\varphi_n = \partial \psi_n / \partial c$. Next we multiply (12.17') by $\varphi_n (U_n - c)^{-1}$, subtract Equation (12.21) (multiplied by ψ_n) from it, and add up the result over n from 1 to $N-1$, using a rule similar to integration by parts, to obtain

$$\delta^{-2} r_{1/2}^2 \psi_0^2 \frac{\partial M}{\partial c} = \sum_{n=1}^{N-1} \frac{B_n \psi_n^2}{(U_n - c)^2}. \quad (12.22)$$

Finally, let us multiply (12.17') by $\psi_n (U_n - c)^{-1}$ and add over n :

$$\begin{aligned} & -k^2 \sum \psi_n^2 - \delta^{-1} \sum r_{n-1/2}^2 (\psi_n - \psi_{n-1})^2 - \delta^{-2} r_{1/2}^2 (\psi_1 - \psi_0) \psi_0 + \\ & + \sum \frac{B_n \psi_n^2}{U_n - c} = 0. \end{aligned} \quad (12.23)$$

For $c \rightarrow \pm \infty$ functions ψ_n have limits satisfying the following limit equation, obtained from (12.17'):

$$-k^2 \psi_n + r_{n-1/2}^2 \frac{\psi_{n-1} - \psi_n}{\delta^2} - r_{n+1/2}^2 \frac{\psi_n - \psi_{n+1}}{\delta^2} = 0, \quad (12.17')$$

and from (12.23) we see that at the limit $\psi_0 \neq 0$ and $\psi_1 \neq \psi_0$, so that $\lim M(c)$ is finite, different from zero, and negative.

For the Rayleigh condition, for example, when all $B_n > 0$, Equation (12.22) yields $\partial M/\partial c > 0$ and, since the denominator of $M(c)$ is a polynomial of degree $N-1$ without nonreal roots, the plot of $M(c)$ has $N-1$ points of discontinuity and thus consists of N increasing branches. The horizontal asymptote of the hyperbola $K(c)$ lies higher than in the case of $M(c)$. If the vertical asymptote of $K(c)$ passes through only one of the discontinuity points of $M(c)$, then $K(c) - M(c)$ is the ratio of a polynomial of degree $N-1$ to $(c - c_1) \dots (c - c_{N-1})$, where c_n are the roots ψ_0 , so that (12.17') has $N-1$ eigenvalues. Here $K(c)$ intersects all the branches of $M(c)$, except, perhaps, two or three adjacent to the common asymptote and one of the edges, so that the number of intersection points is not less than $N-3$, and the number of nonreal eigenvalues is not more than $(N$

$-1) - (N-3) = 2$. If the asymptotes do not coincide, then $K(c) - M(c)$ is the ratio of a polynomial of degree N to $(c-b)(c-c_1)\dots(c-c_{N-1})$ and there are N eigenvalues. In this case $K(c)$ may not intersect the two branches of $M(c)$, that is, there are no less than $N-2$ real roots and no more than $N - (N-2) = 2$ nonreal roots.

For the Fjørtoft condition we multiply (12.22) by $K_0 - c$ and add (11.23) to obtain

$$\begin{aligned} & \delta^{-2} r_{1/2}^2 \psi_0^2 \left[(K_0 - c) \frac{\partial M}{\partial c} - M \right] = \\ & = \Sigma \left[k^2 \psi_n^2 + \delta^{-2} r_{n-1/2}^2 (\psi_n - \psi_{n-1})^2 + \frac{B_n(K_0 - U_n)}{(U_n - c)^2} \psi_n^2 \right], \quad (12.24) \end{aligned}$$

and function $M_1(c) = (K_0 - c) M(c)$ has a positive derivative. The points of discontinuity on it are the same as on $M(c)$, and the zeros are the same plus the N th zero at $c = K_0$. Its points of intersection with $(K_0 - c) K(c)$ are calculated exactly the same as above. Thus Dikii's theorem is completely proven.

13. BIFURCATIONS

The topological singularities of a phase flow $M(t) = F^t M(0)$, describing the evolution of all the currents of a fluid in a given geometry for all possible initial data, usually depend on some parameter μ , characterizing the degree of non-equilibrium of the phase flow. In the case of a viscous fluid this parameter is the *Reynolds number* $Re = \nu^{-1} LU$ (where L and U are typical scales of length and velocity, and ν is the kinematic viscosity), that is, the typical ratio of the inertia forces to the viscosity forces, or some analog of this ratio.

The singularities of a phase flow include: (1) *nonwandering* phase points (that is, points such that any of their neighborhoods intersects some phase trajectory at least twice; these include, in particular, *stationary* points, corresponding to steady-state solutions of the fluid-dynamical equations, and *periodic* points, lying on closed trajectories, corresponding to solutions periodic in time); (2) *limit points of trajectories* $M_\omega = \lim_{t_k \rightarrow \infty} F^{t_k} M$ (provided such limits exist) and the limit sets Ω_M consisting of these (if $M \in \Omega_M$, then point M is called *Poisson-stable*); (3) *invariant sets* (filled with whole trajectories; their boundaries are also such, so that they can be considered closed; limit sets are closed invariant; a nonempty closed invariant set which does not have subsets possessing the same properties is called a *minimal set*); (4) *recurrent* points M , for which for any $\varepsilon > 0$ there exists a $T > 0$ such that the ε , neighborhood of a segment of the trajectory $\{F^t M\}$, $t \in [\tau, \tau + T]$, for any τ contains the entire trajectory (according to Birkhoff's theorem, for the recurrence of M it is necessary and sufficient that the closure of trajectory $F^t M$ be a minimal set); (5) *attractors*, that is, minimal sets Λ of nonwandering points, having neighborhoods in which all the phase trajectories present approach Λ asymptotically (attractors differing from stationary points and closed trajectories are known as *strange attractors*).

Such singularities of phase flows can be very diverse. For instance, for the simplest two-dimensional linear dynamic system $\dot{\mathbf{u}} = A\mathbf{u}$, the stationary point $\mathbf{u} = 0$, depending on the eigenvalues λ_1, λ_2 of matrix A , can be a *node* (λ_1 and λ_2 real and with the same sign), a *saddle* (λ_1 and λ_2 real with different signs), a *focus* (λ_1, λ_2 complex conjugates), or a *center* (λ_1, λ_2 imaginary conjugates); see Figure 13.1. In particular, in the neighborhood of a saddle the trajectories are hyperbolas, both asymptotes of which pass through the saddle point, one serving as the contraction axis (the so-called *unstable manifold* of the saddle, comprising the point $\mathbf{u} = 0$ and the two departing halftrajectories, the *unstable separatrices*), and the other serving as the expansion axis (the *stable manifold* comprising $\mathbf{u} = 0$ and the two incoming trajectories, the *stable separatrices*).

If the determining parameter μ changes (for example, if Re increases), the phase flow is deformed, and at certain critical values $\mu_{1cr}, \mu_{2cr}, \dots$ some of its singularities appear or disappear or undergo a qualitative alteration. Such variations of the topological singularities of a phase flow are called its *bifurcations* (or ‘catastrophes’). We will demonstrate some very simple bifurcations, by considering the evolution of small perturbations $\mathbf{u}(\mathbf{x}, t)$ of a steady flow $\mathbf{u}_0(\mathbf{x})$ of a viscous fluid, for solutions of the linearized fluid-dynamical equations:

$$\mathbf{u}(\mathbf{x}, t) = A(t)\mathbf{f}_0(\mathbf{x}); \quad A(t) = e^{\lambda t}; \quad \lambda = \gamma \pm i\omega. \quad (13.1)$$

For small Re a steady flow $\mathbf{u}_0(\mathbf{x})$ is usually a stable focus in phase space. This means that all the eigenvalues λ of the linearized equations have negative real parts $\gamma < 0$, so that any small perturbations (13.1) become damped in time. With an increase in Re the real parts γ for some eigenvalues increase, and there may exist such a critical value Re_{1cr} that for the first time one of the eigenvalues of the linearized equations $\lambda(Re)$ intersects the imaginary axis in the complex plane of λ , that is, we will have $\gamma(Re_{1cr}) = 0$. The corresponding perturbation (13.1) will neither decay nor grow in time, that is, it will be *neutral*.

It may turn out that simultaneously we will have $\omega(Re_{1cr}) = 0$, so that on the whole $\lambda(Re_{1cr}) = 0$, and hence $A(t) \equiv 1$ and $\mathbf{u}(\mathbf{x}, t) \equiv \mathbf{f}_0(\mathbf{x})$, that is, the perturbed velocity field $\mathbf{u}_0(\mathbf{x}) + \mathbf{u}(\mathbf{x}, t) = \mathbf{u}_0(\mathbf{x}) + \mathbf{f}_0(\mathbf{x})$ describes a new stationary flow; then it is said that for $Re = Re_{1cr}$ a *stability-change bifurcation* occurs. Such a bifurcation is observed, for instance, during the development of *thermal convection* in a fluid layer heated from below (where from the state of rest $\mathbf{u}_0(\mathbf{x}) \equiv 0$ stationary convection appears at first, in the form of rolls or Bénard cells), and also in a *Taylor flow*, that is, a circular Couette flow between two coaxial rotating cylinders

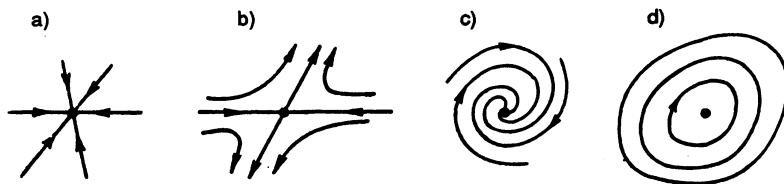


Fig. 13.1 Stationary points $\mathbf{u} = 0$ of two-dimensional linear equation $\dot{\mathbf{u}} = A\mathbf{u}$. (a) node, (b) saddle, (c) focus, (d) center.

(where from the stationary laminar flow *steady-state* toroidal Taylor vortices emerge). Such flows will be considered in detail below.

If, on the other hand, $\omega(\text{Re}_{\text{ler}}) \approx \omega_1 \neq 0$, then perturbation (13.1) for $\text{Re} = \text{Re}_{\text{ler}}$ is a neutral wave. For $\text{Re} > \text{Re}_{\text{ler}}$ there will exist eigenvalues λ with positive real parts $\gamma > 0$, that is, perturbations (13.1) exponentially increasing in time, meaning that flow $\mathbf{u}_0(\mathbf{x})$ will be unstable relative to small perturbations. Hopf (1942) showed that in phase spaces of dynamical systems of quite general form, for Re values in some neighborhood Re_{ler} there exists a single-parameter family of closed trajectories; the applicability of this theorem to the fluid-dynamical equations was demonstrated by Brushlinskaya (1965). The appearance of closed trajectories for $\text{Re} > \text{Re}_{\text{ler}}$ (then they are limit cycles, corresponding to flows that are periodic in time), that is, the transformation of a stable focus to a limit cycle (see Figure 13.2(a)) is called a *normal Hopf bifurcation*.

The transition of an unstable small perturbation (13.1) to a periodic flow was described by Landau (1944). As long as the perturbation is small, its amplitude $A(t)$ satisfies the equation $\partial|A|^2/\partial t = 2\gamma|A|^2$, but for finite $|A|$ its right-hand side must have added to it the subsequent terms of its expansion in powers of A and A^* . In this case it is advisable to eliminate the high-frequency oscillations with a frequency $|\omega_1| \gg \gamma$ by smoothing over time (over the period τ of the interval $[2\pi|\omega_1|^{-1}, \gamma^{-1}]$); then the third-degree terms drop out, and of the fourth-degree terms only the term proportional to $|A|^4$ is retained. With this accuracy we arrive at the following Landau equation and its solution:

$$\frac{\partial|A|^2}{\partial t} = 2\gamma|A|^2 - \delta|A|^4, \quad \delta > 0; \quad (13.2)$$

$$|A(t)|^2 = A_0^2 A_\infty^2 [A_\infty^2 - A_0^2] e^{-2\gamma t}^{-1}; \quad A_\infty = \left(\frac{2\gamma}{\delta} \right)^{1/2}. \quad (13.3)$$

Therefore, for small initial A_0 the amplitude of $|A(t)|$ at first increases exponentially according to a linear law, as $A_0 e^{\gamma t}$, and then this increase slows down, until as $t \rightarrow \infty$ it approaches a limit value A_∞ independent of A_0 ; for small $\text{Re} - \text{Re}_{\text{ler}}$ this limit value is proportional to $(\text{Re} - \text{Re}_{\text{ler}})^{1/2}$ (since $\gamma \sim \text{Re} - \text{Re}_{\text{ler}}$ and $\delta \neq 0$). Thus, for small $\text{Re} - \text{Re}_{\text{ler}} > 0$ perturbation (13.1) approaches (with increasing t) a periodic oscillation $\mathbf{u}_1(\mathbf{x}, t)$ with a finite amplitude A_∞ and an

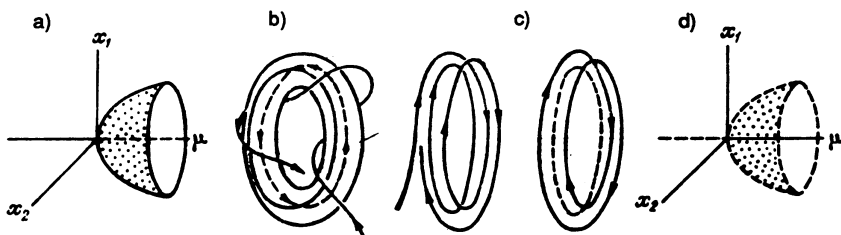


Fig. 13.2. (a) normal Hopf bifurcation; (b) production of two-dimensional invariant torus; (c) period-doubling bifurcation; (d) inverse Hopf bifurcation.

arbitrary phase (determined by the random initial phase of the perturbation and thus representing the degree of freedom of the limit flow).

Let us write the equations of fluid dynamics, linearized with respect to a periodic solution $M(t)$ of period T_1 , symbolically as $\partial M'/\partial t = f'M'$, where f' is a bounded linear operator, continuously and periodically (with a period T_1) dependent on t . Then, for any perturbation $M'(t)$ of the periodic solution, we have $M'(t + T_1) = U(T_1) M'(t)$, where $U(T_1)$ is a linear bounded operator known as the *monodromy operator*. Its eigenvalues $\rho_n(\text{Re})$ are called the *multipliers*; one of these, trivial since it is equal to unity, will be ignored in the following. If all $|\rho_n| < 1$, then all the perturbations during each circuit of a closed trajectory will decrease, so that the periodic motion is stable; if, on the other hand, $|\rho_n| > 1$ even for just one n , then the motion is unstable. Consequently, with increasing Re bifurcations of the periodic motions occur whenever the multipliers $\rho_n(\text{Re})$ passed in the complex ρ plane through the unit circle.

Therefore, as Re increases, a new critical value $\text{Re}_{2\text{cr}}$ may be reached, for which a pair of multipliers assumes the values $\exp(\pm i\alpha)$ (where $\alpha \neq 0, \pi, \pi/2, 2\pi/3$, so as to exclude resonances). Then a second normal Hopf bifurcation occurs: the periodic flow $\mathbf{u}_0(\mathbf{x}) + \mathbf{u}_1(\mathbf{x}, t)$ becomes unstable relative to some perturbation of the form $e^{\lambda t} \mathbf{f}_1(\mathbf{x}, t)$, where $\mathbf{f}_1$ is a periodic (in time) function with a period $2\pi/\omega_1$, and the eigenvalue λ has an imaginary part $\pm i\omega_2$. For small $\text{Re} - \text{Re}_{2\text{cr}}$ this perturbation will increase with time to some final value, corresponding to quasiperiodic motion with two periods ($2\pi/\omega_1$ and $2\pi/\omega_2$) and two degrees of freedom (phases of the oscillations).

Therefore, the closed trajectory evolves into a trajectory on a two-dimensional torus (see Figure 13.2(b)). Then, if the next normal Hopf bifurcation occurs, a trajectory on a three-dimensional torus is formed, etc.

If for $\text{Re} = \text{Re}_{2\text{cr}}$ one of the multipliers passes through the unit circle at the point $\rho = -1$, then $U(T_1) M'(t) = -M'(t)$, that is, a small perturbation after one circuit on the trajectory $\mathbf{u}_0(\mathbf{x}) + \mathbf{u}_1(\mathbf{x}, t)$ simply changes sign. Then after the next circuit we have $M'(t + 2T_1) = -U(T_1) M'(t) = M'(t)$, that is, the perturbed trajectory becomes closed. Consequently, in this case for $\text{Re} = \text{Re}_{2\text{cr}}$ a *period-doubling bifurcation* occurs, whereby a periodic motion with a period T_1 turns into a stable periodic motion with a doubled period $2T_1$, while the initial motion becomes unstable (see Figure 13.2(c)). Similarly, a subsequent period-doubling bifurcation can then occur, etc.

If the single-parameter family of closed trajectories, predicted by the theorem of Hopf bifurcation, still appears for $\text{Re} < \text{Re}_{1\text{cr}}$, then in the Landau expansion (13.2) coefficient δ must be negative, while coefficient $\gamma \sim \text{Re} - \text{Re}_{1\text{cr}}$ will be negative for $\text{Re} < \text{Re}_{1\text{cr}}$ and positive for $\text{Re} > \text{Re}_{1\text{cr}}$. Thus, for $\text{Re} < \text{Re}_{1\text{cr}}$ Equation (13.2) becomes

$$\frac{\partial |A|^2}{\partial t} = -2|\gamma||A|^2 + |\delta||A|^4, \quad (13.3')$$

which implies that the closed trajectory is unstable: trajectories lying inside of it wind around a stationary point (in other words, perturbations with small amplitudes $|A| < (2|\gamma|/|\delta|)^{1/2}$ damp out with time), while trajectories outside of it

unwind from it and move away into other regions of phase space (in other words, perturbations with finite amplitudes $|A| > A_1$ increase with time, so that for $\text{Re}_{1\text{cr}} > \text{Re} > \text{Re}_{1\text{cr}} - \alpha^2|A|^2$ the motion proves to be unstable relative to such perturbations). With an increase in $\text{Re} < \text{Re}_{1\text{cr}}$ the closed trajectory shrinks, and when Re goes through $\text{Re}_{1\text{cr}}$ it disappears, a phenomenon known as an *inverse Hopf bifurcation* (see Figure 13.2(d)). For $\text{Re} > \text{Re}_{1\text{cr}}$ Equation (13.2) with coefficients $\gamma > 0$ and $\delta < 0$ has a solution

$$|A|(t)^2 = A_0^2 A_1^2 [(A_0^2 + A_1^2) e^{-2\gamma t} - A_0^2]^{-1}; \quad A_1 = \left(\frac{2\gamma}{|\delta|} \right)^{1/2}, \quad (13.4)$$

which goes to infinity at some time $t_1 = (\frac{1}{2}\gamma) \ln [1 + (A_1^2/A_0^2)]$. However, it is evident that, even before this, Equation (13.2) stops being applicable and must be supplemented by the next terms of the Landau expansion. Existing examples show that, after the inverse bifurcation, for $\text{Re} > \text{Re}_{1\text{cr}}$ the motions appear to rapidly become nonperiodic.

14. INSTABILITY OF A VISCOUS FLUID

Let us examine first the stability of a stationary plane-parallel flow of an incompressible viscous fluid with a velocity $\mathbf{u}_0 = \{U(z), 0, 0\}$. Perturbations of the velocity field $\mathbf{u}(\mathbf{x}, t)$ will then satisfy Equations (11.7) with the addition on the right-hand side of the equations of motion of a term $\nu \Delta \mathbf{u}$ describing the acceleration of the viscous forces. If we call these Equations (11.7'), then their two-dimensional elementary wave solutions will be given by, instead of the Rayleigh Equation (11.11), the following *Orr—Sommerfeld equation*:

$$\begin{aligned} (U - c) \left(\frac{\partial^2 \psi}{\partial z^2} - k^2 \psi \right) - \frac{\partial^2 U}{\partial z^2} \psi \\ = - \frac{i\nu}{k} \left(\frac{\partial^4 \psi}{\partial z^4} - 2k^2 \frac{\partial^2 \psi}{\partial z^2} + k^4 \psi \right), \end{aligned} \quad (14.1)$$

which in dimensionless variables assumes the same form, provided that ν is replaced by $(\text{Re})^{-1}$. Now let us eliminate the unknowns u , v , and ρ' from the four Equations (11.7'), written for three-dimensional wave solutions, which depend on x , y , and t according to an $\exp[ik_1(x - ct) + ik_2y]$ law. Then, for w we obtain the same Equation (14.1), except that k^2 has been replaced by $k_1^2 + k_2^2$ and ν by $\nu k k_1^{-1}$ (that is, Re has been replaced by $k_1 k^{-1} \text{Re}$). Thus, if for a given value of Re there exists an unstable three-dimensional wave perturbation with an eigenvalue c , then there will also exist a two-dimensional wave perturbation with the same c that is unstable at a lower Reynolds number $k_1 k^{-1} \text{Re}$ (Squire, 1933). We note only that, when excluding the unknowns, we lost the solution of Equations (11.7') for which $w \equiv 0$, and hence $k_1 u + k_2 v \equiv 0$, with u and v satisfying the equation

$$\frac{\partial^2 u}{\partial z^2} - k^2 u - \frac{ik_1}{\nu} (U - c)u = 0 \quad (14.2)$$

with boundary conditions $u = 0$ for $z = 0, h$. For $k_2 \neq 0$ this equation possesses an additional spectrum of eigenvalues c , absent for two-dimensional perturbations, but it corresponds only to stable perturbations, that is, $\text{Im } c < 0$ (Romanov, 1973). Actually, if we multiply (14.2) by u^* , integrate over z , and take the real part of the resulting equation, then we get

$$-\int_0^h \left(\left| \frac{\partial u}{\partial z} \right|^2 + k^2 |u|^2 \right) dz - (\text{Im } c) \frac{k_1}{\nu} \int_0^h |u|^2 dz = 0, \quad (14.3)$$

from which it is evident that $\text{Im } c < 0$. Therefore, with increasing Re two-dimensional wave perturbations lose stability sooner than three-dimensional perturbations, and in order to analyze the stability relative to small wave perturbations it is sufficient to study the spectrum of eigenvalues of the Orr—Sommerfeld Equation (14.1) with the boundary conditions $\psi = \partial\psi/\partial z = 0$ for $z = 0, h$. Moreover, unlike the Rayleigh Equation (11.11), Equation (14.1) possesses no singularities and has only a discrete spectrum, so that undoubtedly any two-dimensional solution of Equations (11.7') will be a superposition of elementary wave solutions (although a complete proof of this statement has apparently not been published).

The following *principle of disappearing viscosity* has been found to be valid: the increasing (in time) and neutral wave solutions of Equations (11.7) represent the limits of the corresponding solutions of Equations (11.7') as $\nu \rightarrow 0$, but the decaying solutions of Equations (11.7') here approach the corresponding solutions of (11.7) only outside of some region of z values (the 'inner boundary layer'), within which their behavior is governed by the viscosity, however low it may be.

In order to formulate this principle in terms of the eigenvalues c of the Orr—Sommerfeld equation, let us assume that function $U(z)$ is determined not only in the segment $0 \leq z \leq h$, but also is extended analytically into some neighborhood of this segment in the complex z plane. We introduce the function $r(z) = [i(U - c)]^{1/2}$, and we call the arcs in the z plane joining points $z = 0$ and $z = h$ *permissible arcs* if along them $U - c \neq 0$ and $\text{Re} \int r dz$ varies monotonically. We will say that c is an eigenvalue along an arc if for this c there exists an analytic (along this arc) solution of the Rayleigh equation, going to zero at the ends of the arc.

Then the principle of disappearing viscosity follows from this theorem: if as $\nu \rightarrow 0$ an eigenvalue of the Orr—Sommerfeld equation approaches some limit c for which in the complex z plane it is possible to draw a permissible arc, then c is an eigenvalue of the Rayleigh equation along this arc; conversely, if c is an eigenvalue of the Rayleigh equation along some permissible arc, then this c is a limit as $\nu \rightarrow 0$ of the eigenvalues of the Orr—Sommerfeld equation (Lin, 1945—1946; Vasov, 1953).

Since the proof is quite laborious, we will just give a brief idea of it here. We introduce a small parameter $\varepsilon = (\nu/k)^{1/2}$ and functions $\psi_1 = \psi$, $\psi_2 = \partial\psi/\partial z$, $\psi_3 = \partial^2\psi/\partial z^2 - k^2\psi$ and $\psi_4 = \varepsilon \partial\psi_3/\partial z$, such that the Orr—Sommerfeld

equation will be equivalent to the following system of four first-order equations:

$$\begin{aligned}\varepsilon\psi'_1 &= \varepsilon\psi_2; \quad \varepsilon\psi'_2 = \varepsilon k^2\psi_1 + \varepsilon\psi_3; \\ \varepsilon\psi'_3 &= \psi_4; \quad \varepsilon\psi'_4 = \varepsilon^2 k^2\psi_3 + i[(U - c)\psi_3 - U'\psi_1].\end{aligned}\quad (14.4)$$

This system can be written in matrix form $\varepsilon\Psi' = P\Psi$, where $\Psi = (\psi_1, \psi_2, \psi_3, \psi_4)$, and P is a square matrix. By linear replacement of the variables $\Psi = (T_0 + \varepsilon T_1 + \varepsilon^2 T_2)\Phi$ this system can, correct to terms of higher order in ε , be put into the form:

$$\begin{aligned}\varepsilon\varphi'_1 &= \left(r - \frac{\varepsilon r'}{2r}\right)\varphi_1; \quad \varepsilon\varphi'_2 = \left(-r - \frac{\varepsilon r'}{2r}\right)\varphi_2; \\ \varphi'_3 &= \varphi_4; \quad \varphi'_4 = [k^2 + U''(U - c)^{-1}]\varphi_3.\end{aligned}\quad (14.5)$$

The last two equations here are equivalent to the Rayleigh equation. If in the complex z plane there exists an arc along which, from one end z^* to the other end z^{**} , the quantity $\text{Re} \int r dz$ increases monotonically, then we can introduce along it the two functions:

$$\chi_1 = C_1 r^{-1/2} \exp\left(\varepsilon^{-1} \int_{z_0}^z r dz\right); \quad \chi_2 = C_2 r^{-1/2} \exp\left(-\varepsilon^{-1} \int_{z_0}^z r dz\right), \quad (14.6)$$

so normalized that $\chi_1(z^{**}) = \chi_2(z^*) = 1$. If $(\varphi_3^{(1)}, \varphi_4^{(1)})$ and $(\varphi_3^{(2)}, \varphi_4^{(2)})$ are a pair of linearly independent solutions for (φ_3, φ_4) , so normalized that their Wronskian is equal to unity (according to the Liouville theorem it is constant), then it follows that system (14.5) possesses along the indicated arc the following four linearly independent solutions:

$$\begin{aligned}\Phi_1 &= (\chi_1 + \bar{\chi}_1, \bar{\chi}_1, \bar{\chi}_1, \bar{\chi}_1); \quad \Phi_2 = (\chi_2, \chi_2 + \bar{\chi}_2, \bar{\chi}_2, \bar{\chi}_2); \\ \Phi_3 &= (\varphi_1, \varphi_2, \varphi_3^{(1)} + \bar{\varphi}_3, \varphi_4^{(1)} + \bar{\varphi}_4); \\ \Phi_4 &= (\bar{\varphi}_1, \bar{\varphi}_2, \varphi_3^{(2)} + \bar{\varphi}_3, \varphi_4^{(2)} + \bar{\varphi}_4),\end{aligned}\quad (14.7)$$

where a single bar indicates multiplication by $O(\varepsilon)$ and a double bar indicates multiplication by $O(\varepsilon^2)$. Any solution of system (14.5) is $\Sigma C_j \Phi_j$. Now, expressing ψ_1 and ψ_2 in terms of it and equating these quantities to zero at $z = 0, h$, the condition for solvability of the obtained system of equations for constants C_j will be that the characteristic determinant $\Delta_4(c)$ must equal zero. But for this we obtain

$$\Delta_4(c) = \varepsilon^2(-r^{-2}\Delta_2(c) + O(\varepsilon)), \quad (14.8)$$

where $\Delta_2(c)$ is the characteristic determinant for the Rayleigh equation. The Lin—Vasov theorem follows from this.

In order to ascertain which arcs are permissible, let us first consider a real c and let us assume that z_c is the root of the equation $U(z) = c$. Then in the neighborhood of z_c we have $U - c \approx U'_c(z - z_c)$. Here $U(z)$ is a monotonically

increasing function, so that $U'_c > 0$. Now we set

$$\eta(z) = \int_{z_c}^z r \, dz \approx e^{\pi i/4} (U'_c)^{1/2} \frac{2}{3} |z - z_c|^{3/2} e^{(3i/2) \arg(z - z_c)}. \quad (14.9)$$

Three curves of $\operatorname{Re} \eta = 0$ (the so-called Stokes *lines*) diverge from the critical point $z = z_c$ at equal angles and divide its neighborhood into sectors I, II, in which the segment $0 \leq z \leq h$ lies, and III, where it does not (see Figure 14.1(a)). Any two sectors are mapped by function $\eta(z)$ onto a plane cut along a ray, the 'sides' of which are images of the Stokes lines. Figure 14.1(b) shows that the image of sectors I + II is mapped onto a plane cut along the positive part of the imaginary axis. In this plane points $\eta(0)$ and $\eta(h)$ can be joined by a curve on which $\operatorname{Re} \eta$ varies monotonically. Its inverse image in the z plane passes below the critical point. It is a permissible arc.

For $U'(z) < 0$, on the other hand, the critical point must be bypassed from above. For a $U(z)$ profile which is symmetric relative to the point $z = h/2$, there are two critical points, one which has to be passed from below (where $U'_c > 0$) and one which has to be passed from above (where $U'_c < 0$). If $\operatorname{Im} c > 0$ (instability) and if at the critical point on the real axis $U' > 0$, then z_c and the entire sector III lie above the real axis, whereas for $U' < 0$ they lie below it, the real axis also being a permissible arc in this case. If, however, $\operatorname{Im} c < 0$ (stability), then segment $0 \leq z \leq h$ intersects all three sectors, $\operatorname{Re} \eta$ varies nonmonotonically along it, and the direction the permissible arc passes the critical point (in all of sector III) is governed by the sign of U' , close to the point as well as above. The eigenfunctions of the Orr–Sommerfeld equation in sectors I and II approach the eigenfunctions of the Rayleigh equation, while in sector III for small ν they apparently oscillate rapidly; in the vicinity of the intersection of the Stokes lines with the real axis, inner boundary layers are created.

Next let us consider some specific plane-parallel flows, beginning with a plane-parallel Couette flow having a linear velocity profile $U(z) = Az$. According to all the calculations carried out, this flow is linearly stable, that is, there are no normal bifurcations in it, and $\operatorname{Re} \lambda_{\text{cr}} = \infty$ (although the complete proof of this has not been published either). At the same time, it is known from experiments that the flow is stable with respect to finite perturbations, apparently, in the region inside some 'neutral surface' in the three-dimensional space of parameters k , Re , and A . It was

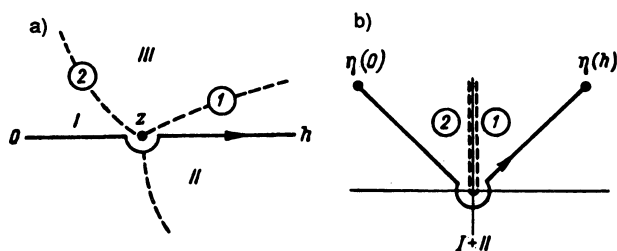


Fig. 14.1. (a) Stokes lines for real c ; (b) their image for mapping of $\eta(z)$.

calculated approximately by Kuwabara (1967) and Ellingsen *et al.* (1970). According to the first of these works instability appears only for $\text{Re} = \frac{1}{2} \nu^{-1} A h^2 > \text{Re}_{\text{cr min}} \approx 45,000$ in quite small regions of the (k, A) plane.

A plane Poiseuille flow with a parabolic velocity profile $U(z) \approx 4U_{\text{max}}(z/h)(1 - z/h)$ in an ideal fluid is linearly stable, and its linear instability in a viscous fluid at large Re (normal bifurcation), which was detected by Heisenberg (1924), at first seemed to be paradoxical (since it appeared that the viscosity can only be a stabilizing factor). However, later this instability was verified by Lin (1945) and by subsequent investigators who considered the 'neutral curve' in the (k, Re) plane, inside of which $\gamma > 0$ (where k is measured in units of $(h/2)^{-1}$, $\text{Re} = (\frac{1}{2}) \nu^{-1} h U_{\text{max}}$ and γ is defined by (13.1)); see solid curve in Figure 14.2. In this case it turned out that $\text{Re}_{\text{1cr}} \approx 5800$, $k_{\text{cr}} \approx 1$, and both branches of the neutral curve approach the abscissa axis $k = 0$ as $\text{Re} \rightarrow \infty$: the upper branch as $\text{Re}^{-1/11}$ and the lower branch as $\text{Re}^{-1/7}$. Therefore, with increasing Re a perturbation with a fixed value of k that is not too high at first is stable, then enters a region of instability, and, finally, for very large Re (including $\text{Re} = \infty$, that is, an ideal fluid) becomes stable again. We should also mention Poiseuille—Couette flow:

$$U(z) = U \left(\frac{h}{2} \right) \left[(4 - \alpha) \left(\frac{z}{h} \right) - (4 - 2\alpha) \left(\frac{z}{h} \right)^2 \right]. \quad (14.10)$$

For $\alpha = 0$ we obtain Poiseuille flow. With increasing α the value of Re_{1cr} increases rapidly in this case and already goes to infinity for $\alpha \approx 0.55$, long before Couette flow is reached. However, a plane Poiseuille flow is nevertheless not a 'suitable' example of a normal bifurcation: experimental data do not verify the above-mentioned calculated 'neutral curve'; rather they show that loss of stability occurs for $\text{Re} = 1000\text{--}2500$, considerably lower than the Re_{1cr} value of linear theory. Thus we ought to suspect the presence of inverse bifurcation and instability with respect to perturbations of finite amplitude. Actually, Reynolds and Potter (1967), as well as Pekeris and Schkoller (1967), after calculating the coefficient $\delta = \delta_1 + \delta_2 + \delta_3$ of the Landau expansion (13.2) (δ_1 describes the influx of energy from the most unstable perturbation, distorting the laminar flow;

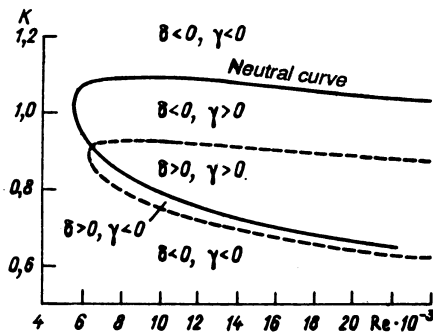


Fig. 14.2. Regions of instability of plane Poiseuille flow according to Pekeris and Schkoller (1967).

δ_2 describes the generation of higher modes by this perturbation; and δ_3 describes the distortion of its form), established that in the (k, Re) plane we have $\delta < 0$ (because of δ_3) everywhere except in the region in Figure 14.2 bounded by the dashed line. Here the $\gamma(k, \text{Re}) = 0$ and $\delta(k, \text{Re}) = 0$ curves divide a plane into four regions with different combinations of the signs of γ and δ . The region $\gamma > 0$, $\delta > 0$, is especially interesting, since in it periodic motions of finite amplitude may exist. According to Meksyn (1964), $\text{Re}_{\text{cr min}} \sim 1000$ is obtained for some three-dimensional perturbations.

But a normal bifurcation is observed perfectly reliably for a flow in a boundary layer near a flat plate, assumed in the calculations to be approximately plane-parallel. The 'neutral curve' for it was first considered by Tollmien in 1929, for a $U(z)$ profile of segments of lines and parabolas, and in 1930 for a Blasius profile (with $\text{Re}_{1\text{cr}} = 420$, where $\text{Re} = \nu^{-1} \delta^* U$, with δ^* the so-called displacement thickness and U the velocity of the oncoming flow). Such calculations were also carried out later by Schlichting and later investigators. It was verified by the experiments of Schubauer and Skramstad (1947) involving vibrator-generated perturbations $\psi(z) e^{i(kx - \omega t)}$ with a fixed (real) frequency ω , which for $\text{Im } k < 0$ increase downstream (but since $\omega = ck$, the 'neutral curves' $\text{Im } k = 0$ and $\text{Im } \omega = 0$ coincide). The 'neutral curve' is here qualitatively the same as in Figure 14.2, but subsequent calculations showed that for a velocity profile $U(z)$ with an inflection point the flow is more unstable: as $\text{Re} \rightarrow \infty$ the upper branch of the 'neutral curve' approaches the asymptote $k = k_\infty > 0$, so that perturbations with $k < k_\infty$, which had become unstable for some Re_{cr} , stop being so for all $\text{Re} > \text{Re}_{\text{cr}}$, right up to an ideal fluid (see Figure 14.6(b)).

The experiments of Klebanov *et al.* (1962) and others have demonstrated that the subsequent evolution of the instability of a Tollmien—Schlichting wave in the boundary layer around a flat plate leads for $\text{Re}_{2\text{cr}}$ to the second normal bifurcation. The latter consists in a superposition onto the primary two-dimensional wave of a secondary three-dimensional wave, periodic with respect to the transverse direction y and possessing a group velocity in x which is close to the phase velocity of the primary wave. The secondary wave grows extremely rapidly downstream and produces eddies with longitudinal axes, bringing about a sudden transverse redistribution of the fluctuation intensity of the longitudinal velocity u (see Figure 14.3), and then it becomes nonlinear. This leads to focusing of the secondary wave packet at the crest of the primary wave, and also to a disappearance in the secondary wave of sections with positive anomalies of the longitudinal velocity $u > 0$, leaving only sections with negative pulses $u < 0$ (this effect was explained by Landahl, 1972). At first one pulse appears in the cycle of oscillations of the vibrator, further downstream two pulses appear, etc., until, apparently after the regime with four pulses in the cycle, the flow becomes chaotic (random).

A flow visualization technique using chains of micron-sized bubbles, periodically generated by an electrical current in a platinum wire, and microinjections of dye (Klein *et al.*, 1963—1975), as well as a method using suspended particles (Corino and Brodkey, 1969), revealed two mutually interacting small-scale flow forms in the viscous sublayer near the wall: eddies with transverse axes and the sign of U' and streamlets of slowed fluid in the troughs of the secondary transverse waves.

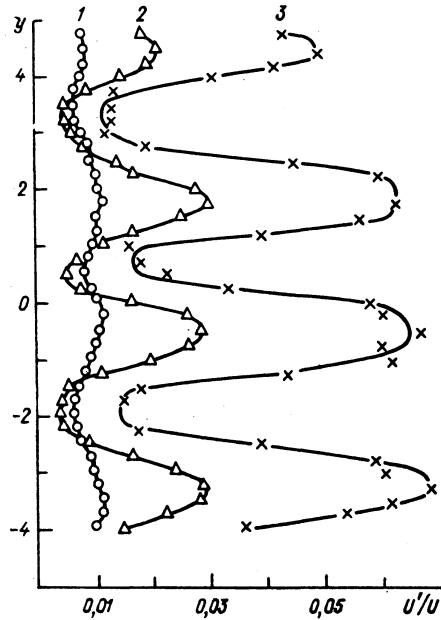


Fig. 14.3. Downstream increase of longitudinal-velocity fluctuations in secondary wave in boundary layer above flat plate. (1) $x = 7.6$ cm; (2) $x = 15.2$ cm; (3) $x = 19$ cm.



Fig. 14.4. Bubble isochrones (that is, lines of constant time; side view) in boundary layer around flat plate. Bubbles generated by vertical wire (left edge of photo, flow from left to right).

The streamlets form at heights $z^+ = \nu^{-1} u_* z = 2.5-10$ (where u_* is the 'friction velocity' at the wall), have a width $\delta y^+ = 10-30$, and are separated by distances $\Delta y^+ \sim 100$. They gradually rise under the influence of the longitudinal eddies, but because of the negative pressure gradients under the higher floating transverse eddies they separate from the wall and penetrate upward into the more rapidly

flowing fluid, producing on the instantaneous velocity profile bends with inflection points (Figure 14.4). After this, oscillations appear on a streamlet, and its end soon 'bursts', creating random motion (primarily at heights $z^+ = 10-30$ and at distances from the separation point $\delta x^+ = 1000-1500$). It can be considered that almost all the production of turbulent energy takes place at the 'bursts'.

When the streamlet separates, the central part of the transverse eddy associated with it rises and moves downstream, forming a 'horseshoe' (see Figure 14.5; the longitudinal section of the 'foot' of such a 'horseshoe' is shown in the photo of Figure 14.4). The upper part of the 'horseshoe', overtaking on the surface the next (further downstream) slowed streamlet, creates in it a negative pressure gradient and causes its separation and the formation of a new 'horseshoe'. The superposition of two 'horseshoes' leads to their unification, but most often they intersect, producing 'bursts', and thus a localized 'stochastization' and transfer of energy over the spectrum toward smaller scales.

With regard to plane-parallel flows in an unbounded space, let us mention, first of all, a two-dimensional jet, for instance, one with a velocity profile $U(z) = U_0(\cosh(z/h))^{-2}$, which is always unstable relative to antisymmetric perturbations. In this case $Re_{cr\min} = (\nu^{-1} h U_0)_{cr} \approx 3.7$ and $hk_{cr} \approx 0.25$, and with increasing Re the upper branch of the 'neutral curve' rises monotonically (see Figure 14.6(c)). The laminar 'mixing zones', smoothing the tangential velocity discontinuity, are unstable for any Re : for these the lower branch of the 'neutral curve' coincides with the entire $k = 0$ semiaxis, while the upper branch rises monotonically with increasing Re (see Figure 16.6(d)).

Of the axisymmetric flows we will confine ourselves here to a consideration just of a Poiseuille flow in a round tube. Since for such a flow the problem of the eigenvalues of the equation of the Orr—Sommerfeld type, generalizing the Rayleigh Equation (11.27) to the case of a viscous fluid, turns out to be mathematically very complicated, only a few particular cases have been considered (axisymmetric perturbations with $n = 0$ and several others). No unstable perturbations were observed, so that a Poiseuille flow in a tube, like a plane Couette flow, is apparently linearly stable, that is, $Re_{1cr} = \infty$.

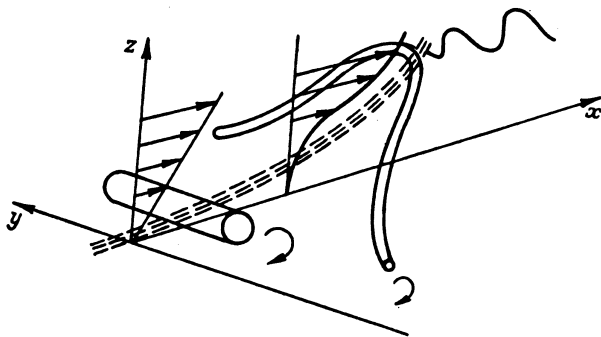


Fig. 14.5. Formation of horseshoe-shaped eddy.

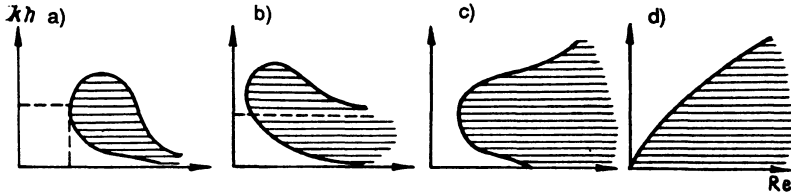


Fig. 14.6. Regions of linear instability in (kh, Re) plane: (a) for plane Poiseuille flow; (b) at boundary layer near flat plate; (c) for plane jet; (d) at plane mixing zone.

At the same time, experiments, beginning with those of Reynolds (1883) himself, showed without doubt that such a flow always remains laminar, regardless of the magnitude of the initial perturbations at the entry to the tube, only for $Re = 2\nu^{-1} RU_{\max} < Re_{cr \min} \sim 2000$. Above some Re_{cr} value the flow loses stability, evidently relative to the final perturbations, since a decrease in the initial perturbations can 'extend' the laminar regime to very high Re (according to Pfenninger (1961), to as high as 100,000). The stability losses take the form of 'turbulent slugs', intermittent and short in the x direction but filling the entire section of the tube; downstream these become longer and merge with one another.

The region of instability in the space of parameters (k, Re, A) for such a flow was considered by Davey and Nguyen (1971). We note that the presence on the tube axis of a rod, at the surface of which the velocity has to go to zero, makes the flow behave like the plane Poiseuille flow in Figure 14.2 (there will be both a region of linear instability and a region of inverse bifurcation).

To conclude this section, we present an example of the behavior of the wake behind a body around which a viscous fluid flows. Let us consider the bifurcations of the wake behind a cylinder of circular section, oriented perpendicular to the oncoming flow. For $Re \sim 10$ a *stability change* occurs, and instead of a monotonic plane flow behind the cylinder a pair of stationary vortices form. For $Re > 40$ they begin one after the other to detach from the cylinder, being replaced by new vortices, and then move downstream, forming a *Kármán vortex street*. For $Re > 100$ the vortices are replaced by rapidly agitated regions of successively detaching boundary layers. For $Re > 10^5$ the boundary layers become turbulent even before detachment, the point of detachment moves downstream, the turbulent wake narrows, and the drag decreases (*drag crisis*). For $Re \sim 10^6$ the turbulent wake broadens and the drag increases. Finally, for $Re \sim 10^7$ the wake begins to oscillate, on the whole. If the fluid has a free surface, then all these phenomena may become modified, and so-called ship waves may be superimposed onto them. In a stratified fluid all these will be accompanied by the generation of different kinds of internal waves.

15. STOCHASTICITY

In a viscous fluid a certain number of the lowest (large-scale) modes of motion completely determine all the other modes, since the viscosity severely damps the higher (small-scale) modes, which only repeat with the reduced amplitude oscilla-

tion of the normal modes (in particular, they have the same type of spectrum: discrete or continuous). In connection with this, Hopf (1948) put forth the hypothesis that as $t \rightarrow \infty$ the entire set of phase trajectories of the Navier—Stokes equations approaches a *finite-dimensional* set. For two-dimensional flows of a viscous fluid this hypothesis has been proven (Foias and Prody, 1967; Ladyzhenskaya, 1972). Note, too, that a reasonable estimate has been made of the number of degrees of freedom for developed (high Re) local three-dimensional turbulence $N \sim (\text{Re}/\text{Re}_{\text{cr}})^{9/4}$ (Landau and Lifshits, 1953). A similar estimate for two-dimensional turbulence with spectral transfer of enstrophy (that is, vorticity squared) toward smaller scales $N \sim \text{Re}/\text{Re}_{\text{cr}}$ is less accurate, since this spectral transfer is not completely local. Therefore, the fluid-dynamical equations for a viscous fluid take the form

$$\dot{\mathbf{u}} = \mathbf{F}(\mathbf{u}, \text{Re}); \quad \mathbf{u} = (u^1(t), \dots, u^N(t)), \quad (15.1)$$

where $u^k(t)$, for instance, are the coefficients of the Galerkin approximation (6.3) (see Equations (6.4) for $f^k \equiv 0$).

The viscosity causes not only finite-dimensionality of the phase space but also *dissipativeness* of the phase flow, that is, a compression (on the average) of the downstream phase volume. If we define

$$\delta V(\mathbf{u}_0, 0) = \prod_{k=1}^N \delta u_0^k$$

as a small initial element of the phase volume around point $\mathbf{u}_0$, and $\delta V(\mathbf{u}_0, t)$ as its value at time t after moving down the phase flow, then for small t we have $\delta V(\mathbf{u}_0, t) = \prod_{k=1}^N \delta u^k$, where $\delta u^k \approx (\partial u^k / \partial u_0^k) \delta u_0^k$. Since

$$\frac{\partial \delta u^k}{\partial t} \approx \frac{\partial \dot{u}^k}{\partial u_0^k} \delta u_0^k,$$

the relative rate of variation of δV with time is

$$\begin{aligned} \Lambda(\mathbf{u}_0) &\equiv (\delta V)^{-1} \frac{\partial(\delta V)}{\partial t} = \sum_{k=1}^N (\delta u^k)^{-1} \frac{\partial(\delta u^k)}{\partial t} \approx \\ &\approx \sum_{k=1}^N \frac{\partial \dot{u}^k}{\partial u_0^k} = \text{div } \mathbf{F}(\mathbf{u}_0). \end{aligned} \quad (15.2)$$

At different phase points $\mathbf{u}_0$ this quantity can be either positive (expansion) or negative (contraction). For a system of the hydrodynamic type, according to the regularity requirements (6.9), $\Lambda \equiv 0$. The phase flow will be called dissipative if for all $\mathbf{u}_0$

$$\Lambda_0(\mathbf{u}_0) = \lim_{t \rightarrow \infty} t^{-1} \ln \left| \frac{\delta V(\mathbf{u}_0, t)}{\delta V(\mathbf{u}_0, 0)} \right| < 0. \quad (15.3)$$

For a Galerkin system (6.4) $\Lambda_0 = -\nu \sum_{k=1}^n \mu_k < 0$. Because of the dissipativeness the attractors have zero phase volumes (and a dimensionality less than N).

Strange attractors are of particular interest for us here, since the phase trajectories on them exhibit properties of *stochasticity*:

- (1) very sensitive dependence on the initial conditions, due to the *exponential divergence* of initially close trajectories (which leads to nonpredictability or nonreproducibility of the trajectories when the initial conditions are specified with an arbitrarily high, but finite, accuracy);
- (2) almost all the trajectories are everywhere dense on the attractor, that is, they pass arbitrarily close to any point of it (and thus have an infinite-fold returnability to it), while any initial nonequilibrium probability distribution (measure) in phase space (more precisely, in the region of attraction of the strange attractor) approaches some limit equilibrium distribution $P(A)$ at the attractor (invariant measure);
- (3) *mixing*, which means that for any (measurable) subsets of attractors A and B the probability a long time t after emergence from A turns out at B to be proportional to the measure of B :

$$\lim_{t \rightarrow \infty} P\{(F^t A) \cap B\} = P(A)P(B), \quad (15.4)$$

where the symbol $\cap$ means intersection of the sets. A consequence of the mixing is nondependence on the initial conditions $\mathbf{u}_0$ (for almost all $\mathbf{u}_0$) of the mean value over the time $\langle \Phi[\mathbf{u}(t)] \rangle$ of any function $\Phi(\mathbf{u})$ on the strange attractor and identity of it to the mean value over the invariant measure $\Phi(\mathbf{u})$ (*ergodicity*):

$$\langle \Phi \rangle \equiv \lim_{T \rightarrow \infty} T^{-1} \int_0^T \Phi[\mathbf{u}(t)] dt = \int \Phi(\mathbf{u}) P(d\mathbf{u}) \equiv \bar{\Phi}. \quad (15.5)$$

An indication of mixing is a quite *rapid decay* (decrease to zero) for long times $\tau \rightarrow \infty$ of the *correlation functions*

$$B^j(\tau) = \langle [u^j(t) - \langle u^j \rangle][u^j(t + \tau) - \langle u^j \rangle] \rangle, \quad (15.6)$$

and thus *continuity* of their Fourier transforms with respect to τ (*spectral functions*).

It is convenient to define *turbulence* as a stochastic (in the sense of properties (1)–(3)) evolution of *vortical* flow of a (viscous) fluid. Stochastic *potential* flows of a fluid are preferably called *random wave fields*, whereas for nonhydrodynamic systems we will reserve, when necessary, the adjective *stochastic*.

Let us now focus on the first property of stochasticity. An exponential divergence of close trajectories with contraction of the phase volume is possible if, in some directions u^k in the phase space $\{\mathbf{u}\}$, expansion takes places, while in others there is contraction, that is, nonwandering phase points must be similar to two-dimensional saddles. Such points are known as *hyperbolic* points. A stationary point $\mathbf{u}_0$ is hyperbolic if the Jacobi matrix $A(\mathbf{u}_0) = \{\partial F^k / \partial u^l\}$ at this point has K

eigenvalues with a positive real part and $N - K$ eigenvalues with a negative real part, where $0 < K < N$. The sets $W_{\mathbf{u}_0}^s$ and $W_{\mathbf{u}_0}^u$ of phase points $\mathbf{u}_0$, through which the trajectories pass, approaching $\mathbf{u}_0$ as $t \rightarrow +\infty$ and as $t \rightarrow -\infty$, are called, respectively, stable and unstable manifolds of point $\mathbf{u}_0$; in the immediate neighborhood of point $\mathbf{u}_0$ these are subspaces, stretched on $N - K$ and K eigenvectors of matrix $A(\mathbf{u}_0)$, corresponding to its eigenvalues with negative and positive real parts.

The points of intersection between stable and unstable manifolds $W_{\mathbf{u}}^s$ and $W_{\mathbf{u}}^u$, differing from the point $\mathbf{u}$ itself, are called *homoclinic points* (whereas intersections of a stable manifold $W_{\mathbf{u}}^s$ of one stationary point $\mathbf{u}$ with an unstable manifold $W_{\mathbf{v}}^u$ of another stationary point $\mathbf{v}$ are called *heteroclinic points*). In the vicinity of these the phase flow may have a particularly complicated structure (see below).

A periodic trajectory is a hyperbolic set if some of the multipliers (see Section 13) lie inside the unit circle, while others lie outside of it (trivial multipliers, equal to unity, are once again ignored). Sets W^s and W^u of trajectories approaching the limit cycle as $t \rightarrow +\infty$ and as $t \rightarrow -\infty$ are called its stable and unstable manifolds. If the sum of their dimensions is $N - 1$, then their intersection is called a *transverse intersection*.

To illustrate homoclinic structures, let us consider the intersections of stable and unstable manifolds of a periodic trajectory in three-dimensional phase space. To do this, it will be convenient to use a so-called *Poincaré mapping*. The latter, in general form in an N -dimensional phase space, consists in recording successive points $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ of intersection of a trajectory (in the same direction) with some *intersecting* $(N - 1)$ -dimensional surface Σ in the phase space, which determines the mapping of $\mathbf{v}_{n+1} = \Pi(\mathbf{v}_n, \text{Re})$ of the surface Σ into itself. Since a solution $\mathbf{u}(\mathbf{u}_0, t)$ of Equation (15.1) exists for all t , this mapping is *reversible*.

In the case $N = 3$ some plane can serve as the intersecting surface Σ . A stationary point of mapping Π corresponds to a periodic trajectory on Σ , while the stable and unstable separatrices correspond, respectively, to the stable and unstable manifolds (see Figure 15.1, on which for the sake of clarity successive points of intersection of Σ with the phase trajectory are joined by solid lines). If the separatrices have a common point, then all of its images as $n \rightarrow \infty$ and inverse images as $n \rightarrow -\infty$ will also be points of intersection of the separatrices,

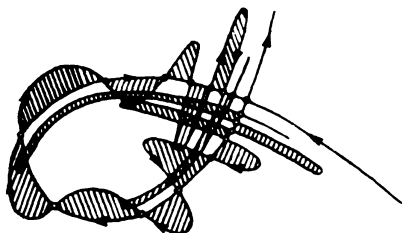


Fig. 15.1. Homoclinic trajectory on intersecting plane. Shaded sections of plane mapped onto one another.

so that they form a denumerable set. These are classified as *homoclinic* trajectories, doubly asymptotic in the sense that as $t \rightarrow \pm \infty$ they either unwind from the initial periodic trajectory or wind around it. With the approach to the stationary point, where the motion is slowed exponentially, the homoclinic points bunch together and the amplitudes of oscillation of the separatrices increase. In the vicinity of a homoclinic trajectory any small phase volume will become deformed in a complicated manner and spread out as $t \rightarrow \pm \infty$, so that almost all of the trajectories diverge exponentially. Such local instability of the trajectories, confined in a limited phase volume, also leads to a particular complexity of the phase flow, which here includes a denumerable set of periodic trajectories and a nondenumerable set of trajectories doubly asymptotic to them.

An invariant set is in general hyperbolic, if at each point of it $\mathbf{u}$ a space $T_{\mathbf{u}}$, tangent to the phase space (a linear space of tangent vectors at point $\mathbf{u}$), is the direct sum of a one-dimensional subspace $E_{\mathbf{u}}^1$, stretched onto the phase-velocity vector, and the stable and unstable subspaces $E_{\mathbf{u}}^s = \{\xi\}$ and $E_{\mathbf{u}}^u = \{\eta\}$, such that the effect on their elements of the differential of the phase flow DF' (linearly mapping $T_{\mathbf{u}}$ onto $T_{F(\mathbf{u})}$) satisfies the inequalities:

$$\begin{aligned} \|\xi\|^{-1} \|DF'\xi\| &\leq a e^{-\alpha t}; & \|\eta\|^{-1} \|DF'\eta\| &\geq b e^{\alpha t} \quad \text{for } t \geq 0; \\ \|\xi\|^{-1} \|DF'\xi\| &\geq b e^{-\alpha t}; & \|\eta\|^{-1} \|DF'\eta\| &\leq a e^{\alpha t} \quad \text{for } t \leq 0, \end{aligned} \quad (15.7)$$

where a, b, c are certain positive constants which are independent of $\mathbf{u}$. The phase points $\mathbf{v}$, for which $F^t \mathbf{v}$ converges indefinitely as $t \rightarrow \infty$ (or as $t \rightarrow -\infty$), form a stable manifold $W_{\mathbf{u}}^s$ (or an unstable manifold $W_{\mathbf{u}}^u$) of point $\mathbf{u}$; it is tangent at this point to $E_{\mathbf{u}}^s$ (or $E_{\mathbf{u}}^u$). Interconnection of these manifolds for all $\mathbf{u}$ on the given trajectory yields the corresponding manifolds W^s and W^u for the trajectory. Their intersections, as above, form homoclinic structures.

Thus, the exponential divergence of close trajectories for dissipative systems is related to the presence in their phase spaces of hyperbolic sets. The question is whether they are typical of many dynamic systems or whether they constitute an exception. If they are indeed an exception, then a small perturbation of such a system (let us say, due to the 'noise' always present in nature) would deprive it of this property. With regard to this, it is convenient to use a term introduced by Andronov and Pontryagin (1937): *structurally stable*. A system which is structurally stable (a 'rough' system) is for (15.1) defined as follows: for any $\varepsilon > 0$ there exists some $\delta > 0$ such that for any perturbed system $\dot{\mathbf{u}} = \mathbf{F}_1(\mathbf{u})$, distant from the initial system by no more than δ (according to some metric $\|\mathbf{F}_1 - \mathbf{F}\|$), there exists a mutually single-valued and mutually continuous transformation of the phase space into itself, shifting its points by no more than ε and transforming the trajectories of the undisturbed system into the trajectories of the disturbed system. Structurally stable systems form an open set in the space of all possible dynamic systems.

In one-dimensional and two-dimensional phase spaces, so-called *Morse—Smale systems* possess structural stability. For these systems the sets of nonwandering points consist only of a certain number of stationary points and closed trajectories, all these being hyperbolic, and the stable and unstable manifolds corresponding to

any such points are transverse (that is, either they do not intersect or the spaces tangent to them at each point of their intersection $\mathbf{u}$ form in sum a complete tangent space $T_{\mathbf{u}}$).

However, in the phase spaces of a large dimension Morse—Smale systems were found to lack structural stability. According to a hypothesis of Smale (1965), for structural stability it is necessary and sufficient that, for each existing phase flow of the transform F^t of the phase space, the set Ω of nonwandering points be hyperbolic, that the set of periodic points be everywhere dense in Ω (this is known as axiom A), and that each stable manifold and each unstable manifold of points from Ω be transverse. The sufficiency of these conditions is proven quite generally, while necessity is proven, for the time being, just for a more limited definition of structural stability. The stochasticity of the attractors in systems satisfying axiom A was demonstrated by Bowen and Ruelle (1975).

Consequently, for $N \geq 3$ for phase flows typically in a definite sense, there will exist an infinite set Ω of hyperbolic nonwandering points with an (everywhere dense in Ω) set of periodic trajectories (Anosov, 1967, even observed flows for which the entire phase space constitutes a hyperbolic set).

Set Ω may have a very complicated geometrical structure. For instance, Smale (1967) showed that in a wide class of typical dynamic systems each homoclinic point belongs to some invariant subset K of set Ω , this subset being a *Cantor discontinuum*, that is, a nowhere dense closed set without any isolated points.

A typical example of a Cantor set is constructed as follows: (1) by discarding the *middle third* of the segment $[0, 1]$, that is, the open (without end points) interval $(\frac{1}{3}, \frac{2}{3})$, we obtain a set K_1 of two segments $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$; (2) by discarding the middle third of each of the two segments of set K_1 , we obtain a set K_2 of four segments; . . . ; (n) by discarding the middle third of each of the 2^{n-1} segments of set K_{n-1} , we obtain a set K_n of 2^n segments; etc. The intersection of all the K_n sets will be a Cantor set K . If each number of segment $[0, 1]$ is written in a ternary system of notation as $\sum_{n=1}^{\infty} a_n 3^{-n}$, where all the a_n assume values 0, 1, and 2, then K will consist of points at which all the a_n are equal to 0 or 2. Therefore, K can be made to have a mutually single-valued relationship with the set of all the binary sequences (with the entire segment $[0, 1]$), that is, set K is nondenumerable (is a continuum).

As an example of the formation of a Cantor set in a phase space, let us consider a dynamic system which, during a fixed time T , produces an onto-mapping $\Pi(\mathbf{u})$ of the interior U of a two-dimensional torus, in such a way that $\Pi(U)$ is the interior of a torus with one loop embedded in U (see Figure 15.2). The circle S , forming a section of U , is then transformed into two small circles $\Pi(S)$ inside S . The next iteration $\Pi^2(S)$ gives two smaller circles inside each of the small circles $\Pi(S)$, and so forth. The intersection of all the iterations $\Pi^n(S)$ yields a Cantor set of points in S , so that the intersection of all the iterations $\Pi^n(U)$ is a Cantor set of lines, known as a *one-dimensional solenoid* of Williams (1970).

A Cantor set is *fractal*, that is, its *Hausdorff* dimension $\dim_H K$ exceeds the usual *topological* dimensional (which for it is equal to zero). Here $\dim_H A$ is defined in terms of the *Hausdorff* α measure of set A

$$\text{mes}_{H, \alpha}(A) = \lim_{\varepsilon \rightarrow 0} \inf_{\Gamma(A)} \sum [D(U)]^\alpha, \quad (15.8)$$

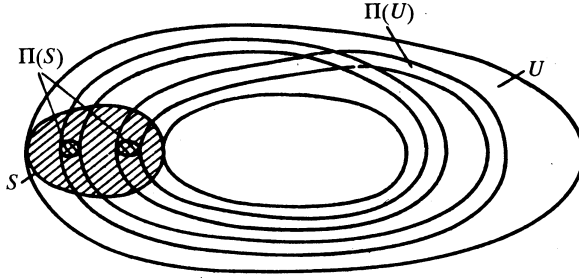


Fig. 15.2. Onto-mapping $\Pi(U)$ of interior U of two-dimensional torus, producing a strange attractor.

where the lower bound is selected according to the finite or denumerable covers Γ of set A by spheres U with diameters $D(U) \leq \varepsilon$. The dimension $\dim_H A$ is defined as a number α_0 such that the measure (15.8) for $\alpha > \alpha_0$ is zero and for $\alpha < \alpha_0$ is infinity. If in (15.8) we use only covers by identical spheres of diameter ε and select the upper or lower limit (that is, the largest or smallest limit according to the subsequences of the values of $\varepsilon \rightarrow 0$), then we obtain the upper $\overline{\text{mes}}_{c, \alpha}(A)$ or the lower $\underline{\text{mes}}_{c, \alpha}(A)$ capacity α measures, and $\overline{\dim}_c A$ or $\underline{\dim}_c A$ is, respectively, the lower bound of the values of α for which the upper or lower measure is zero (or the upper bound of the values of α for which the measure is infinity). Here $\dim_H A \leq \underline{\dim}_c A \leq \overline{\dim}_c A$. If the last two of these quantities are identical, then they define the *capacity* of set A :

$$\dim_c A = \lim_{\varepsilon \rightarrow 0} \frac{\ln N_\varepsilon(A)}{\ln(1/\varepsilon)}, \quad (15.9)$$

where $N_\varepsilon(A)$ is the minimum number of spheres of diameter ε covering set A . The sets K_N used above to construct K consist of $N_\varepsilon = 2^n$ intervals of length $\varepsilon = 3^{-n}$, so that

$$\dim_c K = \lim_{n \rightarrow \infty} \frac{(\ln 2^n)}{(\ln 3^n)} = \frac{\ln 2}{\ln 3} \approx 0.631.$$

The exponential divergence of close trajectories, as a time average, can be characterized quantitatively by the so-called *Lyapunov indices*. For these we introduce, at the points $\mathbf{u}(t)$ of a trajectory with the initial condition $\mathbf{u}(0) = \mathbf{u}_0$, a tangent vector $\mathbf{w} = \mathbf{w}(\mathbf{u}_0, t)$ with the initial condition $\mathbf{w}(\mathbf{u}_0, 0) = \mathbf{w}_0$, so that $\|\mathbf{w}(\mathbf{u}_0, t)\|$ characterizes the projection onto the direction $\mathbf{w}$ of the distance (at time t) between the trajectories with close initial points $\mathbf{u}_0$ and $\mathbf{u}_0 + \mathbf{w}_0$. Vector $\mathbf{w}$ satisfies Equation (15.1), linearized with respect to $\mathbf{u}(t)$, which in this case becomes: $\dot{\mathbf{w}} = A[\mathbf{u}(t)]\mathbf{w}$, where $A = \{\partial F^k / \partial u^l\}$ is the Jacobi matrix at point $\mathbf{u}(t)$. This equation has a complete system of fundamental solutions $\mathbf{w}_1, \dots, \mathbf{w}_N$, and for each of these a Lyapunov index is defined:

$$\sigma_i(\mathbf{u}_0) = \lim_{\substack{t \rightarrow \infty \\ \|\mathbf{w}_i(0)\| \rightarrow 0}} t^{-1} \ln \frac{\|\mathbf{w}_i(\mathbf{u}_0, t)\|}{\|\mathbf{w}_i(\mathbf{u}_0, 0)\|}. \quad (15.10)$$

These are numbered in decreasing order: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_N$. The divergence of

phase flow (15.3) is $\Lambda_0 = \sum_{i=1}^N \sigma_i$, so that for dissipative flows not only is $\sigma_N < 0$ but also the modulus of the sum of the negative σ_i is greater than the sum of the positive σ_i . If trajectory $\mathbf{u}(t)$ approaches a stationary point (or a periodic or quasiperiodic trajectory), then all σ_i (or all σ_i except $\sigma_i = 0$, when vector $\mathbf{w}_1$ is directed along the limit trajectory) are negative.

If the trajectory approaches a strange attractor Λ , then $\sigma_1 > 0$. If at the attractor there exists an invariant measure and if there is ergodicity, then the time averages in (15.10) do not depend on $\mathbf{u}_0$ and can be replaced by the averages over the invariant measure. Let k be a number such that $\sum_{i=1}^k \sigma_i \geq 0$ and $\sum_{i=1}^{k+1} \sigma_i < 0$. Then the *Lyapunov dimension* of attractor Λ is defined as

$$\dim_L \Lambda = k + \sum_{i=1}^k \frac{\sigma_i}{|\sigma_{k+1}|}. \quad (15.11)$$

Li and Yorke (1975) put forth the hypothesis that this quantity is identical to the Hausdorff dimension of set Λ (defined as the lower bound of the Hausdorff dimension of the sets of unit invariant measure, the limit for the Lebesgue measure in phase space). Numerical calculations for some two-dimensional mappings and one three-dimensional flow showed that quantities (15.11) and (15.9) are practically identical.

16. STOCHASTIZATION SCENARIOS

As yet it has not been established rigorously for any flow geometry of a viscous fluid just what sequence of bifurcations with increasing Re causes a steady (laminar) flow to turn into a stochastic (turbulent) flow. The information presented in Section 14 shows that for different flow geometries these bifurcation sequences can apparently differ considerably. Let us call the most plausible hypotheses about the bifurcation sequences leading to stochasticity *scenarios*, and let us discuss some of these.

Scenario of Landau (1944) and Hopf (1948): a sequence of normal bifurcations, producing a limit quasiperiodic flow $\mathbf{u}[\mathbf{x}, \varphi_1(t), \dots, \varphi_N(t)]$, having a period 2π for each of the variables $\varphi_k(t) = \omega_k t + \alpha_k$ with, in general, incommensurable frequencies $\omega_1, \dots, \omega_N$. This flow occupies in the phase space a region corresponding to all possible collections of initial phases $\alpha_1, \dots, \alpha_N$ and is ergodic in the sense that a trajectory winding onto it will in the course of time pass arbitrarily close to any point in the specified region (since at a time $t_k = 2\pi k/\omega_1$, $k = 0, 1, 2, \dots$, when $\varphi_1(t) = \alpha_1$, the phase of any other oscillation $\varphi_2(t_k) = 2\pi k\omega_2/\omega_1$ after reduction to the interval $[0, 2\pi]$ may have a value arbitrarily close to any of the specified value).

The time correlation functions of the velocities here do not in general go to zero at infinity, but at first they decrease rapidly, (as $N^{-1/2}$), and the time T before the next maximum (the *Poincaré recovery period*) is very long: $T \sim e^{\alpha N}$, where $\alpha \sim 1$ (Arnol'd, 1974). What is much worse is that the sequence of normal bifurcations and the resulting quasiperiodic flow are not structurally stable (Eckmann, 1981) and they are not typical even in the sense that the phase flows conforming to this

scenario do not form in the space of all the phase flows of a *Baire* set (a denumerable intersection of open, everywhere dense, sets). In connection with this, Sell (1979) proposed improving this scenario, making it a sequence of bifurcations of k -dimensional tori $T^k \rightarrow T^{k+1}$, without requiring that the flows on them be quasiperiodic. Such bifurcations are also possible in the presence of strange attractors and in a certain sense they possess structural stability.

Scenario of Ruelle and Takens (1971): appearance of a strange attractor after three normal bifurcations. This scenario possesses structural stability: in each sufficiently small neighborhood of a phase flow with an n -dimensional invariant torus for $n \geq 3$ there exists an open set of phase flows with a strange attractor, satisfying axiom A (smallness of a neighborhood being defined in the sense of the C^{n-1} norm

$$\|\mathbf{F}\| = \max_{\substack{(u) \\ 0 \leq k \leq n-1}} |D^{(k)}\mathbf{F}|,$$

where $D^{(k)}\mathbf{F}(\mathbf{u})$ denotes any derivative of the general order k , while for $n \geq 4$ also in the sense of the norm C^∞). In 1971 this theorem was proven for $n \geq 4$ (and attention was focused on processes of stochastization in dissipative systems), while for $n = 3$ it was generalized by these investigators, in collaboration with Newhouse (1978).

The proof of the theorem of Ruelle and Takens is based on approximation of a flow with a torus T^k by a flow with a closed trajectory winding around a torus (where ω_i/ω_k for $i = 1, \dots, k-1$, are small rational numbers), in which it is possible to embed a Cantor attractor, for instance, of the type of a Williams solenoid. In other words, in this mechanism a three-frequency motion is destroyed by nonlinear synchronization (formation of resonances) of its high harmonics. Such a mechanism is apparently too weak (or, to use Chirikov's term, too 'delicate') for turbulence to originate. In the experiments carried out, the stochastization was more like the break up of a two-frequency motion T^2 (that is, of *beats*), probably via their synchronization, followed by either period-doubling bifurcations of the forming loop or merging and disappearance of the stable and saddle loops (plus formation of an attractor from the homoclinic structure of the saddle loop or from the folds of the initial uneven torus).

Scenario of Feigenbaum (1978–1979): the appearance of a strange attractor as a result of an infinite sequence of period-doubling bifurcations. Let us consider such bifurcations first using as an example a one-dimensional irreversible (single-valued and continuous) mapping $x_{n+1} = \Pi(x_n, \mu)$ of segment $0 \leq x \leq 1$, where function Π possesses on this segment one quadratic extremum, a maximum, say. The iterations of such a mapping are conveniently depicted on a Koenigs–Lameras diagram (Figure 16.1). On this diagram x_{n+1} is the ordinate and x_n is the abscissa; $\Pi(x)$ and the bisector of the coordinate angle are plotted, and a *Lameras ladder* is constructed between them, in such a way that the image of x_{n+1} becomes the inverse image for x_{n+2} , $n = 0, 1, 2, \dots$

The point $x_* = \Pi(x_*)$ of intersection between $\Pi(x)$ and the bisector is a fixed mapping point. It is *stable* if it has an arbitrarily small neighborhood, the iterations

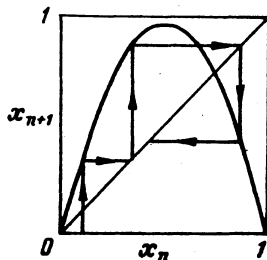


Fig. 16.1. Koenigs—Lamais diagram.

of the points of which converge to x_* ; for this the following condition is sufficient: $|\Pi'(x_*)| < 1$ (here the $\Pi(x)$ plot intersects the bisector at an angle to the abscissa axis of either less than 45° , in which case the iterations converge monotonically to x_* , or more than 135° , in which case the convergence of the iterations is oscillatory). The set of points $x_{i+1} = \Pi(x_i) \neq x_i$, $i = 1, 2, \dots, n-1$; $x_{n+1} = \Pi(x_n) = x_1$ is an n -fold loop (its Lamais ladder is closed). All of its points are fixed points of the n -fold iteration Π_n , and it is stable if at some one of these $|\Pi'_n(x)| < 1$. Retaining only the quadratic part of $\Pi(x)$, we can, without sacrificing generality, reduce the mapping to the *logistic equation*

$$x_{n+1} = \mu x_n(1 - x_n). \quad (16.1)$$

We are interested only in the range $1 < \mu \leq 4$ (since for $\mu \leq 1$ all the trajectories approach the point $x = 0$, while for $\mu > 4$ values of $\Pi(x) > 1$ appear). In this range the mapping has two fixed points: an unstable point $x = 0$ and a point $x_0 = 1 - \mu^{-1}$, the stability of which depends on the eigenvalue λ of Equation (16.1) linearized in it. It is easy to see that $\lambda = 2 - \mu$, so that the range of stability $|\lambda| < 1$ of point x_0 is $1 < \mu < 3$. For $\mu \geq \mu_1 = 3$ this point becomes unstable and, in addition, a twofold loop appears: a pair of roots x_1^0, x_1^1 of the equation $x = \Pi_2(x)$, different from x_0 . The graph of this function has two maxima, and the minimum between them is quite similar to the inverted graph of $\Pi(x)$: for $\mu_1 < \mu < \mu_2 \approx 3.45$ points x_1^0 and x_1^1 are stable, while for $\mu \geq \mu_2$ they become unstable, and in addition in their neighborhoods the twofold loops (x_2^0, x_2^1) and (x_2^2, x_2^3) of mapping Π_2 appear, forming a fourfold loop of mapping Π , which for $\mu_2 < \mu < \mu_3$ is stable, etc.

Therefore, for values μ_n , $n = 1, 2, 3 \dots$, period-doubling bifurcation occurs: the 2^{n-1} -fold loops lose stability and stable 2^n -fold loops appear. Feigenbaum discovered that the sequence μ_n converges (to a limit $\mu_\infty \approx 3.57$) asymptotically as a geometric progression with a quite high common ratio:

$$(\mu_n - \mu_{n-1})(\mu_{n+1} - \mu_n)^{-1} \sim \delta = 4.6692 \dots \quad (16.2)$$

Moreover, with the aid of functional equations he showed that this law and the value of δ are *universal*: they are valid not only for mapping (16.1) but also for any mapping $\Pi(x)$ of the above-indicated general form. Accordingly, knowing μ_0

and μ_1 , we can predict

$$\mu_\infty \approx \mu_0 + \frac{\delta}{\delta - 1} (\mu_1 - \mu_0).$$

Feigenbaum also demonstrated universality for the disposition over the segment $0 \leq x \leq 1$ of periodic points of the next $2^0, 2^1, 2^2, 2^3, \dots$ fold loops $x_0, (x_1^0, x_1^1), (x_2^0, x_2^1, x_2^2, x_2^3), \dots$. These points are stable in the intervals $3 \leq \mu \leq \mu_1, \mu_1 \leq \mu \leq \mu_2, \mu_2 \leq \mu \leq \mu_3, \dots$, and on the plot of $x(\mu)$ are represented by segments of curves, which at points $\mu_1, \mu_2, \mu_3, \dots$ split (unsymmetrically), in sum forming a 'tree' (see Figure 16.2, constructed numerically for a quadratic mapping equivalent to (16.1)). The universality of this 'tree' consists in the fact that each successive splitting is similar to the preceding one, with similarity factors for different branches $(-\alpha^{-1})$ or α^{-2} , where $\alpha = 2.5029 \dots$ (the similarity is also asymptotic, but it is already established practically after the first iterations:

$$\frac{x_n^k - x_n^{k+2^{n-1}}}{x_{n+1}^k - x_{n+1}^{k+2^n}} = \begin{cases} -\alpha, & 0 \leq k < 2^{n-1}, \\ \alpha^2, & 2^{n-1} \leq k < 2^n. \end{cases} \quad (16.3)$$

For large n we replace (by interpolation) the sequence $x_n^0, x_n^1, \dots$ by a function

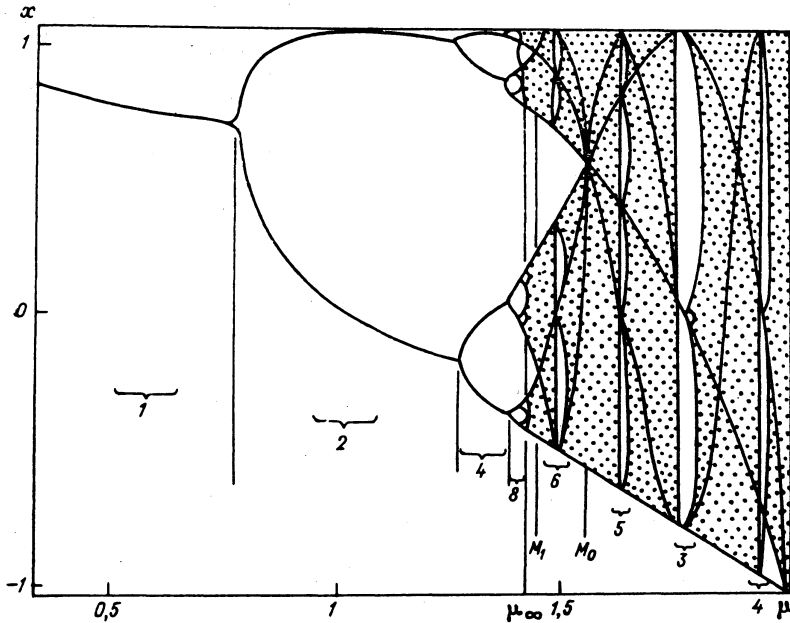


Fig. 16.2 Bifurcation diagram of mapping $x_{n+1} = 1 - \mu x_n^2$ of segment $-1 \leq x \leq 1$ for $0.35 \leq \mu \leq 2$.

with a continuous argument $x_n(t)$ and represent it as a Fourier series

$$x_n(t) = \sum_k X_n^k \exp \left[\frac{2\pi i k t}{2^n} \right], \quad (16.4)$$

in order to arrive at a universal scaling law for the amplitudes $|X_n^k|$:

$$\begin{aligned} X_{n+1}^k &= \frac{1}{2^{n+1}} \int_0^{2^{n+1}} x_{n+1}(t) \exp \left[-\frac{2\pi i k t}{2^{n+1}} \right] dt \\ &= \frac{1}{2 \cdot 2^n} \int_0^{2^n} [x_{n+1}(t) + (-1)^k x_{n+1}(t + 2^n)] \exp \left[-\frac{\pi i k t}{2^n} \right] dt. \end{aligned} \quad (16.5)$$

For even $k = 2l$ from (16.3) we get $x_{n+1}(t) \approx x_{n+1}(t + 2^n) \approx x_n(t)$, so that (16.5) becomes $X_{n+1}^{2l} \approx X_n^l$, indicating that at this frequency the Fourier amplitude for all the subsequent bifurcations does not vary. For odd $k = 2l + 1$ from (16.3)–(16.5) we obtain the recursion relation

$$\begin{aligned} X_{n+1}^{2l+1} &= -\frac{1}{2\alpha} [1 - (-1)^l i] \left[1 + (-1)^l \frac{i}{\alpha} \right] S; \\ S &= \frac{1}{\pi i} \sum_m \left[(2m + 1) - \frac{1}{2} (2l + 1) \right]^{-1} X_n^{2m+1}. \end{aligned} \quad (16.6)$$

Assuming that the amplitude modulus X_n^{2m+1} depends smoothly on m and that its phase is random, if $2l + 1$ is replaced by the continuous argument ξ , then from (16.6) we obtain the scaling law

$$\begin{aligned} |X_{n+1}(\xi)| &= \frac{1}{\gamma} \left| X_n \left(\frac{\xi}{2} \right) \right|; \\ \gamma &= 2\alpha^2(1 + \alpha^2)^{-1/2} \approx 4.65. \end{aligned} \quad (16.7)$$

For $\mu > \mu_\infty$ for certain μ there will be created (in pairs, one stable and one unstable) trajectories of periodic motion (successively with periods of 1, 6, 5, 3, ..., see Figure 16.2), each of which then undergoes a sequence of period-doubling bifurcations with its own concentration point. Moreover, here on the segment $0 \leq x \leq 1$ there will exist bands of stochastic motion, where for the values $\mu_\infty < \dots < \mu_n^* < \mu_{n-1}^* < \dots < \mu_1^*$ they undergo inverse period-doubling bifurcations, via which the number of bands is halved, while the bands themselves broaden (and merge), following a scaling law with the same constants δ and α as above. For example, after the $(n + 1)$ st bifurcation the rms band width $W_{n+1} = (\alpha^{-2}/2 + \alpha^{-4}/2)^{1/2} W_n$, which yields $W_n = W_0 \beta^{-n}$, where $\beta = \gamma/\sqrt{2} \approx 3.29$. Therefore, the variance of the stochastic part of the motion is proportional to $W^2 \sim \beta^{-2n} \sim (\mu_n^* - \mu_\infty)^m$, where $m = 2 \ln \beta / \ln \delta \approx 1.544 \dots$ (here we have

used scaling law (16.2)). Consequently, the stochasticity appears not in jumps, but rather grows gradually with increasing μ (but not monotonically, since there are interspersed regions where the periodic motions are self-organized).

A one-dimensional mapping has a single Lyapunov index (15.10):

$$\sigma(x_0, \mu) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \ln \left| \frac{\partial \Pi(x_n, \mu)}{\partial x_n} \right|. \quad (16.8)$$

It is almost everywhere independent of x_0 , is negative for $\mu < \mu_\infty$ (since a limit loop exists), and for $\mu - \mu_\infty > 0$ is predominantly positive and at first proportional to $(\mu - \mu_\infty)^k$, where $k = \ln 2 / \ln \delta \approx 0.4498 \dots$ (Huberman and Rudnik, 1980), after which it turns out to be a complicated function of μ .

The invariant measure (probability density) $p(x)$ for a mapping $\Pi(x)$ with one extremum will, because of conservation of the 'number of trajectories', satisfy the relation

$$p(x) dx = p(x_1) dx_1 + p(x_2) dx_2,$$

where x_1 and x_2 are the two inverse images of point x . This leads to the functional equation

$$p(x) = p(x_1) \left| \frac{\partial \Pi(x_1)}{\partial x_1} \right|^{-1} + p(x_2) \left| \frac{\partial \Pi(x_2)}{\partial x_2} \right|^{-1}, \quad (16.9)$$

which can be solved numerically (using the iteration method). For a triangular mapping $\Pi(x)$ with a maximum $\Pi(\frac{1}{2}) = 1$, the obvious solution is $p(x) = 1$. Mapping (16.1) for $\mu = 4$ becomes a triangular mapping $\Pi(\bar{x})$ for the variable change to $\bar{x} = 2\pi^{-1} \arcsin x^{1/2}$, so that for it we have

$$p(x) = \frac{\partial \bar{x}}{\partial x} = \pi^{-1} [x(1-x)]^{-1/2} \quad (16.10)$$

(the Lyapunov index being equal to $\sqrt{2}$). Stochastic motion with such an invariant measure is a mixing motion. For $\mu_\infty \leq \mu < 4$ distributions $p(x)$ turn out to be complicated (in particular, they have discontinuities). For $\mu = \mu_\infty$ the attractor is ergodic, but is not characterized by mixing (Eckmann, 1981), which constitutes a certain disadvantage of the scenario being considered.

Now let us consider two-dimensional mappings $\mathbf{x}_{n+1} = \mathbf{\Pi}(\mathbf{x}_n, \mu)$, where $\mathbf{x} = (x, y)$ and the Jacobian $\partial(x_{n+1}, y_{n+1}) / \partial(x_n, y_n)$ is different from zero (reversibility of mapping) and has a modulus less than unity (dissipativeness). The following mapping may serve as an example:

$$x_{n+1} = \Pi(x_n, \mu) + y_n; y_{n+1} = \beta x_n; |\beta| < 1, \quad (16.11)$$

where $\Pi(x)$ is an irreversible one-dimensional mapping. The Jacobian for (16.11) is equal to $(-\beta)$, and since for the k th iteration the Jacobian $(-\beta)^k \rightarrow 0$ ($k \rightarrow \infty$) it is evident that the limit attractor lies in the neighborhood of the curve $x = \Pi(y/\beta)$. The bifurcations of two-dimensional mappings are subject to the same universal laws of Feigenbaum, and thus we can confine ourselves here to quadratic

mappings, which, as is easily demonstrated, for a constant Jacobian can always be reduced to form (16.11) with a function $\Pi(x)$ of form (16.1).

This mapping has two fixed points: one stable point $x = y = 0$ and one unstable point $x = 1 - (1 - \beta)\mu^{-1}$, $y = \beta x$. For some $\mu = \mu_1(\beta)$ the second of these points loses its stability and, in addition, a twofold loop appears, that is, the fixed point of the second iteration $\mathbf{x}_{n+2} = \Pi_2(\mathbf{x}_n)$. In the neighborhood of its fixed point this second iteration (by renormalizing the values of $\mathbf{x}$, μ , and β) reduces to the same functional form as the initial mapping (the renormalization of β takes the form $\beta_2 = \beta^2$). Therefore, there ensues a sequence of period-doubling bifurcations with asymptotic scaling laws of Feigenbaum for the same parameters δ and α , with the same accumulation point μ_∞ (for $\beta_\infty = 0$) and with a further behavior for $\mu > \mu_\infty$ that is similar to that described above. For the following mapping equivalent to (16.11), (16.1):

$$x_{n+1} = 1 - \mu x_n^2 + y_n; \quad y_{n+1} = \beta x_n \quad (16.12)$$

Henon (1976) calculated (for $\mu = 1.4$ and $\beta = 0.3$) a total of 5×10^6 iterations, which appear in the (x, y) plane as a set of lines having, to all appearances, a Cantor structure, although the stochasticity of this attractor had not yet been rigorously demonstrated (this was done by Misyurevich (1980) with x_n^2 in (16.12) replaced by $|x_n|$). The first Lyapunov index was here determined numerically, the second was found to be $\sigma_2 = \ln \beta - \sigma_1$, and the Lyapunov dimension of the attractor (15.11) was equal to $1 + \sigma_1|\sigma_2|^{-1} \approx 1.26$.

The functional equations of Feigenbaum are also generalized to the case of $(N - 1)$ -dimensional Poincaré mappings $\mathbf{x}_{n+1} = \Pi(\mathbf{x}_n, \mu)$ for N -dimensional dissipative phase flows: if for some μ_1 there is a period-doubling bifurcation in the flow, then a subsequent increase in μ will be accompanied by an infinite sequence of such bifurcations, satisfying scaling laws with universal constants δ and α and with some concentration point μ_∞ at which stochastic motion appears (at first ergodic, but without mixing). We present as an example the three-dimensional phase flow of Rösler (1976):

$$\dot{x} = -y - z; \quad \dot{y} = x + \frac{y}{5}; \quad \dot{z} = \frac{1}{5} - \mu z + xz. \quad (16.13)$$

The results of a similar modeling of this system at a projection onto the (x, y) plane, together with the spectral densities for $z(t)$ for three subcritical and three supercritical values of μ (here $\mu_0 \approx 4.20$), are shown in Figure 16.3; the direct and inverse period-doubling bifurcations are clearly evident on the spectra. On a section of this projection along the line $y = 0$ we obtain an almost one-dimensional, approximately quadratic, mapping $x_{n+1} = \Pi(x_n)$. Actually, this 'line' has a thickness, since the attractor consists of an infinite set of sheets with a Cantor structure: topologically it is derived from a sheet that expands to twice its width transverse to the trajectories and folds up to half its width along the longitudinal axis, after which its right edge adheres to its left edge (see Figure 16.4).

Scenario of Pomeau and Manneville (1980): Appearance of intermittence in time of periodic and stochastic motion after inverse tangential bifurcation (merging and disappearance of stable and unstable fixed points of Poincaré mapping, that

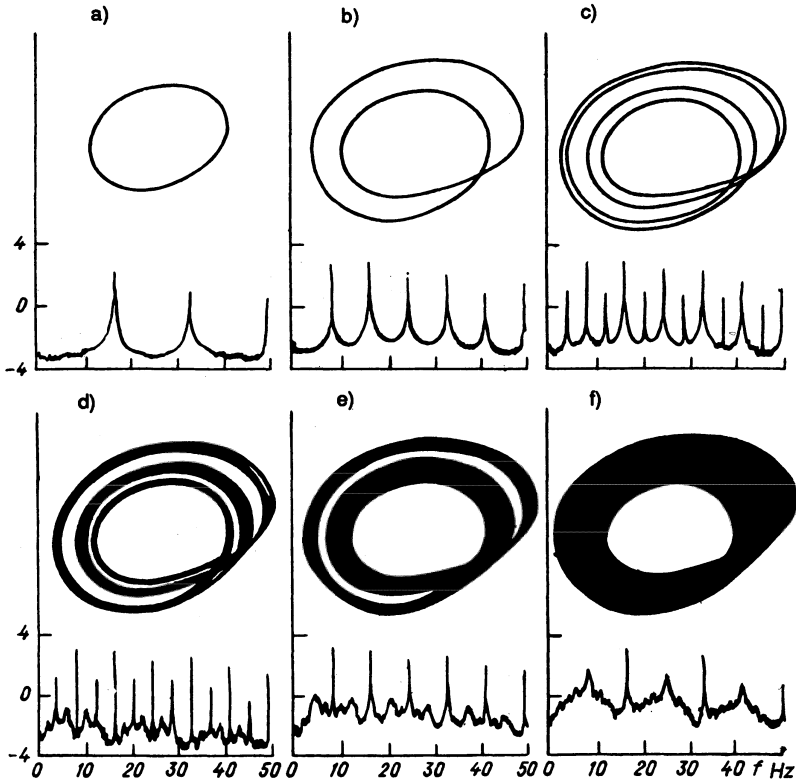


Fig. 16.3. Phase flow of Rössler in projection onto (x, y) plane and spectral densities for $z(t)$ for $\mu = 2.6$ (a), 3.5 (b), 4.1 (c), 4.23 (d), 4.30 (e), and 4.60 (f).

is, of stable and unstable periodic phase trajectories), the time intervals with stochastic motion being random and those with periodic motion being proportional to $|\mu - \mu_{cr}|^{-1/2}$. A good example of this is a bifurcation of merging of stable and unstable trajectories with a period of 3 for mapping (16.1) for the value $\mu_c \approx 3.83$, serving as an accumulation point for the sequence of *inverse* bifurcations of doubling of this period.

Rabinovich and Trubetskov (1984) showed that it is possible to explain the intermittence in terms of a special kind of Poincaré mapping $x_{n+1} = \Pi(x_n, \mu)$. Let the plot of $\Pi(x)$ consist of three curved parts: (1) steeply ascending, (2) section with minimum, (3) steeply descending. For $\mu < \mu_c$ let part 2 intersect the bisector at two points; these correspond to the stable and unstable periodic trajectories of the phase flow being mapped. We assume that with increasing μ this part rises upward, and that for $\mu = \mu_c$ the points of its intersection with the bisector merge. When $\mu > \mu_c$ very slightly, then between part 2 and the bisector the gap is so narrow that the Lamerai's ladder inscribed in it has many steps: during the course of many iterations (for a long time) the flow trajectory remains close to the initial

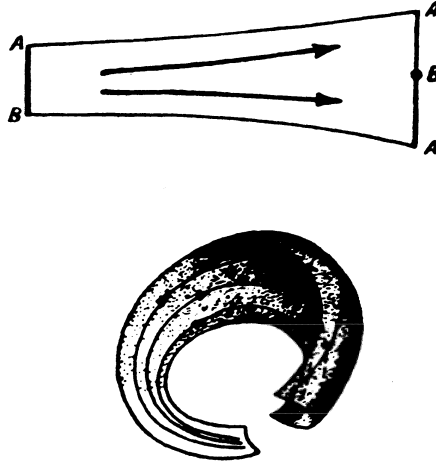


Fig. 16.4. Topology of Rösler's attractor.

periodic trajectory. At the end of part 2 the ladder starts to have large steps (sharp fluctuations of the trajectory), but with the returning approaches to part 2, prolonged intervals of close to periodic motion appear.

17. TAYLOR AND LORENZ FLOWS

Some very detailed illustrations of the above scenarios were obtained during special measurements of the loss of stability of a circular Couette flow (in the space between two coaxial rotating cylinders) and of a fluid layer heated from below (development of thermal convection).

A circular Couette flow in cylindrical coordinates r , φ , z , with the z axis directed along the axis of the cylinders, had velocity components:

$$u_{0r} = u_{0z} = 0; \quad u_{0\varphi} = U(r) = Ar + Br^{-1}; \quad (17.1)$$

$$A = (\Omega_2 R_2^2 - \Omega_1 R_1^2)(R_2^2 - R_1^2)^{-1}; \quad B = -R_1^2 R_2^2 (\Omega_2 - \Omega_1)(R_2^2 - R_1^2)^{-1},$$

where $R_1 < R_2$ are the radii, and Ω_1 and Ω_2 are the angular velocities of rotation of the inner and outer cylinders. In an *ideal fluid* the law of conservation of angular momentum of a fluid particle $mrU = \text{const}$ dictates that, for particle displacement from radius r_0 to $r > r_0$, its velocity becomes equal to $r_0 U(r_0)/r$, and if the centrifugal acceleration $r_0^2 U^2(r_0)/r^3$ acting on it is greater than its equilibrium value $U^2(r)/r$, that is, on condition that $\partial(rU)^2/\partial r < 0$, the flow will be unstable (Rayleigh, 1916). This criterion can be reduced to the condition $(\Omega_2^2 R_2^2 - \Omega_1^2 R_1^2)U < 0$, so that for rotation of the cylinders in different directions the flow is always unstable (since somewhere in the space between them U changes sign), whereas if they rotate in the same direction the instability criterion becomes $(\Omega_2/\Omega_1) < (R_1/R_2)^2$ (see Figure 17.1).

In a viscous fluid the complex amplitudes of small-wave perturbations of the

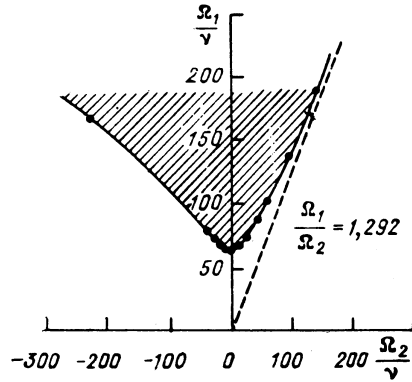


Fig. 17.1 Region of instability of circular Couette flow in the (Ω_1, Ω_2) plane for $R_2/R_1 = 1.13$.

velocity $\mathbf{u}(r) \exp[i(kz + n\varphi - \omega t)]$ with arbitrary k and integral n are given by the equations:

$$\left[\frac{\partial}{\partial r} \mathcal{L} \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) - k^2 \left(\mathcal{L} - \frac{1}{r^2} \right) \right] u_r = \left[\frac{2k^2}{vr} \left(U - \frac{inv}{r} \right) - \right. \\ \left. - in \frac{\partial}{\partial r} \mathcal{L} \frac{1}{r} \right] u_\varphi; \\ \left[\frac{n^2}{r} \mathcal{L} \frac{1}{r} + k^2 \left(\mathcal{L} - \frac{1}{r^2} \right) \right] u_\varphi = \left\{ \frac{2k^2}{vr} \left[\frac{r}{2} \left(\frac{\partial U}{\partial r} + \frac{U}{r} \right) \right] - \right. \\ \left. - \frac{inv}{r} \right\} + \frac{in}{r} \mathcal{L} \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \Bigg\} u_r; \\ \mathcal{L} = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} - \frac{n^2}{r^2} - k^2 + \frac{i\omega}{\nu} - \frac{inU}{vr}$$

with boundary conditions $u_r = (\partial u_r / \partial r) = u_\varphi = 0$ for $r = R_1, R_2$. Since these equations do not have singularities, it is obvious that for fixed parameters (including k and n) there will always exist a discrete spectrum of eigenvalues ω_j and the corresponding collection of eigenfunctions will be complete. For a long time it was assumed that with increasing Re the axisymmetric perturbations (with $n = 0$) would be the first to lose stability. For these it was established (although even now not yet proven with sufficient completeness) that the first bifurcation constitutes a *stability change*: a transition from flow (17.1) to a new *steady* flow of form $u(r) \exp(ik_c z)$, namely to toroidal roll eddies (known as *Taylor eddies*, since they were first calculated by Taylor, 1923, for small d/R_0 , where $d = R_2 - R_1$ and $R_0 = \frac{1}{2}(R_1 + R_2)$). According to these, instability regions like the one shown in Figure 17.1 were also constructed (indicating that here viscosity plays a stabilizing

role, whereas for $\Omega_2/\Omega_1 < (R_1/R_2)^2$ and sufficiently large Re instability nevertheless develops). For a long time it was thought that these showed an excellent fit with the experimental data (see points on boundary of shaded region in Figure 17.1).

However, later it was discovered that for not too small *negative* Ω_2/Ω_1 perturbations with $n \neq 0$ lose stability a little earlier than axisymmetric perturbations. For instance, the calculations of Kruger *et al.* (1966) for $d/R_1 = 1/20$ on the radial lines $\Omega_2/\Omega_1 = -0.7, -0.8, -0.9, -1.0$, and -1.25 showed that with increasing Re the first to lose stability are perturbations with $n = 0, 1, 3, 4$, and 5 . This was verified by experimental data, in particular by that of Snyder (1968).

Both theoretical calculations (Stewart, 1958; Davey, 1962; Davy *et al.*, 1968) and experimental data (Donnelly and Schwartz, 1963, 1965; Snyder and Lambert, 1966) indicate that for $\text{Re} = \nu^{-1}\Omega_1 R_1 d > \text{Re}_{1cr}$ the increase in the intensity of axisymmetric Taylor eddies is described excellently by Landau's Equation (13.2).

Experiments were specially devised in order to check the scenarios, by Gollub and Swinney (1975), Fenstermacher *et al.* (1978), and L'vov and Predtechenskii (1979, 1980). These revealed that at Re_{2cr} the Taylor eddies lose stability, and azimuthal flexural waves appear at them. With increasing Re, at subsequent bifurcations (apparently normal Hopf bifurcations) one to three more independent frequencies appear (while, in addition, on the spectrum of the velocity fluctuations $u_r(t)$ in the space between the cylinders the discrete peaks broaden and the background continuum increases, which is not foreseen by the Landau–Hopf and Ruelle–Takens scenarios). After the next bifurcation at $\text{Re} = \text{Re}_c$, only stochastic motion with a continuous spectrum remains.

The quantitative characteristics of the bifurcations observed in these experiments vary, in particular, as a function of the sizes of the cylinders and the initial data. For instance, in the first of these $R_1 = 22.54$ mm, $R_2 = 25.40$ mm, $d = R_2 - R_1 = 2.86$ mm, $d/R_1 = 0.14$, the cylinder height $h = 20d$, $\Omega_2 = 0$, and $R_c = 2501$. Taylor eddies appeared at $R^* = \text{Re}/\text{Re}_c = 0.051$ (in an amount $kh/2\pi = 17$); for $R^* = 0.064$ flexural waves appeared (four at the periphery, with a dimensionless frequency $f_1^* = 2\pi f_1/\Omega_1 = 1.30$, while in the $u_r(t)$ spectrum six of its harmonics were seen); for $R^* = 0.54$ a second (small) frequency f_2 appeared; with increasing R^* it dropped to zero at $R^* = 0.78$, where a third frequency $f_3^* \approx \frac{2}{3}f_1^*$ appeared; at $R^* \sim 1$ stochastization occurred reversibly and without hysteresis.

In the experiments of L'vov and Predtechenskii, $R_1 = 17.5$ mm, $R_2 = 27.5$ mm, $d = 10$ mm, $d/R_1 = 0.57$, $h = 30d$, and $\Omega_2 = 0$. Taylor eddies appeared at $\text{Re} \approx 74$ (in an amount ranging from 22 to 36 and, for gradual acceleration, from 28 to 30). For 30 eddies flexural waves appeared in the interval $\text{Re} = 995\text{--}1015$ (six at the periphery, $f_1 = 1.93\Omega_1$, two harmonics); then at $\text{Re} = 1040$ in the $u_\varphi(t)$ spectrum there appeared, and alternated with the first frequency, a second frequency $f_2 = 0.55\Omega_1$, as well as combination harmonics; next a third frequency $f_3 = 0.95\Omega_1$ appeared; at $\text{Re} = 1901$ there were already no well-defined maxima in the spectrum (but individual sharp peaks did appear and disappear with a further increase in Re). A flow with 28 eddies was found to behave quite differently (in particular, at $\text{Re} = 1100\text{--}1200$ it passed over into a state with 29 eddies).

On the whole, the evolution of the Taylor eddies was found to conform

generally to the scenario of Ruelle and Takens, but in its details it turned out to be considerably more complicated.

Now let us discuss thermal convection in a horizontal layer $0 \leq z \leq h$, described by the Boussinesq equations:

$$\begin{aligned} \operatorname{div} \mathbf{u} &= 0; & \frac{d\mathbf{u}}{dt} &= -\frac{\nabla p'}{\rho_0} - g\alpha T' + \nu \Delta \mathbf{u}', \\ \frac{dT'}{dt} &= \chi \Delta T', \end{aligned} \quad (17.3)$$

(where α is the coefficient of thermal expansion) with fixed boundary values of the temperature $T'_{z=0} = T_0$ and $T'_{z=h} = T_1$. The steady-state solution of these equations has the form $\mathbf{u} = 0$, $T' \equiv T'_s = T_0 - (T_0 - T_1)z/h$. By introducing the Rayleigh number $Ra = g\alpha(T_0 - T_1)h^3/\nu\chi$, the Prandtl number $\sigma = \nu/\chi$, the length scale h , and the time scale h^2/χ , and setting $T' - T'_s = (T_0 - T_1)\vartheta/Ra$, we can put (17.3) in dimensionless form:

$$\begin{aligned} \operatorname{div} \mathbf{u} &= 0; & \frac{d\mathbf{u}}{dt} &= -\nabla \Pi + \sigma \vartheta \nabla r + \delta \Delta \mathbf{u}; \\ \frac{d\vartheta}{dt} &= \omega Ra + \Delta \vartheta, \end{aligned} \quad (17.4)$$

where Π is the dimensionless deviation of the pressure from its steady-state hydrostatic distribution. Now, by linearizing these equations (that is, replacing d/dt by $\partial/\partial t$) and seeking the unknown functions in the form $f(z)\varphi(x, y)\exp(-i\omega t)$, where $\Delta\varphi + k^2\varphi = 0$, we arrive at the following relation for the complex amplitude $\vartheta(z)$:

$$\begin{aligned} &\left(\frac{\partial^2}{\partial z^2} - k^2\right)\left(\frac{\partial^2}{\partial z^2} - k^2 + i\omega\right)\left(\frac{\partial^2}{\partial z^2} - k^2 + i\sigma\omega\right)\vartheta + \\ &+ k^2 Ra \vartheta = 0 \end{aligned} \quad (17.5)$$

with boundary conditions $\vartheta = \vartheta^{\text{II}} = \vartheta^{\text{III}} - (k^2 - i\sigma\omega)\vartheta^{\text{I}} = 0$ at one wall and $\vartheta = \vartheta^{\text{II}} = \vartheta^{\text{IV}} = 0$ at the other wall. The first bifurcation here is a *stability change* (Pellieu and Southwell, 1940; Rayleigh, 1916, established this for a layer with two free surfaces, where it was found that $Ra_{\text{Icr}} = 27\pi^4/4 \approx 657.5$ and $k_{\text{cr}} = \pi/\sqrt{2} \approx 2.2$; later for a layer with rigid surfaces the values $Ra_{\text{Icr}} \approx 1708$ and $k_{\text{cr}} \approx 3.12$ were obtained, while for a rigid lower surface and a free upper surface $Ra_{\text{Icr}} \approx 1100$). In this case the stability change involves the appearance of a *stationary motion* periodic in x and y . Its form $\varphi(x, y)$ (convective rolls, square or hexagonal cells, etc.) is not determined from the linear theory.

Schluter *et al.* (1965) established that for very small $Ra - Ra_{\text{Icr}} > 0$ rolls are the only stationary convective motions which are linearly stable (in a narrow band of the (k, Ra) plane). Palm (1960) and later investigators arrived at more general results, having ascertained that the formation of hexagonal convective *Bénard cells* (Bénard, 1900) is determined by the temperature dependence of the material

properties of the fluid, primarily the viscosity: $\nu' = |\partial \nu / \partial T|$. For a perturbation of the form

$$w = \left[A_1(t) \cos ky + A_2(t) \cos \left(kx \frac{\sqrt{3}}{2} \right) \cos \left(\frac{ky}{2} \right) \right] \sin lz \quad (17.6)$$

Palm derived a system of equations of the Landau type:

$$\begin{aligned} \dot{A}_1 &= \gamma A_1 - \frac{\nu'}{4} A_2^2 - \delta_1 A_1^3 - \left(2\delta_2 - \frac{\delta_1}{2} \right) A_1 A_2^2; \\ \dot{A}_2 &= \gamma A_2 - \nu' A_1 A_2 - \delta_2 A_2^3 - (4\delta_2 - \delta_1) A_1^2 A_2, \end{aligned} \quad (17.7)$$

having steady-state solutions $A_2 = \pm 2A_1$, corresponding to hexagonal cells. It was established that for $Ra_{1cr} < Ra_1 < Ra < Ra_2$ only hexagons are stable, for $Ra_2 < Ra < Ra_3$ both hexagons and rolls are stable, and for $Ra > Ra_3$ only rolls are stable. As $\nu' \rightarrow 0$, Ra_1 , Ra_2 , and Ra_3 all approach Ra_{1cr} , bringing us back to the result obtained by Schluter, Lorz, and Busse.

Let us consider convective rolls extended along the y axis, so that for them Equations (17.4) are two-dimensional and, after introducing the stream function ψ in the (x, z) plane, take the form

$$\begin{aligned} \frac{\partial \Delta \psi}{\partial t} + \frac{\partial(\psi, \Delta \psi)}{\partial(x, z)} &= \sigma \frac{\partial \vartheta}{\partial x} + \sigma \Delta^2 \psi; \\ \frac{\partial \vartheta}{\partial t} + \frac{\partial(\psi, \vartheta)}{\partial(x, z)} &= Ra \frac{\partial \psi}{\partial x} + \Delta \vartheta. \end{aligned} \quad (17.8)$$

With increasing $Ra > Ra_{1cr}$ the motion in the rolls stops being steady and increases with time. At first it can be described in terms of the following three modes:

$$\begin{aligned} \psi &= \pi^{-1} \left(\frac{R_0}{k} \right)^{1/3} X \sqrt{2} \sin kx \sin \pi z; \\ \vartheta &= R_0 (Y \sqrt{2} \cos kx \sin \pi z - Z \sin 2\pi z), \end{aligned} \quad (17.9)$$

where $R_0 = k^{-2}(\pi^2 + k^2)^3$. Busse (1972) showed that for infinitesimal perturbations the amplitudes of other modes are small quantities of higher order. Taking this into account, substituting (17.9) into (17.8), and neglecting interactions with all the other modes, we obtain for X, Y, Z the system of equations:

$$\begin{aligned} \dot{X} &= -\sigma X - \sigma Y; & \dot{Y} &= -Y + rX - XZ; \\ \dot{Z} &= -bZ + XY, \end{aligned} \quad (17.10)$$

where the dot over a quantity signifies the derivative with respect to $(k^2 R_0)^{1/3} t$, where $b = 4\pi^2(k^2 R_0)^{-1/3}$ is a geometrical factor (for a Rayleigh factor $k = \pi/\sqrt{2}$ it is equal to $\frac{8}{3}$), and where $r = R_0^{-1} Ra$ is the relative Rayleigh number. This system,

first obtained by Lorenz (1963), became well known as the first system with a strange attractor (an attractor of infinitely many sheets with a Cantor structure was already suspected by Lorenz himself, but a rigorous mathematical proof was not yet forthcoming).

It should be stressed that Lorenz's equations (to which we will return later) describe real convection only for small r . The experiments of Willis and Dierdorff (1970) established that for $Ra_{2cr} \approx 3Ra_{1cr}$ transverse waves appear on convective rolls, just as above on toroidal Taylor rollers. These were calculated within the framework of the linear theory by Busse (1972), and according to the nonlinear theory by MacLaughlin and Martin (1975). They computed, first of all, an eight-mode motion with rolls (17.9) and one harmonic along the y axis, and then they constructed for it a Landau expansion (13.2) (this gave $\delta > 0$, which corresponds to a normal bifurcation). Secondly, they calculated numerically a 39-mode motion with rolls (17.9) and four harmonics along the y axis. For $k_1/\pi = 0.072$, $k_2/\pi = 0.1$, and $\sigma = 1$ they obtained $r_{2cr} \approx 1.25$; for $r = 1.4$ this gave a periodic regime, for $r = 1.45$ a slightly nonperiodic regime, for $r = 1.5$ and 1.55 again a periodic regime, and for $r = 1.6$ a sharply nonperiodic regime (and, excluding the fourth harmonic, for $r = 1.6, 2$, and even 20 , periodic regimes).

Now let us consider the Lorenz Equations (17.10). The divergence of the phase flow (15.2) in this case is negative: $\Lambda = -(\sigma + b + 1)$, so that all the trajectories approach some set of zero volume. The quantity

$$W = [X^2 + Y^2 + (Z - r - \sigma)^2]^{1/2}$$

satisfies the condition $\dot{W} \leq -C_1 W + C_2$ with positive C_1 and C_2 , so that all the trajectories are located in the sphere $W \leq 2C_2/C_1$. The system does not change if we make the change $(X, Y, Z) \rightarrow (-X, -Y, Z)$. For $r < 1$ the only fixed point will be the stable node 0 at the coordinate origin. For $r \geq 1$ (start of convection) it loses stability (becomes a saddle with a two-dimensional stable manifold and two unstable one-dimensional manifolds: the separatrices Γ^+ , Γ^-), and two additional fixed points appear:

$$C^+, C^- = (\pm \sqrt{b(r-1)}, \pm \sqrt{b(r-1)}, r-1),$$

toward which the separatrices tend. For $\sigma < b + 1$ they are stable; for $\sigma > b + 1$ (later, following Lorenz, we will examine the case $\sigma = 10$, $b = \frac{8}{3}$) they are stable if $1 < r < r_3 = \sigma(\sigma + b + 3)(\sigma - b - 1)^{-1} \approx 24.74$ but they lose their stability when $r > r_3$. According to the linear theory, for $r = r_3$ neutral perturbations are possible here, with a frequency $\omega = [2b\sigma(\sigma + 1)(\sigma - b - 1)^{-1}]^{1/2}$, say, $\delta X = A \cos \omega t$. If r is slightly less than r_3 , then we must add a small nonlinear correction of the order of $|A|^2$ to δX , for which MacLaughlin and Martin (1975) constructed (with an accuracy of the order of σ^{-1}) the Landau Equation (13.2) and obtained $\gamma = (b/2\sigma^{1/2})(\sqrt{r-1} - \sqrt{r_3-1})$ and $\delta = -\frac{37}{2}\sigma$, meaning that the bifurcation for $r = r_3$ was inverse.

Next let us look at the results of Afraimovich *et al.* (1977) (see Figure 17.2). For $r = r_1 \approx 13.92$ they observed a bifurcation for which the separatrices return to a saddle point. For $r > r_1$ from the loops of the separatrices saddle periodic motions L^+ , L^- around the foci C^+ , C^- originate (and simultaneously a non-

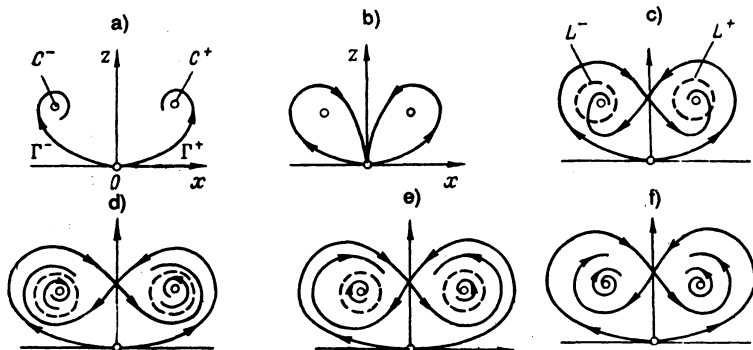


Fig. 17.2. Bifurcations in Lorenz system. (a) $1 < r < r_1$; (b) $r = r_1$; (c) $r_1 < r < r_2$; (d) $r = r_2$; (e) $r_2 < r < r_3$; (f) $r = r_3$.

attractor invariant set of lines Ω_1 of Cantor structure appears, which includes a denumerable set of saddle periodic motions); separatrices Γ^+ , Γ^- intersect and tend toward foci C^- , C^+ . For $r = r_2 \approx 24.06$ separatrices Γ^+ , Γ^- wind around the loops L^- , L^+ , instead of the foci, while a Lorenz attractor of infinitely many sheets Ω_2 appears instead of Ω_1 , the region of attraction being bounded by the stable manifolds of the loops L^- , L^+ (so that the excitement of stochasticity is hard). For $r > r_2$ it is stable, includes 0, Γ^+ , Γ^- , and therefore is not structurally stable, and at it the periodic motions are everywhere dense (motions which may be subject to sequences of period-doubling bifurcations and which may disappear with increasing r only via entanglement in the loop of the separatrices). For $r = r_3$ loops L^+ , L^- contract about points C^+ , C^- , and the latter lose their stability. For $r_3 < r < r_4 \approx 220$ the single stable limit set is the Lorenz attractor (note that, as r decreases from r_4 to r_2 , phase point $M(t)$ remains in the attractor, while for $r < r_2$ it loses stability and $M(t)$ approaches C^+ or C^-).

Figure 17.3 shows an example of a trajectory on an attractor (for $r = 28$ intersecting the plane $Z = 27$). It starts at the coordinate origin, goes around C^+ , then unwinds and approaches C^- , leaves C^- around a spiral to C^+ , etc., the period of revolution around C^+ or C^- being 0.62 and the radii of the spirals varying by 6% during one turn. Lorenz himself gave for this example the Poincaré mapping $Z_{n+1} = \Pi(Z_n)$ of successive maxima of Z ; it has a triangular form, with everywhere $|\Pi'(Z)| > 1$, and it is ergodic and mixing (Bunimovich and Sinai, 1980). The topological Lorenz attractor is made up of two sheets, extending across the trajectories, the right edge of each of these being attached to the left edges of the two sheets (Figure 17.4). Mori and Fujisaka (1980) calculated for a Lorenz attractor, as a function of r , the Lyapunov index σ_1 (the second index σ_2 is here zero, the third $\sigma_3 = \Lambda - \sigma_1$ is negative, and the Lyapunov dimension of the attractor is $2 + \sigma_1|\sigma_3|^{-1}$). For $r < 1$ it is negative, for $r = 1$ (appearance of convective rolls) it goes to zero, for $1 < r < r_3$ it is again negative, and for $r = r_3$ it reaches zero; however, for $r = r_2$ above zero a new branch σ_1 appears, which increases for $r > r_2$. For very large r this branch is repeatedly chopped by 'lacunae' of zero values, corresponding to periodic motions (see in Figure 17.5 the

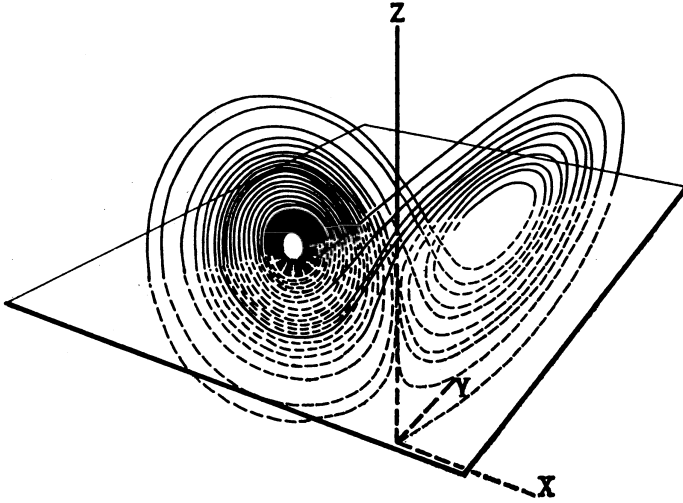


Fig. 17.3. Example of trajectory on Lorenz attractor for $r = 28$ ((X, Y) plane corresponds to $Z = 27$).

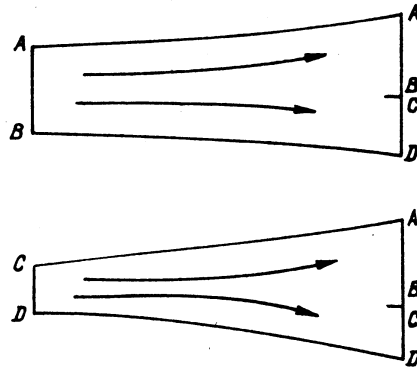


Fig. 17.4. Topology of Lorenz attractor.

example for $b = 4$, $\sigma = 16$; here for $r = 40$ a Lyapunov dimension of 2.06 is obtained).

In recent years the experimental conditions for convection studies have improved, because of the availability of classical liquid helium with a low heat capacity of the walls (compared to the fluid) and with a thermal sensitivity of 10^{-7} (groups of Ahlers and Liebchaber), a laser Doppler anemometer with an accuracy of 10^{-4} cm/s (groups of Gollub and Berger) and computers. It then became possible to observe (on the spectra) quasiperiodic motions with 2 or 3 incommensurable frequencies, sequences of period-doubling bifurcations with scaling law (16.7) for the amplitudes (see in Figure 17.6 the example from Gollub *et al.*, 1980), intermittence, absolute instability of thin layers, fringe effects (rolls perpendicular walls), etc.

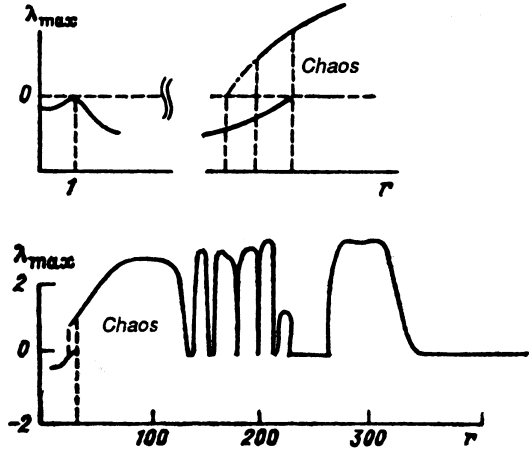


Fig. 17.5. Plots of Lyapunov index $\sigma_1(r)$ for $b = 4$, $\sigma = 16$.

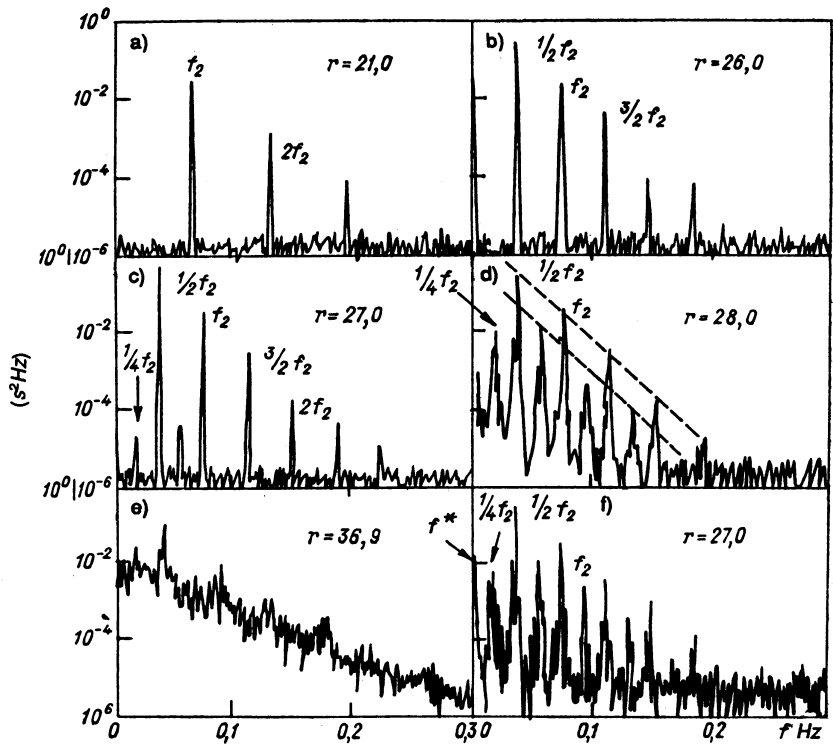


Fig. 17.6. Period-doubling bifurcations in convective rolls.

PART II

PROCESSES

- 4. SURFACE WAVES
- 5. INTERNAL WAVES
- 6. GEOPHYSICAL TURBULENCE
- 7. ROSSBY WAVES

Surface Waves

18. LINEAR THEORY

Waves on the surface of the Earth's water basins, gravitational waves as well as capillary waves, can be described very accurately by assuming water to be incompressible, with a constant density ρ and a motion that is nondivergent. Putting aside for the moment the longest waves, such as tidal waves and tsunamis, we can ignore the Earth's rotation and write the fluid-dynamical equations for water as

$$\begin{aligned} \operatorname{div} \mathbf{u} &= 0; & \frac{\partial \mathbf{u}}{\partial t} + \nabla \left(\frac{p}{\rho} + gz + \frac{|\mathbf{u}|^2}{2} \right) &= \\ \mathbf{u} \times \operatorname{curl} \mathbf{u} + \nu \Delta \mathbf{u}, & \end{aligned} \quad (18.1)$$

where at the bottom $z = -h(x, y)$ the velocity $\mathbf{u}$ goes to zero and at the surface $z = \zeta(x, y, t)$ kinematic boundary condition (5.9) is satisfied. Generally speaking, these equations should be considered together with the fluid-dynamical equations of the air over the water. For the latter equations, the dynamic boundary conditions at the surface $z = \zeta$ have to be satisfied, primarily condition (5.11) for the forces, which in the adiabatic approximation becomes

$$\hat{p} - \hat{p}_a = \gamma \rho \operatorname{div} [(1 + |\nabla \zeta|^2)^{-1/2} \nabla \zeta]. \quad (18.2)$$

Let us represent the velocity field as $\mathbf{u} = \nabla \varphi + \mathbf{v}$, where φ is the velocity potential and $\mathbf{v} = \operatorname{curl} \mathbf{A}$ is its solenoidal (vortical) component ($\operatorname{curl} \mathbf{u} = -\Delta \mathbf{A}$). Then $\operatorname{div} \mathbf{u} = \Delta \varphi = 0$, and $\Delta \mathbf{u} = \Delta \mathbf{v}$, that is, the viscosity force is determined only by the vortical component. Usually it plays a role only in thin boundary layers at the water surface and at the bottom; thus the viscosity can be allowed for in small corrections to the *potential approximation* $\mathbf{u} \approx \nabla \varphi$. In this approximation the right-hand side of Equation (18.1) goes to zero, and we arrive at the *Bernoulli equation*

$$\frac{\partial \varphi}{\partial t} + \frac{p}{\rho} + gz + \frac{|\nabla \varphi|^2}{2} = 0.$$

It need be used only at the water surface $z = \zeta$, after eliminating $\hat{p}$ from it with the

aid of (18.2). Consequently, the potential approximation reduces to the equations:

$$\begin{aligned} \Delta\varphi &= 0; & \left(\frac{\partial\varphi}{\partial t}\right)_{\xi} + g\xi + \frac{|\nabla\varphi|_{\xi}^2}{2} - \gamma \operatorname{div}[(1 + |\nabla\xi|^2)^{-1/2} \nabla\xi] = \\ &= -\frac{p_a}{\rho}; & \frac{\partial\xi}{\partial t} = \left(\frac{\partial\varphi}{\partial z}\right)_{\xi} - (\nabla\varphi)_{\xi} \nabla\xi. \end{aligned} \quad (18.3)$$

The second and third of these equations include three unknown functions of the horizontal coordinates and time: ξ , $\Phi = (\varphi)_{\xi}$, and $\partial\varphi(\partial z)_{\xi}$; the third of these can be expressed in terms of the other two, provided that we solve the boundary problem for the equation $\Delta\varphi = 0$. The total (kinetic + potential, but not internal) energy of a vertical water column of unit section is then ρE , where

$$E = \frac{1}{2} \int_{-h}^{\xi} |\nabla\varphi|^2 dz + \frac{1}{2} g\xi^2 + \gamma[(1 + |\nabla\xi|^2)^{1/2} - 1]. \quad (18.4)$$

The mean (over the horizontal) total energy $\mathcal{H} = \bar{E}$ is a functional of two functions ξ and Φ (note that the contribution of the kinetic energy to $\mathcal{H}$ can be reduced to the form $\frac{1}{2}\Phi \partial\xi/\partial t$). It is easy to see that $\mathcal{H}$ varies in time at a rate $\sim \rho^{-1}\rho_a \partial\xi/\partial t$. For *free waves* (that is, with $p_a = \text{const}$) $\mathcal{H}$ is a Hamiltonian with canonical variables Φ (generalized coordinate) and ξ (generalized momentum), that is,

$$\frac{\partial\Phi}{\partial t} = -\frac{\delta\mathcal{H}}{\delta\xi}; \quad \frac{\partial\xi}{\partial t} = \frac{\delta\mathcal{H}}{\delta\Phi} \quad (18.5)$$

(Zakharov, 1968).

For free waves of small amplitude we can linearize (18.3) with respect to the state of rest (and remove the conditions for $z = \xi$ at the equilibrium level $z = 0$), to obtain

$$\Delta\varphi = 0; \quad \left(\frac{\partial\varphi}{\partial t}\right)_0 + g\xi - \gamma\Delta\xi = 0; \quad \frac{\partial\xi}{\partial t} = \left(\frac{\partial\varphi}{\partial z}\right)_0 \quad (18.6)$$

with the boundary condition at the bottom $(\partial\varphi/\partial z)_{z=-h} = 0$. For a flat bottom $h = \text{const}$ these equations have the elementary wave solutions:

$$\xi = a \cos(\mathbf{k} \cdot \mathbf{x} - \omega t); \quad \varphi = \frac{a\omega \operatorname{ch} k(z+h)}{k \operatorname{sh}(kh)} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t); \quad (18.7)$$

$$\omega^2 = (gk + \gamma k^3) \operatorname{th}(kh) = \frac{1}{2} k^2 c_m^2 \left(\frac{k}{k_m} + \frac{k_m}{k} \right) \operatorname{th}(kh). \quad (18.8)$$

The phase velocity of the waves $c = \omega/k$ here depends on k , that is, the waves exhibit *dispersion*, so that the wave packets must spread out with time. For large

kh (actually even for $kh > \pi/2$) it can be assumed that $\text{th}(kh) \approx 1$ and the dependence of φ and ω on h disappears (waves on deep water). For small kh , on the other hand, we can set $\text{th}(kh) \approx kh$ (waves on shallow water; for $\gamma = 0$ here we obtain $c = (gh)^{1/2}$, so that gravity waves on shallow water do not exhibit dispersion). For $k \ll k_m = (g/\gamma)^{1/2}$ capillary effects are negligible, and only gravity waves remain (on a clean water surface $\lambda_m = 2\pi/k_m \approx 1.7$ cm; then $\min c = c_m = (4g\gamma)^{1/4} \approx 23.2$ cm/s). For these the group velocity $c_g = \partial\omega/\partial k$ has an absolute value less than c (negative dispersion): on deep water $c_g = c/2$, while for $kh \rightarrow 0$ we get $c_g \rightarrow c$. For $k \gg k_m$ gravitational effects are negligible, and only capillary waves remain; for these $c_g > c$ (positive dispersion): on deep water $c_g = 3c/2$. For $k \sim k_m$ the waves are called gravity-capillary waves. Figure 18.1 shows a plot of $c/c_m = f(k/k_m)$ for deep water.

Using the velocity potential (18.7), it is easy to compute the mean (over the volume) kinetic-energy density of linear waves; it is also equal to the potential-energy density (as in any conservative system executing small oscillations), so that the total energy density is

$$E = \frac{\rho a^2 \omega^2}{2k} \text{cth}(kh). \quad (18.9)$$

For pure gravity waves this expression reduces to

$$E = \frac{1}{2} \rho g a^2 = \rho g \langle \xi^2 \rangle,$$

and for pure capillary waves to

$$E = \frac{1}{2} \rho \gamma k^2 a^2 = \rho \gamma \langle |\nabla \xi|^2 \rangle.$$

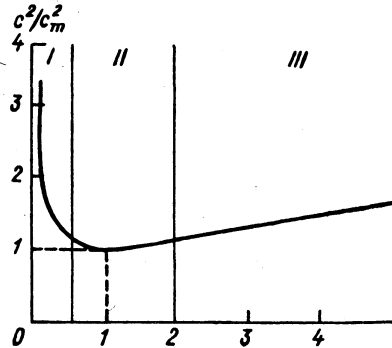


Fig. 18.1. Dispersion relation for waves on deep water: (I) capillary waves; (II) gravity-capillary waves; (III) gravity waves. The abscissa is k/k_m .

Let us also consider, from the vertical projection of the equations of motion in approximation (18.7), the mean (over the horizontal) pressure at a level z :

$$\langle p \rangle = -\rho g z - \frac{1}{2} \rho a^2 \omega^2 \frac{\text{sh}^2 k(z+h)}{\text{sh}^2(kh)}. \quad (18.10)$$

The dynamic pressure (second term) on deep water ($kh \gg 1$) turns out to be equal to $-\frac{1}{2} \rho a^2 \omega^2 e^{2kz}$, so that with increasing depth it rapidly drops to zero.

Now let us discuss the trajectories of fluid particles in the waves (18.7). Let the x axis be directed along the wave vector $\mathbf{k}$, and let $x' = x - x_0$ and $z' = z - z_0$ be the coordinates of *displacement* of a fluid particle located at point (x_0, z_0) at the initial time $t = 0$. Then, if we introduce the quantities $\mu = k(z_0 + h)$ and $\chi = kx_0 - \omega t$, we can, using (18.7), write the equations of motion of the particle $dx'/dt = \partial\varphi/\partial x$, $dz'/dt = \partial\varphi/\partial z$ in the following form:

$$\begin{aligned} \frac{dx'}{dt} &= \frac{a\omega \text{ch}(\mu + kz')}{\text{sh}(kh)} \cos(\chi + kx') = \\ &= \frac{a\omega}{\text{sh}(kh)} (\text{ch} \mu \cos \chi - kx' \text{ch} \mu \sin \chi + kz' \text{sh} \mu \cos \chi + \dots); \\ \frac{dz'}{dt} &= \frac{a\omega \text{sh}(\mu + kz')}{\text{sh}(kh)} \sin(\chi + kx') = \\ &= \frac{a\omega}{\text{sh}(kh)} (\text{sh} \mu \sin \chi + kx' \text{sh} \mu \cos \chi + kz' \text{ch} \mu \sin \chi + \dots). \end{aligned} \quad (18.11)$$

If we retain on the right-hand sides only the first terms, then we get $x' = -A \text{ch} \mu \sin \chi$ and $z' = A \text{sh} \mu \cos \chi$, where $A = a/\text{sh}(kh)$, so that the particle trajectories are ellipses with semiaxes $A \text{ch} \mu$ and $A \text{sh} \mu$ (on deep water they are circles with radii ae^{kz_0}). Now, substituting these expressions for x' and z' into the subsequent terms and averaging over the time, we obtain (with an accuracy of the order of a^2)

$$\langle u \rangle \approx \frac{a^2 k \omega \text{ch}(2\mu)}{2 \text{sh}^2(kh)}; \quad \langle w \rangle = 0. \quad (18.12)$$

Therefore, the particles experience drift, on the average over time (Stokes, 1847), so that their trajectories turn out to be not closed. Note, however, that the calculation of the drift velocity of particles in waves constitutes one of the few problems in which the potential approximation proves to be insufficient, it being necessary to take into account (in the linear theory, simply to add on) the contributions of the vortical component of the velocity $\mathbf{v} = \mathbf{u} - \nabla\varphi$, produced by the molecular viscosity (or, rather, the small-scale eddy viscosity) of the water in the boundary layers at the water surface and at the bottom. These contributions were calculated by Longuet-Higgins (1953).

For instance, in the boundary layer near the bottom $\Omega_y = \partial v_x / \partial z - \partial v_z / \partial x \approx \partial v_x / \partial z$ satisfies the diffusion equation $\partial \Omega_y / \partial t = \nu \partial^2 \Omega_y / \partial z^2$, the solution of which, for the boundary condition at the bottom $v_x = -(\partial \varphi / \partial x)_{z=-h} = -a\omega \cos \chi \operatorname{sh}(kh)$ (where φ is taken according to formula (18.7)) and on condition that it goes to zero far from the bottom, takes the form

$$v_x = -\frac{a\omega e^{-\beta z'} \cos(\chi + \beta z')}{\operatorname{sh}(kh)}, \quad \beta = \left(\frac{\omega}{2\nu}\right)^{1/2}, \quad (18.13)$$

where $z' = z + h$ is the distance from the bottom. The value of v_z is found from the continuity equation. Then, accurate to $O(a^2)$, the Reynolds stress $\langle uw \rangle$ is calculated. Single integration over z of the Reynolds equation $(\partial/\partial z) \langle uw \rangle = \nu(\partial^2 \langle u \rangle / \partial z^2)$ yields $\nu(\partial \langle u \rangle / \partial z) = \langle uw \rangle - \langle uw \rangle_{z=\infty}$, and the next integration on condition that $\langle u \rangle = 0$ at the bottom yields $\langle u(z') \rangle$; as a result, the contribution $\langle u(\infty) \rangle$ to the particle drift is equal to the Stokes drift (18.12), multiplied by $\frac{3}{2}$ (curiously enough, this contribution turns out to be independent of the value of ν , provided it is not zero).

The boundary layer near the free surface may be thin, in comparison with the amplitude of a wave, so that it is not a good idea when describing it to carry the boundary conditions down to the equilibrium level $z = 0$. In this case it is better to use a frame of reference moving with the wave, employing orthogonal curvilinear coordinates $\xi = x - x'$, $\eta = z - z'$ (where x' and z' are first-order expressions derived from (18.11) for $\chi = kx$). In this frame at the free surface $\eta = 0$ and at the bottom $\eta = -h$. The stream function will have the form $\Psi = -c\eta + \psi$, where c is the wave velocity and ψ is the vortical contribution. The vorticity $\Omega_y \approx \partial u / \partial \eta = -\partial^2 \psi / \partial \eta^2$ will satisfy approximately the diffusion equation $-c \partial \Omega_y / \partial \xi = \nu \partial^2 \Omega_y / \partial \eta^2$.

The dissipation $\rho \dot{E}$ of wave energy caused by the viscosity is accompanied by a reduction of the wave momentum $M = \rho E / c$, and thus (in the absence of wind) by a vertical momentum flux $S = \dot{M} = \rho \dot{E} / c$. For a clean water surface this can be compensated only by a mean viscous stress of second order $S = -\rho \nu \partial \langle u \rangle / \partial \eta$. To determine it in the case of deep water, $\dot{E}$ can be calculated sufficiently precisely in terms of the potential motion (18.7), which yields $\dot{E} = -2\nu k \omega^2 a^2$ (this leads to a decrease in the wave amplitude a according to an $\exp(-2\nu k^2 t)$ law; note, however, that the effect of the atmosphere makes a more substantial contribution to the damping of waves that are not too short). Therefore, $\partial \langle u \rangle / \partial \eta = -E / c\nu = 2\omega k^2 a^2$, once again independently of the value of ν . Exactly the same expression is obtained in the case of deep water from (18.2) for $z_0 = 0$, so that the upper boundary layer doubles the Stokes drift. For shallow water, assuming that surface films are present and wind effects are felt, the results are more complicated.

19. RAY THEORY

Let us consider waves on the surface of a water basin of variable depth $h(x)$, assuming the presence of a nonhomogeneous unsteady flow with a velocity $U(x, t)$

at the water surface. Under these conditions refraction of the waves may occur (nonparallel wave crests, that is, curved wave rays), as well as a variation of the wave amplitude along a ray, blocking (no propagation beyond certain boundaries), the formation of turning points, and the formation of envelopes (caustics) at bunches of wave rays.

Let us confine ourselves here to cases in which the scales of the spatial inhomogeneities h and U , and the variations of U with time, are large in comparison with, respectively, the wavelength and wave period. Then the wave can be assumed to be locally flat, as described by the function $\zeta(\mathbf{x}, t) = a(\mathbf{x}, t) \exp(i\psi(\mathbf{x}, t))$, where the amplitude a is a slowly varying function, and the *eikonal* (phase) ψ is an approximately linear function of $\mathbf{x}$ and t (the velocity potential and the other wave characteristics can be described similarly). In such cases the *ray theory* provides a suitable approximation: at each point $(\mathbf{x}, t)$ we define a wave vector $\mathbf{k}$ and a wave frequency σ :

$$\mathbf{k} = \nabla \psi; \quad \sigma = -\frac{\partial \psi}{\partial t}, \quad (19.1)$$

such that the field of the wave vectors $\mathbf{k}(\mathbf{x}, t)$ is potential and the following *equation of conservation of wave number* is valid:

$$\frac{\partial \mathbf{k}}{\partial t} + \nabla \sigma = 0 \quad (19.2)$$

(for a single wave $\psi = \mathbf{k}(\mathbf{x} - \mathbf{c}t)$, where $\mathbf{c}$ is the phase velocity, directed along vector $\mathbf{k}$, so that accurate to the factor $\pi/2$ the number of waves per unit length is equal to $|\mathbf{k}|$; the quantity $\sigma = \mathbf{k} \cdot \mathbf{c}$ is here proportional to the wave-number flux). The frequency σ is the sum of the quantity ω , defined by dispersion relation (18.8) for $h = h(\mathbf{x})$, and possibly also for $\gamma = \gamma(\mathbf{x}, t)$, plus the Doppler shift $\mathbf{k} \cdot \mathbf{U}$:

$$\sigma = \sigma(\mathbf{k}, \mathbf{x}, t) = \omega(\mathbf{k}, \mathbf{x}, t) + \mathbf{k} \cdot \mathbf{U}(\mathbf{x}, t). \quad (19.3)$$

Assuming that here $\mathbf{k} = \mathbf{k}(\mathbf{x}, t)$, we substitute (19.3) into (19.2) and introduce the group velocity of the wave in the absence of a current $\mathbf{c}_g = \partial \omega / \partial \mathbf{k}$, to obtain

$$\frac{\partial \mathbf{k}}{\partial t} + (\mathbf{c}_{ga} + \mathbf{U}_a) \cdot \nabla k_a = -\mathbf{F}; \quad \mathbf{F} = \nabla \omega + k_a \nabla U_a. \quad (19.4)$$

The ray theory is completely analogous to the Hamilton mechanics of a material particle: the eikonal ψ is analogous to the *action* S , the wave vector $\mathbf{k}$ to the *momentum* $\mathbf{p}$, the frequency σ to the *Hamiltonian* $\mathcal{H}$, and the rays to the particle trajectories. Therefore, the equations for the rays are analogs of the Hamilton equations, that is, they have the form

$$\dot{\mathbf{k}} = -\frac{\partial \sigma}{\partial \mathbf{x}} (= -\mathbf{F}); \quad \dot{\mathbf{x}} = \frac{\partial \sigma}{\partial \mathbf{k}} (= \mathbf{c}_g + \mathbf{U}) \quad (19.5)$$

(the first of these equations is identical to (19.4), in view of the condition $\text{curl } \mathbf{k} =$

0). From (19.5) we can derive the variations of the frequency σ along the rays:

$$\dot{\sigma} = \frac{\partial \sigma}{\partial t} \left(= \mathbf{k} \cdot \frac{\partial \mathbf{U}}{\partial t} \right) \quad (19.6)$$

(where the equation inside the parentheses is valid if the dependence of the surface-tension coefficient γ on t need not be considered, in particular for pure gravity waves). Equations (19.5), (19.6) are called the kinematic conservation laws of the waves. In addition, of course, the dynamic laws of conservation of mass, momentum, and energy hold true as well. For instance, taking into account that the presence of currents and waves may change the mean water level $\langle \xi \rangle = d$, we can write the law of mass conservation as

$$\begin{aligned} \frac{\partial}{\partial t} \rho(d+h) + \operatorname{div} \tilde{\mathbf{M}} &= 0; \\ \tilde{\mathbf{M}} &\equiv \rho \tilde{\mathbf{U}}(d+h) = \rho \mathbf{U}(d+h) + \mathbf{M}, \end{aligned} \quad (19.7)$$

where $\mathbf{M}$ is the total (over the depth) mean (over the area) mass flux produced by the fluctuating (wave) motions. Similarly, the law of conservation of momentum can be written as

$$\frac{\partial \tilde{M}_a}{\partial t} + \frac{\partial}{\partial x_\beta} (\tilde{U}_a \tilde{M}_\beta + S_{a\beta}) = T_a \left(= -\rho g(d+h) \frac{\partial d}{\partial x_a} \right), \quad (19.8)$$

where $S_{a\beta}$ is the excess momentum flux produced by the fluctuating motions. If, of these, we take into account only waves, then $S_{a\beta}$ is known as the tensor of the *radiation stresses*; now, accurate to the square of the wave steepness ka , we can write this tensor as

$$S_{a\beta} = \rho \int_{-h}^d \left\langle \frac{\partial \varphi}{\partial x_a} \frac{\partial \varphi}{\partial x_\beta} \right\rangle dz + \left[\int_{-h}^d \langle p \rangle dz - \frac{1}{2} \rho g(d+h)^2 \right] \delta_{a\beta}. \quad (19.9)$$

After calculating this expression with the aid of formulas (18.7) for φ and (18.10) for $\langle p \rangle$, using formula (18.9) for E and the definition of $\mathbf{c}_g$, in the case of pure gravity waves we can simplify tensor (19.9) to

$$S_{a\beta} = E \frac{c_g}{c} \frac{k_a k_\beta}{k^2} + E \left(\frac{c_g}{c} - \frac{1}{2} \right) \delta_{a\beta}. \quad (19.10)$$

If the x axis is in the direction of vector $\mathbf{k}$, then this tensor will be diagonal, and on deep water only one of its components $S_{11} = E/2$ will be nonzero. The capillary effects, firstly, supplement the surface tension $\rho\gamma \cos |\nabla \xi|$, and thus the

excess momentum flux

$$\rho\gamma \langle 1 - \cos |\nabla \xi| \rangle \approx \frac{1}{2} \rho\gamma \langle |\nabla \xi|^2 \rangle;$$

and, secondly, they reduce the subsurface pressure by

$$\rho\gamma \operatorname{div}[(1 + |\nabla \xi|^2)^{-1/2} \nabla \xi];$$

thus we have in toto

$$\begin{aligned} S_{11} &= E \left[\frac{2k(d+h)}{\operatorname{sh} 2k(d+h)} + \frac{1+3\alpha}{2(1+\alpha)} \right]; \\ S_{22} &= E \frac{k(d+h)}{\operatorname{sh} 2k(d+h)}; \quad \alpha = \frac{\gamma k^2}{g} \end{aligned} \quad (19.11)$$

(Longuet-Higgins and Stewart, 1964). For pure capillary waves on deep water ($kh \rightarrow \infty$, $\alpha \rightarrow \infty$) it follows that $S_{11} \rightarrow 3E/2$. Finally, the equation for the wave energy E with $[a/(d+h)]^2 \ll 1$, neglecting energy dissipation, becomes

$$\frac{\partial E}{\partial t} + \operatorname{div} E (\mathbf{c}_g + \mathbf{U}) + S_{\alpha\beta} \frac{\partial U_\alpha}{\partial x_\beta} = 0 \quad (19.12)$$

(Longuet-Higgins and Stewart, 1960–1961).

As an example, we consider stationary gravity waves in a nearshore region with a gently sloping bottom. If the x_1 axis is perpendicular to the shore and the x_2 axis lies along the shore, then from (19.7) we get $\tilde{M}_1 = \text{const} = 0$, so that there exists a mean return current with a speed $U_1 = -M_1/\rho(d+h)$, compensating the mass transport by the waves; outside the wave-breaking zone it will be a small quantity of second order in ka , so that it can be neglected in the other conservation laws. From (19.2) we then obtain $\sigma = \omega = \text{const} = \omega_\infty$. From the condition $\operatorname{curl} \mathbf{k} = 0$ we see that $\partial k_2/\partial x_1 = \partial k_1/\partial x_2 = 0$, which yields

$$k_2 = k \sin \varphi = \text{const} = k_\infty \sin \varphi_\infty, \quad (19.13)$$

where φ is the angle between the wave crests and the shore. From (19.12) outside of the wave-breaking zone we obtain

$$Ec_g \cos \varphi = \text{const} = E_\infty c_{g\infty} \cos \varphi_\infty. \quad (19.14)$$

From (19.13), (19.14) and (18.8), (18.9) we can find k , φ , and E as functions of the dimensionless depth $k_\infty h$. The results of such a calculation have been plotted in Figure 19.1. Moreover, from (19.8) and (19.10) we have

$$\begin{aligned} \frac{\partial S_{11}}{\partial x_1} &= -\rho g (d+h) \frac{\partial d}{\partial x_1} \approx -\rho g h \frac{\partial d}{\partial x_1}; \\ S_{11} &= E \left(2 \frac{c_g}{c} - \frac{1}{2} \right), \end{aligned} \quad (19.15)$$

which implies that with a reduction of the depth the mean level of the water decreases (Longuet-Higgins and Stewart, 1962). In addition, with an increase in $a/(d + h)$ (according to empirical data, up to about 0.4), the waves break, and further from the shore d increases (actually this effect is usually superimposed onto the wind-caused surges).

Another example is the motion of stationary gravity waves on deep water in the x direction on a varying current $U(x)$. From (19.2) we get in this case $\sigma = \omega + kU = \text{const} = \omega_0 = (gk_0)^{1/2}$, where the subscript zero indicates the values at the point where $U = 0$. This gives

$$\frac{k_0}{k} \equiv \frac{c^2}{c_0^2} = \frac{c}{c_0} + \frac{U}{c_0}; \quad \frac{c}{c_0} = \frac{1}{2} \left[1 + \left(1 + \frac{4U}{c_0} \right)^{1/2} \right]. \quad (19.16)$$

For $U < -c_0/4$ the solution has no meaning. At the critical point $c = c_0/2$ and $U/c = -\frac{1}{2}$, the current speed differs in sign only from the local group velocity of the waves, and their energy cannot propagate against the current. Energy equation (19.12) for $S = E/2$ and the result of its integration (after multiplication by c and the use of Equation (19.16) for c) can be written as

$$\frac{\partial}{\partial x} E \left(\frac{c}{2} + U \right) + \frac{1}{2} E \frac{\partial U}{\partial x} = 0; \quad E \left(\frac{c}{2} + U \right) c = \text{const} = \frac{1}{2} E_0 c_0^2. \quad (19.17)$$

If the waves do not break, then their amplitudes will increase upstream according to the law $a = a_0 c_0 [c(c + 2U)]^{-1/2}$ (Longuet-Higgins and Stewart, 1961). Actually, short waves (with frequencies $\omega > g/4 U_m$) do break and do not pass through the maximum U_m of the countercurrent, while longer waves pass through, but become attenuated and turn into moderate swells.

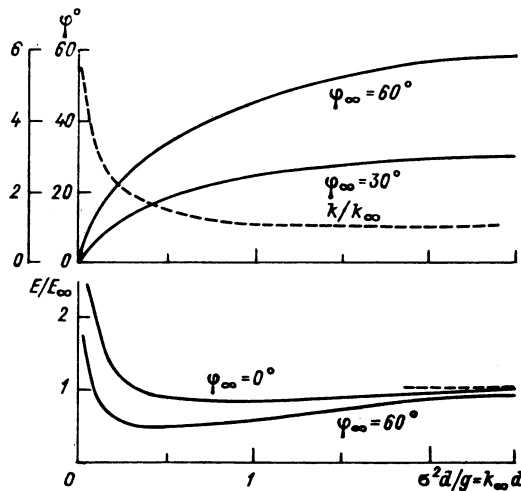


Fig. 19.1. Variations of wave characteristics in the nearshore zone of the sea.

20. STATIONARY WAVES

Stationary waves are waves that propagate without changing shape. In order to ascertain a number of the general properties of such waves, it is convenient to use their Lagrangian description. Assume that stationary waves, propagating in the xz plane through an incompressible ideal heavy fluid with a free surface, move along the horizontal x axis with a velocity c (determinable from the equations of fluid dynamics). The coordinates of the fluid particles, located at points x_0, z_0 in the absence of waves, and the pressure p at these particles, are given by the formulas $w = x' + iz = w(x'_0, z_0)$; $p = p(x'_0, z_0)$, where $x' = x - ct$, $x'_0 = x_0 - ct$, and t is time. In the following, for simplicity we will write the horizontal coordinates x' and x'_0 without the primes. Functions w and p satisfy the Lagrange equations of continuity and motion:

$$\frac{\partial(w, w^*)}{\partial(x_0, z_0)} = -2iM; \quad \frac{\partial^2 w}{\partial x_0^2} = \frac{i}{M} \frac{\partial(\Pi, w)}{\partial(x_0, z_0)}, \quad (20.1)$$

where the asterisk signifies a complex-conjugate quantity, $c^2\Pi = p/\rho + gz$, and $M = \partial(x, z)/\partial(x_0, z_0)$ is an arbitrary real function of z_0 . In the absence of waves, function w must degenerate into $x_0 + iz_0$ and must satisfy the boundary condition $\partial z/\partial x_0 = 0$ at $z_0 = -h$ (the vertical velocity goes to zero at the bottom); as $h \rightarrow \infty$ (deep water) the more rigid condition $\partial w/\partial x_0 \rightarrow 1$ holds true. Moreover, at the curved free surface, at which we place the line $z_0 = 0$, the pressure discontinuity $p - p_a$ caused by the surface tension (the atmospheric pressure p_a will be assumed to be constant, which corresponds to free waves) is proportional to the curvature of the surface:

$$p - p_a = \gamma\rho \left| \frac{\partial w}{\partial x_0} \right|^{-3} \operatorname{Im} \left(\frac{\partial w}{\partial x_0} \frac{\partial^2 w^*}{\partial x_0^2} \right) \quad \text{for } z_0 = 0. \quad (20.2)$$

Note that the vorticity in stationary waves is defined by the formula

$$\Omega = \frac{\partial u_x}{\partial z} - \frac{\partial u_z}{\partial x} = \frac{c}{M} \frac{\partial(\partial w^*/\partial x_0, w)}{\partial(x_0, z_0)}, \quad (20.3)$$

where u_x, u_z are the Euler components of the velocity. In this case the following six theorems are valid (Monin, 1972).

THEOREM 1. *For both potential waves and vortical waves, Equations (20.1) have a first integral*

$$\frac{1}{2} \left| \frac{\partial w}{\partial x_0} \right|^2 + \Pi = \frac{1}{2} + \frac{p_a}{\rho c^2} + f(z_0), \quad (20.4)$$

analogous to the Bernoulli integral, where $f(z_0)$ is an arbitrary function (satisfying

the condition $f(0) = 0$; the constant term on the right-hand side is here so chosen that in the absence of waves condition (20.2) is satisfied).

To prove this theorem, it is sufficient to multiply the second equation of (20.1) by $\partial w^*/\partial x_0$, add the result to the complex-conjugate expression, and then integrate over x_0 .

THEOREM 2. *The field of motion in stationary waves, that is, function w , is found from the equation*

$$\frac{\partial(w, \partial w^*/\partial x_0)}{\partial(x_0, z_0)} = f'(z_0) \quad (20.5)$$

with the boundary condition at the free surface

$$\frac{c^2}{2} \left| \frac{\partial w}{\partial x_0} \right|^2 + gz + \gamma \left| \frac{\partial w}{\partial x_0} \right|^{-3} \operatorname{Im} \left(\frac{\partial w}{\partial x_0} \frac{\partial^2 w^*}{\partial x_0^2} \right) = \frac{c^2}{2} \quad \text{for } z_0 = 0, \quad (20.6)$$

the boundary condition at the bottom, and the condition of degeneracy in the absence of waves. After w has been determined, the pressure p is found from (20.4).

Equation (20.5) is obtained by eliminating $\partial \Pi / \partial z_0$, with the aid of (20.4), from the expression for

$$\frac{\partial w^*}{\partial z_0} \frac{\partial^2 w}{\partial x_0^2} + \frac{\partial w}{\partial z_0} \frac{\partial^2 w^*}{\partial x_0^2},$$

found using the second equation of (20.1). In view of the realness of $f'(z_0)$, Equation (20.5) yields the first equation of (20.1) in the form

$$\frac{\partial}{\partial x_0} \frac{\partial(w, w^*)}{\partial(x_0, z_0)} = 0.$$

Condition (20.6) was derived by excluding p from (20.2) with the aid of (20.4).

THEOREM 3. *There is only one kind of stationary waves in which the pressure p depends only on the vertical Lagrange coordinate z_0 : the trochoidal waves described by Gerstner (1804), that is, gravity waves on deep water ($\gamma = 0$, $h = \infty$). These are described by the formulas*

$$w = x_0 + i \left(z_0 - \frac{1}{2} ka^2 \right) + ia e^{kz_0 + ikx_0}, \quad c^2 = \frac{g}{k}, \quad (20.7)$$

where a and $k = -(i/\rho c^2 M) (\partial p / \partial z_0)$ are arbitrary constants.

For the proof we substitute $\Pi = p(z_0)/\rho c^2 + gz/c^2$ into the second equation of (20.1), which then becomes the following equation, linear with respect to w and ordinary with respect to x_0 : $w'' = -(ig/c^2) + ikw'$. Its solution, which satisfies the

first equation of (20.1) and the boundary condition at the deep bottom, has the form of (20.7). It should be remarked that for Gerstner waves (20.7)

$$M = 1 - k^2 a^2 e^{2kz_0}; \quad \Omega = -\frac{c}{M} 2a^2 k^3 e^{2kz_0}. \quad (20.8)$$

From this it is evident, in particular, that they are vortical waves. Therefore, Gerstner waves constitute an exception: in all other stationary waves the pressure must be a function of the horizontal Lagrange coordinate x_0 .

THEOREM 4. *In stationary waves the vorticity may depend only on the vertical Lagrange coordinate z_0 .*

This theorem is proven by a comparison of (20.3) and (20.5): $\Omega = -cf'(z_0)/M(z_0)$. In particular, according to (20.8), for Gerstner waves $f'(z_0) = 2a^2 k^3 e^{2kz_0}$, so that $f(z_0) = a^2 k^2 (e^{2kz_0} - 1)$.

THEOREM 5. *All vortical waves with different distributions $f'(z_0)$ of the normalized vorticity $(-\Omega M/c)$ can be obtained from waves with $f'(z_0) \equiv 1$ by replacing the vertical coordinate z_0 by f , that is, the universal equation*

$$\frac{\partial(w, \partial w^*/\partial x_0)}{\partial(x_0, f)} = 1, \quad (20.9)$$

is satisfied, where the free surface turns out to be the line $f = 0$.

THEOREM 6. *In the case of potential waves Equation (20.5) has a first integral*

$$\frac{\partial w^*}{\partial x_0} = F(w), \quad (20.10)$$

where $F(w)$ is an arbitrary analytical function, which must in the absence of waves degenerate to unity and must satisfy the condition at the bottom $\text{Im } F = 0$ for $z = -h$ (or $F \rightarrow 1$ as $z \rightarrow -\infty$) and the condition at the free surface

$$\frac{c^2}{2} |F|^2 + gz + \gamma |F|^{-3} \text{Im}(F' F^{*2}) = \frac{c^2}{2} \quad \text{for } z_0 = 0. \quad (20.11)$$

This theorem is proven as a corollary of Theorem 2 on condition that $f'(z_0) \equiv 0$ (and thus $f(z_0) \equiv 0$), which distinguishes potential waves. Function $F(w)$ differs only by a constant factor $(-c)$ from the Euler complex velocity $u - iv$ in a coordinate system x, z moving with a velocity c . Consequently, the use of (20.10) actually constitutes a transition to a Eulerian description of the motion.

21. STATIONARY POTENTIAL WAVES

As an object of general consideration, let us here take a look at periodic gravity-

capillary waves (with a finite wave number $k = 2\pi/\lambda$, $0 < k < \infty$) on the surface of water of a certain depth ($0 < h < \infty$). At the limit as $k \rightarrow 0$, if it exists, a solitary wave (soliton) must be formed from these. For both periodic waves and solitons the limits $h \rightarrow \infty$ (deep water) and $kh < 1$ (shallow water or long waves) are of considerable interest. Finally, the limits of pure gravity waves and pure capillary waves should also be considered. All these waves are symmetrical (this was proven for gravity waves by Levi-Civita, 1925).

Periodic gravity waves were first studied by Stokes, in 1847 on deep water and in 1880 on water of a finite depth. For the characteristics of these waves, he found the first terms of the series expansions in powers of their steepness ka . For instance, for waves on deep water, accurate to the fourth power of the steepness, he got

$$\begin{aligned}\zeta = -a \left[\cos \chi - \left(\frac{1}{2} ka + \frac{17}{24} k^3 a^3 \right) \cos 2\chi + \frac{3}{8} k^2 a^2 \cos 3\chi + \right. \\ \left. + \frac{1}{3} k^3 a^3 \cos 4\chi + \dots \right]; \\ c^2 = \frac{g}{k} \left(1 + k^2 a^2 + \frac{5}{4} k^4 a^4 + \dots \right); \quad \chi = kx,\end{aligned}\tag{21.1}$$

an expression for ζ which differs from the corresponding expansion for a trochoid only by the term $\frac{1}{3} k^3 a^3 \cos 2\chi$ (in 1980 Chen and Safman also detected other, so-called *irregular*, waves). Solitary gravity waves (also only the first terms of the expansions and actually only on shallow water) were first considered by Boussinesq (1871, 1890) and Rayleigh (1876). An equation describing gravity waves on shallow water, having both periodic and soliton solutions, was derived by Korteweg and de Vries (1895); it is now known as the KdV equation.

However, in the theory of periodic and solitary gravity waves, two major problems remained unsolved for a long time: the problem of convergence of the series expansions describing them (that is, of their existence in general) and the problem of their hydrodynamic stability (and thereby of their realizability in nature). A proof of the existence of periodic gravity waves, on deep water and on water of a finite depth, was given by Nekrasov (1921, 1922). The methods developed by him were subsequently used widely by others. For example, Slezkin (1935, 1937) employed Nekrasov's methods to demonstrate the existence of periodic capillary waves, both on water of finite depth and for $h \rightarrow \infty$ (more recently, Krepper, 1957, obtained in explicit analytical form an exact particular solution for such waves). The existence of periodic gravity-capillary waves was proven by Sekerzh-Zen'kovich (1972). Later, Emik and Toland (1981) showed that the equations of Nekrasov allow a limit transition $\lambda \rightarrow \infty$ from a periodic wave to a solitary wave for a fixed h (which here has the sense of an asymptotic bottom depth as $|x| \rightarrow \infty$).

In Nekrasov's theory the waves are described in an Eulerian frame of reference (x, z) , moving with the waves, and the complex coordinate $w = x + iz$ is used. The

Eulerian velocities can be expressed in terms of the potential φ or the stream function ψ using the formulas $u_x = -\partial\varphi/\partial x = -\partial\psi/\partial z$ and $u_z = -\partial\varphi/\partial z = \partial\psi/\partial x$ (where $\Delta\varphi = \Delta\psi = 0$), so that the complex potential $W = \varphi + i\psi$ is an analytic function of the complex variable w , the derivative of which is the complex velocity:

$$-\frac{dW}{dw} = u_x - iu_z = |u|e^{-i\theta}. \quad (21.2)$$

Let us consider in the w plane a water-filled region D which is exactly one wavelength $\lambda = 2\pi/k$ long in the x direction. The region is bounded by two segments (1) and (2) of vertical lines passing through adjacent lowest points of the wave and by the parts of the free surface of the fluid $z = \zeta$ (3) and the bottom $z = -h$ (4) located between them. Since the motion is steady, the free surface and the bottom are streamlines; let us set $\psi = 0$ at (3) and $\psi = \psi_1 > 0$ at (4) (this quantity requires further definition). In view of the symmetry of the wave at a vertical segment (5) passing through the top of the wave, as well as at segments (1) and (2), the potential φ has constant values; let us set $\varphi = 0$ at (5) and $\varphi = \pm c\lambda/2 = \pm \pi c/k$ at (1) and (2) (where c requires further definition).

In the W plane region D will be a rectangle with sides: (1) $\varphi = \pi c/k$; $0 < \psi < \psi_1$; (2) $\varphi = -\pi c/k$; $0 < \psi < \psi_1$; (3) $-\pi c/k < \varphi < \pi c/k$; $\psi = 0$; (4) $-\pi c/k < \varphi < \pi c/k$; $\psi = \psi_1$. Let us so map plane W onto an auxiliary complex plane $w_1 = |w_1| e^{i\chi}$ that region D is mapped onto an annulus D_1 with its center at the origin, with an outer circumference (3) of unit radius and an inner circumference (4) of radius $r_1 < 1$, and with a cut in the negative part of the real axis following shorelines (1) and (2); the segment of the positive part of the real axis in annulus D_1 here corresponds to segment (5) in region D . This mapping is described by the formulas:

$$W = -i \frac{c}{k} \ln w_1; \quad \frac{\partial W}{\partial w_1} = \frac{i}{k} \left[\frac{1}{w_1} + f(w_1) \right], \quad (21.3)$$

where $f(w_1)$ is some Laurent expansion. Note that the imaginary part of the first of these expressions for $z = -h$ yields the following relation between the parameters:

$$\psi_1 = -\frac{c}{k} \ln r_1. \quad (21.4)$$

In order to determine c , let us calculate the integral $S = -\int (dW/dw) dw$ along the line $z = -h$ within region D . According to (21.2), it is equal to $\int_{-\lambda/2}^{\lambda/2} u_x dx$. However, according to the first formula of (21.3) it is

$$-\int \frac{dW}{dw_1} dw_1 = \frac{ic}{k} \int \frac{dw_1}{w_1} = \frac{c}{k} \int_{-\pi}^{\pi} d\chi = c\lambda.$$

If we equate these two expressions, we see that c is the mean value of u_x at the

bottom. Thus the water flow rate (discharge) $\psi_1 = ch$ and, from (21.4), we obtain $r_1 = e^{-kh}$. The complex velocity (21.2) can, with the aid of (21.3), be reduced to

$$-\frac{dW}{dw} = \frac{c}{1 + w_1 f(w_1)} = c e^{-\Omega(w_1)},$$

$$\Omega(w_1) = a_0 + \sum_{n=1}^{\infty} a_n \left[\left(\frac{w_1}{r_1} \right)^n + \left(\frac{w_1}{r_1} \right)^{-n} \right] = \ln R + i\Theta. \quad (21.5)$$

At the bottom $z = -h$ we have $|w_1| = r_1$, so that $w_1 = r_1 e^{i\chi}$. The requirement that the vertical velocity u_z go to zero at the bottom according to (21.2) means that dW/dw is real and according to (21.5) means that all the coefficients a_n of the Laurent expansion for $\Omega(w_1)$ are real.

At the free surface $|w_1| = 1$, so that $w_1 = e^{i\chi}$ and the quantities R and Θ turn out to be functions $R(\chi)$ and $\Theta(\chi)$ only of χ . The relation between them can be ascertained by separating in the second formula of (21.5) for $w_1 = e^{i\chi}$ the real part from the imaginary part and writing these (using (21.4)) in the form

$$\Theta(\chi) = \sum_{n=1}^{\infty} b_n \sin n\chi; \quad b_n = 2a_n \operatorname{sh}(nkh); \quad (21.6)$$

$$\frac{\partial \ln R(\chi)}{\partial \chi} = -\frac{1}{3} \sum_{n=1}^{\infty} b_n \mu_n \sin n\chi; \quad \mu_n = 3n \operatorname{cth}(nkh). \quad (21.6')$$

Then via direct integration we can show the validity of the following relation between series (21.6) and (21.6'):

$$\Theta(\chi) = -3 \int_{-\pi}^{\pi} \frac{\partial \ln R(\chi')}{\partial \chi'} G(\chi, \chi') d\chi'; \quad (21.7)$$

$$G(\chi, \chi') = \sum_{n=1}^{\infty} \frac{\varphi_n(\chi) \varphi_n(\chi')}{\mu_n}, \quad \varphi_n(\chi) = \frac{\sin n\chi}{\sqrt{\pi}}. \quad (21.7')$$

It should be stressed that the kernel $G(\chi, \chi')$ depends on a single parameter kh . Nekrasov discovered that series (21.7') for kernel $G(\chi, \chi')$ can be summed and expressed in terms of elliptic functions, which we will write, somewhat more simply than in the original, as

$$G(\chi, \chi') = \frac{1}{6\pi} \ln \left| \frac{\operatorname{sn} \left[\frac{K}{\pi} (\chi + \chi') \right]}{\operatorname{sn} \left[\frac{K}{\pi} (\chi - \chi') \right]} \right|; \quad \frac{K'(\kappa)}{K(\kappa)} = \frac{2kh}{\pi}, \quad (21.8)$$

where $\operatorname{sn} v$ is the elliptic Jacobi sine with modulus κ , while $K(\kappa)$ is a complete elliptic integral of the first kind and $K'(\kappa) = K(\sqrt{1 - \kappa^2})$. For the proof of formula

(21.8) we will use the following representation of $\operatorname{sn} v$ (Gradshtein and Ryzhik, 1962, formula 8.146.20):

$$\ln \operatorname{sn} v = \ln \frac{2K}{\pi} + \ln \sin \frac{\pi v}{2K} - 2 \sum_{n=1}^{\infty} \frac{1}{n} \frac{q^n}{1 - q^n} \left(1 - \cos \frac{n\pi v}{K} \right), \quad (21.9)$$

where

$$q = \exp \left(-\frac{\pi K'}{K} \right) = \exp(-2kh) = r_1^2.$$

Substitution of (21.9) into (21.8) then yields

$$G(\chi, \chi') = G_0(\chi, \chi') + \frac{1}{3\pi} \sum_{n=1}^{\infty} \frac{1}{n} \frac{r_1^{2n}}{1 + r_1^{2n}} [\cos n(\chi + \chi') - \cos n(\chi - \chi')];$$

$$G_0(\chi, \chi') = \frac{1}{3\pi} \sum_{n=1}^{\infty} \frac{\sin n\chi \sin n\chi'}{n} = \frac{1}{6\pi} \ln \left| \frac{\sin \frac{\chi + \chi'}{2}}{\sin \frac{\chi - \chi'}{2}} \right| \quad (21.10)$$

where $G_0(\chi, \chi')$ is the kernel of Equation (21.7) of the Nekrasov theory for periodic waves on deep water. Formulas (21.7') can already be derived from (21.10) in an elementary manner.

Relation (21.7) between functions $\Theta(\chi)$ and $R(\chi)$ can be inverted. To do this, we rewrite formulas (21.6) and (21.6') as

$$\ln R(\chi) = a_0 + \sum_{n=1}^{\infty} c_n \cos n\chi; \quad c_n = 2a_n \operatorname{ch}(nkh); \quad (21.11)$$

$$\frac{\partial \Theta(\chi)}{\partial \chi} = \sum_{n=1}^{\infty} c_n v_n \cos n\chi; \quad v_n = n \operatorname{th}(nkh) \quad (21.11')$$

and then instead of (21.7) and (21.7') we get

$$\ln R(\chi) = a_0 + \int_{-\pi}^{\pi} \frac{\partial \Theta(\chi')}{\partial \chi'} \gamma(\chi, \chi') d\chi'; \quad (21.12)$$

$$\gamma(\chi, \chi') = \sum_{n=1}^{\infty} \frac{\psi_n(\chi) \psi_n(\chi')}{v_n}; \quad \psi_n(\chi) = \frac{\cos n\chi}{\sqrt{\pi}}. \quad (21.12')$$

In such a form this relation was used by Slezkin to verify the existence of periodic

capillary waves. Kernel (21.12') can be expressed as

$$\begin{aligned}\gamma(\chi, \chi') &= \gamma_0(\chi, \chi') + \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \frac{r_1^{2n}}{1 - r_1^{2n}} \times \\ &\times [\cos n(\chi + \chi') + \cos n(\chi - \chi')]; \\ \gamma_0(\chi, \chi') &= \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{\cos n\chi \cos n\chi'}{n} = \\ &= -\frac{1}{2\pi} \ln 4 \left| \sin \frac{\chi + \chi'}{2} \sin \frac{\chi - \chi'}{2} \right|,\end{aligned}\quad (21.13)$$

where γ_0 corresponds to the limit of the deep water as $r_1 \rightarrow 0$. If this formula for $\gamma(\chi, \chi')$ is differentiated with respect to χ and χ' , then the series obtained can be summed completely:

$$\begin{aligned}\Gamma(\chi, \chi') &= \frac{\partial^2 \gamma(\chi, \chi')}{\partial \chi \partial \chi'} = \Gamma_0(\chi, \chi') - \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{nr_1^{2n}}{1 - r_1^{2n}} \times \\ &\times [\cos n(\chi + \chi') - \cos n(\chi - \chi')]; \\ \Gamma_0(\chi, \chi') &= \frac{1}{8\pi} \left[\sin^{-2} \left(\frac{\chi + \chi'}{2} \right) - \sin^{-2} \left(\frac{\chi - \chi'}{2} \right) \right].\end{aligned}\quad (21.14)$$

Now we can apply the following formula 8.146.27 of Gradshtein and Ryzhik:

$$\begin{aligned}\operatorname{sn}^{-2} v &= \frac{\pi^2}{4K^2} \sin^{-2} \frac{\pi v}{2K} + \frac{K - E}{K} - \\ &- \frac{2\pi^2}{K^2} \sum_{n=1}^{\infty} \frac{nq^{2n}}{1 - q^{2n}} \cos \frac{n\pi v}{K},\end{aligned}\quad (21.15)$$

where

$$q = \exp \left(-\frac{\pi K'}{K} \right) = r_1 = \exp(-kh),$$

so that $K'/K = kh/\pi$, in contrast to (21.8). Thus we have

$$\Gamma(\chi, \chi') = \frac{K^2}{2\pi^3} \left\{ \operatorname{sn}^{-2} \left[\frac{K}{\pi} (\chi + \chi') \right] - \operatorname{sn}^{-2} \left[\frac{K}{\pi} (\chi - \chi') \right] \right\}. \quad (21.16)$$

For direct application of this kernel, we integrate (21.12) by parts over χ for the condition $\Theta(\pi) = \Theta(-\pi) = 0$ and then we differentiate this formula with respect to χ :

$$\frac{\partial \ln R(\chi)}{\partial \chi} = - \int_{-\pi}^{\pi} \Theta(\chi') \Gamma(\chi, \chi') d\chi'. \quad (21.17)$$

This relation, together with (21.16) and the equation $K'/K = kh/\pi$ for modulus κ , is the inverse of (21.7).

Relations (21.7) and (21.17) can be put into a form suitable for application to the case of solitons, by changing from χ to another independent variable $\chi_1 = \chi/k$ and denoting

$$\Theta(k\chi_1) \equiv \Theta_1(\chi_1), \quad R(k\chi_1) \equiv R_1(\chi_1), \quad G(k\chi, k\chi') \equiv G_1(\chi_1, \chi'_1) \quad \text{and} \\ k^2 \Gamma(k\chi_1, k\chi'_1) \equiv \Gamma_1(\chi_1, \chi'_1),$$

so that (21.7) and (21.17) become:

$$\Theta_1(\chi_1) = -3 \int_{-\pi/k}^{\pi/k} \frac{\partial \ln R_1(\chi'_1)}{\partial \chi'_1} G_1(\chi_1, \chi'_1) d\chi'_1; \quad (21.18)$$

$$\frac{\partial \ln R_1(\chi_1)}{\partial \chi_1} = - \int_{-\pi/k}^{\pi/k} \Theta_1(\chi'_1) \Gamma_1(\chi_1, \chi'_1) d\chi'_1. \quad (21.18')$$

These relations permit a limit transition $k \rightarrow 0$ from a periodic wave to a solitary wave for fixed h (which takes on the meaning of an asymptotic bottom depth for $|x| \rightarrow \infty$). At the limit the integrals over χ'_1 are taken from $-\infty$ to $+\infty$, and in (21.8) the modulus κ of the elliptic sines goes to unity, $2K/\lambda = K'/2h \rightarrow \pi/4h$ and $\text{sn } v \rightarrow \text{th } v$, so that

$$G_1(\chi_1, \chi'_1) \rightarrow G_*(\chi_1, \chi'_1) = \frac{1}{6\pi} \ln \left| \frac{\text{th} \left[\frac{\pi}{4h} (\chi_1 + \chi'_1) \right]}{\text{th} \left[\frac{\pi}{4h} (\chi_1 - \chi'_1) \right]} \right|. \quad (21.19)$$

For solitons on deep water $h \rightarrow \infty$ (if they exist), which yields

$$G_*(\chi_1, \chi'_1) \rightarrow G_{0*}(\chi_1, \chi'_1) = \frac{1}{6\pi} \ln \left| \frac{\chi_1 + \chi'_1}{\chi_1 - \chi'_1} \right|. \quad (21.20)$$

Exactly the same expression is, of course, obtained from function $G_0(\chi, \chi')$ of formula (21.10), if in it we set $(\chi, \chi') = k(\chi_1, \chi'_1)$ and then pass to the limit as $k \rightarrow 0$.

In (21.16) at the limit as $k \rightarrow 0$, the modulus κ of the elliptic sines approaches

unity, $2K/\lambda = K'/h \rightarrow \pi/2h$, and we obtain

$$\Gamma_1(\chi_1, \chi'_1) \rightarrow \Gamma_*(\chi_1, \chi'_1) = \frac{\pi}{8h^2} \left\{ \operatorname{cth}^2 \left[\frac{\pi}{2h} (\chi_1 + \chi'_1) \right] - \right. \\ \left. - \operatorname{cth}^2 \left[\frac{\pi}{2h} (\chi_1 - \chi'_1) \right] \right\}. \quad (21.19')$$

Thus at the limit as $h \rightarrow \infty$ or from (21.14) (after multiplication by k^2 , a variable change $\chi = k\chi_1$, and taking the limit $k \rightarrow 0$), for a 'soliton' kernel on deep water we get

$$\Gamma_*(\chi_1, \chi'_1) \rightarrow \Gamma_{0*}(\chi, \chi') = \frac{1}{2\pi} [(\chi_1 + \chi'_1)^{-2} - (\chi_1 - \chi'_1)^{-2}]. \quad (21.20')$$

Let us also consider the asymptote of shallow water $h \rightarrow 0$ for fixed k . Here we can use for periodic waves relations (21.18) and (21.18') (and for solitons these same relations, but with integrals in infinite limits), retaining in expressions (21.19) and (21.19') only the principal terms as $h \rightarrow 0$:

$$G_1(\chi_1, \chi'_1) \sim \frac{1}{3\pi} \left(\exp \left[-\frac{\pi}{2h} |\chi_1 - \chi'_1| \right] - \exp \left[-\frac{\pi}{2h} |\chi_1 + \chi'_1| \right] \right); \quad (21.21)$$

$$\Gamma_1(\chi_1, \chi'_1) \sim \frac{\pi}{2h^2} \left(\exp \left[-\frac{\pi}{h} |\chi_1 + \chi'_1| \right] - \exp \left[-\frac{\pi}{h} |\chi_1 - \chi'_1| \right] \right). \quad (21.21')$$

At the free surface $w_1 = e^{i\chi}$, in addition to the relation between functions $R(\chi)$ and $\Theta(\chi)$ derived from the second formula of (21.5), different forms of which were given above, we can also obtain a kinematic relation and a dynamic relation between these functions and the function $z = \zeta(x)$, characterizing the shape of the wave surface. The kinematic relation is arrived at by writing for $w_1 = e^{i\chi}$ the second equation of (21.3) using the notation of (21.5):

$$\frac{\partial x}{\partial \chi} = -\frac{R}{k} \cos \Theta; \quad \frac{\partial \zeta}{\partial \chi} = -\frac{R}{k} \sin \Theta, \quad (21.22)$$

from which it is evident that $\partial \zeta / \partial x = \operatorname{tg} \Theta$ is the slope of the wave profile. The dynamic equation in this case is the Bernoulli equation (18.3) or (20.6), derived for stationary waves with $z = \zeta$, in which according to (21.2) and (21.5) we can set $|\mathbf{u}|_\zeta = cR$; then with the aid of (21.22) we can put this equation into the form

$$\frac{c^2}{2R^2} + g\zeta + \frac{\gamma k}{R} \frac{\partial \Theta}{\partial \chi} = \frac{c^2}{2}. \quad (21.23)$$

Here it is convenient to eliminate ζ , after differentiating this equation with respect

to χ and using the second formula of (21.22), to obtain

$$\frac{c^2}{R^3} \frac{\partial R}{\partial \chi} + \frac{gR}{k} \sin \Theta = \frac{\gamma k}{R} \left(\frac{\partial^2 \Theta}{\partial \chi^2} - \frac{1}{R} \frac{\partial R}{\partial \chi} \frac{\partial \Theta}{\partial \chi} \right). \quad (21.24)$$

22. STATIONARY GRAVITY WAVES

For pure gravity waves we set $\gamma = 0$ in (21.24) and then integrate this equation over R , to obtain

$$R^3(0) R^{-3}(\chi) = 1 + \mu \int_0^\chi \sin \Theta(\chi') d\chi';$$

$$\mu = \frac{3gR^3(0)}{kc^2} = \frac{3gc}{k|\mathbf{u}|_0^3}. \quad (22.1)$$

Substitution of this expression for R into (21.7) then yields for $\Theta(\chi)$ the following integral equation of Nekrasov:

$$\Theta(\chi) = \mu \int_{-\pi}^{\pi} \frac{\sin \Theta(\chi') G(\chi, \chi') d\chi'}{1 + \mu \int_0^{\chi'} \sin \Theta(\chi'') d\chi'}. \quad (22.2)$$

Nekrasov sought the solution $\Theta(\chi)$ of this equation in the form of a series in powers of $\mu' = \mu - \mu_1$ (where μ_1 is defined in (21.6'), and using the majorant method he showed convergence of this series for $\mu' > 0$, that is, the existence of a solution (whereas a nontrivial solution, regular in the region $-\pi \leq \chi \leq \pi$, $0 \leq \mu \leq \mu_1$, does not exist). Once $\Theta(\chi)$ has been determined, from (21.6) we can find all the coefficients a_n for $n \neq 0$ (a_0 is found by integrating the first relation (21.22) over χ from $-\pi$ to π , giving the wavelength λ), and then (21.5) defines the entire current.

In the notation of (21.18) formulas (22.1) and (22.2) take the form:

$$R_1(\chi_1) = \left(\frac{c^2}{3g} \right)^{1/3} \left[\mu_*^{-1} + \int_0^{\chi_1} \sin \Theta_1(\chi'_1) d\chi'_1 \right]^{-1/3}; \quad \mu_* = \frac{3gc}{|\mathbf{u}|_0^3}, \quad (22.1')$$

$$\Theta_1(\chi_1) = \int_{-\pi/k}^{\pi/k} \frac{\sin \Theta_1(\chi'_1) G_1(\chi_1, \chi'_1) d\chi'_1}{\mu_*^{-1} + \int_0^{\chi'_1} \sin \Theta_1(\chi''_1) d\chi''_1}. \quad (22.2')$$

Equations (21.22) are then integrated as follows:

$$x(\chi_1) = \left(\frac{c^2}{3g} \right)^{1/3} \int_{\chi_1}^0 \frac{\cos \Theta_1(\chi_1') d\chi_1'}{\left[\mu_*^{-1} + \int_0^{\chi_1'} \sin \Theta_1(\chi_1'') d\chi_1'' \right]^{1/3}}; \quad (22.3)$$

$$\xi(0) - \xi(\chi_1) = \frac{3}{2} \left(\frac{c^2}{3g} \right)^{1/3} \left\{ \left[\mu_*^{-1} + \int_0^{\chi_1} \sin \Theta_1(\chi_1'') d\chi_1'' \right]^{2/3} - \mu_*^{-2/3} \right\}. \quad (22.4)$$

The Bernoulli equation (21.23) is an alternative to (22.4) for $\gamma = 0$, and it yields

$$\xi = \frac{c^2 - |\mathbf{u}|^2}{2g}; \quad \xi(0) = \frac{c^2 - |\mathbf{u}|_0^2}{2g}, \quad (22.5)$$

from which it is evident, in particular, that the greatest height of the wave $\xi(0)$ is reached at $|\mathbf{u}|_0 = 0$ (that is, for $\mu_*^{-1} = 0$), and that its value is $\xi(0) = c^2/2g$. Such a limit configuration was predicted by Stokes himself, who noted that in it the top of the wave comes to a point and the angle between the right-hand and left-hand tangents to the wave profile at the top of the wave must be 120° . This result turns out to be valid for both periodic waves and solitons, and it does not depend on the depth h . Actually, we can assume that the limit configuration is determined by, aside from the coordinates, a single parameter g . Then, in polar coordinates (r, θ) with the origin at the corner point and with the $\theta = 0$ axis directed vertically downward, the stream function $\psi(r, \theta)$, which equated to zero yields the formula for the water surface, must for reasons of dimensionality have the form $(gr^3)^{1/2} \Psi(\theta)$ in the vicinity of a crest. However, it satisfies the Laplace equation $\Delta\psi = 0$, or in polar coordinates $\partial^2\psi/\partial r^2 + \partial\psi/r \partial r + \partial^2\psi/r^2 \partial\theta^2 = 0$. This gives $\Psi(\theta) = \cos(3\theta/2)$, so that $\Psi = 0$ gives $\theta = \pm 60^\circ$, *quod erat demonstrandum*. We remark that the indicated singularity of function $w(W) \sim (iW)^{2/3}$ at the top of the wave is not regular: the next term of the asymptote w contains an irrational power of W ; moreover, for waves of prelimit configuration the singularity at the wave top has a different order (Grant, 1973).

We note, too, that for solitons $R_1 \rightarrow 1$, $\xi \rightarrow 0$, and $|\mathbf{u}| \rightarrow c$ as $|\chi_1| \rightarrow \infty$, so that from (22.1') we get

$$\frac{c^2}{3g} = \mu_*^{-1} + \int_0^\infty \sin \Theta_1(\chi_1) d\chi_1. \quad (22.6)$$

For solitons of the limit configuration we must here set $\mu_*^{-1} = 0$.

According to the assumption of Benjamin and Lighthill (1954), which is easily verified for shallow water (see below), stationary gravity waves are completely determined by, aside from the free-fall acceleration g , three constant parameters:

the discharge Q , the energy density R , and the momentum flux per unit mass S (see (19.9)). These parameters are defined as follows:

$$\begin{aligned} Q &= \int_0^\eta u_x dy = \psi_0 - \psi_\eta; \\ R &= \frac{p}{\rho} + \frac{1}{2} |\mathbf{u}|^2 + gy = \frac{1}{2} |\mathbf{u}_\eta|^2 + g\eta; \\ S &= \int_0^\eta \left(\frac{p}{\rho} + u_x^2 \right) dy = \int_0^\eta \left[R - gy - \frac{1}{2} \left(\frac{\partial \psi}{\partial x} \right)^2 + \frac{1}{2} \left(\frac{\partial \psi}{\partial y} \right)^2 \right] dy, \end{aligned} \quad (22.7)$$

where $y = h + z$ and $\eta = h + \zeta$. The second expression for R is here obtained from the condition $\partial S / \partial x = 0$:

$$\begin{aligned} \frac{\partial S}{\partial x} &= \left[R - g\eta - \frac{1}{2} \left(\frac{\partial \psi}{\partial x} \right)_\eta^2 + \frac{1}{2} \left(\frac{\partial \psi}{\partial y} \right)_\eta^2 \right] \frac{\partial \eta}{\partial x} + \\ &+ \int_0^\eta \left(\frac{\partial \psi}{\partial y} \frac{\partial^2 \psi}{\partial y \partial x} - \frac{\partial \psi}{\partial x} \frac{\partial^2 \psi}{\partial x^2} \right) dy. \end{aligned} \quad (22.7')$$

Since $\Delta \psi = 0$, that is, $\partial^2 \psi / \partial x^2 = -\partial^2 \psi / \partial y^2$, the integral is equal to

$$\left(\frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial y} \right)_\eta = -(u_x u_z)_\eta = -(u_x^2)_\eta \frac{\partial \eta}{\partial x} = -\left(\frac{\partial \psi}{\partial y} \right)_\eta^2 \frac{\partial \eta}{\partial x},$$

since the kinematic boundary condition at the free surface yields $u_z = u_x (\partial \eta / \partial x)$. Now, by substituting this value of the integral into (22.8) and setting $\partial S / \partial x$ equal to zero, we obtain for R the second equation of (22.7). The third equation of (22.7) reduces to

$$g\eta^3 - 2R\eta^2 + 2S\eta = \eta \int_0^\eta \left[\left(\frac{\partial \psi}{\partial y} \right)^2 - \left(\frac{\partial \psi}{\partial x} \right)^2 \right] dy. \quad (22.8)$$

The harmonic function $\psi(x, y)$ can be written, accurate to a constant term, as

$$\begin{aligned} \psi &= yf(x) - \frac{y^3}{3!} f^{(2)}(x) + \frac{y^5}{5!} f^{(4)}(x) - \dots \\ -Q &= \eta f(x) - \frac{\eta^3}{3!} f^{(2)}(x) + \frac{\eta^5}{5!} f^{(4)}(x) - \dots \end{aligned} \quad (22.9)$$

The second line here corresponds to normalization of function ψ , specified in the first expression of (22.7). If series (22.9) converge (as in the case of shallow water), then from equation (22.8) for η it is seen that function $\eta(x)$ actually

depends only on the above-indicated parameters. It is easy to show, too, that in order to calculate the right-hand side of (22.8) with an accuracy of $O(\eta^4)$ it is sufficient to take only the first term of each of the series in (22.9). Then the right-hand side of (22.8) turns out to be $Q^2[1 - \frac{1}{3}(\partial\eta/\partial x)^2] + O(\eta^4)$, so that Equation (22.8) becomes, to this same degree of accuracy,

$$\frac{1}{3} Q^2 \left(\frac{\partial\eta}{\partial x} \right)^2 + U(\eta) = 0; \quad U(\eta) = g\eta^3 - 2R\eta^2 + 2S\eta - Q^2. \quad (22.10)$$

This expression is completely analogous to the equation of motion of a particle with a mass of $\frac{1}{3}Q^2$ and a coordinate η (with a 'time' x) in a force field with a potential $U(\eta)$. In this equation it is convenient to change to dimensionless variables, taking as the units of measurement the characteristics (denoted by the subscript 0) of a uniform flow (with a constant depth) having a specified discharge Q and a Froude number $F = c(gh)^{-1/2}$ equal to unity:

$$h_0 = g^{-1/3} Q^{2/3}; \quad c_0 = g^{1/3} Q^{2/3} = (gh_0)^{1/2};$$

$$R_0 = \frac{3}{2} g^{2/3} Q^{2/3} = \frac{3}{2} gh_0; \quad S_0 = \frac{3}{2} g^{1/3} Q^{4/3} = \frac{3}{2} gh_0^2. \quad (22.11)$$

Then, if all the lengths are measured to the scale h_0 , Equation (22.10) becomes

$$\frac{1}{3} \left(\frac{\partial\eta}{\partial x} \right)^2 + \eta^3 - 3r\eta^2 + 3s\eta - 1 = 0;$$

$$r = \frac{R}{R_0}; \quad s = \frac{S}{S_0}, \quad (22.12)$$

and its solution is governed by the two parameters r and s (so that it can be represented as a point on the (r, s) plane; a uniform flow is represented as a point $r = s = 1$). Here the potential can be put into the form $U(\eta) = (\eta - \eta_1)(\eta - \eta_2)(\eta - \eta_3)$, where $\eta_1 < \eta_2 < \eta_3$ are the roots, given by the formulas:

$$\eta_1 = r + 2(r^2 - s)^{1/2} \cos \left(\varphi + \frac{2\pi}{3} \right);$$

$$\eta_2 = r + 2(r^2 - s)^{1/2} \cos \left(\varphi - \frac{2\pi}{3} \right);$$

$$\eta_3 = r + 2(r^2 - s)^{1/2} \cos \varphi; \quad \cos 3\varphi = \frac{r^3 - \frac{3}{2}rs + \frac{1}{2}}{(r^2 - s)^{3/2}}, \quad (22.13)$$

with $0 \leq \varphi \leq \pi/3$. In the general case, when all these roots are different, the

solution of Equation (22.12) has the form

$$\eta = \eta_2 + (\eta_3 - \eta_2) \operatorname{cn}^2 \left[\frac{x}{2} \sqrt{3(\eta_3 - \eta_1)} \right];$$

$$\kappa = \left(\frac{\eta_3 - \eta_2}{\eta_3 - \eta_1} \right)^{1/2}, \quad (22.14)$$

where $\operatorname{cn} v$ is the Jacobi elliptic cosine with a modulus κ . According to this function, the wave (22.14) is called a *cnoidal wave*. Its propagation velocity $c = c_0 r$, the wavelength

$$\lambda = 4h_0[3(\eta_3 - \eta_1)]^{-1/2} K(\kappa),$$

the wave height is $\eta_3 - \eta_2$, and the mean water level is $\eta_1 - (\eta_3 - \eta_1) E/K$ (where E is a complete elliptic integral of the second kind with the same modulus). As $\varphi \rightarrow \pi/3$ we have $\eta_2 \rightarrow \eta_3$, $\kappa \rightarrow 0$, $\operatorname{cn} v \rightarrow \cos v$, and $K \rightarrow \pi/2$. These are waves of infinitesimal amplitude. The fourth relation (22.13) for these waves becomes $4(r^3 + s^3) = 3r^2s^2 + 6rs - 1$; it is plotted in the (r, s) plane of Figure 22.1 as the upper curve, having a parametric representation $r = \frac{1}{3}(F^{4/3} + 2F^{-2/3})$, $s = \frac{1}{3}(F^{-4/3} + 2F^{2/3})$ for Froude numbers $F < 1$ (the values of which are shown on the curve). As $\varphi \rightarrow 0$ we have $\eta_2 \rightarrow \eta_1$, $\kappa \rightarrow 1$, $\operatorname{cn} v \rightarrow \operatorname{ch}^{-1} v$, $K \rightarrow \infty$, and $\lambda \rightarrow \infty$, that is, at the limit we get the solitons

$$\eta = \eta_1 + (\eta_3 - \eta_1) \operatorname{ch}^{-2} \left[\frac{x}{2} \sqrt{3(\eta_3 - \eta_1)} \right]. \quad (22.15)$$

The fourth relation (22.13) for these is depicted in Figure 22.1 by the lower curve, having the same parametric representation as the upper curve, but for $F > 1$.

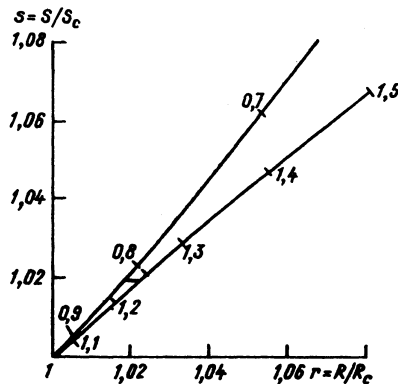


Fig. 22.1. Region of existence of gravity waves in the (r, s) plane.

If cnoidal waves (22.14) and solitons (22.15) are represented as steady-state solutions of some evolution equation, then with its help we can study the stability of these solutions with respect to small perturbations. In order to derive such an equation, for shallow water let us use the small parameter $\varepsilon = kh$, let us set $c_0 = (gh)^{1/2}$, let us introduce the slowly varying abscissa $\xi = \varepsilon^{1/2} (x - c_0 t)$ and the time $\tau = \varepsilon^{3/2} t$, and let us seek the velocity potential in the form $\varphi = \varepsilon^{1/2} \Phi(\xi, y, \tau)$ and the elevation of the water surface in the form $\eta = h + \varepsilon \zeta(\xi, \tau)$. The Laplace equation for φ then becomes $\varepsilon \Phi'' + \Phi_{yy} = 0$ (a prime indicates the derivative with respect to ξ), and for a boundary condition at the bottom $\Phi_y \rightarrow 0$ ($y \rightarrow 0$) it has the solution

$$\begin{aligned} \Phi &= f - \frac{\varepsilon y^2}{2!} f'' + \frac{\varepsilon^2 y^4}{4!} f'''' - \dots; \quad f = f_0 + \varepsilon f_1 + \dots; \\ (\Phi')_\eta &= f'_0 + \varepsilon \left(f'_1 - \frac{h^2}{2} f''_0 \right) + O(\varepsilon^2); \\ (\Phi_y)_\eta &= -\varepsilon h f''_0 + \varepsilon^2 \left(-\zeta_0 f''_0 - h f'_1 + \frac{h^3}{6} f'''_0 \right) + O(\varepsilon^3). \end{aligned} \quad (22.16)$$

The kinematic and dynamic boundary conditions at the water surface $y = \eta$ have the form

$$\begin{aligned} (\Phi_y)_\eta &= \varepsilon^2 \dot{\zeta} - \varepsilon c_0 \zeta' + \varepsilon^2 (\Phi')_\eta \zeta'; \\ \varepsilon (\dot{\Phi})_\eta - c_0 (\Phi')_\eta + \frac{1}{2} [\varepsilon (\Phi')_\eta^2 + (\Phi_y)_\eta^2] + g \zeta &= 0, \end{aligned} \quad (22.17)$$

where the dot indicates the derivative with respect to τ . Next, setting $\zeta = \zeta_0 + \varepsilon \zeta_1$... and using (22.16), from the terms of two lowest orders in these equations we obtain

$$\begin{aligned} -h f''_0 &= -c_0 \zeta'_0; \quad -\zeta_0 f''_0 - h f'_1 + \frac{h^3}{6} f'''_0 = \dot{\zeta}_0 - c_0 \zeta'_1 + f'_0 \zeta'_0; \\ -c_0 f'_0 + g \zeta_0 &= 0; \quad f'_0 - c_0 \left(f'_1 - \frac{h^2}{2} f''_0 \right) + \frac{1}{2} f'^2_0 + g \zeta_1 = 0. \end{aligned} \quad (22.18)$$

The first of these equations is the derivative with respect to ξ of the third, which takes the form $f'_0 = (g/c_0) \zeta_0$. Using this relation, after differentiating the fourth equation with respect to ξ and eliminating $g \zeta'_1 - c_0 f''_1$ from it with the aid of the third equation, we obtain for ζ_0 the KdV equation [named after Korteweg and de Vries]:

$$\dot{\zeta}_0 + \frac{3c_0}{2h} \zeta_0 \zeta'_0 + \frac{c_0 h^2}{6} \zeta'''_0 = 0. \quad (22.19)$$

If a *steady-state* solution of this equation is sought in the form $\zeta_0 = h^3(\eta - r)$, then integration of (22.19) first with respect to ξ and then with respect to η yields (22.12). If, on the other hand, we look for a nonsteady-state solution in the form $\zeta_0 = h^3(\eta_0 - r + \eta_*)$, where η_0 is a soliton (22.15) and η_* is a small perturbation, and if we linearize (22.19) with respect to η_* , then with the aid of this linearized equation we can show that the solitons are stable (Jeffrey and Kakutani, 1970). The stability of one-dimensional solitons with respect to small transverse perturbations is described with the aid of a two-dimensional generalization of the KdV equation, derived by Kadomtsev and Petviashvili (1970) in the form

$$\left(\xi + \frac{3c_0}{2h} \zeta \xi' + \frac{c_0 h^2}{6} \zeta'' \right)' + \frac{c_0}{2} \frac{\partial^2 \zeta}{\partial y^2} = 0. \quad (22.19')$$

In this respect the solitons also turn out to be stable. Equation (22.19), as well as (22.19'), is solved using the method (described below) of the inverse scattering problem (Zakharov and Shabat, 1974).

The KdV equation is by no means only useful for finding cnoidal waves and solitons (22.14) and (22.15): in addition, it describes the nonlinear waves originating in a number of diverse physical phenomena, including, as is evident from the foregoing, nonstationary nonlinear waves $\zeta(x, t)$ on the surface of shallow water for an arbitrary initial condition $\zeta(x, 0)$. This equation was applied especially widely after the remarkable discovery of Gardner, Green, Kruskal and Miura (GGKM, 1967) that the solution of the problem with initial data for the KdV equation

$$\dot{u} - 6uu' + u'' = 0 \quad (22.19'')$$

(where the standard normalization of all variables is assumed, a dot indicates the derivative with respect to t , and a prime the derivative with respect to x) reduces to a solution of some *linear* problem, known as the inverse scattering problem for the Schrödinger equation

$$L\psi = k^2\psi; \quad L = -\frac{\partial^2}{\partial x^2} + u(x, t) \quad (22.20)$$

for the conditions

$$\sum_{m=0}^4 \int_{-\infty}^{\infty} \left| \frac{\partial^m u(x, 0)}{\partial x^m} \right| dx < \infty; \\ \int_{-\infty}^{\infty} (1 + |x|) |u(x, 0)| dx < \infty. \quad (22.20')$$

First for $t = 0$ we solve the direct scattering problem: (1) find the discrete (simple) eigenvalues $k = i\kappa_n$, $\kappa_n > 0$, $n = 1, \dots, N$, of operator L (for which the Schrödinger equation has quadratically integrable solutions $\psi_n(x)$; in view of the

second equation of (22.20') there are only a finite number of these); (2) for the normalization $\int_{-\infty}^{\infty} \psi_n^2(x) dx$ find the normalization constants

$$b_n = \lim_{x \rightarrow \infty} e^{x_n x} \psi_n(x);$$

(3) for real eigenvalues $k \neq 0$ (forming a doubly degenerate continuous spectrum) find a solution $\psi(x, k)$ with asymptotes

$$\psi(x, k) \sim \begin{cases} e^{-ikx} & \text{for } x \rightarrow -\infty, \\ a(k)e^{-ikx} + b(k)e^{ikx} & \text{for } x \rightarrow +\infty, \end{cases} \quad (22.21)$$

where $1/a(k)$ and $r(k) = b(k)/a(k)$ are the coefficients of transmission and reflection for a wave e^{-ikx} , incident upon the potential $u(x, 0)$ from the right. Here $\psi_n(x) = \psi(x, i\kappa_n)$ and $a(i\kappa_n) = 0$. The quantities $s = \{r(k), \kappa_n, b_n, n = 1, \dots, N\}$ are called the *scattering data*. They can also be found according to $u(x, t)$, and if the potential satisfies the KdV equation (11.19"), then $a(k, t)$, $b(k, t)$, and $\psi(x, k, t)$ must satisfy the following GGKM equations:

$$\dot{a} = 0; \quad \dot{b} = 8ik^3 b; \quad (22.22)$$

$$\dot{\psi} = (-A + 4ik^3) \psi; \quad A = 4 \frac{\partial^3}{\partial x^3} - 3 \left(u \frac{\partial}{\partial x} + \frac{\partial}{\partial x} u \right). \quad (22.23)$$

These equations imply that

$$s(t) = \left\{ r(k, 0)e^{8ik^3 t}, \kappa_n, b_n e^{8\kappa_n^3 t}, n = 1, \dots, N \right\}.$$

On the other hand, if the scattering data vary with time, like the solution of the GGKM equations (22.22), then the potential $u(x, t)$ defined by them will satisfy the KdV equation (22.19"). Actually, it is equivalent to the equation of Lax (1968) for the operators

$$\dot{L} = [L, A], \quad (22.24)$$

where the square brackets denote a commutator ($\dot{L}$ is the operator of multiplication by $\dot{u}$, while $[L, A]$ is the operator of multiplication by $6uu' - u'''$). Finally, we still have to solve the inverse scattering problem: to reconstruct the potential according to the scattering data. To do this, let us consider the solution of the Schrödinger equation $\psi_1(x, k) \sim e^{-ikx}$ ($x \rightarrow +\infty$), and with the aid of a Green function let us write for $\chi_1(x, k) = \psi_1(x, k) e^{ikx}$ the integral equation

$$\begin{aligned} \chi_1(x, k) &= 1 - \int_x^\infty \frac{e^{2ik(x-x')}-1}{2ik} u(x') \chi_1(x', k) dx' = \\ &= 1 + \frac{1}{2ik} \int_x^\infty u(x') dx' + O(k^{-2}) \text{ for } \text{Im } k < 0, \quad |k| \rightarrow \infty. \end{aligned} \quad (22.25)$$

The Fourier transform of this function terminates, that is,

$$\begin{aligned}\chi_1(x, k) &= 1 + \int_0^\infty K(x, x+y) e^{-iky} dy = \\ &= 1 + \frac{1}{ik} K(x, x) + O(k^{-2}) \quad \text{for } |k| \rightarrow \infty,\end{aligned}\quad (22.25')$$

where $K(x, y)$ is the kernel of the linear operator transforming the solution e^{-ikx} of the Schrödinger equation with zero potential into the solution of the equation with potential $u(x)$. From a comparison of asymptotes (22.25) and (22.25'), we obtain $u(x) = -2(\partial/\partial x) K(x, x)$. In order to find the kernel, we multiply the equation

$$\psi(x, k) - a(k) e^{-ikx} = a(k) \psi_1(x, k) + b(k) \psi_1^*(x, k) - a(k) e^{-ikx}$$

by $e^{iky}/a(k)$ and integrate over all real k . Using the equation $\psi(x, i\kappa_n) = b_n \psi_1(x, -i\kappa_n)$ and (22.25'), we arrive at the following result:

$$\begin{aligned}K(x, y) + F(x+y) + \int_x^\infty K(x, z) F(y+z) dz &= 0; \\ F(x) &= \sum_{n=1}^N \frac{b_n e^{-\kappa_n x}}{ia' (i\kappa_n)} + \frac{1}{2\pi} \int_{-\infty}^\infty r(k) e^{ikx} dk.\end{aligned}\quad (22.26)$$

This equation was derived by Gel'fand *et al.* (1951). Thus, the KdV equation is solved according to the following scheme: $u(x, 0) \rightarrow s(0) \rightarrow s(t) \rightarrow u(x, t)$. An especially easy solution is obtained in the important particular case of reflectionless potentials $r(k) \equiv 0$:

$$\begin{aligned}u(x, t) &= -2 \frac{\partial^2}{\partial x^2} \ln \det A; \\ A_{mn} &= \delta_{mn} + \frac{\beta_n(0) e^{8\kappa_n^2 t - (\kappa_m + \kappa_n)t}}{\kappa_m + \kappa_n},\end{aligned}\quad (22.27)$$

where $A = \|A_{mn}\|$ is an $N \times N$ matrix, and $\beta_n(0)$ are arbitrary constants. This formula can be verified as follows: here in (22.26) $F(x) = \sum \beta_n e^{-\kappa_n x}$ and the kernel $K(x, y)$ can be sought in the form $\sum K_n(x) e^{-\kappa_n y}$; then for $\{K_n(x)\}$ we obtain linear algebraic equations with a matrix of coefficients A and a column of right-hand sides $\beta_n e^{-\kappa_n x}$, which yields $K_n = \det A_n / \det A$, where A_n is obtained from A by replacing the n th column by the right-hand sides. This gives

$$K(x, x) = \sum e^{-\kappa_n x} \frac{\det A_n}{\det A} = \frac{\partial}{\partial x} \ln \det A,$$

which yields (22.27). This solution constitutes a superposition of N solitons of type (22.15) with amplitudes $2\kappa_n^2$, velocities $4\kappa_n^2$, and arbitrary initial phases. An

analysis of the two-soliton solution shows that, if the first soliton is more rapid and overtakes the second, then its velocity and amplitude both decrease, while those of the second soliton increase. At the moment of closest approach the amplitudes and velocities will be equal, after which these quantities so increase for the second soliton and decrease for the first that they turn out to be asymptotically interchanged velocities and amplitudes (their phases being shifted by $\alpha\kappa_1^{-1}$ and $1 - \alpha\kappa_2^{-1}$, where $\alpha = \ln |(\kappa_1 + \kappa_2)/(\kappa_1 - \kappa_2)|$, so that solitons with similar velocities interact the most intensively; note that in the multisoliton solution the total phase shifts are exactly equal to the sums of the pairs of shifts, that is, multiparticle effects are absent).

The inverse scattering problem can be used to solve not only the KdV equation, but also several other nonlinear evolution equations, representable in the form of the Lax equation (22.24). These include the so-called *nonlinear Schrödinger equation*, describing, in particular, the peculiar instability exhibited by Stokes waves toward small perturbations producing a long-wave modulation of these waves. Here we will confine ourselves just to Stokes waves on deep water, which have a dispersion relation $\omega^2 = gk(1 + k^2a^2 + \dots)$, and we will examine a wave packet in which the bulk of the energy is concentrated around the wave number k_0 , so that in the packet, assuming $k = k_0 + k'$ and $\omega_0 = (gk_0)^{1/2}$, we have

$$\zeta(x, t) = \text{Re}[A(x, t) e^{i(k_0x - \omega_0t)}]; \quad A(x, t) = a(x, t) e^{i\theta(x, t)}; \quad (22.28)$$

$$\omega' = \omega - \omega_0 = \frac{\omega_0}{2k_0} k' - \frac{\omega_0}{8k_0^2} k'^2 + \frac{1}{2} \omega_0 k_0^2 a^2 + O(k'^2, a^2), \quad (22.28')$$

where $e^{i(k_0x - \omega_0t)}$ describes the *carrier* harmonic wave, and $A(x, t)$ is the complex *envelope* modulating it. Since in equation (22.28') the nonlinear (in a) term does not contain k' , the relations $-i\omega' \rightarrow (\partial/\partial t)$ and $ik' \rightarrow (\partial/\partial x)$ valid for linear waves remain in force for (22.28') as well, and thus operator (22.28') applied to $A(x, t)$ yields the nonlinear Schrödinger equation

$$i \left(\frac{\partial A}{\partial t} + \frac{\omega_0}{2k_0} \frac{\partial A}{\partial x} \right) - \frac{\omega_0}{8k_0^2} \frac{\partial^2 A}{\partial x^2} - \frac{1}{2} \omega_0 k_0^2 |A|^2 A = 0 \quad (22.29)$$

(Zakharov, 1968; he also showed that in the two-dimensional case on the left-hand side of this equation we have to add the term $(\omega_0/4k_0^2) (\partial^2 A/\partial y^2)$; a generalization for water of finite depth was given first by Binney and Roskis, 1969, and then by Davey and Stewartson, 1974). The terms in parentheses indicate that the envelope moves at the group velocity of the carrier wave $\omega_0/2k_0$. Solutions of this equation were constructed with the aid of the inverse problem of scattering by Zakharov and Shabat (1971). If as $|x| \rightarrow \infty$ the initial value $|A(x, 0)|$ decreases sufficiently rapidly, so that the quantity $\int |A(x, 0)| dx$ is finite (but not too small!), then *solitons of the envelope*

$$A = a \exp \left[-\frac{i}{4} \omega_0 k_0^2 a^2 t - \frac{4ik_0^2}{\omega_0} (x - ct + \varphi) \right] \text{ch}^{-1} [\sqrt{2} k_0^2 a (x - ct)], \quad (22.30)$$

will exist, where velocity c (relative to $\omega_0/2k_0$) does not depend on amplitude a (these solitons are determined by the real and imaginary parts of the discrete eigenvalues). The solitons of the envelope are stable. Let us consider, however, the stability of the periodic envelope $A = A_0(t) = a \exp(-(i/2) \omega_0 k_0^2 a^2 t)$, the frequency of which, accurate to $O(a^2)$, equals the Stokes correction to the frequency of the carrier wave, so that with the indicated accuracy we can analyze the stability of a Stokes wave. In a frame of reference moving with a velocity $\omega_0/2k_0$, let us consider a small perturbation in the form

$$A' = A_0 e^{-i\alpha\omega_0 t} (b_+ e^{i\beta k_0 x} + b_- e^{-i\beta k_0 x}).$$

After linearizing (22.29) with respect to it, we obtain for this perturbation the dispersion relation $\alpha^2 = (\beta^2/8) (\beta^2/8 - k_0^2 a^2)$, from which it is evident that for $0 < \beta < 2\sqrt{2} \cdot k_0 a$ the periodic envelope turns out to be unstable (imaginary α), with a maximum rate of growth $\alpha\omega_0 = \frac{1}{2}\omega_0 k^2 a^2$ (Benjamin and Fair, 1967). This phenomenon was called 'instability at the side frequencies'.

However, laboratory and numerical experiments have shown that, beyond the limits of application of the linearized equation, the growth of the modulating perturbation slows down, stops, and then becomes a decrease. Thus the wave returns to its initial, almost unmodulated, state, in which the perturbation again starts to grow; such oscillations of the modulation depth repeat periodically (provided energy losses do not occur, for instance, when waves break). This phenomenon is known as 'FPU recovery' in honor of Fermi, Pasta, and Ulam (1952), who discovered it during numerical modeling of vibrations of atomic lattices.

23. STATIONARY CAPILLARY WAVES

In the case of purely capillary waves, in boundary condition (21.23) at the water surface we must set $g = 0$, and in addition it is convenient to change from χ to the independent variable $\varphi = c\chi/k$ and from R to the dependent variable $\tau = -\ln R$. Then Equation (21.5) and boundary condition (21.53) become:

$$\frac{dW}{dw} = -c e^{\tau - i\theta}; \quad \frac{\partial\theta}{\partial\varphi} = -\frac{c}{\gamma} \operatorname{sh} \tau \quad \text{for } \psi = 0. \quad (23.1)$$

It is evident from (21.3) and (21.5) that the quantity $\tau - i\theta$ can be considered to be a function of the complex variable $W = \varphi + i\psi$, so that the Cauchy—Riemann equations $\partial\theta/\partial\varphi = \partial\tau/\partial\psi$ and $\partial\theta/\partial\psi = -\partial\tau/\partial\varphi$ are valid. Consequently, function τ (and also θ) must satisfy the Laplace equation $\partial^2\tau/\partial\varphi^2 + \partial^2\tau/\partial\psi^2 = 0$, while boundary condition (23.1) for $\psi = 0$ becomes $\partial\tau/\partial\psi = -(c/\gamma) \operatorname{sh} \tau$. Following Krepper (1957), let us seek this function as a solution of the equation with separated variables $\partial\tau/\partial\psi = f'(\psi) \operatorname{sh} \tau$, where we must set $f'(0) = -c/\gamma$. Thus, $\ln \operatorname{th} (\tau/2) \approx f(\psi) - h(\varphi)$, and then

$$\tau = \ln \left| \frac{F(\psi) + H(\varphi)}{F(\psi) - H(\varphi)} \right|; \quad F(\psi) = e^{f(\psi)}, \quad H(\varphi) = e^{h(\varphi)}. \quad (23.2)$$

Substitution of this function τ into the Laplace equation given above then yields

$$2(H'^2 + F'^2) + (F^2 - H^2) \left(\frac{H''}{H} - \frac{F''}{F} \right) = 0; \quad (23.3)$$

$$\frac{1}{FF'} \left(\frac{F''}{F} \right)' = - \frac{1}{HH'} \left(\frac{H''}{H} \right)', \quad (23.3')$$

and the second of these expressions (derived from the first by differentiating it with respect to ψ and with respect to φ) indicates that its right-hand side and left-hand side are equal to a constant. This in turn implies that F'^2 is a quadratic polynomial in F^2 , while H'^2 is a quadratic polynomial in H^2 (so that $F(\psi)$ and $H(\varphi)$ are elliptic Jacobi functions); from (23.3) it follows that the corresponding coefficients of these two polynomials differ only in sign. In other respects they are arbitrary, and one of them can be used as an arbitrary common multiplier of F and H (it cancels out everywhere), the second can be an arbitrary common multiplier α of φ and ψ , while the third can be an arbitrary modulus κ of the elliptic functions. Considering deep water only, when $\tau \rightarrow 0$ and $\psi \rightarrow \infty$, we obtain

$$\begin{aligned} \tau &= \ln \frac{\operatorname{cn} i\alpha (\psi - \psi_0) + \operatorname{cn} \alpha\varphi}{\operatorname{cn} i\alpha (\psi - \psi_0) - \operatorname{cn} \alpha\varphi} = \ln(TT^*); \\ T &= \frac{\operatorname{cn} W_1}{\operatorname{sn} W_1 \operatorname{dn} W_1}; \quad W_1 = \frac{\alpha}{2} (W - W_0); \quad W_0 = i\psi_0. \end{aligned} \quad (23.4)$$

The second equation here is obtained from the first by calculating $\operatorname{cn}(u - v) \pm \operatorname{cn}(u + v)$ with the aid of the formula

$$\operatorname{cn}(u \pm v) = (\operatorname{cn} u \operatorname{cn} v \pm \operatorname{sn} u \operatorname{sn} v \operatorname{dn} u \operatorname{dn} v) (1 - \kappa^2 \operatorname{sn}^2 u \operatorname{sn}^2 v)^{-1}. \quad (23.5)$$

Krepper considered solution (23.4) only in the particular case of a zero modulus $\kappa = 0$ (when $\operatorname{cn} u \rightarrow \cos u$); we will, however, assume the modulus to be arbitrary. Since $F(\psi) = e^{f(\psi)} = \operatorname{cn} i\alpha(\psi - \psi_0)$, we can use the formulas:

$$\frac{\partial \ln \operatorname{cn} u}{\partial u} = - \frac{\operatorname{sn} u \operatorname{dn} u}{\operatorname{cn} u}; \quad \frac{\operatorname{sn} iu \operatorname{dn} iu}{\operatorname{cn} iu} = i \frac{\operatorname{sn}' u \operatorname{dn}' u}{\operatorname{cn}' u}, \quad (23.5')$$

where the prime denotes elliptic functions with a modulus $\kappa' = \sqrt{1 - \kappa^2}$, to put the condition $f'(0) = -c/\gamma$ into the form

$$\frac{\operatorname{sn}' \alpha\psi_0 \operatorname{dn}' \alpha\psi_0}{\operatorname{cn}' \alpha\psi_0} = \frac{c}{\alpha\gamma}. \quad (23.6)$$

Moreover, using the equation $\partial\theta/\partial\varphi = \partial\tau/\partial\psi$, we obtain from (23.4) the relation $\theta = i \ln (T/T^*)$, and thus also $\tau - i\theta = \ln(T^2)$, so that the first equation

(23.1) becomes $dW/dw = -cT^2$, with $w = 0$ for $W = 0$. This gives

$$x + iz = -\frac{1}{c} \int_0^{\varphi + i\psi} T^{-2} dW = -\frac{2}{\alpha c} \int_{-i\alpha\psi_0/2}^{\alpha\varphi/2 + i\alpha(\psi - \psi_0)/2} T^{-2}(W_1) dW_1. \quad (23.7)$$

Here the integral is calculated using the formula

$$\int \frac{\operatorname{sn}^2 u \operatorname{dn}^2 u}{\operatorname{cn}^2 u} du = u + \frac{\operatorname{sn} u \operatorname{dn} u}{\operatorname{cn} u} - 2E(\operatorname{am} u), \quad (23.5'')$$

where $E(\varphi)$ is an elliptic integral of the second kind with the same modulus κ . The use of this formula in the general case, however, makes the separation of the real and imaginary parts on the right-hand side of (23.7) a very cumbersome operation, so that we will confine ourselves here to just two particular cases. In Krepper's case for $\kappa = 0$ we get $\operatorname{sn} u \rightarrow \sin u$, $\operatorname{cn} u \rightarrow \cos u$, $\operatorname{dn} u \rightarrow 1$, and $E(\operatorname{am} u) \rightarrow u$, so that for the water surface ($\psi = 0$, $z = \xi$) formulas (23.6) and (23.7) become:

$$\begin{aligned} \operatorname{th} \alpha\psi_0 &= \frac{c}{\alpha\gamma}; & \alpha c(x + i\xi) &= \alpha\varphi - 2 \left[\operatorname{tg} \left(\frac{\alpha\varphi}{2} - \frac{i\alpha\psi_0}{2} \right) + \right. \\ & & & \left. + \operatorname{tg} \left(\frac{i\alpha\psi_0}{2} \right) \right]. \end{aligned} \quad (23.8)$$

Now, introducing the wave number $k = \alpha c$ and the velocity of infinitesimal capillary waves $c_0 = (\gamma k)^{1/2}$, we obtain from (23.8) the formulas

$$\begin{aligned} kx &= \frac{k\varphi}{c} - \frac{2\Phi(1 - \Psi^2)}{1 + \Phi^2\Psi^2}; & k\xi &= \frac{2\Phi^2\Psi(1 - \Psi^2)}{1 + \Phi^2\Psi^2}; \\ \frac{2\Psi}{1 + \Psi^2} &= \frac{c^2}{c_0^2}; & \Phi &= \operatorname{tg} \frac{k\varphi}{2c}; & \Psi &= \frac{k\psi_0}{2c}. \end{aligned} \quad (23.9)$$

These formulas describe parametrically periodic capillary waves of a finite amplitude a , defined as the value of ξ for $k\varphi/c \rightarrow \pi$. Thus we obtain

$$a = \frac{2(1 - \Psi^2)}{k\Psi}; \quad \frac{c^2}{c_0^2} = \left(1 + \frac{k^2 a^2}{16} \right)^{-1/2}. \quad (23.10)$$

By changing the scale of measurement of all the lengths via the formulas $ka = A$, $kx/A = \xi$, $k\xi/A = \eta$ and eliminating parameters φ and ψ from (23.9), we

arrive at the following explicit equation for the wave profile:

$$\pm \xi = -\sqrt{\eta(1-\eta)} + \frac{2}{A} \operatorname{arctg} \left\{ \left[\frac{A}{4} + \left(1 + \frac{A^2}{16} \right)^{1/2} \right] \left(\frac{\eta}{1-\eta} \right)^{1/2} \right\}. \quad (23.11)$$

As shown by Krepper, for $A < A_{\max} \approx 4.58$ a value of $\xi = 0$ is obtained only for $\eta = 0$; the maximum $A = A_{\max}$ corresponds to a wave of limit configuration (with self-tangency of the profile in a trough, under which an air bubble can apparently form). Prelimit configurations are also obtained from (23.7) for $\psi > 0$, but these will not be discussed here.

In the opposite particular case of $\kappa = 1$ we obtain $\operatorname{sn} u \rightarrow \operatorname{th} u$; $\operatorname{cn} u \rightarrow \operatorname{ch}^{-1} u$; $\operatorname{dn} u \rightarrow \operatorname{ch}^{-1} u$; and $E(\operatorname{am} u) \rightarrow \operatorname{th} u$, so that the right-hand side of (23.5") becomes $u - \operatorname{th} u$. Any streamline $\psi = \psi_1 \geq 0$ can be taken as the water surface, having specified at it boundary condition (23.6), with ψ_0 in it replaced by $\psi_0 - \psi_1$; in this case formulas (23.6) and (23.7) become

$$\begin{aligned} \operatorname{tg} \alpha(\psi_0 - \psi_1) &= \frac{c}{\alpha\gamma}; \\ \alpha c(x + iz) &= -\alpha(\varphi + i\psi_1) + 2 \left\{ \operatorname{th} \left[\frac{\alpha\varphi}{2} + \frac{i\alpha(\psi_1 - \psi_0)}{2} \right] + \operatorname{th} \frac{i\alpha\psi_0}{2} \right\}. \end{aligned} \quad (23.8')$$

Using the same definitions of k and c_0 as above, we obtain

$$\begin{aligned} kx &= -\alpha\varphi + \frac{2\Phi_1(1 + \Psi_1^2)}{1 + \Phi_1^2\Psi_1^2}; & k(z - z_0) &= \frac{2\Phi_1^2\Psi_1(1 + \Psi_1^2)}{1 + \Phi_1^2\Psi_1^2}; \\ kz_0 &= -\alpha\psi_1 + 2(\Psi_0 - \Psi_1); & \frac{2\Psi_1}{1 - \Psi_1^2} &= \frac{c^2}{c_0^2}; \\ \Phi_1 &= \operatorname{th} \frac{\alpha\varphi}{2}; & \Psi_0 &= \operatorname{tg} \frac{\alpha\psi_0}{2}; & \Psi_1 &= \operatorname{tg} \frac{\alpha(\psi_0 - \psi_1)}{2} \end{aligned} \quad (23.9')$$

(if here we set $\psi_1 = 0$ only in the formula for c , then these formulas will be the equations for the streamlines $\psi = \psi_1 \geq 0$ under the water surface $\psi = 0$). The quantity z_0 in this case is the ordinate of the streamline in question for $\varphi = 0$ (and therefore for $x = 0$). Now, if we define the dimensionless amplitude $A = ka$ as the limit of $k(z - z_0)$ as $|x| \rightarrow \infty$ (that is, as $|\varphi| \rightarrow \infty$ or $|\Phi_1| \rightarrow 1$), then we get

$$A = 2\Psi_1; \quad \frac{c^2}{c_0^2} = A \left(1 - \frac{A^2}{4} \right)^{-1}, \quad (23.10')$$

from which it is evident that for $\Psi_1 > 0$ (when the wave is a single trough) we must have $A < 2$ (and as $A \rightarrow 2 - 0$ the velocity c of this wave increases without

limit), whereas for $\Psi_1 < 0$ (when the wave is a single crest) we must have $|A| > 2$ (and as $|A| \rightarrow 2 + 0$ the velocity c of the wave increases without limit, but as $|A| \rightarrow \infty$ it drops to zero according to the law $c \sim 2c_0 |A|^{-1/2}$). Using the first two formulas (23.9'), we can put the condition $x = 0$ into the following form:

$$2 \frac{\text{sh } \varphi_0}{\varphi_0} - \text{ch } \varphi_0 = \left(1 - \frac{A^2}{4}\right) \left(1 + \frac{A^2}{4}\right)^{-1};$$

$$\eta_0 = \frac{k(z - z_0)}{A} = \frac{\varphi_0}{2} \text{th } \frac{\varphi_0}{2}, \quad (23.12)$$

where $\varphi_0 = \alpha\varphi$. For $\varphi_0 = 0$ the left-hand side of the first of these equations goes to unity, and with an increase in $|\varphi_0|$ it decreases monotonically. The root $|\varphi_0(A)|$ of this equation goes to zero at $A = 0$, and with increasing $|A|$ it increases monotonically (reaching for $|A| = 2$ a value $|\varphi_0| \approx 1.9$). It corresponds to a point of self-intersection of the wave profile, the relative ordinate of which $\eta_0(A)$ for $A = 0$ will naturally be zero, while with increasing $|A|$ it increases monotonically (reaching at $|A| = 2$ value $\eta_0 \approx 0.7$, while as $|A| \rightarrow \infty$ it approaches a certain limit, at which the right-hand side of the first of Equations (23.12) goes to -1). As an illustration, let us examine the case $\psi_1 = 0$, in which, by eliminating parameters Φ_1 and Ψ_1 from (23.9') and using the variables $\xi = kx/A$ and $\eta = kz/A$, we are able to arrive at an explicit equation for the wave profile:

$$\pm \xi = \frac{2}{A} \eta^{1/2} \left[1 + \frac{A^2}{4} (1 - \eta) \right]^{1/2} +$$

$$+ \frac{1}{A} \ln \frac{1 - \eta^{1/2} \left[1 + \frac{A^2}{4} (1 - \eta) \right]^{-1/2}}{1 + \eta^{1/2} \left[1 + \frac{A^2}{4} (1 - \eta) \right]^{-1/2}}. \quad (23.11')$$

Figure 23.1 shows plots of this function for various values of A . Only the branch for $\eta \geq \eta_0(A)$ has any real physical meaning; on this branch the velocity potential φ has a discontinuity at the point $\xi = 0$ (so that in the family of solutions being considered there are no solitary waves with a velocity potential varying continuously everywhere along the streamlines).

Numerous studies have been made of stationary gravity-capillary waves, and these include both analytical studies, with the aid of the first terms of certain series (Wilton, 1915, was one of the first to do this), and numerical studies. In this case the most general boundary condition (21.24) must be imposed at the water surface. Confining ourselves only to the case of deep water and introducing into the indicated boundary condition the parameter η , defined as the normalized difference between the squares of the propagation velocities of waves of finite and infinitesimal amplitude (the latter is defined as ω/k , where ω is given by formula

(18.8) as $h \rightarrow \infty$, and the plot of its dependence on k is given in Figure 18.1), we obtain:

$$c^2 - \frac{1}{2} c_m^2 \left(p + \frac{1}{p} \right) = \frac{1}{2} c_m^2 \eta; \quad p = \frac{k_m}{k}. \quad (23.13)$$

Parameter η is introduced into Equation (21.24) by replacing c^2 in the first term on the left-hand side by $\gamma k (\eta p + p^2 + 1)$. We note at this point that two waves (see Figure 18.1), whose lengths differ by a factor of p^2 , correspond to each fixed value of the propagation velocity of small-amplitude waves. In the general case this number is not an integer, and then the formulas describing a gravity-capillary wave can be constructed by representing function $\Omega(w_1)$ in (21.5) and parameter η as series: $\varepsilon \Omega_1 + \varepsilon^2 \Omega_2 + \varepsilon^3 \Omega_3 + \dots$ and $\varepsilon \eta_1 + \varepsilon^2 \eta_2 + \varepsilon^3 \eta_3 + \dots$. The coefficients of these series are determined successively from the requirement of holomorphicity of functions $\Omega_n(w_1)$ for $|w_1| \leq 1$ and the boundary conditions obtained from (21.24). Leaving out the details of the derivation, let us just write the result:

$$\begin{aligned} \Omega(w_1) = & \varepsilon w_1 + \frac{3}{2} \frac{p^2}{p^2 - 2} \varepsilon^2 w_1^2 + \\ & + \frac{1}{2} \frac{\frac{17}{3} p^4 - \frac{41}{24} p^2 + \frac{1}{4}}{(p^2 - 2)(p^2 - 3)} \varepsilon^3 w_1^3 + \dots, \end{aligned} \quad (23.14)$$

where $\eta_1 = 0$ and $\eta_2 = (8p^4 + p^2 + 2)/8p(p^2 - 2)$.

Now, taking (21.5) into account and integrating (21.3), we get

$$\begin{aligned} -\frac{i}{k} w = & \ln w_1 + \varepsilon(w_1 - 1) + \frac{\varepsilon^2}{2} \frac{2p^2 - 1}{p^2 - 2} (w_1^2 - 1) + \\ & + \frac{3\varepsilon^3}{4} \frac{p^4 - \frac{11}{8} p^2 + \frac{1}{4}}{(p^2 - 2)(p^2 - 3)} (w_1^3 - 1) + \dots, \end{aligned} \quad (23.15)$$

from which for $w_1 = \exp(i\chi)$ we arrive at a parametric representation of the wave surface, while the wave velocity is found from (23.13) for $\eta \approx \varepsilon^2 \eta_2$. In cases where $p^2 = 2, 3, 4, \dots$, of the two above-mentioned waves, the longer one will contain p^2 of the shorter ('Wilton ripples'). In these instances series (23.15) are not valid, and the solutions for holomorphic functions $\Omega_n(w_1)$ and numbers η_n must be selected otherwise (Pearson and Fife, 1961), but we will not discuss this here.

Like Stokes waves, periodic gravity-capillary waves of finite amplitude are modulationally unstable: with the aid of a nonlinear Schrödinger equation, analogous to (22.29), for the complex envelope, we see that for deep water modulation stability occurs only in a narrow interval of wave numbers $(2/\sqrt{3} -$

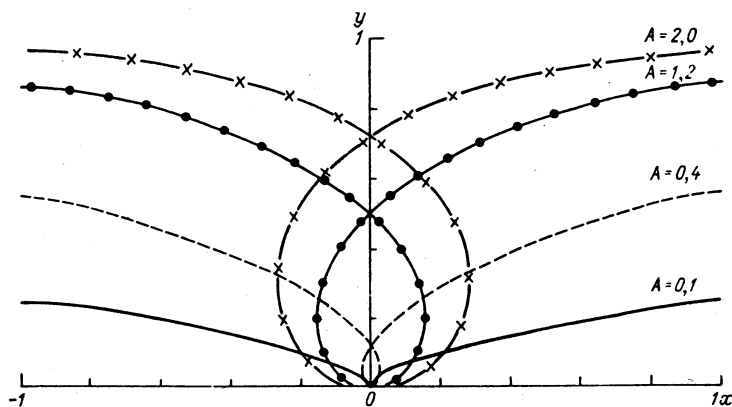


Fig. 23.1. Normalized profiles (23.11) of capillary solitons on deep water for various dimensionless amplitudes A .

$1)^{1/2} < k/k_m < 2^{-1/2}$; however, the instability gives rise to 'FPU recovery'. The case of shallow water for capillary effects under natural conditions is not very interesting. All the same, it should be noted that in this case it is easy to find gravity-capillary soliton solutions of the (analogous to (22.19')) Kadomtsev–Petviashvili equation:

$$\left(\xi + \frac{3c_0}{2h} \xi \xi' - q \xi''' \right)' + \frac{c_0}{2} \frac{\partial^2 \xi}{\partial y^2} = 0; \quad q = \frac{c_0 h^2}{6} - \frac{\gamma h}{2c_0} \quad (23.16)$$

(Krasitskii, 1982), which for $q > 0$ has stable one-dimensional soliton solutions ('oblique solitons'):

$$\begin{aligned} \xi &= \xi_m \operatorname{ch}^{-2} \left[\left(\frac{c_0 \xi_m}{2qh} \right)^{1/2} (x - \kappa y - Ut) \right]; \\ U &= c_0 \left(1 + \frac{\kappa^2}{2} + \frac{\xi_m}{2h} \right). \end{aligned} \quad (23.17)$$

For $q < 0$ one-dimensional solitons are unstable, but there exist stable two-dimensional, so-called *rational*, solitons

$$\xi = -\xi_m \left(\frac{h}{a} \right)^2 \frac{\left(\frac{h}{a} \right)^2 - (x' + by')^2 + (ay')^2}{\left[\left(\frac{h}{a} \right)^2 + (x' + by')^2 + (ay')^2 \right]^2}, \quad (23.18)$$

where x', y' are certain normalized (differently) Galilean coordinates, namely

$$x' = -\gamma^{1/6} h^{1/2} |q|^{-1/3} (x - Ut) \quad \text{and} \quad y' = 6^{1/2} \gamma^{1/12} c_0^{-1} h^{-1/4} (y - Vt),$$

in which $U = c_0 (1 - 3\xi_m/16h - V^2/2c_0^2)$. During collisions such solitons do not interact at all (not even phase changes occur).

24. HAMILTONIAN FORMALISM

For a theoretical description of *nonstationary* free waves, we can use Hamilton equations (18.5), derived by Zakharov (1968). In the following, it will be convenient to use in these equations the Fourier transforms of the generalized coordinate Φ and the generalized momentum ξ with respect to the horizontal coordinates:

$$\Phi_k(t) = \frac{1}{2\pi} \int \Phi(x, t) e^{-ik \cdot x} dx; \quad \xi_k(t) = \frac{1}{2\pi} \int \xi(x, t) e^{-ik \cdot x} dx. \quad (24.1)$$

Considering only the case of deep water, we can represent the solution of the Laplace equation for the velocity potential as

$$\begin{aligned} \varphi_k(z) = e^{kz} \left[\Phi_k + \int \Phi_{k_1} \xi_{k_2} k_1 \delta_{k-k_1-k_2} dk_1 dk_2 - \frac{1}{2} \int \Phi_{k_1} \xi_{k_2} \xi_{k_3} k_1 \times \right. \\ \left. \times (|k-k_3| + |k-k_2| - k) \delta_{k-k_1-k_2-k_3} dk_1 dk_2 dk_3 + \dots \right]. \end{aligned} \quad (24.2)$$

In the Hamilton equations it is advisable to perform a *canonical* transformation to *normal* variables: the complex wave amplitudes ia_k^* , a_k ; these are classical analogs of the quantum operators of creation and annihilation of particles:

$$\xi_k = \left(\frac{k}{2\omega_k} \right)^{1/2} (a_k + a_{-k}^*); \quad \Phi_k = -i \left(\frac{\omega_k}{2k} \right)^{1/2} (a_k - a_{-k}^*), \quad (24.3)$$

where ω_k is the frequency of an elementary wave, defined in the general case by formula (18.8). Then the Hamilton equations take the form:

$$\begin{aligned} \frac{\partial a_k}{\partial t} = -i \frac{\delta \mathcal{H}}{\delta a_k^*}; \quad \mathcal{H} = \mathcal{H}_2 + \mathcal{H}_3 + \mathcal{H}_4 + \dots; \quad \mathcal{H}_2 = \int \omega_k a_k a_k^* dk; \\ \mathcal{H}_3 = \int V_{kk_1k_2} (a_k^* a_{k_1} a_{k_2} + a_k a_{k_1}^* a_{k_2}^*) \delta_{k-k_1-k_2} dk dk_1 dk_2 + \\ + \frac{1}{3} \int U_{kk_1k_2} (a_k^* a_{k_1}^* a_{k_2}^* + a_k a_{k_1} a_{k_2}) \delta_{k+k_1+k_2} dk dk_1 dk_2; \\ \mathcal{H}_4 = \frac{1}{2} \int W_{kk_1k_2k_3} a_k^* a_{k_1}^* a_{k_2}^* a_{k_3} \delta_{k+k_1-k_2-k_3} dk dk_1 dk_2 dk_3 + \dots, \end{aligned} \quad (24.4)$$

where only the most important term of $\mathcal{H}_4$ is given. The coefficients of three-wave

interactions in the case of gravity waves on deep water are homogeneous functions of degree 7/4:

$$8\pi V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} = (\mathbf{k} \cdot \mathbf{k}_1 - k k_1) \left(\frac{g k_2}{2 k k_1} \right)^{1/4} + (\mathbf{k} \cdot \mathbf{k}_2 - k k_2) \left(\frac{g k_1}{2 k k_2} \right)^{1/4} + (\mathbf{k}_1 \cdot \mathbf{k}_2 + k_1 k_2) \left(\frac{g k}{2 k_1 k_2} \right)^{1/4}, \quad (24.5)$$

and $U_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}$ is obtained from this by changing the two minuses to pluses. The coefficients of four-wave interactions $W_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}$ are homogeneous functions of the third degree. If we retain only one Hamiltonian $\mathcal{H}_2$, then for $a_{\mathbf{k}}$ we obtain the linear equation $\dot{a}_{\mathbf{k}} + i\omega_{\mathbf{k}} a_{\mathbf{k}} = 0$, that is, it agrees with the linear wave theory of Section 18. In the case of gravity waves on deep water, the three-wave interactions described by Hamiltonian $\mathcal{H}_3$ may not be decay interactions (that is, they may not satisfy the resonance conditions $\omega_{\mathbf{k}} = \omega_{\mathbf{k}_1} + \omega_{\mathbf{k}_2}$ for $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$, for which either a wave $\mathbf{k}$ decays into waves $\mathbf{k}_1$ and $\mathbf{k}_2$ or else the opposite: rather, they merge to form wave $\mathbf{k}$), and thus they may be eliminated by a canonical (that is, retaining the form of the Hamilton equations (24.4)) representation of the normal variables $a_{\mathbf{k}} \rightarrow b_{\mathbf{k}}$, the principal terms of which for this must have the form

$$a_{\mathbf{k}} = b_{\mathbf{k}} + \int \frac{V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} b_{\mathbf{k}_1} b_{\mathbf{k}_2}}{\omega_{\mathbf{k}} - \omega_{\mathbf{k}_1} - \omega_{\mathbf{k}_2}} \delta_{\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2} d\mathbf{k}_1 d\mathbf{k}_2 + \int \frac{V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} b_{\mathbf{k}_1} b_{\mathbf{k}_2}^*}{\omega_{\mathbf{k}_1} - \omega_{\mathbf{k}} - \omega_{\mathbf{k}_2}} \times \\ \times \delta_{\mathbf{k}_1 - \mathbf{k} - \mathbf{k}_2} d\mathbf{k}_1 d\mathbf{k}_2 + \frac{1}{3} \int \frac{U_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} b_{\mathbf{k}_1}^* b_{\mathbf{k}_2}^*}{\omega_{\mathbf{k}} + \omega_{\mathbf{k}_1} + \omega_{\mathbf{k}_2}} \delta_{\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2} d\mathbf{k}_1 d\mathbf{k}_2. \quad (24.6)$$

In terms of variables $b_{\mathbf{k}}$, the Hamiltonian has the form $\mathcal{H} = \bar{\mathcal{H}}_2 + \bar{\mathcal{H}}_4 + \dots$, where now

$$\bar{\mathcal{H}}_4 = \frac{1}{2} \int \bar{W}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} b_{\mathbf{k}}^* b_{\mathbf{k}_1}^* b_{\mathbf{k}_2} b_{\mathbf{k}_3} \delta_{\mathbf{k} + \mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3} d\mathbf{k} d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3, \quad (24.7)$$

and $\bar{W}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}$ is a new homogeneous function of third degree (calculated by Zakharov, 1968), symmetric with respect to $\mathbf{k}_1$ and $\mathbf{k}$, with respect to $\mathbf{k}_3$ and $\mathbf{k}_2$, and with respect to the pairs $(\mathbf{k}, \mathbf{k}_1)$ and $(\mathbf{k}_2, \mathbf{k}_3)$. The case in which in the Hamiltonian it is sufficient to retain only the first two terms $\bar{\mathcal{H}}_2 + \bar{\mathcal{H}}_4$, terms of higher order in $b_{\mathbf{k}}$ being dropped, is called the *approximation of weak nonlinearity*. In this approximation the Hamiltonian is invariant with respect to the gauge transformation $b_{\mathbf{k}} \rightarrow b_{\mathbf{k}} e^{i\alpha}$, as a result of which the corresponding equation of motion

$$\frac{\partial b_{\mathbf{k}}}{\partial t} + i\omega_{\mathbf{k}} b_{\mathbf{k}} = -i \int \bar{W}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} b_{\mathbf{k}_1}^* b_{\mathbf{k}_2} b_{\mathbf{k}_3} \delta_{\mathbf{k} + \mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3} d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3 \quad (24.8)$$

conserves, in addition to the total momentum and total energy, the so-called *integral of wave action*

$$n = \int |b_k|^2 dk \quad (24.9)$$

(we stress, however, that it can be considered conserving only in the approximation of weak nonlinearity). Let us introduce the space *spectrum of wave action* $N_k(t)$ via the relation

$$\langle a_k(t) a_{k_1}^*(t) \rangle = N_k(t) \delta_{k-k_1}; \quad N_k = \frac{F_k}{\omega_k}, \quad (24.10)$$

where the angle brackets denote the statistical mean, and F_k is the space spectrum of the wave field ξ , so that the space-time spectrum of the waves is written as

$$\begin{aligned} \langle \xi_{k\omega} \xi_{k_1\omega_1}^* \rangle &= F_{k\omega} \delta_{k-k_1} \delta_{\omega-\omega_1}; \\ F_{k\omega} &= \frac{1}{2} (F_k \delta_{\omega-\omega_k} + F_{-k} \delta_{\omega+\omega_k}), \end{aligned} \quad (24.11)$$

(if the waves are not free, then their space-time spectrum will be concentrated not exactly at the dispersion surface, but rather *close* to this surface, that is, the δ functions $\delta_{\omega \pm \omega_k}$ should be assumed to be somewhat blurred). In the following we will approximately equate (24.10) to the correlation $\langle b_k b_{k_1}^* \rangle$ (in spite of the slight difference ensuing from (24.6)). Then, after expressing b_k from (24.8) in terms of the right-hand side with the aid of the Green function, we see that the rate of evolution of the wave-action spectrum $\partial N_k / \partial t$, in view of the wave interactions can be expressed in terms of the weighted integral of the sixfold correlation of the amplitudes b_k . If this evolution rate is expressed in terms of pairwise correlations according to a formula valid for a multidimensional Gaussian probability distribution, then finally we arrive at the so-called *kinetic equation*

$$\begin{aligned} \frac{\partial N_k}{\partial t} = I_k &= 2\pi \int |\bar{W}_{kk_1k_2k_3}|^2 (N_{k_1}N_{k_2}N_{k_3} + N_kN_{k_2}N_{k_3} - N_kN_{k_1}N_{k_2} - \\ &- N_kN_{k_1}N_{k_3}) \delta_{k+k_1-k_2-k_3} \delta_{\omega_k+\omega_{k_1}-\omega_{k_2}-\omega_{k_3}} dk_1 dk_2 dk_3 \end{aligned} \quad (24.12)$$

(Hasselmann, 1962, 1963; Zakharov and Filonenko, 1966; Zakharov and L'vov, 1975). Integral I_k can be written as $A_k - B_k N_k$, where A_k is defined by the first term in parentheses and $B_k N_k$ by the other three terms.

Now we can define the blurred δ function in relation (24.11) using the formula

$$\begin{aligned}\delta_{\omega-\omega_k} &= \pi B_k [(\omega - \omega_k - \delta\omega_k)^2 + B_k^2]^{-1}, \\ \delta\omega_k &= 2 \int W_{kk,kk_1} N_{k_1} d\mathbf{k}_1,\end{aligned}\quad (24.13)$$

where $\delta\omega_k$ is a nonlinear correction to the dispersion law. Integral I_k satisfies the identities

$$\int \mathbf{k} I_k d\mathbf{k} = \int \omega_k I_k d\mathbf{k} = \int I_k d\mathbf{k} = 0, \quad (24.14)$$

which imply that the kinetic equation conserves the momentum of the wave field $\mathbf{m} = \int \mathbf{k} N_k d\mathbf{k}$, its energy $e = \int \omega_k N_k d\mathbf{k}$, and the wave action $n = \int N_k d\mathbf{k}$. Outside the intervals of supply and dissipation of energy, the wave spectra can be determined only by the spectral fluxes of these conserved quantities $\mathbf{q} = \dot{\mathbf{m}}$, $\varepsilon_e = \dot{e}$, and $\varepsilon_n = \dot{n}$. Such spectra, known as *Kolmogorov spectra*, can be found as steady-state solutions of the kinetic equation (24.12), that is, in the isotropic case (when $\mathbf{m} = 0$, and thus $\mathbf{q} = 0$ as well) as the solutions N_k of the equation $I_k = 0$. Changing in it to the new independent variables $\omega = \omega_k$, we can put this equation into the form

$$\begin{aligned}\int S_{\omega\omega_3\omega_1\omega_2} (N_{\omega_1} N_{\omega_2} N_{\omega_3} + N_{\omega} N_{\omega_1} N_{\omega_2} - \\ - N_{\omega} N_{\omega_1} N_{\omega_3} - N_{\omega} N_{\omega_2} N_{\omega_3}) d\omega_1 d\omega_2 = 0,\end{aligned}\quad (24.15)$$

where $\omega_3 = \omega_1 + \omega_2 - \omega$, and $S_{\omega\omega_3\omega_1\omega_2}$ is the result of averaging over the angles of the expression

$$|\bar{W}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}|^2 \delta_{\mathbf{k}+\mathbf{k}_1-\mathbf{k}_2-\mathbf{k}_3} k_1 k_2 k_3 \frac{\partial k_1}{\partial \omega_1} \frac{\partial k_2}{\partial \omega_2} \frac{\partial k_3}{\partial \omega_3}, \quad (24.15')$$

the integration in (24.15) being carried out in the $\omega_1\omega_2$ plane over a symmetrical region with asymptotes $\omega_1 = \omega$ and $\omega_2 = \omega$. Function $S_{\omega\omega_3\omega_1\omega_2}$ is also symmetrical with respect to the indices, like $\bar{W}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}$, and it is a homogeneous function of 20th degree, so that Equation (24.15) has precise power-law solutions $N_{\omega} \sim \omega^{-a}$, for which it reduces to

$$\begin{aligned}\int \frac{S_{\omega\omega_3\omega_1\omega_2}}{(\omega_1\omega_2\omega_3)^a} \left[1 + \left(\frac{\omega_3}{\omega} \right)^a - \left(\frac{\omega_1}{\omega} \right)^a - \left(\frac{\omega_2}{\omega} \right)^a \right] \times \\ \times \left[1 + \left(\frac{\omega_3}{\omega} \right)^{3a-23} - \left(\frac{\omega_1}{\omega} \right)^{3a-23} - \left(\frac{\omega_2}{\omega} \right)^{3a-23} \right] d\omega_1 d\omega_2 = 0,\end{aligned}\quad (24.16)$$

the integration in this case being over the subregion in which $\omega_1 < \omega$, and $\omega_2 < \omega$ (Zakharov and Filonenko, 1966). The integrand here goes to zero for $\alpha = 0, 1, \frac{23}{3}$, and 8 (and is of constant sign for other α). The first two of these solutions (thermodynamically stable) do not have any physical meaning, since for them the integrals

$$n = \int_0^\infty I(\omega) d\omega \quad \text{and} \quad e = \int_0^\infty \omega I(\omega) d\omega = \int_0^\infty S(\omega) d\omega$$

(where

$$I(\omega) = N_\omega \left(k \frac{\partial k}{\partial \omega} \right)_{k=\omega^2/g} = 2\omega^3 N_\omega g^{-1},$$

and $S(\omega)$ is the usual frequency spectrum of the waves) diverge at $\omega = \infty$. The other two solutions correspond to the Kolmogorov spectra:

$$S(\omega) \sim g^{4/3} \varepsilon_n^{1/3} \omega^{-11/3}; \quad F_k \sim \varepsilon_n^{1/3} k^{-10/3}; \quad (24.17)$$

$$S(\omega) \sim g^{4/3} \varepsilon_e^{1/3} \omega^{-4}; \quad F_k \sim g^{-1/6} \varepsilon_e^{1/3} k^{-7/2}. \quad (24.18)$$

The first of these models the action flux from the high-frequency region of energy supply to the low-frequency region, and the second models the energy flux from the low-frequency supply region to the high-frequency region. If supply with an increment β is concentrated in a region $\omega \sim \omega_0$ with a width of the same order, then by equating it to I_k we get $S(\omega) \sim (\beta\omega_0)^{1/2} |\bar{W}|^{-1}$. If for $\omega < \omega_0$ energy supply and dissipation are absent, then spectrum (24.17) becomes established there, its lower limit being determined by the effects of nonstationarity, nonhomogeneity, or low-frequency dissipation. If the high-frequency dissipation is located in a region $\omega \sim \omega_1 \gg \omega_0$ (which, of course, is hardly ever realized in practice, since I_k is not usually able to compensate the supply, and the excess energy is released in the breaking of waves), then for $\omega_0 < \omega < \omega_1$ spectrum (24.18) is established.

For gravity waves on water of finite depth, and also for any gravity-capillary waves or purely capillary waves, decay three-wave interactions are possible; these interactions satisfy the resonance conditions $\omega_k = \omega_{k_1} + \omega_{k_2}$ for $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$, and in (24.4) the main role is assumed by the three-wave Hamiltonian $\mathcal{H}_3$. The kinetic equation in these cases has the form

$$\begin{aligned} \frac{\partial N_k}{\partial t} = I_k = & \int [T_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} (N_{\mathbf{k}_1} N_{\mathbf{k}_2} - N_{\mathbf{k}} N_{\mathbf{k}_1} - N_{\mathbf{k}} N_{\mathbf{k}_2}) \delta_{\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2} \times \\ & \times \delta_{\omega_k - \omega_{k_1} - \omega_{k_2}} + 2T_{\mathbf{k}_1\mathbf{k}\mathbf{k}_2} (N_{\mathbf{k}_1} N_{\mathbf{k}_2} + N_{\mathbf{k}} N_{\mathbf{k}_1} - N_{\mathbf{k}} N_{\mathbf{k}_2}) \times \\ & \times \delta_{\mathbf{k}-\mathbf{k}_1+\mathbf{k}_2} \delta_{\omega_k - \omega_{k_1} + \omega_{k_2}}] d\mathbf{k}_1 d\mathbf{k}_2. \end{aligned} \quad (24.12')$$

It conserves momentum and energy, but not wave action. Thus in the isotropic

case there will be only one Kolmogorov spectrum. For pure capillary waves it becomes

$$S(\omega) \sim \gamma^{1/3} \varepsilon_e^{1/2} \omega^{-13/6}, \quad F_k \sim \gamma^{-1/4} \varepsilon_e^{1/2} k^{-15/4} \quad (24.18')$$

and it models the energy flux in the high-frequency region.

In the case of a weakly horizontally nonuniform and nonstationary wave field, described by the ray theory of Section 19, taking into account the supply rate P_k^+ and dissipation rate P_k^- , the following generalization of the kinetic equation will be valid:

$$\frac{\partial N_k}{\partial t} + \frac{\partial \sigma}{\partial k_a} \frac{\partial N_k}{\partial x_a} - \frac{\partial \sigma}{\partial x_a} \frac{\partial N_k}{\partial k_a} = I_k + P_k^+ - P_k^-, \quad (24.19)$$

where $\sigma = \sigma(\mathbf{k}, \mathbf{x}, t)$ is defined by formula (19.3). The refraction terms on the left-hand side will be important, for instance, for capillary waves, which under actual oceanic conditions are always present on larger gravity waves, being concentrated on their windward sides.

25. WIND WAVES

In nature, waves on a water surface originate and develop under the influence of a turbulent wind. These waves in turn introduce perturbations into the fields of the wind speed and pressure in the surface layer of air. These fields can therefore be represented as $\mathbf{u}_a = \mathbf{U} + \mathbf{u}' + \mathbf{u}''$ and $p_a = P + p' + p''$, where $\mathbf{U} = \mathbf{U}(z)$ and $P = P(z)$ are the mean (horizontally uniform and quasistationary) fields of the wind speed and atmospheric pressure in the absence of waves, $\mathbf{u}'$ and p' are the turbulent fluctuations of these fields in the absence of waves, and $\mathbf{u}''$ and p'' are the wave-induced perturbations.

During the initial period of wave development from a state of rest, we will describe all the perturbations in the linear approximation (even though it will probably rapidly become unsuitable for the description of capillary ripples). Therefore, all the boundary conditions at the water surface can be imposed for $z = 0$, and, in particular, turbulent fluctuations of the atmospheric pressure p' and of the tangential wind stresses $\tau'_i = \rho_a \nu_a (\partial u'_i / \partial z)$ can be considered as specified fields over a plane interface. Here, instead of (24.3), it will be more convenient to use the (differently normalized) normal variables:

$$a_k^+ = \omega_k^{-1} \dot{\xi}_k - i \xi_k; \quad a_k^- = \omega_k^{-1} \dot{\xi}_k + i \xi_k. \quad (25.1)$$

During the initial period of wave development, when it is still possible to ignore the wave-caused perturbations of the atmosphere, the linear theory yields the equation

$$\dot{a}_k^+ + i \omega_k a_k^+ = - \frac{\omega_k}{\rho g \Delta_k} \left(p_k' + i \frac{k_a}{k} \tau'_{ak} - 2 \tau'_{3k} \right) - \gamma_k (a_k^+ + a_k^-), \quad (25.2)$$

where $\Delta_k = 1 + \gamma k^2/g$, and γ_k is the damping constant of the waves, related to the eddy viscosity of water and air (which has to be taken into account in order to satisfy the boundary conditions at the water surface); this form of the equation of the linear theory was introduced by Krasitskii (1980). From (25.2) we obtain

$$\frac{\partial F_k}{\partial t} = \alpha_k - 2\gamma_k F_k; \quad \alpha_k = \frac{\pi \omega_k^2}{\rho^2 g^2 \Delta_k^2} \Pi_{k\omega_k}, \quad (25.3)$$

where $\Pi_{k\omega_k}$ is the space-time spectrum of the force exerted by the atmosphere on the water surface (this force is represented by the expression in the first parentheses on the right-hand side of (25.2)). If we neglect viscous damping, then from (25.3) we get the linear increase of the wave spectrum with time $F_k = \alpha_k t$. Since spectrum $\Pi_{k\omega_k}$ is here taken at the frequency of free waves ω_k , the generation of wind waves primarily turns out to be a resonance reaction of free waves to fluctuating atmospheric effects (these were first described by Phillips, 1957, who took into account, of all the effects, only p').

With a further growth of the waves, the wave perturbations in the water-adjacent air layer become significant, and they turn out to be hydrodynamically unstable. Attempts can be made to describe their instability with the aid of linearized *quasilaminar* equations (without turbulence and viscosity), for the boundary conditions $p'_\zeta = p - \gamma \rho \Delta \zeta$ and $u'_z = \partial \zeta / \partial t + U_a (\partial \zeta / \partial x_a)$ with $z = 0$ (such a theory was first devised by Miles, 1957). Here instead of (25.2) we obtain the following equation (reduced to this form by Krasitskii and Zaslavskii, 1978):

$$\begin{aligned} \dot{a}_k^+ + i\omega_k a_k^+ &= \frac{1}{2} (\beta_k^+ a_k^+ - \beta_k^- a_k^-); \\ \beta_k^s &= -\frac{\pi \rho_a \omega_k}{\rho g \Delta_k} \left[\frac{U''}{U'} |w_k^s|^2 \right] z_k^s \operatorname{sign} \cos \theta_k, \end{aligned} \quad (25.4)$$

where θ_k is the angle between the directions of $\mathbf{U}$ and $\mathbf{k}$; $s = \pm$; the expression for β_k^s is valid for $s \cos \theta > 0$ (and for $s \cos \theta \leq 0$ it is zero), and $w_k^s(z)$ is the solution of the Rayleigh equation (11.11)

$$(U_k - c_k^s) \left(\frac{\partial^2}{\partial z^2} - k^2 \right) w_k^s - U_k'' w_k^s = 0; \quad (25.4')$$

$$w_k^s = U_k - c_k^s \quad \text{for } z = 0; \quad w_k^s \rightarrow 0 \quad \text{for } z \rightarrow \infty,$$

while

$$U_k = U \cos \theta_k; \quad c_k^s = sk^{-1} \omega_k + ik^{-1} \lambda_k^s$$

(where λ_k^s is a small real constant, which in (25.4) is replaced by zero); and z_k^s is a root of the equation $U \cos \theta_k = sk^{-1} \omega_k$. Instability appears (the waves grow) if $U' > 0$, $U'' < 0$, for example, for a logarithmic profile $U(z)$. Taking into account the

two right-hand sides of (25.2) and (25.4), we obtain

$$F_k = \frac{\alpha_k}{\beta_k - 2\gamma_k} [e^{(\beta_k - 2\gamma_k)\tau} - 1], \quad (25.5)$$

where $\beta_k = \beta_k^+$. Note, however, that the quasilinear theory gives a value of the rate β_k of exponential growth of wind waves which is several times less than that given by the experimental data. This shortcoming can be rectified if we allow for the interaction of wave perturbations with the energy-bearing large-scale components of the atmospheric turbulence (this approximation is called *adiabatic*). Since they are quasihorizontal, they can simply be added to $\mathbf{U}$; but since at the limits of the water-adjacent air layer they vary little with height, this additive factor can be assumed to have the approximate form ξu_* , where u_* is the *friction velocity*, while ξ is a Gaussian random quantity with a zero mean value (Zaslavskii, 1974).

The viscous damping of the waves, even allowing for the effect of the atmosphere, is minimal, so that the principal mechanism of dissipation of wave energy is the breaking of the crests when stability is lost; this happens if the local accelerations at the crests exceed some definite fraction of g (for Stokes waves this is $g/2$). Assuming such a dissipation mechanism, we find that the spectral intervals of energy supply $\Delta k^+ = (k < k^+)$ and dissipation $\Delta k^- = (k > k^-)$ (which make the main contribution, for instance, 80%, to the integral values of supply and dissipation) overlap, that is, $k^+ < k^-$; thus in the region $k^+ < k < k^-$ a statistical equilibrium is established between supply and dissipation, determined by parameter g (and to a lesser degree by the wind friction velocity u_* , governing the surface drift, which may facilitate breaking). The spectrum F_k of a surface with sharp crests must be universal, and proportional to k^{-4} , since it is the mean square of the two-dimensional transform of the Fourier function describing the surface, and this transform is proportional to k^{-2} . Dimensional considerations indicate that $F_k = k^{-4} f(\theta)$, where $\theta = \text{arctg}(k_2/k_1)$ (the weak dependence on ku_*^2/g will not be discussed further). The fact arrived at here, that the angular distribution of energy is independent of k , can apparently be satisfied approximately only for very high k . Therefore, it is better not to assume isotropy and to use the spectrum of the wave-number moduli $E(k)$, together with the spectrum of frequencies $S(\omega)$, defined in terms of $E(k)$ with the aid of the dispersion relation. These are given by the formulas:

$$E(k) = \int_{-\pi}^{\pi} k F_k d\theta; \quad S(\omega) = \left[E(k) \left| \frac{\partial \omega_k}{\partial k} \right|^{-1} \right]_{k=\kappa(\omega)}, \quad (25.6)$$

where $k = \kappa(\omega)$ is the solution of the equation $\omega = \omega(k)$. Now, after determining $E(k)$ from dimensional considerations and $S(\omega)$ from the second formula of (25.6) for $\omega_k^2 = gk$, we obtain

$$E(k) = \frac{1}{2} \beta k^{-3}; \quad S(\omega) = \beta g^2 \omega^{-5} \quad (25.7)$$

(Phillips, 1958). Empirical data give $\beta \approx 6 \times 10^{-3}$. This coefficient exhibits some variability, apparently due to the influence on the equilibrium waves of accelerations produced by longer waves, and possibly also due to the stratification of the upper layer of the ocean. Note that, in the equilibrium interval, the spectral tensor of the sea-surface inclinations $\partial \zeta / \partial x_i$ must have the form $C_{ij} \omega^{-1}$, where in view of the symmetry with respect to the wind direction we should have $C_{12} = C_{21} = 0$; this prediction was verified by measurements, which gave $C_{11} + C_{22} = 8 \times 10^{-3}$.

For gravity waves on a water surface of finite depth, instead of the second formula of (25.7) we have

$$S(\omega) = \beta g^2 \omega^{-5} \varphi(\omega_h); \quad \varphi(\omega_h) = \frac{\omega h}{\kappa^2} \left[1 + \frac{2\omega_h^2 \kappa}{\text{sh}(2\omega_h^2 \kappa)} \right]^{-1}, \quad (25.7')$$

where $\omega_h = \omega (h/g)^{1/2}$, while $\kappa = \kappa(\omega_h)$ is a root of the equation $\kappa \text{th}(\omega_h^2 \kappa) = 1$. For $h \rightarrow \infty$ we have $\varphi \rightarrow 1$, but for $\omega_h \rightarrow 0$ we get $\varphi \sim \omega_h^2/2$ and $S(\omega) \sim \frac{1}{2} \beta g h \omega^{-3}$, which is corroborated by observations on a shallow sea. For capillary waves the existence of an equilibrium interval is hypothetical; if it indeed exists, then in it

$$E(k) = \beta_1 k^{-3}; \quad S(\omega) = \beta_1 \gamma^{2/3} \omega^{-7/3}, \quad (25.8)$$

with a slight deviation from (24.18'), while the spectral tensor of the slopes must have the form $C'_{ij} \omega^{-1}$ (which is corroborated by measurements, giving $C'_{11} + C'_{22} = 1 \times 10^{-2}$).

The overall picture of the development of wind waves is obtained as follows (see Figure 25.1). The waves originate in the energy-supply interval Δk^+ at the wave number k_β of the maximum of the function $P_k^+ - P_k^-$ (probably at the boundary k^- of the dissipation interval). For $k^+ < k < k^-$ (interval I) the wave spectrum should be described by the Phillips law $E(k) \sim k^{-3}$. With a growth of the waves, the energy-bearing part of the spectrum moves into the region $k_e < k < k^+$ (interval II), where the complete kinetic equation (24.19) applies, and then into the region $k_1^- < k < k_e$ (interval III) outside the supply interval, where the simplified kinetic equation (24.12) applies, and the wave spectra have the form (24.17). Let us use the dimensionless wave numbers $\tilde{k} = kU^2/g$ and frequencies $\tilde{\omega} = \omega U/g$. According to experimental lab data, $\tilde{\omega}_\beta \sim 30$ and $\tilde{k}_\beta \approx \tilde{\omega}_\beta^2 \sim 10^3$ (whereas in the quasilaminar model of Miles, according to the lower boundary of the logarithmic velocity profile for smooth flow, we have only $\tilde{\omega}_\beta \sim 3$ and $\tilde{k}_\beta \sim 10$). Empirical estimates yield: $\tilde{\omega}^+ = 2-2.5$ and $\tilde{k}^+ = 4-6$; $\tilde{\omega}_e = 1.5-2$ and $\tilde{k}_e = 2-4$; the dimensionless spectrum (24.17) for $\tilde{\omega} \sim \tilde{\omega}_m \sim 1$, $\tilde{\omega} < \tilde{\omega}^+$ is then

$$\left(\frac{g^3}{U^5} \right) S(\omega) = \tilde{S}(\tilde{\omega}) = 2a\tilde{\varepsilon}_n^{1/3} \tilde{\omega}^{-11/3} \Theta_s \left(\frac{\tilde{\omega}}{\tilde{\omega}_m} \right), \quad (25.9)$$

where $\Theta_s(x)$ is the slightly blurred Heaviside function (jump from 0 to 1 at point x), while $b_s = 2a\tilde{\varepsilon}_n^{1/3} \approx 4 \times 10^{-3}$, $\tilde{\varepsilon}_n = g^2 \varepsilon_n / U^4 \approx 1.4 \times 10^{-8}$, and $a \approx 0.83$. The dependence of the single parameter of this spectrum $\tilde{\omega}_m$ on the dimensionless

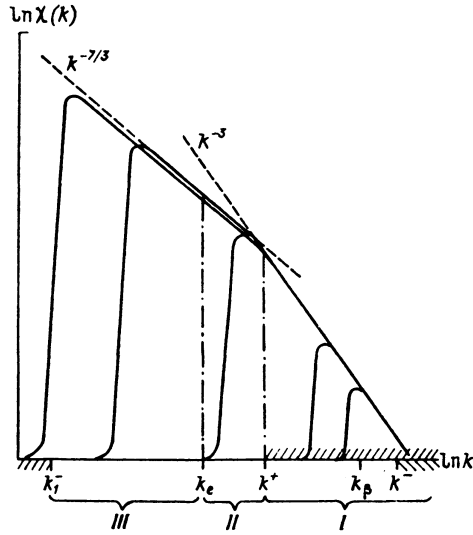


Fig. 25.1. Development of wind-wave spectrum.

fetch length $\tilde{x} = gx/U^2$ and on the time of action of the wind $\tilde{t} = gt/U$ is obtained after integrating the kinetic equation (24.19) over k , taking into account the narrow directivity of the wave spectrum in the energy-bearing region:

$$2 \frac{\partial \tilde{\omega}_m}{\partial \tilde{t}} + \frac{1}{\tilde{\omega}_m} \frac{\partial \tilde{\omega}_m}{\partial \tilde{x}} = - \frac{1}{a} \tilde{\epsilon}_n^{2/3} \tilde{\omega}_m^{14/3} \quad (25.10)$$

(Zakharov and Zaslavskii, 1982). For established horizontally nonuniform waves, developing over a fetch length x from the windward shore $x = 0$, we thus have

$$\tilde{\omega}_m = b_\omega(\tilde{x})^{-3/14}, \quad b_\omega = \frac{3}{14} a^{2/14} (\tilde{\epsilon}_n)^{-1/7} \approx 9.2. \quad (25.11)$$

This formula for $\tilde{x} \approx 10^2 - 10^3$ is verified satisfactorily by data obtained under natural conditions. With the aid of (25.11), we can use formula (25.9) to calculate the dimensionless mean-square height of the waves $\tilde{h} = (g/U^2) (2\pi e)^{1/2}$:

$$\tilde{h} = b_h(\tilde{x})^{2/7}, \quad b_h = \left(\frac{3\pi}{2} \right)^{1/2} \left(\frac{14}{3} \right)^{3/7} a^{3/14} (\tilde{\epsilon}_n)^{5/14} \approx 5.1 \times 10^{-3} \quad (25.12)$$

in good agreement with the field data. Formulas (25.11) and (25.12) can serve as a basis for forecasting wind waves. The calculation of the angular energy distribution in a two-dimensional wind-wave spectrum constitutes a problem which has not yet been thoroughly solved.

Internal Waves

26. LINEAR THEORY

In this chapter we will describe internal *gravity waves* (capable of propagating in stably stratified media, so that they turn out to be a typical phenomenon in the World Ocean, frequently in the Earth's atmosphere, and possibly also in the liquid core, if it is stably stratified; such waves are related to the gravitational modes of the natural oscillations of the Sun and the large planets). We will make use of, as a rule, the following three simplifications: (1) the adiabatic approximation, including Equation (1.17), but with the simplification $c^2 \approx c_0^2(z)$; (2) the Boussinesq approximation, that is, formula (3.4) or (3.5) and the condition of nondivergence of the velocity field; (3) the sphericity of the Earth is neglected, and its rotation is taken into account only in the 'f-plane' approximation, that is, for the 'traditional approximation' in the Coriolis acceleration and ignoring the latitude dependence of the Coriolis parameter $f = 2\omega_z$ (except for some cases which will be discussed specially). Therefore, we will use the following equations, together with the boundary conditions appropriate to the ocean:

$$\begin{aligned} \frac{d_h \mathbf{u}}{dt} + w \frac{\partial \mathbf{u}}{\partial z} + f \nabla r \times \mathbf{u} &= -\frac{\nabla p}{\rho_*} - N^2 \xi \nabla r; \\ \text{div } \mathbf{u} &= 0; \quad \frac{d_h N^2 \xi}{dt} = N^2 w; \end{aligned} \quad (26.1)$$

$$\begin{aligned} z = 0: \quad w &= \frac{d_h \xi}{dt}; \quad \frac{d_h p}{dt} - g \rho_0 w = \frac{d_h p_a}{dt}; \\ z = -h: \quad w &= -\mathbf{u} \cdot \nabla h, \end{aligned} \quad (26.2)$$

where d_h/dt is the individual derivative with respect to the horizontal motion, and $\xi = \xi(\mathbf{x}, z, t)$ are the vertical displacements of the fluid particles ($\xi = -g\eta''/N^2$, where η'' is defined by (3.5), so that the third of Equations (26.1) is equivalent to Equation (1.17)). In the first two conditions of (26.2) it is taken into account that internal waves produce only very small oscillations of the ocean surface $z = \xi$; thus, these conditions are shifted at the $z = 0$ level. The linearization of Equations (26.1) and (26.2) relative to a *horizontally uniform* (whereby $N = N_0(z)$ and $h = \text{const}$) state of rest reduces to everywhere replacing $d_h/dt + w \partial/\partial z$ by $\partial/\partial t$. Elimination of $\mathbf{u}$ and p from these expressions yields the following equations and

boundary conditions in terms of ξ :

$$\begin{aligned} \frac{\partial^2}{\partial t^2} \left(\Delta_h \xi + \frac{1}{\rho_*} \frac{\partial}{\partial z} \rho_* \frac{\partial \xi}{\partial z} \right) + \frac{f^2}{\rho_*} \frac{\partial}{\partial z} \rho_* \frac{\partial \xi}{\partial z} + N_0^2 \Delta_h \xi = 0; \\ z = 0: \quad \left(\frac{\partial^2}{\partial t^2} + f^2 \right) \frac{\partial \xi}{\partial z} - g \Delta_h \xi = \frac{\Delta_h p_a}{\rho_0}; \quad z = -h: \quad \xi = 0. \end{aligned} \quad (26.3)$$

The solution of this equation can be represented as a superposition of modes, that is, of standing (relative to the vertical) elementary plane waves:

$$\xi(\mathbf{x}, z, t) = \int e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} \xi_{k\omega}(z) dZ_{k\omega}; \quad (26.4)$$

$$dZ_{k\omega} = \sum_{n=0}^{\infty} (A_{nk} \delta_{\omega - \omega_{nk}} + A_{-nk} \delta_{\omega + \omega_{nk}}) d\mathbf{k} d\omega,$$

where A_{nk} and A_{-nk} are the complex amplitudes of the modes propagating in the positive and negative directions of vector $\mathbf{k}$; they are determined by the initial conditions, they depend on the particular normalization of the eigenfunctions $\xi_{k\omega}(z)$ (assumed in the following to be real and dimensionless), and they satisfy the realness condition $A_{nk} = A_{-n, -\mathbf{k}}^*$. Function $\xi_{k\omega}(z)$ is a solution of the equation

$$\begin{aligned} (D + k^2 \beta^2 \rho_*) \xi = 0; \quad D = \frac{\partial}{\partial z} \rho_* \frac{\partial}{\partial z}; \\ \beta^2 = \frac{N_0^2(z) - \omega^2}{\omega^2 - f^2}, \end{aligned} \quad (26.5)$$

satisfying one of the following two boundary conditions:

$$z = 0: \quad (\omega^2 - f^2) \frac{\partial \xi}{\partial z} = g k^2 \xi; \quad z = -h: \quad \xi = 0. \quad (26.5')$$

The velocity components u , v , w and the fluctuations of the pressure p and the density ρ in the field of the internal waves are represented by formulas of the type of (26.4), with $\xi = \xi_{k\omega}(z)$ replaced by, respectively, the functions:

$$\begin{aligned} u = \frac{k_x \omega + i k_y f}{k^2} \frac{\partial \xi}{\partial z}; \quad v = \frac{k_y \omega - i k_x f}{k^2} \frac{\partial \xi}{\partial z}; \quad \omega = -i \omega \xi; \\ p = \frac{\omega^2 - f^2}{k^2} \rho_* \frac{\partial \xi}{\partial z}; \quad \rho = \frac{\omega^2 - f^2}{k^2} \frac{\rho_*}{c_0^2} \frac{\partial \xi}{\partial z} - \frac{\rho_* N_0^2}{g} \xi. \end{aligned} \quad (26.6)$$

It should be noted that internal waves are vortical waves: if the x axis is in the direction of wave propagation $\mathbf{k}$, then the amplitude of the y component of the vorticity $\partial u / \partial z - \partial w / \partial x$ is given by the expression $(\omega / k)(\partial^2 \xi / \partial z^2 - k^2 \xi)$, which goes to zero at the water surface but is, in general, not zero at any depths.

After substituting $\zeta = \zeta_{k\omega}(z)$ in the second boundary condition, we get for a fixed horizontal wave number k an equation in ω having in general a denumerable number of roots: the eigenfrequencies $\omega = \pm \omega_{nk}$ of the different modes $n = 0, 1, 2, \dots$. Let $k = k_{n\omega}$ be the inverse function of $\omega = \omega_{nk}$, and let us introduce the frequency eigenfunctions $\zeta_{n\omega}(z) \equiv \zeta_{k_{n\omega}\omega}(z)$ of the boundary-value problem (26.5) and (26.5'), orthonormalized by the condition $\zeta_{m\omega} \otimes \zeta_{n\omega} = \delta_{mn}$, where the functional scalar product is defined by the formula

$$\zeta_{m\omega} \otimes \zeta_{n\omega} = \frac{\omega^2}{\omega^2 - f^2} \left[g\rho_0 \zeta_{m\omega}(0) \zeta_{n\omega}(0) + \int_{-h}^0 (N^2 - f^2) \zeta_{m\omega}(z) \zeta_{n\omega}(z) \rho_* dz \right]. \quad (26.7)$$

Instead of $\zeta_{n\omega}(z)$ we can introduce the space eigenfunctions $\zeta_{nk} \equiv \zeta_{k\omega_{nk}}(z)$ with the orthogonality condition

$$\zeta_{mk}(z) \otimes \zeta_{nk}(z) = g\rho_0 \zeta_{mk}(0) \zeta_{nk}(0) + \int_{-h}^0 (N^2 - f^2) \zeta_{mk}(z) \zeta_{nk}(z) \rho_* dz = \|\zeta_{nk}\|^2 \delta_{mn}, \quad (26.7')$$

where $\|\zeta_{nk}\| = (\zeta_{nk} \otimes \zeta_{nk})^{1/2}$ is the norm. Using these functions, we can now write the spectral representation of the field $\zeta(\mathbf{x}, z, t)$ as

$$\zeta(\mathbf{x}, z, t) = \sum_{n=0}^{\infty} \int e^{ik \cdot \mathbf{x}} \zeta_{nk}(z) (A_{nk} e^{-i\omega_{nk}t} + A_{-nk} e^{i\omega_{nk}t}) dk. \quad (26.4')$$

The functions (26.6) corresponding to $\zeta_{nk}(z)$ are, in general, complex; if we set $\omega_{-nk} = -\omega_{nk}$, then these functions will satisfy the condition $\eta_{nk} = \eta_{-n, -k}^*$ and they will determine the fluid-dynamical fields

$$\eta(\mathbf{x}, z, t) = \sum_{n=0}^{\infty} \int e^{ik \cdot \mathbf{x}} [\eta_{nk}(z) A_{nk} e^{-i\omega_{nk}t} + \eta_{-nk}(z) A_{-nk} e^{i\omega_{nk}t}] dk. \quad (26.7'')$$

Boundary-value problem (26.5) and (26.5') differs from the Sturm—Liouville problem in that the eigenfrequency enters into the boundary condition. For $\min N_0^2(z) \geq f^2$ with the substitution $\omega^2 - f^2 = \lambda^2$ it reduces to the problem considered by Yanovich (1962), from which it follows that the spectrum of eigenvalues is purely discrete, all the eigenfrequencies $\pm \omega_n$ are real (since adiabaticity dictates that the integral total energy does not vary with time); $|\omega_n| > |\omega_{n+1}| > |f|$ ($n \rightarrow \infty$), but the concentration points themselves $\pm f$ are not eigenvalues; in order for extrema of the eigenfunctions to exist at $z \neq 0$, that is, in order for internal waves to exist, the inequality $|f| < |\omega| < \max N_0(z)$ must be observed; for $N_0^2(z) > f^2$ the system of eigenvectors is complete and they are orthogonal with a weight ρ_* (the discreteness of the spectrum and the ortho-

gonality of the eigenvectors with a weight ρ_* are, with certain reservations, also valid for an arbitrary sign of $N_0^2(z) - f^2$.

It can be verified that the frequency ω_{0k} , obtained for $n = 0$, increases monotonically with increasing k , from an absolute value equal to $|f|$ at $k = 0$ to $(gk)^{1/2}$ as $k \rightarrow \infty$. The corresponding eigenfunction $\xi_{0k}(z)$ decreases monotonically (and quite rapidly) with depth, dropping from $\xi_{0k}(0) \neq 0$ to $\xi_{0k}(-h) = 0$. Functions ω_{0k} and $\xi_{0k}(z)$ represent surface gravity waves; they depend slightly on the stratification of the water. Modes $n = 1, 2, \dots$, on the other hand, are determined wholly by the shape of function $N_0(z)$.

For the World Ocean low values of $N_0(z)$ are typical in the upper mixed layer $0 \geq z \geq -H$, with a maximum $N_0 \sim 10^{-2} \text{ s}^{-1}$ (period $2\pi/N_0 \sim 10$ min) in the pycnocline $-H \geq z \geq -H - \delta$ and then low values again deeper than the pycnocline (at $z < -H - \delta$). Such a distribution $N_0(z)$ is described in a simplified manner by a three-layer model with a constant value of $N_0 > 0$ in the pycnocline and $N_0 = 0$ outside of it, in which for $f = 0$ and $h = \infty$ a continuous solution of Equations (26.5), (26.5') is obtained in the following form (Monin, 1974):

$$\begin{aligned}\xi^{(1)}(z) &= \xi(-H) \frac{\omega^2 \operatorname{ch} kz + gk \operatorname{sh} kz}{\omega^2 \operatorname{ch} kH - gk \operatorname{sh} kH}; \\ \xi^{(2)}(z) &= \xi(-H) e^{(z+H)N_0^2/2g} \times \\ &\quad \times \frac{l \cos l(z+H+\delta) + k(1 - N_0^2/2gk) \sin l(z+H+\delta)}{l \cos l\delta + k(1 - N_0^2/2gk) \sin l\delta}; \\ \xi^{(3)}(z) &= \xi(-H) e^{-\delta N_0^2/2g} \frac{l e^{k(z+H+\delta)}}{l \cos l\delta + k(1 - N_0^2/2gk) \sin l\delta},\end{aligned}\quad (26.8)$$

where indices (2), (1), (3) pertain to, respectively, the pycnocline, the layer above it, and the layer below it, while l is the vertical wave number, which is related to the frequency ω by the expression

$$\omega^2 = N_0^2 \left(1 + \frac{l^2}{k^2} + \frac{N_0^2}{4g^2 k^2} \right)^{-1}. \quad (26.8')$$

If we set $\omega^2 = gk$, then the continuity condition $\partial \xi / \partial z$ for $z = -H$ is satisfied for any N_0 and k ; thus we obtain $\xi(z) = \xi(0) e^{-kz}$, that is, surface waves undistorted by the stratification exist. However, for internal waves, with $\omega^2 \neq gk$, this continuity condition yields the following equation in terms of the vertical wave numbers l (or, according to (26.8'), in terms of the frequencies ω):

$$\operatorname{tg} l\delta = F(l) \equiv \frac{(l/k)(1 + \operatorname{cth} kH)}{(l/k)^2 - (1 + N_0^2/2gk)(\operatorname{cth} kH - N_0^2/2gk)}. \quad (26.9)$$

It has a denumerable set of real positive roots $l_1 < l_2 < \dots \rightarrow \infty$, extracted at the intersections of the curve $y = \operatorname{tg} l\delta$ with the curve $y = F(l)$ and for large n approaching at the upper limit the points $l_n = \pi n / \delta$. The frequencies corresponding to these roots satisfy the inequalities $N_0 > \omega_1 > \omega_2 > \dots \rightarrow 0$, while the corresponding modes $\xi_n(z)$ for not too low n each have n nodes (that is, depths where ξ changes sign) in the pycnocline.

In the case of waves which are long compared to the thickness of the pycnocline, that is, for $k\delta \ll 1$, and thus $k/l \ll 1$, we can replace function $F(l)$ in (26.9) approximately by $(k/l)(1 + \text{cth } kH)$, so that $\text{tg } l\delta$ will also be small, and for the lower mode of the internal waves we can set $\text{tg } l_1\delta \approx l_1\delta$. Then (26.9) takes the form $(l_1\delta)^2 \approx k\delta(1 + \text{cth } kH)$, and for the square of the corresponding frequency $\omega_1^2 \approx N_0^2 k^2 / l^2$ with $N_0^2 \approx g \Delta\rho / \rho\delta$ (where $\Delta\rho$ is the density difference in the pycnocline) we obtain

$$\omega_1^2 \approx \left(\frac{\Delta\rho}{\rho} \right) gk(1 + \text{cth } kH)^{-1}. \quad (26.10)$$

For small $k\delta$ the variations of $\xi(z)$ with depth are also small in the pycnocline, so that in the lower mode of a long wave it oscillates, as a whole. Therefore, using the first formula of (26.6) for $k_x = k$, $k_y = 0$, and then Equation (26.5) for $\rho_* = \text{const}$, $f = 0$, we arrive at the shear of the longitudinal velocity in the wave in the form $\partial u / \partial z = (\omega/k) \partial^2 \xi / \partial z^2 \approx (k/\omega)(\omega^2 - N_0^2)\xi$. This yields a Richardson number in the lower mode in the pycnocline equal to

$$\text{Ri} = N_0^2 \left(\frac{\partial u}{\partial z} \right)^{-2} \approx \left(\frac{\omega_1}{N_0} - \frac{N_0}{\omega_1} \right)^{-2} k^{-2} \xi_H^{-2} \quad (26.11)$$

(Phillips, 1969). In addition, since the longitudinal velocity on the ocean surface is $u_s \approx \omega_1 \xi_H / \text{sh } kH$, the maximum convergence of u_s will occur over the troughs of the wave in the pycnocline (and, actually, bands of foam perpendicular to the direction of wave propagation are sometimes observed there).

Let us recall that even slight shears in the velocity of the mean currents may be comparable to the phase velocities of internal waves and may create critical levels for them (boundaries of the wave channels). Linearization of Equations (26.1) and (26.2) with respect to a horizontal current with a shear $\mathbf{U}(z)$ yields, instead of (26.5) and (26.5'), equations in which ω is replaced by $k(c - U_k)$ (where c is the phase velocity of the waves and $U_k = \mathbf{k} \cdot \mathbf{U}/k$). In the simplest formulation of the problem for $f = 0$, operator $D\xi$ is replaced by $\omega^{-2}(\partial/\partial z)\omega^2\rho_*(\partial\xi/\partial z)$. For $w = -i\omega\xi$ these equations for $\rho_* \approx \text{const}$ reduce to

$$\frac{\partial^2 w}{\partial z^2} + \left[\frac{N_0^2}{(U_k - c)^2} - \frac{U_k''}{U_k - c} - k^2 \right] w = 0; \quad (26.12)$$

$$z = 0: \quad \frac{\partial w}{\partial z} = \frac{g + (U_k - c)U_k'}{(U_k - c)^2} w; \quad z = -h: \quad w = 0. \quad (26.12')$$

A shear in the mean current not only changes the usual modes of internal waves (which are reestablished for $U_k \equiv 0$), but also, because of singularities at the critical levels z_c , where $U_k(z) = c$, new modes appear either with complex eigenvalues $c = c_r + ic_i$ or with real c but with discontinuities of the eigenfunctions at the critical levels. Equation (26.12) is equivalent to (11.24), so that for it the theorem of Miles and Howarth (1961) is valid: unstable modes ($c_i > 0$) may appear only if somewhere it turns out that $\text{Ri}(z) = N_0^2(U_k')^{-2} < \frac{1}{4}$; in this case

according to Howarth's theorem points c in the complex plane lie in a circle with a diameter on the segment $U_{k \min} < c_r < U_{k \max}$. According to Banks *et al.* (1976), the number of unstable modes can only be finite; the complex conjugates of these are damped modes, and the stable modes include a discrete set of usual modes (with $c > U_{k \max}$ or $c < U_{k \min}$), a finite number of modes with $U_{k \min} < c < U_{k \max}$ and with branch points at the critical levels, and a continuous spectrum $U_{k \min} < c < U_{k \max}$ with discontinuities at the critical levels.

Let us examine the behavior of a solution $w(z)$ of Equation (26.12) near the critical level $z = z_c$. For low $c_i = \text{Im } c$ (that is, for $|U''|(U')^{-2}c_i \ll 1$; here and in the following subscript k of U_k will be dropped), the general solution can be written as

$$w = A \left(z - z_c - \frac{ic_i}{U'} \right)^{1/2 + i\mu} + B \left(z - z_c - \frac{ic_i}{U'} \right)^{1/2 - i\mu}, \quad (26.13)$$

where $\mu = (\text{Ri} - \frac{1}{4})^{1/2} > 0$; for $c_i = 0$ both terms here have branch points $z = z_c$. If we let $c_i > 0$, then as $z - z_c$ decreases from a positive value (larger than c_i/U') to a negative value, $\arg(z - z_c - ic_i/U')$ varies from 0 to $-\pi$, and for $z > z_c$ we obtain

$$(z - z_c)^{1/2 + i\mu} = |z - z_c|^{1/2} e^{i\mu|z - z_c|},$$

while for $z < z_c$ we have

$$(z - z_c)^{1/2 \pm i\mu} = -i|z - z_c|^{1/2} e^{\pm \pi\mu \mp i\mu|z - z_c|}$$

(from which we see that for $c_i = 0$ function $w(z)$ at $z \approx z_c$ oscillates rapidly). For quasiconstant N_0 and U solution (26.13) becomes

$$w = A e^{ilz} + B e^{-ilz}; \quad l^2 = N_0^2(U - c)^{-2} - k^2. \quad (26.13')$$

Assuming $\text{Im } l > 0$ for $c_i > 0$, we get $l \approx N_0(U - c)^{-1}$ for $k^2 \ll N_0^2(U - c)^{-2}$ and $l \approx ik$ for $k^2 \gg N_0^2(U - c)^{-2}$. The first term in (26.13') describes a plane wave with a vertical component of the phase velocity pointing downward (for $U - c < 0$, that is, wave motion in the direction of increasing x , we have $l < 0$, meaning that the lines of constant phase $kx + lz - kct = \text{const}$ move downward; similarly, for $U - c > 0$, we have $l > 0$, and the lines of constant phase move upward). The second term, correspondingly, describes a wave with a phase-velocity component directed upward. From formula (26.13') for l^2 we obtain $\sigma = kc = kU \pm N_0 k(k^2 + l^2)^{-1/2}$, where the upper sign is selected for $U - c < 0$, $l < 0$, and the lower sign for $U - c > 0$, $l > 0$. The vertical component of the group velocity for the first term in (26.13') takes the form $c_{gz} = \partial\sigma/\partial l = kl(U - c)^3 N_0^{-2} > 0$, that is, the energy in this wave propagates upward; in the second wave, on the other hand, it propagates downward.

Near the critical point $z = z_c$ these results are valid for not only quasiconstant, but also arbitrary, $N_0(z)$ and $U(z)$: then for $c_i = 0$ in (26.13) we can introduce the local vertical wave number

$$l = \mu(z - z_c)^{-1} \approx N_0(U - c)^{-1} \left(1 - \frac{1}{4\text{Ri}} \right)^{1/2},$$

for large Ri close to the above-considered value. Here w is small, while $u \sim \partial w / \partial z \sim (z - z_c)^{-1/2}$, on the other hand, is large, so that the motions of particles in the wave become almost horizontal, while the wave frequency is very low.

The equation of horizontal motion $(U - c) \partial u / \partial x = -\partial p / \rho \partial x$ yields the mean vertical energy flux of the wave $\langle pw \rangle = -\rho(U - c)\langle uw \rangle$. With the aid of (26.6), the Reynolds stress is found as

$$\langle uw \rangle = \frac{1}{4ik} \left(w \frac{\partial w^*}{\partial z} - w^* \frac{\partial w}{\partial z} \right)$$

and for (26.13') it is $(\mu/2k)(|B|^2 - |A|^2)$ for $z > z_c$ and $(\mu/2k)(|A|^2 e^{2\pi\mu} - |B|^2 e^{-2\pi\mu})$ for $z < z_c$, that is, for $z = z_c$ it goes through a discontinuity and also changes in sign and absolute value. The energy flux in this case does not change sign, but its absolute value goes through a discontinuity, so that in the critical layer absorption of wave energy takes place (Booker and Bretherton, 1967), which in the absence of dissipation and nonlinear effects is evidently imparted to the mean flow.

As opposed to conditions of hydrodynamic stability, when $Ri > \frac{1}{4}$, if $0 < Ri < \frac{1}{4}$, then according to Jones (1968) there may be strong reflection of internal waves from the critical level, whereas for $Ri < Ri_{cr}(k, \omega)$ 'superreflection' may even take place, with the reflected wave obtaining momentum and energy from the mean flow and proving to be more intense than the incident wave.

As long as the system of eigenfunctions $\zeta_{nw}(z)$ (or $\zeta_{nk}(z)$) is complete, the field of the internal waves can always be represented in form (26.4) (or (26.4')), that is, in the form of a superposition of elementary waves (modes) propagating only horizontally, while vertically they are standing waves. However, individual vertically standing waves may be physically unrealistic, since their formation requires repeated reflections of long coherent wave trains from the surface and bottom of the ocean or else from the boundaries of the wave channel, and fixed phase ratios would have to be established between the ascending and descending components. The typical times of vertical wave propagation, especially of the higher modes (with short vertical wavelengths), may be comparable to the times of nonlinear interactions preventing the formation of the modes (for example, randomizing phases).

If the vertical wavelength $2\pi/l$ is small compared to the scales of inhomogeneities in the profiles $N_0(z)$ and $U(z)$, then wave trains propagating with a vertical component (upward or downward) become quite realistic, their local dispersion relation being

$$\begin{aligned} \sigma &= \omega + \mathbf{k} \cdot \mathbf{U}; & \omega^2 &= N_0^2 \cos^2 \theta + f^2 \sin^2 \theta; \\ \cos \theta &= k(k^2 + l^2)^{-1/2}; & \sin \theta &= l(k^2 + l^2)^{-1/2}; \\ \operatorname{tg} \theta &= \frac{l}{k} = \left(\frac{N_0^2 - \omega^2}{\omega^2 - f^2} \right)^{1/2}, \end{aligned} \quad (26.14)$$

so that $f < \omega < N_0$ and ω is independent of the modulus $(k^2 + l^2)^{1/2}$ of the

three-dimensional wave vector, being rather a function of its inclination to the horizontal θ (which also shows the dispersion of these waves).

For $f = 0$ and $U = 0$ this formula is derived from (26.8') in cases where $l\delta \gg \Delta\rho/\rho$, which is satisfied for higher modes of the waves (26.8), and frequently for the lower mode as well; in this case the phase velocity $c \approx N_0 k(k^2 + l^2)^{-1}$ is always less than N_0/k and it decreases with increasing mode number (for example, for $\tau = 2\pi/N_0 \approx 10$ min waves of kilometer length $2\pi/k$ propagate with velocities no greater than several m/s).

The group velocity $\mathbf{c}_g$ of waves with a frequency ω and the polarization vector $\mathbf{V}$ of the velocity field in these waves (orthogonal to the three-dimensional wave vector) are:

$$\begin{aligned}\mathbf{c}_g &= \frac{N_0^2 - \omega^2}{\omega(k^2 + l^2)} \left(k_x, k_y, -\frac{k^2}{l} \right); \\ \mathbf{V} &= C \left(k_x \omega + i k_y f, k_y \omega - i k_x f, -\frac{\omega k^2}{l} \right); \\ C &= (\omega k)^{-1} (N_0^2 - \omega^2)^{1/2} (N_0^2 - f)^{-1/2}.\end{aligned}\tag{26.15}$$

The polarization is elliptical; close to the inertial frequency ($\omega \lesssim f$, $l^2 \gg k^2$) particles in the waves move almost horizontally and approximately along circles, whereas close to the Brunt–Väisälä frequency ($\omega \lesssim N_0$, $l^2 \ll k^2$) they oscillate almost entirely along the vertical. Note, too, that the energy flux density of a wave is equal to $E \mathbf{c}_g$, where $E = \frac{1}{2} \langle u^2 + v^2 + w^2 + N_0^2 \xi^2 \rangle$ is the density of the total energy of an internal wave, averaged over the period.

27. RAY THEORY

If the scales of the spatial inhomogeneities of the ocean depth h , the Coriolis parameter f , the Brunt–Väisälä frequency N , and the mean current $\mathbf{U}$, and of the variations in N and $\mathbf{U}$ with time, are large in comparison with the length and period of an internal wave of a fixed mode, then as in Section 19 the wave can be described as a train $\zeta(\xi, t) = a(\xi, t) \exp[i\theta(\xi, t)]$ of locally plane waves with a gradually varying amplitude a and an eikonal θ that depends approximately linearly on ξ and t (where $\xi = (\mathbf{x}, z)$ is a three-dimensional coordinate). We can then determine the three-dimensional wave vector $\kappa = (k_x, k_y, l)$ as $\nabla\theta$, and the frequency $\sigma = \omega + \kappa \cdot \mathbf{U}$ (where ω is given by formula (26.13)) as $(-\partial\theta/\partial t)$; the equation of conservation of wave number (19.2) is thus valid. The ray equations have the form

$$\dot{\xi} = \frac{\partial\sigma}{\partial\kappa} = \mathbf{c}_g + \mathbf{U}; \quad \dot{\kappa} = -\frac{\partial\sigma}{\partial\xi} = \mathbf{r} - \kappa_\alpha \frac{\partial U_\alpha}{\partial\xi}; \quad \dot{\sigma} = \frac{\partial\sigma}{\partial t}, \tag{27.1}$$

where $\mathbf{r} = -\partial\omega/\partial\xi$ is the so-called *refraction velocity*. With the aid of these equations and an equation of type (19.12) for the energy density of the waves E ,

we arrive at the following equation of conservation of wave action E/ω :

$$\frac{\partial}{\partial t} \frac{E}{\omega} + \operatorname{div} \frac{E}{\omega} (\mathbf{c}_g + \mathbf{U}) = 0 \quad (27.2)$$

(Witham, 1965; Bretherton and Garrett, 1968). In the case of a single monochromatic wave, the spatial distribution of the action E/ω is stationary, and the first term here vanishes. Let us also cite the law of reflection of a wave from an inclined rigid wall $z = x \operatorname{tg} \alpha + \text{const}$: since the frequencies of the incident and reflected waves are the same, the angles of inclination $\phi = \pi/2 - \theta$ of their wave vectors κ and κ' to the vertical are equal to one another. Consequently, the angle of incidence $\phi + \alpha$ is not equal to the angle of reflection $\phi - \alpha$, but since the velocity normal to the bottom is zero the components of the wave vectors in the (s_g, s'_g) plane must be:

$$\kappa \sin(\varphi + \alpha) = \kappa' \sin(\varphi - \alpha). \quad (27.3)$$

High-frequency waves with $\omega > \omega_s = (N^2 \sin^2 \alpha + f^2 \cos^2 \alpha)^{1/2}$, that is, $\varphi > \alpha$, are reflected forward (Figure 27.1(a)), so that they propagate obliquely upward, and their energy may be stored at underwater peaks. On the other hand, low-frequency waves with $\omega < \omega_s$, that is, $\varphi < \alpha$, are reflected backward (Figure 27.1(b)), so that they propagate obliquely downward, and their energy may be stored in submarine valleys.

Let us first apply the equations of ray theory (27.1) and (27.2) to the simplest case $h = \text{const}$, $f = \text{const}$, $\mathbf{U} = 0$, with N varying only in the z direction and moreover slowly, so that $N^{-1} \partial N / \partial z \sim \varepsilon \ll 1$. According to the equations of ray theory, k and ω are here constant. Let us write for this case Equation (26.5) for

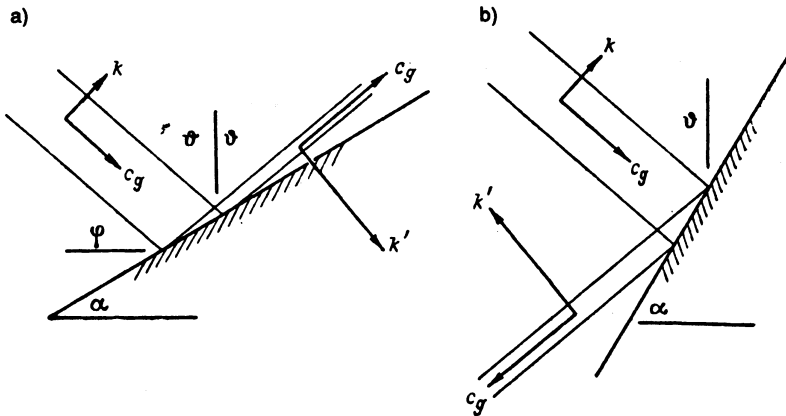


Fig. 27.1. Reflection of short-wavelength internal waves from sloping bottom: (a) for $\omega > \omega_s$, (b) for $\omega < \omega_s$.

$\rho_* \approx \text{const}$ in the form

$$\frac{\partial^2 \xi}{\partial z^2} + \alpha^2 \gamma^2(z) \xi = 0; \quad \alpha^2 = k^2 \omega^2 (\omega^2 - f^2)^{-1};$$

$$\gamma^2 = \frac{N^2}{\omega^2} - 1$$
(27.4)

and let us seek its approximate solution using the WKB method (the Wentzel—Kramers—Brillouin method). To do this, we introduce the “slow” depth $\eta = \varepsilon z$, we define $\lambda = \alpha/\varepsilon$ and $\gamma_1(\eta) \equiv \gamma(\eta/\varepsilon)$, and we look for $\xi(\eta)$ in the form $A(\eta) \exp[\pm i\lambda \int_0^\eta \gamma_1(\eta) d\eta]$. Then for $A(\eta)$ from (27.4) we get the equation

$$\lambda^{-1} \frac{\partial^2 A}{\partial \eta^2} \pm 2\gamma_1 \frac{\partial A}{\partial \eta} \pm iA \frac{\partial \gamma_1}{\partial \eta} = 0. \quad (27.5)$$

Neglecting the first term (the one involving the WKB approximation in this case), we get $A \sim \gamma_1^{-1/2}$. Consequently, the approximate WKB solution of Equation (27.4) can be written as

$$\xi \approx \left(\frac{N^2}{\omega^2} - 1 \right)^{-1/4} (c_1 \cos \alpha z_1 + c_2 \sin \alpha z_1);$$

$$z_1 = \int_0^z \left(\frac{N^2}{\omega^2} - 1 \right)^{1/2} dz.$$
(27.6)

Satisfaction of boundary conditions (26.5') yields an equation giving ω as a function of k (dispersion relation). According to solution (27.6), as the wave packet approaches the depth $z = z_c$, at which $N(z_c) = \omega$ (*turning point*), the vertical wave number and the group velocity both go to zero, but the packet nevertheless reaches this depth in some finite time, and it is reflected there. Near depth $z = z_c$, the WKB approximation is no longer valid, and solution (27.6), which possesses a weak singularity at this depth, must be replaced by a more exact solution (Desobier, 1973). To accomplish this, we make the following variable change in (27.4):

$$\xi(z) = \left[\frac{3ia}{2} \int_{z_c}^z \gamma(z) dz \right]^{2/3}; \quad V(\xi) = \gamma^{1/2} \xi^{-1/4} \xi;$$

$$\frac{\partial^2 V}{\partial \xi^2} - \xi V = \left(\frac{1}{2} \frac{\xi'''}{\xi'^3} - \frac{3}{4} \frac{\xi''^2}{\xi'^4} \right) V. \quad (27.7)$$

Since $\varepsilon \sim \alpha^{2/3}$, the coefficient of V on the right-hand side of this equation is of the order of $\alpha^{-4/3}$, so that for large α the right-hand side can be equated to zero and (27.7) becomes the Airy equation. It has solutions in cylindrical functions

$$\xi^{1/2} Z_{\pm 1/3} \left(\frac{2i}{3} \xi^{3/2} \right);$$

the linearly independent solutions (Airy functions) are denoted as $\text{Ai}(\xi)$ and $\text{Bi}(\xi)$, which gives

$$\zeta \approx \gamma^{-1/2} \xi^{1/4} [C_1 \text{Ai}(\xi) + C_2 \text{Bi}(\xi)]. \quad (27.6')$$

Because near the turning point $\gamma \sim (z - z_c)^{1/2}$ and $\xi \sim z - z_c$, while $Z_v(x) \sim x^\nu$, function $\zeta(z)$, defined by formula (27.6'), no longer has a singularity at this point.

If there is a shift in the flow velocity $U(z)$, then the behavior of a wave packet in the vicinity of the critical level z_c , at which $U(z) = c$, that is, $\omega = \sigma - \mathbf{k} \cdot \mathbf{U} = 0$, will be different from the behavior at the turning point. The group velocity in this case is the sum of the local flow velocity $\mathbf{U}(z)$ and velocity $\mathbf{c}_g$ of formula (26.15). Its vertical component c_{gz} near the critical level, where $\omega \approx (\mathbf{k} \cdot \mathbf{U}')(z - z_c)$, can be represented as

$$c_{gz} = \pm \omega^2 \frac{(1 - \omega^2/N^2)^{1/2}}{kN} \sim \pm (\mathbf{k} \cdot \mathbf{U}')^2 \frac{(z - z_c)^2}{kN}. \quad (27.8)$$

From this it is seen that for $z_1 > z_c > z_2$ the wave packet moves from level z_1 to level z_2 in a time

$$t_2 - t_1 = \int_{z_1}^{z_2} \frac{dz}{c_{gz}} = \frac{kN}{(\mathbf{k} \cdot \mathbf{U}')^2} \left(\frac{1}{z_1 - z_c} + \frac{1}{z_c - z_2} \right), \quad (27.8')$$

which as $z_2 \rightarrow z_c$ increases without limit. This means that, in contrast to the turning point, the wave packet never reaches the critical level z_c (Bretherton, 1966). According to (26.15), with the approach to z_c the vertical wave number l increases without limit (as kN/ω), while the horizontal components of the group velocity approach $\mathbf{U}(z_c)$, and they stop depending on z . This constitutes a 'capture' of the waves at the critical level, so that their energy density increases (and the linear theory stops being applicable).

In the case of a horizontally inhomogeneous ocean exact solutions of the linear equations for internal waves can be found only in particular cases. For instance, for an ocean depth $h = \gamma y$, a constant coefficient $\beta^2 = (N^2 - \omega^2)(\omega^2 - f^2)^{-1}$, and no current, Wunsch (1969) found for the stream function ψ in the (y, z) plane for $\beta\gamma < 1$ the solution

$$\psi \sim e^{iq \ln(\beta^{-1}y - z)} - e^{iq \ln(\beta^{-1}y + z)}, \quad q = 2\pi n \left(\ln \frac{1 - \beta\gamma}{1 + \beta\gamma} \right)^{-1}, \quad (27.9)$$

describing the increase in wave energy upward along the underwater slope (or for $\beta\gamma > 1$, on the other hand, downward along the slope). If, however, the scales of the horizontal inhomogeneities in the ocean are much larger than those of the waves, then the ray theory is suitable to describe the latter and the WKB approximation can be used almost everywhere (one of the first to apply it was Miropol'skii, 1974, in the case $N = N(x)$).

Let us examine, following Olbers (1981), the case $N = N(x, z)$ with a horizontal plane-parallel geostrophic current $U_y = V(x, z)$, which combines a number of problems considered earlier. In this case both the component of the

wave vector along the current k_y and the 'frequency of encounters' $\sigma = \omega + k_y V$ for a fixed wave are constant. The equations of ray theory (27.1) can be solved analytically in the useful particular case when the isopycnal surfaces are inclined planes (with a constant normal $\mathbf{n} = (n_x, n_z)$, $|\mathbf{n}| = 1$), so that $V = (n_x/fn_z)\chi(\eta) + \text{const}$ and $N^2 = n_z\chi'(\eta)$, where $\eta = \mathbf{n} \cdot \boldsymbol{\xi} = n_x x + n_z z$, and $\chi(\eta)$ is an arbitrary function. Here the orthogonal (to $\mathbf{n}$) projection $\kappa_\perp = n_z k_x - n_x k_z$ of the three-dimensional wave vector $\boldsymbol{\kappa}$ is constant, and variations of $\kappa_\parallel = \mathbf{n} \cdot \boldsymbol{\kappa}$ are found from the dispersion relation (26.14), put into the form

$$(L^2 - \omega^2)\kappa_\parallel^2 + (M^2 - \omega^2)\kappa_\perp^2 + (N^2 - \omega^2)k_y^2 + 2(N^2 - f^2)n_x n_z \kappa_\parallel \kappa_\perp = 0; \quad (27.10)$$

$$L^2 = N^2 n_x^2 + f^2 n_z^2 \geq f^2; M^2 = N^2 n_z^2 + f^2 n_x^2 \leq N^2.$$

For $\omega^2 \neq L^2$ we thus have

$$\begin{aligned} \kappa_\parallel &= [-(N^2 - f^2)n_x n_z \kappa_\perp \pm (N^2 - \omega^2)^{1/2}(\omega^2 - \omega_c^2)^{1/2}\alpha] (L^2 - \omega^2)^{-1}; \\ \omega_c^2 &= L^2 \frac{k_y^2}{\alpha^2} + f^2 \frac{\kappa_\perp^2}{\alpha^2}; \quad k^2 = \frac{\omega^2 - f^2}{N^2 - \omega^2} (n_z \kappa_\parallel - n_x \kappa_\perp)^2, \end{aligned} \quad (27.11)$$

where $\alpha^2 = \kappa^2 + k_y^2 = \text{const}$. At the same time, $f^2 \leq \omega_c^2 \leq L^2$, and the inequality $\omega_c^2 \leq \omega^2 \leq N^2$, ensuring the realness of $\kappa_\parallel$, defines the wave channel for waves with constant σ , k_y , and $\kappa_\perp$ (note, too, that $\omega_c^2 \sim f^2$ for $k_y/\kappa_\perp \sim 0$ and that $\omega_c^2 \sim L^2$ for $k_y/\kappa_\perp \sim \infty$, while $L^2 \sim f^2$ for $n_z/n_x \sim \infty$ and $L^2 \sim N^2$ for $n_z/n_x \sim 0$). The components of the group velocity are obtained in the following form:

$$\begin{aligned} c_{g\parallel} &= \pm \frac{\alpha(\omega^2 - f^2)(N^2 - \omega^2)^{1/2}(\omega^2 - \omega_c^2)^{1/2}}{\omega k^2(N^2 - f^2)}; \\ c_{g\perp} &= \frac{(\omega^2 - f^2)(N^2 - \omega^2)^{1/2}}{\omega k^2(N^2 - f^2)(L^2 - \omega^2)} [-\kappa_\perp(\omega^2 - f^2)(N^2 - \omega^2)^{1/2} \pm \\ &\quad \pm n_x n_z \alpha(N^2 - f^2)(\omega^2 - \omega_c^2)^{1/2}]. \end{aligned} \quad (27.12)$$

We will use the slope $(n_z/n_x)^2 = (N^2 - L^2)(L^2 - f^2)^{-1}$ of the refraction vector $\boldsymbol{\kappa}$ (or $\mathbf{n}$) to evaluate its orientation with respect to the asymptotic cone $l = \pm k_x(N^2 - \omega^2)^{1/2}(\omega^2 - f^2)^{-1/2}$ of the dispersion surface $\omega = \omega(\boldsymbol{\kappa})$ with fixed ω and k_y . For $\omega^2 > L^2$ vector $\boldsymbol{\kappa}$ lies inside this cone, and the two solutions of (27.12) for $\mathbf{c}_g$ have different vertical directions, but the same horizontal directions; for $\omega^2 < L^2$ the opposite will be true.

Let us examine the behavior of a wave near the layers where $c_{g\parallel} \sim 0$, that is, near the boundaries of the wave channels $\omega = \pm N$; $\pm \omega_c$ and the layers $\omega = \pm L$ inside the channels. We assume that $c_{g\parallel} \sim \eta$ there, so that in view of the law (27.2) of conservation of wave action $E(\eta) = \text{const} \cdot \omega(\eta)/c_{g\parallel}(\eta) \sim \eta^{-\mu}$, and the time for the wave to progress from η_1 to η_2 turns out to be, by analogy with (27.8'), proportional to $\eta^{2-\mu}|_1^2$ for $\mu \neq 1$ and proportional to $\ln |\eta|_1^2$ for $\mu = 1$. Thus, for $\mu \geq 1$ the wave cannot reach point $\eta = 0$, and its energy density near this point increases in a nonintegrable manner, while for $\mu < 1$ point $\eta = 0$ is

reached in a finite time, and the wave either passes through it or is reflected back from it with an integrable energy density.

For $\omega^2 \sim N^2$ in the case $n_x n_z \neq 0$, from (27.11) we get $\kappa_{||} \sim (n_x/n_z)\kappa_{\perp}$, and hence $l \sim 0$ and $k_x \sim \kappa_{\perp}/n_z$, that is, the ray becomes vertical, just as in a horizontally homogeneous ocean. Here $c_{g||} \sim \eta^{1/2}$; $c_{g\perp}$, $c_{gn} \sim \eta$, so that $\mu = \frac{1}{2}$, and the planes $\omega = \pm N$ are reflecting (see Figure 27.2; in the less realistic case of a vertically homogeneous ocean $n_z \equiv 0$ we have $c_{g||} = n_x c_{gx} \sim \eta^{3/2}$, so that layers $\omega = \pm N$ become critical). For $\omega^2 \sim \omega_c^2$ let $\kappa_{\perp} \neq 0$, $k_y \neq 0$, $n_x \neq 0$, so that $\omega^2 \neq f^2$; at the same time, both values of $\kappa_{||}$ in (27.11) are equal to $-n_z \alpha^2 / n_x \kappa$ and $c_{g||} \sim \eta^{1/2}$, so that layers $\omega = \pm \omega_c$ are reached in a finite time and are reflecting ($c_{g\perp}$ and c_{gy} are finite); for $k_y \rightarrow 0$ we get $\omega_c^2 \rightarrow f^2$, but the curves $\omega^2 = \omega_c^2$ and $\omega^2 = f^2$ vanish, while for $n_x = 0$ we have a critical layer in a horizontally homogeneous ocean.

Finally, for $\omega^2 \sim L^2$ above and below each of the layers $\omega = \pm L$ two possible waves exist with given $k_{\perp}$, one of which intersects this layer and the other of which moves away from it (see Figure 27.2). The refraction vector κ is orthogonal to the layer and parallel to one of the asymptotes of the dispersion curve. Wave A, which has at the layer a finite wave vector and a nonzero group velocity, reaches the layer in a finite time and intersects it. Wave B for $\omega^2 < L^2$ and B' for $\omega^2 > L^2$ has at this layer

$$\kappa_{||} \sim -\frac{2(N^2 - f^2)n_x n_z \kappa_{\perp}}{L^2 - \omega^2} \sim \infty, \quad (27.13)$$

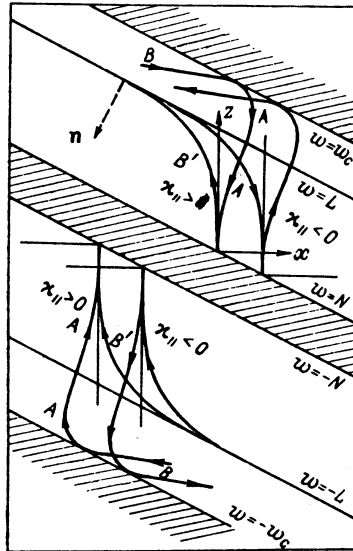


Fig. 27.2. Ray diagram for wave $(\omega, k_y, \kappa_{||})$ in projection onto (x, z) plane. Forbidden regions shaded. Reflection takes place for $\omega = \pm N, \pm \omega_c$; for $\omega = \pm L$ wave A passes, but for B' this layer turns out to be critical.

and it turns out to be critical for this wave: $k^2 \approx (L^2 - \omega^2)^{-2} \sim \eta^{-2}$; $c_{g\parallel}, c_{gy} \sim \eta^2$; $c_{g\perp} \sim \eta$. Consequently, this layer acts as a valve, admitting the wave in one direction only. For $n_x \rightarrow 0$ we return to a horizontally homogeneous ocean with a critical layer $\omega_c^2 = L^2 = f^2$, and for $n_z \rightarrow 0$ we have a vertically homogeneous ocean with a critical layer $\omega^2 = L^2 = N^2$.

We should also mention the planetary wave channel, bounded by latitudes φ_c , at which a wave of frequency ω moving toward the pole becomes reflected (the turning points φ_c are found from the equation $2\Omega \sin \varphi = \omega$, where Ω is the angular velocity of rotation of the Earth). Waves in the planetary channel with $f = f(\varphi)$ can be described in the WKB approximation, except in the vicinity of the turning points, where analogously to (27.7) for the latitude dependence of the meridional velocity component v the Airy equation is approximately true:

$$\frac{\partial^2 v}{\partial \eta^2} - \eta v = 0; \quad \eta = \frac{\varphi - \varphi_c}{L} + \left(\frac{sL}{\cos \varphi_c} \right)^2 - \frac{\omega - f}{2\Omega L \cos \varphi_c}; \quad (27.14)$$

$$L = (\gamma^2 \sin 2\varphi_c)^{-1/3}; \quad \gamma = \frac{2\pi n a \Omega}{Nh},$$

where a is the Earth's radius, and n and s are the numbers of the vertical and longitudinal modes (Munk and Phillips, 1968).

Next let us consider the propagation of internal waves of fixed frequency in an inhomogeneous ocean with $N = N(\mathbf{x}, z)$ (ignoring currents, for the time being). If $N(\mathbf{x}, z)$ has a local maximum N_0 at a point $(\mathbf{x}_0, z_0)$, then for $N_1 \leq N_0$ a surface Γ_1 with an equation $N(\mathbf{x}, z) = N_1$ will bound a region V_1 , within which a wave with a frequency $\omega = N_1$ will propagate 'indefinitely', being reflected back inside for each incidence upon boundary Γ_1 , so that V_1 will for this wave act as a trap. On the other hand, a surface Γ_1^* with an equation $N(\mathbf{x}, z) = N_1^* \geq N_0^*$, where $N_0^* = N(\mathbf{x}_0^*, z_0^*)$ is a local minimum of field $N(\mathbf{x}, z)$, will bound a region V_1^* (for example, a patch of mixed liquid, formed after the breaking of an internal wave), impenetrable to the propagation of waves outside it with a frequency $\omega = N_1^*$; when these waves are incident upon boundary Γ_1^* , they are always reflected backward. The impenetrable regions may also lie inside the trap.

The presence of traps creates spatial inhomogeneities of the spectra of internal waves: in a trap the energy spectrum should have a maximum in the frequency interval $N_1 \leq \omega \leq N_0$, since waves with such frequencies present in a trap cannot leave it, while waves of lower frequency only pass through it; at different points in space, even nearby ones, these maxima on the spectra may be different. In the neighborhoods of the maximum N_0 an infinite number of traps, located one inside the other, will appear for waves with a continuous range of frequencies $N_k < N_{k-1} < \dots < N_1 < N_0$, and in each layer $N_k < N_{k-1}$ the spectrum is enriched with frequencies ω from the interval $N_k \leq \omega \leq N_{k-1}$; at the same time, in the traps standing waves or quasistanding waves may sometimes appear. A trap may break up into parts, constituting traps for waves of different frequencies.

The statistical regime of waves propagating inside a trap has been investigated (within the framework of the ray theory) by Bunimovich (1980). Following him, let

us consider first the two-dimensional problem (x, z) . If l is the distance along the trap boundary from some fixed point on it (so normalized that the whole length of the boundary is equal to unity); $\theta = 1, 2$ are the numbers of two rays along which a wave of fixed frequency $\omega = \pm N_0 \kappa \nu / \kappa$ may move from a boundary point inside the trap (ν is the unit horizontal vector); r is the distance along the ray from a point inside the trap to the point of closest reflection of the ray from the boundary. The dependence of 'coordinates' (l, θ, r) on time will be described by some dynamic system, concerning which we assume that it retains an invariant measure μ in its phase space. Let us also carry out a time change such that the motion along all the trajectories of the dynamic system takes place with unit velocity.

Let us consider further a transform T of a circle homeomorphic to a closed curve, specified in phase space by the condition $r = 0, \theta = 1$. This transform, generated by the indicated dynamic system and similar to the mapping of a Poincaré sequence (see Section 15 above), juxtaposes to each point (l_1, θ_1) a point $(l_2, \theta_2) = T(l_1, \theta_1)$ at which is incident a trajectory emerging from l_1 along ray θ_1 at its second incidence upon the boundary (at the first, this trajectory lies on curve θ_2). Assume, for instance, a trap, that is, a circle inside of which $N = N_0$ is constant, while at the boundary N drops suddenly to the value N_1 . Then T reduces to a rotation of the circle at twice the angle α between the two directions of possible wave motion (Figure 27.3) and conserves the Lebesgue measure generated by the length. If angle α is incommensurable with π , then transform T is ergodic (and the ray during the course of time passes through all the points of the trap); in the opposite case, all the trajectories are periodic with the same period (Sinai, 1979), and standing waves may appear. Let T_* be an onto mapping of line R^1 , with projection of the line onto the circle $x \rightarrow \exp(2\pi i x)$ changing into a homeomorphism T conserving the orientation; then $\tau(T) = \lim_{|n| \rightarrow \infty} n^{-1} T_*^n(x)$ is known as the rotation number; in our example $\tau = \alpha$. Rationality of τ is equivalent to the presence of a periodic point in T .

If the boundary of the trap consists of horizontal and vertical line segments (while the trap interior may contain regions impenetrable to waves with boundaries of the same type), then the dynamic system is in general ergodic. If the trap is convex and has a smooth boundary, then for an irrational rotation number $\tau(T)$

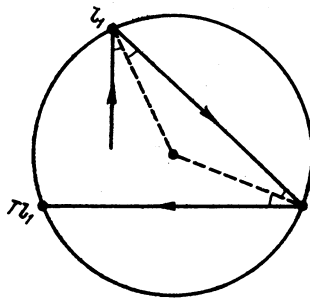


Fig. 27.3. Transform T of boundary of plane trap.

the system possesses an absolutely continuous invariant measure and is ergodic (but does not exhibit mixing, according to Kochergin, 1972), whereas for rational τ there are a finite number of periodic trajectories, to which all the others approach and on which the invariant measure is concentrated. If the trap is not convex, then it disintegrates into parts not joined by trajectories.

In the three-dimensional case there exists an entire cone of possible directions of motion of a wave of frequency ω . If the trap is a sphere, then T reduces to a rotation of each great circle perpendicular to the horizontal plane by an angle 2γ , where γ is the angle at the vertex of the intersection between the cone and the vertical plane. If the boundary of the trap consists of vertical and horizontal planes, then the motion takes place in a vertical section of the trap and is, as a rule, ergodic. In general the motion will apparently be close to ergodic for a long time. In the presence of a current with a slight shear, the trap moves along the current and because of the Doppler shift of the wave frequency is deformed somewhat.

Variations of the bottom relief $h(\mathbf{x})$, and of the fields $N(\mathbf{x}, z, t)$ and $\mathbf{U}(\mathbf{x}, z, t)$, with the formation of turning points, critical layers, and traps (characterized by predominantly ergodic motion of the waves in them), together with numerous processes of generation, dissipation, and nonlinear interactions of internal waves, must lead to sharp inhomogeneities of their distribution in space and time (and probably to a relatively infrequent formation of vertically standing waves).

28. STATIONARY WAVES

Let us consider here plane stationary internal waves, propagating in the (x, z) plane in the horizontal x direction, so that all their characteristics depend only on two variables $x' = x - ct$ and z (in the following only these two coordinates x' and z will be used, the prime of x' being dropped). Employing the nondivergence of the velocity field specified in (26.1), let us introduce the stream function $\psi(x, z)$, and let us set $u = -\partial\psi/\partial z$ and $w = \partial\psi/\partial x$, which yields the following individual derivative of any quantity A :

$$\frac{dA}{dt} = (u - c) \frac{\partial A}{\partial x} + w \frac{\partial A}{\partial z} = \frac{\partial(\Psi, A)}{\partial(x, z)}; \quad \Psi = \psi + cz. \quad (28.1)$$

Thus we see that any quantity A which is conserved in the fluid particles (that is, for which $dA/dt = 0$) may be a function solely of Ψ . In particular, this holds true for the *potential density* $\rho_* = \rho_*(\Psi)$, which is conserved during adiabatic processes. This exact condition will be used here instead of the approximate latter equation of (26.1). We will consider only short or high-frequency waves, for which the Earth's rotation can be ignored. The equations of motion (26.1) then become

$$\begin{aligned} \rho_* \left[(u - c) \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z} \right] &= - \frac{\partial p}{\partial x}; \\ \rho_* \left[(u - c) \frac{\partial w}{\partial x} + w \frac{\partial w}{\partial z} \right] &= - \frac{\partial p}{\partial z} - g\rho'_*, \end{aligned} \quad (28.2)$$

where p , normalized as in (3.5), is the nonhydrostatic part of the pressure, and $\rho'_* = -\eta''$ can be interpreted as the nonhydrostatic part of the potential density (it is not difficult to derive an exact equation for it, but we need only the equation $\partial\rho'_*/\partial x \approx \partial\rho_*/\partial x$). The pressure can be eliminated from (28.2) by cross differentiation, to give

$$\partial \left\{ \Psi, \Delta\Psi + \frac{\partial \ln \rho_*}{\partial \Psi} \left[\frac{1}{2} \left(\frac{\partial \Psi}{\partial x} \right)^2 + \frac{1}{2} \left(\frac{\partial \Psi}{\partial z} \right)^2 + gz \right] \right\} / \partial(x, z) = 0, \quad (28.3)$$

from which we have

$$\Delta\Psi + \frac{\partial \ln \rho_*}{\partial \Psi} \left[\frac{1}{2} \left(\frac{\partial \Psi}{\partial x} \right)^2 + \frac{1}{2} \left(\frac{\partial \Psi}{\partial z} \right)^2 + gz \right] = F'(\Psi), \quad (28.4)$$

where $F(\Psi)$ is an arbitrary function (Dubrais—Jacotin, 1932). This equation can be simplified considerably if we change from ψ to the new variable $\Phi = \Phi(\Psi)$, defined by the relation $\partial\Phi/\partial\Psi = \rho_*^{1/2}$:

$$\Delta\Phi + gz \frac{\partial \rho_*}{\partial \Phi} = \rho_* \frac{\partial F}{\partial \Phi}. \quad (28.5)$$

If we so choose the arbitrary functions $\rho_*(\Phi)$ and $F(\Phi)$ in (28.5) that, for instance, $\rho_* \partial F/\partial \Phi = -a^2\Phi$ and $\partial\rho_*/\partial\Phi = -b$, where a and b are constants, then for Φ it is easy to obtain wave solutions. For example, for the boundary conditions $w = 0$, and thus also $\partial\Phi/\partial x = 0$ for $z = 0, -h$, the following functions can serve as these solutions:

$$\Phi_n = \frac{bgz}{a^2} + A_n \sin \frac{\pi n z}{h} \sin k_n x; \quad k_n^2 = a^2 - \left(\frac{\pi n}{h} \right)^2 \quad (28.6)$$

(Magaard, 1965). For $n = 1$ in the upper layer of the ocean at the isopycnal we get waves with flat tops and sharp troughs, whereas in the near-bottom layer the opposite is true.

Function $F(\Psi)$ in (28.4) is uniquely defined in the case of a soliton, for which it is advisable to change from the argument Ψ of functions ρ_* and F to the argument $s = \Psi/c = z + \psi/c$, so that Equation (28.4) becomes

$$\begin{aligned} \Delta\psi + \frac{1}{c} \frac{\partial \ln \rho_*}{\partial s} \left[\frac{1}{2} \left(\frac{\partial \psi}{\partial x} \right)^2 + \frac{1}{2} \left(\frac{\partial \psi}{\partial z} + c \right)^2 + gz \right] \\ = \frac{1}{c} \frac{\partial F}{\partial s}. \end{aligned} \quad (28.7)$$

For a soliton, as $|x| \rightarrow \infty$ we have $\psi \rightarrow 0$ and $\rho_* \rightarrow \rho_{*0}(z)$, and so at this limit

(28.7) becomes

$$\frac{\partial F}{\partial z} = - \left(\frac{c^2}{2g} + z \right) N^2(z); \quad N^2 = -g \frac{\partial \ln \rho_{*0}}{\partial z}. \quad (28.8)$$

Now, using this function F , we can put Equation (28.7) for a soliton into the form

$$\begin{aligned} \Delta \psi + \left\{ \frac{\psi}{c^2} - \frac{1}{2gc} \left[\left(\frac{\partial \psi}{\partial x} \right)^2 + \right. \right. \\ \left. \left. + \left(\frac{\partial \psi}{\partial z} + c \right)^2 - c^2 \right] \right\} N^2 \left(z + \frac{\psi}{c} \right) = 0 \end{aligned} \quad (28.9)$$

(Leonov and Miropol'skii, 1974). In the Boussinesq approximation (when ρ_* on the left-hand sides of Equations (28.2) is taken to be quasiconstant) the expression in square brackets in (28.9) vanishes.

Let us consider a long-wave asymptote of the solution of Equation (28.9), that is, wave solutions with a wavelength λ much larger than the ocean depth h , so that $\varepsilon = h/\lambda$ is small (for instance, propagation of long internal waves in a shallow sea of the Baltic type). Let us measure x in units of λ , the depth z in units of h , stream function ψ in units of ch , and frequency N in units of its maximum value N_0 . Retaining for the dimensionless quantities measured according to these scales the same notation as for the initial dimensional quantities, we can rewrite (28.9) as

$$\begin{aligned} \varepsilon^2 \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial z^2} + \left\{ \beta \psi - \frac{\mu}{2} \left[\varepsilon^2 \left(\frac{\partial \psi}{\partial x} \right)^2 + \right. \right. \\ \left. \left. + \left(\frac{\partial \psi}{\partial z} \right)^2 + 2 \frac{\partial \psi}{\partial z} \right] \right\} N^2(z + \psi) = 0, \end{aligned} \quad (28.9')$$

where $\beta = N_0^2 h^2 / c^2$ is the dimensionless inverse square of the velocity of wave propagation, while $\mu = N_0^2 h / g$ is the dimensionless inverse of the free-fall acceleration. Let us impose, for example, boundary conditions $\partial \psi / \partial x = 0$ for $z = 0, -1$. We will seek a solution in the form $\psi = \varepsilon^2 \psi_1 + \varepsilon^4 \psi_2 + \dots$, $\beta = \beta_0 + \varepsilon^2 \beta_1 + \dots$, so that from (28.9') we obtain

$$\begin{aligned} \psi_1'' - \mu N^2 \psi_1' + \beta_0 N^2 \psi_1 = 0; \quad \psi_2'' - \mu N^2 \psi_2' + \beta_0 N^2 \psi_2 = Q; \\ Q = - \left[\frac{\partial^2 \psi_1}{\partial x^2} + \beta_1 N^2 \psi_1 + \beta_0 (N^2)' \psi_1^2 - \right. \\ \left. - \mu (N^2)' \psi_1 \psi_1' - \frac{\mu}{2} N^2 \psi_1'^2 \right], \end{aligned} \quad (28.10)$$

where a prime indicates differentiation with respect to z , and $N^2 = N^2(z)$.

The first of these equations has a solution with separated variables $\psi_1 = \eta(x)\xi(z)$, where $\xi(z)$ satisfies the same equation and zero boundary conditions, while $\eta(z)$ can be found from the condition stating that the second equation of (28.10) must not contain secular terms, that is, its right-hand side Q is orthogonal to $\xi(z)$. The condition $\int_{-1}^0 \xi Q \rho_* dz = 0$ with the normalization $\int_{-1}^0 \xi \rho_*^2 dz = 1$ (where ρ_* is assumed to be dimensionless) yields

$$\frac{\partial^2 \eta}{\partial x^2} + \gamma \eta^2 + \alpha \beta_1 \eta = 0; \quad \alpha = \int_{-1}^0 N^2 \xi^2 \rho_* dz;$$

$$\gamma = \int_{-1}^0 \left[\beta_0 (N^2)' \xi^2 - \mu (N^2)' \xi \xi' - \frac{\mu}{2} N^2 \xi'^2 \right] \rho_* \xi dz. \quad (28.11)$$

This equation can be obtained by single integration over x of the steady-state KdV equation (or, conversely, by differentiation with respect to x of the Benjamin—Lighthill equation (22.10) for $S = 0$), and thus it has the same solutions: cnoidal waves (22.14) and solitons (22.15), any sign of the latter being possible (it is determined by the sign of γ). By defining $\eta_0 = -3\alpha\beta_1/2\gamma$, we arrive at this form of the equation:

$$\gamma^{-2} \frac{\partial^2 \eta}{\partial x^2} = -\frac{\partial U}{\partial \eta}; \quad U = \frac{1}{3} \eta^2 (\eta - \eta_0), \quad (28.11')$$

and thus it is exactly analogous to the equation of motion of a particle with a mass γ^{-1} and a coordinate η (with a 'time' x) in a force field having a potential $U(\eta)$, $0 \leq \eta \leq \eta_0$. Figure 28.1 shows the 'potential well' $U(\eta)$ and the phase trajectories of Equation (28.11') in the (η, η') plane. The center $U \approx U_{\min}$ (that is, $\eta \approx \frac{2}{3}\eta_0$) corresponds to infinitesimal waves, and in order to give them the form $\eta \sim e^{ix}$ we have to set $\frac{2}{3}\eta_0\gamma = 1$. With increasing amplitude, the cnoidal waves (closed cycles in Figure 28.1) become increasingly asymmetrical: for $\gamma > 0$ the 'particle' is retarded more and more in the region of low η , whereas for high η it 'moves out fast'. The soliton constitutes the limit (loop of separatrix in Figure

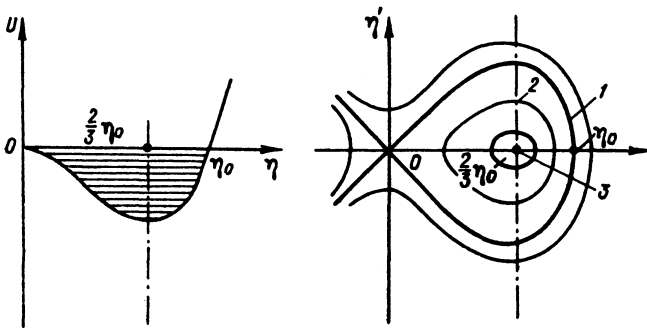


Fig. 28.1. Potential $U(\eta)$ and phase trajectories (η, η') of KdV equation (28.11'): (1) soliton, (2) cnoidal waves, (3) infinitesimal waves.

28.1), with $\eta = \eta_0 \text{ch}^{-2}(x/2)$. Note that infinitesimal waves propagate at 'subsonic' velocities $c < (gh)^{1/2}$, which is an effect of their negative dispersion; positive solitons are 'supersonic' and negative solitons are 'subsonic'.

Similarly to the derivation of Equations (22.19) and (28.11), for long internal waves we derive the unsteady-state KdV equation and then use it to show that stationary long waves are stable with respect to small *longitudinal* perturbations (that is, not prone to *self-modulation*). For these waves a two-dimensional generalization of the KdV equation is also derived (Leonov, 1976). It is an analog of the Kadomtsev—Petviashvili equation (22.19'). In this case it indicates that long internal waves are unstable with respect to transverse perturbations (phenomenon of *self-focusing*). For a complex envelope of a slightly nonlinear packet of internal waves with a narrow spectrum, a nonlinear Schrödinger equation (22.29) is derived. The latter is used to establish a criterion for self-modulation instability (in the case of a three-dimensional wave packet, this criterion is more complicated than for surface waves; see Miropol'skii, 1975).

Now let us apply Equation (28.9') to the case of a thin pycnocline (thickness $2h \ll \lambda$), which can be idealized as being surrounded by an unbounded unstratified fluid (here z should be reckoned from the midpoint of the pycnocline). At its wavy boundaries $z = \psi \pm 1$ quantities ψ and $\partial\psi/\partial z$ must be continuous (while as $z \rightarrow \pm\infty$ these quantities must go to zero). Let us seek the solution of Equation (28.9') in the pycnocline in the form of series $\psi = \varepsilon\psi_1 + \varepsilon^2\psi_2 \dots$, $\beta = \beta_0 + \varepsilon\beta_1 + \dots$ and outside of the pycnocline in the same form, but only after changing to stretched vertical coordinates $z_{\pm} = \varepsilon(z \pm 1)$, where the upper sign pertains to the layer above the pycnocline and the lower sign to the layer below it; here for $\psi_{\pm}(x, z_{\pm})$ we have $\partial^2\psi_{\pm}/\partial x^2 + \partial^2\psi_{\pm}/\partial z_{\pm}^2 = 0$. In the pycnocline we obtain for ψ_1 and ψ_2 the same equations (28.10) as above, except that now the term $\partial^2\psi_1/\partial x^2$ of Q vanishes. The boundary conditions for ψ_1 will now have the form $\psi_1 = \psi_{1\pm}$ and $\psi'_1 = 0$ for $z = \pm 1$, while for ψ_2 we get $\psi_2 = \psi_{2\pm}$ and $\psi'_2 = Q_{\pm} \equiv \partial\psi_{1\pm}/\partial z + \beta_0 N^2 \psi_1^2$ for $z = \pm 1$ and $z_{\pm} = 0$. The solution for ψ_1 in the pycnocline may, as above, be taken in the form $\psi_1 = \eta(x)\xi(z)$, where $\xi(z)$ satisfies the same equation and boundary conditions $\xi'(\pm 1) = 0$. The equation for ψ_2 can be written as

$$(\rho_*\psi'_2)' + \beta_0\rho_*N^2\psi_2 = \rho_*Q; \quad \psi'_2(z = \pm 1) = Q_{\pm}. \quad (28.12)$$

This equation can be put into the form of a nonhomogeneous equation for function $F = \rho_*\psi'_2 - \psi_0$ with homogeneous boundary conditions $F(z = \pm 1) = 0$, if we set

$$\psi_0 = \frac{\rho_+Q_+ - \rho_-Q_-}{\rho_+ - \rho_-} [\rho_*(z) - \rho_-] + \rho_-Q_-; \quad \rho_{\pm} = \rho_*(\pm 1); \quad (28.12')$$

it is easy to see that the condition for solvability of this nonhomogeneous equation for F reduces to

$$\xi_+Q_+\rho_+ - \xi_-Q_-\rho_- - \int_{-1}^{+1} \xi Q \rho_* dz = 0; \quad \xi_{\pm} = \xi(\pm 1). \quad (28.13)$$

Substitution here of the above-indicated expressions for $Q(x, z)$ and $Q_{\pm}(x)$ then yields

$$\rho_+ \zeta_+^2 \left(\frac{\partial \Phi}{\partial z} \right)_{z=+0} - \rho_- \zeta_-^2 \left(\frac{\partial \Phi}{\partial z} \right)_{z=-0} + \gamma \eta^2 + \alpha \beta_1 \eta = 0, \quad (28.13')$$

where $\Phi(x, z)$ is the solution of the Laplace equation $\Delta \Phi = 0$, which goes to zero as $z \rightarrow \pm \infty$ and becomes η for $z = 0$. This solution has the form of the potential of a double layer

$$\Phi_{\pm} = \pm \frac{z}{\pi} \int_{-\infty}^{\infty} \frac{\eta(x') dx'}{(x - x')^2 + z^2};$$

$$\left(\frac{\partial \Phi}{\partial z} \right)_{z=\pm 0} = \mp \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{\eta(x') dx'}{x - x'}, \quad (28.14)$$

so that Equation (28.13') finally becomes

$$- \frac{r}{\pi} \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{\eta(x') dx'}{x - x'} + \gamma \eta^2 + \alpha \beta_1 \eta = 0; \quad (28.15)$$

$$r = \rho_+ \zeta_+^2 + \rho_- \zeta_-^2$$

(Benjamin, 1967; Ono, 1975). It differs from the stationary KdV equation (28.11) in that the differential dispersion operator is replaced by an integrodifferential operator (with an integral Hilbert transform). The corresponding nonstationary equation (Benjamin—Ono equation) has the form

$$\frac{\partial \eta}{\partial t} + c_0 \frac{\partial \eta}{\partial z} + \alpha \eta \frac{\partial \eta}{\partial x} - \frac{r}{\pi} \frac{\partial^2}{\partial x^2} \int_{-\infty}^{\infty} \frac{\eta(x', t) dx'}{x - x'} = 0, \quad (28.15')$$

and its linearized variant corresponds to the dispersion relation $\omega = c_0 k - rk|k|$ (showing, incidentally, that $c = c_0 - r|k| < c_0$, that is, as in the KdV case, the infinitesimal waves are 'subsonic'); operator $rk|k|$ also produces integral dispersion. The solutions of Equation (28.15) include both periodic waves and solitons. The latter have the very simple form $\eta = \eta_m(1 + x^2/L^2)^{-1}$, and their amplitude $\eta_m = -2\alpha\beta_1\gamma^{-1}$, while their halfwidth $|L| = -r/\alpha\beta_1$ (so that $\beta_1 < 0$, that is, all the solitons are 'supersonic', their dimensionless relative velocity $\gamma\eta_m/2 = -\alpha\beta_1$, and their sign is the same as the sign of γ). Such solitons were produced in the laboratory so easily by Davis and Akrivos (1967) that it stands to reason that they must occur frequently in nature.

29. GENERATION OF MICROSTRUCTURE

Let us consider, as in the preceding section, plane waves propagating in the (x, z) plane in the x direction, but in this case let the waves be nonstationary. They are

described by the nonstationary analogs of Equations (28.2) (differing from the latter by the substitution $-c \partial/\partial x \rightarrow \partial/\partial t$) and by the equation $d\rho_*/dt = 0$. Using these formulas, after introducing the stream function $\psi(x, z, t)$ in the Boussinesq approximation (to which we will confine ourselves here), we obtain

$$\frac{\partial \Delta \psi}{\partial t} + \frac{\partial(\psi, \Delta \psi)}{\partial(x, z)} = -g \frac{\partial \ln \rho_*}{\partial x}; \quad \frac{\partial \rho_*}{\partial t} + \frac{\partial(\psi, \rho_*)}{\partial(x, z)} = 0. \quad (29.1)$$

We consider a packet of slightly nonlinear waves, for which the ray-theory approximation is suitable, so that ψ and ρ_* can be represented as functions of a 'slow' abscissa $\xi = \varepsilon^2 x$ and a 'slow' time $\tau = \varepsilon^2 t$, the usual depth z , and a 'fast' phase $\theta = \varepsilon^{-2} \vartheta(\xi, \tau)$ (where ε is a small parameter of the order of the wave steepness) and can be expanded in series:

$$\psi = \sum_1^\infty \varepsilon^n \psi_n(\xi, \tau, z, \theta) \quad \text{and} \quad \rho_* = \rho_0(z) + \sum_1^n \varepsilon^n \rho_n(\xi, \tau, z, \theta)$$

(Borisenko *et al.*, 1976). Let us introduce the horizontal wave number $k = \partial \vartheta / \partial \xi = \partial \theta / \partial x$ and the frequency $\omega = -\partial \vartheta / \partial \tau = -\partial \theta / \partial t$, so that the equation of conservation of wave number $\partial k / \partial \tau + \partial \omega / \partial \xi = 0$ is valid. Substitution of the indicated series into (29.1) yields

$$\begin{aligned} \omega^2 \frac{\partial}{\partial \theta} \left(k^2 \frac{\partial^2 \psi_n}{\partial \theta^2} + \frac{\partial^2 \psi_n}{\partial z^2} \right) + k^2 N^2 \frac{\partial \psi_n}{\partial \theta} &= Q_n; \\ \omega \frac{\partial \rho_n}{\partial \theta} - k \frac{\partial \rho_0}{\partial z} \frac{\partial \psi_n}{\partial \theta} &= S_n, \end{aligned} \quad (29.2)$$

with $Q_1 = S_1 = 0$, and for ψ_1 and ρ_1 we arrive at solutions with separated variables:

$$\begin{aligned} \psi_1 &= \eta_1(\xi, \tau, \theta) \zeta_1(z); \quad \rho_1 = \eta_1 \frac{k}{\omega} \frac{\partial \rho_0}{\partial z} \zeta_1; \\ \eta_1 &= A(\xi, \tau) e^{i\theta} + A^*(\xi, \tau) e^{-i\theta}; \quad \zeta_1'' + \frac{N^2 - \omega^2}{\omega^2} k^2 \zeta_1 = 0, \end{aligned} \quad (29.3)$$

where A is the complex envelope, and the boundary conditions for $\zeta_1(z)$ can be taken, for instance, in the form $\zeta_1(0) = \zeta_1(-h) = 0$. In the second approximation to the wave solutions with separated variables we have to add nonlinear corrections of the order of $|A|^2$, analogous in their nature to the Stokes current in the theory of surface waves:

$$\begin{aligned} \psi_2 &= \eta_2(\xi, \tau, \theta) \zeta_2(z) + |A|^2 \psi(z); \\ \rho_2 &= \eta_2 \left(\frac{k}{\omega} \frac{\partial \rho_0}{\partial z} \zeta_2 + \frac{k^2}{2\omega^2} \frac{\partial^2 \rho_0}{\partial z^2} \zeta_1^2 \right) + |A|^2 R(z); \\ \eta_2 &= A^2 e^{2i\theta} + A^{*2} e^{-2i\theta}; \quad \zeta_2'' + \frac{N^2 - 4\omega^2}{\omega^2} k^2 \zeta_2 = \frac{k^2}{\omega^2} (N^2)' \zeta_1^2, \end{aligned} \quad (29.4)$$

with the same zero boundary conditions for $\xi_2(z)$.

The equation for the complex envelope $A = \frac{1}{2}a \exp(ib)$ (where a and b are real) is derived with the requirement that Equations (29.2) for ψ_3 and ρ_3 do not contain secular terms. The resulting expressions are

$$\frac{\partial A}{\partial \tau} + c_g \frac{\partial A}{\partial \xi} + \frac{1}{2} \frac{\partial c_g}{\partial k} \frac{\partial k}{\partial \xi} A = 4il(k)A^2 A^*, \quad (29.5)$$

$$\frac{\partial a^2}{\partial \tau} + \frac{\partial c_g a^2}{\partial \xi} = 0; \quad \frac{\partial b}{\partial \tau} + c_g \frac{\partial b}{\partial \xi} = I(k)a^2, \quad (29.5')$$

where $c_g = \partial\omega_1/\partial k$ is the group velocity of the waves in the linear approximation, that is, for the dispersion relation $\omega = \omega_1(k)$ obtained from boundary-value problem (29.3) for $\xi_1(z)$ (these equations are valid for a detuning of the wave number in the packet $\delta k/k$ of the order of ε^2 ; if, however, it is of the order of ε , then we must use a different, 'slow', abscissa $\xi = \varepsilon(x - c_g t)$, so that the left-hand side of (29.5) becomes $\partial A/\partial \tau - (i/2)(\partial c_g/\partial k)\partial^2 A/\partial \xi^2$, meaning that we arrive at a nonlinear Schrödinger equation for A). Equations (29.5') are direct corollaries of (29.5); the first of these represents the conservation of energy. These equations imply that the nonlinear dispersion relation has the form $\omega = \omega_1(k) - \varepsilon^2 a^2 I(k)$. The function $I(k)$ entering into the latter formulas is expressed in terms of $N(z)$ and the factors $\Psi(z)$ and $R(z)$ in the nonlinear corrections, the equations for which are obtained from the requirement that Equations (29.2) for ψ_4 and ρ_4 do not contain secular terms. In the particular case when in the linear approximation a monochromatic wave is specified ($\omega = \text{const}$, $k = \text{const}$, $c = \omega/k$), these equations have the following form:

$$\begin{aligned} \Psi'' + c_g^{-2} N^2 \Psi &= (2c^{-3} - c^{-2} c_g^{-1} - c^{-1} c_g^{-2}) N^2 (\xi_1^2)' + 2c^{-3} \xi_1^2 (N^2)'; \\ \frac{g}{\rho_0} R &= c_g^{-1} N^2 \Psi + (c^{-1} c_g^{-1} - c^{-2}) N^2 (\xi_1^2)' - c^{-2} \xi_1^2 (N^2)'; \\ 4I &= \left(\int_{-h}^0 N^2 \xi_1^2 dz \right)^{-1} \left[k \int_{-h}^0 (N^2)' \xi_1^2 \xi_2 dz - \right. \\ &\quad \left. - \frac{\omega^2}{k} \int_{-h}^0 (\xi_1'^2 + k^2 \xi_1^2) \Psi dz - \frac{g\omega}{\rho_0} \int_{-h}^0 \xi_1 \xi_1' R dz \right], \end{aligned} \quad (29.6)$$

with the same zero boundary conditions for $\Psi(z)$ as above. If we know $I(k)$, we can evaluate the stability of the waves according to the nonlinear dispersion relation: for $(\partial c_g/\partial k)I(k) > 0$ the wave is unstable with respect to longitudinal perturbations (self-modulation; except for the degenerate case $N = \text{const}$, long waves are unable to self-modulate); for $I(k) > 0$ the wave is unstable with respect to transverse perturbations (self-focusing).

Consequently, in the considered approximation (29.3), (29.4) a wave packet contains a fundamental wave, a second harmonic, and a nonlinear contribution to the mean current. When the packet encounters an immobile medium, the fluid particles acquire mean horizontal velocities $\varepsilon^2 |A|^2 \Psi'(z)$ and, in addition, become shifted along the vertical over a distance $\varepsilon^2 |A|^2 R(z) (\partial \rho_0/\partial z)^{-1}$. Since the total

flow through a cross section of the ocean is zero, the ocean is divided into layers with alternating signs of the horizontal velocity (and of the density perturbations); thus a *vertical thinly laminated microstructure* appears. From (29.6) we see that its wavelength along the vertical is of the order of $c_g N^{-1}$, and that for small lengths of the carrier wave ($k \rightarrow \infty$, and thus $c_g \rightarrow 0$, in view of the dispersion relation for internal waves) it is small. Moreover, the right-hand sides of Equations (29.6) oscillate vertically with a wavelength inversely proportional to the number of the mode.

After the wave packet has passed (after a time Lc_g^{-1} of the order of a day to a week, where L , its length, is of the order of 1 to 10 km), the fluid particles return to their initial levels (although they have been shifted horizontally over a distance $\varepsilon^2 |A|^2 \Psi'(z) L c_g^{-1}$) and the microstructure disappears (provided that in it a secondary instability did not appear, that is, the onset of turbulence in regions with $Ri < \frac{1}{4}$, or even the breaking of ultrashort internal waves, which can have irreversible effects).

30. HAMILTONIAN FORMALISM

Following Voronovich (1979), let us introduce for internal waves in the Boussinesq approximation (although we could just as easily do without it) the Hamiltonian

$$\mathcal{H} = \bar{E}; \quad E = \int_{-h}^0 \left[\frac{|\mathbf{u}|^2}{2} + \pi(\rho, z) \right] dz; \\ \mathbf{u} = \nabla \varphi + \lambda \nabla(\rho_0 + \rho); \quad (30.1)$$

$$\pi(\rho, z) = \rho g z + g \int_{\rho_0 + \rho}^{\rho_0} \eta(R) dR = -\frac{g}{\rho'_0} \frac{\rho^2}{2!} + \frac{g \rho''_0}{\rho'^3_0} \frac{\rho^3}{3!} + \dots,$$

where we have neglected the Earth's rotation and assumed the boundary condition $w \rightarrow 0$ ($z \rightarrow 0$); for $\mathbf{u}$ in (30.1) we have used the Clebsch representation; $\rho_0 + \rho$ is the dimensionless potential density, measured in units of the mean standard density ρ_w of seawater, $\rho_0 = \rho_0(z)$ being its equilibrium value; $\eta(R)$ is a function which is the inverse of $\rho_0(z)$, that is, defined by the condition $\eta[\rho_0(z)] \equiv z$; the quantity $\rho_w \pi(\rho, z)$ is the potential energy density, equal to the work done to move through an unperturbed medium a fluid particle with a potential density $\rho_0 + \rho$ from its equilibrium level $\eta(\rho_0 + \rho)$ to a level z (retention in $\pi(\rho, z)$ of only the terms given here in ρ^2 and ρ^3 is sufficient to approximate the weak nonlinearity); $\rho_w E$ is the total (kinetic + potential) energy of a vertical water column of unit cross section; $\bar{E}$ is a quantity averaged over the horizontal. Functions λ and ρ are the canonical variables in Hamiltonian $\mathcal{H}$. Function φ is expressed in terms of the latter with the aid of the condition $\text{div } \mathbf{u} = 0$, that is, from the equation $\Delta \varphi = -\nabla[\lambda \nabla(\rho_0 + \rho)]$, or

$$\frac{\partial \lambda}{\partial t} = \frac{\delta \mathcal{H}}{\delta \rho}; \quad \frac{\partial \rho}{\partial t} = -\frac{\delta \mathcal{H}}{\delta \lambda}. \quad (30.2)$$

In what follows it will be convenient to change from variables λ and ρ to new canonical variables: the Fourier transforms $\lambda_{\mathbf{k}}(z, t)$ and $\rho_{\mathbf{k}}^*(z, t)$ with respect to the horizontal coordinates (cf. Equations (24.1)). In view of the realness of λ and ρ , we have $\lambda_{\mathbf{k}} = \lambda_{-\mathbf{k}}^*$ and $\rho_{\mathbf{k}} = \rho_{-\mathbf{k}}^*$. If we set

$$\lambda_{\mathbf{k}} = \xi_{\mathbf{k}}(z) e^{-i\omega t}, \quad \rho_{\mathbf{k}} = \frac{i\omega\rho'_0}{g} \xi_{\mathbf{k}}(z) e^{-i\omega t},$$

then from the linearized canonical equations (obtained keeping in $\mathcal{H}$ only the terms quadratic in λ and ρ) for $\xi_{\mathbf{k}}(z)$ we arrive at a simplified Equation (26.5) in the form $\xi'' + k^2(N^2/\omega^2 - 1)\xi = 0$, where $N^2 = -g\rho'_0$. For boundary conditions $\xi = 0$ ($z = 0, -h$) it has solutions in the form of discrete modes $\omega = \omega_{nk}$, $\xi = \xi_{nk}(z)$, $n = 0, 1, 2, \dots$. In unbounded space the discrete modes are replaced by a continuum of vertical wave numbers $n \rightarrow l$, and locally we obtain $\omega_{lk} = Nk(k^2 + l^2)^{-1/2}$ and $\xi_{lk}(z) \sim e^{-ilz}$. We will normalize the eigenfunctions $\xi_{\nu k}(z)$ using the condition $\int_{-h}^0 N^2 \xi_{\mu k} \xi_{\nu k}^* dz = \delta_{\mu-\nu}$, from which we have $N^2 \int_{-\infty}^{\infty} \xi_{\nu k}(z) \xi_{\nu k}^*(z') d\nu = \delta_{z-z'}$ (in the case of discrete modes the dimensional wave numbers ν are replaced by dimensionless numbers; $\delta_{\mu-\nu}$ is replaced by $\delta_{\mu\nu}$, so that the dimension of functions $\xi_{\nu k}$ is also altered; integration over all ν is replaced by summation over all integral ν ; functions $\xi_{\nu k}$ can here be taken to be real). A change from variables $\lambda_{\mathbf{k}}, \rho_{-\mathbf{k}}$ to *normal* canonical variables $ia_{\nu\mathbf{k}}^*, a_{\nu\mathbf{k}}$ yields

$$a_{\nu\mathbf{k}}(t) = -\frac{i}{g} \left(\frac{\omega_{\nu k}}{2} \right)^{1/2} \int_{-h}^0 \left(N^2 \lambda_{\mathbf{k}} + \frac{ig^2}{\omega_{\nu k}} \rho_{\mathbf{k}} \right) \xi_{\nu k} dz; \quad (30.3)$$

$$\lambda_{\mathbf{k}}(z, t) = ig \int (2\omega_{\nu k})^{-1/2} [\xi_{\nu k}^*(z) a_{\nu\mathbf{k}}(t) - \xi_{\nu k}(z) a_{\nu, -\mathbf{k}}^*(t)] d\nu; \quad (30.3')$$

$$\rho_{\mathbf{k}}(z, t) = -\rho'_0 \int (\omega_{\nu k/2})^{1/2} [\xi_{\nu k}^*(z) a_{\nu\mathbf{k}}(t) + \xi_{\nu k}(z) a_{\nu, -\mathbf{k}}^*(t)] d\nu.$$

Then the Hamilton equations become:

$$\frac{\partial a_{\nu\mathbf{k}}}{\partial t} = -i \frac{\delta \mathcal{H}}{\delta a_{\nu\mathbf{k}}^*}; \quad \mathcal{H} = \mathcal{H}_2 + \mathcal{H}_3 + \dots; \quad (30.4)$$

$$\mathcal{H}_2 = \int \omega_{\nu k} a_{\nu\mathbf{k}} a_{\nu\mathbf{k}}^* d\mathbf{k} d\nu;$$

$$\begin{aligned} \mathcal{H}_3 = & \int (V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{\nu\nu_1\nu_2} a_{\nu\mathbf{k}}^* a_{\nu_1\mathbf{k}_1} a_{\nu_2\mathbf{k}_2} + V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{\nu\nu_1\nu_2} a_{\nu\mathbf{k}} a_{\nu_1\mathbf{k}_1}^* a_{\nu_2\mathbf{k}_2}^*) \delta_{\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2} \times \\ & \times d\mathbf{k} d\mathbf{k}_1 d\mathbf{k}_2 d\nu d\nu_1 d\nu_2 + \frac{1}{3} \int (U_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{\nu\nu_1\nu_2} a_{\nu\mathbf{k}} a_{\nu_1\mathbf{k}_1} a_{\nu_2\mathbf{k}_2} + \\ & + U_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{\nu\nu_1\nu_2} a_{\nu\mathbf{k}}^* a_{\nu_1\mathbf{k}_1}^* a_{\nu_2\mathbf{k}_2}^*) \delta_{\mathbf{k}+\mathbf{k}_1+\mathbf{k}_2} d\mathbf{k} d\mathbf{k}_1 d\mathbf{k}_2 d\nu d\nu_1 d\nu_2, \end{aligned}$$

where the coefficients of the three-wave interactions are found from the formulas

$$\begin{aligned}
 (V, U)_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{v_1v_2} = & \frac{1}{8\pi} \left(\frac{\omega_{vk}\omega_{v_1k_1}\omega_{v_2k_2}}{2} \right)^{1/2} \left\{ \left[\frac{\omega_{vk}}{k^2} \left(\frac{\mathbf{k} \cdot \mathbf{k}_1}{\omega_{v_1k_1}} + \frac{\mathbf{k} \cdot \mathbf{k}_2}{\omega_{v_2k_2}} \right) - \right. \right. \\
 & \left. \left. - \frac{\omega_{vk} \mp \omega_{v_1k_1} \mp \omega_{v_2k_2}}{\omega_{vk}} \right] \int N^2 \frac{\partial(\zeta, \zeta^*)_{vk}}{\partial z} \zeta_{v_1k_1}^* \zeta_{v_2k_2}^* dz + \right. \\
 & + \left[\frac{\omega_{v_1k_1}}{k_1^2} \left(\frac{\mathbf{k}_1 \cdot \mathbf{k}}{\omega_{vk}} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{\omega_{v_2k_2}} \right) + \right. \\
 & + \left. \frac{\pm \omega_{vk} - \omega_{v_1k_1} - \omega_{v_2k_2}}{\omega_{v_1k_1}} \right] \int N^2(\zeta, \zeta^*)_{vk} \times \\
 & \times \frac{\partial \zeta_{v_1k_1}^*}{\partial z} \zeta_{v_2k_2}^* dz + \left[\frac{\omega_{v_2k_2}}{k_2^2} \left(\frac{\mathbf{k}_2 \cdot \mathbf{k}}{\omega_{vk}} + \frac{\mathbf{k}_2 \cdot \mathbf{k}_1}{\omega_{v_1k_1}} \right) - \right. \\
 & \left. \left. - \frac{\mp \omega_{vk} + \omega_{v_1k_1} + \omega_{v_2k_2}}{\omega_{v_2k_2}} \right] \int N^2(\zeta, \zeta^*)_{vk} \zeta_{v_1k_1}^* \frac{\partial \zeta_{v_2k_2}^*}{\partial z} dz \right\}. \quad (30.5)
 \end{aligned}$$

the superscripts pertaining to V and the subscripts to U . For internal waves the resonance conditions $\omega_{vk} = \omega_{v_1k_1} + \omega_{v_2k_2}$ for $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$ are satisfied, so that of the three-wave interactions described in $\mathcal{H}_3$ by the terms with coefficients V some are resonance interactions. At the same time, since all the eigenfrequencies ω_{vk} are positive, the condition $\omega_{vk} + \omega_{v_1k_1} + \omega_{v_2k_2} = 0$ is not satisfied, even for $\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2 = 0$, meaning that all the three-wave interactions described in $\mathcal{H}_3$ by the terms with coefficients U are nonresonance interactions, and these terms can be eliminated by a canonical transformation of type (24.6). After this, the Hamilton equations, with only their principal terms retained (in the so-called *weak-nonlinearity approximation*), can be written as

$$\begin{aligned}
 \frac{\partial a_{v\mathbf{k}}}{\partial t} + i\omega_{v\mathbf{k}}a_{v\mathbf{k}} = & -i \int V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{v_1v_2} a_{v_1\mathbf{k}_1} a_{v_2\mathbf{k}_2} \delta_{\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2} d\mathbf{k}_1 d\mathbf{k}_2 dv_1 dv_2 - \\
 & - 2i \int V_{\mathbf{k}_1\mathbf{k}\mathbf{k}_2}^{*v_1v_2} a_{v_1\mathbf{k}_1} a_{v_2\mathbf{k}_2}^* \delta_{\mathbf{k}_1-\mathbf{k}-\mathbf{k}_2} d\mathbf{k}_1 d\mathbf{k}_2 dv_1 dv_2. \quad (30.6)
 \end{aligned}$$

Let us examine, with the aid of these equations, a resonance interaction of three waves: the case $a_{v\mathbf{k}} = \sum_1^3 a_n(\mathbf{k}, t) e^{-i\omega_n t} \delta(\nu - \mu_n)$, where functions $a_n(\mathbf{k}, t)$, which vary slowly with t , describe narrow spectral peaks at points $\mathbf{k} = \boldsymbol{\kappa}_n$ satisfying together with the frequencies $\omega_n = \omega_{\mu_n \boldsymbol{\kappa}_n}$ the resonance conditions $\boldsymbol{\kappa}_3 = \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_2$ and $\omega_3 = \omega_1 + \omega_2$. Then for $\mathbf{k} \approx \boldsymbol{\kappa}_3$ in (30.6) the second integral goes to zero and

we have

$$\left[\frac{\partial}{\partial t} + i(\omega_{\mu_3 \mathbf{k}} - \omega_3) \right] a_3(\mathbf{k}) = -2iV \int a_1(\mathbf{k}_1) a_2(\mathbf{k} - \mathbf{k}_1) d\mathbf{k}_1, \quad (30.7)$$

where $V = V_{\kappa_3 \kappa_1 \kappa_2}^{\mu_3 \mu_1 \mu_2}$. Here we can set $\omega_{\mu_3 \mathbf{k}} - \omega_3 \approx \mathbf{c}_{g3} \cdot (\mathbf{k} - \kappa_3)$, and it is convenient to carry out the Fourier transformation $\psi_n(\kappa, t) = \int e^{i(\mathbf{k} - \kappa_n) \cdot \mathbf{x}} (a_n(\mathbf{k})/2\pi) d\mathbf{k}$ (in view of the relation $\int |\psi_n|^2 d\mathbf{x} = \int |a_n|^2 d\mathbf{k}$ and the narrowness of the spectral peaks, the quantity $\omega_n |\psi_n|^2$ will have the sense of the averaged energy $\bar{E}$ for each wave).

The factor $e^{-i\kappa_n \cdot \mathbf{x}}$ describes rapid variations in the $\mathbf{x}$ direction, so that the remaining quantity ψ_n depends on $\mathbf{x}$ slowly, and the terms $\mathbf{c}_{gn} \cdot \nabla \psi_n$ obtained by the Fourier transformation from $i\mathbf{c}_{gn} \cdot (\mathbf{k} - \kappa_n) a_n$ can be neglected. As a result, we have

$$\begin{aligned} \dot{\psi}_1 &= -4i\pi V^* \psi_2^* \psi_3; & \dot{\psi}_2 &= -4i\pi V^* \psi_3 \psi_1^*; \\ \dot{\psi}_3 &= -4i\pi V \psi_1 \psi_2, \end{aligned} \quad (30.8)$$

where the derivation of the third equation was given above, and the first and second equations are obtained similarly from (30.6) for $\mathbf{k} \approx \kappa_1$ and $\mathbf{k} \approx \kappa_2$. It should be noted, too, that multiplication of all the ψ_n by the same factor of form $e^{i\varphi}$ makes coefficient V in Equations (30.8) real, and this will be assumed in the following. Such equations were considered with regard to nonlinear optics by Blombergen (1965), they were applied to plane internal waves by Thorpe (1966), Davis and Akrivos (1967), and subsequent investigators, and to three-dimensional waves by Voronovich (1975). It can be shown that they have two integrals $\Sigma_1^3 \omega_n |\psi_n|^2 = \text{const}$ (law of conservation of the total energy of a triad of waves) and $\Sigma_1^3 \kappa_n |\psi_n|^2 = \text{const}$ (law of conservation of momentum).

If $|\psi_3| \gg |\psi_{1,2}|$ for $t = 0$, then for small t we have $\psi_3 \approx \text{const}$ and from (30.8) it follows that $\dot{\psi}_{1,2} = 4\pi V \psi_3 |\psi_{1,2}|$, so that $\psi_{1,2}$ will increase exponentially with an increment $\tau^{-1} = 4\pi V |\psi_3| \sim \omega_3 \zeta_3 l_3$, where ζ_3 is the vertical displacement, and l_3 is the vertical wave number of wave ψ_3 : it turns out to be unstable with respect to decay into two waves with lower frequencies (and the described theory is applicable if $\zeta_3 l_3$ is small). If, on the other hand, $|\psi_1| \gg |\psi_{2,3}|$ for $t = 0$, then we get $\dot{\psi}_{2,3} = -4\pi V \psi_1 |\psi_{2,3}|$, so that $\psi_{2,3}$ do not increase with time: wave ψ_1 turns out to be stable with respect to perturbations $\psi_{2,3}$, one of which has a frequency higher than ω_1 . According to Voronovich, in the case of discrete modes for decay instability of wave μ_3 it is necessary and sufficient that $\mu_3 < \max(\mu_1, \mu_2)$. We note, too, that, in view of the conservation of total energy, the exponential instability for large times changes into a periodic exchange of energy between modes.

Interactions of resonance triads in the presence of a current with a velocity shear are completely different in nature. If the velocity is $\mathbf{u}_0[\rho_0(z)]$, then in Hamiltonian (30.1) we have to replace $|\mathbf{u}|^2$ by $|\mathbf{u}_0(\rho_0 + \rho) + \mathbf{u}|^2 - |\mathbf{u}_0(\rho_0 + \rho)|^2$ (in order for it to represent the work done to change the system from the equilibrium state to the given state). Moreover, instead of $\zeta_k(z)$ we will use functions $w_{\mathbf{k}}(z)$ described

by Equation (26.12), we will use the Doppler frequencies $\omega_d \approx \omega - \mathbf{k} \cdot \mathbf{u}_0$, and we will integrate the wave action

$$I = \frac{1}{2} \int \left(\frac{N^2}{\omega_d^3} - \frac{\omega_d''}{2k^2 \omega_d^2} \right) w_{vk}^2 dz,$$

over the vertical, being careful in all instances to select for description of the waves that pair $(\mathbf{k}, \omega)$ or $(-\mathbf{k}, -\omega)$ for which $I > 0$ (for the other pair $I < 0$). Note that I appears in the denominator of the formula for $\mathbf{c}_g$ and goes to zero wherever $\partial \omega_{vk} / \partial \mathbf{k} = 0$.

If in the resonance triad the first two waves have the same sign of I , while the third has a different sign, then the resonance conditions have to be written as $\kappa_1 + \kappa_2 + \kappa_3 = 0$ and $\omega_1 + \omega_2 + \omega_3 = 0$, and in Hamiltonian $\mathcal{H}_3$ in (30.4) only terms with coefficients U are retained. Then instead of (30.8) we have the equations $\dot{\psi}_i = -2i\pi U \psi_j \psi_k$, where all three subscripts are different. They have the same two integrals of motion, but because of the different signs of ω_i this no longer signifies conservation of energy: energy exchange with the mean current is possible. For instance, there is a solution $\psi_1 = \psi_2 = \psi_3 \sim (t_0 - t)^{-1}$ which goes to infinity after a certain time (*explosive* instability; around time t_0 it is, of course, smoothed out by nonlinear effects of higher orders, but it is possibly able to produce turbulence of the flow).

Let us now introduce, by analogy with (24.10), the space spectrum of the wave action $N_{vk}(t) = F_{vk} / \omega_{vk}$ (where F_{vk} is the so-called *unsymmetrical* space spectrum of the v th mode of the field $\zeta(\mathbf{x}, z, t)$). Here we assume, because of the statistical homogeneity of the internal-wave field, horizontally as well as vertically (in the second respect, even due to a weak correlation of the different modes),

$$\langle a_{vk}(t) a_{v_1 k_1}^*(t) \rangle = N_{vk}(t) \delta_{\mathbf{k}-\mathbf{k}_1} \delta_{v-v_1}, \quad (30.9)$$

where the angle brackets indicate statistical averaging. After calculating $\partial N_{vk} / \partial t$ with the aid of equations (30.6), we can express this quantity via triple correlations of the normal variables. For the triple correlations, with the aid of (30.6) we arrive at equations with quadruple correlations. The latter are expressed in terms of the double correlations, as in the case of a multidimensional Gaussian distribution. We neglect slow variations of the triple correlations with time and, temporarily, introduce a slight damping of these; then, with the aid of the formula $(x - i\varepsilon)^{-1} \rightarrow \pi \delta(x)$ ($\varepsilon \rightarrow 0$), we replace the quantities $(\omega_{vk} - \omega_{v_1 k_1} - \omega_{v_2 k_2})^{-1}$ by delta functions. As a result of all these operations we obtain the following kinetic equation, which is similar to (24.12')

$$\begin{aligned} \frac{\partial N_{vk}}{\partial t} = I_{vk} = 4\pi \int & [|V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{v v_1 v_2}|^2 (N_{v_1 k_1} N_{v_2 k_2} - N_{vk} N_{v_1 k_1} - N_{vk} N_{v_2 k_2}) \times \\ & \times \delta_{\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2} \delta_{\omega_{vk}-\omega_{v_1 k_1}-\omega_{v_2 k_2}} + 2 |V_{\mathbf{k}_1 \mathbf{k} \mathbf{k}_2}^{v_1 v v_2}|^2 (N_{v_1 k_1} N_{v_2 k_2} + N_{vk} N_{v_1 k_1} - \\ & - N_{vk} N_{v_2 k_2}) \delta_{\mathbf{k}-\mathbf{k}_1+\mathbf{k}_2} \delta_{\omega_{vk}-\omega_{v_1 k_1}+\omega_{v_2 k_2}}] d\mathbf{k}_1 d\mathbf{k}_2 dv_1 dv_2. \end{aligned} \quad (30.10)$$

This equation conserves the total energy $\int \omega_{vk} N_{vk} d\mathbf{k} dv$ and the total momentum $\int \mathbf{k} N_{vk} d\mathbf{k} dv$ (but does not conserve the total action $\int N_{vk} d\mathbf{k} dv$). Its right-

hand side $I_{\nu\mathbf{k}}$ has the form $A_{\nu\mathbf{k}} - B_{\nu\mathbf{k}}N_{\nu\mathbf{k}}$, where the first term describes the inflow of energy, and the second describes the energy outflow from mode $\nu\mathbf{k}$ because of the nonlinear interactions. Outside the spectral intervals of supply and dissipation of energy (if, of course, they are separated) Kolmogorov spectra are possible, that is, steady-state solutions of the kinetic equation, determined by the spectral fluxes of energy and momentum. For example, according to Pelinovskii and Raevskii (1977), the isotropic Kolmogorov spectrum of modules of the three-dimensional wave vectors $\kappa = (k^2 + l^2)^{1/2}$ has the form

$$E(\kappa) \sim (\varepsilon N)^{1/2} \kappa^{-2}, \quad (30.11)$$

where N is the local Brunt–Väisälä frequency, while in the case of long waves (on shallow water or with a thin pycnocline of thickness h)

$$E_n(k) \sim (\varepsilon_n N_{\max} h)^{1/2} k^{-3/2}. \quad (30.11')$$

Note, by the way, that the concept of weak nonlinearity is applicable only if the time of the nonlinear interactions $N_{\nu\mathbf{k}}/I_{\nu\mathbf{k}}$ is long in comparison with the period of a wave $\omega_{\nu\mathbf{k}}^{-1}$, that is, on condition that $I_{\nu\mathbf{k}} \omega_{\nu\mathbf{k}}^{-1} N_{\nu\mathbf{k}}^{-1} \ll 1$. Apparently, for small-scale waves this condition may not be satisfied (in particular, Miropol'skii noted that the theory of resonance interactions does not describe the self-modulation instability of quite long internal waves, which may play a significant role in nature).

Kinetic equation (30.10) is incomplete not only because not all the nonlinear wave interactions are taken into account in $I_{\nu\mathbf{k}}$, but also because, as in (24.19), on the left-hand side of the equation it is necessary in general to take refraction terms into account, while on the right-hand side the energy supply and dissipation naturally must figure in. Beginning with the energy supply, we note first of all perturbations at the upper and lower boundaries of the ocean. According to the calculations of a number of investigators, similar to (25.2)–(25.5), linear resonance atmospheric effects, such as fluctuations in atmospheric pressure, heat exchange with the atmosphere (producing pressure oscillations at the lower limit of the Ekman boundary layer), and especially frictional wind stress, are quite able to generate sizable internal waves, increasing to the limits of the linear theory $ak \sim 10^{-1}$ during a time of the order of a day (although for reliable estimates we do not know enough about the spectra of atmospheric effects in the appropriate intervals of wave number and frequency); if currents with a vertical shear are present, then these effects may be enhanced severalfold.

Thorpe (1966), and in a more general form Brekhovskikh *et al.* (1972, 1975), established that internal waves can be generated effectively during nonlinear interactions of surface waves, especially when the surface waves $(\mathbf{k}_1, \omega_1)$ and $(\mathbf{k}_2, \omega_2)$ form a resonance triad with an internal wave $\mathbf{k} = \mathbf{k}_1 - \mathbf{k}_2$, $\omega = \omega_1 - \omega_2$ (for instance, for $k_1 \approx k_2$); the latter here is damped slightly with increasing depth (as e^{kz}).

Internal waves are generated when a stratified current flows over irregularities of the bottom relief. This effect was calculated for the atmosphere by Dorodnitsyn (1950) and for the ocean by Tareev (1965). Wave generation by tsunamis from submarine earthquakes is a rare, but important, phenomenon. Internal waves with tidal periods are generated by tides and by tidal currents over inclined shelves (Rattray, 1960).

Of all the possible mechanisms of dissipation of the energy of internal waves, we will list three. Gravitational breaking (for $a\omega^2 \sim g$) is extremely improbable, because frequencies ω are usually low, but it becomes quite likely if a current with a velocity shear is present (produced, for instance, by longer waves), because it washes away the wave tops at depths where $U_k(z) > c_k$ (convective instability, Orlanskii and Bryan, 1969). Regular shear instability appears for $Ri < \frac{1}{4}$; apparently, Woods (1968) observed precisely this visually. According to the model calculations of Garrett and Munk (1972), who obtained $Ri = 2\pi^{-2}j^{-2}E^{-1}(N/N_{\max})^{-1}$, where $j = 20$ is the equivalent number of modes, while $E = 2\pi \times 10^{-5}$, shear instability is more probable than convective instability (determined mainly by the inclination), since in their models a considerable portion of the energy is supplied at a frequency around f , where the ratio of the shear to the inclination increases. The specific rate of energy dissipation is here found to be $\frac{5}{2}\nu(2\pi N)^2$, where $\nu \sim 1 \text{ cm}^2/\text{s}$ is the eddy viscosity (N is in cycles/s), and it is of the order of 10^{-6} W/m^3 . The third dissipation mechanism (of comparable, but apparently slightly lower, mean effectiveness) is the absorption of waves in the critical layers.

31. SPECTRA

Let us consider spectra (known as *unsymmetrical* spectra) which are different for waves propagating along the positive and negative directions of the wave vector $\mathbf{k}$. In addition, let us consider waves satisfying the boundary condition at the ocean surface: the set of modes ν is discrete (denumerable) and the wave field does not possess statistical homogeneity along the vertical (so that for its dependence on z we have to construct other statistical models). In this case the following three theorems are valid (Monin and Piterbarg, 1977).

THEOREM 1. *For statistical stationarity and horizontal homogeneity of a random field $\xi(\mathbf{x}, z, t)$ specified by formula (26.4), it is necessary and sufficient that the random complex amplitudes of its different modes be uncorrelated:*

$$\langle A_{\nu\mathbf{k}} A_{\nu_1\mathbf{k}_1}^* \rangle = F_{\nu\mathbf{k}} \delta_{\mathbf{k}-\mathbf{k}_1} \delta_{\nu\nu_1}, \quad (31.1)$$

where $F_{\nu\mathbf{k}}$ are certain real functions (unsymmetrical space spectra of the modes); here the correlation function of field (26.4) has the form

$$\begin{aligned} B_{rr}(z_1, z_2) &= \langle \xi(\mathbf{x} + \mathbf{r}, z_1, t + \tau) \xi(\mathbf{x}, z_2, t) \rangle \\ &= \int e^{i(\mathbf{k} \cdot \mathbf{r} - \omega\tau)} \xi_{k\omega}(z_1) \xi_{k\omega}(z_2) F(\mathbf{k}, \omega) d\mathbf{k} d\omega; \end{aligned} \quad (31.2)$$

$$F(\mathbf{k}, \omega) = \frac{1}{2} \sum_{\nu} [F_{\nu\mathbf{k}} \delta(\omega - \omega_{\nu\mathbf{k}}) + F_{\nu_1, -\mathbf{k}} \delta(\omega + \omega_{\nu\mathbf{k}})],$$

The correlation $\langle \varphi'(M_1) \psi^*(M_2) \rangle$ of two linear functionals φ and ψ of field ξ (represented in form (26.4) with $\xi_{k\omega}$ replaced by $\varphi_{\mathbf{k}\omega}$ or $\psi_{\mathbf{k}\omega}$) is represented by formula (31.2) with the product $\xi_{k\omega}(z_1) \xi_{k\omega}(z_2)$ replaced by $\varphi_{\mathbf{k}\omega}(z_1) \psi_{\mathbf{k}\omega}^*(z_2)$.

COROLLARY. *The spectral densities of fields u , v , w , p at a depth z and of the total energy of the internal waves on a unit area differ from $F(\mathbf{k}, \omega)$ by, respectively, the following factors:*

$$\begin{aligned} & \frac{k_x^2 \omega^2 + k_y^2 f^2}{k^4} \left(\frac{\partial \xi}{\partial z} \right)^2; \quad \frac{k_y^2 \omega^2 + k_x^2 f^2}{k^4} \left(\frac{\partial \xi}{\partial z} \right)^2; \quad \omega^2 \xi^2; \\ & \left(\frac{\omega^2 - f^2}{k^2} \rho_* \frac{\partial \xi}{\partial z} \right)^2; \quad \frac{\omega^2}{\omega^2 - f^2} \left[g \rho_0 \xi_0^2 + \int_{-h}^0 (N^2 - f^2) \xi^2 \rho_* dz \right], \end{aligned} \quad (31.3)$$

while the mutual spectral density of fields u and v differs by the factor $[k_x k_y (\omega^2 - f^2) - ik^2 \omega f] k^{-4} (\partial \xi / \partial z)^2$.

THEOREM 2. *If a random field $\xi(\mathbf{x}, z, t)$ is horizontally isotropic, so that all $F_{\mathbf{v}\mathbf{k}}$ depend only on k , its correlation function will have the form*

$$\begin{aligned} B_{rr}(z_1, z_2) &= \int_f^{N_{\max}} e^{-i\omega\tau} \sum_{\mathbf{v}} J_0(rk_{v\omega}) \xi_{v\omega}(z_1) \xi_{v\omega}(z_2) G_{v\omega} d\omega; \\ G_{v\omega} &= 2\pi k_{v\omega} k'_{v\omega} F_v(k_{v\omega}), \end{aligned} \quad (31.4)$$

where $J_0(x)$ is a Bessel function; the inverse of this formula is:

$$\begin{aligned} G_{v\omega} &= \xi_{v\omega}(z_1) \circ \Phi_{0\omega}(z_1, z_2) \circ \xi_{v\omega}(z_2); \\ \Phi_{r\omega}(z_1, z_2) &= (2\pi)^{-1} \int e^{i\omega\tau} B_{rr}(z_1, z_2) d\tau, \end{aligned} \quad (31.5)$$

(the sign $\circ$ denotes the convolution), so that the statistical structure of a horizontally isotropic field of internal waves, characterized by a set of spectra $F_v(k)$, $k = 0, 1, 2, \dots$, is completely determined by the values of this field on one vertical (for a known solution of the boundary-value problem (26.5), (26.5')).

Example 1. For $N = \text{const}$ (and, for simplicity, assuming a boundary condition $\xi = 0$ at $z = 0$) in a shallow ocean ($hN^2/g \ll 1$) the solution of the boundary-value problem (26.5), (26.5') will be approximately $\xi_v(z) = C_v \sin(v\pi z/h)$, with $k_{v\omega} = v\pi/\beta h$, where β is defined in (26.5). Using this solution and an integral representation of the Bessel function, we obtain from (31.4)

$$\begin{aligned} \Phi_{r\omega}(z_1, z_2) &= \frac{1}{8\pi} \int_0^{2\pi} \left[\varphi \left(\alpha + \frac{z_1 - z_2}{2h} \right) + \varphi \left(\alpha - \frac{z_1 - z_2}{2h} \right) - \right. \\ &\quad \left. - \varphi \left(\alpha + \frac{z_1 + z_2}{2h} \right) - \varphi \left(\alpha - \frac{z_1 + z_2}{2h} \right) \right] d\theta; \end{aligned} \quad (31.6)$$

$$\alpha = \frac{r \cos \theta}{2\beta h}; \quad \varphi(\alpha) = \sum_{\mathbf{v}} C_v^2 G_{v\omega} e^{2iv\pi\alpha}$$

(dependence of φ on ω is understood). Thus, for the frequency spectrum at a fixed depth $\Phi_\omega(z) = \Phi_{0\omega}(z, z)$ we obtain the value $\frac{1}{2}[\varphi(0) - \varphi(z/h)]$. Now we can express φ in terms of Φ_ω with the aid of this relation, and then rewrite formula (31.6) as

$$\Phi_{r\omega}(z_1, z_2) = \frac{1}{4\pi} \int_0^{2\pi} \left[\Phi_\omega \left(ah + \frac{z_1 + z_2}{2} \right) + \Phi_\omega \left(ah - \frac{z_1 + z_2}{2} \right) - \right. \\ \left. - \Phi_\omega \left(ah + \frac{z_1 - z_2}{2} \right) - \Phi_\omega \left(ah - \frac{z_1 - z_2}{2} \right) \right] d\theta. \quad (31.6')$$

Therefore, here for the complete determination of the statistical structure of the field ξ even nonsynchronous measurements at different depths along one vertical will be sufficient. Note, too, that if r is replaced by $(x^2/a^2 + y^2/b^2)^{1/2}$ then formula (31.6') also remains valid for an elliptically anisotropic field (since ξ_r is independent of k).

Example 2. For a horizontally isotropic field $\xi(\mathbf{x}, z, t)$ function $G_{v\omega}$ has the form $E_\omega(\alpha_\omega^2 + k_{v\omega}^2)^{-1}$ if, and only if, for any fixed $\mathbf{x}$ and ω the harmonics ξ_r of this field are random functions of z with uncorrelated increments; these constitute the simplest analogs of steady-state processes of autoregression (having spectra $|P_k|^{-2}$, where P_k is a polynomial). If such a field is measured at $z \geq z_1$, then its best extrapolation $\hat{\xi}$ into the region $z < z_1$ is the solution of the equation $(D - \alpha^2 \beta^2 \rho_*) \hat{\xi} = 0$ for the conditions $\hat{\xi}(z_1) = \xi(z_1)$, $\hat{\xi}(-h) = 0$, while the maximum likelihood estimate for parameter α is the solution of the equation $\Sigma \xi_v^2 = \int_{-h}^0 \Gamma(z, z; \alpha) \beta^2 \rho_* dz$, where $\Gamma(z_1, z_2; \alpha)$ is the Green function of operator $D - \alpha^2 \beta^2 \rho_*$. In the self-similarity case $\alpha = 0$, with an accuracy to terms of the order of $hN_{\max}^2/g$, we obtain

$$\Phi_{0\omega}(z_1, z_2) = E_\omega \left[\min(z_1, z_2) - \frac{z_1 z_2}{h} \right], \quad (31.7)$$

from which it follows that the frequency spectra of fields ξ and u at depth z and of the total energy of the internal waves differ from E_ω by, respectively, the factors

$$z \left(1 - \frac{z}{h} \right); \quad \frac{1}{2} (\omega^2 + f^2) \left[\int_z^0 \left(1 - \frac{2z}{h} \right) \beta^2 dz + \int_{-h}^0 \left(\frac{z}{h} \right)^2 \beta^2 dz \right]; \\ \omega^2 (\omega^2 - f^2)^{-1} \int_{-h}^0 z \left(1 - \frac{z}{h} \right) (N^2 - f^2) dz. \quad (31.7')$$

Example 3. Garrett and Munk (1972, 1975) selected for internal waves a horizontally isotropic climatic energy spectrum $F_E(\mathbf{k}, \omega) = E(k, \omega) (2\pi k)^{-1}$. They assumed self-similarity for the spectrum of wave numbers and frequencies $E(k, \omega) = k_*^{-1} A(k/k_*) B_\omega$, where $k_* = k_*(\omega)$ is the band width of the wave numbers k for waves of fixed frequency ω , and functions $A(p)$ and B_ω are normalized by the conditions $\int_0^\infty A(p) dp = \int_0^\infty B_\omega d\omega = 1$. The power law $B_\omega \sim$

$(2f/\pi\omega)(\omega^2 - f^2)^{-1/2}$ was selected for the frequency factor, since this formula agrees with the experimental ω^{-2} dependence for $\omega \gg f$ and gives the experimentally observed peak for $\omega \approx f$. Under average conditions in the ocean an exponential stratification $N(z) = N_0 \exp(z/L_0)$ with parameters $N_0 = 3$ cycles/h and $L_0 = 1.3$ km was assumed; then, between the local vertical wave number ν and the mode number j the approximate relation $\nu = j\pi N(z)/N_0 L_0$ is obtained, while the dispersion relation (26.14) for $\omega \ll N$ yields $k = \nu(\omega^2 - f^2)^{1/2}/N(z) \approx j\pi(\omega^2 - f^2)^{1/2}/N_0 L_0$. Similarly, it was assumed that $k_* \sim j\pi(\omega^2 - f^2)^{1/2}$. Consequently, the Garrett—Munk (GM) model of the dimensionless spectrum of internal waves has the form

$$F_E(\mathbf{k}, \omega) = \frac{Ef}{\pi^2 \omega(\omega^2 - f^2)^{1/2}} \frac{A(k/k_*)}{kk_*}; \quad (31.8)$$

$$f \leq \omega \leq N; \quad 0 \leq k \leq k_* = j\pi(\omega^2 - f^2)^{1/2},$$

where $E = 2\pi \times 10^{-5}$ is an empirical constant, the frequencies are measured in units of N_0 , and the wave numbers are measured in units of $M = (2\pi L_0)^{-1} = 1.22 \times 10^{-6}$ cycles/cm. The dimensional spectrum then becomes $\rho_0 L_0^5 N_0 F_E(kL_0, \omega/N_0)$, measured in joule-seconds. In model GM-72 it was assumed that $j = 20$ and $A(p) = 1$ for $0 \leq p \leq 1$; this gave a total wave energy per unit area of ocean surface equal to 0.382 J/cm^2 . According to a three-layer model of stratification of type (26.8), but with the boundary condition $\xi = 0$ for $z = 0$, we obtain from (31.3) and (31.8) at the ocean surface the relations

$$\sigma_u^2 \approx \sigma_0^2 \xi^{-1} \arctg \xi; \quad \sigma_0^2 = \frac{3E}{4\pi} \frac{N_0^2}{M^3 h}; \quad (31.9)$$

$$\xi = j\pi MH \left(\frac{\delta}{h} \right)^{1/2}.$$

Here it turns out that $\sigma_0 \approx 3.5 \text{ cm/s}$, about the same as σ_u , so that the horizontal velocities produced at the ocean surface by the internal waves are quite appreciable and may well create effects causing the emergence to the ocean surface of internal waves which are observable using aerospace techniques. Note, too, that in model GM-75 it was assumed that $j = 6$ and $A(p) = \frac{3}{2}(1 + p)^{-5/2}$, and that more recently other modifications have also been proposed (Munk, 1981).

The kinetic integral (30.10) was calculated on the basis of the GM models (Pomfey, 1981). With regard to this, one of the examples in Figure 31.1 indicates that in the absence of energy input and dissipation nonlinear interactions shift the energy from low modes with a broad frequency interval to intermediate and higher modes with low frequencies. For high ν there is an influx of energy to frequencies ω_{ν_k} around the inertial frequency f , while for low ν there is an outflow (indicating that inertial oscillations, with the low vertical coherence typical of them, may indeed be nonlinear interactions of internal waves). The energy flow over the spectrum is of the order of $\varepsilon \sim 10^{-3} \text{ W/m}^2$, while the time of energy transfer N_{ν_k}/I_{ν_k} is of the order of several days.

THEOREM 3. *If all the components $E_{ij}(\mathbf{k}, \omega)$ of the spectral tensor of the field of*

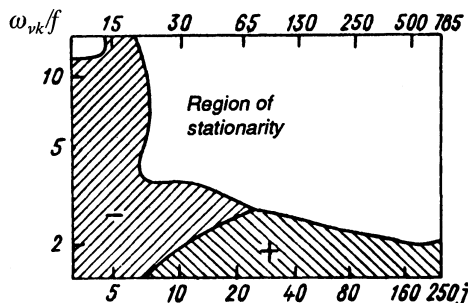


Fig. 31.1. Signs of kinetic integral of I_{vk} in j (or vL_0) ω_{vk}/f plane, according to Pomfrey (1981). Region of stationarity defined by condition $|I_{vk}|A_{vk} < 0.1$.

the horizontal velocity at the ocean surface $\mathbf{u}(\mathbf{x}, t)$ in the low-frequency region $\omega < \omega_0$ go to zero for all $\mathbf{k}$, then in this range $\omega < \omega_0$ the spectral tensor $E_{ij}^L(\omega)$ of the Lagrangian velocity $\dot{\mathbf{x}}(t)$ will also be equal to zero (this velocity satisfies the equation $\dot{\mathbf{x}} = \mathbf{u}(\mathbf{x}, t)$ with the condition $\mathbf{x}(0) = 0$).

COROLLARY. For the variance of the particle displacement $\sigma_i^2(t) = \langle x_i^2(t) \rangle$, the upper limit can be set at $\sigma_i^2 \leq 2\sigma_\omega^2/f^2$. Assuming that the Lagrangian correlation function of the velocity is not less than the Eulerian function, we obtain for the limit variance of the displacements an estimate of the lower limit as well:

$$\sigma_i^2(\infty) \geq 2 \int E_{ii}(\mathbf{k}, \omega) \omega^{-2} d\mathbf{k} d\omega \approx \frac{(7/6)\sigma_u^2}{f^2}, \quad (31.10)$$

where the latter equation corresponds to spectrum GM-72. Therefore, the limit rms radius σ_∞ of, for instance, an oil slick from a instantaneous point source on the ocean surface in an isotropic field of internal waves satisfies the inequalities $(7/6)^{1/2}\sigma_u/f \leq \sigma_\infty \leq 2^{1/2}\sigma_u/f$. For $\sigma_u \sim 3.5$ cm/s and $\max f = 1.46 \times 10^{-4}$ s $^{-1}$, we get $250 \text{ m} \leq \sigma_\infty \leq 350 \text{ m}$, so that fields of pollutants on the ocean surface must typically contain patches hundreds of meters in size. It is clear that at such distances r the structure functions of the velocity field are small in comparison with σ_ω^2 , and in the equation of convective diffusion $\partial s/\partial t = -\mathbf{u} \cdot \nabla s$ the dependence of $\mathbf{u}$ on $\mathbf{x}$ can be neglected. After writing the solution in the form $s(\mathbf{x}, t) = e^{-A} s(\mathbf{x}, 0)$, where $A = \int_0^t dt \mathbf{u} \cdot \nabla$, and averaging this result on the assumption that $\mathbf{u}$ is a Gaussian random function with a zero mean value and with uncorrelated Cartesian components, we obtain $\langle e^{-A} \rangle = e^{\langle A^2 \rangle/2}$, where $\langle A^2 \rangle = \sigma^2(t)\Delta$. Then the averaged equation of convective diffusion in the field of internal waves at the ocean surface becomes

$$\begin{aligned} \frac{\partial \langle s \rangle}{\partial t} &= K(t) \Delta \langle s \rangle; \\ K(t) &= \frac{1}{2} \frac{d\sigma^2(t)}{dt}; \quad \sigma^2(t) = \int_0^t \int_0^t R(t_1 - t_2) dt_1 dt_2, \end{aligned} \quad (31.11)$$

where $R(\tau)$ is the correlation function of the Eulerian velocity $u(0, t)$.

32. INERTIAL OSCILLATIONS

When we retain in the equations of the theory of internal waves terms with the Coriolis parameter f , we take into account not only gravitational oscillations produced by buoyancy forces (which in the case of stable stratification act as restoring forces), but also inertial oscillations caused by a disturbance of the equilibrium between the inertial force in the motion of the water relative to the rotating Earth and the inertial force in the transfer motion (the rotation of the planet), that is, the Coriolis force. At frequencies $\omega \gg f$ gravity waves are observed, slightly distorted by inertial effects. For $\omega \sim f$, on the other hand, the inertial oscillations are slightly distorted by the effect of stratification. These inertial oscillations, according to (26.5), take place practically without pressure perturbations and, according to (26.14), are waves with approximately vertical three-dimensional wave vectors $\boldsymbol{\kappa}$. Therefore, they must be local (any internal wave with a fixed frequency ω , propagating along the meridian, will, when it approaches a latitude at which the value of f equals ω , be transformed there into an inertial oscillation, described near this latitude by the Airy equation (27.14)).

Because of the local nature of the inertial oscillations, when describing them we can ignore the sphericity of the Earth, but we can easily take into account the total Coriolis acceleration $2\boldsymbol{\Omega} \times \mathbf{u}$ with $\boldsymbol{\Omega} = \text{const}$ without using the 'traditional approximation'. Then the fundamental equation (26.3), with $\rho_* \approx \text{const}$ for simplicity, becomes

$$\frac{\partial^2 \Delta \xi}{\partial t^2} + [4(\boldsymbol{\Omega} \cdot \nabla)^2 + N_0^2 \Delta_h] \xi = 0. \quad (32.1)$$

For $N_0 = 0$ we thus arrive at an equation of purely inertial oscillations (Equation (7.17) for gyroscopic waves). It has the elementary wave solutions $e^{i\theta}$, $\theta = \boldsymbol{\kappa} \cdot \boldsymbol{\xi} - \omega t$ with frequencies $\omega = \pm 2\boldsymbol{\Omega} \cos \alpha$, where α is the angle between $\boldsymbol{\Omega}$ and $\boldsymbol{\kappa}$. The velocity field is then

$$\begin{aligned} \mathbf{u} &= \mathbf{U} \cos \theta \pm \frac{\boldsymbol{\kappa} \times \mathbf{U}}{\kappa} \sin \theta; \\ \mathbf{c} &= 2 \frac{\boldsymbol{\Omega} \cdot \boldsymbol{\kappa}}{\kappa^3} \boldsymbol{\kappa}; \quad \mathbf{c}_g = \frac{2}{\kappa^3} \boldsymbol{\kappa} \times (\boldsymbol{\Omega} \times \boldsymbol{\kappa}), \end{aligned} \quad (32.2)$$

where $\mathbf{U}$ is an arbitrary constant vector, orthogonal to $\boldsymbol{\kappa}$, which describes transverse waves with circular polarization. At vertical $\boldsymbol{\kappa}$ vector $\mathbf{u}$ rotates both with time and with depth, clockwise in the Northern Hemisphere with $\omega = \pm f$; the phase propagates upward and the energy downward. The highest frequency of waves (32.2) is reached for $\boldsymbol{\kappa} \parallel \boldsymbol{\Omega}$ and is equal to $2\boldsymbol{\Omega}$ (semidiurnal period); here $p' = 0$, that is, the motion actually proceeds by inertia. Three-wave resonance interactions may occur between the gyroscopic waves.

Assuming quasiconstant N_0 , for gyroscopic waves modified by the effect of stratification, from (32.1) we arrive at the local dispersion relation $\omega^2 = N_0^2 \cos^2 \theta + 4\boldsymbol{\Omega}^2 \cos^2 \alpha$. The extrema ω_m of function $\omega(\boldsymbol{\kappa})$ are the roots of the equation $\omega_m^2(\omega_m^2 - 4\boldsymbol{\Omega}^2) = N_0^2(\omega_m^2 - 4\boldsymbol{\Omega}^2 \sin^2 \varphi)$, where φ is the geographical latitude; for $N_0 \gg \boldsymbol{\Omega}$ we obtain $\omega_{\min} \approx f$ and $\omega_{\max} \approx N_0$.

The precise nonlinear theory of horizontal inertial motion on a rotating sphere (Monin, 1972) is based on the equations

$$\begin{aligned}\frac{du}{dt} - \frac{uv \operatorname{tg} \varphi}{a} - 2\Omega v \sin \varphi &= 0; \\ \frac{dv}{dt} + \frac{u^2 \operatorname{tg} \varphi}{a} + 2\Omega u \sin \varphi &= 0,\end{aligned}\tag{32.3}$$

where $u = a(d\lambda/dt) \cos \varphi$ and $v = a(d\varphi/dt)$ are the velocity components directed eastward and northward, respectively; $\varphi(t)$ and $\lambda(t)$ are the latitude and longitude of the moving fluid particle, and a is the radius of the sphere. These equations have the following integrals of energy and angular momentum:

$$\begin{aligned}u^2 + v^2 &= u_0^2 + v_0^2; \\ (u + \Omega a \cos \varphi)a \cos \varphi &= (u_0 + \Omega a \cos \varphi_0)a \cos \varphi_0,\end{aligned}\tag{32.4}$$

where the zero subscripts denote the initial values of the functions of time. Setting $u_0 = U \cos \beta$, $v_0 = U \sin \beta$, $\varepsilon = U/2\Omega a$, for $\mu = \sin \varphi$ we obtain from (32.4) the equation

$$\left(\frac{d\mu}{dt} \right)^2 = \Omega^2 (\mu^2 - \mu_1^2) (\mu_2^2 - \mu^2);\tag{32.5}$$

$$\mu_{1,2}^2 = 1 - (\varepsilon \pm \delta)^2; \quad \delta = [(\cos \varphi_0 + \varepsilon \cos \beta)^2 + \varepsilon^2 \sin^2 \beta]^{1/2},$$

which is integrated in elliptic functions. The obtained trajectories of fluid particles are situated in the latitude zone $|\varepsilon - \delta| \leq \cos \varphi \leq \varepsilon + \delta$ (or $\mu_1^2 \leq \mu^2 \leq \mu_2^2 \geq 0$), which at its boundaries is tangent to the latitude circles. For $\varepsilon + \delta \leq 1$ or $\mu_1^2 \geq 0$ (class A₁) they do not cross the equator (but on these trajectories at the minimum distance from the equator the particle velocity is directed toward the west), while in the opposite case (class A₂) they *do* cross the equator, and periodically the same figures are executed alternately in each hemisphere. Trajectories for which $\delta \leq \varepsilon$ (class B₁) have everywhere a westerly component of the motion and describe loops around the poles or else pass through them; for $\delta > \varepsilon$ (class B₂) the trajectories at a maximum distance from the equator have an easterly component of motion.

If on the trajectory there are points at which $u = 0$ (loops on the trajectories are, as a rule, connected with these), then their latitudes φ_1 will, according to (32.4), be determined by the formula $\cos \varphi_1 = [\cos \varphi_0 (\cos \varphi_0 + 2\varepsilon \cos \beta)]^{1/2}$. The conditions for the existence of such points are: $-2\varepsilon \cos \beta \leq \cos \varphi_0 \leq (1 + \varepsilon^2 \cos^2 \beta)^{1/2} - \varepsilon \cos \beta$. If the initial motion of a particle has an easterly component ($\cos \beta > 0$), then the left-hand inequality is satisfied, while the right-hand one shows that loops are possible only on trajectories beginning not too close to the equator. If, on the other hand, $\cos \beta < 0$, then the right-hand inequality will be valid, while the left-hand one shows that loops are possible only on condition that $2\varepsilon |\cos \beta| \leq 1$ on trajectories beginning not too far from the equator.

For trajectories passing through a pole ($\cos \varphi_0 = 0$, $\delta = \varepsilon$) Equations (32.4) take the form $u = -\Omega a \cos \varphi$, $v^2 = U^2 - \Omega^2 a^2 \cos^2 \varphi$; the first of these gives $\lambda =$

$\lambda_0 - \Omega t$ (that is, a particle on a polar trajectory does not participate in the rotation of the planet), while according to the second we have $\cos \varphi \leq 2\varepsilon$ (and geophysical trajectories with $\varepsilon \leq \frac{1}{2}$ do not cross the equator; these are class A_1B_1 ; satellite trajectories with $\varepsilon > \frac{1}{2}$ pass alternately through the two poles and are class A_2B_1). Trajectories crossing the equator ($\cos \varphi_0 = 1$, class A_2) for $\cos \beta \geq \frac{1}{2}\varepsilon$ (with $\varepsilon \leq \frac{1}{2}$ always) are located in the equatorial zone $\cos \varphi \geq (1 + 2\varepsilon \cos \beta + \varepsilon^2)^{1/2} - \varepsilon$ and are class A_2B_2 . For $\cos \beta > 0$ these are without loops, having only meanders (for instance, for small ε at equatorial deep countercurrents in a very narrow zone, where for $y = a\varphi$ the second equation of (32.3) has the approximate form $\ddot{y} + (2\Omega U/a)y = 0$, so that the oscillation period of the meanders $\tau = 2\pi(a/2\Omega U)^{1/2}$ and their length $U\tau = \pi(2aU/\Omega)^{1/2}$ are large), while for $\cos \beta < 0$ they have loops, which are repeated alternately in each hemisphere (instability in cases where equatorial currents cross the equator).

For trajectories of class A_1 , Equation (32.5) yields

$$\mu = \pm \mu_1 \operatorname{dn}^{-1}(\mu_2 \Omega t); \quad k = \left(1 - \frac{\mu_1^2}{\mu_2^2}\right)^{1/2}, \quad (32.6)$$

where $\operatorname{dn} \theta$ is a Jacobi elliptic function (delta of amplitude with a modulus k), having with respect to θ a period $2K(k)$, so that the period of the latitude oscillations of the particles is

$$\tau = \frac{2K(k)}{\Omega \mu_2} \approx \frac{\pi}{\Omega |\sin \varphi_0|} \left(1 + \varepsilon \frac{\cos \varphi_0 \cos \beta}{\sin^2 \varphi_0}\right) \quad (32.7)$$

(the latter equation, valid for $\varepsilon \ll 1$, explains the observable slight deviations of the period of inertial oscillations from the pendulum day; this is a nonlinear effect). For trajectories of class A_2 instead of (32.6), (32.7) we have

$$\mu = \pm k |\mu_1| \frac{\sin \theta}{\operatorname{dn} \theta}; \quad \theta = \frac{\Omega \mu_2 t}{k}; \quad k = \left(1 + \frac{|\mu_1^2|}{\mu_2^2}\right)^{-1/2}; \quad (32.6')$$

$$\tau = \frac{4kK(k)}{\Omega \mu_2} \approx 4 \left(\frac{a}{2\Omega U}\right)^{1/2} K\left(\sin \frac{\beta}{2}\right). \quad (32.7')$$

Finally, the dependence of the particle longitude on t has the form

$$\lambda = \lambda_0 - \Omega t + (\delta^2 - \varepsilon^2)\Omega \int_0^t (1 - \mu^2)^{-1} dt \quad (32.8)$$

being expressed in terms of elliptic functions. For trajectories of class B_1 this dependence is monotonic (that is, the motion always has a westerly component), whereas for class B_2 the longitude period $\lambda(\tau) - \lambda_0$ may be equal to zero (inertial oscillations without a mean drift, for example, on A_1B_2 trajectories in the middle latitudes for $\cos \beta = -1$ and $\varepsilon \ll 1$ this is valid with an accuracy of ε^2).

Geophysical Turbulence

33. STATISTICAL FLUID MECHANICS

Turbulence was defined in Section 15 as an vortical motion of a viscous fluid, which is *stochastic* in the sense that it possesses the following three properties: (1) it is highly sensitive to changes in the initial conditions (because of the exponential divergence of initially close phase trajectories); (2) almost all the phase trajectories are everywhere dense, and any initial nonequilibrium probability distribution in phase space approaches a limit equilibrium distribution; (3) it is characterized by mixing in phase space (and, as a result, by ergodicity, by a rapid decay of the time correlation functions, and by continuity of the frequency spectra). The *developed* turbulence also possesses: (4) a multimode nature (and hence a random spatial structure at any fixed moment of time).

Qualitatively, a turbulent flow differs from other flows in that its thermodynamic and fluid-dynamical characteristics (velocity vector, pressure, temperature, concentrations of admixtures, density, sound velocity, refractive index, electrical conductivity, etc.) exhibit random fluctuations, caused by the presence of numerous eddies of different sizes. As a consequence, a turbulent flow varies quite irregularly in both time and space. Moreover, unlike stochastic wave fields, the Fourier components of turbulent thermodynamic and fluid-dynamical fields do not possess single-valued dispersion relations between the frequencies and the wave vectors, or frequency dependences of the phase shifts between the values of different fields.

By definition, geophysical turbulence is observed in natural flows occurring in the Earth's atmosphere, the World Ocean, seas, rivers, lakes, etc., and also beyond the Earth: on other planets, in the Sun's atmosphere and interior, in other stars, and in cosmic gas clouds. This definition has even been extended to flows of stars in galaxies, and of galaxies in the Metagalaxy, provided these are idealized as flows of fluids in which peculiar remote (gravitational) interactions between 'particles' take place.

Geophysical turbulence is affected specifically by, firstly, the stratification of the fluid, and, secondly, the rotation of the Earth (or of the celestial object involved). The local characteristic of the stratification is the quantity

$$N^2(z) = -g \frac{\partial \rho_*}{\partial_* \partial z},$$

where ρ_* is the potential density. If $N^2 > 0$, then the buoyancy force is a restoring force, so that the stratification is hydrostatically stable, and the turbulence must consume energy by working against the buoyancy force, so that it only develops slightly, often being concentrated in individual thin layers. For $N^2 < 0$ the buoyancy force becomes an accelerating force, and the stratification is unstable; in this case a source of turbulent-convection energy exists, and turbulence develops intensively. If there is a shear in the mean flow velocity $U(z)$, then the hydrodynamic instability is enhanced, and it will be observed for $N^2 > 0$ as well as for $N^2 < 0$, as long as the Richardson number $Ri = N^2|U'|^{-2}$ does not exceed some critical value $Ri_{cr} > 0$ (according to formula (11.26), in an ideal fluid $Ri_{cr} = \frac{1}{4}$; see Miles and Hobard, 1961).

The Earth's rotation is characterized by, first of all, the Coriolis parameter $f = 2\omega_z$. Locally its effect reduces to the formation of turbulent boundary layers (*Ekman layers*) in the atmosphere near the Earth's surface and also at the surface and bottom of the ocean. These layers have a thickness $h_0 \approx u_*/f$, where u_* is the 'friction velocity', and in them rotation of the mean flow occurs with increasing height or depth ('Ekman spiral'). Globally, however, variations of f along the meridian $\beta = \partial f / \partial y$ produce a meridional restoring force, giving rise to Rossby waves, forming quasi-two-dimensional large-scale turbulence (cyclones and anticyclones in the atmosphere, synoptic eddies in the ocean, eddies in the atmospheres of Jupiter and Saturn, including the Great Red Spot, and, apparently, giant convective cells on the Sun as well).

In the formulation of Kolmogorov (1941) the theory of turbulence is a statistical fluid-mechanics in which the thermodynamic and fluid-dynamical fields are considered to be random functions of points $M = (\mathbf{x}, t)$ in space-time, these fields being averaged using the techniques of mathematical expectation (in this sense, any theory of turbulence will be statistical).

The complete statistical description of a turbulent flow (with a fixed geometry of its boundaries) reduces to a determination of the probability measure in its phase space. The latter consists of all possible individual realizations of the random thermodynamic and fluid-dynamical fields characterizing it (so that turbulence theory becomes one of the applied chapters of the theory describing probability distributions in functional spaces). If we specify this probability measure in its most compact form, then the characteristic functional satisfies a *linear* equation in variation derivatives, following from the equations of fluid mechanics (Hopf, 1952); however, a constructive solution of such an equation has not yet been devised (Vishik and Fursikov, 1980, showed that a common method can be used to construct probability measures for a sequence of Galerkin approximations of the fluid-dynamical equations: the family of these measures possesses weak compactness).

The probability measure can be specified less compactly but more constructively if we determine all the finite-dimensional probability distributions for the values $u_i(M_1), \dots, u_n(M_n)$ of the thermodynamic and fluid-dynamical fields in all possible finite sets of points in space-time $M_1 = (\mathbf{x}_1, t_1), \dots, M_n = (\mathbf{x}_n, t_n)$. Such finite-dimensional distributions satisfy an infinite series of *linear* equations following from the equations of fluid mechanics (Monin, 1967). For example, for an

n -point probability density of the velocity field $f_n = p_{M_1 \dots M_n}(\mathbf{u}_1, \dots, \mathbf{u}_n)$ we obtain the equations

$$\frac{\partial f_n}{\partial t_l} = - \left(\frac{\partial u_{la} f_n}{\partial x_{la}} + \frac{\partial w_{la} f_n}{\partial u_{la}} \right), \quad (33.1)$$

where w_l is the conditional mathematical expectation (shifted to point $\mathbf{x}_l$) of the fluid-dynamical acceleration at point $(\mathbf{x}, t_l)$, on condition that the values $\mathbf{u}_1 = \mathbf{u}(M_1), \dots, \mathbf{u}_n = \mathbf{u}(M_n)$ are fixed (in geophysical fluid dynamics w_l includes terms corresponding to the accelerations of the buoyancy force and the Coriolis force). The quantities $w_{la} f_n$ contain integrals over $f_{n+1} d\mathbf{u}_{n+1}$, so that the equations for f_n , $n = 1, 2, \dots$, form an infinite coupled series, analogous to the series of equations of Bogolyubov for n -particle distribution functions of molecular velocities in the kinetic theory of gases; however, as yet no convincing methods have been worked out for deriving from (33.1) any appropriate finite system of equations, similar for instance to the derivation of the Boltzmann kinetic equation. Integration of (33.1) over $F(\mathbf{u}_1, \dots, \mathbf{u}_n) d\mathbf{u}_1 \dots d\mathbf{u}_n$ yields

$$\frac{\partial \langle F \rangle}{\partial t_l} = - \frac{\partial \langle u_{la} F \rangle}{\partial x_{la}} + \left\langle w_{la} \frac{\partial F}{\partial u_{la}} \right\rangle. \quad (33.2)$$

For

$$F = u_{1a_1}^{m_1} \dots u_{na_n}^{m_n}$$

the quantity $\langle F \rangle$ is the n -point statistical moment ($N = (m_1 + \dots + m_n)$ th order) of the velocity field. Equations (33.2) for the moments were first derived by Fridman and Keller (1924), and they constitute the first formulation of the total statistical description of turbulent flows. Like (33.1), they are *linear* and form an infinite coupled series: if $\langle F \rangle$ is a moment of order N , then the right-hand side of (33.2) contains moments of order $N + 1$. The solvability of the infinite system of Equations (33.2) is demonstrated with the aid of the Galerkin approximations of the fluid-dynamical equations.

For $N = 1$, Equations (33.2) are obtained by direct averaging of the fluid-dynamical equations; since they were first obtained by Reynolds (1894), they are known as the *Reynolds equations*. Here and throughout this chapter a bar over a quantity will denote the average value, that is, the mathematical expectation $\bar{A}$, while a prime $A' = A - \bar{A}$ will denote the turbulent fluctuations, that is, the departures of the individual realizations from the mathematical expectations. Equations (33.2) for $N = 1$ for the average velocity field $\bar{\mathbf{u}}$ differ from the initial equations for $\mathbf{u}$, in that new unknowns appear: the single-point second moments $\tau_{ji} = -\rho \overline{u'_j u'_i}$, which can be interpreted as additional (turbulent) stresses and are thus called *Reynolds stresses*. Similarly, the averaged continuity equation for a compressible fluid must include the turbulent mass flux $\mathbf{m} = \overline{\rho' \mathbf{u}'}$, the averaged heat-conduction equation must include the turbulent heat flux $\mathbf{q} = c_p \rho_0 \overline{T' \mathbf{u}'}$, and the

averaged equation of diffusion of a passive admixture must include the turbulent flux of the admixture $\mathbf{I} = \rho_0 s' \mathbf{u}'$, and so forth.

For the local description of geophysical turbulent boundary layers the SHH model is usually applicable, where SHH refers to a statistically Stationary and Horizontally Homogeneous turbulent flow, in which all the finite-dimensional probability distributions are invariant relative to any shifts of the coordinate origin on the time axis or any horizontal shifts of the frame of reference in space. Then, in particular, all single-point moments (including the mean pressure gradient) may depend only on the vertical coordinate z of the observation point, and the mean flow is horizontal: $\bar{u} = u(z)$, $\bar{v} = v(z)$, $\bar{w} = 0$. For the SHH model the Reynolds equations become

$$\begin{aligned} \frac{1}{\rho_0} \frac{\partial}{\partial z} \left(\tau_{xz} + \rho_0 \nu \frac{\partial \bar{u}}{\partial z} \right) + f \left(\bar{v} - \frac{1}{f \rho_0} \frac{\partial \bar{p}}{\partial x} \right) &= 0; \\ \frac{1}{\rho_0} \frac{\partial}{\partial z} \left(\tau_{yz} + \rho_0 \nu \frac{\partial \bar{v}}{\partial z} \right) - f \left(\bar{u} + \frac{1}{f \rho_0} \frac{\partial \bar{p}}{\partial y} \right) &= 0; \\ \frac{\partial}{\partial z} (\rho_0 \sigma_w^2 + \bar{p}) &= g \bar{p} + 2 \rho_0 \omega_y \bar{u}; \\ q_z - c_p \rho_0 X \frac{\partial \bar{T}}{\partial z} = \text{const}; \quad I_z - \rho_0 D \frac{\partial \bar{s}}{\partial z} &= \text{const}. \end{aligned} \tag{33.3}$$

Here and in the following we use the notation: $\sigma_A^2 = \overline{A'^2}$, χ is the molecular thermal diffusivity, and D is the kinematic coefficient of molecular diffusion. The first two equations characterize the Ekman boundary layers, while the third shows that the mean SHH flow is not, strictly speaking, hydrostatic; the fourth (derived from the entropy equation, taking into account only the temperature dependence of the entropy) and the fifth show that the total vertical fluxes of heat and admixture in an SHH flow do not vary along the vertical (this will remain true if the radiative heat flux is also taken into account).

The turbulent flows entering into the Reynolds equations in the so-called *semiempirical* theories of turbulence are described, by analogy with the corresponding molecular flows, as linear functions of the gradients of the mean fluid-dynamical fields. The coefficients of these linear functions have the sense of coefficients of eddy viscosity, thermal conduction, and diffusion. To determine these, the theories of Taylor (1915, 1932) and Prandtl (1925) introduced the concept of the *mixing length* l , analogous to the mean free path of a molecule in the kinetic theory of gases. Thus, for example, in a plane-parallel mean SHH flow (when $\bar{v} = 0$) we obtain $w' = l' \partial \bar{u} / \partial z$, and the kinematic coefficient of eddy

viscosity is given by the formula

$$K = \overline{l'w'} = l^2 \left| \frac{\partial \bar{u}}{\partial z} \right|,$$

where $l = (\overline{l'^2})^{1/2}$ is known as the *scale of turbulence*. In the theory of von Kármán (1930) it is defined as $l = -\kappa \Psi (\partial \Psi / \partial z)^{-1}$, where $\Psi = \partial \bar{u} / \partial z$ and κ is a numerical constant (equal to about 0.4).

By subtracting the corresponding averaged equations from the equations of fluid dynamics, we arrive at evolution equations for the turbulent fluctuations u'_j , T' , and s' . Then, cross multiplication and addition readily yield equations for the single-point second moments of these fluctuations. For instance, in the SHH model the equations for the diagonal elements of the Reynolds-stress tensor are:

$$\begin{aligned} \frac{1}{2} \frac{\partial \rho_0 \sigma_u^2}{\partial t} &= \tau_{xz} \frac{\partial \bar{u}}{\partial z} - \rho_0 \varepsilon_u - \overline{p' \frac{\partial v'}{\partial y}} - \overline{p' \frac{\partial w'}{\partial z}} - \\ &\quad - \frac{1}{2} \frac{\partial}{\partial z} \left(\rho_0 \overline{u'^2 w'} - \rho_0 \nu \frac{\partial \sigma_u^2}{\partial z} \right) = 0; \\ \frac{1}{2} \frac{\partial \rho_0 \sigma_v^2}{\partial t} &= \tau_{yz} \frac{\partial \bar{v}}{\partial z} - \rho_0 \varepsilon_v + \overline{p' \frac{\partial v'}{\partial y}} - \\ &\quad - \frac{1}{2} \frac{\partial}{\partial z} \left(\rho_0 \overline{v'^2 w'} - \rho_0 \nu \frac{\partial \sigma_v^2}{\partial z} \right) = 0; \\ \frac{1}{2} \frac{\partial \rho_0 \sigma_w^2}{\partial t} &= g m_z - \rho_0 \varepsilon_w + \overline{p' \frac{\partial w'}{\partial z}} - \\ &\quad - \frac{1}{2} \frac{\partial}{\partial z} \left(\rho_0 \overline{w'^2} + 2 \overline{p' w'} - \rho_0 \nu \frac{\partial \sigma_w^2}{\partial z} \right) = 0, \end{aligned} \quad (33.4)$$

where $\varepsilon_i = \overline{v |\nabla u'_i|^2}$ is the 'rate of dissipation' of energy $\sigma_{u'_i}^2/2$. Terms $\tau_{xz} \partial \bar{u} / \partial z$ and $\tau_{yz} \partial \bar{v} / \partial z$ enter into the right-hand sides of the equations for components $|\bar{u}_i|^2/2$ of the kinetic energy of the mean flow with opposite signs; consequently, they describe the generation of turbulence because of the mean flow. It is clear from (33.4) that in just this way only horizontal velocity fluctuations can be generated; vertical fluctuations can be generated only via buoyancy forces, at the expense of the potential energy of stratification of the medium (if it is unstable, that is, if $m_z > 0$; otherwise the vertical fluctuations will lose energy by having to work against the buoyancy forces). Terms with pressure fluctuations $\overline{p' \partial v' / \partial y}$ and $\overline{p' \partial w' / \partial z}$ enter into the equation for σ_u^2 , and the equations for σ_v^2 and σ_w^2 , with

opposite signs; consequently, these terms describe the energy exchange between the longitudinal and transverse velocity fluctuations.

In the case of a plane-parallel flow along the x axis with neutral stratification ($\bar{v} = m_z = 0$), integration of Equation (33.4) over a volume, at the boundaries of which there are no flows of energy of turbulence, yields

$$\int p' \frac{\partial v'}{\partial y} dV = \int \rho_0 \varepsilon_v dV > 0; \quad \int p' \frac{\partial w'}{\partial z} dV = \int \rho_0 \varepsilon_w dV > 0; \quad (33.5)$$

$$\int \tau_{xz} \frac{\partial \bar{v}}{\partial z} dV = \int \left(p' \frac{\partial v'}{\partial y} + p' \frac{\partial w'}{\partial z} + \rho_0 \varepsilon_u \right) dV = \int \rho_0 \varepsilon dV > 0,$$

where $\varepsilon = \varepsilon_u + \varepsilon_v + \varepsilon_w$, so that the energy of the mean flow is directly transferred only to the longitudinal fluctuations u' , while the transverse fluctuations v' and w' receive energy from u' via the work of the pressure fluctuations (which thus serve to redistribute energy among the turbulent motions in different directions, making them tend to become isotropic). By adding the three Equations (33.4), we arrive at an equation for the kinetic energy of turbulence per unit mass $b^2 = \frac{1}{2}|\mathbf{u}'|^2$ (Kolmogorov, 1942; Monin, 1950):

$$\begin{aligned} \frac{\partial \rho_0 b^2}{\partial t} = & \left(\tau_{xz} \frac{\partial \bar{u}}{\partial z} + \tau_{yz} \frac{\partial \bar{v}}{\partial z} \right) (1 - \text{Rf}) - \rho_0 \varepsilon - \\ & - \frac{\partial}{\partial z} \left(\rho_0 \overline{b'^2 w'} + \overline{p' w'} - \rho_0 \nu \frac{\partial b^2}{\partial z} \right) = 0; \end{aligned} \quad (33.6)$$

$$\text{Rf} = -gm_z \left(\tau_{xz} \frac{\partial \bar{u}}{\partial z} + \tau_{yz} \frac{\partial \bar{v}}{\partial z} \right)^{-1},$$

where $b'^2 = \frac{1}{2}|\mathbf{u}'|^2$, and Rf is the *dynamic* Richardson number (the ratio of the work of the buoyancy forces to the work of the Reynolds stresses). In order for the first term on the right-hand side to be able to compensate the viscous dissipation, the condition $\text{Rf} < 1$ (the Richardson criterion) must be satisfied. This is only a very rough criterion, and a more realistic one can be specified in the form $\text{Rf} < \text{Rf}_{\text{cr}} < 1$. If we assume that the third term in (33.6) (the divergence of the flow of turbulent energy) for not very low Rf is proportional to gm_z , and if we call the proportionality coefficient $\sigma - 1$ (> 0), then we get $\text{Rf}_{\text{cr}} = 1/\sigma$.

Kolmogorov (1942) suggested expressing the single-point characteristics of any (not too anisotropic) turbulence in terms of its energy b^2 and scale l . Then, we can introduce into (33.6), using the laws of the semiempirical theory of turbulence, the eddy viscosity K , the thermal diffusivity $K_T = \alpha K$, and the diffusion coefficient

$K_1 = \alpha_1 K$, to obtain:

$$\begin{aligned} \frac{\partial b^2}{\partial t} = K \left[\left(\frac{\partial \bar{u}}{\partial z} \right)^2 + \left(\frac{\partial \bar{v}}{\partial z} \right)^2 \right] (1 - \alpha \text{ Ri}) - \\ - \varepsilon + \frac{\partial}{\partial z} \alpha_1 K \frac{\partial b^2}{\partial z} = 0. \end{aligned} \quad (33.7)$$

$$\text{Ri} = N^2 \left[\left(\frac{\partial \bar{u}}{\partial z} \right)^2 + \left(\frac{\partial \bar{v}}{\partial z} \right)^2 \right]^{-1}; \quad K = bl; \quad \varepsilon = b^3 l^{-1}.$$

This equation, together with (33.3), forms a closed system, provided it is supplemented by some definite scale of turbulence l , for example, as given by the following generalization of the previously cited formula of von Kármán:

$$l = -\kappa \Psi \left(\frac{\partial \Psi}{\partial z} \right)^{-1}; \quad \Psi^2 = \left[\left(\frac{\partial \bar{u}}{\partial z} \right)^2 + \left(\frac{\partial \bar{v}}{\partial z} \right)^2 \right] (1 - \alpha \text{ Ri}). \quad (33.8)$$

Equations like (33.4) can also be derived without difficulty for all the other single-point second moments of the fields u'_i , T' , and s' ; we will return to these below.

According to the basic idea of Kolmogorov (1941), developing turbulence (with a sufficiently high Reynolds number) must be *locally isotropic*, that is, for any fixed values $\mathbf{u}(M_0) = \mathbf{u}_0$, $T(M_0) = T_0$, $s(M_0) = s_0$ all the finite-dimensional probability distributions for the differences $\mathbf{u}(M_k) - \mathbf{u}(M_0)$, $T(M_k) - T(M_0)$, $s(M_k) - s(M_0)$ in all possible finite sets of points in space-time $M_1, \dots, M_n$ from a sufficiently small space-time neighborhood G around point M_0 are independent of $\mathbf{u}_0$, T_0 , s_0 and are invariant with respect to the choice of point M_0 in region G and arbitrary rotations and reflections in the space of vectors $\mathbf{r} = \mathbf{x} - \mathbf{x}_0 - \mathbf{u}_0(t - t_0)$.

For a neutral density stratification this is also valid for geophysical turbulence, so that in view of the weakness of its small-scale components the influence on these of the rotation of the planet will also be small (compared to the effect of the pressure gradients), so that it practically does not disturb their isotropy. If, on the other hand, the stratification is not neutral, then the direction of the buoyancy forces (vertical) turns out to be preferred, and the turbulence is not locally isotropic, being only *locally axisymmetric relative to the vertical* (that is, the above-mentioned probability distributions for the differences in the values of the fluid-dynamical fields are invariant with respect to rotations of the frame of reference only about the vertical axis and reflections only in vertical planes). The spatial *correlation functions* (that is, the simultaneous two-point moments) of the fluctuations of the fluid-dynamical fields in this case are

$$\overline{u'_{j_1}(\mathbf{x}_1, t) u'_{j_2}(\mathbf{x}_2, t)} = B_{j_1 j_2}(\mathbf{x}_{2h} - \mathbf{x}_{1h}, z_1, z_2), \quad (33.9)$$

where subscript h denotes the horizontal component; the spatial *structure functions*

(that is, the two-point moments of the spatial differences) are

$$[u'_{j_1}(\mathbf{x}_2, t) - u'_{j_1}(\mathbf{x}_1, t)] [u'_{j_2}(\mathbf{x}_2, t) - u'_{j_2}(\mathbf{x}_1, t)] = D_{j_1 j_2}(\mathbf{r}), \quad (33.10)$$

where $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$. Note that the scalar structure functions depend only on the two variables r_h and r_z , while the structure functions with tensor indices are linear combinations of constant tensors and tensors consisting of vector $\mathbf{r}$ and the unit vector of the vertical λ . In particular, the single-point moments $\overline{\nabla u'_{j_1} \cdot \nabla u'_{j_2}}$, expressed in terms of the derivatives at zero of the structure functions, are

$$\begin{aligned} \overline{\nabla u'_i \cdot \nabla u'_j} &= |\overline{\nabla u'}|^2 \delta_{ij} + (|\overline{\nabla w'}|^2 - |\overline{\nabla u'}|^2) \lambda_i \lambda_j; \\ \overline{\nabla T' \cdot \nabla u'_i} &= \overline{\nabla T' \cdot \nabla w'} \lambda_i, \end{aligned} \quad (33.11)$$

from which it follows that

$$|\overline{\nabla v'}|^2 = |\overline{\nabla u'}|^2; \quad \overline{\nabla u'_i \cdot \nabla u'_j} = 0 \quad \text{for } i \neq j \quad \text{and} \quad \overline{\nabla T' \cdot \nabla u'_i} = 0$$

for $i \neq 3$ (where here T' can everywhere be replaced by s').

34. TURBULENCE IN STRATIFIED MEDIA

The effect of stratification on turbulence will be examined first using the simplest example of an atmospheric surface layer so thin that in it the effect of the Coriolis force on a mean stationary horizontally uniform flow is still practically imperceptible (the second terms in the first two equations of (33.3) are negligibly small, indicating that the vertical flux of momentum is constant). In this case the mean flow does not rotate with height (in a plane-parallel manner). If the x axis is along the flow, then $\bar{u} = \bar{u}(z)$ and $\bar{v} = \bar{w} = 0$. Then the two indicated equations reduce to

$$\tau_{xz} + \rho_0 \nu \frac{\partial \bar{u}}{\partial z} = \text{const}; \quad \tau_{yz} = 0. \quad (34.1)$$

Now we note that practically everywhere, except in a very thin sublayer adjacent to the underlying surface, the molecular fluxes of momentum and heat are always negligibly small in comparison with the turbulent fluxes. Consequently, the first equation of (34.1) and the fourth equation of (33.3) can be written in the simplified form $\tau_{xz} = \rho_0 u_*^2 = \text{const}$ and $q_z = q = \text{const}$ (the quantity $u_* = (\tau_{xz}/\rho_0)^{1/2}$ introduced here, which has units of velocity, is known as the *friction velocity*). We take into account, too, recalling formula (1.11), that in the atmosphere $\rho'/\rho_0 \approx -T'/T_0$, so that the specific work of the buoyancy forces gm_z/ρ_0 entering into the third equation of (33.4) or energy equation (33.6) turns out to be equal to $-(g/T_0)(q/c_p \rho_0)$, and thus is determined not only by the quantity q but also by the so-called *buoyancy parameter* g/T_0 (we note that, when generalizing this theory to the ocean, we have to set $\rho'/\rho_0 \approx -\alpha T' + \gamma s'$); the work will thus be

$$\frac{gm_z}{\rho_0} = -\alpha g \left(\frac{q}{c_p \rho_0} \right) + \gamma g \left(\frac{I}{\rho_0} \right).$$

From the three constant parameters u_* , $q/c_p\rho_0$, and g/T_0 entering into Equations (33.4) and (33.6), we can construct the scales of length L and temperature T_* in the following form (accurate to a numerical factor, this form is unique):

$$L = \frac{u_*^3}{\kappa g m_z / \rho_0} = - \frac{u_*^3}{\kappa (g/T_0) (q/c_p\rho_0)}; \quad T_* = - \frac{q/c_p\rho_0}{\kappa u_*}, \quad (34.2)$$

where κ is the above-mentioned *von Kármán constant*, introduced for the sake of convenience of certain subsequent calculations. Turbulence in stratified SHH flows can (except for its very small-scale components, in the regime of which the effects of molecular viscosity and thermal conduction are substantial) be described with the aid of the similarity theory of Monin and Obukhov (1953). According to the latter theory, the single-point probability distribution for the dimensionless velocity $\mathbf{u}/u_*$ and dimensionless temperature T/T_* may depend just on one parameter: the dimensionless height of the observation point $\xi = z/L$:

$$p_M(\mathbf{u}, T) = u_*^{-3} |T_*|^{-1} F\left(\frac{\mathbf{u}}{u_*}, \frac{T}{T_*}; \frac{z}{L}\right). \quad (34.3)$$

This is also valid for the *form* of the mean profiles of the velocity $u(z)$ and temperature $T(z)$, whereas their *absolute values* are determined by the boundary conditions at the Earth's surface and have the form:

$$\begin{aligned} \bar{u}(z) &= \frac{u_*}{\kappa} \left[f\left(\frac{z}{L}\right) - f\left(\frac{z_0}{L}\right) \right]; \\ \bar{T}(z) &= \bar{T}(z_0) + T_* \left[f_1\left(\frac{z}{L}\right) - f_1\left(\frac{z_0}{L}\right) \right], \end{aligned} \quad (34.4)$$

where z_0 is the *roughness height* of the underlying surface, defined by the condition $\bar{u}(z_0) = 0$, while $f(\xi)$ and $f_1(\xi)$ are certain universal functions. According to functions (34.4), the coefficients of turbulent exchange for momentum and heat can be defined as

$$\begin{aligned} K &= \frac{u_*^2}{\partial \bar{u} / \partial z} = \frac{\kappa u_* L}{f'(\xi)}; \quad K_T = \frac{-q/c_p\rho}{\partial \bar{T} / \partial z} = \frac{\kappa u_* L}{f'_1(\xi)}; \\ \frac{K_T}{K} &= \frac{f'(\xi)}{f'_1(\xi)} \equiv \alpha(\xi), \end{aligned} \quad (34.5)$$

where function $\alpha(\xi)$, in general, is different from unity. In terms of these same functions, the Richardson number Ri and the dynamic Richardson number Rf can

be expressed as

$$\begin{aligned} \text{Ri} &= \frac{(g/T_0)(\partial \bar{T}/\partial z)}{(\partial \bar{u}/\partial z)^2} = \frac{1}{\alpha(\xi)f'(\xi)}; \\ \text{Rf} &= \frac{-(g/T_0)(q/c_p\rho_0)}{u_*^2\partial \bar{u}/\partial z} = \frac{1}{f'(z)}. \end{aligned} \quad (34.6)$$

The number Rf is obviously related simply to the kinematic eddy viscosity K :

$$K = \kappa u_* L \cdot \text{Rf}. \quad (34.7)$$

Below it will be seen that functions $\text{Ri}(\xi)$ and $\text{Rf}(\xi)$ are monotonic and thus have single-valued inverse functions. From this it is clear that the quantities Ri, Rf, and ξ can with equal validity be used as local parameters characterizing the vertical stability of a stratified SHH flow.

Formulas (34.3) and (34.4) are valid for stable stratification, when $q < 0$, $L > 0$, $T_* > 0$, and $\xi \geq 0$, and also for unstable stratification, when, on the other hand, $q > 0$, $L < 0$, $T_* < 0$, and $\xi \leq 0$. Their asymptotes as $\xi \rightarrow 0$, $\pm\infty$ can be established with the aid of qualitative considerations. The asymptote for $\xi \rightarrow 0$ corresponds to taking the limit $q \rightarrow 0$ for neutral stratification. But it is also obtained at the limit $z \rightarrow 0$ for any stratification, that is, it must occur in the sublayer $z \ll |L|$. Thus, in this sublayer the effect of stratification on turbulence is still invisible (giving it the name: *sublayer of dynamic turbulence*). This means that in it the buoyancy parameter g/T_0 does not play a role, that is, it must drop out of the asymptotes of formulas (34.3) and (34.4). Thus, after rewriting (34.4) in the form

$$\frac{\partial \bar{u}}{\partial z} = \frac{u_*}{\kappa z} \xi f'(\xi) \sim \frac{u_*}{\kappa z}; \quad \frac{\partial \bar{T}}{\partial z} = \frac{T_*}{\alpha_0 z} \frac{\alpha_0 \xi f'(\xi)}{\alpha(\xi)} \sim \frac{T_*}{\alpha_0 z}, \quad (34.4')$$

where $\alpha_0 = \alpha(0)$; in order for the buoyancy parameter to drop out of the asymptote, the condition $\xi f'(\xi) \sim 1$ must be satisfied (assigning to this asymptotic constant a value of 1, we thereby reserve for the constant κ the same meaning as in the familiar logarithmic law $\bar{u}(z) = (u^*/\kappa) \ln z + \text{const}$ in turbulent boundary layers in an incompressible fluid). Note that the roughness height z_0 , which was defined above, can generally be assumed in geophysics to always satisfy the condition $z_0 \ll |L|$, that is, it is independent of the stratification of the medium, and thus characterizes not the state of the medium but rather the underlying surface itself (in the atmospheric surface layer z_0 usually ranges from 0.1 to 0.2 of the typical height h_0 of roughnesses of this surface). From (34.4') it is clear that for $|\xi| \ll 1$ we get the asymptotes:

$$f(\xi) \sim \ln|\xi| + \beta\xi + \dots; \quad f_1(\xi) \sim \frac{1}{\alpha_0} (\ln|\xi| + \beta_1\xi + \dots), \quad (34.8)$$

where both β and β_1 for different signs of ξ may assume different values (but positive values, in order for the profiles of velocity and temperature for stable

stratification, when turbulent exchange is hampered, to be steeper, and for unstable stratification, when the exchange is enhanced, to be less steep, than for neutral stratification). Taking into account that for $|\xi| \ll 1$ the parameter g/T_0 must drop out, and in addition that for neutral stratification there are no temperature fluctuations, from (34.3) we obtain the asymptote

$$p_M(\mathbf{u}', T') \sim u_*^{-3} F_0 \left(\frac{\mathbf{u}'}{u_*} \right) \delta(T'). \quad (34.9)$$

Now let us go on to the asymptote for strong instability $\xi \rightarrow -\infty$. We note that the condition $z \gg |L|$ may imply, for fixed $q > 0$, g/T_0 , and z , passage to the limit $u_* \rightarrow 0$ and to purely thermal turbulence with free convection (in the absence of a mean flow). In this case the function characterizing the mean temperature profile $f'_1(\xi) = f'(\xi)/\alpha(\xi)$ has to be such that parameter u_* drops out of the first formula of (34.4') for $\partial \bar{T} / \partial z$:

$$f'_1(\xi) \sim -\frac{1}{3} C_1 \kappa^{4/3} \xi^{-4/3}; \quad \bar{T}(z) \sim \bar{T}_\infty + C_1 \left(\frac{q}{c_p \rho_0} \right)^{2/3} \left(\frac{gz}{T_0} \right)^{-1/3}; \quad (34.10)$$

$$K_T \sim \frac{3}{C_1} \left(\frac{g}{T_0} \frac{q}{c_p \rho_0} \right)^{1/3} z^{4/3},$$

where C_1 is a positive numerical constant. Therefore, the heat-exchange coefficient increases rapidly with height (due to the appearance of increasingly larger convective eddies, comparable in size to the distance z from the Earth's surface), while the temperature distribution approaches isothermy. The free convection here is only an asymptotic form, so that a mean flow is not absent. Its profile can be calculated, starting from the fact that no dimensionless combination of the quantities determining the asymptote, namely $q/c_p \rho_0$, g/T_0 , and z , can be formed, so that any dimensionless characteristic of the asymptotic regime has to be constant. In particular, $f'(\xi)/f'_1(\xi) \sim \alpha_{-\infty} = \text{const}$, which yields

$$f'(\xi) \sim -\frac{C_2}{3} \xi^{-4/3}; \quad f(\xi) \sim C_2 \xi^{-1/3} + \text{const}; \quad C_2 = \alpha_{-\infty} \kappa^{4/3} C_1. \quad (34.11)$$

Function $\bar{u}(z)$ may not be independent of u_* . According to (34.4), for asymptote (34.11) it turns out to be proportional to u_*^2 . As for the single-point probability distribution for the fluctuations, when u_* drops out of (34.3) we have to set

$$p_M(\mathbf{u}', T') \sim \left(\frac{q}{c_p \rho_0} \right)^{-5/3} \left(\frac{gz}{T_0} \right)^{-2/3} F_1 \left[\mathbf{u}' \left(\frac{q}{c_p \rho_0} \frac{gz}{T_0} \right)^{-1/3}, \right. \\ \left. T' \left(\frac{q}{c_p \rho_0} \right)^{-2/3} \left(\frac{gz}{T_0} \right)^{1/3} \right]. \quad (34.12)$$

This leads us, in particular, to the following formulas for the rms values of the

fluctuations (Obukhov, 1960):

$$\begin{aligned}\sigma_u \approx \sigma_v \sim c_1 \left(\frac{q}{c_p \rho_0} \frac{gz}{T_0} \right)^{1/3}; \quad \sigma_w \sim c_2 \left(\frac{q}{c_p \rho_0} \frac{gz}{T_0} \right)^{1/3}; \\ \sigma_T \sim c_3 \left(\frac{q}{c_p \rho_0} \right)^{2/3} \left(\frac{gz}{T_0} \right)^{-1/3},\end{aligned}\quad (34.13)$$

where c_1, c_2, c_3 are certain numerical constants.

Finally, let us consider the asymptote for strong stability $\zeta \rightarrow +\infty$, starting from the fact that with increasing stability the dynamic Richardson number Rf must increase (monotonically, of course), albeit within bounds (as was shown above during the discussion of energy Equation (33.6), $Rf < Rf_{cr} < 1$). Consequently, with increasing $\zeta > 0$ the value of Rf approaches some constant $R \leq Rf_{cr}$ (even as, in general, because of the weakening of the turbulent exchange between different fluid layers, the turbulence becomes localized, that is, its characteristics stop depending on z . For $Rf = Rf_{cr}$ it degenerates completely, and fluctuating motions of the medium become possible only in the form of internal gravity waves. However, when the latter lose their stability, turbulent patches appear, which then spread into thin layers, so that a thinly laminated microstructure is formed, see Section 35). Taking (34.6) into account, we then have

$$\begin{aligned}f'(\zeta) \sim \frac{1}{R}; \quad f(\zeta) \sim \frac{\zeta}{R} + \text{const}; \\ \bar{u}(z) \sim \frac{u_* z}{\kappa R L} + \text{const}.\end{aligned}\quad (34.14)$$

Consequently, the velocity profile asymptotically becomes linear. This will not be the case, however, for the temperature profile, the gradient of which is defined by the function $f'_1(\zeta) = f'(\zeta)/\alpha(\zeta) \sim 1/R\alpha(\zeta)$. It is to be expected that $\alpha(\zeta) \rightarrow 0$ as $\zeta \rightarrow +\infty$, since the limit stable stratification is the interface between the lower heavy fluid and the upper light fluid, through which momentum can be transferred by turbulent pressure fluctuations, so that $K \neq 0$, whereas heat is transferred only by molecular processes and radiation, that is, $K_T = 0$ (the semiempirical theory will be given below, according to which $\alpha \sim \zeta^{-2}$). From (34.6) we see that here $Ri \rightarrow \infty$, so that, in contrast to Rf , no finite Ri_{cr} exists. Note, finally, that the probability distribution for the fluctuations (34.3) must stop depending on ζ as $\zeta \rightarrow +\infty$.

The above asymptotic regularities were verified empirically by Monin and Obukhov (1953), who selected the values of u_* and L on the basis of the measured profiles of the wind velocity and temperature (coefficient β in (34.8) was, averaged over the interval of ζ considered by them, equal to 0.6). They also constructed the empirical function

$$f(\zeta) - f\left(\pm \frac{1}{2}\right) = \frac{\kappa}{u} \left[\bar{u}(z) - \bar{u}\left(\frac{|L|}{2}\right) \right] \quad (34.15)$$

(see Figure 34.1, which depicts the above-mentioned asymptotes as dashed lines). The predictions of similarity theory were also verified in many later studies.

The indicated similarity theory suggests that the probability distribution for the values of $\mathbf{u}'/u_*$, T'/T_* at points $(\mathbf{x}_{h1}, z_1, t_1), \dots, (\mathbf{x}_{hn}, z_n, t_n)$ may depend only on the arguments $(\mathbf{x}_{h2} - \mathbf{x}_{h1})/|L|, \dots, (\mathbf{x}_{hn} - \mathbf{x}_{h1})/|L|; z_1/L, \dots, z_n/L, (t_2 - t_1)u_*/|L|, \dots, (t_n - t_1)u_*/|L|$ (provided that all the heights $z_1, \dots, z_n$ lie in the air layer near the ground and are much greater than z_0 , while the distances between different points $(\mathbf{x}_{hk}, z_k)$ and the nonzero intervals $|t_k - t_l|$ are not too small, so that the requirements of SHH flow are met). The asymptotes for such a multipoint probability distribution, approaching neutral, highly unstable, and very stable stratifications, can be established with the aid of considerations analogous to those presented above.

The statistical characteristics of the derivatives of the fields $\mathbf{u}(M)$ and $T(M)$, that is, of the differences of values of these fields at points with infinitely close values of any of the space coordinates or time, may, generally speaking, already depend on the molecular viscosity ν and the thermal conductivity χ . This does not, however, pertain to the specific rates of dissipation of turbulent energy ε and decay of inhomogeneities of the temperature field ε_T , which are defined as

$$\varepsilon = \frac{\nu}{2} \sum_{\alpha, \beta} \overline{\left(\frac{\partial u'_\alpha}{\partial x_\beta} + \frac{\partial u'_\beta}{\partial x_\alpha} \right)^2} = \frac{u_*^3}{\kappa z} \varphi_\varepsilon(\zeta); \quad (34.16)$$

$$\varepsilon_T = \chi \overline{(\nabla T')^2} = \frac{u_* T_*^2}{z} \varphi_{\varepsilon_T}(\zeta),$$

where φ_ε and φ_{ε_T} are certain universal functions, the asymptotes of which can be established on the basis of the above-indicated rules. Actually, ε can be expressed in terms of the single-point moments using the equation for the energy of turbulence (33.6), while ε_T can be expressed using the analogous equation for σ_T^2 (which will be presented below).

Now let us go on to determine the spatial structure functions (33.10) of turbulence in stratified media or, equivalently, their Fourier transforms $F_{j_1 j_2}(\mathbf{k})$, entering into the formulas:

$$D_{j_1 j_2}(\mathbf{r}) = 2 \int (1 - \cos \mathbf{k} \cdot \mathbf{r}) F_{j_1 j_2}(\mathbf{k}) d\mathbf{k} \quad (34.17)$$

and known as the three-dimensional *spectral densities*. In particular, let us consider the three-dimensional spectral density of the specific kinetic energy of turbulence

$$F(\mathbf{k}) = \frac{1}{2} \sum_{\alpha=1}^3 F_{\alpha\alpha}(\mathbf{k})$$

and the result of its integration over all the *directions* of the wave vector $\mathbf{k}$, the so-called *wave-number spectrum* $E(k)$. First we will examine the locally isotropic turbulence (recalling that this includes any developed geophysical turbulence for

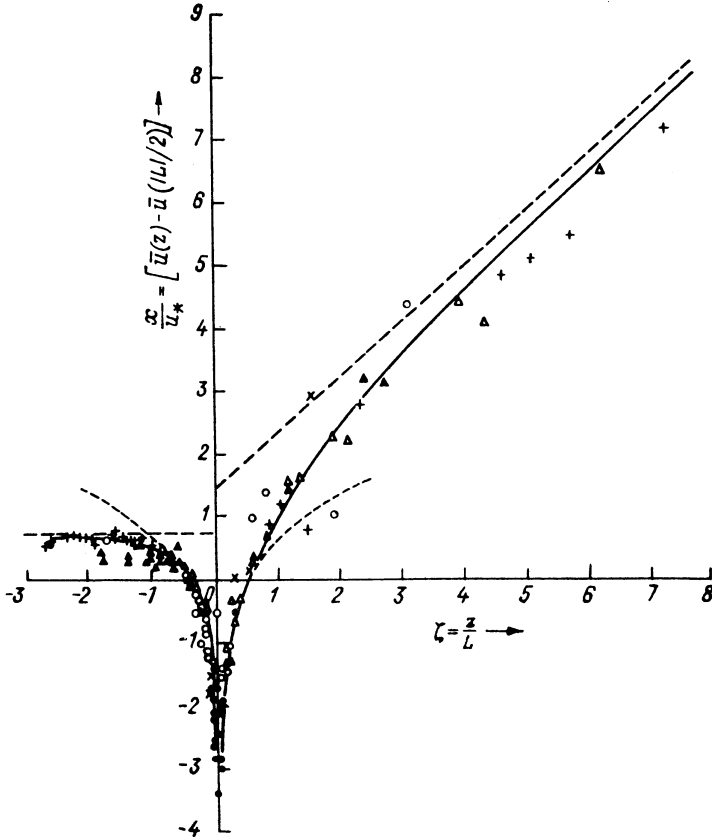


Fig. 34.1. Empirical universal function $f(\xi) - f(\pm \frac{1}{2})$, according to Monin and Obukhov (1953).

neutral stratification); in this case $F(\mathbf{k})$ may depend only on the wave number $k = |\mathbf{k}|$ and $E(k) = 4\pi k^2 F(k)$.

According to the similarity hypotheses of Kolmogorov (1941), for locally isotropic turbulence the finite-dimensional probability distributions for the relative velocities $\mathbf{u}(M) - \mathbf{u}(M_0)$ in the case of sufficiently small (relative to the scales L_0 and L_0/U of the flow as a whole) distances r and δt between observation points may depend only on the two parameters ε and ν , so that when measuring all the lengths to a scale $\eta = \varepsilon^{-1/4} \nu^{3/4}$, the velocities to a scale $v_\eta = \varepsilon^{1/4} \nu^{1/4}$, and the times to a scale $t_\eta = \eta / v_\eta$, we find that these probability distributions become universal; if, however, in addition, all $r \gg \eta$ and $|\delta t| \gg t_\eta$, then these distributions stop being functions of ν and thus are determined uniquely by a single parameter ε . According to these similarity hypotheses, for $k \gg L_0^{-1}$ or $r \ll L_0$

$$E(k) = \varepsilon^{1/4} \nu^{5/4} \varphi(\eta k) \sim C_1 \varepsilon^{2/3} k^{-5/3};$$

$$D_{ij}(r) \sim \frac{27}{55} \Gamma\left(\frac{1}{3}\right) C_1 (\varepsilon r)^{2/3} \left(-\frac{1}{3} \frac{r_i r_j}{r^2} + \frac{4}{3} \delta_{ij}\right), \quad (34.18)$$

where the asymptote corresponds to the conditions $k \ll \eta^{-1}$ or $r \gg \eta$ (that is, to the so-called *inertia interval* $L_0^{-1} \ll k \ll \eta^{-1}$ or $L_0 \gg r \gg \eta$).

The second of these formulas expresses the familiar *two-thirds law* of Kolmogorov, and the asymptote in the first formula corresponds to its spectral formulation (*five-thirds law*) according to Obukhov (1941). These laws have repeatedly been verified empirically (with a coefficient $C_1 \approx 1.4$). For fields of passive admixtures, which in the case of neutral stratification also includes the temperature T , according to Obukhov (1949) for large Reynolds numbers Re and Péclet numbers $Pe = Re \cdot Pr$, where $Pr = \nu/\chi$ is the Prandtl number, we must add to the controlling parameters ε and ν the parameters ε_T and χ , which for $k \gg L_0^{-1}$ yields

$$E_T(k) = \varepsilon_T \varepsilon^{-3/4} \chi^{5/4} \varphi_T(\eta k, Pr) \sim C_T \varepsilon_T \varepsilon^{-1/3} k^{-5/3}, \quad (34.19)$$

where the asymptote corresponds to the condition $k \ll \eta_T^{-1} = \eta^{-1}(Pr)^{3/4}$ (the interval $L^{-1} \ll k \ll \eta_T^{-1}$ is called the *convective interval*).

If the medium is stratified but not neutrally, then, as explained above, the turbulence will no longer be locally isotropic, being just locally axisymmetric about the vertical. Here the scalar spectral densities $F(\mathbf{k})$ no longer depend just on $k = |\mathbf{k}|$, but on k_z as well. Moreover, the relation between $E(k)$ and $F(\mathbf{k})$ varies, and the structure of the tensor $D_{ij}(\mathbf{r})$ turns out to be more complicated than in (34.18), while the similarity hypotheses of Kolmogorov and Obukhov vary little: according to Obukhov (1959), to the controlling parameters given in them we need to add only the buoyancy parameter g/T_0 . On the other hand, the dependence of the finite-dimensional probability densities on u_* and $q/c_p \rho_0$ predicted by the similarity theory of Monin and Obukhov for both the small-scale and large-scale turbulence components will in the case of just the small-scale components alone, to which the similarity hypotheses of Kolmogorov and Obukhov apply, be realized via ε and ε_T in accordance with formulas (34.16).

Let us confine ourselves to the inertial-convective interval of the scales, in which now there are three controlling parameters ε , ε_T , and g/T_0 . Using these, we can construct a single combination having the dimension of length: the so-called *buoyancy scale*

$$L_* = \varepsilon^{5/4} \varepsilon_T^{-3/4} \left(\frac{g}{T_0} \right)^{-3/2}. \quad (34.20)$$

If it falls within the bounds of the inertial-convective interval, then the turbulence spectra will be determined by asymptotes (34.18) and (34.19), in which now coefficients C_1 and C_T are certain functions of the dimensionless wave number kL_* . They also describe the effect of stratification on the spectra in the indicated interval. In it for $k \gg L_*^{-1}$ these functions approach certain limits (their values for neutral stratification).

For $k \ll L_*^{-1}$ (the 'buoyancy interval') in the case of stable stratification, according to the hypothesis of Bolgiano (1959), parameter ε should drop out of the formulas for the spectra (since the main part of the turbulent energy is spent on working against the buoyancy forces in the region of large scales, only a very

small part going for the small-scale region and being dissipated there as heat; thus, ε turns out to be very small and cannot in any way influence much the statistical regime of the large-scale components of the turbulence). For this, we must have $C_1 \sim (kL_*)^{-8/15}$ and $C_T \sim (kL_*)^{4/15}$, and spectra (34.18) and (34.19) at the large-scale end of the inertial-convective interval, accurate to numerical factors, will be

$$E(k) \sim \varepsilon_T^{2/5} \left(\frac{g}{T_0} \right)^{4/5} k^{-11/5}; \quad E_T(k) \sim \varepsilon_T^{4/5} \left(\frac{g}{T_0} \right)^{-2/5} k^{-7/5}. \quad (34.21)$$

These predictions were verified by measurements of turbulence spectra in the ocean (Nozdrin, 1975). Note, too, that in the case of locally axisymmetric turbulence the fluctuations of the velocity $\mathbf{u}'$ and the temperature T' correlate with one another (whereas in a field of locally isotropic turbulence they are uncorrelated). In the buoyancy interval it follows from dimensionality considerations that the spectrum of the turbulent heat flux $c_p \rho_0 E_{u,T}(k)$ must be proportional to $[E(k) E_T(k)]^{1/2}$.

A semiempirical theory of spectra of geophysical turbulence was constructed by Monin (1962) with the aid of spectral analogs of the equations for the second moments:

$$E_3(k) - \frac{g}{T_0} \int_k^\infty E_{wT}(k) dk = 2\nu \int_k^\infty k^2 E(k) dk; \quad (34.22)$$

$$E_{uTT}(k) = \chi \int_k^\infty k^2 E_T(k) dk,$$

The spectra of the third moments and the heat flux appearing on the left-hand sides of these equations were represented in the manner of the semiempirical theory of Heisenberg (1948) as

$$E_3(k) = K(k) \left[\left(\frac{\partial \bar{u}}{\partial z} \right)^2 + 2 \int_{k_0}^k k^2 E(k) dk \right];$$

$$E_{uTT}(k) = \alpha K(k) \left[\left(\frac{\partial \bar{T}}{\partial z} \right)^2 + \int_{k_0}^k k^2 E_T(k) dk \right]; \quad (34.23)$$

$$\int_k^\infty E_{wT}(k) dk = \pm \alpha K(k) \left[\left(\frac{\partial \bar{T}}{\partial z} \right)^2 + \int_{k_0}^k k^2 E_T(k) dk \right]^{1/2};$$

$$K(k) = \gamma \left[\int_k^\infty E(k) k^{-2} dk \right]^{-1/2},$$

where $K(k)$ is the eddy viscosity, produced by the small-scale components of the turbulence (with wave numbers higher than k). For this we will use here an expression suggested by Howells (1960). The rest of the notation is explained as follows: γ is a numerical constant; $k_0 = 2\pi/L_0$ corresponds to the outer scale of turbulence; the upper sign (plus or minus) refers to unstable stratification, and the lower sign to stable stratification. In the inertial-convective interval and the buoyancy interval the right-hand sides of (34.22) are replaced, respectively, by ε and ε_T . For $k = k_0$ we arrive at equations for the mean fields $\bar{u}(z)$ and $\bar{T}(z)$, considered above and assumed to be known for the spectrum calculations. In the following it will be convenient to introduce the dimensionless quantities:

$$\begin{aligned} f(kL_1) &= \left(\frac{\varepsilon}{\gamma} \right)^{-2/3} L_1^{-5/3} E(k); \\ f_T(kL_1) &= \frac{\alpha\gamma}{\varepsilon_T} \left(\frac{\varepsilon}{\gamma} \right)^{1/3} L^{-5/3} E_T(k); \end{aligned} \quad (34.24)$$

$$F(x) = \left[\int_x^\infty f(x)x^{-2} dx \right]^{1/4};$$

$$L_1 = \left[\frac{g}{T_0} \left(\frac{\alpha\varepsilon_T}{\gamma} \right)^{1/2} \right]^{-3/2} \left(\frac{\varepsilon}{\gamma} \right)^{5/4},$$

and then the solution of Equations (34.22), (34.23) becomes

$$\begin{aligned} x &= \frac{F^{-3/2}(4 \mp 2F)^{1/4}}{2}; & f &= \frac{4F^{5/2}(4 \mp 2F)^{5/4}}{12 \mp 5F}; \\ f_T &= \frac{32F^{5/2}(4 \mp 2F)^{1/4}}{12 \mp 5F}. \end{aligned} \quad (34.25)$$

Thus, for low F we obtain asymptotic five-thirds laws (34.18), (34.19), and for high F and the lower sign (plus or minus) we have asymptote (34.21).

35. VERTICAL THINLY LAMINATED MICROSTRUCTURE

As stated above, for very stable stratification the turbulence is not able to encompass the entire fluid for very long, because its energy is consumed in working against buoyancy forces, and it can exist only in individual layers, or 'blini', as they were called by Kolmogorov, who predicted this phenomenon back in 1948–1949 (the term 'blini' has been used in works on turbulence in English as well). Due to mixing, the turbulent fluid layers must be quasihomogeneous along

the vertical and they are separated by very thin films (sheets) with microdiscontinuities ('jumps') of temperature, density, and other thermodynamic parameters.

This prediction was tentatively verified in the 1950s during observations of 'clouds of turbulence' in the free atmosphere by manned balloon observations carried out by the Geophysical Institute of the Soviet Academy of Sciences. Striking verification was later achieved, in the 1970s, with the discovery of a vertical thinly laminated microstructure of the ocean. This microstructure was found to be present practically everywhere in the World Ocean at almost all times (it should be absent only in regions of large-scale convection, which are encountered in the ocean only quite rarely, at least in the low and middle latitudes).

The stability is especially high in the 'layer of density discontinuity' (*pycnocline*) beneath the upper mixed layer of the ocean. There it is usually true that $|\partial\rho/\partial z| \sim 10^{-6} \text{ g/cm}^4$ and $N \sim 3 \times 10^{-2} \text{ s}^{-1}$ (whereas at the greatest temperature inversions in the atmosphere near the ground $|\partial\rho/\partial z| \sim 4 \times 10^{-8} \text{ g/cm}^4$ and, averaged over the bottom 10 km of the atmosphere, $N \sim 10^{-2} \text{ s}^{-1}$). The microstructure layers in the ocean have thicknesses $h \approx 10^0$ to 10^3 cm , the turbulence in them is weak (incapable of breaking up these layers), it is local in nature (does not vary directly with the depth), and it is characterized by low Reynolds numbers $\nu^{-1}hU \sim 10^2$ to 10^5 .

The main mechanism of formation of the thinly laminated microstructure, which operates practically everywhere in the ocean, is spreading of the mixed patches which form in regions of shear instability and convective hydrodynamic instability of the internal waves, at their troughs and crests, when the local values of Ri there are less than $\frac{1}{4}$ (and also in shear flows $U(z)$ at critical levels z_c , where $U(z_c) = c$, see (27.8')), as well as due to 'explosive' instability, see Section 30).

For stable stratification, when the density of the fluid (or, more precisely, its potential density) increases with depth, in the upper half of a mixed volume of fluid the density turns out to be higher, being lower in the lower half, than at the same levels in the surrounding medium. Under the influence of buoyancy forces, the upper half of a mixed volume should descend, and the lower half should rise, to its mean level. Therefore, the fluid volume should collapse, spreading out flat due to the weakness of the compressibility and thus becoming a thin 'blin'. This intrusion into the surroundings produces in it a new layer of the microstructure.

If the initial internal wave was long (and with a long period, as, for instance, in the case of forced internal waves with tidal periods), then the patches of turbulence produced by it will be large, and the layers of microstructure originating from them will be thick. In these, internal waves of shorter length can develop, forming patches of turbulence of smaller size, spreading out into thinner layers of the microstructure, etc., until internal waves of minimum length are reached, along with the smallest patches of turbulence and the thinnest layers of microstructure. Therefore, in this case the answer to the question: "Which came first, the chicken or the egg?" takes the form of the cascade process 'stable stratification $\rightarrow$ internal waves $\rightarrow$ patches of turbulence $\rightarrow$ layers of microstructure $\rightarrow$ internal waves, etc.'

The initial stages of the spreading of a turbulent patch were studied experimentally by Wu (1969) and then theoretically described quantitatively by Kao

(1976). A patch spreads out, at first easily overcoming the forces of resistance and intensely emitting internal waves and then reaching a quasisteady state in which the driving force of spreading (collapse) is equalized by the sum of the form drag and the wave drag. This stage is completed in a time equal to several tens of Brunt–Väisälä periods. The relative rate of increase in the horizontal area S of a patch $S^{-1} \partial S / \partial t$ is at first proportional to the collapse rate $N^2 t$ (since the buoyant acceleration is proportional to N^2). In this case the horizontal size of a patch increases according to the law $(L - L_0)/L_0 \sim N^2 t^2$. Then a regime $\partial L / \partial t \sim Nh$ is reached, where h is the thickness of the intrusion layer; for a large patch $h \sim V/L^2$ (where V is the patch volume, constant in time), so that $L \sim [NV(t - t_0)]^{1/3}$.

The next stage is a prolonged final stage of spreading of the mixed patch, during which the driving force of spreading is compensated by the viscous resistance (Barenblatt, 1978; Barenblatt and Monin, 1979). The equation describing the conservation of mass of the patch here has the form

$$\frac{\partial h}{\partial t} + \operatorname{div} h \mathbf{v} = 0; \quad \mathbf{v} = -\frac{h}{c\mu} \nabla(ph), \quad (35.1)$$

where the second equation represents equilibrium between the driving force $-(\delta S)\nabla(ph)$ and the resistance $(\delta S)c\mu\mathbf{v}/h$ (where μ is the viscosity; the numerical factor c , for example, in a flow between flat walls, is equal to 12). Here p is the excess pressure, that is, the difference between the pressure in the fluid

$$p_1 - \rho_1 g(z - z_1) + \frac{1}{2} \rho_1 N^2 (z - z_1)^2$$

and the pressure in the intrusion

$$p_1 - \rho_1 g(z - z_1) + \frac{1}{2} \rho_1 N^2 \left(\frac{h}{2} \right)^2,$$

equal on the average to $\frac{1}{12} \rho_1 N^2 h^2$. Consequently, Equations (35.1) become

$$\frac{\partial h}{\partial t} = \kappa \Delta h^5; \quad \kappa = \frac{\rho_1 N^2}{20c\mu}. \quad (35.2)$$

For an axisymmetric patch the solution has the form

$$h = \left(\frac{V}{2\pi\kappa} \right)^{1/5} (t - t_1)^{-1/5} f(\xi); \quad \xi = \left(\frac{\kappa V^4}{16\pi^4} \right)^{-1/10} (t - t_1)^{-1/10} r; \quad (35.3)$$

$$f(\xi) = \left(\frac{10^{1/5}}{6} \right)^{1/4} \left(1 - \frac{\xi^2}{\xi_0^2} \right)^{1/4}; \quad 0 \leq \xi \leq \xi_0 = \frac{10^{3/5}}{2},$$

so that the patch radius ($\xi = \xi_0$) increases as $(t - t_1)^{1/10}$. It may be that such mixed lenses in the atmosphere, with the dust retained in them by the turbulence, can under certain illumination conditions be taken to be ‘flying saucers’. For a cylindrical turbulence path with a horizontal axis (the y axis, say), produced by a plane internal wave, instead of (35.3) we obtain the following solution of Equation

(35.2):

$$h = \left(\frac{V^2}{4\kappa H^2} \right)^{1/6} (t - t_1)^{-1/6} f_1(\xi_1); \quad \xi_1 = \left(\frac{\kappa V^4}{16H^4} \right)^{-1/6} (t - t_1)^{-1/6} x;$$

$$f_1(\xi_1) = \left(\frac{\xi_0^2}{15} \right)^{1/4} \left(1 - \frac{\xi^2}{\xi_0^2} \right)^{1/4}; \quad (35.4)$$

$$0 \leq \xi \leq \xi_0 = (15)^{1/6} \frac{[2\Gamma(5/4)\Gamma(1/2)]^{2/3}}{\Gamma(7/4)},$$

where H is the length of the patch along the y axis. The laws of expansion of an axisymmetric patch $r \sim (t - t_1)^{1/10}$ and a cylindrical patch $x \sim (t - t_0)^{1/6}$ have been verified by laboratory experiments.

Another mechanism of formation of a thinly laminated microstructure by internal waves was described in Section 29, where it was mentioned that a packet of slightly nonlinear waves in the (x, z) plane can be described by the stream function

$$\psi = \varepsilon(Ae^{i\theta} + A^*e^{-i\theta})\xi_1(z) + \varepsilon^2(A^2e^{2i\theta} + A^*e^{-2i\theta})\xi_2(z) + \varepsilon^2|A|^2\Psi(z) + o(\varepsilon^2). \quad (35.5)$$

Here ε is a small parameter of the order of the wave steepness; $A = A(\varepsilon^2x, \varepsilon^2t, \theta)$ is the complex envelope; $\theta = \varepsilon^{-2}\vartheta(\varepsilon^2x, \varepsilon^2t)$ is the 'fast' phase; the first two terms describe the fundamental wave and the second harmonic, while the third corresponds to a nonlinear contribution of the Stokes drift type to the mean flow. This contribution (and the density disturbances corresponding to it) has an alternating sign along the vertical with a wavelength of the order of $c_g N^{-1}$; for small lengths of the carrier wave it is itself small, so that it forms a thinly laminated microstructure (existing for the time Lc_g^{-1} of passage of a wave packet of length L). The existence of a finely laminated microstructure of the mean flow in the ocean was discovered during measurements carried out by Monin *et al.* (1973). According to (26.14) and (26.15), for a carrier wave with an approximately vertical wave vector, and thus an approximately inertial frequency $\omega \approx f$, we obtain

$$\mathbf{c}_g \approx \frac{\omega^2 - f^2}{fk^2} \mathbf{k},$$

that is, the wave packet propagates approximately horizontally. Apparently, this is the scenario corresponding to the measurements of Sanford (1975), who detected an inertial period in the oscillations in time of the thinly laminated flow microstructure.

Still another possible mechanism of formation of a thinly laminated microstructure in the ocean, which may by no means operate everywhere, is microconvection in a layer of salt water with a stable density stratification

$$\frac{\partial \rho_*}{\rho \partial z} \approx -\alpha \frac{\partial T}{\partial z} + \gamma \frac{\partial s}{\partial z} < 0,$$

but with opposite signs of the contributions to it of the gradients of temperature and salinity (this is anomalous; usually the temperature drops, while the salinity increases, with increasing depth, that is, $\partial T/\partial z > 0$ and $\partial s/\partial z < 0$); this microconvection appears as a result of molecular processes of *double diffusion*, that is, due to a difference in the coefficients of diffusion of heat and salt in water $\chi/D \approx 100$ (Stommel *et al.*, 1956; Stommel, 1962).

For instance, if warm salt water lies over denser, less saline water ($\partial T/\partial z > 0$; $\partial s/\partial z > 0$), the relatively rapid compensation of the temperature fluctuations, the salinity fluctuations being conserved due to the slowness of the diffusion, produces density fluctuations that lead to laminar convection in the form of narrow and long vertical cells, known as *salt fingers*. These were discovered, and even photographed using the shadow method, under Mediterranean waters in the Atlantic (Williams, 1974). Laboratory measurements indicated that in this case $\alpha q_z/\gamma I_z \approx 0.56$, that is, more than half of the potential energy liberated during the concentration (salinity) convection is used for the transfer of heat.

On the other hand, if cold, relatively fresh water lies over denser warm saline water, then microconvection due to the rapid downward diffusion of heat is hindered by the microdiscontinuities of density, and *layering* takes place, whereby the dimensionless fluxes of heat and salt $q_z(c_p \rho_0 \chi \delta T/h)^{-1}$ and $I_z(\rho_0 D \delta s/h)^{-1}$ are proportional to $(Ra)^{1/3}$ (where Ra is the Rayleigh number, see Section 17), with the proportionality factors depending on $\gamma \delta s/\alpha \delta T$. According to laboratory experiments, the ratio $\gamma I_z/\alpha q_z$ at first drops rapidly with increasing $\gamma \delta s/\alpha \delta T$, and then for $\gamma \delta s/\alpha \delta T \geq 2$ it becomes approximately equal to 0.15; if the salinity makes a major contribution to the density stratification, then 15% of the potential energy liberated during thermal convection goes to lift salt.

However, the microconvection caused by double diffusion can by no means take place everywhere, while thinly laminated microstructure is observed even in fresh-water bodies, where double diffusion does not occur at all. For example, Simpson and Woods (1970) detected microstructure in the famous Scottish lake Loch Ness.

36. BOUNDARY LAYERS

Let us now apply the similarity theory for the atmospheric boundary layer, presented in Section 34, to the entire planetary boundary layer of the atmosphere (or the upper mixed layer of the the ocean, or the near-bottom boundary layer, etc.). In the case of SHH flow this layer differs from the atmospheric boundary layer only by the effect of the Coriolis force (then it is called the *Ekman boundary layer*), so that we just have to add the Coriolis parameter f to the controlling parameters of the theory presented in Section 34 (a similarity theory so generalized was formulated by Monin, 1950).

Then we can introduce the thickness of the neutrally stratified Ekman layer $h_0 = \kappa u_* / f$ and the internal stratification parameter $\mu_0 = h_0 / L$ (Kazanskii and Monin, 1960, 1961), so that the thickness of the stratified Ekman boundary layer will be $h = h_0 \zeta(\mu_0)$, where $\zeta(\mu_0)$ is some universal function. Defining the latter as the thickness of momentum loss, we can set $h \sim (K/f)^{1/2}$, where K is the effective value of the kinematic eddy viscosity. According to the rules set forth in Section

34, for strong instability K asymptotically stops depending on u_* , so that

$$K \sim \left(\frac{g}{T_0} \right)^{1/3} \left(\frac{q}{c_p \rho_0} \right)^{1/3} h^{4/3},$$

from which we have

$$h \sim \left(\frac{g}{T_0} \right)^{1/2} \left(\frac{q}{c_p \rho_0} \right)^{1/2} f^{-3/2} \sim h_0 |\mu_0|^{1/2}. \quad (36.1)$$

For strong stability

$$Rf \sim \left(\frac{g}{T_0} \right) \left(\frac{q}{c_p \rho_0} \right) \left(\frac{K}{u_*^4} \right) \sim R.$$

After finding K from this, we get

$$h \sim h_0 \mu_0^{-1/2}. \quad (36.2)$$

With increasing stability, and also with increasing instability, $|L|$ decreases and $|\mu_0|$ increases. Therefore, with a weakening of the stability and an increase in instability, the thickness h of the layer of attenuation of the momentum flow in a stratified Ekman boundary layer becomes larger.

The same generalization of the similarity theory of Section 34 is obtained if to the controlling parameters we add, instead of f , the boundary-layer thickness h (and if we introduce the internal stratification parameter $\mu = h/L$). This approach has the advantage of an even greater generality: it applies not only to the Ekman boundary layer, but also to quasi-SHH boundary layers, depending on $(\mathbf{x}_h, t)$ through h (and through the other controlling parameters of similarity theory). Instead of (34.3), we now have

$$p_M(\mathbf{u}, T) = u_*^{-3} |T_*|^{-1} F \left(\frac{\mathbf{u}}{u_*}, \frac{T}{T_*}, \frac{z}{h}, \frac{z_0}{h}, \mu \right), \quad (36.3)$$

where the only functions of z_0 are the *absolute* values of the *mean* flow velocity and *mean* temperature (but not their gradients and not their fluctuations):

$$\begin{aligned} \bar{u}(z) &= \bar{u}(z_0) + \frac{u_*}{\kappa} \left[f_u \left(\frac{z}{h}, \mu \right) - f_u \left(\frac{z_0}{h}, \mu \right) \right]; \\ \bar{v}(z) &= \bar{v}(z_0) + \frac{u_*}{\kappa} \left[f_v \left(\frac{z}{h}, \mu \right) - f_v \left(\frac{z_0}{h}, \mu \right) \right]; \\ \bar{T}(z) &= \bar{T}(z_0) + T_* \left[f_T \left(\frac{z}{h}, \mu \right) - f_T \left(\frac{z_0}{h}, \mu \right) \right]. \end{aligned} \quad (36.4)$$

For $z \ll h$ the dependence on h in (36.3), (36.4) disappears, so that universal functions F , f_u , f_v , and f_T will no longer depend on all three of the variables z/h , z_0/h , and μ , but rather just on two products of these $\mu z/h = z/L$ and $\mu z_0/h = z_0/L$, that is, we return to the similarity theory of Section 34. If, in addition, $z \ll |L|$, then formulas (36.4) must become the 'wall' logarithmic laws

$$\begin{aligned}\bar{u}(z) &\approx U \cos \gamma + \frac{u_*}{\kappa} \ln \frac{z}{z_0}; & \bar{v}(z) &\approx -U \sin \gamma; \\ \bar{T}(z) &\approx \bar{T}(z_0) + \frac{T_*}{\alpha_0} \ln \frac{z}{z_0},\end{aligned}\tag{36.5}$$

where the x axis is in the direction of the frictional stress vector at the level $z = z_0$ (lying at an angle γ to the flow outside the boundary layer), and U is the modulus of the velocity at this level. At the upper surface of the boundary layer, on the other hand, the following 'laws of the defect' must hold true:

$$\begin{aligned}\bar{u}(z) &= \frac{u_*}{\kappa} \psi_u \left(\frac{z}{h}, \mu \right); & \bar{v}(z) &= \frac{u_*}{\kappa} \psi_v \left(\frac{z}{h}, \mu \right); \\ \bar{T}(z) - \bar{T}(h) &= T_* \psi_T \left(\frac{z}{h}, \mu \right).\end{aligned}\tag{36.6}$$

For low z these 'laws of the defect' will be valid simultaneously with the 'wall' laws (36.5). Equating these to one another, we arrive at the following limit values of the dimensionless quantities as $z/h \rightarrow 0$:

$$\begin{aligned}\frac{\kappa U \cos \gamma}{u_*} - \ln \frac{z_0}{h} &= \psi_u \left(\frac{z}{h}, \mu \right) - \ln \frac{z}{h} \rightarrow A(\mu); \\ -\frac{\kappa U \sin \gamma}{u_*} &= \psi_v \left(\frac{z}{h}, \mu \right) \rightarrow B(\mu); \\ \alpha_0 \frac{\delta T}{T_*} + \ln \frac{z_0}{h} &= -\alpha_0 \psi_T \left(\frac{z}{h}, \mu \right) + \ln \frac{z}{h} \rightarrow C(\mu),\end{aligned}\tag{36.7}$$

where $\delta T = \bar{T}(h) - \bar{T}(z_0)$ is the temperature difference across the boundary layer. This yields the following laws of resistance and heat transfer, expressing the drag coefficient u_*/U , the angle of turn of the flow γ , and the heat-transfer coefficient $T_*/\delta T$ in terms of parameters z_0/h and μ :

$$\begin{aligned}\frac{u_*}{\kappa U} &= \left\{ B^2(\mu) + \left[A(\mu) + \ln \frac{z_0}{h} \right]^2 \right\}^{-1/2}; & \sin \gamma &= -\frac{u_*}{\kappa U} B(\mu); \\ \frac{T_*}{\delta T} &= \alpha_0 \left[C(\mu) - \ln \frac{z_0}{h} \right]^{-1}.\end{aligned}\tag{36.8}$$

The external Richardson number can then be written as

$$\text{Ri} = \frac{gh\delta T}{T_0 U^2} = \frac{\mu}{\alpha_0} \left[C(\mu) - \ln \frac{z_0}{h} \right] \left\{ B^2(\mu) + \left[A(\mu) + \ln \frac{z_0}{h} \right]^2 \right\}^{-1}. \quad (36.9)$$

With the aid of these formulas, we can use the external parameters of the boundary layer h , z_0 , U , δT and the buoyancy parameter g/T_0 (and thus Ri) to determine the internal parameters u_* , $q/c_p \rho_0$ (and thus T_* , L , and μ), and γ . To do this, we must, of course, know functions A , B , C . These have been measured empirically, and also calculated with the aid of semiempirical theories, one of which will be presented here in accordance with the work of Monin (1965) on the Ekman boundary layer. In this case in all the formulas (36.3)–(36.9) we have to replace h and μ by h_0 and μ_0 (here we set $h_0/z_0 = (\kappa u_*/U)\text{Ro}$, where $\text{Ro} = U/z_0 f$ is the Rossby number, introduced by C.—G. Rossby in 1932). In the semiempirical theory being presented here, we will neglect the molecular momentum flux in the first two Reynolds equations (33.3), so that they become

$$\frac{1}{\rho_0} \frac{\partial \tau_{xz}}{\partial z} = -f(\bar{v} + U \sin \gamma); \quad \frac{1}{\rho_0} \frac{\partial \tau_{yz}}{\partial z} = f(\bar{u} - U \cos \gamma), \quad (36.10)$$

while the fourth of Equations (33.3) takes the same form as in Section 34:

$$\frac{q}{c_p \rho_0} = \text{const.} \quad (36.11)$$

Next we derive equations analogous to (33.4), albeit somewhat simplified, for all ten single-point second moments of the fluctuations $\mathbf{u}'$ and T' . The main simplification here is the representation of the single-point second moments of the *derivatives* of the fluctuation fields (including moments (33.11)) in terms of b and l (this was suggested by Kolmogorov, 1942), as in (33.7), taking into account the tensor dimension of these moments:

$$\begin{aligned} \nu \overline{\nabla u'_i \cdot \nabla u'_j} &= [(c_1 - 3c_2)\lambda_i \lambda_j + c_2 \delta_{ij}] b^3 l^{-1}; \\ \chi \overline{(\nabla T')^2} &= c_3 b \sigma_T^2 l^{-1}; \quad (\nu + \chi) \overline{\nabla u'_i \cdot \nabla T'} = c_4 b^2 \sigma_T l^{-1} \lambda_i; \\ \frac{1}{\rho_0} p' \left(\frac{\partial u'_i}{\partial x_j} + \frac{\partial u'_j}{\partial x_i} \right) &= c_5 b l^{-1} \left(\overline{u'_i u'_j} - \frac{b^2}{3} \delta_{ij} \right) - \\ &\quad - c_6 b^3 l^{-1} \left(\lambda_i \lambda_j - \frac{1}{3} \delta_{ij} \right); \\ \frac{1}{\rho_0} p' \frac{\partial T'}{\partial x_i} &= -c_7 b l^{-1} \overline{u'_i T'} - c_8 b^2 \sigma_T l^{-1} \lambda_i, \end{aligned} \quad (36.12)$$

where $c_1, \dots, c_8$ are numerical constants (some of these formulas were first

offered by Rotta, 1951). As an additional simplification, we will neglect the 'fluxes of the second moments' produced by the turbulence (that is, the third moments of the fluctuations) and the molecular effects, that is, quantities such as those in parentheses on the right-hand sides of Equations (33.4) (or we will assume that these quantities are also described by expressions (36.12), so that they will be taken into account by a judicious choice of the numerical constants $c_1, \dots, c_8$). A third, nonessential, simplification involves neglecting the direct effect of the Coriolis force on the fluctuating motion. After all these simplifications, Equations (33.4) for the diagonal elements of the tensor of the Reynolds stresses take the form:

$$\begin{aligned}\frac{\tau_{xz}}{\rho_0} \frac{\partial \bar{u}}{\partial z} &= \left(c_2 - \frac{c_5 + c_6}{6} \right) b^2 l^{-1} + \frac{c_5}{2} b l^{-1} \sigma_u^2; \\ \frac{\tau_{yz}}{\rho_0} \frac{\partial \bar{v}}{\partial z} &= \left(c_2 - \frac{c_5 + c_6}{6} \right) b^3 l^{-1} + \frac{c_5}{2} \sigma_v^2; \\ \frac{g}{T_0} \frac{q}{c_p \rho_0} &= \left(c_1 - 2c_2 - \frac{c_5 - 2c_6}{2} \right) b^3 l^{-1} + \frac{c_5}{2} b l^{-1} \sigma_w^2.\end{aligned}\tag{36.13}$$

Equations $\partial \tau_{xz} / \partial t = \partial \tau_{yz} / \partial t = 0$ then become

$$\begin{aligned}c_5 b l^{-1} \frac{\tau_{xz}}{\rho_0} + \frac{g}{T_0} \frac{q_x}{c_p \rho_0} &= \sigma_w^2 \frac{\partial \bar{u}}{\partial z}; \\ c_5 b l^{-1} \frac{\tau_{yz}}{\rho_0} + \frac{g}{T_0} \frac{q_y}{c_p \rho_0} &= \sigma_w^2 \frac{\partial \bar{v}}{\partial z},\end{aligned}\tag{36.14}$$

while the equation $\partial \tau_{xy} / \partial t = 0$ becomes

$$c_5 b l^{-1} \frac{\tau_{xy}}{\rho_0} = - \left(\frac{\tau_{xz}}{\rho_0} \frac{\partial \bar{v}}{\partial z} + \frac{\tau_{yz}}{\rho_0} \frac{\partial \bar{u}}{\partial z} \right).\tag{36.15}$$

The equations for the horizontal components of the turbulent heat flux $\partial q_x / \partial t = \partial q_y / \partial t = 0$ have the form:

$$\begin{aligned}\frac{\tau_{xz}}{\rho_0} \frac{\partial \bar{T}}{\partial z} - c_7 b l^{-1} \frac{q_x}{c_p \rho_0} &= \frac{q}{c_p \rho_0} \frac{\partial \bar{u}}{\partial z}; \\ \frac{\tau_{yz}}{\rho_0} \frac{\partial \bar{T}}{\partial z} - c_7 b l^{-1} \frac{q_y}{c_p \rho_0} &= \frac{q}{c_p \rho_0} \frac{\partial \bar{v}}{\partial z},\end{aligned}\tag{36.16}$$

while the equation $\partial q / \partial t = 0$ becomes

$$\frac{g}{T_0} \sigma_T^2 = \sigma_w^2 \frac{\partial \bar{T}}{\partial z} + c_7 b l^{-1} \frac{q}{c_p \rho_0} + (c_4 + c_8) b^2 l^{-1} \sigma_T.\tag{36.17}$$

Finally, the equation $\partial\sigma_T^2/\partial t = 0$ becomes

$$c_3 b l^{-1} \sigma_T^2 = -\frac{q}{c_p \rho_0} \frac{\partial \bar{T}}{\partial z}. \quad (36.18)$$

The twelve equations (36.10), (36.13)–(36.18) contain 13 unknowns: $\bar{u}(z)$, $v(z)$, $\bar{T}(z)$, nine second moments of the fluctuations (the tenth second moment (36.11) is assumed to be given), and the scale of turbulence

$$l = \kappa z \Lambda(\eta, \mu_0), \quad \eta = \frac{z}{h_0},$$

for the determination of which we can invoke, for example, hypothesis (33.8).

From (36.18) it is evident that for $q = 0$ we have $\sigma_T = 0$; then from (36.17) we obtain the equilibrium temperature gradient (or, more precisely, the potential temperature) $\partial\bar{T}/\partial z = 0$. Thus, for the relation between q and $\partial\bar{T}/\partial z$ it is convenient to use the familiar formula $q/c_p \rho_0 = -K_T \partial\bar{T}/\partial z$ (the second formula of (34.5)). From (36.14)–(36.16) we then get

$$\begin{aligned} \tau_{xz} &= \rho_0 K \frac{\partial \bar{u}}{\partial z}; & \tau_{yz} &= \rho_0 K \frac{\partial \bar{v}}{\partial z}; & \tau_{xy} &= -\frac{2\rho_0 K l}{c_5 b} \frac{\partial \bar{u}}{\partial z} \frac{\partial \bar{v}}{\partial z}; \\ K &= \left(\frac{q}{c_p \rho_0} + \frac{c_7 b l^{-1}}{g/T_0} \sigma_w^2 \right) \left(\frac{\partial \bar{T}}{\partial z} + \frac{c_5 c_7 b^2 l^{-2}}{g/T_0} \right)^{-1}. \end{aligned} \quad (36.19)$$

This is a generalization of the first formula of (34.5). The relations $\text{Rf} = \alpha \text{Ri}$ and $\alpha = K_T/K$ are valid here. Moreover, with the aid of the first two formulas of (36.19) and the definition of Rf , we obtain from (36.13) the relations:

$$\begin{aligned} \frac{\sigma_u^2}{b^2} &= \frac{\sigma_{v0}^2}{b_0^2} + \frac{\sigma_{u0}^2 - \sigma_{v0}^2}{b_0^2} \frac{\varphi_u^2/\varphi^2}{1 - \text{Rf}}; & \frac{\sigma_v^2}{b^2} &= \frac{\sigma_{v0}^2}{b_0^2} + \frac{\sigma_{u0}^2 - \sigma_{v0}^2}{b_0^2} \frac{\varphi_v^2/\varphi^2}{1 - \text{Rf}}; \\ \frac{\sigma_w^2}{b^2} &= \frac{\sigma_{w0}^2}{b_0^2} - \frac{\sigma_{u0}^2 - \sigma_{v0}^2}{b_0^2} \frac{\text{Rf}}{1 - \text{Rf}}; & \frac{b^2}{b_0^3 u^3} &= \lambda \frac{1 - \text{Rf}}{\text{Rf}}; \\ \varphi_u &= \frac{l}{u} \frac{\partial \bar{u}}{\partial z}; & \varphi_v &= \frac{l}{u} \frac{\partial \bar{v}}{\partial z}; & \varphi^2 &= \varphi_u^2 + \varphi_v^2, & \lambda &= \frac{l}{\kappa L}, \end{aligned} \quad (36.20)$$

where σ_{u0}^2 , σ_{v0}^2 , σ_{w0}^2 are the values of quantities σ_u^2/u_*^2 , σ_v^2/u_*^2 , σ_w^2/u_*^2 in a neutrally stratified atmospheric surface layer (these are related by the formula $(\sigma_{u0}^2 - \sigma_{v0}^2)\sigma_{w0}^2 = 2$; measurements set their values at 5.8, 2.9, and 0.7), and $b^2 = \sigma_u^2 + \sigma_v^2 + \sigma_w^2$. Similarly, from (36.16) and (36.18) we obtain

$$\begin{aligned} \frac{(q_x, q_y)}{q} &= -q_{x0} \frac{u_* b_0}{b} \frac{\alpha_0}{\alpha} \frac{1 + \alpha}{1 + \alpha_0} (\varphi_u, \varphi_v); \\ \left(\frac{\sigma_T}{\kappa T_*} \right)^2 &= \sigma_{T_0}^2 \left(\frac{u_* b_0}{b} \right)^2 \frac{\alpha_0}{\alpha} \varphi^2 (1 - \text{Rf}). \end{aligned} \quad (36.21)$$

where q_{x0} , $\sigma_{T_0}^2$, and α_0 are numerical constants (the limits of the corresponding

quantities in a neutrally stratified air layer near the ground). From (36.14), the definition of Rf , and the fourth equation (36.20), we obtain

$$\frac{\lambda}{\varphi^3} = \frac{Rf}{1 - Rf} \left(\frac{\alpha_0}{\alpha} \psi \right)^{3/2}; \quad K = \frac{u_* l \varphi (\alpha_0/\alpha) \psi^{3/2}}{1 - Rf}; \quad (36.22)$$

$$\frac{\alpha_0}{\alpha} \psi = 1 - \left(1 + \frac{2}{\sigma_{w0}^4} + \frac{q_{x0}}{\sigma_{w0}^2} \frac{\alpha_0}{\alpha} \frac{1 + \alpha}{1 + \alpha_0} \right) Rf.$$

Next we eliminate σ_T^2 from (36.17)–(36.18) to arrive at a quadratic equation for $\psi = \psi(Rf)$:

$$(1 - A^2)(1 - Rf)\psi = [\psi - A(1 - B \cdot Rf)]^2; \quad (36.23)$$

$$A = \frac{q_{x0} \sigma_{w0}^2}{1 + \alpha_0}; \quad B = 1 + \frac{2}{\sigma_{w0}^4} + \frac{\alpha_0 \sigma_{T0}^2}{\sigma_{w0}^2},$$

where we must choose a solution for ψ which goes to unity as $Rf \rightarrow 0$. This solution corresponds to Rf_{cr} . Using it, we can find $\alpha = \alpha(Rf)$ from the third formula of (36.22). Turning now to the Reynolds equations (36.10) we introduce into them the dimensionless complex function

$$\tau(\eta) = \frac{\tau_{xz} + i\tau_{yz}}{\rho_0 u_*^2},$$

and then after applying the operator $K \partial/\partial z$ to these equations we obtain them in the elegant form

$$Rf |\tau|^2 \tau'' = i\mu_0 \tau. \quad (36.24)$$

The solution has to satisfy the boundary conditions $\tau \rightarrow 1$ at the air layer near the ground and $\tau' \rightarrow 0$ at the upper limit of the boundary layer (for stable neutral stratification, evidently, at a certain height). Equation (36.24) contains two unknown functions τ and Rf . Additional relations between $|\tau|$ and Rf can be obtained by equating the values of l from the first equation of (36.22) and from the generalized Kármán formula (33.8):

$$\frac{|\tau|^{3/2} Rf (1 - Rf)^{1/2}}{((\alpha_0/\alpha)\psi)^{3/4}} = -\mu_0 \left(\frac{\partial \ln \Psi}{\partial \eta} \right)^{-1}; \quad \Psi = \frac{(1 - Rf)^{1/2}}{|\tau| Rf}. \quad (36.25)$$

Setting $\tau = |\mu_0| \tau_1$ and $\eta = |\mu_0|^{1/2} \eta_1$, for functions $\tau_1(\eta_1)$ and $Rf(\eta_1)$ we arrive at the universal equations (36.24) and (36.25) with values $\mu_0 = \pm 1$. In this sense, all stratified Ekman boundary layers with the same sign of the stratification possess dynamic similarity.

37. GEOSTROPHIC TURBULENCE

Large-scale turbulence (in the Earth's atmosphere and ocean, and in the atmos-

pheres of other planets and the Sun) with horizontal scales L that are large compared to the effective thickness H of the layer encompassing it is called *quasi-two-dimensional* turbulence. In spite of the fact that the fluid-dynamical fields characteristic of such turbulence vary considerably over the height, that a major role in its dynamics is played by vertical motions (thereby relating the fields of horizontal motions at different heights), and that even for adiabatic processes vorticity is not conserved in the fluid particles (while, on the contrary, stretching of the vortex lines occurs), still it is described by the *two-dimensional* equation of horizontal transfer of potential vorticity (12.6).

In order to better explain the provenance of this equation (and also to convert it from coordinates x, y, p to Cartesian coordinates x, y, z , which will be more convenient here), let us derive it in a manner somewhat different from that in Section 12. We consider adiabatic processes, for which in the fluid particles the entropy η and the potential vorticity Ω_* are conserved, that is,

$$\frac{d_h \eta}{dt} + w \frac{\partial \eta}{\partial z} = 0; \quad \frac{d_h \Omega_*}{dt} + w \frac{\partial \Omega_*}{\partial z} = 0, \quad (37.1)$$

where d_h/dt is the symbol for the individual derivative with respect to the horizontal motion. After eliminating w , we obtain

$$\frac{d_h \Omega_*}{dt} - \frac{\partial \Omega_*/\partial z}{\partial \eta/\partial z} \frac{d_h \eta}{dt} = 0. \quad (37.2)$$

Large-scale motions can, firstly, be considered with good accuracy to be nondivergent, so that the conserved quantity will be $\Omega_* = \Omega_a \cdot \nabla \eta$, where $\Omega_a = \text{curl } \mathbf{u} + 2\boldsymbol{\omega}$ is the vorticity of the absolute velocity of motion. Secondly, for such motions the field $\eta(\mathbf{x})$ is quasihorizontal, meaning that vector $\nabla \eta$ is quasivertical, so that $\Omega_* \approx \Omega_{az} \partial \eta / \partial z$. Then (37.2) reduces to

$$\begin{aligned} \frac{d_h \ln \Omega_{az}}{dt} + \left(\frac{\partial \eta}{\partial z} \right)^{-1} \frac{d_h}{dt} \frac{\partial \eta'}{\partial z} - \\ - \left(\frac{\partial \eta}{\partial z} \right)^{-2} \left(\frac{\partial^2 \eta}{\partial z^2} + \frac{\partial n}{\partial z} \frac{\partial \ln \Omega_{az}}{\partial z} \right) \frac{d_h \eta'}{dt} = 0, \end{aligned} \quad (37.2')$$

where $\eta' = \eta - \eta_0$ is the part of field η which depends on $\mathbf{x}_h$ and t . Since it is small compared to η_0 , it can be neglected in the multipliers in front of derivatives d_h/dt (in addition, we neglect the term $\partial \ln \Omega_{az} / \partial z$ in comparison with $\partial(\ln \partial \eta_0 / \partial z) / \partial z$). Then from (37.2') we arrive at the following approximate conservation law:

$$\frac{d_h}{dt} \left(\ln \Omega_{az} + \frac{\partial}{\partial z} \frac{\eta'}{\partial \eta_0 / \partial z} \right) = 0 \quad (37.3)$$

(Gavrilin and Monin, 1969). Taking into account that large-scale processes are

quasihydrostatic, we then obtain

$$\begin{aligned}\frac{\partial \eta_0}{\partial z} &= \left(\frac{\partial \eta}{\partial p} \right)_0 \left(\frac{\partial p_0}{\partial z} - c_0^2 \frac{\partial \rho_0}{\partial z} \right) = \left(\frac{\partial \eta}{\partial p} \right)_0 \frac{\rho_0 c_0^2 N^2}{g}; \\ \eta' &= \left(\frac{\partial \eta}{\partial p} \right)_0 (p' - c_0^2 \rho') \approx - \left(\frac{\partial \eta}{\partial p} \right)_0 c_0^2 \rho' \approx \left(\frac{\partial \eta}{\partial p} \right)_0 \frac{c_0^2}{g} \frac{\partial p'}{\partial z},\end{aligned}\quad (37.4)$$

so that in (37.3) we can set $(\partial \eta_0 / \partial z)^{-1} \eta' \approx (\rho_0 N^2)^{-1} \partial p' / \partial z$. Finally, we invoke the quasigeostrophicity of the large-scale motions to obtain $|\text{curl}_z \mathbf{u}| \ll 2\omega_z = f$ and $p' \approx \rho_0 f \psi$, where ψ is the stream function of the horizontal flow. Then Equation (37.3) reduces to the form $d_h \omega_* / dt = 0$, where

$$\omega_* = \Omega_{az} + \frac{\partial}{\partial z} \frac{H^2}{L_R^2} \frac{\partial \psi}{\partial z}$$

is a quantity proportional to the potential vorticity, taken in the quasigeostrophic approximation (H is the thickness of the layer, and $L_R = H N f^{-1}$ is the Rossby deformation radius). Since we know that $\Omega_{az} = \Delta \psi + f$ (where $\Delta = \Delta_h$ is the horizontal Laplacian), this equation can also be rewritten as

$$\frac{\partial \mathcal{L} \psi}{\partial t} + \frac{\partial(\psi, \mathcal{L} \psi)}{\partial(x, y)} + \beta \frac{\partial \psi}{\partial x} = F; \quad \mathcal{L} = \Delta + \frac{\partial}{\partial z} \frac{H^2}{L_R^2} \frac{\partial}{\partial z}, \quad (37.5)$$

where $\beta = \partial f / \partial y$; for generality, we have here written the right-hand side as F , characterizing the effect of nonadiabatic processes. This equation also describes the evolution of individual realizations of what is known as *geostrophic turbulence*. The term $\beta \partial \psi / \partial x$ on its left-hand side describes *Rossby waves* (points 4⁺ and 5⁺ in Figure 10.3), which will be the subject of the next chapter. The ratio δ of the second term to the third, that is, the ‘nonlinearity parameter’ of these waves, is equal, as usual, to the ratio of the typical velocity of fluid motion U to the phase velocity of the waves, here equal to $c(L) \sim \beta L^2 / 2$ (where L is the wavelength), that is, $\delta = 2U / \beta L^2$. For $\delta < 1$ waves dominate in the flow, and for $\delta > 1$ eddies dominate, the boundary value $\delta = 1$ corresponding to the horizontal scale

$$L_\beta = \left(\frac{2U}{\beta} \right)^{1/2}. \quad (37.6)$$

On the Earth for $\beta \sim 2 \times 10^{-8} \text{ km}^{-1} \text{ s}^{-1}$ in the atmosphere, where $U \sim 10 \text{ m/s}$, we get $L_\beta \sim 1000 \text{ km}$, while in the ocean, where $U \sim 10 \text{ cm/s}$, the scale $L_\beta \sim 100 \text{ km}$. For $L \ll L_\beta$ the third term on the left-hand side of Equation (37.5) can be neglected.

Moreover, for $L > L_R$ a significant role in operator $\mathcal{L}$ is played by the second term, so that different fluid layers will interact strongly along the vertical, the vertical inhomogeneities become blurred, and the entire fluid evolves as a single layer (‘barotropic’ structure). For $L \ll L_R$, on the other hand, we have $\mathcal{L} \approx \Delta$, so that vertical interactions between different fluid layers play a minor role, and these layers evolve more or less independently of one another (vertical inhomogeneities do not become blurred: ‘baroclinic’ structure). In the Earth’s atmosphere

for $f \sim 10^{-4} \text{ s}^{-1}$, $H \sim 10 \text{ km}$, and $N \sim 2 \times 10^{-3} \text{ s}^{-1}$ (in a layer of the jet stream of the upper troposphere and lower stratosphere) we get $L_R \sim 2000 \text{ km}$; in the ocean for $H \sim 5 \text{ km}$ and $N \sim 10^{-3} \text{ s}^{-1}$ we get $L_R \sim 50 \text{ km}$.

Consequently, for $L \ll (L_\beta, L_R)$ from (37.5) we obtain an equation for the stream function ψ of the two-dimensional flow of an incompressible fluid:

$$\frac{\partial \Delta \psi}{\partial t} + \frac{\partial(\psi, \Delta \psi)}{\partial(x, y)} = \nu \Delta \Delta \psi, \quad (37.7)$$

where as the right-hand side F we take the usual expression describing the effect of molecular viscosity. Therefore, in the region of smallest scales the geostrophic turbulence decays into ordinary two-dimensional turbulence. It differs considerably from three-dimensional turbulence, first of all, in that for an ideal fluid in the fluid particles not only is kinetic energy conserved, as in three-dimensional flows, but vorticity as well (and thus the vorticity squared, which divided by two is the *enstrophy*); the existence of this additional conservation law makes it possible to demonstrate the existence and uniqueness of the solutions of the equations of two-dimensional fluid dynamics. In a viscous fluid the mean kinetic energy $E = \frac{1}{2} |\nabla \psi|^2$ and the enstrophy $\Omega = \frac{1}{2} (\Delta \psi)^2$ decay according to the laws

$$\frac{\partial E}{\partial t} = -2\nu\Omega (= -\varepsilon); \quad \frac{\partial \Omega}{\partial t} = -\nu |\nabla \Delta \psi|^2 (= -\varepsilon_\omega). \quad (37.8)$$

Thus it is evident that for any finite initial values of E and Ω the value of Ω can only decrease with time, and for $\nu \rightarrow 0$ we have $\varepsilon \rightarrow 0$, so that for large Re the energy E will be approximately constant. Moreover, the Kolmogorov cascade transfer of any substantial portion of E from large scales to small scales (in the three-dimensional case leading to turbulence spectrum (34.18)) will be impossible. Here the energy spectrum in the energy-bearing interval of the scales has the form

$$E(k, t) = E^{3/2} t f(E^{1/2} k t), \quad (37.9)$$

from which it follows that $\Omega = \frac{1}{2} A t^{-2}$ and $\varepsilon_\omega = A t^{-3}$, where A is a numerical constant (Batchelor, 1969). At the same time, a nonzero limit of the quantity ε_ω as $\nu \rightarrow 0$, as well as a cascade process of transfer of enstrophy from large scales to small scales, are not excluded.

Let us next introduce the mean wave numbers for the energy spectrum $E(k, t)$ and the enstrophy spectrum $k^2 E(k, t)$:

$$k_1(t) = E^{-1} \int_0^\infty k E(k, t) dk; \quad k_2(t) = \Omega^{-1} \int_0^\infty k^3 E(k, t) dk. \quad (37.10)$$

If we set

$$kE(k) = [kE^{1/2}(k)] \cdot [E^{1/2}(k)] \quad \text{and} \quad k^2 E(k) = [k^{1/2} E^{1/2}(k)] \cdot [k^{3/2} E^{1/2}(k)]$$

and apply to the integrals of these functions the Cauchy—Bunjakowski—Schwarz inequality, then we can show that

$$k_1(t)k_2(t) \geq k_*^2 = \frac{\Omega}{E}; \quad k_1(t) \leq k_*; \quad k_2(t) \geq k_*. \quad (37.11)$$

These inequalities indicate that a transfer of energy toward small k (decrease in k_1) must be accompanied by a transfer of enstrophy toward large k (increase in k_2) (the 'eddy viscosity' has to be considered negative in this case!) and also that in cascade processes energy cannot be stored in small-scale motions, while enstrophy cannot be stored in large-scale motions. If k_1 is calculated according to the self-similar spectrum (37.9) obtained for $E \approx \text{const}$, then the rate of increase with time of the turbulence scale k_1^{-1} turns out to be $\partial k_1^{-1} / \partial t = aU$, where $U = (2E)^{1/2}$ is the rms velocity of fluid motion, while $a = (\int xf(x) dx)^{-1} \int f(x) dx$ is a positive constant.

For spectral transfer of enstrophy in the inertial interval of the wave numbers k (or, more precisely, for $L_0^{-1} \ll k \ll \min(\varepsilon^{1/4} \nu^{-3/4}, \varepsilon_\omega^{1/6} \nu^{-1/2})$) the energy spectrum will be determined, in general, by both parameters ε and ε_ω . These can be used to construct a length scale $L_\omega = (\varepsilon/\varepsilon_\omega)^{1/2}$ and, in accordance with dimensionality considerations, the energy spectrum will be determined by the right-hand formula of (34.18) with a coefficient $C_1 = C_1(kL_\omega)$, depending on the dimensionless wave number kL_ω . According to (37.11), at the large-scale end of the spectrum (that is, for low k) only parameter ε is important, so that there we have $C_1(kL_\omega) \approx C_1(0)$, and spectrum (34.18) satisfies the 'five-thirds law', as in three-dimensional turbulence, except with an opposite direction of the spectral energy transfer: from small scales to large. At the small-scale end of the inertial interval, on the other hand, only parameter ε_ω is important, and parameter ε can be dropped from the formula for $E(k)$. Here for large kL_ω there should be an asymptote $C_1(kL_\omega) \sim (kL_\omega)^{-4/3}$, which yields an 'inverse cube law'

$$E(k) = C_{1\infty} \varepsilon^{2/3} k^{-3} \quad (37.12)$$

(Kraichnan, 1967). This law was confirmed by numerical experiments on two-dimensional turbulence (with a coefficient $C_{1\infty} \approx 1.5$) and by empirical data on the spectra of large-scale meteorological fields in the Earth's atmosphere (starting with the studies of Ogura, 1952, who first drew attention to this law). The predictions implied by (37.11) concerning the relation between the spectral transfer of kinetic energy and the transfer of enstrophy can be illustrated by a calculation of the rates $\Pi(k)$ and $\Pi_\omega(k)$ of this transfer (from wave numbers less than k to wave numbers greater than k) in the so-called diffusion approximation (Leith, 1968):

$$\Pi(k) = -ak^{-1} \frac{\partial}{\partial k} k^{9/2} E^{3/2}(k); \quad \Pi_\omega(k) = -ak^3 \frac{\partial}{\partial k} k^{5/2} E^{3/2}(k), \quad (37.13)$$

where a is a positive numerical constant. For $E(k) \sim k^{-5/3}$ we have $\Pi = \text{const} < 0$ and $\Pi_\omega = 0$, while for $E(k) \sim k^{-3}$ it turns out that $\Pi_\omega = \text{const} > 0$ and $\Pi = 0$.

Similarly to what was said above for the concentration ϑ of a passive admixture, satisfying in two-dimensional fluid dynamics the equation

$$\frac{\partial \vartheta}{\partial t} + \frac{\partial(\psi, \vartheta)}{\partial(x, y)} = \chi \Delta \vartheta, \quad (37.14)$$

in the inertial-convective interval of wave numbers (more precisely, for $L_0^{-1} \ll k \ll \min(\varepsilon^{1/4} \nu^{-3/4}, \varepsilon^{1/4} \chi^{-3/4}, \varepsilon^{1/6} \nu^{-1/2})$, the spectrum will be determined by formula (34.19) with a coefficient $C_\delta = C_\delta(kL_\omega)$. For low kL_ω here we should get $C_\delta(kL_\omega) \approx C_\delta(0)$, that is, spectrum (34.19) satisfies the 'five-thirds law'. For large kL_ω , on the other hand, parameter ε has to be dropped from the formula for $E_\delta(k)$, which corresponds to the asymptote $C_\delta(kL_\omega) \sim (kL_\omega)^{2/3}$, yielding a 'minus-first-power law'

$$E_\delta(k) = C_{\delta\infty} \varepsilon_\delta \varepsilon_\omega^{-1/3} k^{-1} \quad (37.15)$$

(Gavrilin *et al.*, 1972; Saunders, 1972). Note, however, that for laws (37.12) and (37.15) the total enstrophy $\int_{k_0}^{k_m} k^2 E(k) dk$ and the measure of inhomogeneity of the field of the passive admixture $\int_{k_0}^k E_\delta(k) dk$ diverge logarithmically, for both $k_0 \rightarrow 0$ and $k_m \rightarrow \infty$. In order to avoid these divergences, we have to introduce in spectra (37.12) and (37.15) corrections in the form of logarithmic multipliers. For this purpose we will set

$$\begin{aligned} \Pi(k) &\sim kE(k)\sigma(k); & \Pi_\omega(k) &\sim k^3 E(k)\sigma(k); \\ \Pi_\delta(k) &\sim kE_\delta(k)\sigma(k); & \sigma(k) &\sim \left[\int_{k_0}^k k'^2 E(k') dk' \right]^{1/2}. \end{aligned} \quad (37.16)$$

From this it is evident that satisfaction of the condition $\Pi(k) = \text{const}$ can yield only a 'five-thirds law', and in this case the main contribution to the integral in the formula for $\sigma(k)$ is made by wave numbers $k' \sim k$, demonstrating the localization of the spectral energy transfer. For spectral laws (37.12) and (37.15) we would obtain $\sigma(k) \sim (\ln(k/k_0))^{1/2}$, and the values of $\Pi_\omega(k)$ and $\Pi_\delta(k)$ would increase with k , whereas to ensure the requirements $\Pi_\omega(k) = \text{const}$ and $\Pi_\delta = \text{const}$ we have to set $\sigma(k) \sim \varepsilon_\omega^{1/3} (\ln(k/k_0))^{1/3}$, and instead of formulas (37.12) and (37.15) we get

$$E(k) = C_{1\infty} \varepsilon_\omega^{2/3} k^{-3} \left(\ln \frac{k}{k_0} \right)^{-1/3}; \quad E_\delta(k) = C_{\delta\infty} \varepsilon_\delta \varepsilon_\omega^{-1/3} k^{-1} \left(\ln \frac{k}{k_0} \right)^{-1/3}. \quad (37.17)$$

Now the main contribution to the integral in the formula for $\sigma(k)$ is made by wave numbers k' from the interval $k \gg k' \gg k_0$, so that the spectral transfers of enstrophy and the passive admixture turn out to be nonlocal processes. The first of formulas (37.17) was obtained by Kraichnan and the second by Mirabel' and Monin (1983), where spectra (37.17) were calculated with the aid of the fluid-dynamical equations, the third moments of the velocity field being expressed in terms of the second moments in the Markov approximation (which gave $C_{\theta\infty} = C_{1\infty} = (4)^{2/3}$). Here a modification of formula (37.17) for $E_\delta(k)$ in the region of small scales (large k) was also found, in which molecular diffusion of the admixture already has an effect, and the fluid-dynamical equation for $E_\delta(k)$ becomes

$$\frac{\partial}{\partial k} [kE_\delta(k)\sigma(k)] + 2\chi k^2 E_\delta(k) = 0. \quad (37.18)$$

For large k , the factor $\sigma(k)$ can be taken out from under the derivative sign, and then the solution of this equation with the aid of the asymptotic formula

$$\int_{k_0}^k k \left(\ln \frac{k}{k_0} \right)^{-1/3} dk \sim \frac{k^2}{2} \left(\ln \frac{k}{k_0} \right)^{-1/3} \quad (k \rightarrow \infty)$$

becomes

$$E_\delta(k) = C_\delta \varepsilon_\delta \varepsilon_\omega^{-3} k^{-1} \exp[-\chi k^2 \sigma^{-1}(k)], \quad (37.19)$$

where constant C_δ is determined from the normalization condition $2\chi \int_0^\infty k^2 E_\delta(k) dk = \varepsilon_\delta$. Note that, if $\sigma(k)$ were determined without the local logarithmic factor, but as a constant $\sigma_0 \sim \varepsilon_\omega^{1/3}$ (the mean rate of deformation of a fluid particle), then formula (37.19) would become the formula of Batchelor (1959) for the admixture spectrum in a field of three-dimensional turbulence for a high Prandtl number ν/χ . Therefore, here, too, the two-dimensionality shows up as a nonlocal logarithmic correction. The same is true of the eddy diffusion $K_\delta(k)$ produced by the small-scale components of two-dimensional turbulence (with wave numbers higher than the value of k from the inertial-convective interval), which fluid-dynamical calculations show is equal to

$$K(k) \sim \int_k^\infty E(k) \sigma^{-1}(k) dk = A \varepsilon_\omega^{1/3} k^{-1} \left(\ln \frac{k}{k_0} \right)^{-2/3} \left[1 + O \left(\ln \frac{k}{k_0} \right)^{-1} \right], \quad (37.20)$$

where A is the calculated numerical constant (Mirabel' and Monin, 1983). For diffusion from a symmetrical source we obtain the equation

$$\frac{\partial \delta}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} r K_0(r) \frac{\partial \delta}{\partial r}; \quad K_0(r) \approx A \varepsilon_\omega^{1/3} r^2 \left(\ln \frac{r_0}{r} \right)^{-2/3}, \quad (37.21)$$

in which because of the nonlocal logarithmic factor it follows that $r^{-2} K_0(r) \rightarrow 0$ ($r \rightarrow \infty$). This means that there must exist a solution $\delta(r, t)$ corresponding to an instantaneous point source of admixture (whereas in the absence of a logarithmic correction we would have $K_0(r) = A \varepsilon_\omega^{1/3} r^2$, and such a solution would not exist, since admixture would be carried away to infinity too rapidly; instead of it, we have here a solution $\delta \sim r^{-1} \tau^{-1/2} \exp[-\tau - (\ln r)^2/4\tau]$, $\tau = A \varepsilon_\omega^{1/3} t$, with the initial admixture distribution in a single ring). Finally, taking into account that the mean-square distance between two diffusing particles l^2 is found from the equation $\frac{1}{2} \partial l^2 / \partial t = K_0(l)$, with the aid of (37.21) we have

$$l^2(t) = l_0^2 \exp \left[- \left(\frac{t_0 - t}{T} \right)^{3/5} \right]; \quad T \sim \varepsilon_\omega^{-1/3}, \quad (37.22)$$

where t_0 is the time when the value of l becomes equal to the outer scale of turbulence l_0 (large-scale boundary of the inertial interval).

For $L \ll L_\beta$ all the above-mentioned regularities of two-dimensional turbulence remain in force, regardless of the value of L_R , the only stipulation being that if $L_R < L_\beta$ (as in the ocean) then for $L_R < L < L_\beta$ the enstrophy $\frac{1}{2} (\Delta\psi)^2$ must everywhere be replaced by the potential enstrophy $\frac{1}{2} (\mathcal{L}\psi)^2$.

From (37.11) and (37.9) it follows that the scale of two-dimensional turbulence k_1^{-1} increases with time according to the law $\partial k_1^{-1} / \partial t = aU$ (and that, correspondingly, the typical frequency $\omega \sim Uk$ decreases), during a time of the order of L_β / aU . Thus a scale L_β is reached, for which the eddies moving along with the fluid are transformed into Rossby waves traveling along the fluid (in view of the dispersion relation of these waves $\omega = -\beta k_x k^{-2}$, where k_x is the zonal wave number, a tendency appears for inhomogeneities of the field to shift toward the west). After this, the mean scale of the inhomogeneities (now already waves) will continue to increase, while their typical frequency ω decreases, but much more slowly, since the wave interactions call for both the superposition of three waves in space and a resonance relation between their wave numbers and frequencies. The frequency decrease as a result of these interactions can be explained by the fact that in a resonance triad the wave with the highest frequency turns out to be unstable *vis-à-vis* the transfer of energy by the other two waves.

In addition, a slowing of the increase with time of the mean scale of the waves is caused by the tendency of the wave interactions to impart energy primarily not to other wavelengths, but rather to other directions in space, more precisely, to the zonal direction: the drop in the frequency $|\omega| = |\beta k_x k^{-2}|$ occurs because the waves become anisotropic, there being a preferential growth of their zonal scales compared to their meridional scales (the ratio $|k_x / k_y|$ decreases), that is, zonal currents develop (at the limit, stationary zonal currents appear in narrow latitude bands with directions varying toward the west or toward the east as we pass from pole to pole).

When the scale of turbulence (or of the waves) reaches a value L_R , there should appear a tendency for horizontal inhomogeneities to 'become barotropic', that is, for differences between motions at different levels to become smaller. The effects of an increase in the scales of inhomogeneities, the slowing of this increase, the appearance of a tendency toward a westward shift and anisotropy when the scale value L_β is reached, as well as a beginning of barotropy when L_R is reached, were all predicted by Rhines (1975) and observed by him in numerical experiments; on the barotropic model he obtained $a \approx 3 \times 10^{-2}$ for turbulence and $a \approx 6 \times 10^{-3}$ for waves, while on the two-layer model he demonstrated that the inhomogeneities very rapidly become completely barotropic.

In the Earth's atmosphere $L_\beta < L_R$; eddies are created mostly due to the baroclinic instability of the global currents (and to a lesser degree due to topographic factors) and they evolve according to the following pattern: baroclinic eddies $\rightarrow$ baroclinic waves $\rightarrow$ barotropic waves. In the World Ocean $L_R < L_\beta$; marine eddies originate because of the baroclinic instability of large-scale currents, to a fairly large degree because of topographic factors (the relative height of the bottom

irregularities h/H is comparable to the magnitude of the β effect L/a , the critical value $h = HL/a$ being only some tenths of the value in the atmosphere), also because of the reflection of waves from the shore (at the western shores the enstrophy increases, while it decreases at the eastern shores; since small-scale eddies move more slowly, both enstrophy and energy tend to accumulate at the western shores, and they become scattered at the eastern shores), and finally as a result of atmospheric effects. The evolution of the eddies in this case is primarily as follows: baroclinic eddies $\rightarrow$ barotropic eddies $\rightarrow$ barotropic waves; more precisely, for large L/L_R we obtain $\delta \sim (L_\beta/L_R)^2 > 1$, so that linear barotropic waves are hardly attained.

The above-mentioned limit of stationary zonal currents is apparently reached in the atmospheres of Jupiter and the other large planets, in which the typical times of energy conversions $E(\partial E/\partial t)^{-1}$ are much longer (of the order of several years) than $\omega^{-1} = (Uk)^{-1}$. The numerical experiments of Williams (1978, 1979) on geostrophic turbulence in the atmosphere of Jupiter at the indicated limit revealed zonal jets with velocities of the order of 50 m/s, alternating with bands of zonally elongated eddies having scales of the order of 10^4 to 2×10^4 km. Isolated eddies in the Jovian atmosphere (such as the Great Red Spot) can be explained by baroclinic instability or by theory of Rossby solitons (see Chapter 7).

Rossby Waves

38. LINEAR THEORY

Variations along the meridian of the Coriolis parameter $f = 2\omega \cos \theta$ produce a restoring force, for motions with an easterly component of the velocity. For instance, in the Northern Hemisphere for motion toward the northeast, f increases, and the augmented Coriolis force turns the flow toward the right, until it becomes southeasterly; then f decreases, and this weakening of the Coriolis force produces a leftward turning of the flow, etc. The waves generated by this restoring force were first discovered by Margules (1893) and Hough (1898), but they were generally known as ‘Rossby waves’, after C.-G. Rossby (1939), who derived the ‘trough formula’, giving the velocity of zonal displacement of the troughs of monochromatic barotropic waves of this type. Later, Blinova (1943) formulated a general solution of the linear equation for barotropic waves on a sphere for arbitrary initial data. [In Soviet works Rossby waves are called ‘Rossby-Blinova waves’ (Translator).]

In Section 10 linear Rossby waves were represented as families of solutions of the Laplace tidal equation on a sphere, the frequencies σ of which correspond to the intersections of the ‘horizontal’ eigencurves $\sigma = \sigma(\varepsilon)$, $\varepsilon = 4r_0^2\omega^2/g h_0$ of the second type (with finite values of $\sigma \sim 2\omega m/n(n+1)$ for low $|\varepsilon|$ and with asymptotes $\sigma \rightarrow 0$ as $\varepsilon \rightarrow \infty$ and $\sigma \rightarrow 2\omega$ as $\varepsilon \rightarrow -\infty$) with the zero ‘vertical’ curve (with the asymptotes $\sigma \sim \pm(g/h_0)^{1/2}$ as $\varepsilon \rightarrow \infty$; these frequencies correspond to barotropic waves) and with ‘positive’ curves (with the asymptotes $\sigma \sim \pm N_{\max}$ as $\varepsilon \rightarrow \infty$, which corresponds to baroclinic waves), see Figure 10.3.

The properties of linear Rossby waves are explained most simply using the ‘ β -plane’ approximation, that is, in local Cartesian coordinates (x, y, z) , where the origin is at some selected point on the Earth’s surface, the x axis points eastward, the y axis points northward (with $f = f_0 + \beta y$), and the z axis points upward. Using these coordinates, we can describe free Rossby waves of low amplitude with the aid of the linearized equation (37.5):

$$\frac{\partial \mathcal{L}\psi}{\partial t} + \beta \frac{\partial \psi}{\partial x} = 0; \quad \mathcal{L} = \Delta + \frac{\partial}{\partial z} \frac{H^2}{L_R^2} \frac{\partial}{\partial z}; \quad (38.1)$$

$$\begin{aligned}
 z = 0: \quad & \frac{\partial}{\partial t} \left(\frac{1}{N^2} \frac{\partial \psi}{\partial z} + \frac{1}{g} \psi \right) = 0; \\
 z = -h: \quad & \frac{f_0}{N^2} \frac{\partial^2 \psi}{\partial t \partial z} = \frac{\partial(\psi, h)}{\partial(x, y)}.
 \end{aligned} \tag{38.2}$$

Boundary conditions (38.2) at a free surface and at the bottom of the ocean are the usual conditions (26.2), simplified just as for the derivation of (37.1)–(37.5) and linearized (where the vertical velocity has the form $w = -fN^{-2}\partial^2\psi/\partial t \partial z$, and the level of the free surface $z = \zeta(\mathbf{x}_h, t)$ is found from the condition $\partial\zeta/\partial t = w$ for $z = 0$). Let us examine in this section the case of a plane bottom $h = \text{const}$, at which there exist solutions with separated variables $\psi = \psi_m(\mathbf{x}_h, t)\Psi_m(z)$. Here factors ψ_m have to satisfy the ‘horizontal’ equations

$$\frac{\partial}{\partial t} \left(\Delta - \frac{f_0^2}{gh_m} \right) \psi_m + \beta \frac{\partial \psi_m}{\partial x} = 0, \tag{38.3}$$

where h_m and $\Psi_m(z)$ are the eigenvalues and eigenfunctions of the ‘vertical’ equation

$$\begin{aligned}
 \frac{\partial}{\partial z} \frac{1}{N^2} \frac{\partial \Psi_m}{\partial z} + \frac{1}{gh_m} \Psi_m &= 0; \\
 \left(\frac{1}{N^2} \frac{\partial \Psi_m}{\partial z} + \frac{1}{g} \Psi_m \right)_{z=0} &= \left(\frac{\partial \Psi_m}{\partial z} \right)_{z=-h} = 0,
 \end{aligned} \tag{38.4}$$

obtained from (10.9) in the quasistatic approximation. The barotropic mode $\Psi_0(z)$, corresponding to $h_0 = h$, depends slightly on z and does not change sign in the interval $0 > z > -h$. Baroclinic modes $\Psi_m(z)$, $m = 1, 2, \dots$; $h > h_1 > h_2, \dots$, possess in the indicated interval of m zero values (in particular, for $N(z) = \text{const}$ we have $h_m = g^{-1}(Nh/m\pi)^2$; these so-called *equivalent depths* are the thicknesses of those homogeneous fluid layers within which the barotropic waves coincide with the corresponding baroclinic waves in the initial stratified layer). For the corresponding wave solutions $\psi_m \sim \exp[i(\mathbf{k} \cdot \mathbf{x}_h - \sigma_m t)]$, formula (38.3) yields the frequencies

$$\sigma_m = -\beta k_x \left(k^2 + \frac{f_0^2}{gh_m} \right)^{-1}. \tag{38.5}$$

The minus sign here means that the linear Rossby waves move westward (for barotropic waves on deep water, superimposed onto a westerly transport with a velocity U , the phase velocity $c = U - \beta k^{-2}$, which is a form of Rossby’s ‘trough formula’). The dispersion curves $\sigma = \sigma(k_x)$ for two different k_y and for $m = 0, 1, 2, 3$ are shown in Figure 38.1. Inspection of the curves shows that, for $k_x = k_{\max} = f(gh_m)^{-1/2}$ and $k_y = 0$, frequencies σ go through maxima $\sigma_m = \beta(gh_m)^{1/2}f_0^{-1}$, so that $k_{\max}/k_z \sim f_0/N \ll 1$, that is, for $k \sim k_{\max}$ Rossby waves are *long waves*.

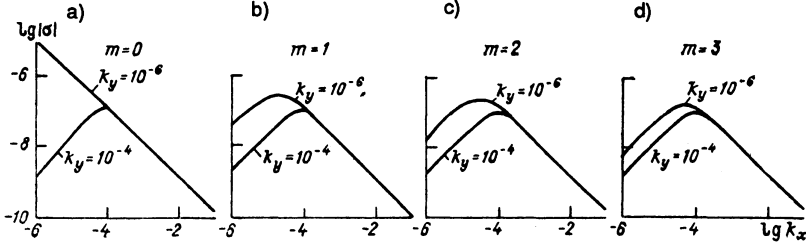


Fig. 38.1. Dispersion curves for Rossby waves. σ is measured in s^{-1} , k_x and k_y in m^{-1} , $\omega = 7.29 \times 10^{-5} \text{ s}^{-1}$. The example was plotted for latitude $\varphi = \pi/9$ at $h = 4 \text{ km}$ and $N = 2 \times 10^{-3} \text{ s}^{-1}$.

Recalling the concepts of ray theory outlined in Section 19, let us consider a wave packet consisting of plane waves with wave vectors $\mathbf{k}$ in the region $\mathbf{k}_0 \pm \delta \mathbf{k}$ with very small $\delta k/k_0$. Then $\sigma_m(\mathbf{k}) \approx \sigma_m(\mathbf{k}_0) + (\mathbf{k} - \mathbf{k}_0) \nabla \sigma_m(\mathbf{k}_0)$, and the solution of the Cauchy problem, obtained with the aid of a Fourier transform for Equation (38.3) in unbounded space, is

$$\begin{aligned} \psi_m(\mathbf{x}, t) &= \frac{1}{2\pi} \int \tilde{\psi}_m(\mathbf{k}) \exp\{-i[\mathbf{k} \cdot \mathbf{x} - \sigma_m(\mathbf{k})t]\} d\mathbf{k} \approx \\ &\approx \psi_m[\mathbf{x} - \mathbf{c}_g(\mathbf{k}_0)t, 0] \exp\{-i[\mathbf{k}_0 \cdot \mathbf{x} - \sigma_m(\mathbf{k}_0)t]\}; \end{aligned} \quad (38.6)$$

$$\mathbf{c}_g(\mathbf{k}) = \frac{\partial \sigma}{\partial \mathbf{k}} = \frac{\beta}{(k^2 + f_0^2/gh_m)^2} \left(k_x^2 - k_y^2 - \frac{f_0^2}{gh_m}, 2k_x k_y \right).$$

Therefore, the wave packet is made up of a plane wave moving in direction $\mathbf{k}_0$ at the phase velocity $\sigma_m(\mathbf{k}_0)/k_0$ and a modulated envelope $\psi_m(\mathbf{x} - \mathbf{c}_g(\mathbf{k}_0)t, 0]$, moving without changing its shape at the group velocity $\mathbf{c}_g(\mathbf{k}_0)$. Unlike the phase velocity, the zonal component c_{gx} of the group velocity (describing the propagation of the wave-packet energy) may be directed either to the west (if $k_x^2 < k_y^2 + f_0^2/gh_m$) or to the east (if $k_x^2 > k_y^2 + f_0^2/gh_m$). Let us consider the reflection of a plane wave from a wall $x = 0$. Assume that at $x > 0$ we have an incident wave $\psi_i = \cos(k_i x + k_y y - \sigma t)$, the energy of which propagates westward ($k_i^2 < k_y^2 + f_0^2/gh_m$). Because of the condition of nontransmission $(\partial/\partial y)(\psi_i + \psi_r) = 0$ for $x = 0$, the values of k_y and σ do not change upon reflection. The energy of the reflected wave $\psi_r = -\cos(k_r x + k_y y - \sigma t)$ must propagate eastward ($k_r^2 > k_y^2 + f_0^2/gh_m$). Then (38.5) yields

$$k_r = \frac{\beta}{2\sigma} - \left(\frac{\beta^2}{4\sigma^2} - k_y^2 - \frac{f_0^2 \sigma^2}{gh_m} \right)^{1/2}. \quad (38.7)$$

The wave vectors and the group-velocity vectors of the incident and reflected waves $(\mathbf{k}_i, \mathbf{k}_r)$ and $(\mathbf{c}_{gi}, \mathbf{c}_{gr})$ are shown in Figure 38.2, from which it is seen that for the group velocity the law that 'the angle of incidence equals the angle of reflection' holds.

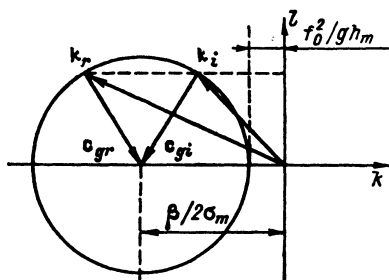


Fig. 38.2. Wave vectors and group-velocity vectors of wave $(\mathbf{k}_i, \mathbf{c}_{gi})$ incident upon western shore and of reflected wave $(\mathbf{k}_r, \mathbf{c}_{gr})$.

tion' is satisfied. Note here that $|\mathbf{k}_r| > |\mathbf{k}_i|$, that is, for reflection from the west side the spatial scales of the disturbances decrease (it is easy to show that for reflection from the east wall, on the other hand, $|\mathbf{k}_r| < |\mathbf{k}_i|$). In the real ocean atmospheric effects engender primarily almost zonal wave packets ($k_x^2 \ll k_y^2$) with westward group velocities, and their reflection from western shores, according to the foregoing, produces there intense boundary currents with a small transverse scale: much smaller than $|k_y|^{-1}$ (Pedlosky, 1965; Lighthill, 1969).

When considering the field of Rossby waves on the scale of the entire ocean, or all the more the entire atmosphere, we should give up the ' β -plane' approximation and return to analyzing the required families of solutions of the Laplace tidal equation on a sphere, considered in Section 10. We will confine ourselves here to a study of not very long waves, for which the ratio f_0^2/ghk^2 is small. In this case the boundary condition at the ocean surface (38.2) for $z = 0$ can be replaced by the condition $w \sim \partial^2 \psi / \partial t \partial z = 0$ (the 'rigid cover' approximation), and the barotropic waves become nondivergent. But we will linearize the spherical analog of Equation (37.5) relative to not a state of rest but rather a westerly transport with a velocity $\bar{u}_\lambda = \alpha r_0 \sin \theta$, where α is the constant velocity of zonal rotation of the atmosphere (known as the *circulation index*). Then instead of (38.3) we obtain

$$\left(\frac{\partial}{\partial t} + \alpha \frac{\partial}{\partial \lambda} \right) \Delta \psi + \frac{2(\alpha + \omega)}{r_0} \frac{\partial \psi}{\partial \lambda} = 0, \quad (38.8)$$

where Δ is the two-dimensional Laplace operator at the surface of a sphere of radius r_0 . It has particular solutions of the form $P_n^m(\mu) e^{i(m\lambda - \sigma_n^m t)}$, where $\mu = \cos \theta$; P_n^m are associated Legendre polynomials; $m = 1, 2, \dots$; $n = m, m + 1, \dots$, and frequencies σ_n^m have the form

$$\sigma_n^m = \alpha m - \frac{2(\alpha + \omega)m}{n(n+1)} \quad (38.9)$$

(Haurwitz, 1940). In Section 10 these solutions (for $\alpha = 0$) were obtained from (10.16), that is, as asymptotes for small ε for those solutions for which in this case the frequencies remain finite. Note that the inclusion of index α in this formula

plays an important role on Venus. For its extremely slow spin (one rotation takes about 243 days) the horizontal scale of synoptic processes $L_0 = c_0/2\omega$, introduced by Obukhov, turns out to be much larger than the dimensions of the planet. However, because of the four-day circulation observed on Venus, $\alpha + \omega \approx 2\pi$ (4 days) $^{-1}$, so that synoptic processes with a scale $L_0 \sim c_0/2(\alpha + \omega)$ turn out to be possible (Monin, 1981), and these are actually observed.

Blinova (1943) summarized the indicated particular solutions, to obtain from the linearized equations of motion (9.2) the following solution of the problem with initial data for the nonzonal part of the pressure field on a sphere:

$$\begin{aligned} (z')_{t=0} &= \sum_{n=m}^{\infty} \sum_{m=1}^{\infty} (A_n^m \cos m\lambda + B_n^m \sin m\lambda) P_n^m(\mu); \\ z' &= \sum_{n=m}^{\infty} \sum_{m=1}^{\infty} [C_n^m \cos(m\lambda - \sigma_n^m t) + D_n^m \sin(m\lambda - \sigma_n^m t)] Z_n^m(\mu); \quad (38.10) \\ Z_n^m &= \frac{n(n-m+1)}{(n+1)(2n+1)} P_{n+1}^m + \frac{(n+1)(n+m)}{n(2n+1)} P_{n-1}^m, \end{aligned}$$

where coefficients C_n^m , D_n^m of the prognostic formula are expressed in terms of coefficients A_n^m , B_n^m of the series expansion of the initial pressure field in spherical harmonics from the equations

$$\begin{aligned} (C_{n-1}^m, D_{n-1}^m) \frac{(n-1)(n-m)}{n(2n-1)} + (C_{n+1}^m, D_{n+1}^m) \frac{(n+2)(n+m-1)}{(n+1)(2n+3)} \\ = (A_n^m, B_n^m). \end{aligned} \quad (38.11)$$

Consequently, formula (38.10) provided a linear adiabatic forecast of the barotropic mode of the pressure field on a sphere, thereby anticipating much of the fluid-dynamical theory of weather forecasting worked out since the Second World War.

As mentioned above, the approximation yielding solutions with frequencies (38.9) is an asymptote for low $\varepsilon = (2r_0\omega/c_0)^2$, where $c_0 = (gh_0)^{1/2}$. However, in the Earth's atmosphere $r_0\omega/c_0 \approx 1.4$ and $\varepsilon \approx 8$, and for small m , n formula (38.9) turns out to be a rather coarse approximation (although with increasing m , n its accuracy increases rapidly). At the same time, on Jupiter, for instance, $r_0\omega/c_0 \approx 17$, and we have to use the asymptote for large ε , which is derived from (10.17') by analogy with the quasiclassical approximation in quantum mechanics. For $\alpha = 0$, as stated in Section 10, this asymptote is found from the relation $E\varepsilon^{1/2} = 2n+1$, that is,

$$\varepsilon^{1/2} \left(\frac{\sigma}{2\omega} \right)^2 + m\varepsilon^{-1/2} \left(\frac{2\omega}{\sigma} \right) = 2n+1, \quad (38.12)$$

analogous to Bohr's quantization formula (Golitsyn and Dikii, 1966–1967; in these works the corresponding eigenfunctions were also calculated).

39. TOPOGRAPHIC WAVES

Now let us consider an ocean with an irregular bottom, the relief of which is described by the equation $z = -h + h'(\mathbf{x}_h)$, where h is the mean depth and h' is the elevation of the bottom above this mean level. For linear waves, assuming h' to be relatively small, boundary condition (38.2) should be moved to the mean level and written as

$$z = -h: \quad \frac{f_0}{N^2} \frac{\partial^2 \psi}{\partial t \partial z} = - \frac{\partial(\psi, h')}{\partial(x, y)}. \quad (39.1)$$

First let us look at, following Rhines (1970), the case of an *inclined* bottom with constant values of $h_x = \partial h' / \partial x$ and $h_y = \partial h' / \partial y$ for $N = \text{const}$ and in the 'rigid cover' approximation at the ocean surface. In this case for the amplitudes $\Psi(z)$ of linear waves we obtain the following equations and boundary conditions:

$$\begin{aligned} \frac{\partial^2 \Psi}{\partial z^2} - \mu^2 \Psi &= 0; \\ z = 0: \quad \frac{\partial \Psi}{\partial z} &= 0; \quad z = -h: \quad \frac{\partial \Psi}{\partial z} = L_R^2 \frac{k_x \beta}{h \sigma} \gamma \Psi; \\ \gamma &= \frac{f_0}{h \beta} \left(h_y - \frac{k_y}{k_x} h_x \right); \quad \mu = \frac{N}{f_0} \left(k^2 + \frac{k_x \beta}{\sigma} \right)^{1/2}. \end{aligned} \quad (39.2)$$

For $k^2 + k_x \beta / \sigma > 0$ it is found that $\Psi = \text{ch } \mu z$, while for μ the boundary condition at the bottom yields the equation

$$\text{th } \mu h = -\gamma \mu h \left[1 - \left(\frac{k L_R}{\mu h} \right)^2 \right],$$

which always has only one root μ_0 . To this root there corresponds a zero topographic mode with a frequency

$$\sigma_0 = -\beta k_x \left(k^2 - \frac{\mu_0^2 h^2}{L_R^2} \right)^{-1}. \quad (39.3)$$

For $k^2 + k_x \beta / \sigma < 0$ it is found that $\mu = i\lambda$ with a real λ , so that $\Psi = \cos \lambda z$, while for λ the boundary condition at the bottom yields the equation

$$\text{tg } \lambda h = -\gamma \lambda h \left[1 + \left(\frac{k L_R}{\lambda h} \right)^2 \right]$$

with a denumerable set of roots $\lambda_1, \lambda_2, \dots$. To these correspond topographic waves with frequencies

$$\sigma_m = -\beta k_x \left(k^2 + \frac{\lambda_m^2 h^2}{L_R^2} \right)^{-1}; \quad m = 1, 2, \dots \quad (39.4)$$

For $\gamma > 0$, modes $\Psi_m(z)$, $m = 0, 1, 2, \dots$, each have m zeros on segment $[-h, 0]$. For $\gamma \rightarrow 0$ and finite kL_R , they become the corresponding modes for an ocean with a flat bottom. For $\gamma \sim 1$ in the case of long waves $kL_R \ll 1$ we obtain

$$\sigma_0 \approx -k^2 \left[k_x \left(\beta + \frac{f_0 h_y}{h} \right) - \frac{k_y f_0 h_x}{h} \right],$$

that is, the contribution of the bottom relief is here linearly added to the β -effect; in (39.4) we can drop k^2 , so that long baroclinic waves turn out to be non-dispersive. For $\gamma \sim 1$ in the case of short waves $kL_R \gg 1$ we have $\mu_0 h \approx kL_R$ and $\sigma_0 \approx -Nk^{-1}(k_x h_y - k_y h_x)$, that is, the zero mode is baroclinic and does not depend on the β -effect; moreover, here $\lambda_m h \approx \pi(m - \frac{1}{2})$ and $\sigma_m \approx -\beta k_x k^{-2}$, that is, these modes are barotropic and the bottom relief has the sole effect of replacing m by $m - \frac{1}{2}$.

For $\gamma < 0$, modes $\Psi_0(z)$ and $\Psi_1(z)$ do not have zeros, while each of the other modes $\Psi_m(z)$ have $m - 1$ zeros. For $\gamma \rightarrow 0$ and finite kL_R the zero mode is concentrated at the bottom and $\sigma_0 \rightarrow 0$, while modes $m = 1, 2, \dots$ become modes $m - 1$ for an ocean with a flat bottom. For finite γ in the case of long waves modes $m = 2, 3, \dots$ turn out to be nondispersive, modes $\Psi_0(z)$ for $\gamma < -1$ and $\Psi_1(z)$ for $\gamma > -1$ are barotropic, while $\Psi_0(z)$ for $\gamma \geq -1$ and $\Psi_1(z)$ and $\gamma \geq -1$ are, on the other hand, baroclinic. In the case of short waves the results are the same as for $\gamma > 0$.

For all modes the bottom slope can either enhance or weaken the β -effect, but for $m = 1, 2, \dots$ it is always true that $\sigma_m/k_x < 0$; with regard to the zero mode, for $\gamma < 0$ it moves eastward, while for $\gamma > 0$ it moves westward.

In the rest of this section we will consider only barotropic waves, for which, by definition, inhomogeneities of the density field are insignificant. Then, in view of the quasihydrostatic nature $\partial p/\partial z \approx -g\rho \approx \text{const}$, so that the geostrophic horizontal velocity is almost independent of z . Because of the nondivergence, this is also valid for $\partial w/\partial z$, and thus

$$\frac{\partial w}{\partial z} = \frac{w(z = \xi) - w(z = -h + h')}{H} = \frac{1}{H} \frac{d_h H}{dt}; \quad (39.5)$$

$$H = \xi + h - h',$$

where H is the total thickness of the fluid layer. On the other hand, the first of Equations (37.1) yields $w \approx -(\partial \eta_0/\partial z)^{-1} d_h \eta'/dt$, and, in view of this result and (39.5), Equation (37.3) in the case of barotropic motions becomes

$$\frac{d_h}{dt} \frac{\Delta \psi + f}{H} = 0. \quad (39.6)$$

Linearizing this equation and using the 'rigid cover' approximation ($\xi = 0$), we obtain

$$\frac{\partial \Delta \psi}{\partial t} + \beta \frac{\partial \psi}{\partial x} + \frac{f_0}{h} \frac{\partial(\psi, h')}{\partial(x, y)} = 0. \quad (39.7)$$

If L_h is the horizontal scale of the bottom irregularities and L is the typical

wavelength, then for $\varepsilon = L/L_h \ll 1$ the bottom can be termed *sloping*, while during the calculation of the horizontal gradients $\varepsilon h'$ can be assumed to be a function of the 'slow' coordinates $\xi = \varepsilon x$, $\eta = \varepsilon y$ (above we considered the particular case in which this function is linear, although baroclinicity was taken into account). Introducing as well the 'slow' time $\tau = \varepsilon t$, we seek a wave solution of Equation (39.7) in the form $\psi = A \exp(i\varepsilon^{-1}\Phi)$ (where the amplitude A and the 'fast' phase Φ are functions of the 'slow' variables), so that $\mathbf{k} = (\partial\Phi/\partial\xi, \partial\Phi/\partial\eta)$ and $\sigma = -\partial\Phi/\partial\tau$ play the roles of the local wave vector and frequency. Substituting this solution into (39.7) and equating to zero the coefficients of ε^0 and ε^1 , we arrive at the equations:

$$|\nabla\Phi|^2 \frac{\partial\Phi}{\partial\tau} = \frac{\partial(\Phi, K)}{\partial(\xi, h)}; \quad K = \beta\eta + \varepsilon \frac{f_0}{h} h'; \quad (39.8)$$

$$\frac{\partial}{\partial\tau} (A |\nabla\Phi|^2) = \frac{\partial(A, K)}{\partial(\xi, \eta)} - (A \nabla\Phi + 2\nabla A \cdot \nabla\Phi) \frac{\partial\Phi}{\partial\tau}. \quad (39.9)$$

The first of these equations, which locally can be rewritten as $\sigma k^2 = k_y \partial K / \partial \xi - k_x \partial K / \partial \eta$, is a dispersion relation. It shows that for $h' = 0$ the phase velocity of the waves is directed toward the west, and also that the waves being considered here can also exist for $\beta = 0$, being produced actually by the same mechanism as Rossby waves.

If $\varepsilon_1 = L_h/L \ll 1$, then the bottom is called *rough*. Let us consider a cylindrical bottom relief (Rhines and Bretherton, 1973) and for simplicity let us assume that $h' = h'(y)$. We introduce a 'fast' ordinate $\eta = \varepsilon_1^{-1}y$ and we denote $(f_0/h) \partial h' / \partial \eta = b(\eta)$. Using this notation, we seek for (39.7) a wave solution $\psi = \Psi(\eta) \exp[i(k_x x - \sigma t)]$, and we get

$$\frac{\partial^2 \Psi}{\partial \eta^2} - \varepsilon_1 \frac{k_x}{\sigma} b(\eta) \Psi - \varepsilon_1^2 \left(k_x^2 + \frac{\beta k_x}{\sigma} \right) \Psi = 0. \quad (39.10)$$

Since we are interested in large-scale motions over a small-scale bottom relief, let us look for function Ψ in the form of a series $\Psi_0(y) + \varepsilon_1 \Psi_1(\eta, y) + \dots$ with the requirement that all Ψ_m remain finite as $|y|, |\eta| \rightarrow \infty$. For Ψ_1 from (39.10) we obtain the equation $\partial^2 \Psi_1 / \partial \eta^2 = (k_x / \sigma) b(\eta) \Psi_0(y)$, from which it follows that $\Psi_1 = (k_x / \sigma) \beta(\eta) \Psi_0(y) + R_1(y)$, where $B(\eta) = (f_0/h) \int h' d\eta$ (because Ψ_1 remains finite as $|\eta| \rightarrow \infty$, the term $\eta C(y)$ does not have to be added on here). For Ψ_2 (39.10) yields

$$\begin{aligned} \frac{\partial^2 \Psi_2}{\partial \eta^2} + 2 \frac{\partial^2 \Psi_1}{\partial \eta \partial y} \frac{k_x}{\sigma} b(\eta) \Psi_1 + \\ + \frac{\partial^2 \Psi_0}{\partial y^2} - \left(k_x^2 + \frac{\beta k_x}{\sigma} \right) \Psi_0 = 0. \end{aligned} \quad (39.11)$$

By applying to this expression the averaging operation $\lim_{D \rightarrow \infty} (1/D) \int_{\eta_0}^{\eta_0+D} d\eta$,

we obtain

$$\frac{\partial^2 \Psi_0}{\partial y^2} + \left(-k_x^2 - \frac{\beta k_x}{\sigma} + \frac{f_0^2}{h^2} \frac{k_x^2}{\sigma^2} \overline{h'^2} \right) \Psi_0 = 0 \quad (39.12)$$

(according to Volosov, 1976, we can, similarly, determine Ψ_2 from (39.11) and average the equation for Ψ_3 to arrive at term R_1 in the expression Ψ_1 , and so forth, while in the nonlinear case instead of $\overline{h'^2}$ the dynamics of the large-scale motions is governed by the correlation function of the bottom relief $\overline{h'(\eta)h'(\eta + \xi)}$). Function Ψ_0 turns out to be finite and equal to $\exp(ik_y y)$, if the expression in parentheses in (39.12) is positive; the value of k_y in this case corresponds to the dispersion relation

$$\sigma^2 + \sigma \beta k_x k^{-2} - f_0^2 k_x^2 k^{-2} \overline{h'^2} h^{-2} = 0. \quad (39.13)$$

The waves being considered here for $\beta = 0$ do not disappear, so that the bottom roughness is able to generate large-scale waves. This capability is comparable to the β -effect, if $h'/h \sim L/r_0$. Correction $\varepsilon_1 \Psi_1$ is small, but it is a 'fast' function of y , that is, $\varepsilon_1 d\Psi_1/d\eta$ is comparable to $\partial\Psi_0/\partial\eta$. In other words, the velocities (and thus the energy) of the large-scale and small-scale motions are found to be comparable. This serves as a warning against smoothing the bottom relief during calculations of large-scale flows.

For a wavy bottom $L_h \sim L$ analytical calculations succeed only in the simplest particular cases. Individual bottom features (peculiarities) may give rise either to partial reflection of waves or to the appearance of new wave types: *trapped modes*. Assume, as previously, a cylindrical bottom with a peculiarity within a band $0 \leq \eta \leq 1$; in the band the depth difference h_0 is a maximum, and outside it the depths $h_1 = h$ and $h_2 < h$ are constant, so that $b(\eta) = 0$. For an incident wave with an amplitude $\Psi = A \exp(ik_y y)$ and a frequency $\sigma = -\beta k_x k^{-2}$ the amplitude $\Psi(y)$ of the entire wave field will satisfy the equation

$$\frac{\partial^2 \Psi}{\partial \eta^2} + \left[\varepsilon_1^2 k_y^2 - \varepsilon_1 \frac{k_x}{\sigma} b(\eta) \right] \Psi = 0. \quad (39.14)$$

Let us seek its solution for $\eta < 0$ in the form $\Psi_1^{(0)} + \varepsilon_1 \Psi_1^{(1)} + \dots = A \exp(ik_y y) + (B_0 + \varepsilon_1 B_1 + \dots) \exp(-ik_y y)$, for $0 \leq \eta \leq 1$ in the form $\Psi_2^{(0)}(\eta) + \varepsilon_1 \Psi_2^{(1)}(\eta) + \dots$, and for $\eta > 1$ in the form $\Psi_3^{(0)} + \varepsilon_1 \Psi_3^{(1)} + \dots = (C_0 + \varepsilon_1 C_1 + \dots) \exp(ik_y y)$, requiring at the boundaries continuity of Ψ and $\partial\Psi/\partial\eta$. A first approximation is easily obtained in the form

$$C_0 = A \left[1 + (ik_x/2k_y\sigma) \int_0^1 b(\eta) d\eta \right]^{-1}; \quad B_0 = C_0 - A; \quad (39.15)$$

$$T = |C_0|^2 |A|^{-2} = \left[1 + \frac{1}{4} \frac{k_x^2}{k_y^2} \frac{f_0^2}{\sigma^2} \left(1 - \frac{h_2}{h_1} \right)^2 \right]^{-1},$$

so that the *transmission coefficient* T as a first approximation is independent of the width L_h and the structure of the bottom feature $b(\eta)$, being a function only of the relative depth difference $1 - h_2/h_1$. Strong reflection $T \ll 1$ is possible even for small depth differentials. If, however, $h_2 = h_1$, that is, the bottom feature is a ridge or a depression on a level bottom, then $T = [1 + O(\varepsilon_1^2)]^{-1}$ (Rhines, 1969).

To describe trapped modes, we will use Equation (39.10), denoting in it $k_x^2 + \beta k_x/\sigma = k_y^2$, and we will seek a solution which goes to zero exponentially with increasing distance from the bottom feature, that is,

$$\Psi_{1,3} = \Psi_{1,3}^{(0)} + \varepsilon_1 \Psi_{1,3}^{(1)} + \dots = (A_{1,3}^{(0)} + \varepsilon_1 A_{1,3}^{(1)} + \dots) \exp(-k_y |y|).$$

This yields

$$A_{1,3}^{(0)} = -A_1^{(0)} \left(\frac{k_x}{2k_y \sigma} \right) \int_0^1 b(\eta) d\eta, \quad (39.16)$$

$$\sigma = -\frac{\beta}{2k_x} + \frac{1}{2} \operatorname{sign}(h_2 - h_1) \left[\frac{\beta^2}{k_x^2} + f_0^2 \left(1 - \frac{h_2}{h_1} \right)^2 \right]^{1/2}.$$

From this it is seen that trapped modes also exist for $\beta = 0$, and then $\sigma = f_0(h_2 - h_1)/2h_1$ does not depend on k_x , that is, purely topographical trapped waves are as a first approximation nondispersive, but the β -effect introduces dispersion into them.

40. STATIONARY WAVES

First of all, let us define a stationary wave, or field, as one which propagates without any change in its form at some constant velocity c , directed, say, toward the east. Then, by analogy with (28.1), any stationary field $A = A(x - ct, y)$ will be a Lagrangian invariant of the horizontal solenoidal motion, that is, a function of $\psi + cy$ (so that for it the equation $d_h A/dt = 0$ becomes $\partial(A, \psi + cy)/\partial(x, y) = 0$). Therefore, all the steady-state solutions of Equation (37.5), in which $A = \mathcal{L}\psi + \beta y$, must satisfy the condition

$$\mathcal{L}\psi + \beta y = F(\psi + cy), \quad (40.1)$$

where F is an arbitrary function. The boundary condition at the free surface, applied at the equilibrium surface $z = 0$, reduces to the condition that there $\partial\psi/\partial z + N^2\psi/g$ is an arbitrary function $F_0(\psi + cy)$; in the 'rigid cover' approximation this reduces just to $\partial\psi/\partial z$. The boundary condition at a cylindrical bottom $z = -h + h'(y)$, applied at its mean level $z = -h$, reduces to the condition that there $\partial\psi/\partial z + N^2 h'/f_0$ is an arbitrary function $F_h(\psi + cy)$; for a flat sea bottom this once again reduces just to $\partial\psi/\partial z$. For simplicity, let us confine ourselves now to the 'rigid cover' approximation and to waves over a flat bottom. A linear function $F = -k_0^2$ and zero F_0 and F_h are selected as simplest. A pure wave solution (without a zonal current superimposed on it) is obtained for $k_0^2 =$

$-\beta/c$, that is, from the equation

$$\left(\mathcal{L} - \frac{\beta}{c} \right) \psi = 0. \quad (40.2)$$

Its solutions with separated variables $\psi = \psi_m(x - c_m t, y) \Psi_m(z)$ are obtained from (38.3), (38.4) (in the 'rigid cover' approximation), so that harmonic waves with phase velocities $c_m = -\beta(k^2 + f_0^2/gh_m)^{-1}$ satisfy a nonlinear equation as well as a linear equation.

Of the nonlinear stationary waves let us first examine slightly nonlinear and weakly dispersive waves on a zonal current with a meridional velocity shear; for simplicity, the waves are assumed to be barotropic, but moving over a cylindrical bottom relief $z = -h + h'(y)$. We will consider long waves with typical scales L in the x direction and εL , where $\varepsilon \ll 1$, in the y direction. If we measure coordinates x and y by these scales, and the stream function by the scale $\beta \varepsilon^3 L^3$, then the above-mentioned Lagrangian invariant A can be represented in dimensionless variables as

$$A = \left(\varepsilon^2 \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) (\bar{\psi} + \varepsilon^2 \psi') + y + b,$$

where $\bar{\psi}(y)$ describes the westerly zonal current, $\psi'(x, y)$ describes the wave, and $b(y)$ describes the bottom relief. The quantity A will be an arbitrary function of $\bar{\psi} + \varepsilon^2 \psi' + cy$, where c is the phase velocity of the wave, measured in terms of scale $\beta \varepsilon^2 L^2$. After equating to zero the Jacobian of these two quantities with respect to (x, y) and setting $\psi' = \psi_0 + \varepsilon^2 \psi_1 + \dots$ and $c = c_0 + \varepsilon^2 c_1 + \dots$, in the two lowest orders of ε^2 we arrive at the equations

$$\frac{\partial}{\partial x} \left(\frac{\partial^2 \psi_0}{\partial y^2} + G \psi_0 \right) = 0; \quad (40.3)$$

$$G(y) = - \left(1 + \frac{\partial^3 \bar{\psi}}{\partial y^3} + \frac{\partial b}{\partial y} \right) \left(c_0 + \frac{\partial \bar{\psi}}{\partial y} \right)^{-1};$$

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial^2 \psi_1}{\partial y^2} + G \psi_1 \right) = & - \frac{\partial^3 \psi_0}{\partial x^3} - c_1 \left(c_0 + \frac{\partial \bar{\psi}}{\partial y} \right)^{-1} \frac{\partial^3 \psi_0}{\partial x \partial y^2} + \\ & + \left(c_0 + \frac{\partial \bar{\psi}}{\partial y} \right)^{-1} \frac{\partial(\psi_0, \partial^2 \psi_0 / \partial y^2)}{\partial(x, y)}. \end{aligned} \quad (40.4)$$

If we seek ψ_0 in the form $\Phi(x)\Psi(y)$, then Ψ can be found from the relation $\partial^2 \Psi / \partial y^2 + G \Psi = 0$, with, say, zero boundary conditions at the channel walls $y = 0, Y$ (which also gives the eigenvalue c_0). Multiplying (40.4) by Ψ and integrating

the result over y over the width of the channel (integrating by parts where necessary), we obtain

$$\frac{\partial \Phi}{\partial x} + \frac{a_1}{c_1} \Phi \frac{\partial \Phi}{\partial x} + \frac{a_2}{c_1} \frac{\partial^3 \Phi}{\partial x^3} = 0; \quad (40.5)$$

$$a_1 = - \left[\int_0^y \Psi^2 G \left(c_0 + \frac{\partial \bar{\Psi}}{\partial y} \right)^{-1} dy \right]^{-1} \int_0^y \Psi^3 \frac{\partial G}{\partial y} \left(c_0 + \frac{\partial \bar{\Psi}}{\partial y} \right)^{-1} dy;$$

$$a_2 = - \left[\int_0^y \Psi^2 G \left(c_0 + \frac{\partial \bar{\Psi}}{\partial y} \right)^{-1} dy \right]^{-1} \int_0^y \Psi^2 dy.$$

Thus, it is seen (Long, 1964) that the stationary field Ψ' as a first approximation satisfies an equation of the KdV type (22.19), indicating that such fields can include cnoidal waves (22.14) and solitons (22.15). For example, a soliton with an amplitude A has the form

$$\Phi(x) = \text{sign}(a_1 a_2) A \text{ch}^{-2} \left[A^{1/2} \left| \frac{a_1}{a_2} \right|^{1/2} \frac{x}{2} \sqrt{3} \right], \quad (40.6)$$

while the nonlinear correction to the soliton velocity in this case is $c_1 = -\text{sign}(a_1 a_2) a_1 A/3$. Note that for $G(y) = \text{const}$ the effect of nonlinearity vanishes, so that to provide it here we must have either a shear of the zonal velocity or nonlinear variations of the depth. Interestingly enough, this is also valid for the opposite asymptote: short waves with scales L in the x and y directions, on a zonal current with a scale $\varepsilon^{-1} L$, $\varepsilon \ll 1$, in the y direction. Following Larichev and Reznik (1976), let us consider barotropic waves over a flat bottom. If the stream function of the zonal current is measured to a scale $\beta L^3 \varepsilon^{-1}$ and that of the waves to a scale $\beta L^3 \varepsilon^{-1} \alpha$, while the phase velocity of a wave is measured to a scale βL^2 , and if we introduce 'slow' coordinates $\xi = \varepsilon x$ and $\eta = \varepsilon y$, then we obtain (40.1) in dimensionless variables:

$$\varepsilon^2 \frac{\partial^2 \bar{\psi}}{\partial \eta^2} + \alpha \left(\frac{\partial^2 \psi'}{\partial x^2} + \frac{\partial^2 \psi'}{\partial y^2} \right) + \eta = F(\bar{\psi} + \alpha \psi' + c\eta). \quad (40.7)$$

Taking into account that ψ' depends continuously on α in the interval $[0, O(1)]$, at the limit as $\alpha \rightarrow 0$ we arrive at an equation for function F in terms of the specified function $\bar{\psi}(\eta)$:

$$F(\bar{\psi} + c\eta) = \eta + \varepsilon^2 \frac{\partial^2 \bar{\psi}}{\partial \eta^2}. \quad (40.8)$$

Invoking the smallness of ε , let us seek a solution of Equations (40.7), (40.8) in the form $F = F_0 + \varepsilon^2 F_2$ and $\alpha \psi' = \psi_0 + \varepsilon \psi_1 + \dots$, where all ψ_n are functions of the 'slow' coordinates (ξ, η) and 2π -periodic functions of the 'fast' phase $\theta = \varepsilon^{-1} \varphi(\xi, \eta)$ (so that ψ' is sought in the form of a quasilplane wave with a

slowly varying wave vector $\mathbf{k}(\xi, \eta) = (\partial\varphi/\partial\xi, \partial\varphi/\partial\eta)$, frequency $\sigma = c/k$, and an amplitude). From (40.8) we then have

$$F_0(\bar{\psi} + c\eta) = \eta; \quad F_2(\bar{\psi} + c\eta) = \frac{\partial^2 \bar{\psi}}{\partial \eta^2}. \quad (40.9)$$

Equation (40.7) yields the following equations for ψ_0 and ψ_1 :

$$k^2 \frac{\partial^2 \psi_0}{\partial \theta^2} + \frac{\partial Q}{\partial \psi_0} = 0; \quad Q(\psi_0) = \eta\psi_0 - \int F_0(\bar{\psi} + \psi_0 + c\eta) d\psi_0; \quad (40.10)$$

$$k^2 \frac{\partial^2 \psi_1}{\partial \theta^2} - \frac{\partial F_0}{\partial \psi_0} \psi_1 = - \frac{\partial \psi_0}{\partial \theta} \operatorname{div} \mathbf{k} - 2\mathbf{k} \nabla \frac{\partial \psi_0}{\partial \theta}. \quad (40.11)$$

For Equation (40.10) to have a periodic solution $\psi_0(\theta)$, function $Q(\psi_0)$ must go through a minimum. Clearly, this occurs at $\psi_0 = 0$, so that the following conditions are satisfied:

$$\frac{\partial Q}{\partial \psi_0} = \eta - F_0(\bar{\psi} + c\eta) = 0$$

and

$$\frac{\partial^2 Q}{\partial \psi_0^2} = - \left[\frac{\partial F_0(\bar{\psi} + \bar{\psi}_0 + c\eta)}{\partial \psi_0} \right]_{\psi_0=0} - \left(c + \frac{\partial \bar{\psi}}{\partial \eta} \right)^{-1} > 0$$

(the first of these is identical to the first of Equations (40.9), while the second is derived from it by differentiating with respect to η). The inequality $c + \partial \bar{\psi} / \partial \eta < 0$ signifies that the wave moves westward relative to the mean current, just as in the linear theory. Moreover, taking (40.10) to be an equation in function $\theta(\psi_0)$, we obtain its solution in quadratures:

$$\begin{aligned} \theta - \theta_0 &= \pm \frac{k}{\sqrt{2}} \int_{\tau_1}^{\tau_0} (E - Q)^{-1/2} d\psi_0; \\ 2\pi &= k\sqrt{2} \int_{\tau_1}^{\tau_2} (E - Q)^{-1/2} d\psi_0, \end{aligned} \quad (40.12)$$

where E does not depend on ψ_0 (but may depend on ξ, η), while τ_1, τ_2 are the minimum and maximum roots of the equation $Q(\psi_0) = E$. The second of these equations is a simple statement of the 2π periodicity of function $\psi_0(\theta)$; it plays the part of the dispersion relation for the nonlinear wave. Multiplying (40.11) by $\partial \psi_0 / \partial \theta$ and integrating over θ from 0 to 2π , we obtain

$$\operatorname{div} \left\{ \left[\int_0^{2\pi} \left(\frac{\partial \psi_0}{\partial \theta} \right)^2 d\theta \right] \mathbf{k} \right\} = 0. \quad (40.13)$$

Equations (40.12), (40.13) can be used to find functions $\varphi(\xi, \eta)$ and $E(\xi, \eta)$

and as a first approximation they describe a periodic nonlinear wave. As a specific example, Larichev and Reznik calculated a wave on a zonal current with a constant shear s , so that $\bar{\psi} = \frac{1}{2}s\eta^2$. Because of the inequality $c + \partial\bar{\psi}/\partial\eta < 0$, a solution is possible here only in the region $\eta < -c/s$. If we set $c + s\eta = \xi$, then from the first of Equations (40.9) we obtain $F_0(\psi_0, \eta) = -cs^{-1} - s^{-1}(\xi^2 + 2s\psi_0)^{1/2}$. Passing in (40.10) from the unknown ψ_0 to $\Psi_0 = \xi^{-1}(\xi^2 + 2s\psi_0)^{1/2}$, multiplying this by $\partial\Psi_0/\partial\theta$, and integrating over θ , we obtain

$$\frac{3}{8} k^2 |\xi| \left(\frac{\partial\Psi_0}{\partial\theta} \right)^2 = (\gamma - \Psi_0)(\Psi_0 - \beta_1)(\Psi_0 - \beta_2), \quad (40.14)$$

where γ is an arbitrary constant and

$$\beta_{1,2} = \frac{3}{4} - \frac{\gamma}{2} \pm \frac{\sqrt{3}}{2} \left(\frac{3}{4} + \gamma - \gamma^2 \right)^{1/2},$$

so that $\gamma > \beta_1 > \beta_2$. This equation is integrated in elliptic functions with a modulus $\kappa = (\gamma - \beta_1)^{1/2}(\gamma - \beta_2)^{-1/2}$ and an argument

$$\chi = \arcsin[\kappa^{-1}(\Psi_0 - \beta_1)^{1/2}(\Psi_0 - \beta_2)^{-1/2}]:$$

Equations (40.12) here become:

$$\begin{aligned} \frac{\theta - \theta_0}{k\sqrt{6\xi}} &= \beta_2(\gamma - \beta_2)^{-1/2}F(\chi, \kappa) + (\gamma - \beta_2)^{1/2}E(\chi, \kappa) - \\ &\quad - (\gamma - \Psi_0)^{1/2}(\Psi_0 - \beta_1)^{1/2}(\Psi_0 - \beta_2)^{-1/2}; \\ c + \frac{\partial\bar{\psi}}{\partial\eta} &= - \left(\frac{\pi^2}{6k^2} \right) [\gamma(\gamma - \beta_2)^{-1/2}F(\kappa) + (\gamma - \beta_2)^{1/2}E(\kappa)]^{-2}, \end{aligned} \quad (40.15)$$

where F and E are incomplete and complete elliptic integrals of the first and second kinds. Finally, the integral in (40.13) is found to be

$$\begin{aligned} \int_0^{2\pi} \left(\frac{\partial\psi_0}{\partial\theta} \right)^2 d\theta &= \frac{2\sqrt{2}}{\sqrt{3}} \frac{\xi^{7/2}}{s^2k} \left[(\gamma - \beta_2)^{1/2} \left(\frac{18}{35} + \frac{6}{7} \gamma\beta_1\beta_2 \right) E(\kappa) + \right. \\ &\quad \left. + 6(\gamma - \beta_2)^{-1/2} \left(\frac{3}{35} \beta_2 + \frac{1}{35} \gamma\beta_1\beta_2 + \frac{1}{7} \gamma\beta_1\beta_2^2 \right) F(\kappa) \right]. \end{aligned} \quad (40.16)$$

Calculations have shown that strong nonlinearity ($\kappa \rightarrow 1$) produces deviations of up to 25% from the sinusoidal wave shape $\psi_0(\theta)$, broadening the maxima and narrowing the minima, but the deviation $k^2|c + \partial\bar{\psi}/\partial\eta|$ of the dispersion relation from linear as κ varies from 0 to 1 increases above 1 only up to $\pi^2/9 \approx 1.1$. The asymptote of weak nonlinearity $\alpha \rightarrow 0$ can be derived from (40.14)–(40.16) for low ψ_0 , but it is simpler to obtain it directly from (40.10) with the aid of the Taylor series $\partial Q/\partial\psi_0 = b_1\psi_0 + b_2\psi_0^2 + \dots$ and the series $\psi_0 = \alpha\psi_{01} + \alpha^2\psi_{02} +$

$\dots$, where $\psi_{01}, \psi_{02}, \dots$ are periodic functions of $2\pi\theta/T$ and the period $T = 2\pi(b_1^{1/2}k^{-1} + \alpha^2\omega_1 + \dots)^{-1}$ is equal to 2π . If we assume here that $\partial\psi_0/\partial\theta = 0$ for $\theta = \theta_0$ and if we specify the wave energy $\alpha^2 E = k^2[\frac{1}{2}(\partial\psi_0/\partial\theta)^2 + Q]$, then for $\theta = \theta_0$ we get $\psi_{0n} = \psi_{0n}^{(0)}(E)$ and $\partial\psi_{0n}/\partial\theta = 0$. Let us introduce the notation:

$$\bar{c}(\eta) = -\frac{\partial\bar{\psi}}{\partial\eta}, \quad A = \alpha[2E(\bar{c} - c)]^{1/2},$$

$$B = \frac{1}{3}(\bar{c} - c)^{-4} \left(\frac{\partial\bar{c}}{\partial\eta} \right)^2 + \frac{1}{4}(\bar{c} - c)^{-3} \frac{\partial^2\bar{c}}{\partial\eta^2}.$$

Finally, to ensure that ψ_{03} is finite, we require that the coefficient go to zero for the resonance term in the equation for ψ_{03} proportional to $\cos(2\pi\theta/T)$. In summation we then have

$$\begin{aligned} \psi_0 = & A \cos(\theta - \theta_0) + \frac{A^2}{12} \frac{\partial\bar{c}}{\partial\eta} (\bar{c} - c)^{-2} \cos 2(\theta - \theta_0) - \\ & - \frac{A^2}{4} \frac{\partial\bar{c}}{\partial\eta} (\bar{c} - c)^2 + O(\alpha^3); \end{aligned} \quad (40.12')$$

$$k^2(c - \bar{c}) = - \left[1 + \frac{A^2}{2} B + O(\alpha^4) \right];$$

$$\text{div} \left\{ A^2 \left[1 - \frac{3}{8} A^2 B + O(\alpha^4) \right] \mathbf{k} \right\} = 0. \quad (40.13')$$

Consequently, to the nonlinear wave we have to add a multiple harmonic and a wave-produced mean flow, analogous to the Stokes drift, while in the dispersion relation a dependence on the wave amplitude makes its appearance. As mentioned above, for all these effects to occur, there must be a shear in the velocity of the zonal current $\partial\bar{c}/\partial\eta$.

Equations (40.10), (40.11), etc., for $\psi_n(\theta)$ may also have soliton solutions, exponentially going to zero as $|\theta| \rightarrow \infty$. Equation (40.10) then asymptotically becomes

$$k^2 \frac{\partial^2\psi_0}{\partial\eta^2} - \left(c + \frac{\partial\bar{\psi}}{\partial\eta} \right)^{-1} \psi_0 = 0,$$

so that a soliton solution is possible only on condition that $c + \partial\bar{\psi}/\partial\eta > 0$. This means that in contrast to a periodic wave a soliton has to move eastward relative to the mean current. Function $Q(\psi_0)$ has in this case a maximum at $\psi_0 = 0$. From (40.10) it follows that for finite solutions $\psi_0(\theta)$ function $Q(\psi_0)$ must have at least one minimum at some $\psi_0 = \psi_{\min}$. If in (40.12) we set $E = Q(0)$, then asymptotically as $|\theta| \rightarrow \infty$ we will have

$$\psi_0 \sim |\psi_0| \exp[-(c - \bar{c})^{-1/2} k^{-1} |\theta|].$$

In Equation (40.13) the integral in this case is taken from $-\infty$ to $+\infty$ rather than from 0 to 2π .

Function $F_0(\xi)$, describing the distribution of vorticity on the streamlines $\xi = \bar{\psi} + \psi_0 + c\eta$, was a monotonically decreasing function for periodic waves, but in the case of a soliton it goes through a minimum. It is seen from (40.12) that ψ_0 varies from $\psi_c = \tau_1 < 0$ to $\tau_2 = 0$. From the definition of Q we get

$$\int_{\psi_c}^0 F_0(\Psi + \psi_0) d\psi_0 + \eta\psi_c = 0, \quad \Psi = \bar{\psi} + c\eta,$$

where $F_0(\Psi) = \eta$. Now, passing to the variable $\xi = \Psi + \psi_0$ and differentiating with respect to η , we obtain

$$[F_0(\Psi_1) - \eta] \frac{\partial \Psi_1}{\partial \eta} = \Psi_1 - \Psi; \quad \Psi_1 = \Psi + \psi_c. \quad (40.17)$$

Thus, for the boundary condition $\Psi_1 = \Psi$ at $\eta = \eta_1$, according to the specified function $F_0(\xi)$ we can determine function $\Psi_1(\eta)$ and thus the soliton amplitude $|\psi_c(\eta)|$ as well. On the other hand, if $\psi_c(\eta)$ is known, then we can find $F_0(\xi)$, $\xi < \xi_{\min}$. For instance, on a zonal current *without a velocity shear* ($c(\eta) = \text{const}$) we have $F_0(\xi) = \xi/\xi_0$ for $\xi > 0$, where $\xi_0 = c - \bar{c} > 0$. Assume, for instance, that $\psi_c = -a\eta$. Then $\Psi = \xi_0\eta$ and $\Psi_1 = (\xi_0 - a)\eta$, so that (40.17) yields $F_0(\xi) = -\xi/\xi_1$ for $\xi < 0$, where $\xi_1 = (2a - \xi_0)(a - \xi_0)^{-2} > 0$. From (40.12) we then have

$$\frac{\psi_0}{\eta} = \begin{cases} -\xi_1 \left(1 + \frac{\xi_0}{\xi_1} \right)^{1/2} \cos(\xi_1^{-1/2} k^{-1} \theta) - \xi_0 + \xi_1; & |\theta| \leq \theta_0; \\ -\xi_0 \exp[-\xi_0^{-1/2} k^{-1} (\theta - \theta_0)]; & \\ |\theta| > \theta_0 = k\xi_1^{1/2} \arccos \left(1 + \frac{\xi_0}{\xi_1} \right)^{-1/2}, & \end{cases} \quad (40.18)$$

while from the soliton analog of Equation (40.13) for $k_x = \text{const}$ we obtain $k_y = \pm a(\eta^4/b^4 - 1)^{-1/2}$, where b is a constant. Therefore, here we have also constructed an example of a soliton without a shear of the velocity of the zonal current. Other such examples can be constructed for a barotropic flow over a cylindrical bottom, in which the Jacobian of the functions

$$\Delta\psi - \frac{f_0^2}{gh} \psi + \beta y + \frac{f_0}{h} h'(y) \quad \text{and} \quad \psi + cy$$

is equal to zero. If we measure the horizontal lengths to a scale L , the stream function of the soliton $\psi' \sim UL$ and of the zonal current $\bar{\psi} \sim \varepsilon UL$, where $\varepsilon = \beta L^2/U \sim c/U \ll 1$ (that is, if we consider the case of strong nonlinearity

$U \gg c$) and if we measure c to a scale βL^2 , then in dimensionless variables we have

$$\begin{aligned} \frac{\partial(\psi', \Delta\psi')}{\partial(x, y)} + \varepsilon \frac{\partial}{\partial x} [\alpha(y) \Delta\psi' + \gamma(y) \psi'] &= 0; \\ \alpha(y) &= - \left(c + \frac{\partial \bar{\psi}}{\partial y} \right); \\ \gamma(y) &= 1 + \frac{f_0^2 L^2}{gh} c + \frac{\partial^3 \bar{\psi}}{\partial y^3} + \frac{f_0}{\beta L h} + \frac{\partial h'}{\partial y}. \end{aligned} \quad (40.19)$$

Let us seek a solution for ψ' and c , using the series $\psi' = \psi_0 + \varepsilon \psi_1 + \dots$, $\alpha = \alpha_0 + \varepsilon \alpha_1 + \dots$, $\gamma = \gamma_0 + \varepsilon \gamma_1 + \dots$, where α_0 and γ_0 do not depend on y , while ψ_0 is a function only of $r = (x^2 + y^2)^{1/2}$, that is, the soliton being sought is almost axisymmetric (Larichev and Reznik, 1976; Berestov, 1981). Then from (40.19) it follows that function $\psi_0(r)$ can be arbitrary, whereas for ψ_1 in polar coordinates (r, θ) after a single integration over θ we obtain

$$\Delta \psi_1 + \gamma_0 r \sin \theta = \Psi_0(r) (\psi_1 - \alpha_0 r \sin \theta) + \Psi_1(r), \quad (40.20)$$

where $\Psi_0 = (\partial \Delta \psi_0 / \partial r) (\partial \psi_0 / \partial r)^{-1}$ and Ψ_1 is an arbitrary function. The solution can be looked for in the form $\Psi = \Phi_0(r) \sin \theta + \Phi_1(r)$, and for coefficients Φ_0 and Φ_1 we get the equations

$$\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial \Phi_0}{\partial r} - (1 + \Psi_0) \Phi_0 = -r(\alpha_0 \Psi_0 + \gamma_0); \quad (40.21)$$

$$\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial \Phi_1}{\partial r} - \Psi_0 \Phi_1 = \Psi_1. \quad (40.22)$$

We require of Φ_0 that this function go to zero as $r \rightarrow \infty$ more rapidly than any power of r . This entails, first of all, the condition $\alpha_0 \gamma_0 < 0$; for $\bar{\psi} = h' = 0$ it means that the soliton moves eastward. The condition that the right-hand side of (40.21) go to zero rapidly as $r \rightarrow \infty$ yields the asymptote

$$\psi_0 \sim r^{-1/2} \exp(-|\alpha_0 \gamma_0|^{1/2} r).$$

Continuity of the velocity at the center of the soliton requires the condition $\partial \psi_0 / \partial r = 0$ there. At the other zeros of the r_k of this function the conditions of smoothness must be satisfied:

$$\left(\frac{\partial^3 \psi_0}{\partial r^3} \right) \left(\frac{\partial^2 \psi_0}{\partial r^2} \right)^{-1} + r_k^{-1} = 0.$$

Finally, the total angular momentum of the soliton must be equal to zero, which yields $\int_0^\infty (\partial \psi_0 / \partial r) r^2 dr = 0$, so that the angular velocity $\partial \psi_0 / \partial r$ has to change sign: in the soliton circulation in both directions must take place. Coefficient Φ_0 is found from Ψ_0 , but Φ_1 remains arbitrary (because of the arbitrariness of Ψ_1). The

total stream function of the soliton in its central part differs only slightly from $\psi_0(r)$, although crossclosure of the boundary streamline is observed, probably because this soliton is only an asymptote of the exact soliton solution of Equation (40.1). We will now go on to consider some of these exact solutions.

The first of these was constructed by Stern (1975) in the simplest case of a barotropic ocean, using the 'rigid cover' approximation and assuming a flat bottom, without a mean current and assuming a zero velocity of the soliton c . Equation (40.1) in this case simplifies to $\partial(\Delta\psi + \beta y, \psi)/\partial(x, y) = 0$. Converting to polar coordinates via the relations $x = r \cos \theta$, $y = r \sin \theta$, and seeking a solution in the form $\psi = \Psi(r) \sin \theta$, we arrive at an equation having the form

$$\xi \frac{\partial \eta}{\partial \xi} - \eta = 0; \quad \eta = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} - \frac{1}{r^2} \right) \Psi + \beta r; \quad (40.23)$$

$$\xi = \Psi.$$

Stern considered a solution in the form $\eta = -k^2 \xi$ for $\xi < 0$, which yields $\Psi = AJ_1(kr) - \beta k^{-2}r$, where J_1 is a Bessel function of the first kind and A is an arbitrary constant. Let us so select it that $\Psi(a) = 0$, that is, we assume $A = \beta k^{-2}a/J_1(ka)$. Next we so select k that $\Psi'(a) = 0$. For this the equation $J_1'(ka)/J_1(ka) = 1/ka$ must be satisfied, which with the aid of the recurrence formula $zJ_1'(z) = J_1(z) - zJ_2(z)$ leads to $J_2(ka) = 0$, that is, ka must be one of the roots of function $J_2(z)$ (it can be shown that then for $r = a$ the quantity η also goes to zero, and thus the absolute vorticity $\Delta\psi + \beta y$ as well). These roots are extrema of the function $z^{-1}J_1(z)$, and if we select as ka the first minimum of this function (it is negative), that is, the first root of function $J_2(z)$, then for $r < a$ we will have $\Psi < 0$; however, *the other roots of this condition are not obtained thereby*. Consequently, for this selection of k we can set

$$\psi = \frac{\beta a}{k^2} \left[\frac{J_1(kr)}{J_1(ka)} - \frac{r}{a} \right] \sin \theta \quad \text{for } r < a, \quad (40.24)$$

while for $r > a$ we can choose the trivial solution $\psi = 0$. Thus, in such a soliton, the velocity field and the absolute vorticity will be continuous for $r = a$ (Stern, however, did not note that the simple choice of different solutions ψ for $r < a$ and $r > a$, some conditions of their matching at $r = a$ being specified, would be incorrect; he also did not offer a means of so selecting those roots ka that η would depend only on ξ , rather than on both ξ and r ; this was first pointed out by Berestov, 1979).

The next step was taken by Larichev and Reznik (1976), who considered a barotropic ocean, assuming a free surface and an inclined flat bottom. In this case β is replaced by

$$\beta' = \beta + \frac{f_0^2 c}{gh} + \frac{f_0}{h} \frac{\partial h'}{\partial y}$$

without a mean current, but with a nonzero soliton velocity c . Here

$$\eta = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \right) \Psi(r) + \beta' r$$

is found to be a piecewise-linear function of $\xi = \Psi + cr$:

$$\eta(\xi) = \begin{cases} -k^2 \xi, & \xi < 0, \\ p^2 \xi, & \xi > 0. \end{cases} \quad (40.25)$$

In order to ensure an exponential decrease of ψ as $\xi \rightarrow \infty$, we will set $p^2 = \beta'/c$; then ψ will be asymptotically proportional to

$$r^{-1/2} \exp \left[- \left(\frac{\beta'}{c} \right)^{1/2} r \right]$$

(it can be shown that for $\beta' = 0$ no stationary waves exponentially decaying as $r \rightarrow \infty$ will exist at all). The condition $\beta'/c > 0$ implies that either $c > 0$ (that is, the soliton moves eastward) or

$$c < - \left(\beta + \frac{f_0}{h} \frac{\partial h'}{\partial y} \right) \frac{gh}{f_0^2}$$

(that is, the soliton moves westward quite rapidly). Therefore, the soliton velocity must lie outside of the interval of permissible velocities of periodic waves (otherwise their generation would lead to decay of the soliton). From (40.25) we obtain the following continuous solution for $\Psi(r)$:

$$\Psi = \begin{cases} \frac{\beta' a}{k^2} \frac{J_1(kr)}{J_1(ka)} - \frac{\beta' + ck^2}{k^2} r & \text{for } \Psi + cr < 0, \\ -ca \frac{K_1(pr)}{K_1(pa)} & \text{for } \Psi + cr > 0, \end{cases} \quad (40.26)$$

where K_1 is the MacDonald function. The requirement of continuous $\Psi'(r)$ for $r = a$ can, with the aid of recurrence formulas for cylindrical functions, be reduced to

$$- \frac{1}{ka} \frac{J_2(ka)}{J_1(ka)} = \frac{1}{pa} \frac{K_2(ka)}{K_1(ka)}. \quad (40.27)$$

Since $p = (\beta'/c)^{1/2}$, this condition is a dispersion relation making it possible to determine c for given k and a . It ensures continuity for $r = a$ of the second derivative $\Psi''(r)$ as well, and thus also of the vorticity (but $\Psi'''(r)$ for $r = a$ exhibits a finite discontinuity). It should be noted that the condition $\Psi + cr > 0$ in the second line of (40.26) is satisfied for $r > 0$, for any p and a , but the condition $\Psi + cr < 0$ in the first line is satisfied for $r < a$ only if ka lies in the

interval between the first zero and the first minimum of function $J_1(kr)/kr$. According to (40.27), a single value of ka from this interval corresponds to each pa .

A soliton $\psi = \Psi(r) \sin \theta$ with $\Psi(r)$ from formula (40.26) is a vortex dipole, the streamlines in which are depicted in Figure 40.1. Interactions of such solitons were investigated in the numerical experiments of Makino *et al.* (1981) and also those of Larichev and Reznik (1982). The results were varied. For a central collision of solitons moving in the same direction, at certain values of the parameters both solitons 'survived', as in Figure 40.2, whereas sometimes only one 'survived'. In the case of oncoming collisions both solitons always 'survived', with the vortexes of the weaker soliton passing around the peripheries of the more intense one.

Flierl (1979) and Larichev *et al.* (1980) showed that to the solution $\psi = \Psi(r) \sin \theta$ of formula (40.26) we can add the function $\Psi_1(r)$, equal to $LJ_0(kr) - M$ for $r < a$ and $K_0(pr)$ for $r > a$, where

$$L = \frac{pK_1(pa)}{kJ_1(ka)} \quad \text{and} \quad M = 2paK_1(pa)[(ka)^{-2} + (pa)^{-2}],$$

but in this case $\partial^2 \psi / \partial r^2$ has a discontinuity at $r = a$. Berestov (1979) indicated a three-zoned soliton, for which

$$\Psi = \begin{cases} -\frac{\beta'a}{p^2} \frac{I_1(pr)}{I_1(pa)} & \text{for } r < a, \\ BJ_1(kr) + E \frac{\beta'}{k^3} N_1(kr) - \frac{\beta' + ck^2}{k^2} r & \text{for } a < r < b \quad (40.28) \\ -\frac{\beta'b}{p^2} \frac{K_1(pr)}{K_1(pb)} & \text{for } r > b, \end{cases}$$

where the parameters must satisfy the conditions:

$$\begin{aligned} B &= \frac{\beta'a}{k^2 J_1(ka)} \left[1 - E \frac{N_1(ka)}{ka} \right] = \frac{\beta'b}{k^2 J_1(kb)} \left[1 - E \frac{N_1(kb)}{kb} \right]; \\ \frac{1}{ka} \frac{J_2(ka)}{J_1(ka)} \left[1 - \frac{2E}{\pi ka J_2(ka)} \right] &= \frac{1}{pa} \frac{J_2(pa)}{J_1(pa)}; \\ -\frac{1}{kb} \frac{J_2(kb)}{J_1(kb)} \left[1 - \frac{2E}{\pi kb J_2(kb)} \right] &= \frac{1}{pb} \frac{K_2(pb)}{K_1(pb)}, \end{aligned} \quad (40.28')$$

while the conditions stating that η depends only on ξ have the form

$$\frac{kr - EN_1(kr)}{J_1(kr)} < (>) \frac{ka - EN_1(ka)}{J_1(ka)} \quad \text{for } J_1(kr) < (>) 0, \quad (40.28'')$$

$$a < r < b.$$

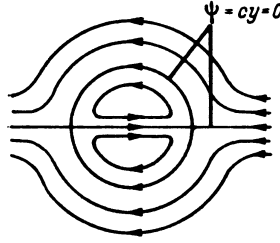


Fig. 40.1. Barotropic soliton of Larichev and Reznik (1976).

Here to $\Psi(r) \sin \theta$ we can also add some function $\psi_1(r)$.

This same investigator constructed for the case $N = \text{const}$ a three-dimensional baroclinic soliton; to do this, he changed from Cartesian coordinates to spherical coordinates via the relations

$$x = r \sin \theta \cos \lambda, \quad y = r \sin \theta \sin \lambda, \quad z = \left(\frac{f_0}{N} \right) r \cos \theta$$

and he sought a solution of the equation

$$\frac{\partial(\mathcal{L}\psi + \beta y, \psi + cy)}{\partial(x, y)} = 0$$

in the form $\psi = \Psi(r, \theta) \sin \lambda$. As a result, this equation reduces to a form analogous to (40.23): $\xi \partial \eta / \partial \xi = \eta$, where $\eta = \mathcal{L}_1 \Psi + \beta r \sin \theta$ with the operator

$$\mathcal{L}_1 = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} - \frac{1}{r^2 \sin^2 \theta}$$

and

$$\xi = \Psi + cr \sin \theta.$$

For $\eta(\xi)$ we again choose a piecewise-linear solution (40.25). For an exponential decrease of ψ as $r \rightarrow \infty$ we should set $p^2 = \beta/c > 0$ (so that $c > 0$, that is, the soliton moves eastward). Then ψ is found to be asymptotically proportional to $r^{-1} \exp[-(\beta/c)^{1/2} r]$ (for $c < 0$ and $p^2 < 0$ a solution also exists, but it does not go to zero exponentially; for instance, for $-k^2 = p^2 = \beta/c$ it goes to zero as r^{-1}). Function Ψ is sought in the form $\Psi_1(r) \sin \theta$, which yields

$$\Psi_1(r) = \begin{cases} \frac{\beta a}{k^2} \frac{r^{-1/2} J_{3/2}(kr)}{a^{-1/2} J_{3/2}(ka)} - \frac{\beta + ck^2}{k^2} r & \text{for } r < a, \\ -ca \frac{r^{-1/2} K_{3/2}(pr)}{a^{-1/2} K_{3/2}(pa)} & \text{for } r > a. \end{cases} \quad (40.29)$$

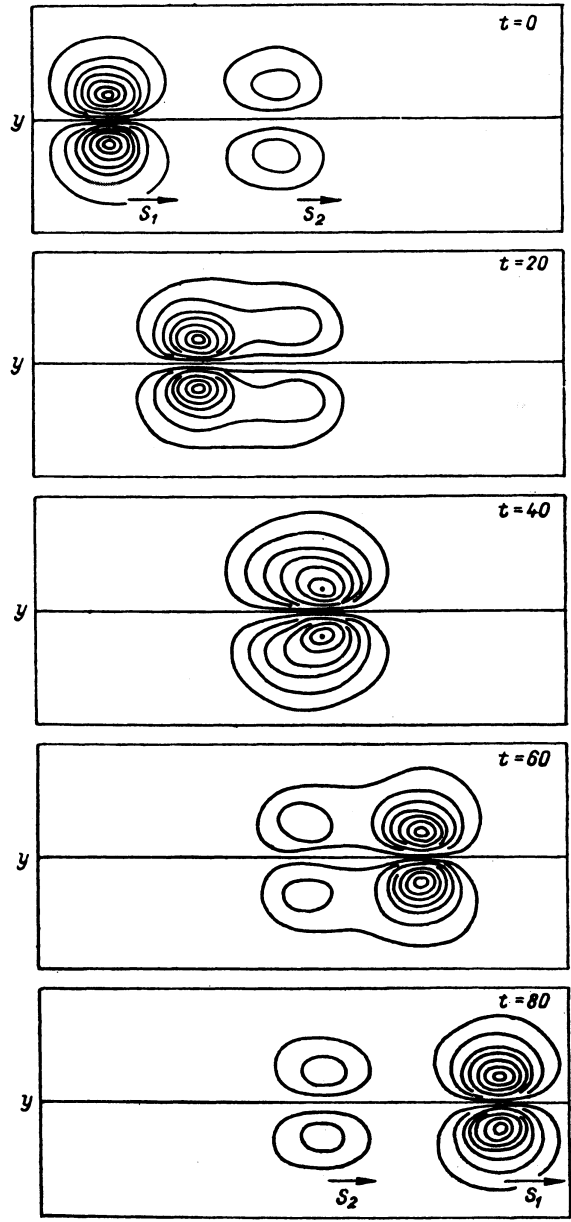


Fig. 40.2. Passage of one barotropic soliton through another, according to numerical experiment of Makino *et al.* (1981).

The condition of continuity of $\Psi_1'(r)$ reduces to

$$-\frac{1}{ka} \frac{J_{5/2}(ka)}{J_{3/2}(ka)} = \frac{1}{pa} \frac{K_{5/2}(pa)}{K_{3/2}(pa)}, \quad (40.30)$$

which is analogous to (40.27). The requirement that $\xi > 0$ for $r > 0$ is satisfied automatically, but the requirement that $\xi < 0$ for $r < a$ calls for a choice of just one of the roots ka of Equation (40.30), namely the root lying in the interval between the first zero and the first minimum of function $(kr)^{-3/2} J_{3/2}(kr)$. For $r = a$ function $\Psi_1''(r)$ is here continuous, but $\Psi_1'''(r)$ has a finite discontinuity. The streamlines in the xy and yz planes in this soliton are shown in Figure 40.3 (in the xz plane $\psi \equiv 0$). We note, too, that a three-dimensional soliton at rest similar to the solution of Stern (40.24) is obtained if we set $c = 0$ in the first line of formula (40.29), if the second line is replaced by the condition $\Psi_1 = 0$ for $r > a$, and if the first root of the equation $J_{5/2}(ka) = 0$ is used.

Berestov also constructed a three-layered baroclinic soliton, in which

$$\Psi_1(r) = \begin{cases} -ca \frac{r^{-1/2} I_{3/2}(pr)}{a^{-1/2} I_{3/2}(pa)} & \text{for } r < a; \\ Br^{-1/2} J_{3/2}(kr) + E \frac{\beta r^{-1/2}}{k^{7/2}} N_{3/2}(kr) - \frac{\beta + ck^2}{k^2} r & \text{for } a < r < b, \\ -cb \frac{r^{-1/2} K_{3/2}(pr)}{b^{-1/2} K_{3/2}(pb)} & \text{for } r > b, \end{cases} \quad (40.31)$$

while the parameters have to satisfy the conditions

$$\begin{aligned} B &= \frac{\beta a}{k^2} \frac{1}{a^{-1/2} J_{3/2}(ka)} \left[1 - \frac{E}{(ka)^{3/2}} N_{3/2}(ka) \right] \\ &= \frac{\beta b}{k^2} \frac{1}{b^{-1/2} J_{3/2}(kb)} \left[1 - \frac{E}{(kb)^{3/2}} N_{3/2}(kb) \right]; \end{aligned} \quad (40.31')$$

$$\begin{aligned} \frac{1}{ka} \frac{J_{5/2}(ka)}{J_{3/2}(ka)} \left[1 - \frac{2E}{\pi(ka)^{5/2} J_{5/2}(ka)} \right] &= \frac{1}{pa} \frac{I_{5/2}(pa)}{I_{3/2}(pa)}; \\ -\frac{1}{kb} \frac{J_{5/2}(kb)}{J_{3/2}(kb)} \left[1 - \frac{2E}{\pi(kb)^{5/2} J_{5/2}(kb)} \right] &= \frac{1}{pb} \frac{K_{5/2}(pb)}{K_{3/2}(pb)}, \end{aligned}$$

and the conditions of dependence of η solely on ξ are in this case

$$\frac{(kr)^{3/2} - EN_{3/2}(kr)}{J_{3/2}(kr)} < (>) \frac{(ka)^{3/2} - EN_{3/2}(ka)}{J_{3/2}(ka)} \quad (40.31'')$$

for $J_{3/2}(ka) < (>) 0, \quad a < r < b.$

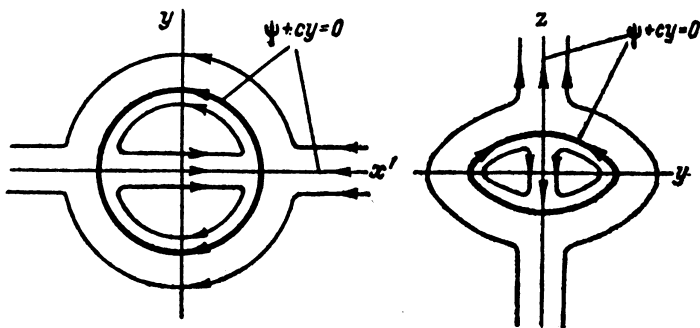


Fig. 40.3. Three-dimensional soliton of Berestov (1979).

It is apparently impossible, in general, to construct similar piecewise-linear solutions $\eta(\xi)$ with more than three layers, since the number of parameters in them would be less than the number of conditions which must be satisfied. Let us note that, as in the barotropic case, the stream functions ψ of two-layered and three-layered baroclinic solitons may be supplemented by some definite purely radial terms $\Psi_2(r)$ (although this disrupts the continuity of $\partial^2\psi/\partial r^2$). We note, too, that if the coordinate origin is placed at the ocean surface $z = 0$, then the lower halves of all the above-mentioned three-dimensional solitons will satisfy at this surface the 'rigid-cover' condition, which means that they can be considered to be independent solitons (Kizner, 1984).

Kizner (1984) proposed considering baroclinic solitons of another kind, having noted that arbitrary functions of z can be added to the quantities $\mathcal{L}\psi + \beta'y$ and $\psi + cy$ in (40.1). Let us select, for instance, $\mathcal{L}\psi + \beta'y - CF(z)$ in the form of a piecewise-linear function (40.25) of $\psi + cy$, while as function $F(z)$ we take a solution of the Sturm—Liouville problem of type (38.4):

$$\frac{\partial}{\partial z} \frac{f_0^2}{N^2} \frac{\partial F}{\partial z} + m^2 F = 0; \quad F'(0) = F'(-h) = 0. \quad (40.32)$$

Then the solution of Equation (40.1) in cylindrical coordinates r, θ, z can be sought tentatively in the form $\psi = \psi_0(r, \theta) + \Phi(r)F(z)$, where $\psi_0(r, \theta)$ is the barotropic soliton of Larichev and Reznik. For $\Phi(r)$ in this case we then have

$$\Phi'' + \frac{1}{r} \Phi' - m^2 \Phi - C = \begin{cases} -k^2 \Phi & \text{for } r < a, \\ p^2 \Phi & \text{for } r > a. \end{cases} \quad (40.33)$$

The solution of these equations for $k^2 - m^2 > 0$ has the form

$$\Phi(r) = \begin{cases} A_0 J_0(r \sqrt{k^2 - m^2}) + (k^2 - m^2)^{-1} C & \text{for } r < a, \\ B_0 K_0(r \sqrt{p^2 + m^2}) - (p^2 + m^2)^{-1} C & \text{for } r > a. \end{cases} \quad (40.34)$$

Note that as $r \rightarrow \infty$ function Φ does not decrease to zero; rather it approaches some constant value as $r^{-1/2} \exp(-r\sqrt{p^2 + m^2})$, that is, more rapidly than ψ_0 goes to zero. Therefore, ψ is asymptotically proportional to $F(z)$, but this does not disrupt the localization of the soliton, since as $r \rightarrow \infty$ the horizontal velocities u, v go to zero exponentially. The conditions of continuity of Φ , Φ' , and Φ'' at $r = a$ enable us to determine ratios A_0/C and B_0/C and to arrive at a dispersion relation for the baroclinic mode. The latter relation has the form

$$-\frac{a\sqrt{k^2 - m^2}J_0(a\sqrt{k^2 - m^2})}{J_1(a\sqrt{k^2 - m^2})} = \frac{a\sqrt{p^2 + m^2}K_0(a\sqrt{p^2 + m^2})}{K_1(a\sqrt{p^2 + m^2})}. \quad (40.35)$$

We recall that $p = (\beta'/c)^{1/2}$ and that ak is defined in terms of ap in (40.27). Consequently, for each eigenvalue m of problem (40.32) and for a specified a from (40.35), we can in principle determine ap , and thus also find the soliton velocity $c = a^2\beta'/(ap)^2$. It should be noted, however, that the possibility of ensuring the dependence of $\mathcal{L}\psi + \beta'y - CF(z)$ just on $\psi + cy$ here remained unsolved.

41. HAMILTONIAN FORMALISM

This powerful method was used in Section 24 to describe surface waves and in Section 30 to describe internal gravity waves. It is also applicable to the theory of Rossby waves, since Equation (37.5) characterizing them is a Hamiltonian system (for barotropic waves this was pointed out independently by Zakharov and Kuznetsov, 1982, and by Vainshtein, 1983). The use of this method provides a uniform orderly exposition of the theory of waves of different nature (especially in the approximation of weak nonlinearity): the dynamical equations (24.4) in normal canonical variables can be distinguished only by the form of the dispersion relations α_k and the interaction coefficients $V_{kk_1k_2}$, $U_{kk_1k_2}$, etc.

However, the introduction of normal canonical variables in the theory of Rossby waves proved to be a nontrivial matter. It was accomplished (by our request, for the purposes of this book) for barotropic waves by Zakharov and Piterbarg (1986). For barotropic waves, Equation (37.5) takes the form of (12.11):

$$\frac{\partial \Omega}{\partial t} + \frac{\partial(\psi, \Omega)}{\partial(x, y)} + \beta \frac{\partial \psi}{\partial x} = 0; \quad \Omega = \Delta \psi - k_0^2 \psi, \quad (41.1)$$

where $k_0^2 = f_0^2/gh$ (the quantity $\Omega + f_0 + \beta y$ is the potential vorticity). This equation can be put into Hamiltonian form:

$$\frac{\partial \Omega}{\partial t} = \{\Omega, \mathcal{H}\}; \quad \mathcal{H} = \frac{1}{2} \int [|\nabla \psi|^2 + k_0^2 |\psi|^2] dx dy, \quad (41.2)$$

where $\mathcal{H}$ is the Hamiltonian (proportional to the sum of the kinetic energy and the barotropic available potential energy, see (12.10)), and the braces represent the so-called Poisson bracket, in the given case defined for any two functionals

$F[\Omega]$ and $G[\Omega]$ by the formula

$$\begin{aligned}\{F, G\} &= \{F, G\}_0 + \{F, G\}_1 \\ &= \int (\Omega + \beta y) \frac{\partial(\delta F / \delta \Omega, \delta G / \delta \Omega)}{\partial(x, y)} dx dy; \\ \{F, G\}_1 &= \beta \int \frac{\delta F}{\delta \Omega} \frac{\partial}{\partial x} \frac{\delta G}{\delta \Omega} dx dy.\end{aligned}\quad (41.3)$$

Here $\{F, G\}_0$ is the value of $\{F, G\}$ for $\beta = 0$; this Poisson bracket specifies the Hamiltonian structure of the two-dimensional fluid dynamics of an incompressible fluid (and it is found to be a Lie bracket for a group of area-conserving diffeomorphisms of the plane). Braces $\{F, G\}_1$ with constant (that is, independent of Ω) coefficients represent the so-called Gardner bracket from the theory of integrable systems, for which a means of introducing normal canonical variables is known.

The idea of Zakharov and Piterbarg is as follows: bracket $\{F, G\}$ can be put into the form $\{F, G\}_1$ by replacing the functional argument $\Omega(x, y)$ by some new argument $\zeta(x, y)$. To do this, we can set

$$\Omega(x, y) = \zeta \left[x, y + \frac{1}{\beta} \Omega(x, y) \right]; \quad (41.4)$$

$$\zeta(x, y) = \Omega \left[x, y - \frac{1}{\beta} \zeta(x, y) \right] \quad (41.4')$$

(these relations are equivalent), for which functions

$$w = y - \frac{1}{\beta} \zeta(x, y); \quad (41.5)$$

$$y = w + \frac{1}{\beta} \Omega(x, w) \quad (41.5')$$

must be convertible (they can be converted into each other). This requirement means that the isolines of the potential vorticity are not closed, that is, solitons are absent, which will be the case for waves of sufficiently small amplitude. Relation (41.4') can with the aid of (41.5) be rewritten in the form $\zeta(x, y) = \Omega(x, w)$, and after differentiation with respect to y and with respect to x we have

$$\frac{\partial \zeta}{\partial y} = \frac{\partial \Omega}{\partial w} \left(1 - \frac{1}{\beta} \frac{\partial \zeta}{\partial y} \right); \quad \frac{\partial \zeta}{\partial x} = \frac{\partial \Omega}{\partial x} \left(1 - \frac{1}{\beta} \frac{\partial \zeta}{\partial y} \right). \quad (41.6)$$

Now the proof of the theorem of Zakharov and Piterbarg is obtained as follows:

$$\begin{aligned}
 \{F, G\}_{1\xi} &= -\beta \int \frac{\delta G}{\delta \xi(x, y)} \frac{\partial}{\partial x} \frac{\delta F}{\delta \xi(x, y)} dx dy \\
 &= -\beta \int \frac{\delta G}{\delta \Omega(x, w)} \left[\frac{\partial}{\partial x} \frac{\delta F}{\delta \Omega(x, w)} - \right. \\
 &\quad \left. - \frac{1}{\beta} \frac{\partial \xi}{\partial x} \frac{\partial}{\partial w} \frac{\delta F}{\delta \Omega(x, w)} \right] \frac{dx dw}{1 - (1/\beta) (\partial \xi / \partial y)} \\
 &= -\beta \int \frac{\delta G}{\delta \Omega(x, w)} \frac{\partial}{\partial x} \frac{\delta F}{\delta \Omega(x, w)} dx dw + \\
 &\quad + \int \frac{\delta G}{\delta \Omega(x, w)} \left[\frac{\partial \Omega}{\partial x} \frac{\partial}{\partial w} \frac{\delta F}{\delta \Omega(x, w)} - \right. \\
 &\quad \left. - \frac{\partial \Omega}{\partial w} \frac{\partial}{\partial x} \frac{\delta F}{\delta \Omega(x, w)} \right] dx dw = \{F, G\}_{\Omega}, \quad (41.7)
 \end{aligned}$$

where in the second line it is taken into account that during the variable change $(x, y) \rightarrow (x, w)$ it must be assumed that

$$\frac{\partial}{\partial x} \rightarrow \frac{\partial}{\partial x} + \left(\frac{\partial w}{\partial x} \right) \frac{\partial}{\partial w}, \quad \frac{\partial w}{\partial x} = - \left(\frac{1}{\beta} \right) \frac{\partial \xi}{\partial x},$$

while in the following equation relations (41.6) are used. We note, too, that the implicit formulas (41.4) and (41.4'), for the change of the functional variable $\Omega(x, y) \rightarrow \xi(x, y)$ at small Ω , ξ , can be converted into explicit series expansions of these quantities:

$$\begin{aligned}
 \Omega(x, y) &= \sum_{n=0}^{\infty} \frac{1}{(n+1)! \beta^n} \frac{\partial^n}{\partial y^n} \xi^{n+1} \\
 &= - \frac{1}{2\pi i} \oint_{|z|=1} \ln \left[1 - \frac{1}{z} \xi \left(x, y + \frac{1}{\beta} z \right) \right] dz; \quad (41.8)
 \end{aligned}$$

$$\begin{aligned}
 \xi(x, y) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)! \beta^n} \frac{\partial^n}{\partial y^n} \Omega^{n+1} \\
 &= - \frac{1}{2\pi i} \oint_{|z|=1} \ln \left[1 - \frac{1}{z} \Omega \left(x, y - \frac{1}{\beta} z \right) \right] dz \quad (41.8')
 \end{aligned}$$

(the representation in the form of contour integrals is obtained assuming functions Ω , ξ to be analytic in the band $-1/\beta < \text{Im } y < 1/\beta$). Now the normal canonical variables $a_k(t)$ for Equation (41.1) are introduced with the aid of the formula

$$\xi(\mathbf{k}, t) = \frac{1}{\pi} \sqrt{\frac{\beta}{2}} \int_{k_x < 0} [a_k(t) e^{i\mathbf{k} \cdot \mathbf{x}} + a_k^*(t) e^{-i\mathbf{k} \cdot \mathbf{x}}] |k_x|^{1/2} d\mathbf{k}, \quad (41.9)$$

and in these variables the Hamilton equations and the representation of the Hamiltonian in the form of a functional power series (24.4) are valid, where all the integrations are carried out over the halfplane $k_x < 0$. The dispersion relation for linear barotropic Rossby waves has the form $\sigma_k = -\beta k_x (k^2 + k_0^2)^{-1}$, while the coefficients of the three-wave interactions turn out to be in this case:

$$\begin{aligned} V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} &= -\frac{i}{2\pi} (-2\beta k_x k_{1x} k_{2x})^{1/2} \left(\frac{k_y}{k^2 + k_0^2} + \frac{k_{1y}}{k_1^2 + k_0^2} + \frac{k_{2y}}{k_2^2 + k_0^2} \right); \\ U_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} &= -\frac{3i}{2\pi} (-2\beta k_x k_{1x} k_{2x})^{1/2} \frac{k_y}{k^2 + k_0^2}. \end{aligned} \quad (41.10)$$

Barotropic Rossby waves can enter into three-wave resonance (decay) interactions. This was first pointed out, and such interactions were calculated in the weak-nonlinearity approximation, by Kenyon (1964) for nondivergent waves and by Longuet-Higgins and Gill (1967) for the general case. The corresponding calculation is similar to the derivation of the equations of Blombergen (30.8), but we will present it here in more detail, following the indicated works. Using the small parameter ε for the nonlinear term in Equation (41.1), let us seek a solution of this equation in the form $\psi_0 + \varepsilon\psi_1 + \dots$, where all ψ_n are functions of x , t , and the 'slow' variables $\mathbf{X} = \varepsilon\mathbf{x}$, $T = \varepsilon t$. Then ψ_0 will satisfy the linear equation and it can be represented as the sum of three wave packets $\psi_0 = \sum_1^3 \psi_{k_n} \cos \theta_n$, where ψ_{k_n} depends arbitrarily on $\mathbf{X}$, T , while $\theta_n = \mathbf{k}_n \cdot \mathbf{x} - \sigma_{k_n} t$, and $\mathbf{k}_n$ satisfies resonance conditions $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3$, $\sigma_{k_1} + \sigma_{k_2} = \sigma_{k_3}$, so that $\theta_1 + \theta_2 = \theta_3$. For ψ_1 we then get the following equation:

$$\begin{aligned} \frac{\partial}{\partial t} (\Delta\psi_1 - k_0^2\psi_1) + \beta \frac{\partial\psi_1}{\partial x} &= -\frac{\partial(\psi_0, \Delta\psi_0)}{\partial(x, y)} - \left[\frac{\partial}{\partial T} (\Delta\psi_0 - k_0^2\psi_0) + \right. \\ &\quad \left. + \beta \frac{\partial\psi_0}{\partial X} + 2 \frac{\partial}{\partial t} \left(\frac{\partial^2\psi_0}{\partial x \partial X} + \frac{\partial^2\psi_0}{\partial y \partial Y} \right) \right] \\ &= -D_{\mathbf{k}_2\mathbf{k}_3} \psi_{k_2} \psi_{k_3} \cos(\theta_2 + \theta_3) + \\ &\quad + \left[D_{\mathbf{k}_2\mathbf{k}_3} \psi_{k_2} \psi_{k_3} + (k_1^2 + k_0^2) \left(\frac{\partial\psi_{k_1}}{\partial T} + \mathbf{c}_{g\mathbf{k}_1} \nabla \psi_{k_1} \right) \right] \times \\ &\quad \times \cos \theta_1 - D_{\mathbf{k}_1\mathbf{k}_3} \psi_{k_1} \psi_{k_3} \cos(\theta_3 + \theta_1) + \end{aligned}$$

$$\begin{aligned}
& + \left[D_{\mathbf{k}_3 \mathbf{k}_1} \psi_{\mathbf{k}_3} \psi_{\mathbf{k}_1} + (k_2^2 + k_0^2) \times \right. \\
& \times \left(\frac{\partial \psi_{\mathbf{k}_2}}{\partial T} + \mathbf{c}_{g \mathbf{k}_2} \cdot \nabla \psi_{\mathbf{k}_2} \right) \left. \right] \times \\
& \times \cos \theta_2 + D_{\mathbf{k}_1 \mathbf{k}_2} \psi_{\mathbf{k}_1} \psi_{\mathbf{k}_2} \cos(\theta_1 - \theta_2) + \\
& + \left[-D_{\mathbf{k}_1 \mathbf{k}_2} \psi_{\mathbf{k}_1} \psi_{\mathbf{k}_2} + (k_3^2 + k_0^2) \times \right. \\
& \times \left(\frac{\partial \psi_{\mathbf{k}_3}}{\partial T} + \mathbf{c}_{g \mathbf{k}_3} \cdot \nabla \psi_{\mathbf{k}_3} \right) \left. \right] \cos \theta_3, \quad (41.11)
\end{aligned}$$

where operator Δ corresponds to coordinates $\mathbf{x}$, and operator ∇ to coordinates $\mathbf{X}$; $\mathbf{c}_{g \mathbf{k}}$ is the group velocity of the waves (38.6), and the interaction coefficients are defined as

$$D_{\mathbf{k}_m \mathbf{k}_n} = \frac{1}{2} (k_m^2 - k_n^2) (k_{n_x} k_{m_y} - k_{m_x} k_{n_y}). \quad (41.12)$$

The terms in (41.11) with multipliers $\cos \theta_n$ are resonance terms, and in order to cancel these we have to set the multipliers of $\cos \theta_n$ in square brackets equal to zero. This yields equations for the amplitudes $\psi_{\mathbf{k}_n}$ of the wave packets from field ψ_0 . Having determined these, we can seek the solution ψ_1 of Equation (41.11) as the sum of the three corresponding wave packets, the amplitude equations of which are found by canceling out the resonance terms on the right-hand side of the equation for ψ_2 , etc. Consequently, the dependence of quantities ψ_n on the 'slow' variables $\mathbf{X}$, T allows us to cancel out the resonance terms (without which the series $\psi_0 + \varepsilon \psi_1 + \dots$ for times of the order of ε^{-1} would become inapplicable).

The above-indicated equations for the amplitudes $\psi_{\mathbf{k}_n}$ are quite complicated, but if these amplitudes do not depend on $\mathbf{X}$ at the initial time $T = 0$, then this will remain the case for other T as well. Then the equations for the amplitudes are simply:

$$\begin{aligned}
\frac{\partial \psi_{\mathbf{k}_1}}{\partial T} &= b \frac{k_3^2 - k_2^2}{k_1^2 + k_0^2} \psi_{\mathbf{k}_2} \psi_{\mathbf{k}_3}; \\
\frac{\partial \psi_{\mathbf{k}_2}}{\partial T} &= b \frac{k_1^2 - k_3^2}{k_2^2 + k_0^2} \psi_{\mathbf{k}_3} \psi_{\mathbf{k}_1}; \\
\frac{\partial \psi_{\mathbf{k}_3}}{\partial T} &= b \frac{k_2^2 - k_1^2}{k_3^2 + k_0^2} \psi_{\mathbf{k}_1} \psi_{\mathbf{k}_2}, \quad (41.13)
\end{aligned}$$

where

$$2b = k_{3x} k_{2y} - k_{3y} k_{2x} = k_{1x} k_{3y} - k_{1y} k_{3x} = k_{1x} k_{2y} - k_{1y} k_{2x}$$

is the area of the parallelogram constructed using vectors $\mathbf{k}_1$ and $\mathbf{k}_2$. These

equations are equivalent to the triplet (6.13), as is demonstrated by the conversion to the new unknowns $(k_n^2 + k_0^2)^{1/2} \psi_{k_n}$ and the new time

$$(k_1^2 + k_0^2)^{-1/2} (k_2^2 + k_0^2)^{-1/2} (k_3^2 + k_0^2)^{-1/2} bT,$$

and thus they are solved in Jacobi elliptic functions (6.15). Therefore, in the weak-nonlinearity approximation the decay interactions of a triad of packets of Rossby waves are found to be periodic (since the nonlinearity is weak, the period will be long, of the order of ε^{-1}). The above calculations are generalized for N wave packets with resonance conditions $\sigma_{k_m} + \sigma_{k_n} = \sigma_{k_m \pm k_n}$ (Kamenkovich and Reznik, 1978).

In the above works of Kenyon (1964) and Longuet-Higgins and Gill (1967) in the weak-nonlinearity approximation a kinetic equation for Rossby waves was derived that is analogous to Equations (24.12') and (30.10). Let ψ_k be the Fourier transform of a statistically homogeneous field $\psi(\mathbf{x})$ (which here is represented in the form $\varepsilon\psi_1 + \varepsilon^2\psi_2 + \dots$), let $\tau = \varepsilon^2 t$ be the 'slow' time, and let the spectral energy density F_k be given by the relation

$$\frac{1}{2} (k^2 + k_0^2) \langle \psi_k(t) \psi_{k_1}^*(t) \rangle = F_k(\tau) \delta_{k-k_1}. \quad (41.14)$$

In a manner similar to that described for Equation (30.10), we derive the following equation for F_k :

$$\begin{aligned} \frac{\partial F_k}{\partial \tau} = 8\pi \int & \frac{D_{k_1 k_2} F_{k_1} F_{k_2} + D_{k k_1} F_k F_{k_1} + D_{k k_2} F_k F_{k_2}}{(k^2 + k_0^2) (k_1^2 + k_0^2) (k_2^2 + k_0^2)} \times \\ & \times \delta_{\mathbf{k}+\mathbf{k}_1+\mathbf{k}_2} \delta_{\sigma_k+\sigma_{k_1}+\sigma_{k_2}} D_{\mathbf{k}_1 \mathbf{k}_2} d\mathbf{k}_1 d\mathbf{k}_2. \end{aligned} \quad (41.15)$$

It is easy to see that if $\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2 = 0$, then for any a and b it will be true that

$$(ak^2 + b) D_{\mathbf{k}_1 \mathbf{k}_2} + (ak_1^2 + b) D_{\mathbf{k} \mathbf{k}_2} + (ak_2^2 + b) D_{\mathbf{k} \mathbf{k}_1} = 0.$$

This implies that equation (41.15) conserves the total energy $\int F_k d\mathbf{k}$ and the potential enstrophy $\int (k^2 + k_0^2) F_k d\mathbf{k}$ (however, these are conserved not just in the weak-nonlinearity approximation, but also according to the exact equation (41.1)). From the conservation of enstrophy it follows that $F_k = O(k^{-4})$ as $k \rightarrow \infty$. Conservation of enstrophy is equivalent to conservation of the zonal component of the momentum $\int k_x \sigma_k^{-1} d\mathbf{k}$. If the resonance condition $\sigma_k + \sigma_{k_1} + \sigma_{k_2} = 0$ is still satisfied, then

$$k_y \sigma_k^{-1} D_{\mathbf{k}_1 \mathbf{k}_2} + k_{1y} \sigma_{k_1}^{-1} D_{\mathbf{k} \mathbf{k}_2} + k_{2y} \sigma_{k_2}^{-1} D_{\mathbf{k} \mathbf{k}_1} = 0,$$

which suggests that Equation (41.15) also conserves the meridional component of the momentum $\int k_y \sigma_k^{-1} F_k d\mathbf{k}$.

The integration in the kinetic equation (41.15) is carried out for each fixed $\mathbf{k}$, actually according to the so-called resonance line $\mathcal{L}_k$ in the four-dimensional space $(\mathbf{k}_1, \mathbf{k}_2)$, on which the three conditions $\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2 = \sigma_k + \sigma_{k_1} + \sigma_{k_2} = 0$ are satisfied. However, a careful derivation of Equation (41.5) reveals (Reznik, 1984) that it should be valid for all $\mathbf{k}$ only provided that all the values $|\mathbf{c}_{g\mathbf{k}_1} - \mathbf{c}_{g\mathbf{k}_2}|$ on all $\mathcal{L}_k$ be bounded uniformly from below [have the same lower limit]. But for Rossby

waves $|\mathbf{c}_{gk}| \rightarrow 0$ as $k \rightarrow \infty$. Therefore, the indicated condition is violated for very short waves $\mathbf{k}_1$ and $\mathbf{k}_2$ which are in resonance with wave $\mathbf{k}$.

If in this case k_x is small, that is, the resonance interaction occurs between two short waves and one long wave, then it reduces essentially to an interaction between the short waves, there being almost no effect on the long wave (which serves, therefore, just as a 'catalyst'). The kinetic equation remains valid in this case.

If, on the other hand, all three waves $\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2$ are very short (so that $k, k_1, k_2 \geq \varepsilon^{-2}$), then the resonance interactions between them will not differ from any of the nonresonance interactions satisfying the condition $\sigma_k + \sigma_{k_1} + \sigma_{k_2} \sim \varepsilon^2$. The approximation of weak nonlinearity is inapplicable here, so that for $k \sim \varepsilon^{-2}$ the kinetic equation does not hold true.

However, very short waves exchange energy mainly just with each other, and they make only a very small contribution to the total energy of the developed wave field. Therefore, for $k \ll \varepsilon^{-2}$ the kinetic energy can be used, provided we just integrate over $\mathbf{k}_1, \mathbf{k}_2$ a finite region $\Omega_{k_1} \times \Omega_{k_2}$ with k_1, k_2 values that are not too high (but still such that outside Ω_k the condition $F_k = O(k^{-4})$ is already satisfied). Such a 'cut off' kinetic equation conserves the 'cut off' values of the energy, enstrophy, and momentum of the wave field. Below we will deal only with this equation.

Using the following equations, valid on $\mathcal{L}_k$:

$$\sigma_k^{-1} D_{k_1 k_2} = \sigma_{k_1}^{-1} D_{k k_2} = \sigma_{k_2}^{-1} D_{k k_1},$$

we see with the aid of the kinetic equation that the rate of variation in time of the entropy of the wave field $\int \ln F_k d\mathbf{k}$ is equal to some weighted integral of the square of the quantity

$$A = \sigma_k F_k^{-1} + \sigma_{k_1} F_{k_1}^{-1} + \sigma_{k_2} F_{k_2}^{-1}.$$

Therefore, the thermodynamically stable spectrum is the solution of the equation $A = 0$ on $\mathcal{L}_k$. The sole *differentiable* solution is, just as for statistically stable two-dimensional turbulence, the isotropic spectrum $(a + bk^2)^{-1}$, where a and b are constants ('temperatures'); this spectrum ensues from the conservation of energy and enstrophy.

However, in contrast to two-dimensional turbulence, a random field of Rossby waves becomes anisotropic with time (see Section 37). Therefore, it is important to remember that the singular spectrum $\Phi(k_y)\delta(k_x)$ of an arbitrary stationary zonal flow also serves as a steady-state solution of the kinetic equation. Note that the sum of this solution plus any solution $F_k(t)$ of the kinetic equation that is symmetric with respect to k_x also satisfies this equation; if, however, the solution has the form $\Phi(k_y)\delta(k_x) + F_k(t)$, where the added term is asymmetric, then it will become symmetric with time (to a certain degree, as prescribed by the conservation of the meridional component of the momentum).

Finally, it can be shown that a steady-state solution $\Phi(k_y)\delta(k_x) + (a + bk^2)^{-1}$ is statistically in equilibrium. Of course, such a solution is, generally speaking, unattainable, since according to the kinetic equation the spectrum does not vary in time on the k_y axis. However, the tendency of solutions of the kinetic equation for

different initial data to evolve into such a statistical equilibrium form has been verified by several numerical experiments (Reznik *et al.*, 1981–1984).

Consequently, *anisotropization* of the wave field occurs, namely concentration of a considerable part (but not 100%) of its energy around the k_y axis, that is, in the zonal current (the phenomenon of *negative viscosity*, produced by the β -effect and the nonlinearity), with a simultaneous reduction of the entropy of the field in this region of the spectrum. In addition, there will be an *isotropization* of the field far away from the k_y axis, that is, a uniform distribution of the remaining part of the energy among all the directions of the wave vectors with a wave-number spectrum $(a + bk^2)^{-1}$ in the region $|k_x| > \delta > 0$ (which, moreover, governs the evolution of the total entropy of the field).

A generalization for baroclinic Rossby waves was proposed by Monin and Piterbarg (1987). Instead of (41.1) they used Equation (37.5), and instead of (41.2) they introduced the following Hamiltonian:

$$H = \frac{1}{2} \int \left[|\nabla \psi|^2 + \frac{f_0^2}{N^2} \left(\frac{\partial \psi}{\partial z} \right)^2 \right] dx dy dz \quad (41.2')$$

Formulas (41.3) remain valid, while in (41.4)–(41.4') and (41.5)–(41.5') the argument z must be added for functions Ω and ζ , so that these functions are now represented in the following manner:

$$\zeta(x, y, z) = \frac{1}{2\pi} \sum_{\nu} \varphi_{\nu}(z) \int_{k_x < 0} (\zeta_{\nu k} e^{ik \cdot x} + \zeta_{\nu k}^* e^{-ik \cdot x}) dk \quad (41.9')$$

where $\{\varphi_{\nu}(z)\}$ is an orthonormalized system of eigenfunctions of operator $L = \mathcal{L} - \Delta$, so that $L\varphi_{\nu} = -\lambda_{\nu}^2 \varphi_{\nu}$, where $\varphi_{\nu}(0) = \varphi_{\nu}(h) = 0$ and $(1/h) \int_0^h \varphi_{\mu} \varphi_{\nu} (f_0^2/N^2) dz = \delta_{\mu\nu}$. Using this notation, we can write Hamiltonian (41.2') as

$$H = \frac{1}{2} \sum_{\nu} \int (\lambda_{\nu}^2 + k^2) |\psi_{\nu k}|^2 dk = \frac{1}{2} \sum_{\nu} \int \frac{|\Omega_{\nu k}|^2 dk}{\lambda_{\nu}^2 + k^2}, \quad (41.16)$$

and the normal canonical variables are introduced with the aid of the formula $a_{\nu k} = (-2\beta k_x)^{-1/2} \zeta_{\nu k}$. Since three-wave resonance interactions are possible for Rossby waves, in the functional power series (24.4) (summed over ν) it is sufficient to calculate the quadratic and cubic terms. Then, instead of (41.4), we can invoke the following approximation relation between Ω and ζ , valid with an accuracy of the order of β^{-2} :

$$\begin{aligned} \Omega(x, y, z) &= \zeta(x, y, z) + \frac{1}{2\beta} \frac{\partial}{\partial y} \zeta^2(x, y, z); \\ \Omega_{\nu k} &= \zeta_{\nu k} + \frac{i}{4\pi\beta} \sum_{\nu_1, \nu_2} \int \alpha_{\nu\nu_1\nu_2} k_y \zeta_{\nu_1 k_1} \zeta_{\nu_2 k_2} \delta_{k-k_1-k_2} dk_1 dk_2; \\ \alpha_{\nu\nu_1\nu_2} &= \frac{1}{h} \int_c^h \varphi_{\nu} \varphi_{\nu_1} \varphi_{\nu_2} \left(\frac{f_0^2}{N^2} \right) dz. \end{aligned} \quad (41.17)$$

Then in series (24.4) (with doubled sums over all ν and integrations over regions with negative k_x) the coefficients become

$$\begin{aligned}\sigma_{\nu k} &= -\beta k_x (\lambda_\nu^2 + k^2)^{-1}; \\ V_{kk_1 k_2}^{\nu \nu_1 \nu_2} &= \frac{i}{4\pi} (-2\beta k_x k_{1x} k_{2x})^{1/2} (S_{kk_1 k_2}^{\nu \nu_1 \nu_2} - S_{k_2 - k_1 k}^{\nu_2 \nu_1 \nu} - S_{k_1 k - k_2}^{\nu_1 \nu \nu_2}); \\ U_{kk_1 k_2}^{\nu \nu_1 \nu_2} &= -\frac{i}{4\pi} (-2\beta k_x k_{1x} k_{2x})^{1/2} S_{kk_1 k_2}^{\nu \nu_1 \nu_2}, \quad S_{kk_1 k_2}^{\nu \nu_1 \nu_2} = \frac{k_y \alpha_{\nu \nu_1 \nu_2}}{\lambda_\nu^2 + k^2}\end{aligned}\quad (41.18)$$

The kinetic equation for the wave-action spectrum $N_{\nu k}$ here has a form similar to (24.12'), except with summation over ν_1, ν_2 , integration over regions with negative k_{1x}, k_{2x} , and coefficients $T_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} = 4\pi |V_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}^{\nu \nu_1 \nu_2}|^2$. In particular, for $N^2 = \text{const}$, we get $\lambda_\nu = \pi \nu f_0 / Nh$ and $\varphi_\nu(z) = \sin(\pi \nu z / h)$, while the coefficients of (41.17) become

$$\begin{aligned}\alpha_{\nu \nu_1 \nu_2} &= \frac{f_0^2}{N^2} [(-1)^{\nu + \nu_1 + \nu_2} - 1] \nu \nu_1 \nu_2 (\nu + \nu_1 + \nu_2)^{-1} \times \\ &\quad \times (\nu_2 - \nu + \nu_1)^{-1} (\nu_2 + \nu - \nu_1)^{-1} (\nu_2 - \nu - \nu_1)^{-1}\end{aligned}\quad (41.19)$$

In the case of barotropic Rossby waves this expression is replaced by unity, and λ_ν is replaced by the constant k_0 .

42. GENERATION AND DISSIPATION

Eddies and Rossby waves are generated mainly as a result of the baroclinic instability of large-scale currents, at the expense of their available potential energy (see Section 12, where this process was examined, using as an example the baroclinic instability of zonal currents in the Earth's atmosphere).

In addition, synoptic eddies and waves (of scale L) can be generated in a flow around irregularities h' of the bottom relief; their importance becomes comparable to the β -effect for $h'/h \sim L/r_0$, and the critical value $h' = hL/r_0$ in the World Ocean is found to be some tenths of that in the atmosphere, which produces a quite large abundance in it of topographic eddies and Rossby waves. Finally, for the ocean the possibility of generating synoptic processes as a result of direct atmospheric factors is typical.

From formula (12.10) for the energy of the geostrophic currents it follows that the ratio Π/K of the available potential energy to the kinetic energy is of the order of $(L/L_R)^2$. In the ocean $L_R \sim 50$ km, so that for large-scale currents with $L \sim 1000$ km we get $\Pi \sim 400K$ (of the order of 7×10^2 J/m² according to the empirical estimate of Vulis and Monin, 1975) and for synoptic currents with $L \sim L_R$ we have $\Pi \sim K$ (of the order of 20 J/m² according to the empirical estimate of Wyrte *et al.*, 1976). Therefore, the total energy of the synoptic eddies in the ocean amounts to about 10% of the available potential energy of the large-scale currents, and the generation of eddies does not vary much the large-scale inclinations of the isentropic surfaces.

With the aid of the fluid-dynamical equations for synoptic disturbances of the zonal current (12.15), and the equation for their energy (12.16), it can be shown that in baroclinically unstable disturbances the fluid particles must move in a

meridional plane with an inclination to the horizontal $w'/v' \sim (Uh/L)/f_0L$, bounded by the inclination of the isentropes f_0U/N^2h , which leads to the inequality $(L_R/L)^2 \lesssim 1$. This means that the scales L of the baroclinically unstable disturbances cannot be substantially smaller than L_R .

In Section 12 we established the *sufficient* conditions for the baroclinic *stability* of a zonal current $U(y, p)$, and in the simplest case $U = U(p)$ we demonstrated the theorem of Dikii stating that for the conditions of Rayleigh or Fj\o rtoft and fixed values of the parameters there cannot be more than one baroclinically unstable synoptic disturbance. Here we add to these results the *necessary* conditions for the baroclinic *instability* of zonal currents, using, as almost everywhere in this chapter, the coordinates (x, y, z) and limiting ourselves, to keep the formulas simple, to the boundary condition of a 'rigid cover' at the ocean surface, but considering a more general form of the zonal current $U = U(y, z)$ and the synoptic disturbances $\psi = \psi(y, z) \exp[ik_1(x - ct)]$. For their amplitudes $\psi(y, z)$, instead of (12.17), (12.18) we obtain the equations:

$$(U - c) \left(\frac{\partial^2 \psi}{\partial y^2} + \frac{\partial}{\partial z} \frac{f_0^2}{N^2} \frac{\partial \psi}{\partial z} - k_1^2 \psi \right) + B\psi = 0; \quad (42.1)$$

$$B = \beta - \frac{\partial^2 U}{\partial y^2} - \frac{\partial}{\partial z} \frac{f_0^2}{N^2} \frac{\partial U}{\partial z};$$

$$(U - c) \frac{\partial \psi}{\partial z} - \frac{\partial U}{\partial z} \psi = 0 \quad \text{for } z = 0, -h, \quad (42.2)$$

where B is the meridional gradient of the quasigeostrophic potential vorticity of the zonal current. Assume that for these equations a nonzero solution with a value $\text{Im } c > 0$ exists. If we divide (42.1) by $U - c$, multiply by ψ^* , integrate over y and z , and then separate the real and imaginary parts, we obtain

$$\int \frac{B |\psi|^2 dy dz}{|U - c|^2} + \int \left[\frac{f_0^2}{N^2} \frac{\partial U}{\partial z} \frac{|\psi|^2}{|U - c|^2} \right]_{z=-h}^{z=0} dy = 0; \quad (42.3)$$

$$\begin{aligned} & \int \frac{UB |\psi|^2 dy dz}{|U - c|^2} + \int \left[\frac{f_0^2}{N^2} U \frac{\partial U}{\partial z} \frac{|\psi|^2}{|U - c|^2} \right]_{z=-h}^{z=0} dy \\ &= \int \left(\left| \frac{\partial \psi}{\partial y} \right|^2 + \frac{f_0^2}{N^2} \left| \frac{\partial \psi}{\partial z} \right|^2 + k_1^2 |\psi|^2 \right) dy dz. \end{aligned} \quad (42.4)$$

The first of these equations shows either that B changes sign somewhere or that on some part of the bottom $(\partial U / \partial z)_{-h}$ has the same sign as B or that on some part of the ocean surface the sign of $(\partial U / \partial z)_0$ differs from that of B . According to (42.4), either in some region $UB > 0$ or on some part of the bottom $(U \partial U / \partial z)_{-h} < 0$ or on some part of the surface $(U \partial U / \partial z)_0 > 0$. If none of these six conditions is satisfied, then the zonal current is stable (but the opposite

state of instability is not yet guaranteed). We also note the inequality

$$\begin{aligned}
 (2k_1 \operatorname{Im} c)^2 &\leq k_1^2 (k_1^2 + R_y^{-2} + R_z^{-2})^{-1} \left[\left(\frac{\partial U}{\partial y} \right)^2 + \frac{f_0^2}{N^2} \left(\frac{\partial U}{\partial z} \right)^2 \right]; \\
 R_y^{-2} &= R^{-2} \int \left| \frac{\partial \chi}{\partial y} \right|^2 dy dz; \\
 R_z^{-2} &= R^{-2} \int \frac{f_0^2}{N^2} \left| \frac{\partial \chi}{\partial z} \right|^2 dy dz; \\
 R^2 &= \int |\chi|^2 dy dz; \chi = \psi(U - c)^{-1/2}
 \end{aligned} \tag{42.5}$$

(Pedlosky, 1979), limiting the growth of unstable disturbances the more, the lower the value of k_1 . In particular, for baroclinic instability the scale L of the increasing disturbances cannot be significantly lower than L_R , but according to (42.5) with increasing $L \sim k_1^{-1}$ the growth of the disturbances slows down, so that it is to be expected that for developed disturbances $L \sim L_R$.

A specific calculation of the eigenvalues c of Equations (42.1) and (42.2) is, as usual, not simple. In the two-layer model the necessary and sufficient conditions for *instability* take the form: $U = U_1 - U_2 > \beta g' h_2 f_0^{-2}$ for $U > 0$ and $U < -\beta g' h_2 f_0^{-2}$ for $U < 0$, where $g' = g(\rho_2 - \rho_1)/\rho_1$, the subscript 1 pertaining to the upper layer and the subscript 2 to the lower layer (Pedlosky, 1979).

Topographic Rossby waves were considered in Section 39. Of these, we should mention here those which can appear only because of the bottom relief, such as, for example, the zero topographic mode with frequency (39.3) for $\gamma < 0$. We note, too, that waves over an inclined bottom with dispersion relation (39.8) and trapped modes over a wavy bottom with dispersion relation (39.16) can also exist in the absence of a β -effect.

Huppert and Bryan (1976) considered eddies formed during flow at a velocity U around an isolated underwater protuberance with a height h_m , with conservation in the fluid particles of the quasigeostrophic potential vorticity

$$\omega_* = \omega - \frac{gf_0}{\rho_0 N^2} \frac{\partial \rho'}{\partial z}; \quad \omega = \Delta \psi \tag{42.6}$$

(taken for simplicity at $\beta = 0$ and $N = \text{const}$). For particles lifted up onto the protuberance by the upwelling of cold water, the derivative $(-\partial \rho'/\partial z)$ increases and ω decreases. On the other hand, for particles descending from the protuberance, ω increases. As a consequence, above the protuberance a cold anticyclonic eddy will form, while past it on the downstream side a warm cyclonic eddy appears. Calculations indicate that the anticyclonic eddy always remains 'attached' to the protuberance, whereas for small $Nh_m/U \sim 10^0$ the cyclonic eddy moves downstream (the eddies in this case are stronger), and only for large

$Nh_m/U \sim 10^1$ does it remain 'attached' (the eddies being less strong in this case; this situation is apparently typical for the real ocean).

Of the direct atmospheric effects on the ocean surface capable of generating synoptic motions in the ocean, we should mention the frictional wind stress and the atmospheric pressure, as well as the buoyancy flux M_s , to be discussed in Section 46. Empirical data on the space-time spectra of these quantities in the range of the scales of oceanic synoptic motions were studied by Magaard (1977), Willebrand (1978), and in particular detail by Frankignoul and Müller (1979) (who also formulated the theory of the reaction of the ocean to atmospheric effects, to be presented below). These data revealed that the frictional wind stress alone plays a major role (the atmospheric pressure becomes important only at considerably high frequencies and wavelengths, while the role of the buoyancy flux is even much less).

In the middle latitudes the spectra of the atmospheric effects have maxima in the period range of a few days and the wavelength range of 3000 to 6000 km, typical of atmospheric synoptic processes, but they also contain a significant amount of energy at lower frequencies and larger wave numbers. It was found that at periods longer than 10 or 20 days the time spectra become slightly frequency-dependent, while for wavelengths shorter than 3000 km the space spectra are approximately isotropic. For the space-time spectrum of the vorticity of the frictional wind stress, the following empirical model is suitable for low frequencies:

$$F_k = F_0 k^2 S_k; \quad F_0 = 10^4 H^2 (\text{m}^4 \text{Hz})^{-1};$$

$$2\pi k S_k = \begin{cases} \frac{3}{4} k_0^{-3} k^2, & 0 \leq k \leq k_0 = 2\pi/5000 \text{ km}; \\ \frac{3}{4} k_0 k^{-2}, & k_0 \leq k \leq k_1 \gg k_0; \\ 0, & k > k_1, \end{cases} \quad (42.7)$$

where S_k is the space spectrum normalized to unity (its actual anisotropy for $k < k_0$ is of little importance, since in the region $k < k_0$ atmospheric disturbances move mainly eastward and do not induce a resonance reaction on the ocean; also insignificant is the fact that the actual spectra for $k > k_0$ drop a little more steeply than indicated according to this model). Because of the symmetry $F_{k\sigma} = F_{-k, -\sigma}$, here we also take into account westward-moving atmospheric disturbances, which may cause the resonance excitation of Rossby waves in the ocean.

The reaction of the ocean to atmospheric effects can be calculated with the aid of the linearized equation for the waves (38.1), introducing into the first of boundary conditions (38.2) the right-hand side $Q/g = -\text{curl } \tau / \rho_0 f_0^2$, and in the second equation setting $h = \text{const}$. The solution of this equation can be written as

$$\psi = \sum_{n=0}^{\infty} \int \psi_{nk}(t) \Psi_n(z) e^{ik \cdot x} dk, \quad (42.8)$$

where $\Psi_n(z)$ are the orthonormalized eigenfunctions of Equation (38.4), and amplitudes ψ_{nk} are found from the equations

$$\frac{\partial \psi_{nk}}{\partial t} + i\sigma_{nk} \psi_{nk} = \frac{Q_{nk}}{1 + k^2 L_n^2} = \frac{s_n Q_k}{1 + k^2 L_n^2}, \quad (42.9)$$

where $Q_k(t)$ is the Fourier transform in $\mathbf{x}$ of function Q (here Q_{nk} are the coefficients of the expansion of Q in form (42.8)), $L_n^2 = gh_n f_0^{-2}$ are the normalized eigenvalues of Equation (38.4), and $s_n = \int_{-h}^0 \Psi_n dz = h_n \Psi_n(0)$. Below we will need only the following space spectrum of field ψ , obtained from (42.9) asymptotically as $t \rightarrow \infty$:

$$\langle \psi_{nk}(t) \psi_{nk_1}^*(t) \rangle = \frac{2\pi t s_n^2}{(1 + k^2 L_n^2)^2} \Phi_{k\sigma_n} \delta_{\mathbf{k}-\mathbf{k}_1}, \quad (42.10)$$

where $\Phi_{k\sigma}$ is derived from spectrum $F_{k\sigma}$ of formula (42.7) by multiplying it by $g^2/\rho_0^2 f_0^4$. The proportionality of the time t is here related to the stochasticity of the resonance excitation. The total energy of a water column of unit cross section now will be

$$\begin{aligned} \langle E(t) \rangle &= \frac{1}{2} \rho_0 \int_{-h}^0 \left\langle \left| \nabla \psi \right|^2 + \frac{f_0^2}{N^2} \left| \frac{\partial \psi}{\partial z} \right|^2 + \frac{f_0^2}{gh} (\psi^2)_{z=0} \right\rangle dz \\ &= \int d\mathbf{k} \int_{-h}^0 E_{\mathbf{k}}(z, t) dz; \quad E_{\mathbf{k}}(z, t) = \sum_{n=0}^{\infty} E_{n\mathbf{k}}(z, t); \end{aligned} \quad (42.11)$$

$$E_{n\mathbf{k}}(z, t) = \frac{1}{2} \rho_0 \Psi_n^2(z) \int \langle \psi_{n\mathbf{k}}(t) \psi_{n\mathbf{k}}^*(t) \rangle (\mathbf{k} \cdot \mathbf{k}_1 + L_n^{-2}) d\mathbf{k}_1,$$

where $E_{\mathbf{k}}(z, t)$ is the spectral density of the total energy. The energy flux to a wave number $\mathbf{k}$ at a depth z , produced by the atmospheric effect, is found to be asymptotically equal to

$$\frac{\partial E_{\mathbf{k}}(z, t)}{\partial t} = \pi \rho_0 \sum_{n=0}^{\infty} \frac{s_n^2 L_n^2 \Psi_n^2(z) \Phi_{k\sigma_n}}{1 + k^2 L_n^2}. \quad (42.12)$$

In the n th mode it is proportional to $\Psi_n^2(z)$ and thus depends considerably on the profile $N(z)$; for $n = 1$ the energy is concentrated mainly in the upper layer of the ocean, but with increasing n it becomes distributed more and more widely over the depth. The energy flux $\partial E_{\mathbf{k}}/\partial t$, integrated over the depth (the integral of Ψ_n^2 is equal to unity), is less sensitive to details of the stratification; the analogous flux of kinetic energy $K_{\mathbf{k}}$ has the form

$$\frac{\partial K_{\mathbf{k}}}{\partial t} = \pi \rho_0 \sum_{n=0}^{\infty} \frac{s_n^2 k^2 \Phi_{k\sigma_n}}{(1 + k^2 L_n^2)^2}. \quad (42.13)$$

The stratification influences these fluxes via L_n and s_n ; if $N(z)$ is replaced by $N = \text{const}$, then L_n varies only a little, whereas for $n > 1$ we find that s_n varies significantly. The terms with $n = 0$ constitute the barotropic parts of the energy fluxes (these are practically independent of the stratification), while the sums of the other terms are the baroclinic parts. For $N = \text{const}$ and taking into account that in

the frequency interval considered $\Phi_{k\sigma} \approx \Phi_{k0}$, for the barotropic parts we have

$$\frac{\partial E_k^{bt}}{\partial t} = \frac{\pi \rho_0 f_0^4 \Phi_{k0}}{g^2 h (k^2 + L_0^{-2})}; \quad \frac{\partial K_k^{bt}}{\partial t} = \frac{\pi \rho_0 f_0^4 k^2 \Phi_{k0}}{g^2 h (k^2 + L_0^{-2})^2}. \quad (42.14)$$

Summation of the series for the baroclinic parts in this case then yields

$$\begin{aligned} \frac{\partial E_k^{bc}}{\partial t} &= \frac{\pi \rho_0 f_0^4}{g^2 h k^2} (k L_R \operatorname{cth} k L_R - 1) \Phi_{k0}; \\ \frac{\partial K_k^{bc}}{\partial t} &= \frac{\pi \rho_0 f_0^4 L_R^2}{2 g^2 h} \left(\operatorname{cth}^2 k L_R + \frac{\operatorname{cth} k L_R}{k L_R} - 1 - \frac{1}{k^2 L_R^2} \right) \Phi_{k0}. \end{aligned} \quad (42.15)$$

The barotropic transfer functions are concentrated mainly around $k = L_0^{-1}$, and they go to zero as k^{-2} for $k \rightarrow \infty$. The baroclinic functions are not very concentrated and they go to zero as k^{-1} . The barotropic fluxes (42.14), integrated over $\mathbf{k}$ in the wavelength interval from 4000 to 50 km, are found to be 2.90×10^{-4} and 2.85×10^{-4} W/m², that is, almost all the flux goes to increase the kinetic energy (predominantly of long waves; these numbers are not very sensitive, with regard to the form of spectrum S_k). The analogous baroclinic fluxes are 1.5×10^{-4} and 0.5×10^{-4} W/m², that is, here the main portion of the flux goes to increase the available potential energy (these fluxes and their wavelength distribution are very sensitive to variations of the form of spectrum S_k).

The total flux of 0.44×10^{-3} W/m² is comparable to the estimate, averaged over space, of Gill *et al.* (1974) of the flux from a baroclinic instability, but in the jet-stream regions the latter flux is much greater. Therefore, the generation by atmospheric effects is considerable outside of the jet-stream region, and there a resonance atmospheric effect is responsible mainly for the barotropic synoptic disturbances. For nonresonance effects, instead of (42.10) we obtain

$$\langle \psi_{m\mathbf{k}\sigma} \psi_{n\mathbf{k}_1\sigma_1}^* \rangle = \frac{s_m s_n \Phi_{\mathbf{k}\sigma} \delta_{\mathbf{k}-\mathbf{k}_1} \delta_{\sigma-\sigma_1}}{(\sigma - \sigma_{m\mathbf{k}})(\sigma - \sigma_{n\mathbf{k}})(1 + k^2 L_m^2)(1 + k^2 L_n^2)}, \quad (42.16)$$

and for $\sigma \gg \sigma_m, \sigma_n$ this spectrum varies as σ^{-2} , which is verified by certain empirical data pertaining to the central part of the northern half of the Pacific Ocean.

In the high latitudes level F_0 of spectrum (42.7) turns out to be three times as high, and this spectrum decays more rapidly for large k . This doubles the barotropic energy fluxes and reduces by more than a factor of ten the baroclinic energy fluxes. In the tropics a substantial role is probably played by the effects of hurricanes and typhoons.

If to kinetic equation (41.15) we now add a term describing the viscous dissipation of kinetic energy, we see that its effect is considerable only for large k . Although this term is isotropic, its effect is anisotropic: according to the above-mentioned numerical experiments of Reznik and his coworkers, dissipation suppresses short waves, the direction of propagation of which is quite far from meridional; waves of meridional direction, on the other hand, are almost not suppressed at all (they are even sometimes strengthened a little, compared to the

calculation without viscosity), thereby introducing an additional contribution to the anisotropy of the wave field.

This is also verified by the configuration of the regions in the $\mathbf{k}$ plane with an energy influx $\partial F_{\mathbf{k}}/\partial\tau > 0$ and an energy efflux $\partial F_{\mathbf{k}}/\partial\tau < 0$. In the presence of dissipation the region of influx is concentrated around the k_y axis and it narrows with time, whereas outside of this region the total energy decreases with time (the total energy of the entire field in fact decreases). In the absence of dissipation an additional region of influx appears far from the k_y axis, leading there to isotropization of the field, because of its proximity to statistical equilibrium in this region.

If steady-state sources and sinks of energy are present, and if these are dispersed over the spectrum, then according to Pelinovskii (1978) for the inertial interval the isotropic spectra exist $F_{\mathbf{k}} \sim k^{-5/2}$, k^{-3} , and $k^{-7/2}$, corresponding to constant spectral fluxes of energy, enstrophy, and momentum. More detailed predictions, taking anisotropy into account, were made by Sazontov (1981):

$$F_{\mathbf{k}} \sim k_x^{-3/2} k_y^{-2}, \quad k_x^{-3/2} k_y^{-3} \quad \text{and} \quad k_x^{-1} k_y^{-7/2}$$

(the last spectrum corresponds to a constant spectral flux of the meridional component of the momentum).

PART III

GLOBAL PROBLEMS

- 8. GENERAL CIRCULATION OF THE ATMOSPHERE AND OCEAN
- 9. THEORY OF CLIMATE
- 10. FLUID DYNAMICS OF PLANETARY INTERIORS

General Circulation of the Atmosphere and Ocean

43. GENERAL CIRCULATION OF THE ATMOSPHERE

The word *circulation* here simply refers to motion (sometimes it is used in a narrower sense, to mean motions around closed trajectories, or ‘circulation wheels’). It is convenient to divide atmospheric motions into *small-scale motions*, with scales L much smaller than the effective thickness H of the atmosphere (defined, for example, by formula (2.6)), *meso-scale motions*, with L of the order of H or several times larger than H , and *large-scale motions*, with $L \gg H$ (Kolesnikova and Monin, 1965). The latter include global circulations, of the zonal and monsoon types, as well as synoptic processes, that is, eddies and Rossby waves, barotropic with scales of the order of $L_0 = c_0/f$, where $c_0 = (gH)^{1/2}$, and baroclinic with scales of the order of $L_R = HN/f$. The statistical ensemble of large-scale atmospheric motions constitutes the general circulation of the atmosphere.

The atmospheric circulation on Earth, Mars, and Venus originates because of the baroclinicity of the atmospheric gases (the dependence of ρ on T of type (1.7)), due to the spatially inhomogeneous heating of these gases by the long-wave radiation of the underlying surface and by the direct short-wave radiation of the Sun. Air expands when heated, and air masses rise, so that at a given height the pressure increases. This pressure rise at elevations in warm regions forces air at high levels to flow into cool regions. At lower levels there will then be a compensating downflow of air into warmer regions. This circulation between the equatorial zone and the middle latitudes on the Earth is known as the *trade winds* (a representation first formulated by Hadley in 1735; subsequent investigators introduced additional ‘wheels’ of the meridional circulation).

On Mars and Earth the Coriolis force turns the flow of air from the equator to the poles at high levels toward the east, producing a westerly transport of air in the upper atmosphere of the middle latitudes (and a westward component of the surface trade winds in the tropics). Meteorological measurements have shown, however, that the ‘wheels’ of the meridional circulation are too weak to account for the very strong zonal circulation observable in the Earth’s atmosphere (including the subtropical jet streams at latitudes of around $\pm 35^\circ$ and heights of about 12 km, which have an effective width of 300 to 400 km, a thickness of 1 to 2 km, and a velocity of 60 to 80 m/s, although a speed as high as 190 m/s has been recorded).

A quantitative explanation of the zonal circulation is obtained only if the role of synoptic processes is taken into account. The first to do this was Defant (1921), who interpreted the synoptic eddies as being a kind of large-scale turbulence (see Section 37), causing meridional heat transfer (Defant estimated the effective transfer coefficient to be $\rho\nu \sim 10^8 \text{ g/(cm s)}$). Jeffreys (1926) suggested that this also makes a major contribution to the meridional fluxes of angular momentum.

One of the first to offer a systematic fluid-dynamical description of the general circulation of the atmosphere was Kochin (1936). He assigned a major role to the forces of eddy viscosity, which he took to be comparable to the Coriolis force (rather than to the inertial forces, as in boundary-layer theory). Moreover, Kochin showed that the appropriate simplified equations of fluid dynamics can be used to calculate the three-dimensional fields of the velocity and density, in terms of the specified field of the surface pressure p_s and the specified temperature field T . The next step was taken by Blinova (1947), who appended to the equations of Kochin the equation of heat influx and, taking the continents and oceans into account, calculated the mean annual three-dimensional field of T and p_s (in which, in particular, all the so-called *centers of action of the atmosphere* were obtained theoretically for the first time, these being quasistationary subtropical anticyclones and subpolar cyclones); later she also computed their annual variation.

Any complete theory of the general circulation must include the feedback between the zonal circulation and synoptic processes. On the one hand, the zonal circulation is baroclinically unstable, with respect to small nonzonal disturbances (see Dikii's theorem in Section 12), which grow at the expense of the available potential energy to become synoptic eddies (Bjerknes, 1937). On the other hand, the statistical ensemble of synoptic eddies may transfer angular momentum to the latitude zones with jet streams, that is, it may act like a 'negative viscosity' and thus transfer its kinetic energy to the zonal circulation (Rossby, 1941, 1947).

The unusual effect of a 'negative viscosity' can be explained as follows: the axes of the crests of the Rossby waves, which are oriented from the tropics toward the subtropical jet stream (toward the north in the Northern Hemisphere), tend to slope predominantly toward the current (they point toward the northeast). Thus angular momentum ahead of the Rossby troughs, where it is higher (since the flow is close to zonal), is transferred to the jet stream (northward), whereas behind the troughs, where it is lower (flow close to meridional), angular momentum leaves the stream (southward). The total effect over the entire wave amounts to a transfer of angular momentum to the jet stream.

Following Monin and Seidov (1982), we will present a simplified qualitative model describing the enhancement of a jet stream by negative viscosity in a fluid layer of constant depth that is uniform along the vertical. The density ρ is assumed to be a function only of the temperature T . The origin of the horizontal coordinates (x, y) is on the stream axis, the x axis points east, and the y axis points north. We assume that

$$T = \text{const} - T_0 \sin \frac{\pi y}{2Y}; \quad U = -\frac{\alpha g z}{f} \frac{\partial T}{\partial y}; \quad |y| \leq Y, \quad (43.1)$$

where U is the velocity of the thermal wind, and α is the coefficient of thermal

expansion of the fluid. For the kinetic energy K of the jet stream we will use the equation

$$\frac{\partial K}{\partial t} = \overline{\rho u'v'} \frac{\partial U}{\partial y} + \varepsilon; \overline{u'v'} = -k \frac{LU^2}{T_0} \frac{\partial T}{\partial y}, \quad (43.2)$$

where ε includes the rate of generation of energy K by external forces, the rate of its exchange with the potential energy due to the work of the buoyancy forces, and the rate of its dissipation by small-scale processes. The second formula of (43.2) gives the parametrization of the Reynolds stresses produced by synoptic motions. This parametrization is analogous to that proposed by Williams and Davis (1965) (based on the fact that synoptic processes are generated by the baroclinic instability of the zonal current, so that their effects must be determined by the value of $\partial T/\partial y$). Here L is the scale of the synoptic processes, and k is a positive numerical constant. If in the region of the jet stream $\varepsilon \approx 0$, then we obtain

$$\frac{\partial K}{\partial t} = m^2 \left(\frac{\partial T}{\partial y} \right)^3 \frac{\partial^2 T}{\partial y^2} = - \left(\frac{\pi}{2Y} \right)^5 m^2 T_0^4 \cos^3 \frac{\pi y}{2Y} \sin \frac{\pi y}{2Y}, \quad (43.3)$$

where $m^2 = k\rho(agz/f)^3 LT_0^{-1}$. Therefore, in the zone $-Y \leq y \leq 0$ south of the jet stream it follows that $\partial K/\partial t > 0$, that is, the jet stream is enhanced at the expense of the synoptic motions: negative viscosity is acting. In the zone $0 \leq y \leq Y$ north of the stream, on the other hand, $\partial K/\partial t < 0$: there the synoptic motions act as a dissipative factor, that is, they produce positive viscosity. After some time the stream decays in the northern zone, while in the southern zone it intensifies, so that as a result the stream narrows to half its original width, and its channel is shifted southward by a distance of approximately $Y/2$. Then all these processes can recur, as long as they are not hindered by the factors ε , which were neglected at the outset (and possibly the effect of the external forces may return the jet stream to its initial location). These predictions have been verified by numerical experiments with 'eddy-resolving' models, which will be discussed in Section 47.

A model in which the jet streams delimit zones of positive and negative viscosity can be illustrated using as an example a stream in a flat channel $|y| \leq Y$ with a constant pressure gradient $\partial p/\partial x$. We assume that at the walls $y = \pm Y$ there are boundary layers of thickness h with a positive viscosity μ (and zero velocity at the walls) and a negative viscosity ($-\mu_1$) in the stream interior $|y| \leq Y - h$. The profile of the velocity $U(y)$ will here consist of segments of parabolas and it has cusps at the points $|y| = Y - h$:

$$U(y) \begin{cases} \frac{1}{2\mu} \frac{\partial p}{\partial x} (y^2 - Y^2), & |y| \geq Y - h; \\ \frac{1}{2\mu_1} \frac{\partial p}{\partial x} [(Y - h)^2 - y^2] + \frac{1}{2\mu} \frac{\partial p}{\partial x} [(Y - h)^2 - Y^2], & |y| \leq Y - h. \end{cases} \quad (43.4)$$

The profile of the dimensionless velocity

$$\tilde{U} = - \left(\frac{Y^2}{2\mu} \frac{\partial p}{\partial x} \right)^{-1} U$$

as a function of the dimensionless transverse coordinate $\bar{y} = y/Y$ for $\mu = \mu_1$ is shown in Figure 43.1.

The role of the zonal flow in the general circulation of the atmosphere depends considerably on the angular velocity of rotation of the planet ω , the value of which can be characterized conveniently by the *Mach rotation number* $\Pi_\omega = \omega r_0/c_r$, where r_0 is the radius of the planet, and c_r is the typical velocity of sound in its atmosphere. For slow rotation of the planet (in the case of Venus, for example, $\Pi_\omega = 6 \times 10^{-3}$) its direct effect on the atmospheric circulation is small, and the latter is controlled by the difference between the heating at the subsolar and antisolar points. If the planet rotates rapidly (as in the case of Jupiter: $\Pi_\omega = 17$), then this imparts to the zonal flow a decisive role in the circulation of the atmosphere.

On the Earth $\Pi_\omega = 1.4$, and this determines the specifics of the general circulation of the terrestrial atmosphere: *comparability of the roles of the zonal and monsoon circulations*. The latter arise from differences in the heating of the atmosphere over the continents and the oceans, which are subject to seasonal fluctuations, the continents being warmer in summer, and cooler in winter, than the oceans (a seasonal effect can also be detected in the meridional circulation: during each season only one tradewind cell exists in it, with ascending motions on the summer side of the equator and descending motions in the subtropics of the winter hemisphere, so that this cell transfers heat to the winter side).

The specifics of the atmospheric circulation on Mars can be summarized as follows: firstly, the main mass M of the thin atmosphere (for a mean atmospheric pressure at the Martian surface $p_s \sim 6$ mbar and a free-fall acceleration $g = 370$ cm/s², the mass $M = p_s/g \approx 16$ g/cm²) is concentrated in low places (primarily in the Northern Hemisphere, which is on the average 2 km lower than the Southern Hemisphere), so that *the topographical effects must be quite substantial*. Secondly, *global dust storms* are typical occurrences on Mars, and these originate on high plateaus in the southern tropics, predominantly at the time of the solstices in the Southern Hemisphere, after which they propagate due to the cooling of the lower atmosphere produced by these storms ('antigreenhouse effect'). This pro-

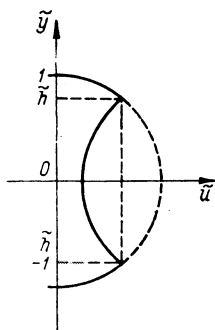


Fig. 43.1. Profile of flow velocity in flat channel with negative viscosity in stream interior and positive viscosity in wall boundary layers.

duces a thermal wind (which eddies cyclonically near the Martian surface and anticyclonically in the upper atmosphere), thereby lifting quantities of new dust into the atmosphere (Girache and Goody, 1972; Golitsyn, 1973; Hess, 1973).

Venus, on the other hand, has a very thick atmosphere ($p_s \sim 90 \text{ kg/cm}^2$, so that for $g = 890 \text{ cm/s}^2$ we have $M \approx 100 \text{ kg/cm}^2$), consisting primarily of carbon dioxide, which produces a very marked greenhouse effect (the temperature at the Venusian surface is about 470°C , the troposphere is 58 km thick, and the cloud layer lies at heights of 35 to 65 km). Venus rotates very slowly, with a period of 243 days, in a sense opposite to its revolution about the Sun (retrograde rotation); thus the solar days on the planet are equal to 117 days. The atmospheric circulation under these conditions should develop mostly between the subsolar point and the antisolar point. This is borne out by calculations based on the numerical model of Zilitinkevich *et al.* (1971). Because of the great thermal inertia of the atmosphere, the region of ascending motions lags behind the subsolar point, being shifted 45° toward the evening terminator. The region of descending motions, on the other hand, lies on the night side, near the morning terminator. The mean difference between the surface temperatures on the dayside and nightside was found to be 1°C . According to photographs taken by Mariner 10 (1974), in the vicinity of the subsolar point a continual disturbance is present, with a cellular cloud structure occupying $\pm 20^\circ$ of latitude and no less than 80° of longitude.

A specific property of the circulation of the Venusian atmosphere is a rapid (with a period of 4 days and in the same sense as the rotation of the planet) *zonal circulation of its upper layers*. This was detected from cloud observations in ultraviolet radiation (with wavelengths of 300 to 400 nm), emitted into space from just the layers above about 60 km with an atmospheric pressure less than 400 mbar (Bayet, 1960; Bayet and Camichelle, 1961). It may well be explained by the thorough heating of the atmosphere during the slow turning of the planet, which creates inclined convective cells that impart momentum to the zonal circulation (Schubert and Whitehead, 1969).

The very term 'atmosphere' does not apply very well to Jupiter. This planet consists, like the Sun, mainly of hydrogen and helium (with a ratio of their specific concentrations of about 4:1). As a first approximation, the internal structure of Jupiter can be described using an adiabatic model with a linear density distribution along the radius $\rho = 4\rho_0(1 - r/r_0)$ for $0.2 < r/r_0 < 0.98$, where $\rho_0 = 1.332 \text{ g/cm}^3$ is the mean density and $r_0 \approx 70,000 \text{ km}$ is the mean radius (Trubitsyn, 1971). On a pT diagram of the phase of the matter, an adiabat drawn through the point $p = 3 \times 10^5 \text{ Pa}$, $T = 150 \text{ K}$, pertaining to the upper cloud limit, is found to lie everywhere considerably above (according to T) not only the regions of solid hydrogen and helium (both normal and metallized), but also the critical points of Mendeleeev, terminating the lines separating the gaseous and liquid phases. Therefore, the matter is everywhere in the same fluid phase (and only for $r/r_0 < 0.2$ is there probably a solid silicate core). Figure 43.2 portrays the radial distributions of p , ρ , and T according to this model.

It should be noted, too, that at some depth (apparently at about $r/r_0 \sim 0.8$) matter becomes ionized and then continuously, without a density discontinuity,

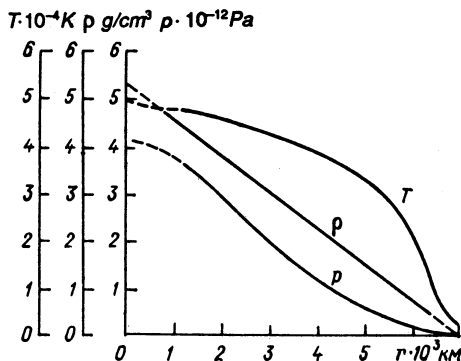


Fig. 43.2. Adiabatic model of internal structure of Jupiter, according to Trubitsyn (1971).

metallized. The flows in this electrically conducting fluid apparently also promote the generation of the observed strong magnetic field of Jupiter.

Since there is no basis for considering as atmosphere only the above-cloud layer of gas, on Jupiter we have to consider the circulation of the matter of the entire planet as a whole (except the relatively small silicate core).

The 'hot' model of Jupiter is verified by measurements of its surface temperature, which turned out to be considerably higher than the equilibrium temperature T_e given by the equation

$$f\sigma T_r^4 = q \left[= \frac{1}{4} I_0 (1 - A) \right], \quad (43.5)$$

where f is the emissivity, which is close to unity; $\sigma = 5.67 \times 10^{-8} \text{ W}/(\text{m}^2 \text{ K}^4)$ is the Stefan—Boltzmann constant; the left-hand side gives the flux of departing radiation; I_0 is the solar constant, which on Jupiter is only 1/27 of that on the Earth; A is the albedo (in the visible range equal to 0.45); q is the flux of the solar radiation taken up. This temperature $T_r \approx 109 \text{ K}$, whereas measurements in the infrared gave an average $T \approx 129 \text{ K}$, and at individual places in the dark belts (where the depths are not covered by clouds) even as high as 200 to 310 K (we should mention, too, the observations of a California group in 1962, according to which in the shadow of a moon the value of T at the disk of Jupiter was 50 K higher, since there apparently the clouds dissolve and the depths beneath the clouds are exposed).

This indicates that Jupiter possesses internal heat sources (possibly the heat of continuing compression). The flux of departing radiation produced by these sources is estimated as $1.8I_0$ to $2.0I_0$. It is natural to assume that it takes place via convection (probably multistage); we recall the excellent photographs of *Pioneer 10* and *Pioneer 11* (1973—1974) and *Voyager 1* and *Voyager 2* (1979), revealing a

cellular convective structure of the polar regions of Jupiter. The convection also produces the adiabatic structure of the interior of the planet.

For Jupiter (and the other giant planets) a very rapid and latitudinally differential rotation is typical (note that its equator is inclined to the orbital plane by only $3^\circ 05'$). Three longitude systems are used for Jupiter: the equatorial PI system with a rotation period of 9 h 55 min 30.003 s; the middle-latitude PII system with a period of 9 h 55 min 40.632 s; the PIII system based on bursts of decameter radio emission, with a period of 9 h 55 min 29.37 s. The circulation observed at its visible surface is naturally zonal, and it consists of bright *zones* parallel to the equator (clouds of ammonia ice) and dark reddish *belts* (Figure 43.3), which vary somewhat in time (as a rule, on a scale measured in years, although sometimes details up to 10^4 km in size vary in only 10^3 s). The Great Red Spot in the Southern tropical zone is a quasipermanent disruption of the zonality; the Spot is $40,000 \times 13,000$ km in size and has an anticyclonic vorticity, being apparently a long-lived synoptic eddy (Golitsyn, 1970). The banded structure may possibly be due to convective rolls, with ascending motions in the zones and descending motions in the belts.

Actually the circulation can be broken up into individual zonal flows (Table 43.1), at the boundaries of which (not always coinciding with the boundaries of the zones and belts) velocity discontinuities exist (sometimes enormous jumps, greater than the velocity of sound). A possible explanation of these is the limit of anisotropization of the large synoptic eddies, mentioned at the end of Section 37. The equatorial acceleration may well be a global effect of the negative viscosity produced by smaller synoptic eddies, while the velocity jumps may be an effect of the alternation of the zones of positive and negative viscosity of type (43.3) and (43.4).

The general circulation on the other giant planets, as well as in the convective layer of the Sun, is to a large degree qualitatively similar to what has been stated here for Jupiter.

TABLE 43.1
Zonal flows on disk of Jupiter

Latitude (degrees)	Period (9h + . . .)		Latitude (degrees)	Period (9h + . . .)		Latitude (degrees)	Period (9h + . . .)		Latitude (degrees)	Period (9h + . . .)	
	min	s		min	s		min	s		min	s
90 to 47	55	42	27	53	17	10 to 3	50	24	-22	55	38
43	55	20	23	49	07	3 to -3	50	24	-27	53	02
40 to 36	55	42	22 to 14	55	29	-3 to -10	50	26	-29	55	20
							51	21			
35	53	55	15	55	32	-19	55	39	-31 to -45	55	07
							58	43			
33 to 29	56	05	13	54	09	-21 to -26	55	36	-45 to -90	55	30

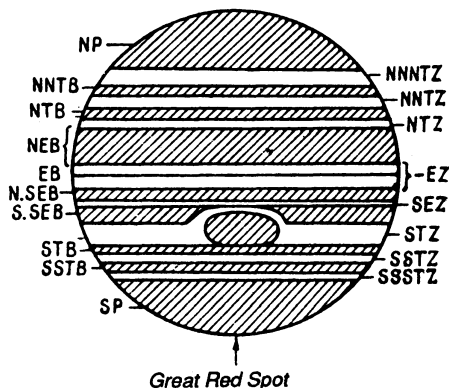


Fig. 43.3. Banded structure of Jovian surface: N north, S south, B belt, Z zone, E equator, P pole, T tropics.

44. SIMILARITY THEORY FOR THE CIRCULATION OF PLANETARY ATMOSPHERES

In order to characterize the atmospheric circulation, it is not enough just to know the equilibrium temperature $T_r = (q/f\sigma)^{1/4}$. At the very least, we also have to know the typical temperature difference δT in the atmosphere between the equator and the poles. This difference enters into the definition of the specific rate ε of generation of kinetic energy in the atmosphere (Golitsyn, 1968):

$$\varepsilon = \eta \frac{q}{M}; \quad \eta = k \frac{\delta T}{T_r}, \quad (44.1)$$

where M is the mass of an air column of unit cross section, η is the efficiency of the atmospheric 'heat engine', and k is the coefficient of heat utilization ($k < 1$, since an actual heat machine is not ideal).

For example, in the Earth's atmosphere for $I_0 \approx 1370 \text{ W/m}^2$ and $A \approx 0.30$, we have $q \approx 240 \text{ W/m}^2$ and, according to (43.5), $T_r \approx 255 \text{ K}$. This temperature exists in the atmosphere at a height of about 5 km. Its ratio to the mean temperature $T_s \approx 288 \text{ K}$ of the atmospheric surface layer $\alpha_r = T_r/T_s$ characterizes the temperature stratification of the lower atmosphere. For $\varepsilon \approx 4 \times 10^{-4} \text{ m}^2/\text{s}^3$ we obtain $\eta \approx 0.018$, which for $\delta T \approx 50 \text{ K}$ yields $k \approx 0.1$.

From T_r we can determine the velocity of sound $c_r = (\kappa R/\mu T_r)^{1/2}$, where $\kappa = c_p/c_v$ is the specific-heat ratio, $R = 8.314$ J/(mole K) is the universal gas constant, and μ is the relative molecular mass of the atmospheric gas. Then we can introduce the scale of synoptic processes $L = c_r/2\omega$, where ω is the angular velocity of rotation of the planet, their typical time $\tau \approx \varepsilon^{-1/3} L^{2/3}$, the typical wind speed $U \sim (\varepsilon L)^{1/3}$, and the coefficient of horizontal mixing due to these processes $K_h \sim 0.1 \varepsilon^{1/3} L^{4/3}$. For instance, for the Earth's atmosphere with $\kappa \approx 1.4$ and $\mu = 29$ we obtain $c_r \approx 320$ m/s, $L \sim 3000$ km, $\tau \sim 3$ days, $U \sim 10$ m/s, and $K_h \sim 3 \times 10^6$ m²/s.

It should be noted here that the value of δT need not be specified, since it can be determined with the aid of relations from the thermodynamics of fluids. For

instance, we can use the condition of equilibrium between the divergence of the meridional heat flux caused by air currents and heat losses to radiation. The condition can be put into the following simplified form:

$$c_p MU \frac{\delta T}{\pi r_0/2} = q [= f\sigma T_r^4]. \quad (44.2)$$

Using the above definitions of T_r , ε , and U , we obtain from (44.2) the relation

$$\delta T = (f\sigma)^{-1/16} c_p^{-3/4} q^{9/16} \left(\frac{\pi r_0}{2M} \right)^{1/2} \left(\frac{\pi}{k} \frac{r_0}{2L} \right)^{1/4} \quad (44.3)$$

The factor with $r_0/2L$ describes the influence on δT of the velocity of rotation of the planet, which was discussed in the preceding section. For rapid rotation we have $2L = c_r/\omega < r_0$, so that $r_0/2L = \omega r_0/c_r = \Pi_\omega > 1$, and, since the heat transport from the equator to the poles by comparatively small eddies is not very great, the temperature difference $\delta T \sim \omega^{1/4}$ turns out to be quite large. On the other hand, for slow rotation $L \sim 2r_0$, and such rotation does not affect the value of δT . Note, too, the marked dependence of a number of characteristics of the atmospheric circulation on the mass M of an air column of unit section:

$$\eta \sim U \sim \delta T \sim M^{-1/2}; \quad \varepsilon \sim M^{-3/2}. \quad (44.4)$$

This has, in particular, the following interesting corollary: the total kinetic energy of the atmospheric circulation $E \sim MU^2$ is independent of M .

The foregoing preliminary statements show that the characteristics of the atmospheric circulation depend on seven dimensional 'external' parameters: the assimilable flux of solar radiation (or endogenic heat) q , the radius of the planet r_0 , its angular velocity of rotation ω , the mass M of an air column of unit cross section, the free-fall acceleration in the atmosphere g (determined by the mass and radius of the planet and, in turn, determining the pressure at the bottom of the atmosphere $p_s = Mg$; in the interior of the giant planets and the Sun we also have to determine the function $g(r)$); the specific heat of the atmospheric air c_p , and the reduced Stefan—Boltzmann constant $f\sigma$. In addition, these characteristics depend on one dimensionless 'internal' parameter: the relative molecular mass of the atmospheric air μ (entering into the formulas only via $R/\mu = c_p - c_v$).

Golitsyn (1970, 1973) stated the hypothesis that these eight parameters completely determine the atmospheric circulation. He constructed a corresponding similarity theory for the circulation of planetary atmospheres. Since there are four independent dimensions (length, time, mass, and temperature), using the seven 'external' dimensional parameters we can construct three independent dimensionless combinations (similarity criteria):

$$\begin{aligned} \Pi_\omega &= \frac{r_0}{2L} = \frac{\omega r_0}{c_r}, & \Pi_g &= \frac{H}{r_0} = \frac{c_r^2}{\kappa g r_0} \left(H = \frac{c_r^2}{\kappa g} \right); \\ \Pi_M &= \frac{\tau_p}{\tau_r} = (f\sigma)^{3/8} c_p^{-3/2} q^{5/8} \frac{r_0}{M} \left(\tau_p = \frac{r_0}{c_p}; \quad \tau_r = \frac{c_p M T_r}{q} \right). \end{aligned} \quad (44.5)$$

We can stipulate that ω must enter into all the formulas only via Π_ω (while, if ω is so small that $L > 2r_0$, then we get $\Pi_\omega = \frac{1}{4}$), that g must enter in only via Π_g (in the expression for Π_g the quantity H is the so-called depth of the homogeneous atmosphere), and that M must enter in only via criterion Π_M , which is the ratio of the relaxation time of disturbances of the pressure (or density) τ_p to the time of thermal relaxation τ_r (to the state of local radiation equilibrium). The similarity criteria for the planets, according to Golitsyn (1973), have the values shown in Table 44.1.

The values of Π_M for the giant planets are only tentative, since they apply to the above-cloud layers. Thus only the order of magnitude is given. The smallness of Π_g and Π_M means that we can rely on self-similarity for these parameters, whereas Π_ω varies over a wide range and can serve as a basis for the classification of circulations of planetary atmospheres.

According to similarity theory, any characteristic of the atmospheric circulation F must have the form

$$F = F_0(f\sigma, c_p, q, r_0) \Psi_F(\Pi_\omega, \Pi_g, \Pi_M), \tag{44.6}$$

where F_0 is a combination of the dimension F of parameters $f\sigma$, c_p , q , and r_0 , while Ψ_F is a dimensionless function of the similarity criteria (for low Π_M it probably depends on this criterion as some power). For instance, the total kinetic energy of the general circulation of the atmosphere should have the form

$$E = \frac{1}{2} (4\pi r_0^2 M) U_0^2 = (f\sigma)^{1/8} c_p^{1/2} q^{7/8} r_0^3 \Psi_E(\Pi_\omega, \Pi_g), \tag{44.7}$$

where U_0 is the rms velocity of the atmospheric motions. Here it is seen that Π_M differs from the square of the Mach number $(\text{Ma})^2 = (U_0/c_r)^2$ only by the factor $2\pi(\kappa - 1)\Psi_E^{-1}$. Function Ψ_E can be found by setting $\varepsilon = E/4\pi r_0^2 \tau_0$, where $\tau_0 = L/U_0$, and by comparing this formula for ε with the first formula of (44.1), which yields Ω , and then finding from the second formula of (44.1) the value of δT and

TABLE 44.1
Similarity criteria for the planets

Planet	Π_ω	Π_g	Π_M
Mercury	8.5×10^{-3}	1.0×10^{-2}	> 1
Venus	7.6×10^{-3}	3.3×10^{-4}	1×10^{-5}
Earth	1.43	1.2×10^{-3}	1.17×10^{-3}
Mars	1.05	3.2×10^{-3}	3.3×10^{-2}
Jupiter	15.6	2.4×10^{-4}	10^{-4}
Saturn	14.7	5.5×10^{-4}	10^{-4}
Uranus	7.5	1.0×10^{-3}	10^{-5}
Neptune	6	6.0×10^{-4}	10^{-5}

comparing it with (44.3). As a result we obtain

$$\Psi_E = (2\pi)^{1/3} k^{1/2} \left(\frac{r_0}{L} \right)^{-1/2}, \quad L < r_0;$$

$$E = (2\pi)^{1/3} k^{1/2} (f\sigma)^{1/8} c_p^{-1/2} q^{7/8} r_0^3 \left(\frac{r_0}{L} \right)^{-1/2}. \quad (44.8)$$

Then $U = (2E/M)^{1/2}$, and from (44.2) we get $\delta T = (\pi r_0/2) q / c_p MU$, while η and ε are found from (44.1).

In order to estimate possible variations of the atmospheric circulation when there are changes in the assimilable heat flux q and in its meridional distribution, we can further develop the above theory as follows (Zilitinkevich and Monin, 1973). Equation (43.5) remains valid and, accurate to some constant (not variable) factor, it takes the form $T \sim q^{1/4}$. In Equation (44.1) quantity q varies by the typical difference δq of the assimilable heat fluxes at the equator and at the pole, and this equation is written, also accurate to a fixed factor, in the form $\varepsilon \sim (\delta T/T) \delta q$. Equation (44.2) takes the form $c_p M (LU/r_0) (L \delta T/r_0) \sim \delta q$, and since $L \sim c_r/\omega \sim T^{1/2}$ it becomes $TU \delta T \sim \delta q$. Finally, we add the definition $\varepsilon \sim U\tau$, where $\tau \sim \rho U^2$ is the typical frictional stress in the boundary layer of the atmosphere, while $\rho \sim T^{-1}$ is the air density, so that we have $\varepsilon \sim U^3 T^{-1}$. From the four above relations we arrive at the following simple formulas:

$$T \sim q^{1/4}; \quad \delta T \sim q^{-3/16} (\delta q)^{1/2}; \quad U \sim q^{-1/16} (\delta q)^{1/2};$$

$$\varepsilon \sim q^{-7/16} (\delta q)^{3/2}. \quad (44.9)$$

Let us use these to estimate the seasonal fluctuations of the atmospheric circulation, taking into account that the summer value of q is greater by a factor of $1 + \alpha$ than the annual mean (according to astronomical data, on the day of the summer solstice $\alpha = 0.33$) and also that the excess flux αq falls mainly in the polar region, that is, δq acquires the multiplier $1 - \alpha$. For the winter hemisphere we have to set $\alpha = -0.39$. The results of the calculations are given in Table 44.2.

The seasonal fluctuations of the mean temperature T obtained here were somewhat high, since the atmosphere in its seasonal fluctuations departs considerably from radiation equilibrium, which is what is described by the formula $T \sim q^{1/4}$. At the same time, estimates of the seasonal fluctuations of all the other characteristics turned out to be quite satisfactory. This reliability of formulas (44.9) enables us to employ them to evaluate the possibilities of an energy effect on the general circulation of the atmosphere.

For instance, for an increase in the solar constant by a factor of $1 + \alpha$ or for some given arrangement over the Earth's surface of artificial heat sources with a strength αq (scaled to a unit area of the entire Earth), we have $q \sim 1 + \alpha$ and $T \sim (1 + \alpha)^{1/4}$. For $\alpha = 0.1$ this gives an increment of the mean temperature of 7.2°C (although this estimate, as explained above, is too high). If these sources are distributed uniformly over the entire surface of the Earth we have $\delta q \sim 1$, over

TABLE 44.2
Results of calculations according to (44.9)

	T (K)	δT (°C)	U (m/s)	$E \times 10^{-20}$ (J)	$\varepsilon \times 10^4$ (m ² /s ³)
Actual annual mean	287	60	20	5.3	5
Multiplier for summer	1.07	0.78	0.80	0.64	0.48
Multiplier for winter	0.88	1.29	1.22	1.49	2.04
Summer extremum, estimate	309.3	46.6	16.0	3.4	2.4
Southern Hemisphere, actual		40.2		3.9	
Winter extremum, estimate	254.6	77.6	24.4	7.8	10.2
Southern Hemisphere, actual		74.2		7.1	

the equator $\delta q \sim 1 + \alpha$, at both poles $\delta q \sim 1 - \alpha/2$, and at one pole in this hemisphere $\delta q \sim 1 - \alpha$ (such an effect would be produced by a lowering of the albedo of the Arctic; for example, a reduction of A from 0.8 to 0.3 would increase the local heat flux by a factor of $(1 - 0.3)/(1 - 0.8) = 3.5$ or by $2.5q$, but referred to a unit area of the entire Earth, which is 40 times greater than the area of the Arctic, we would have $\alpha = 2.5/40 \approx 0.06$, only one-sixth of the value for the seasonal fluctuations). For $\alpha = 0.1$, this is equivalent to an enormous intensity of the artificial heat sources, of the order of 10^{13} kW, which yields only small variations: δT from -4 to 2 °C, U in the range ± 1 m/s, and ε in the range $\pm 0.5 \times 10^{-4}$ m²/s³.

45. NUMERICAL MODELS OF THE GENERAL CIRCULATION OF THE ATMOSPHERE

The discussion in Section 43 was qualitative, while that in Section 44, although quantitative, was only tentative in nature. Since atmospheres are distinguished by considerable spatial inhomogeneity, both vertical and horizontal, which moreover varies in time (for instance, from season to season), a sufficiently complete description of the general circulation calls for a calculation of the corresponding thermodynamic fields of the fluids (hydrothermodynamic fields).

We will confine ourselves here to the comparatively thin atmospheres of Venus, Earth, and Mars. Let us note that the motions in the global and synoptic atmospheric processes to be described individually here are practically quasistatic and nondivergent. They can be described by the appropriate form of simplified fluid-dynamical equations (in which with the aid of the hydrostatic equation we should use the so-called 'traditional approximation' for the Coriolis acceleration, and also, in view of the smallness of the heights $z = r - r_0$ compared to the distance r from the planet center, we can everywhere replace r by the planetary radius r_0). The hydrostatic equation and the condition of nondivergence of the velocity field can be written as

$$\frac{\partial p}{\partial z} = -g\rho; \quad \frac{\partial u_z}{\partial z} + \frac{1}{r_0 \sin \theta} \left(\frac{\partial u_\theta \sin \theta}{\partial \theta} + \frac{\partial u_\lambda}{\partial \lambda} \right) = 0. \quad (45.1)$$

The equations of horizontal motion will have the form:

$$\begin{aligned} \frac{du_\theta}{dt} - \frac{\operatorname{ctg} \theta}{r_0} u_\lambda^2 - fu_\lambda &= -\frac{1}{r_0 \rho} \frac{\partial p}{\partial \theta} + F_\theta; \\ \frac{du_\lambda}{dt} + \frac{\operatorname{ctg} \theta}{r_0} u_\theta u_\lambda + fu_\theta &= -\frac{1}{r_0 \rho \sin \theta} \frac{\partial p}{\partial \lambda} + F_\lambda; \\ \frac{d}{dt} &= \frac{\partial}{\partial t} + u_z \frac{\partial}{\partial z} + \frac{u_\theta}{r_0} \frac{\partial}{\partial \theta} + \frac{u_\lambda}{r_0 \sin \theta} \frac{\partial}{\partial \lambda}, \end{aligned} \quad (45.2)$$

where, as usual, $f = 2\omega \cos \theta$ is the Coriolis parameter, and F_θ and F_λ are the components of the friction force, expressed in terms of the frictional-stress tensor τ_{ij} :

$$\begin{aligned} \rho F_\theta &= \frac{\partial \tau_{\theta z}}{\partial z} + \frac{1}{r_0 \sin \theta} \left(\frac{\partial \tau_{\theta\theta} \sin \theta}{\partial \theta} + \frac{\partial \tau_{\theta\lambda}}{\partial \lambda} \right) - \frac{\operatorname{ctg} \theta}{r_0} \tau_{\lambda\lambda}; \\ \rho F_\lambda &= \frac{\partial \tau_{\lambda z}}{\partial z} + \frac{1}{r_0 \sin \theta} \left(\frac{\partial \tau_{\lambda\theta} \sin \theta}{\partial \theta} + \frac{\partial \tau_{\lambda\lambda}}{\partial \lambda} \right) + \frac{\operatorname{ctg} \theta}{r_0} \tau_{\lambda\theta}. \end{aligned} \quad (45.3)$$

For the velocity field the boundary condition at the land surface $z = h(\theta, \lambda)$ has the form $u_z = u_\theta = u_\lambda = 0$, whereas at the ocean surface, which may be taken to be horizontal ($z = 0$) with respect to large-scale atmospheric processes, we set $u_z = 0$ and require continuity of u_θ , u_λ , p , $\tau_{\theta z}$, and $\tau_{\lambda z}$. Usually, however, Equations (45.2) are of interest only in the region above the atmospheric boundary layer. There the terms $\partial \tau_{\theta z} / \partial z$ and $\partial \tau_{\lambda z} / \partial z$ can frequently be neglected, and the boundary layer itself is 'parametrized', that is, the frictional stress ρu_*^2 and the turbulent fluxes of heat q and moisture E at its lower side are expressed in terms of the external parameters z_0 , U , δT , δs , f , and g/T_0 with the aid of formulas of the type of (36.8) and (36.9).

Equations (45.1), (45.2) have to be supplemented by the equation of state of moist air (1.7). It contains two new fields T and s , and to determine these we invoke the equations of evolution of the entropy and moisture:

$$\frac{d\eta}{dt} = \frac{\varepsilon + \mathcal{L}_m}{T} - \left[\frac{\partial Q_z}{\partial z} + \frac{1}{r_0 \sin \theta} \left(\frac{\partial Q_\theta \sin \theta}{\partial \theta} + \frac{\partial Q_\lambda}{\partial \lambda} \right) \right]; \quad (45.4)$$

$$\frac{ds}{dt} = -m - \left[\frac{\partial I_z}{\partial z} + \frac{1}{r_0 \sin \theta} \left(\frac{\partial I_\theta \sin \theta}{\partial \theta} + \frac{\partial I_\lambda}{\partial \lambda} \right) \right], \quad (45.5)$$

where ε is the rate of *radiative* influx of heat (and dissipation of kinetic energy) per unit mass of air, m is the specific rate of condensation or sublimation of water vapor (for $m < 0$ this is the rate of vaporization or volatilization), $\mathcal{L}$ is the latent heat of condensation or sublimation (for condensation $\mathcal{L} = \mathcal{L}_0 - (c_w - c_{pv})(T - T_0)$, $\mathcal{L}_0 = 2,502,500$ J/kg, $c_w = 4190$, and $c_{pv} = 1810$ J/(kg K), $T_0 = 0^\circ\text{C}$; for sublimation $\mathcal{L} = 2,836,300$ J/kg), and Q and I are the small-scale fluxes of entropy and moisture (see (1.9), (1.10), and (1.13)).

The quantity m is found from the condition of saturation of the air with moisture in clouds, the saturation humidity $s = s_m(T, p)$ being determined from the pressure of saturated water vapor $p_{om}(T)$, satisfying the Clausius—Clapeyron equation $T^2 \partial \ln p_{om} / \partial T = \mathcal{L} / R_v$ (for a more detailed description of clouds and precipitation, we have to add the evolution equations for the concentrations of liquid water and ice).

The boundary condition for Equation (45.4) at the lower bound of the atmosphere is in this case the heat-budget equation:

$$q_s = (1 - A) F_s^i + F_L^i - B(T_s) - q - \mathcal{L}E + q_p + q_{wi}, \quad (45.6)$$

where q_s is the heat flux in the interior of the ocean or soil; F_s^i is the downward flux of direct and scattered short-wave solar radiation; A is the corresponding albedo of the underlying surface; F_L^i is the downward flux of long-wave radiation of the atmosphere; $B(T_s) = \epsilon \sigma T_s^4$ is the long-wave radiation of an underlying surface with a temperature T_s ; q is the turbulent heat flux in the atmosphere (positive upward); $E = \rho I_e$ is the evaporation rate; $q_p = c_w P (T_p - T_s)$ is the heat flux brought in by precipitation (P is the precipitation rate and T_p is the precipitation temperature); q_{wi} is the heat of freezing of water or melting of ice or snow (positive for freezing).

The boundary condition for Equation (45.5) at a water surface is the equation of saturation humidity $s = s_m(T_s, p_s)$ at temperature T_s , and at a land surface it is the moisture-budget equation:

$$I_s = P - E - \frac{q_{wi}}{\mathcal{L}_{wi}} - E_s, \quad (45.7)$$

where I_s is the moisture flux into the soil, $q_{wi}/\mathcal{L}_{wi}$ is the water loss during freezing (for $q_{wi} < 0$ the influx during melting; $\mathcal{L}_{wi} = 334,000$ J/kg is the specific latent heat of freezing of ice), and E_s is the rate of surface runoff.

The mechanism of the general circulation of the atmosphere can be characterized expressively in terms of the conversions of kinetic and available potential energy $K = \bar{K} + K'$ and $\Pi = \bar{\Pi} + \Pi'$ which take place, where, having in mind primarily the Earth, we divide these into parts $(\bar{K}, \bar{\Pi})$ and (K', Π') , pertaining respectively to the zonal circulation and non-zonal processes. Using Equation (5.3) for $\bar{K}$ as an example, let us derive the following exact local equation:

$$\begin{aligned} & \frac{\partial}{\partial t} \left(\bar{\rho} \frac{\bar{v}^2}{2} + \bar{\mathbf{v}} \cdot \overline{\rho' \mathbf{v}'} \right) + \nabla_\alpha [(\bar{\rho} \bar{v}^\alpha + \overline{\rho' v'^\alpha}) \frac{\bar{v}^2}{2} + \bar{\mathbf{v}} \cdot (\overline{\rho' \mathbf{v}'} \cdot \bar{\mathbf{v}}^\alpha + \\ & + \overline{\rho' \mathbf{v}' v'^\alpha}) + \bar{p} \bar{v}^\alpha - \bar{\sigma}^{\alpha\beta} \bar{v}^\beta] = [\overline{\rho \mathbf{v}' v'^\alpha} \cdot \nabla \bar{v}^\alpha - \overline{\rho' v'_\alpha} (\bar{\mathbf{v}'} \cdot \nabla \bar{v}^\alpha + \\ & + \overline{\rho^{-1} \nabla_\alpha p} - \overline{\rho^{-1} \nabla_\beta \sigma^{\alpha\beta}})] + (\bar{\rho} \bar{\mathbf{v}} + \overline{\rho' \mathbf{v}'}) \cdot \mathbf{g} + \bar{p} \operatorname{div} \bar{\mathbf{v}} - \frac{1}{2} \bar{\sigma}^{\alpha\beta} \bar{\varphi}_{\alpha\beta}, \quad (45.8) \end{aligned}$$

where $\varphi_{\alpha\beta}$ is the tensor of the deformation rates. We obtain the equation for K'

similarly:

$$\begin{aligned}
 & \frac{\partial}{\partial t} \left(\bar{\rho} \frac{\overline{v'^2}}{2} \right) + \nabla_\alpha \left(\frac{1}{2} \overline{\rho v'^2 \bar{v}^\alpha} + \frac{1}{2} \overline{\rho v'^2 v'^\alpha} + \overline{p' v'^\alpha} - \overline{\sigma'^{\alpha\beta} v'_\beta} \right) \\
 &= - [\overline{\rho \mathbf{v}' v'_\alpha} \cdot \nabla \bar{v}^\alpha - \overline{\rho' v'_\alpha} (\mathbf{v}' \cdot \nabla v'^\alpha) + \overline{\rho^{-1} \nabla_\alpha p} - \overline{\rho^{-1} \nabla_\beta \sigma'^{\alpha\beta}}] + \\
 &+ \overline{p' \operatorname{div} \mathbf{v}'} - \frac{1}{2} \overline{\sigma'^{\alpha\beta} \varphi'_{\alpha\beta}}. \tag{45.9}
 \end{aligned}$$

The terms in square brackets on the right-hand sides of these two equations differ only in sign, and thus they describe the exchange between $\bar{K}$ and K' . The first, and apparently main, term $\overline{\rho \mathbf{v}' v'_\alpha} \cdot \nabla \bar{v}^\alpha$, describing the work of the Reynolds stresses at the gradients of the zonal velocity, is apparently positive, on the average, that is, it corresponds to the transfer of kinetic energy from nonzonal motions (synoptic processes) to the zonal circulation (negative viscosity, barotropic stability of zonal circulation). The local equation for $\bar{\Pi}$ is obtained in the form

$$\begin{aligned}
 & \frac{\partial}{\partial t} \bar{\rho} (g\bar{z} + c_v \bar{T}) + \nabla_\alpha [(g\bar{z} + c_v \bar{T}) (\bar{\rho} \bar{v}^\alpha + \overline{\rho' v'^\alpha})] \\
 &= \left[\overline{\rho' \mathbf{v}' \cdot \nabla (g\bar{z} + c_v \bar{T})} - \overline{\rho \mathbf{v}' \cdot \nabla (gz' + c_v T')} - \right. \\
 &\quad \left. - \bar{\rho} \overline{\rho^{-1} \nabla_\alpha \left(Q^\alpha + \frac{\partial w}{\partial s} I^\alpha \right)} \right] - \\
 &\quad - \overline{\bar{\rho} \mathbf{v} \cdot \mathbf{g}} - \overline{\bar{\rho} \rho^{-1} p \operatorname{div} \mathbf{v}} + \frac{1}{2} \overline{\bar{\rho} \rho^{-1} \sigma'^{\alpha\beta} \varphi_{\alpha\beta}}. \tag{45.10}
 \end{aligned}$$

The equation for Π' is obtained similarly:

$$\begin{aligned}
 & \frac{\partial}{\partial t} (g \overline{\rho' z'} + c_v \overline{\rho' T'}) + \nabla_\alpha \left[(g \overline{\rho' z'} + c_v \overline{\rho' T'}) \bar{v}^\alpha + \overline{(gz' + c_v T') \rho v'^\alpha} + \right. \\
 &+ \bar{Q}^\alpha + \left. \frac{\partial w}{\partial s} I^\alpha \right] = - \left[\overline{\rho' \mathbf{v}' \cdot \nabla (g\bar{z} + c_v \bar{T})} - \overline{\rho \mathbf{v}' \cdot \nabla (gz' + c_v T')} - \right. \\
 &\quad \left. - \bar{\rho} \overline{\rho^{-1} \nabla_\alpha \left(Q^\alpha + \frac{\partial w}{\partial s} I^\alpha \right)} \right] - \overline{\rho' \mathbf{v}' \cdot \mathbf{g}} - \overline{(1 - \bar{\rho} \rho^{-1}) p \operatorname{div} \mathbf{v}} + \\
 &+ \frac{1}{2} \overline{(1 - \bar{\rho} \rho^{-1}) \sigma'^{\alpha\beta} \varphi_{\alpha\beta}}. \tag{45.11}
 \end{aligned}$$

The terms in square brackets on the right-hand sides of these two equations describe the exchange between $\bar{\Pi}$ and Π' . The first, and apparently main, term $\overline{\rho' \mathbf{v}' \cdot \nabla (g\bar{z} + c_v \bar{T})}$ describes the baroclinic instability of the zonal circulation (see (12.16)). According to empirical estimates, it is negative, on the average, that is, it describes the transfer of available potential energy from the zonal circulation to the synoptic processes. Moreover, $\bar{K}$ is exchanged with $\bar{\Pi}$ at a rate $\overline{\rho \bar{\mathbf{v}} \cdot \mathbf{g}}$ and with Π' at a rate

$$\overline{\rho' \mathbf{v}' \cdot \mathbf{g}} + \bar{p} \operatorname{div} \bar{\mathbf{v}} - \frac{1}{2} \bar{\sigma}^{\alpha\beta} \bar{\varphi}_{\alpha\beta},$$

while K' is exchanged with Π' at a rate $\overline{p' \operatorname{div} \mathbf{v}'} - \frac{1}{2} \overline{\sigma'^{\alpha\beta} \varphi'_{\alpha\beta}}$ (direct exchange with $\bar{\Pi}$ does not exist).

Empirical estimates of the mean annual values of the elements making up the energy cycle of the general circulation of the atmosphere (of the Earth's Northern Hemisphere) were obtained by Oort and Peixoto (1974) according to data for the years 1958 to 1963 (see Figure 45.1). The energy values there were given in 10^5 J/m^2 , and the rates of energy conversions in W/m^2 , arrows pointing outward toward the top and bottom signify energy exchange with the Southern Hemisphere; arrows pointing out to the right signify conversions of $\bar{K}$ and K' to Π' due to the viscous dissipation of kinetic energy. The main branch of the energy cycle has the form $\bar{\Pi} \rightarrow \Pi' \rightarrow K' \rightarrow \bar{K} \rightarrow (\bar{\Pi}, \Pi')$. The above investigators obtained similar estimates for January and July as well, and they concluded that the winter energy cycle is much more intense than the summer cycle.

If we approximate the above differential equations by difference equations on some space grid of points, solvable by finite steps in time, then we arrive at a numerical model of the general circulation of the atmosphere.

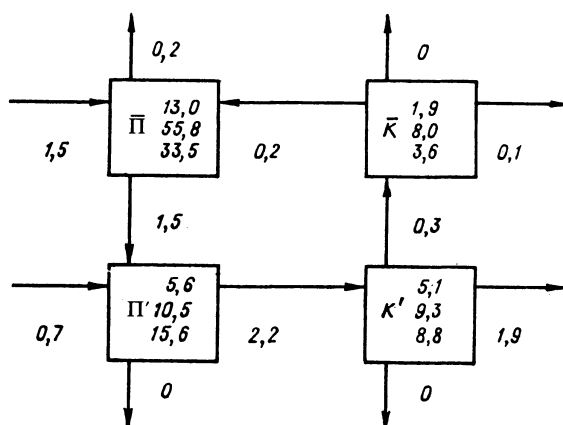


Fig. 45.1. Energy cycle of general circulation of atmosphere of Earth's Northern Hemisphere, according to Oort and Peixoto (1974). Upper number gives mean for year, middle number gives winter mean, lower number gives summer mean.

The first such model was constructed by Phillips (1956). It was a two-level quasigeostrophic model on a grid in the β plane with intervals $\Delta x = 375$ km and $\Delta y = 625$ km. The instantaneous state of the atmosphere was characterized by 450 numbers. The time interval was 0.5 to 2 hours. The radiative influx of heat was specified by the empirical formula $\varepsilon = -2\varepsilon_0 y/L$, where $\varepsilon_0 = 2 \times 10^{-3} \text{ m}^2/\text{s}^3$ and $L = 5000$ km. The condensation influx of heat was taken into account only via the lowering (by 20%) of the parameter of static stability. The dissipative factors taken into account in this model were the horizontal eddy viscosity and the heat conduction, with a coefficient $\mathcal{K}_h = 10^5 \text{ m}^2/\text{s}$, and the friction at the Earth's surface, with a drag coefficient $c_f = \tau/\rho U^2 = 3 \times 10^{-3}$. In the main experiment the equations of the model were integrated, beginning from a slightly disturbed state of rest, over a 31-day period. During this time the model atmosphere reached states with wind velocities of the order of 10 m/s and amplitudes of synoptic pressure fluctuations of the order of 20 mbar.

The success of this model inspired an intensive development of numerical models of planetary atmospheres. For the Earth one of the best was the 9-level model of a dry atmosphere constructed by Smagorinsky *et al.* (1965) and the model of a moist atmosphere of Manabe, Smagorinsky, and Strickler (1965), as well as the 11-level model with a seasonal variation worked out by Manabe and Holloway (1974) (later there occurred a transition to models of the joint circulation of the atmosphere and the ocean, which will be examined in the next chapter).

In these models the vertical coordinate was the normalized pressure $\sigma = p/p_s$; the quantity $\dot{\sigma}$, analogous to the vertical velocity, was taken to be zero for $\sigma = 1$ and $\sigma = 0$. For $\sigma = 1$ the simplified heat-budget equation (45.6) was used (in particular, without any heat flow into the soil). In the 'dry' model the flux of long-wave radiation was calculated as a function of temperature, for specified climatological distributions (over latitude and longitude) of water vapor, CO_2 , ozone, cloudiness, and albedo of the underlying surface, while the flux of short-wave radiation was calculated for the mean daily cosine of the zenith angle of the Sun. The concept of 'convective adjustment', which then enjoyed popularity, was introduced: if $|\partial T/\partial z|$ exceeded the moist-adiabatic gradient $|\partial T/\partial z|_a$, then the quantity $|\partial T/\partial z|$ was replaced by $|\partial T/\partial z|_a$, the total potential energy being conserved. Turbulent mixing was described with the aid of the exchange coefficients $\mathcal{K}_z$ and $\mathcal{K}_h$, expressed in terms of the velocity gradients via the formulas of Prandtl's theory (mentioned in Section 33).

The equations of this model were written in difference form on a space grid of points with 9 levels along the vertical (not counting the boundaries $\sigma = 1$ and $\sigma = 0$) and 20 levels of latitude between the equator and the poles. They were integrated with a time interval of 20 min in a period of 310 days, and the results for the 231st to 300th days were used to compute the mean values of the hydrothermodynamic fields, the mean elements of the energy cycle, the mean longitude spectra of the velocity components, and the mean values of a number of other characteristics of the atmospheric circulation.

The 'moist' model was distinguished by the fact that the humidity field and the thermal effects of phase changes of moisture and precipitation were already

calculated rather than specified. The underlying surface was taken to be wet (and not to have a heat capacity), while the 'convective adjustment' was supplemented by the prohibition of supersaturation of the air with moisture $s \leq s_m$ (that is, for $s > s_m$ the quantity s was replaced by s_m with a release $\mathcal{L}(s - s_m)$ of heat and the formation $s - s_m$ of precipitation). Such a model reproduced the atmospheric circulation somewhat better than the 'dry' model, and it gave a reasonably good description of the elements of the hydrologic cycle (for instance, the mean annual precipitation was found to be 106 cm/year, as compared with an empirical estimate of 104 cm/year).

The zonal average of the results obtained with a three-dimensional model of the atmosphere reveals the zonal circulation, but to describe this circulation it would, at first glance, appear to be simpler to use the zonally averaged equations of the hydrothermodynamics of fluids. Then, of the three space coordinates (θ, λ, z) only two (θ, z) are retained, while the three velocity components $\bar{u}_\theta, \bar{u}_\lambda, \bar{u}_z$ are expressed in terms of two functions of (θ, z, t) , say, the stream function ψ in the plane of the meridian and the angular velocity α of the differential zonal rotation of the atmosphere relative to the underlying surface:

$$\bar{u}_\theta = \frac{1}{\sin \theta} \frac{\partial \psi}{\partial z}; \quad \bar{u}_z = -\frac{1}{r_0 \sin \theta} \frac{\partial \psi}{\partial \theta}; \quad \bar{u}_\lambda = \alpha r_0 \sin \theta. \quad (45.12)$$

However, a very important complication arises: in the averaged equations there are some new unknowns, the second moments $\overline{u'_i u'_j}$, $\overline{u'_i \eta'}$, and $\overline{u'_i s'}$, produced by large-scale nonzonal motions. Their expression in terms of the gradients of the mean fields and the specification of the corresponding coefficients of large-scale eddy viscosity, heat conduction, and diffusion involve considerable arbitrariness. Moreover, in the case of moment $\overline{u'_\theta u'_\lambda}$ it is *necessary* to introduce the negative viscosity, for instance, via a formula of the type of (43.2).

However, despite this arbitrariness, Williams and Davis (1965) and a number of subsequent investigators succeeded in constructing in this way numerical models of the zonal circulation of the Earth's atmosphere which showed a fairly good fit with observational data (see, for instance, the mean annual zonal field of the temperature in the troposphere, according to the calculation of Kirichkov (1978), depicted in Figure 45.2). The temperature field in the figure is practically identical to that observed everywhere, except for the upper troposphere of the polar regions, which turned out to be somewhat warmer than the real field there.

The arbitrariness of the parametrization of the second moments can be reduced substantially if the Fridman—Keller equations (33.2) are used for these moments, and if, for instance, we neglect in them the third moments and supplement these with the semiempirical formulas (36.12), similarly to what was done in Section 36 for the Ekman boundary layer. Such an approach was used by Monin and Simuni (1980) to calculate the zonal circulation in the Sun's convective layer. As a greatly simplified illustration of this approach, we present the following nonstationary [unsteady-state] barotropic model of the zonal circulation of an atmosphere with

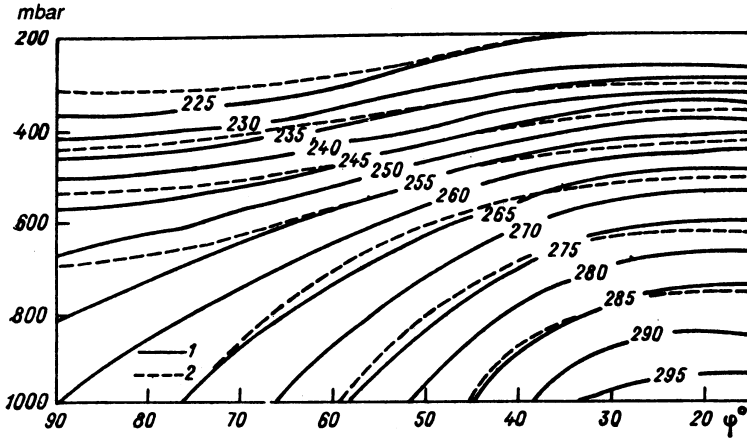


Fig. 45.2. Mean annual zonal field of temperature in troposphere: (1) actual, (2) according to calculation of Kirichkov (1978).

quasigeostrophic quasi-isotropic synoptic processes:

$$\begin{aligned} \frac{\partial \alpha}{\partial t} &= -\frac{1}{r_0^2 \sin^3 \theta} \frac{\partial}{\partial \theta} \overline{u'_\theta u'_\lambda} \sin^2 \theta; & \frac{\partial \overline{u'_\theta u'_\lambda}}{\partial t} &= -\overline{u'^2_\theta} \sin \theta \frac{\partial \alpha}{\partial \theta}; \\ \frac{\partial \overline{u'^2_\theta}}{\partial t} &= -\overline{u'_\theta u'_\lambda} \sin \theta \frac{\partial \alpha}{\partial \theta} \end{aligned} \quad (45.13)$$

(Monin, 1958). From this we see that $U^4 = (\overline{u'^2_\theta})^2 - (\overline{u'_\theta u'_\lambda})^2$ does not vary with time. These equations can be reduced to the following form:

$$\begin{aligned} \alpha &= \frac{\omega}{\sin^3 \theta} \frac{\partial \Psi}{\partial \theta}; & \overline{u'_\theta u'_\lambda} &= -\frac{\omega r_0^2}{\sin^2 \theta} \frac{\partial \Psi}{\partial t}; \\ \frac{\partial^2 \Psi}{\partial t^2} &= \left[\frac{U^4(\theta)}{r_0^4} + \left(\frac{\omega}{\sin^2 \theta} \frac{\partial \Psi}{\partial t} \right)^2 \right]^{1/2} \sin^3 \theta \frac{\partial}{\partial \theta} \frac{1}{\sin^3 \theta} \frac{\partial \Psi}{\partial \theta}. \end{aligned} \quad (45.14)$$

The latter equation describes nonlinear waves propagating along the meridian with a velocity $(\overline{u'^2_\theta})^{1/2}$.

46. CIRCULATION OF THE WORLD OCEAN

It is convenient to apply this term to the statistical ensemble of large-scale ocean currents, quasistationary as well as synoptic. Those quasistationary currents which

are, properly speaking, global probably include just the Circumpolar Antarctic Current (West Wind Drift) and the slow bottom currents which flow northward in the Antarctic and southward in the arctic waters of the Atlantic. The synoptic currents include eddies and Rossby waves, which, as mentioned in Section 37, in the ocean have horizontal scales of the order of $L_R \sim 50$ km, considerably smaller than in the atmosphere, and thus much longer periods $\tau \sim (\beta L_R)^{-1}$, of the order of a month.

The quasistationary circulation is seen to be very different in the upper layer of the ocean (in the top 1.5 km, say) and in the bottom layers. The upper circulation is well coordinated with the dynamic topography of the ocean surface (that is, with the variations in the height of this surface, reckoned from a depth layer with a fixed pressure, say 1.5×10^6 kg/m²): lines of divergence (convergence) of the surface currents correspond to the axes of the topographic troughs (crests).

The principal elements of this circulation are the gigantic anticyclonic gyres of water around the atmospheric subtropical 'centers of action': The Azores and St Helena in the Atlantic, Mauritius in the Indian Ocean, Hawaii and Easter Island in the Pacific. These are asymmetrical: their western branches are intensified and form strong narrow-stream boundary currents such as the Gulf Stream in the North Atlantic and the Kuroshio in the North Pacific.

In the tropical zone these gyres are set apart by trade wind currents flowing toward the equator and to the west (along the equator at depths of 50 to 300 m beneath the gyres, there are narrow-stream countercurrents, meandering slightly like type A_2B_2 of Section 32 with $\cos \beta > 0$). On their polar sides the gyres are bounded by cyclonic circulations, on the south by the largest current in the World Ocean: the West Wind Drift, with a volume transport of 210×10^6 m³/s.

The deep circulation is only weakly connected with the surface circulation (and with the wind field above the ocean). It appears to be primarily of thermohaline origin, that is, it is produced by the sinking and spreading out along the bottom (with a tendency to follow lines of constant depth) of water with a higher density (cooler and more saline water). For instance, antarctic bottom water in the Atlantic penetrates, mainly via the western troughs, as far north as 40°N, where it encounters North-Atlantic deep and arctic bottom water. In the Indian Ocean antarctic bottom water fills all the deep troughs, while in the Pacific its main flow proceeds along the Kermadec Islands and Tonga, with a branch flowing northeast located 10°S of it, whereas the main current in the Northern Hemisphere splits in two and flows along troughs nearly to the tropic.

Synoptic currents in the ocean can be divided, according to the mechanisms of their formation, into *frontal currents* (I), which are meanders on the stream currents and their end form of closed annular currents ('rings') and *free currents*, or eddies of the open ocean, caused by the baroclinic instability of the quasistationary currents (II), by topographic effects (III), and by the action of atmospheric synoptic processes (IV). Depending on the direction of rotation, these are divided into cyclonic and anticyclonic currents. It is convenient to distinguish between *eddies*, moving together with the water contained in them, and *waves*, moving over the water (including individual waves, or solitons). All these formations change considerably as they age.

For a quantitative description of the general circulation of the World Ocean we can use the same fluid-dynamical equations (45.1)–(45.3) and boundary conditions for the velocity field at the bottom $z = -h_w(\theta, \lambda)$ as in the case of the general circulation of the atmosphere. Here, by analogy with the atmospheric boundary layer, we will introduce the ‘parametrized’ near-bottom boundary layer. However, the kinematic boundary condition at the ocean surface $z = \zeta(\theta, \lambda, t)$ has to be written more precisely than in the case of the atmosphere: it must enable us to determine function ζ , which plays a major role in the fluid dynamics of the ocean. The following generalization of Equation (5.9) can serve as this condition:

$$u_z = \frac{\partial \zeta}{\partial t} + \frac{u_\theta}{r_0} \frac{\partial \zeta}{\partial \theta} + \frac{u_\lambda}{r_0 \sin \theta} \frac{\partial \zeta}{\partial \lambda} - \frac{1}{\rho_w} (P - E - E_{wi}), \quad (46.1)$$

where, in view of the smallness of ζ in large-scale currents (decimeters) the values of u_θ , u_λ , u_z can here be taken at $z = 0$ rather than at $z = \zeta$. The dynamic boundary conditions at the ocean surface reduce, with a high degree of accuracy, to requirements of continuity of the velocity, horizontal stresses, and pressure (during the consideration of large-scale currents capillary effects can, of course, be neglected). The condition of continuity of the pressure can with sufficient accuracy be written as

$$p = p_a + g\rho_w \zeta \quad \text{for } z = 0. \quad (46.2)$$

At the shores of the ocean the horizontal velocity of the currents must go to zero (everywhere except at the mouths of rivers, where the velocity normal to the shore depends on the mean flow velocity in the river; the discharge of small rivers and streams, as well as of icebergs, has to be ‘parametrized’; sometimes the entire coastal boundary layer is ‘parametrized’, neglecting in (45.2) the horizontal momentum transfer outside of it and requiring that only the velocity component normal to it go to zero at its inner boundary).

The equations of fluid dynamics must be supplemented by the equation of state of seawater (1.14), and in order to determine the two new functions T and s entering into it we must also invoke the equations of evolution of the entropy and salinity fields (45.4), (45.5) (for $m = 0$), as well as the formula for the entropy (1.2). The boundary condition at the ocean surface for Equation (45.4) will be in this case the heat-budget equation (45.6) (strictly speaking, with certain complications such as the dependence of the albedo A on the waviness of the sea and on the height of the Sun). For Equation (45.5) the boundary condition is the following salt-budget equation analogous to (45.7):

$$J_s = -s \left(P - E - \frac{q_{wi}}{\mathcal{L}_{wi}} \right), \quad (46.3)$$

where J_s is the salt flux into the ocean depths. At the bottom of the ocean and at its shores the normal components of the heat and salt fluxes have to equal zero (everywhere except in hydrothermal regions of the bottom, at the mouths of rivers, and at sites where icebergs break away).

We note the usefulness of calculating the combination $M_s = J_s + (\alpha/c_w) q_s$ of the heat and salt fluxes (45.6) and (46.3), where α is the coefficient of thermal expansion and c_w is the specific heat of water: this combination can be construed as the vertical mass flux at the ocean surface. At places where it is positive, that is, the mass flux is downward, the buoyancy of the surface water is reduced and it sinks; in other words, a thermohaline circulation is set up (Monin, 1970; an empirical chart of the mean annual values of M_s was drawn up by Agafonova *et al.*, 1972, and Galerkin *et al.*, 1985, provided illustrations of the very substantial seasonal variability of the M_s sources of the ocean's thermohaline circulation).

The boundary conditions at the ocean surface will be very different, of course, if this surface is covered with ice. The freezing and melting of ice can be calculated locally with the aid of the equation of vertical heat transfer:

$$c\rho \frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left(c\rho k \frac{\partial T}{\partial z} + F_0^i \exp \left[\int_0^z \kappa \, dz \right] \right), \quad (46.4)$$

where F_0^i is the flux of short-wave solar radiation passing through the upper surface $z = 0$ of the snow and ice layers (their thicknesses are taken to be h_s and h_i), and $\kappa = \kappa(z)$ is the coefficient of absorption of the radiation. The specific heats c , densities ρ , molecular thermal diffusivities k , and values of κ are all different for ice and snow, and at the boundary between the ice and snow $z = -h_s$ there must be continuity of the temperature and the heat flux (described by the expression in parentheses on the right-hand side of (46.4)). In boundary condition (45.6) at $z = 0$ the left-hand side must be equal to

$$q_s = \left(c\rho k \frac{\partial T}{\partial z} \right)_{z=0} + F_0^i.$$

Variations in the thickness of the snow layer are described by the equation

$$\frac{\partial \rho_s h_s}{\partial t} = P_s - E + \frac{q_{wi}}{\mathcal{L}_{wi}}, \quad (46.5)$$

which is valid for $h_s > 0$ (where ρ_s is the density of the snow, and P_s is the rate of snowfall). At temperatures below the freezing point of seawater (which is often taken to be 271.2 K) in (45.6) we have to set $q_{wi} = 0$, and at the freezing point condition (45.6) determines the value of q_{wi} . Finally, at the lower boundary of the ice $z = -(h_s + h_i)$ the difference between the heat fluxes in the ice and in the water determines q_{wi} , and thus the ratio $q_{wi}/\mathcal{L}_{wi}$ as well, defining in water equivalent the rate of freezing or melting of ice at the bottom (sometimes for purposes of simplification a standard value is assumed for the heat flux from the water to the ice: 2 W/m² in the Arctic and 25 W/m² in the Antarctic).

The large-scale horizontal motion of sea ice, with variations in the ice thickness h_i due to the formation of hummocks or polynyas, can be described by the kinetic

equation

$$\frac{\partial \Phi}{\partial t} + \frac{1}{r_0 \sin \theta} \left(\frac{\partial \Phi u_\theta \sin \theta}{\partial \theta} + \frac{\partial \Phi u_\lambda}{\partial \lambda} \right) + \frac{\partial \Phi \varepsilon_i}{\partial h_i} = \Psi, \quad (46.6)$$

where $\Phi(h_i; \theta, \lambda, t)$ is the probability density for h_i (so that $\int_{+0}^{\infty} \Phi \, dh_i$ is the *compactness* of the ice cover, that is, the portion of the surface occupied by it); $\varepsilon_i = \partial h_i / \partial t$ is the rate of freezing or melting of the ice, determined with the aid of Equation (46.4); Ψ is the redistribution function, describing the formation of hummocks or polynyas. Velocity components u_θ, u_λ must be found from the equations of fluid dynamics (45.1)–(45.3), taking into account the sum of the wind-friction forces at the upper surface of the ice cover and the relative flow of water at its lower surface, the sea-level inclinations, and the internal forces of interaction between ice floes (that is, components F_θ and F_λ of the divergence of the horizontal-stress tensor for the specified rheology of the ice cover).

Using such simplified equations, we are able to explain qualitatively a number of the specific features of the general circulation of the World Ocean. The features produced by the presence of meridional ‘walls’ (shores) are the most specific. It was stated above that these consist in, first of all, an east–west asymmetry of the huge subtropical anticyclonic gyres of water, with an intensification of the westerly boundary currents. As first noted by Stommel (1948), this feature can be attributed to the ‘beta effect’, that is, to the latitude variation $\beta = \partial f / \partial y$ of the Coriolis parameter f .

In order to show this, let us first apply to the ocean the theory of the Ekman boundary layer presented in Section 36. From the condition of continuity of the vertical momentum flux at the ocean surface $\tau = \rho_a u_{*a}^2 = \rho_w u_{*w}^2$, for a typical atmospheric value $u_{*a} \sim 30$ cm/s we obtain $u_{*w} = (\rho_a / \rho_w)^{1/2} u_{*a} \sim 1$ cm/s. The thickness of the Ekman boundary layer in the ocean turns out to be of the order of $h_0 = \kappa u_{*w} / f \sim 20$ m, while the time of reaction of the currents in the Ekman layer to the wind effects is $\rho_w h_0 u_{av} / \tau \sim 3$ hours (where $u_{av} = u_{*w} / \kappa$ is the mean value, over the thickness of the Ekman layer, of the velocity of the drift current). This time is short compared to the time of synoptic variations of the wind. Consequently, the currents in the Ekman layer can be considered quasistationary, and the Reynolds equations (36.10) for the Ekman boundary layer will have the form:

$$\begin{aligned} -fv &= -g \frac{\partial \zeta}{\partial x} + \frac{1}{\rho_w} \frac{\partial \tau_{xz}}{\partial z} = -g \frac{\partial \zeta}{\partial x} + \frac{\partial}{\partial x} \mathcal{K} \frac{\partial u}{\partial z}; \\ fu &= -g \frac{\partial \zeta}{\partial y} + \frac{1}{\rho_w} \frac{\partial \tau_{yz}}{\partial z} = -g \frac{\partial \zeta}{\partial y} + \frac{\partial}{\partial y} \mathcal{K} \frac{\partial v}{\partial z}, \end{aligned} \quad (46.7)$$

where on the right-hand sides we used the first two relations of (36.19). For a qualitative analysis, we set $\mathcal{K} = \text{const}$ (a more complete theory was presented in Section 36); then the solution of Equations (46.7) for an ocean of constant depth

h becomes

$$\begin{aligned}
 u + iv = & h \frac{\tau_{x0} + i\tau_{y0}}{\rho_w \mu \mathcal{K}} \frac{\operatorname{sh} \mu (z' + 1)}{\operatorname{ch} \mu} + i \frac{g}{f} \left(\frac{\partial \xi}{\partial x} + i \frac{\partial \xi}{\partial y} \right) \times \\
 & \times \left(1 - \frac{\operatorname{ch} \mu z'}{\operatorname{ch} \mu} \right) \approx h \frac{\tau_{x0} + i\tau_{y0}}{\rho_w \mu \mathcal{K}} e^{\mu z'} + \\
 & + i \frac{g}{f} \left(\frac{\partial \xi}{\partial x} + i \frac{\partial \xi}{\partial y} \right) [1 - e^{-\mu(z'+1)}], \tag{46.8}
 \end{aligned}$$

where $z' = z/h$, $\mu = (1 + i)/E$, and $E = h_E/h$ is the Ekman number (where $h_E = (2\mathcal{K}/f)^{1/2}$ is a new vertical scale of the Ekman layer; for $h_E \sim 20$ m and $f \sim 10^{-4} \text{ s}^{-1}$ we get $\mathcal{K} \sim 2 \times 10^{-2} \text{ m}^2/\text{s}$), which is small, on which result we also based the approximate equation in (46.8). Function $\xi(x, y)$ has to be found from the equation

$$w(z = -h) = \int_{-h}^0 \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dz = 0 \tag{46.9}$$

for the appropriate boundary conditions at the ocean shores. Equations (46.7) enable us to require at the shores nonpassage only on the average over the depth, that is, a zero normal component of the so-called *total flow* $S = \int_{-h}^0 \mathbf{u} \, dz$ (first introduced by Shtokman, 1946):

$$\begin{aligned}
 S_x + iS_y = & \frac{\tau_{x0} + i\tau_{y0}}{if\rho_w} \left(1 - \frac{1}{\operatorname{ch} \mu} \right) + \frac{igh}{f} \left(\frac{\partial \xi}{\partial x} + i \frac{\partial \xi}{\partial y} \right) \times \\
 & \times \left(1 - \frac{\operatorname{th} \mu}{\mu} \right) \approx \frac{\tau_{x0} + i\tau_{y0}}{if\rho_w} + \\
 & + \frac{igh}{f} \left(\frac{\partial \xi}{\partial x} + i \frac{\partial \xi}{\partial y} \right) \frac{\mu - 1}{\mu}. \tag{46.10}
 \end{aligned}$$

Note that, because of the nondivergence of the three-dimensional velocity and the zero boundary conditions for w , the total flow is also nondivergent, that is, $\partial S_x / \partial x + \partial S_y / \partial y = 0$. Therefore, we can introduce the stream function ψ , with $S_x = -\partial \psi / \partial y$ and $S_y = \partial \psi / \partial x$. Now, expressing from (46.10) $\nabla \xi$ in terms of ψ , we can with an accuracy of the order of E reduce the condition $\operatorname{curl} \nabla \xi = 0$ to the form:

$$\frac{fh_E}{2h^2} \Delta \psi + \frac{\partial (\psi, f/h)}{\partial (x, y)} = \operatorname{curl} \frac{\tau_0}{\rho_w h}. \tag{46.11}$$

The boundary condition at the shore can now be written as $\psi = \text{const}$ (if the ocean is multiply connected, then the values of ψ on each closed boundary contour must be found from the condition of uniqueness of function ξ).

Let us consider first an ocean of constant depth ($h = \text{const}$). In its open part the first term in (46.11), describing the effect of friction, can be neglected, so that we have

$$\beta \frac{\partial \psi}{\partial x} = \text{curl} \frac{\tau_0}{\rho_w} \quad (46.12)$$

(Sverdrup, 1947). This shows that ψ should be measured on a scale $\tau_0/\rho_w\beta$. By measuring, moreover, the lengths on a scale $f/2\beta$, we can put (46.11) in dimensionless variables into the form

$$E\Delta\psi + \frac{\partial \psi}{\partial x} = F(x, y), \quad (46.13)$$

where the right-hand side can be assumed to be given. Let us consider a model ocean with a square shape $0 \leq (x, y) \leq 1$ and let us seek in it a solution of Equation (46.13) in the form of a series $\psi = \sum_{n=0}^{\infty} \psi_n E^n$, which must satisfy the boundary conditions $\psi = 0$ on all sides of the square. It is impossible to assume in this case that all the functions ψ_n , $\partial\psi_n/\partial x$, and $\partial\psi_n/\partial y$ are of the order of unity, since then for ψ_n we would obtain first-order equations $\partial\psi_n/\partial x = F_n - \Delta\psi_{n-1}$ and it would be impossible to satisfy both boundary conditions in x at $x = 0$ and $x = 1$.

Therefore, in the vicinities of these shores we have to introduce boundary layers with 'fast' transverse coordinates $x' = x/E$ and $x'' = (1 - x)/E$, and we have to seek a solution in the form $\psi_n = \psi_n(x, y) + \psi'_n(x', y) + \psi''_n(x'', y)$, where functions ψ_n , ψ'_n , and ψ''_n differ considerably from zero, respectively, in the open ocean, in the western boundary layer, and in the eastern boundary layer. This means that functions ψ'_n must go to zero rapidly with increasing x' , and ψ''_n must go to zero rapidly with increasing x'' . From (46.13) we obtain

$$\frac{\partial \psi_0}{\partial x} = F_0; \quad \frac{\partial^2 \psi'_0}{\partial x'^2} + \frac{\partial \psi'_0}{\partial x'} = 0; \quad \frac{\partial^2 \psi''_0}{\partial x''^2} - \frac{\partial \psi''_0}{\partial x''} = 0. \quad (46.14)$$

This shows that a solution ψ''_0 that decreases rapidly as x'' increases does not exist, so that we must set $\psi''_0 = 0$, whereas a solution $\psi'_0 = c e^{-x'}$ is valid. Consequently, there is an east—west asymmetry of the current: the boundary conditions in x take the form $\psi_0 + \psi'_0 = 0$ for $x = 0$ and $\psi_0 = 0$ for $x = 1$; then the solution of Equation (46.13) not too close to the walls $y = 0$ and $y = 1$ can (accurate to a term $O(E)$) be written as $\psi = -(1 - e^{-x/E}) \int_0^1 F_0 dx$ in the boundary layer $x \leq E$, $0 < y < 1$, and as $\psi = -\int_x^1 F_0 dx$ in the main part of the ocean $E \leq x \leq 1$; see in Figure 46.1 the example of the streamlines $\psi = \text{const}$, calculated for $\tau_{x0} = -\cos \pi y$, $\tau_{y0} = 0$.

Therefore, at the western shore there will be a strong boundary current, in which $S_x = O(1)$, $S_y = O(1/E)$, while at the eastern shore no such solution exists; this asymmetry can be traced to the second term in (46.13), that is, to the 'beta effect'. The total volume transport of the current through a circle of latitude $y =$

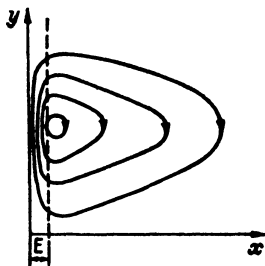


Fig. 46.1. Solution of Equation (46.10) for $\tau_{x0} = \cos \pi y$, $\tau_{y0} = 0$.

const from the western shore $x = x_w(y)$ to the eastern shore $x = x_E(y)$ is equal to $Q = \int_{x_w}^{x_E} S_y dx$, where $S_y = \partial\psi/\partial x$ can be found from Sverdrup's formula (46.12). Outside of the western boundary layer it would be found to be nonzero and a function of y , which would contradict the law of conservation of the mass of water. Consequently, it must be compensated by the transport Q_w of the western boundary layer, that is, $Q_w = -Q$ or

$$Q_w = -\frac{1}{\beta} \int_{x_w}^{x_E} \text{curl} \frac{\tau_0}{\rho_w} dx \quad (46.15)$$

(Munk, 1950). Calculations made using this formula show a fairly good fit with empirical estimates of the volume transport of the Gulf Stream southward of the point of its separation from the shore 33°N .

In an ocean of variable depth, neglecting in its open part the first term in (46.11), we arrive at a generalized Sverdrup relation, in which the role of the circles of latitude $f = \text{const}$ are played by isolines $f/h = \text{const}$. Here a western boundary current exists in the case $(\partial/\partial s)(f/h) > 0$, while an eastern boundary current exists if $(\partial/\partial s)(f/h) < 0$, where ds is directed along the tangent to the shoreline toward the north. If isoline $f/h = \text{const}$ at some point is tangent to the shore, then there $(\partial/\partial s)(f/h)$ changes sign, that is, the boundary current separates from the shore because of the bottom relief.

For a more precise calculation, on the right-hand sides of Equations (46.7) of the Ekman theory we have to add the terms $\mathcal{K}_h \Delta_h u$ and $\mathcal{K}_h \Delta_h v$ describing the horizontal mixing. Let the horizontal Ekman number $E_h = (2\mathcal{K}_h/f)^{1/2}/L$ have the same order of magnitude as the vertical number $E = (2\mathcal{K}/f)^{1/2}/h$. Then the following will be true: (1) in the open ocean u, v, w, ζ have values of the order of E ; (2) a surface Ekman boundary layer exists with a thickness $O(E)$, in which $u, v \sim O(1)$ and $w \sim O(E)$; (3) a thin bottom Ekman layer exists of thickness $O(E)$, in which $u, v \sim O(E)$ and $w \sim O(E^2)$; (4) near the western shore there is a Munk boundary layer of thickness $O(E^{2/3})$, in which $u \sim O(E)$, $v \sim O(E^{1/3})$, $w \sim O(E^{2/3})$, and $\zeta \sim O(E)$; (5) inside this layer, and also near the eastern shore, there are sublayers of thickness $O(E)$, in which $u, v \sim O(E)$, $\zeta \sim O(E^2)$, and $w \sim O(1)$; (6) boundary layers also exist at the northern and southern shores.

For an even more complete treatment, in the equations of horizontal motion (45.2) we retain the nonlinear (inertial) terms. In order to explain the *horizontal* structure of the currents, following Shtokman, we limit ourselves to these equations, approximately averaged over the entire thickness of the ocean $-h \leq z \leq 0$. For $h \approx \text{const}$ these yield the following equation for the relative vorticity $\Omega = \text{curl } \bar{\mathbf{u}}_h$:

$$\bar{\mathbf{u}}_h \cdot \nabla(\Omega + f) = \mathcal{K}_h \left(\Delta_h \Omega + \frac{2\Omega}{r_0^2} \right) + \frac{\text{curl } \boldsymbol{\tau}_0}{\rho_w h}. \quad (46.16)$$

Since $\text{div } \bar{\mathbf{u}}_h = 0$, the integral over any closed contour of the left-hand side of this equation equals zero, while the integral of the second term on the right-hand side does not have to be zero (for instance, in regions of atmospheric anticyclones $\text{curl } \boldsymbol{\tau}_0 < 0$), so that it is impossible, in general, not to take into account the horizontal mixing. In the open ocean we can neglect the advection of vorticity $\bar{\mathbf{u}}_h \cdot \nabla \Omega$ and horizontal mixing, and (46.16) becomes the Sverdrup relation $\beta \bar{u}_\theta = -\text{curl } \boldsymbol{\tau}_0 / \rho_w h$, describing a southward current with a volume transport depending on the latitude. It should be compensated by coastal currents toward the north. In these the vorticity advection and horizontal friction are already substantial, so that near the meridional shores inertial-viscous boundary layers form. The theory of these was formulated by Kamenkovich (1966).

This theory was based on the fact that in the boundary layers the total advection of the relative vorticity $\int v \Omega \, dx$ and the contribution of the term $2\mathcal{K}_h \Omega / r_0^2$ are small. This means that the advection of the vorticity of the Earth's rotation βv must be compensated by the diffusion of relative vorticity from the wall. At the western shore this takes place, since positive vorticity both is transported northward and diffuses from the wall. Near the eastern shore this does not take place in a northward current, since negative vorticity diffuses out from the wall, and in a southward current it is possible only because of the term $\text{curl } \boldsymbol{\tau}_0 / \rho_w h$, which can produce only a weak current.

In order to calculate the boundary layers, it is convenient to change to dimensionless variables in the equations of fluid dynamics, measuring the lengths to a scale L and the velocities to a scale $U = \tau_0 / \rho_w \beta L h$. In these variables the inertial terms in the equations will have a small multiplier $\varepsilon^2 = (L_i/L)^2$, where $L_i = (U/\beta)^{1/2}$ (so that ε^2 is identical to the Kibel' number), while the viscous terms have a multiplier ε^3/Re , where $\text{Re} = UL_i/\mathcal{K}_h$ is the Reynolds number. The solutions of the dimensionless equations are sought in the form of the same series as in the case of Equation (46.13) (except that E is replaced by ε).

As a result it is seen that for large Re at the western shore in the southern half of a basin (where $\partial^2 \tau_0 / \partial y^2 > 0$) there will exist an inertial boundary layer with a viscous near-shore sublayer; in the northern half of a basin (where $\partial^2 \tau_0 / \partial y^2 < 0$), on the other hand, such a layer already does not exist for $\text{Re} \geq 10$ (which apparently indicates the separation of the boundary layer from the shore). For low Re , however, only a viscous boundary layer is present, in which inertial effects are insignificant.

The above-indicated equations made it possible to construct numerical models

of the general circulation of the World Ocean *for specified atmospheric effects*. First, stationary (barotropic) models of the mean annual total flows (Sarkisyan, 1954, model of North Atlantic; Welander, 1959, model of World Ocean between 50°N and 40°S; Il'in *et al.*, 1969, model of World Ocean) were constructed; we should also mention the nonstationary barotropic model of seasonal variability of the World Ocean, devised by Kagan *et al.* (1972).

Next a number of investigators constructed three-dimensional (baroclinic) models of the ocean circulation. Among the first to do this were Bryan and Cox (1967–1969), who computed the mean annual circulation of the Atlantic (modeled in the shape of a parallelepiped). In their models at the ocean surface T and s , or the fluxes of heat and salt, were specified, all the vertical and horizontal exchange coefficients were taken to be constant, and the procedure of hydrostatic adjustment was applied. Of the various later models, we should cite that of Cox (1970) on the seasonal (monsoon) variability of the circulation of the Indian Ocean for assumed values of the wind frictional stress, temperature, and salinity at the ocean surface; the two-layer model of the World Ocean with real geometry of Andreev *et al.* (1976); the six-level models of Marchuk and Zalesnyi (1974) and Kochergin and Shcherbakov (1977); and, finally, the 12-level model of Bryan *et al.* (1975), used by Bryan and Lewis (1979) to calculate (for 1150 years) the state of the ocean for specified atmospheric effects (with a seasonal variation).

The numerical models constructed more recently were based on integration of the nonstationary equations by steps in time from the initial state of rest to the establishment of a quasistationary state of the ocean (for which a long time $h^2/\mathcal{K}_z$, of the order of at least 10^3 years, is needed). Here in the global models the step on the horizontal grid was taken to be, say, 500 km, so that synoptic eddies were not described individually. The horizontal mixing produced by these is traditionally parametrized with the aid of operators of the form $\mathcal{K}_h \Delta_h$ with excessively high, positive coefficients $\mathcal{K}_h$ (values up to 10^8 cm²/s), whereas in order to form the observed narrow-stream major ocean currents we must invoke a 'negative viscosity' (without the latter, the models of the ocean circulation would not be sufficiently realistic).

47. EDDY-RESOLVING MODELS

In order to circumvent the difficulties involved with an *a priori* specification of the 'negative viscosity', it is desirable to have numerical models of the ocean circulation, allowing an individual description of all synoptic eddies. To accomplish this, we must ensure that the step of the horizontal grid is in any case not larger than the Rossby deformation radius L_R , that is, no greater than some tens of kilometers. Such numerical models are called *eddy-resolving models*.

Clearly, calculations of the circulation in a very large water body (and over a long time interval) with the aid of such models call for very high-powered computers. Therefore, the first experiments were carried out only for limited ocean regions, squares or rectangles 1000 to 2000 km in size, with a flat bottom and with some artificial boundary conditions specified. During the past decade, on the other hand, it has already become possible to devise eddy-resolving models of

the circulation in the real ocean (the first example was the North Atlantic), and it now appears that an eddy-resolving model of the entire World Ocean is right around the corner.

The first eddy-resolving model was constructed by Holland and Lin (1975). It was a *two-layer* model (with respect to the water density) employing the usual equations of fluid dynamics (the so-called 'primitive equations', or 'initial equations', without filtering out internal waves via the quasigeostrophic or quasisolonoidal approximation). It was integrated over a 1000×1000 km square on a grid of 51×51 points (the grid interval of 20 km was doubled or halved in some of the experiments) for a constant ocean depth of 4 km (the upper and lower layers were 1 and 3 km thick, respectively). Only the wind activation (steady zonal wind) was taken into account, without differential heating. The wind was so selected as to produce one or two large-scale horizontal gyres.

Numerical experiments using this model indicated that for steady wind activation a nonstationary circulation with large-scale gyres and synoptic eddies is created (with a size $L \approx 400$ km and a period $\tau \approx 64$ days), predetermined by the specified $L_R = (gh\delta\rho/\rho_0)^{1/2} f^{-1}$, where $h = h_1 h_2 (h_1 + h_2)^{-1}$. At first (for about 500 days) there is an increase in the kinetic energy of the gyres $\bar{K}$ and the available potential energy $\bar{\Pi}$, related to the bending of the layer interface caused by the currents in the upper layer. When some critical value $\bar{K}$ is reached, the gyres become baroclinically unstable as a result of the large velocity shear between layers, so that $\bar{\Pi}$ drops; then, in the form of K' and Π' , eddies are turned loose in both layers (currents being induced in the lower layer by the eddies of the upper layer). After a transition period (amounting to 500 to 600 days) statistical equilibrium is established between the gyres and the eddies.

In the case of a single gyre, an intensified western current directed northward changes at the northern wall into a strong eastward current and then into a return current in the eastern and southern regions. At the northern wall and in the return current barotropic eddies appear, and in the return current sometimes baroclinic synoptic eddies also form. A two-ring circulation with a meandering middle jet shows both baroclinic and barotropic instability.

The weakness of the fluctuations of the interface between layers and the slowness of the meandering of the middle jet, compared to what is observed in nature, constituted a quantitative shortcoming of these results. Analogous findings were yielded by the two-layer quasigeostrophic model of Holland (1978), who added bottom friction to his model and who described the horizontal turbulent exchange with the aid of a biharmonic operator instead of the usual Laplacian; this artificially accentuated the synoptic eddies.

More complete (and essentially identical) multilayer eddy-resolving models were constructed by Hahn (1975) and by Robinson *et al.* (1977). In these models both the wind activation and the thermal activation were calculated (the heat flux through the ocean surface was assumed to be proportional to $T_w - T_a$). The length of the basin was about 2000 km, the step along the horizontal was 20 to 35 km, the time interval was 20 min, and the derivatives were approximated as differences, correct to terms of second order in time and space. An eddy-resolving model similar in many respects was constructed by Seidov (1978, 1980). Somewhat

different models were shortly after devised to describe eddies in the open ocean (with fluid boundaries), by Kamenkovich *et al.* (1981), as well as by Neelov and Chalikov (1981).

Let us consider the models of Seidov in more detail. He took into account in the equations the horizontal and vertical eddy viscosity and eddy thermal conductivity (with ρ depending only on T) and he specified at the ocean surface the zonal wind stress (which produced in a rectangular basin one or two large-scale gyres) and the zonal heat flux (proportional to $T_w - T_a$). In order to avoid calculating small differences of large numbers, and also to increase the time interval by filtering out gravitational and inertial fluctuations, Seidov distinguished the mean parts (over the depth) of the fluid-dynamical fields from the shear parts, and he assumed the quasigeostrophic approximation for the shear velocities. As the evolution ('prognostic') equations, he retained only the equations for the vorticity of the total flow and the temperature.

Moreover, at the horizontal walls of the basin slipping was assumed, and at the bottom (of constant depth) linear friction (that is, proportionality of the vertical flux of vorticity to the vorticity itself). The equations were solved numerically using a Eulerian—Lagrangian method for splitting the advective and diffusion operators. The principal numerical experiments were performed in a basin measuring $960 \times 1440 \times 5$ km in the x, y, z directions with $25 \times 37 \times 5$ nodes on the space grid and a time interval of about 6 hours over a period of 10 years (although other variants were calculated as well).

Among his conclusions Seidov cited the following: a substantial role is played by inertial effects for synoptic eddies (in comparison with large-scale gyres); the energy-conversion processes are spatially nonuniform (negative viscosity plays an especially important part in accentuating the jet currents, see formulas (43.1)—(43.3)); eddies cause a "tunnel effect" of heat transfer through zonal jets, as well as concentration of heat in the western boundary currents.

The simplified (mean over the volume) energy cycle of the system of large-scale gyres (currents averaged over time) and synoptic eddies (deviations from the mean current), obtained by Seidov, are shown in Figure 47.1(a). This cycle has the form $\bar{K} \rightarrow \Pi \rightarrow K' \rightarrow \bar{K}$, the feedback $K' \rightarrow \bar{K}$ being relatively weak only in the mean over the volume, whereas in the zones of the jet currents it plays a decisive role. Figure 47.1(b) shows the energy cycle of the system of total flows (mean over the depth) and shear currents (deviations from the mean over the depth). It has the form $K' \rightarrow (\Pi, \bar{K})$, where both Π and $\bar{K}$ receive from K' (that is, as a result of barotropization) considerably more energy than from external sources. Therefore, the system is far from conservative: energy enters into it from outside, passes through it, and becomes dissipated.

In such systems relaxation oscillations may appear, with periods of accumulation and subsequent release of energy. Actually, in the numerical experiments of Seidov it was found that, after the system reached a state of statistical equilibrium, self-oscillations appeared in it, with a sequence of long periods of weak eddy activity, a growth of Π , and a drop in K , and shorter periods (by a factor of 1.5 to 2.5) of strong eddy activity, a drop in Π , and an increase in K .

As an illustration Figure 47.2 shows plots of the oscillations of K and of the

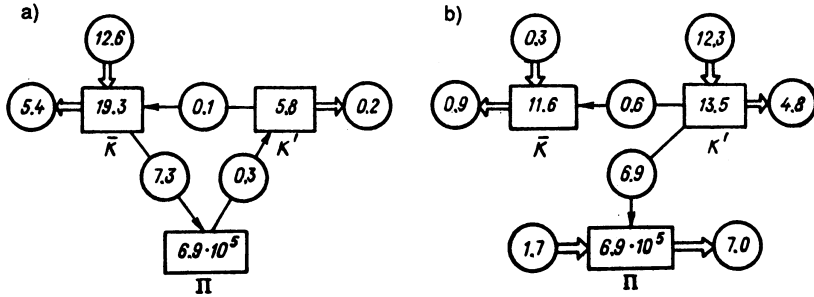


Fig. 47.1. Energy cycles in eddy-resolving model of Seidov: (a) in system of large-scale gyres and synoptic eddies; (b) in system of total flows and shear currents. Energy density in 10^{-1} J/m^3 ; density of energy conversions in $10^{-6} \text{ J/(m}^3 \text{ s)}$.

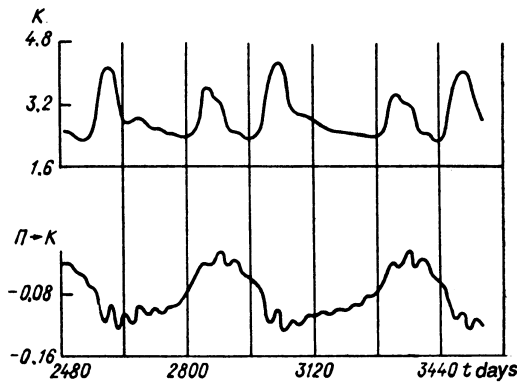


Fig. 47.2. Self-oscillations in eddy-resolving model of Seidov for $\Delta x = 65 \text{ km}$.

rate of energy conversions $\Pi \rightarrow K$ for days 2480 to 3500 of a numerical experiment with a large horizontal grid interval (65 km). The period of oscillations $\Pi \rightarrow K$ is here about 420 days (300 days of increase and 120 days of decrease), and it is interesting that K oscillates with twice the frequency: the smaller max K corresponds to maximum $\Pi \rightarrow K$, and the larger max K to the minimum. A smaller grid interval (40 km) is used to resolve smaller eddies, and the spectrum of the self-oscillation periods becomes more complicated (see Figure 47.3). This pattern is very similar to that actually observed in nature during the POLYMODE experiment in the Sargasso Sea (Figure 47.4).

When eddy-resolving models are used for calculations in regions of the open ocean, that is, for the synoptic forecasting of ocean currents, the normal component of the velocity on the entire boundary has to be specified, as well as the prognostic quantities which are quasiconserved along the trajectories of the fluid particles (in Seidov's model these are the vorticity of the total flow and the

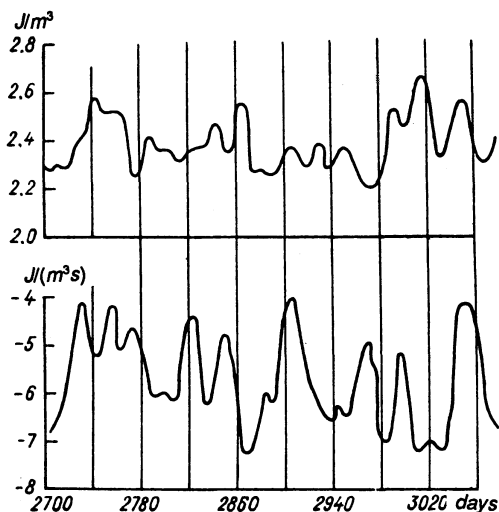


Fig. 47.3. Self-oscillations in eddy-resolving model of Seidov for $\Delta x = 40$ km.

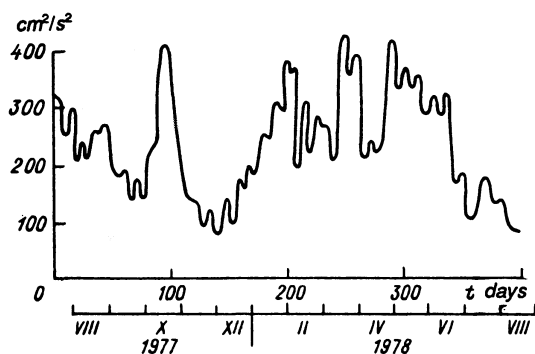


Fig. 47.4. Oscillations of kinetic energy of eddies at depth of 100 m during POLYMODE experiment in Sargasso Sea.

temperature) on those parts of the boundary where fluid enters the region (Charney *et al.*, 1960).

These difficulties can be circumvented only by applying the eddy-resolving model to the entire ocean. The first to do this were Seidov *et al.* (1986), who applied the above-described model of Seidov to the North Atlantic between 13 and 61°N with a real configuration of the coasts and bottom relief and assuming nonpassage and slipping at all boundaries. The model contained seven reference levels (at depths of 0, 200, 500, 800, 1200, 2000, and 3000 m) and a $40 \times 40'$ horizontal grid (for comparison, an 'eddy-free' model on a $2 \times 2^\circ$ grid was also calculated). The time interval was 12 hours, and the calculations were carried out for a five-year period.

The results demonstrated that active eddy production occurs in the region of the Gulf Stream and in the southeastern part near the center of the waters of Mediterranean origin (as well as in the water flowing over the Mid-Atlantic Ridge). The eddies led to an intensification of the jet currents (by a factor of 1.5 to 2 in the Gulf Stream and a factor of 3 in the Labrador Current) and also to a sharpening of the temperature gradients.

The World Ocean can be divided into 11 regions: the northern, tropical, and southern parts of the three oceans, the West Wind Drift, and the Arctic, for each of which an eddy-resolving model can be constructed according to a specified pattern. Then it remains only to match up these 11 models.

Theory of Climate

48. CONCEPT OF CLIMATE

Let us consider planets possessing an *atmosphere* (A), that is, an outer gaseous or gas—liquid envelope bounded at the bottom by a solid or liquid *underlying surface* (US). The layer of liquid and/or solid matter beneath this surface, the states of which change as a result of the effect of the atmosphere and which can, in turn, vary the states of the atmosphere, is called the *active layer* (ALUS). Consequently, A and ALUS together constitute a single variable global system, which we will call the *climate system*.

On Earth the active layer comprises the entire World Ocean and the active layer of the land surface, as well as glaciers which may form on it, so that the climate system is the atmosphere—ocean—land system (sometimes the cryosphere and biosphere are considered separately). On Mars the activity of the active layer manifests itself in the growth and melting of the polar ice caps and the permafrost around them, as well as in the raising of dust during dust storms. On the giant planets the effect of their solid cores on the behavior of their fluid atmospheres is apparently secondary, although this effect probably increases from Jupiter to Neptune. There is no active layer of the underlying surface on the Sun.

The A—ALUS climate system is global, but it possesses internal spatial inhomogeneities of different scales (ranging, say, from fractions of a millimeter all the way up to planetary dimensions), vertical as well as horizontal. When studying the climate system as a whole, it is, of course, impossible to describe all these inhomogeneities individually. The definition of the *states of the climate system* must include the individual characteristics of the large-scale inhomogeneities (global and synoptic) and only the statistical characteristics of the meso-scale and small-scale inhomogeneities (which should be ‘parametrized’, that is, expressed in terms of the individual characteristics of the large-scale inhomogeneities).

The inhomogeneities of each scale vary in time with periods typical of these inhomogeneities. The variations of the states of the climate system have both *external periods*, which are astronomically caused, and *internal periods*, that is, periods of its possible self-oscillations which are intrinsic to the system. In order to arrive at a definition of climate, we will assume that both the external and internal periods *naturally separate* into short and long periods. This means that the short-period oscillations and the long-period oscillations are both intense, but that there is an interval of intermediate periods within which the oscillations are

either of low intensity or completely absent. At any rate, this is how it is on the Earth.

The short astronomical periods include the *diurnal oscillations of the insolation* caused by the rotation of the planet, the insolation being the influx of solar radiation. On Venus, as mentioned above, this period is 117 days; on the distant planets its role diminishes rapidly with increasing distance from the Sun. Other short-period oscillations are *tidal oscillations* due to the gravitational effects of satellites and the Sun. On the Earth these are mainly the principal semidiurnal and diurnal lunar tides M_2 and O_1 , with periods of about 12 h 25 min and 24 h 50 min, and the principal semidiurnal and diurnal solar tides S_2 and P_1 . Finally, a third kind of short-period oscillation is the *seasonal oscillations* of the insolation, associated with the revolution of the planet around the Sun, the inclination of its equator to the ecliptic, and the eccentricity of the orbit.

The short internal periods include *synoptic oscillations*, caused by Rossby waves and eddies, and also *global oscillations*. On the Earth the latter have periods ranging from weeks to months; they include the approximately two-week oscillation cycle of the zonal circulation of the atmosphere (45.13) and (45.14), self-oscillations in the atmosphere—ocean system having periods of the order of months, and self-oscillations of the eddy-formation type (see Figure 47.2). A third kind of short-period internal oscillation has a period measured in years. These *year-to-year oscillations* mostly have periods of 2 to 5 years on the Earth; they include, among others, the quasibiennial cyclicity of the atmospheric circulation mentioned in Section 6, the El Niño phenomenon in the eastern equatorial part of the Pacific, and the 3.5-year self-oscillations of the north branch of the Gulf Stream.

Long astronomical periods are caused by, firstly, perturbations of the motion of the planet due to the pull of the other planets. These produce *conditionally periodic oscillations* of the inclination ε of the equator to the ecliptic, of the eccentricity e of the orbit, and of the longitude Π of its perihelion (in a system of n planets these oscillations have $3n - 1$ independent frequencies). On the Earth the predominant astronomical periods are 100,000, 41,000, 23,000, and 19,000 years, and it is these which show up in the oscillation spectra of a number of climatic indicators for the glacial periods of the Pleistocene (Hays *et al.*, 1976). On Mars the predominant astronomical periods are 117,000 and 52,000 years, and the corresponding oscillations of the insolation are by no means small; at the summer poles they can be 32.5% of the mean insolation, whereas on the Earth they are at the most 5% (Monin, 1979).

Existing historical data about the Earth indicate that oscillations in the atmosphere—ocean—land system with intermediate periods of tens of years (*intrasecular oscillations*) or hundreds of years (*intersecular oscillations*, including the so-called ‘Little Ice Age’ of the seventeenth to nineteenth century) are on the average considerably less intense than both the short-period oscillations mentioned above and the long-period oscillations (periods of thousands or tens of thousands of years) detectable in the formation, growth, and breakup of the continental ice sheets. This is probably also true of Mars, which is apparently now in a state of glaciation (with widespread permafrost).

Let us specify at the short-period end of the interval of intermediate periods a boundary period τ_0 (for the Earth by international agreement $\tau_0 = 30$ years) and then make the following definition: a *statistical ensemble of states of an A—ALUS system (atmosphere plus active layer of underlying surface) during a time of the order of τ_0 is called a climate*. Accordingly, oscillations in the climate system with short periods $\tau < \tau_0$ are to be called changes in the weather, while oscillations with long periods $\tau \gg \tau_0$ are changes in climate.

The most important climate-controlling parameter, external to the climate system, is the heat flux $q_i + q_r$ radiated or absorbed by the system, where q_i is the flux of internal heat and q_r is the flux of assimilable solar radiation given by formula (43.5). For the terrestrial planets q_r plays the main role; for instance, on the Earth $q_r \approx 240 \text{ W/m}^2$, whereas the geothermal heat flux q_i is only about $5 \times 10^{-2} \text{ W/m}^2$. On the giant planets, on the other hand, the main factor is apparently q_i . On the Sun $q_i \approx 6.31 \times 10^7 \text{ W/m}^2$, and flux q_r is, of course, zero.

The climate of a planet is governed, first of all, by the astronomical factors: the irradiance of the Sun (and for satellites the irradiance of the central planet as well), the position and motion of the planet in the Solar System, the inclination of its rotation axis to the orbital plane, and the velocity of rotation. These factors determine the insolation and its distribution over the outer boundary of the atmosphere, as well as the gravitational effects of external bodies, producing both the tides and the oscillations of the characteristics of the orbital motion and the rotation (and thus the oscillations of the insolation, too).

Secondly, the climate is shaped by the size and mass of the planet, by the internal sources of heat, and, in addition, possibly by its gravitational and magnetic fields. Thirdly, the characteristics of the climate system itself are determining factors. These include the mass and composition of the atmosphere (both its constant component parts and its varying amounts of thermodynamically active admixtures, in particular, cloud-forming substances) and the properties of the active layer of the underlying surface (the relief of the solid surface, the distribution of dry land and, provided these exist, as on the Earth, the distribution of basins containing liquid, the mass and composition of the solid surface, and the structure and properties of the active layer under it).

Now it could be asked whether for fixed values of all the climatogenic factors the climate system is *transitive*, that is, *ergodic* in the sense of (15.5), so that the trajectory representing its evolution in the corresponding phase space goes around all the points of this space as time goes by (and that thus the statistics of the states passed through by the system in a sufficiently long time depend little on the initial state)? Or whether the climate system is *intransitive*, so that for the very same values of the climatogenic factors slight differences in the initial states of the system may lead to different climates?

Or, finally, we could ask whether the A—ALUS system is *almost intransitive* (Lorenz, 1960) and may experience 'crossovers' (as defined by Obukhov, see Section 6) from one climatic regime to another, as is the case for the system of equations (17.10) with a Lorenz strange attractor, and as occurred, for example, on the Earth at the time of the alternating glacial and nonglacial climates of the Pleistocene? However, the above-mentioned actual detection of long astronomical

periods in the spectra of these oscillations indicates that these oscillations were caused by external climatogenic factors (perhaps serving as triggering mechanisms for the 'crossovers').

49. SIMPLIFIED CLIMATE MODELS

Models of a climate system must include all of its component parts. On the Earth these are the atmosphere, the ocean, and the active layer of the land. In the models of the atmosphere studied in Sections 44 and 45, the active layer of the underlying surface had to be taken into account via simplified boundary conditions. In the models of the ocean in Sections 46 and 47 atmospheric effects had to be specified. Now we wish to combine these models, that is, at the very least, to devise models of the global *interaction of the atmosphere and ocean*.

The simplest models are those of the similarity theory of Section 44, with equations just for a small number of integral parameters of the system, or one-dimensional models describing, for example, only the vertical inhomogeneity of the system or only the latitude inhomogeneity. Such models are, of course, incomplete, and they generally lead to correct estimates only of some of the characteristics of the system, whereas the other characteristics turn out to be unrealistic (for instance, in (44.9) the seasonal oscillations of δT , U , and ε were satisfactory, but the values obtained for T were too high).

For the Earth a generalization of the similarity theory of Section 44 to include the global interaction of the atmosphere and ocean was constructed by Zilitinkevich and Monin (1976). It consisted in, firstly, introducing along with the meridional heat flux in the atmosphere

$$Q_a = c_p M \frac{\overline{U' T'}}{\pi r_0/2} = \frac{m}{\Pi_M \Pi_\omega^2} \frac{U \delta T}{c_r T_r} q \quad (49.1)$$

the meridional heat flux in the ocean $2Q$, where Q is the typical value of the vertical heat flux from the atmosphere to the ocean in the low latitudes and from the ocean to the atmosphere in the high latitudes (formula (49.1) is completely analogous to (44.2); it differs only in that $U \delta t$ is replaced by $\overline{U' T'} \sim \sigma_U \sigma_T \sim (LU/r_0) (L \delta T/r_0)$; now, using for U the formula for the thermal wind $U = R \delta T/(\omega \sqrt{2}) (\pi r_0/2)$, for the effective Coriolis parameter $\omega \sqrt{2}$ taking (44.5) into account we arrive at the right-hand side of (49.1) with a coefficient $m \approx 1.2$). If we multiply $2Q$ by the relative area of the World Ocean, we get

$$Q_a(1 + \xi) = q; \quad \xi = 2/3 \left(\frac{2Q}{Q_a} \right). \quad (49.2)$$

The quantity Q can be determined with the aid of semiempirical formulas of the type of (34.5), (34.6):

$$\frac{Q}{c_w \rho_w} = \alpha \mathcal{K} \frac{\delta T_w}{h}; \quad u_*^2 = \mathcal{K} \frac{\delta U_w}{h}; \quad \text{Rf} = \frac{\beta Q}{c_w \rho_w} \left(u_*^2 \frac{\delta U_w}{h} \right)^{-1}. \quad (49.3)$$

Here c_w and ρ_w are the specific heat and density of seawater; $\beta \approx 0.2 \text{ cm}/(\text{s}^2 \text{ K})$ is its buoyancy parameter; $\delta T_w \sim \delta T$ and δU_w are typical vertical differences of temperature and flow velocity in the upper active layer of the ocean (of thickness h); $\mathcal{K}$ and $\alpha\mathcal{K}$ are the coefficients of vertical exchange of momentum and heat in the ocean; $u_* = (c_f (\rho_a/\rho_w) U^2)^{1/2}$ is the friction velocity in the ocean ($c_f \sim 10^{-3}$ is the friction factor for the sea surface); Rf is the dynamic Richardson number.

For $\alpha Rf = R_0 \approx 5 \times 10^{-4}$, $u_* \sim 1 \text{ cm/s}$, $\delta T_w \sim 27^\circ\text{C}$, $h \sim 300 \text{ m}$, we obtain $\alpha\mathcal{K} = u_*^2 (R_0 h / \beta \delta T_w)^{1/2} = 5/3 \text{ cm}^2/\text{s}$ and $M_s = \beta Q / c_w g \sim 10 \text{ g}/(\text{cm}^2 \text{ year})$, which is in fairly good agreement with empirical estimates. However, it is advisable to determine h rather than specify it, because of the interaction of the atmosphere and ocean. To do this, let us assume that during the cold halfyear h increases with time according to the law $\delta h / \delta t = b \alpha\mathcal{K} / h$, where b is a numerical constant and the right-hand side is the rate of intake of water by the active layer of the ocean. Then the typical value of h will be $(b \alpha\mathcal{K} t_0)^{1/2}$, where t_0 is the annual period and for agreement with the foregoing estimates we assume that $b \approx 17$. For this h , from (49.3) we now obtain

$$\alpha\mathcal{K} = \left(\frac{R_0 u_*^4}{\beta \delta T_w} \right)^{2/3} (b t_0)^{1/3}, \quad h = \left(\frac{R_0 u_*^4}{\beta \delta T_w} \right)^{1/3} (b t_0)^{2/3},$$

$$\frac{Q}{c_w \rho_w} = \left(\frac{R_0 u_*^4}{\beta \delta T_w} \right)^{1/3} \delta T_w (b t_0)^{-1/3}. \quad (49.4)$$

For a fixed u_* the variation of Q with δT_w is slower than direct proportionality, which can be explained by the stable stratification of the ocean, since this hinders the vertical exchange. However, u_* is not fixed, but rather is proportional to U , and thus to δT . If we do not distinguish between δT_w and δT , then from (49.4) we get $Q \sim \alpha\mathcal{K} \sim (\delta T)^2$ and $h \sim \delta T$. Substituting Q from (49.4) into relations (49.2) and (49.3) and then solving these for δT and ξ , we obtain

$$\frac{\delta T}{T_r} = \left(\frac{2.2}{m} \frac{\Pi_M \Pi_Q}{1 + \xi} \right)^{1/2}; \quad \xi = \Pi_Q \left(\frac{\delta T_w}{\delta T} \right)^{2/3}, \quad (49.5)$$

where Π_Q is a dimensionless factor, which can be construed as the upper limit of the ratio of the oceanic meridional heat flux to the atmospheric flux. It is difficult to express this quantity explicitly in terms of Π_ω , Π_M , and the other parameters introduced above, but this need not be done here; let us remark only that for the above values of the parameters the factor $\Pi_Q \approx 0.42$.

The theory presented here (together with the similarity theory outlined in Section 44) enables us to use the external parameters to calculate the climatic characteristics of the atmosphere: the mean temperature of the surface layer of air $T_s = T_r / \alpha_r$, the meridional temperature difference δT , the temperatures at the equator $T_E = T_s + \frac{1}{3} \delta T$ and at the pole $T_P = T_s - \frac{2}{3} \delta T$, the typical wind speed U ,

and the meridional atmospheric heat flux Q_a . Likewise we can compute the characteristics of the ocean: $\delta T_w = T_E - \max(T_p, T_{wi})$, $\alpha\mathcal{K}$ and h , as well as the characteristics of the small-scale interaction of the atmosphere and ocean: $\tau = \rho_w u_*^2$ and Q .

For instance, for the climate at present, when the Arctic is semi-isolated from the World Ocean, while Antarctica is covered with ice, in formula (49.5) for δT we can set approximately $\xi \approx 0$ and for the present-day values $T_r = 255$ K, $\Pi_\omega = 1.7$, and $\Pi_M = 2.17 \times 10^{-3}$ we obtain $\delta T = 36^\circ\text{C}$ and then $T_s = 15^\circ\text{C}$, $T_E = 27^\circ\text{C}$, $T_p = -9^\circ\text{C}$, $U = 10$ m/s, $\tau = 0.1$ Pa, $Q_a = 244$ W/m², $\frac{4}{3}Q = 84$ W/m², $\delta T_w = 27^\circ\text{C}$, $\alpha\mathcal{K} = 1.7$ cm²/s, and $h = 300$ m. These figures show a general good fit with empirical data.

This success encourages us to make similar estimates for the climates of the past. For example, during the Mesozoic both poles were apparently located in the open ocean, and, as the estimates to be presented just below will show, there was probably no perennial ice at all in the polar regions of the time. Consequently, the total albedo of the Earth was probably somewhat lower than at present, making q somewhat higher. If we assume the value of q to have been 4% higher than now, then T_r and T_s should have been 3°C higher than at present, and for $\xi = 0.42$ we obtain $\delta T = 31^\circ\text{C}$, $T_E = 301$ K (1 K higher than now), and $T_p = 270$ K (6 K higher than at present, which accounts for the absence of perennial ice).

During the substantial Permian—Carboniferous glaciation of Gondwana 300 to 250 million years ago, the South Pole was located on the continent of Gondwanaland, while the North Pole was apparently in the open ocean. If as a first approximation we evaluate the climates separately for the Southern and Northern Hemispheres, then we can assign to the Southern Hemisphere a value $\xi = 0$ and an enhanced mean albedo, say $A = 0.37$ (20% ice surface with an albedo of 0.8 and 80% nonice surfaces with an albedo of 0.26). Then the values of T_r and T_s there turn out to be 7°C lower than at present, $\delta T = 35^\circ\text{C}$, $T_E = 293$ K, and $T_p = 258$ K. In the Northern Hemisphere, on the other hand, for the indicated T_E and the Mesozoic $\delta T = 31^\circ\text{C}$, we obtain a value of $T_p = 262$ K close to the present temperature of the polar region.

It is clear from the outset that *one-dimensional* models of the *vertical* structure of the atmosphere and ocean cannot be complete, since the *horizontal transfer* of heat and other substances always participates in the formation of the vertical structure. Let us demonstrate this using as an example the elegant theory of the vertical structure of the ocean devised by Kamenkovich (1958) and Welander (1959). In this theory it is assumed that the variations in the density of seawater ρ'_w are proportional to the variations of temperature T' , and that the currents are nondivergent and geostrophic:

$$\begin{aligned} \frac{\partial p'}{\partial z} &= g\rho_0\alpha_0 T'; & \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} &= 0; \\ fv &= \frac{1}{\rho_0} \frac{\partial p'}{\partial x}; & fu &= -\frac{1}{\rho_0} \frac{\partial p'}{\partial y}. \end{aligned} \quad (49.6)$$

The elegance of this theory resides in the fact that these equations can be used to express all the unknown fields u , v , w , p' , and T' in terms of a single function M :

$$\begin{aligned} u &= -\frac{1}{f} \frac{\partial^2 M}{\partial y \partial z}; & v &= \frac{1}{f} \frac{\partial^2 M}{\partial x \partial z}; & w &= \frac{\beta}{f^2} \frac{\partial M}{\partial x}; \\ p' &= \rho_0 \frac{\partial M}{\partial z}; & T' &= \frac{1}{g\alpha_0} \frac{\partial^2 M}{\partial z^2}. \end{aligned} \quad (49.7)$$

Then the equation of convective heat conduction with vertical turbulent exchange $\mathbf{u} \cdot \nabla T' = \alpha \mathcal{K} \partial^2 T' / \partial z^2$ and the boundary conditions at the ocean surface $z = 0$ (specified temperature T'_s and specified vertical velocity $w_s = \text{curl } \boldsymbol{\tau}_0 / f$) and at the bottom $z = -h$ (specified temperature $T'_h \approx 0$ and condition of nonpassage $w = -\mathbf{u} \cdot \nabla h$) will have the form

$$\begin{aligned} \alpha \mathcal{K} f \frac{\partial^4 M}{\partial z^4} + \frac{\partial(\partial^2 M / \partial z^2, \partial M / \partial z)}{\partial(x, y)} - \frac{\beta}{f} \frac{\partial M}{\partial x} \frac{\partial^3 M}{\partial z^3} &= 0; \\ z = 0; \quad \frac{\partial^2 M}{\partial z^2} &= g\alpha_0 T'_s; \quad M = M_s = \frac{f^2}{\beta} \int_0^x w_s dx; \\ z = -h; \quad \frac{\partial^2 M}{\partial z^2} &= 0; \quad \frac{\beta}{f} \frac{\partial M}{\partial x} = \frac{\partial(h, \partial M / \partial z)}{\partial(x, y)}. \end{aligned} \quad (49.8)$$

If the scale of horizontal length $L = f/\beta \approx 6 \times 10^3$ km, the vertical velocity scale $W \approx 10^{-4}$ cm/s, and the temperature scale $\Theta \approx 10$ K, then from the third equation of (49.8) for M we obtain a scale $M_h = fL^2 W$, and from the fifth equation for z we get the *advective* scale $z_a = (fL^2 W / g\alpha_0 \Theta)^{1/2} \approx 400$ m (for $\alpha_0 \approx 2.5 \times 10^{-4}$ (K) $^{-1}$). Changing now to dimensionless variables, in (49.8) we retain only the small parameter $\varepsilon = Z_d / Z_a \approx 1/40$ (where $Z_d = \alpha \mathcal{K} / W \approx 10$ m is the *diffusion* scale for $\alpha \mathcal{K} \approx 10^{-1}$ cm ^{2} /s) and $\delta = Z_a / Z_h$ (where Z_h is the scale of variations of the depth h). Therefore, except in the boundary layers at the ocean surface and at the bottom, the vertical advection of heat must be compensated by horizontal advection, and the model cannot be one-dimensional.

However, in order not to be restricted here by this negative conclusion, we offer a particular solution of Equation (49.8), analogous to that obtained by Needler (1967, 1971). This solution, which satisfies both upper boundary conditions, can be written as

$$\begin{aligned} M &= M_s + A(x, y)z - \frac{g\alpha_0 f^2}{b^2} T'_s (1 - e^{bz/f}); \\ \frac{\partial(A, \ln T'_s)}{\partial(x, y)} &= \frac{b^2}{f} \alpha \mathcal{K} - bw_s + \frac{g\alpha_0 \beta}{b} \frac{\partial T'_s}{\partial x}. \end{aligned} \quad (49.9)$$

For large bh it also satisfies approximately the first of the bottom conditions (the second condition is satisfied only for a special choice of w_s , that is, of the wind field). It turns out that this solution in general provides a fairly good description of the vertical distributions of u , v , and T in the ocean.

With regard to the vertical thermal structure of the atmosphere, in addition to turbulent mixing and large-scale transport, an important role is played by radiative heat exchange (since in the air, in contrast to the water, the mean free paths of photons are quite long, for long-wave self-radiation as well as for short-wave solar radiation). Here the radiative heat exchange is so important that it is advisable to calculate the vertical profile of the temperature of *radiation equilibrium*. The first to make such a calculation was Emden (1913), who used the so-called two-flux approximation (that is, he took into account only vertical fluxes of radiant energy, upward $F^\uparrow$ and downward $F^\downarrow$). He assumed that the atmosphere absorbs long-wave radiation like a ‘gray’ body (that is, the absorption coefficient α_a does not depend on the wavelength of the radiation):

$$\frac{\partial F^\uparrow}{\partial z} = \alpha_a \rho_a (F^\uparrow - f\sigma T^4); \quad \frac{\partial F^\downarrow}{\partial z} = \alpha_a \rho_a (f\sigma T^4 - F^\downarrow);$$

$$z = 0: \quad F^\uparrow - F^\downarrow = S; \quad z = \infty: \quad F^\downarrow = 0, \quad (49.10)$$

where ρ_a is the density of the radiation-absorbing substance, and S is the flux of solar radiation. Since for radiation equilibrium we must have $F^\uparrow - F^\downarrow = \text{const}$, these equations are easy to solve for T ; however, they naturally give exaggerated vertical and latitudinal contrasts, which in reality are smoothed out by turbulent mixing and large-scale heat transfer. This was only partly corrected by the rejection of the ‘gray-body’ approximation (that is, the summation over all the spectral lines) by Manabe and Strickler (1964), who also introduced ‘convective adjustment’.

Kibel’ (1943) attempted to compensate for the deficit of radiation equilibrium by turbulent mixing, remaining within the framework just of vertical heat exchange (that is, without taking large-scale heat transfer into account). To do this, he replaced the condition of radiation equilibrium $F^\uparrow - F^\downarrow = \text{const}$ by the equation

$$\alpha_a \rho_a (F^\uparrow + F^\downarrow - 2f\sigma T^4 + \gamma S) + \frac{\partial}{\partial z} \rho_a \mathcal{K} \frac{\partial T}{\partial z} = 0, \quad (49.11)$$

where γ is the ratio of the absorption coefficients for ‘gray’ short-wave and long-wave radiation, the profile of S is described by the equation $\partial S / \partial z = \gamma \alpha_a \rho_a S$, the boundary conditions for $z = 0$ are taken to be $F^\uparrow - F^\downarrow = S + \rho_a \mathcal{K} \partial T / \partial z$ and $F^\uparrow = f_* \sigma T^4$, and for $z = \infty$ the condition $F^\downarrow = S = q$ is added. Now, if instead of z we use the optical thickness $\tau = \int_z^\infty \alpha_a \rho_a dz$, and if we change to dimensionless variables:

$$\xi = q^{-1}(F^\uparrow + F^\downarrow); \quad \eta = q^{-1}(F^\uparrow - F^\downarrow); \quad \zeta = q^{-1}f\sigma T^4;$$

$$\frac{m^2 - 1}{2} = \frac{4f\sigma T^3}{\alpha_a \rho_a \mathcal{K}}, \quad (49.12)$$

then we can rewrite the equations of the Kibel' theory as

$$\begin{aligned} \frac{\partial}{\partial \tau} \frac{2}{m^2 - 1} \frac{\partial \xi}{\partial \tau} + \xi - 2\xi + \gamma e^{-\gamma \tau} &= 0; \\ \frac{\partial \xi}{\partial \tau} &= \eta; \quad \frac{\partial \eta}{\partial \tau} = \xi - 2\xi; \\ \tau = \tau_0: \quad \eta &= e^{-\gamma \tau_0} - \frac{2}{m^2 - 1} \frac{\partial \xi}{\partial \tau}; \quad \xi + \eta = 2 \frac{f_*}{f} \xi; \\ \tau = 0: \quad \xi &= \eta = 1, \end{aligned} \quad (49.13)$$

where τ_0 is the total optical thickness of the atmosphere. For $m \approx \text{const}$ these equations are easily solved. Kibel' assumed $m = 1.75$ and $\rho_a = \rho_{a0}(T_0/T_\infty)(1 + \tau)^{-1/4} e^{-z/L}$, and with a quite reasonable choice of parameters he was able to obtain a good fit with the actual temperature profile (at the latitude of 42° chosen by him). However, for τ_0 the value of 12.6 obtained was much too high, so that in this model the fluxes of radiant energy were far from realistic. Consequently, neglecting large-scale heat transfer keeps us from discerning the real relationships between the vertical profiles of temperature and radiation, as was to be expected.

Similar discrepancies also appear in *one-dimensional* models of the *latitudinal structure* of the climate. These are usually based on the equation of the annual heat budget for the latitude zones

$$QS(y)[1 - A(y)] - I(y) = \mathcal{D}(y), \quad (49.14)$$

where $y = \sin \varphi$ is the meridional coordinate, $Q = \frac{1}{4}I_0$ is one fourth of the solar constant, $QS(y)$ is the meridional distribution of the insolation (given by astronomical formulas), $A(y)$ is the planetary albedo (so that $QS(1 - A)$ is the annual assimilable solar heat), $I(y)$ is the annual flux of outgoing solar radiation (so that the left-hand side of (49.14) is the total annual flux of radiant energy at the outer limit of the atmosphere), and $\mathcal{D}(y)$ is the divergence of the annual meridional heat flux, produced by meridional motions, this flux being integrated over the entire thickness of the atmosphere and ocean.

This equation is self-evident, but any attempts to express the quantities A , I , and $\mathcal{D}$ in it in terms of the temperature of the surface layer of the atmosphere $T(y)$ may involve discrepancies, since the state of the atmospheric surface layer can hardly determine itself. Let us consider this subject, using as examples: (1) the model of Budyko (1968):

$$I(y) = (a - na_1) + (b - nb_1) T(y); \quad (49.15)$$

$$\mathcal{D}(y) = \beta[T(y) - \bar{T}]; \quad (49.16)$$

$$A(y) = \begin{cases} A_i & \text{for } T(y) < T_c, \\ A_w & \text{for } T(y) > T_c, \end{cases} \quad (49.17)$$

where it was assumed that $a = 14$, $a_1 = 125.7 \text{ MJ/m}^2$ per month; $b = 0.14$; $b_1 =$

0.10; $\beta = 9.85 \text{ MJ}/(\text{m}^2 \text{ K})$ per month; $n = 0.5$ is the amount of cloud cover; $T_c = -10^\circ \text{C}$; $A_i = 0.62$, and $A_w = 0.32$ are the albedos of the ice-covered and ice-free zones; $\bar{T}$ is the mean planetary temperature, introducing nonlocalization into the mode; and (2) the model of Sellers (1969):

$$I(y) = \sigma T^4(y) \left\{ 1 - \frac{1}{2} \text{th}[\sigma_1 T^6(y)] \right\}; \quad (49.15')$$

$$\mathcal{D}(y) = -\frac{\partial}{\partial y} (1 - y^2) \mathcal{D} \frac{\partial T(y)}{\partial y}; \quad (49.16')$$

$$A(y) = \begin{cases} A_0 - A_1 T(y) & \text{for } A \leq 0.85 \text{ and } T < 10^\circ \text{C}, \\ \text{const} & \text{in other cases} \end{cases} \quad (49.17')$$

(where it is assumed that $\sigma_1 = 1.9 \times 10^{-15}$; $\mathcal{D} = 0.45 \text{ W}/(\text{m}^2 \text{ K})$; $A_0 = 0.486$; $A_1 = 0.092 \text{ K}^{-1}$; nonlocalization is introduced here via (49.16')). These models are characterized by nonuniqueness of their solutions (indicating, perhaps, 'intransitiveness' of the climate, in Lorenz's sense of the word) and instability of some of them with regard to small variations of the parameters.

For example, for models (49.15)–(49.17) at the critical latitude $y = y_c$, where $T = T_c$, we have

$$F(y_c) = \frac{I_0}{Q} \frac{1 + \beta_1}{\beta_1} \frac{1 - \alpha_c}{\beta_1} S(y_c) + \bar{A}(y_c) - 1 = 0, \quad (49.18)$$

where $\beta_1 \approx \beta/(b - nb_1) \approx 2.61$, and $\bar{A}$ is the weighted-mean (with respect to $S(y)$) value of the albedo. If $S(y)$ is approximated by a piecewise-linear function, then $\bar{A}$ will be a quadratic function of y_c and (49.18) will be a quadratic equation in y_c . For present-day values of the parameters it has two real roots, the larger of which ($\varphi_c = 72^\circ \text{N}$) corresponds to the present climate. For a reduction of Q by only 1.6%, the two roots merge (at a latitude $\varphi_c = 50^\circ$), and at even lower values they become complex (that is, the boundary of the ice moves jumpwise to the equator, and complete glaciation of the entire Earth sets in, a state stable with respect to small variations of Q).

According to the calculations of Golitsyn and Mokhov (1978), the minimum rate of entropy exchange between the climate system and the surroundings ensuring stability of the system is reached at a considerably lower value of β than indicated above, and thus at a larger temperature difference δT between equator and poles than in reality.

By replacing (49.17) with the more suitable formula $A = A_0 - (\delta A/\delta T)T$ (in the temperature interval outside of which A does not depend on T), Golitsyn and Mokhov derived, for the critical latitude φ_* below which complete glaciation of the Earth occurs, the equation

$$\frac{\pi}{4} - \frac{\varphi_*}{2} - \frac{\sin 2\varphi_*}{4} = \frac{b - nb_1 + \beta}{Q \partial A/\partial T} = B. \quad (49.19)$$

Climates with a latitude of ice limit higher than φ_* , and the present-day climate is one, now turn out to be stable, since the negative feedback $I(T)$ is greater than the positive feedback $A(T)$ (in particular, for $\partial A/\partial T = 0$ all climates are stable); if the opposite is true, the climates are unstable. For the sufficient condition of stability $B \geq 1$ the critical latitude φ_* vanishes, and the transition from partial to complete glaciation of the Earth, and vice versa, takes place smoothly with a constant variation of parameters Q and $b - nb_1 + \beta$.

The extreme sensitivity of model (49.15)–(49.17) to small variations of Q and the excessively high value $\varphi_* = 50^\circ$ can be corrected if we introduce the dependence of the albedo $A_{i,w} = nA_0 + (1 - n)A'_{i,w}$ on the amount of cloud cover n . If, in addition, we introduce the dependence of n , A_0 , and $A'_{i,w}$ on the latitude, then the calculations of Golitsyn and Mokhov show that φ_* decreases to 20° and is reached for an 8% drop in Q ; this shows a good fit with results obtained using three-dimensional models of the climate system. However, such a corrected model could no longer be called simple.

The same results were yielded by model (49.15')–(49.17'). Faegre (1972) solved these equations numerically (he replaced (49.15') by the formula $I = K\sigma T^4$, where $K = K(y)$ was selected in conformity with the present climate). He obtained five solutions: solutions 1 (present climate) and 2 (15° lower and with $\varphi_c = 50^\circ$) are unstable relative to a 2% drop in Q ; solution 3 is stable (100° lower); and solutions 4 and 5 with a warmer Southern Hemisphere which vanish when $K(y)$ is replaced by a constant.

North (1975) considered a mixed model (49.15), (49.16'), (49.17), the equations of which for $S(y) = 1 + S_2P_2(y)$ are solved in hypergeometric functions, where $P_2(y)$ is a Legendre polynomial. The function $y_c(Q)$ obtained by North corresponds to five climate regimes. The present climate and complete glaciation of the Earth were found to be stable, while the climate of the glacial periods was unstable, and thus unattainable from the state existing at present (a result attributed by North to the incompleteness of the model, to the absence in it of a number of important feedback relationships $n(T)$, $\mathcal{D}(T)$, etc.).

Golitsyn and Mokhov, setting $A_w = 0.3 - 0.08P_2(y)$, established that in the models of North the rate of entropy exchange between the climate system and the surroundings goes through a minimum at a considerably lower value of $\mathcal{D}$, and thus a greater meridional temperature contrast δT , than for the present climate. In addition, they obtained the same criteria for the stability of different climates as were obtained above.

Consequently, a simplified parametrization of the terms in Equation (49.14) leads to a number of contradictions, indicating that more complete climate models have to be constructed.

50. THREE-DIMENSIONAL NUMERICAL MODELS

In order to devise sufficiently complete models of the Earth's climate system (atmosphere–ocean–land system), we describe the atmosphere using equations (45.1)–(45.5) with the appropriate boundary conditions at the lower boundary, including the heat and moisture budget equations (45.6) and (45.7). The ocean is described by the same fluid-dynamical equations (45.1)–(45.3) as well as by

analogues of the equations of thermodynamics (45.4), (45.5), but with the inclusion of boundary conditions at the ocean surface (46.1) and (46.2) specifying the shape of this surface and also the salt-budget equation (46.3).

Moreover, three additional problems have to be solved: (1) in order to determine the fluxes of heat and moisture into the soil q_s and I_s , entering into conditions (45.6) and (45.7), we have to describe the hydrological processes in the active soil layer on land; (2) we have to overcome the difficulties related to the very great mechanical and thermal inertia of the ocean, compared with the atmosphere, and to integrate satisfactorily the equations of ocean dynamics over very long periods; (3) we have to describe the state and evolution of the continental ice sheets, which have an even greater inertia than the ocean. Let us consider these three problems individually.

A program for portraying the hydrological processes on land was formulated by Monin (1979). According to this program, we specify the mean relief $z = h(\theta, \lambda)$ over the squares of a selected space grid (the *watersheds* being oriented along the crest axes and the *riverbeds* separating the *slopes* lying along the trough axes); the 'roughness height' z_0 ; the albedo A and the emissivity f ; the heat capacity and thermal conductivity of the soil; the relief $z = h_2(\theta, \lambda)$ of the underground water-resistant surface (the relief of the ground water surface $z = h_1(\theta, \lambda, t)$, $h_2 \leq h_1 \leq h$, is not specified, but must be calculated); the capillary m_0 and noncapillary m_1 volume porosities of the soil, as well as the corresponding *seepage coefficients* k_0 and k_1 in the layer $h_1 \leq z \leq h$ (and m and k in the layer $h_2 \leq z \leq h_1$).

For incomplete saturation of the layer $h_1 \leq z \leq h$ (when the volume *soil humidity* σ is less than m_0), the capillary seepage in this layer is found from the equation of moisture conservation $\delta\sigma/\delta t + \partial V_z/\partial z = 0$, the *seepage rate* V_z being defined by *Darcy's law* $V_z = -k_0 \partial Z/\partial z$, where $Z = z + (p_a + p_c)/g\rho_w$ is the *head*. The dependences of k_0 and the capillary pressure p_c on σ are found with aid of empirical formulas. This yields the seepage equation

$$\frac{\partial \sigma}{\partial t} - W \frac{\partial \sigma}{\partial z} = \frac{\partial}{\partial z} \mathcal{D} \frac{\partial \sigma}{\partial z}; \quad W = \frac{\partial k_0}{\partial \sigma};$$

$$\mathcal{D} = \frac{k_0}{g\rho_w} \frac{\partial p_c}{\partial \sigma}. \quad (50.1)$$

The boundary condition for σ at the ground surface $z = h$ when this surface is covered with water has the form $\sigma = m_0$, whereas if there is no water present $\rho_w V_z = E$; at the surface of the ground water $z = h_1$ we again have $\sigma = m_0$. The moisture flux into the soil I_s , which appears in (45.7), is the sum of $\rho_w V_z$ and the *infiltration* $(-\rho_w k_1)$.

The subsurface runoff is calculated with the aid of the *hydraulic approximation*: on the average over the layer $h_1 - h_2$ we set $\mathbf{V} = -k\nabla h_1$. Then the equation of conservation of moisture reduces to

$$\frac{\partial m h_1}{\partial z} = \text{div}(\mathcal{D}_1 \nabla h_1) + \frac{I_1 - E_1}{\rho \omega}; \quad \mathcal{D}_1 = k(h_1 - h_2) \quad (50.2)$$

(the *Boussinesq equation* for h_1), where I_1 and E_1 are the infiltration and evaporation at the upper surface of the ground water, and $\mathcal{D}_1$ is called the *conductivity of the stratum*.

The surface runoff, the source of which E_s is found from (45.7), is the sum of the slope runoff and the channel (river) runoff. The downhill runoff can be characterized as a two-dimensional flow of a compressible fluid with a surface density ρ_s satisfying the continuity equation with sources

$$\frac{\partial \rho_s}{\partial t} + \operatorname{div} \rho_s \mathbf{v} = E_s \quad (50.3)$$

(valid in the regions θ, λ, t where $\rho_s > 0$), the velocity $\mathbf{v} = (v_\theta, v_\lambda)$ being given by the equations:

$$\begin{aligned} \frac{dv_\theta}{dt} - \frac{\operatorname{ctg} \theta}{r_0} v_\lambda^2 - f v_\lambda &= -\frac{g}{r_0} \frac{\partial h}{\partial \theta} - \frac{E_s v_\theta}{\rho_s} - \frac{v v_\theta}{l} + F_\theta; \\ \frac{dv_\lambda}{dt} + \frac{\operatorname{ctg} \theta}{r_0} v_\theta v_\lambda + f v_\theta &= -\frac{g}{r_0 \sin \theta} \frac{\partial h}{\partial \lambda} - \frac{E_s v_\lambda}{\rho_s} - \frac{v v_\lambda}{l} + F_\lambda. \end{aligned} \quad (50.4)$$

The second terms on the right-hand sides appeared after we went from the equations for the momentum density $\rho_s \mathbf{v}$ to the equations for $\mathbf{v}$ with the aid of (50.3). The third terms on the right describe the frictional resistance squared (according to the *Chezy formula*), while the fourth terms describe the diffusion of the downhill runoff (with a viscosity that depends on the rms size of the small-scale inclinations of the relief, as well as on the large-scale curvature of the slope, determining the sign of the viscosity; for instance, for a positive curvature the downhill flow diverges, and for a negative curvature it converges). Because of the sources or sinks E_s , the downhill flow may have variable edges $\rho_s = 0$ (to be calculated), which may contain 'holes' and in which 'drops' can be distinguished.

Channel flow can be described as a *one-dimensional* flow of a compressible fluid with a linear density $\rho_l = \rho_w \Sigma$ (where Σ is the area of the active section of the river filled with water), satisfying the *St. Venant equations*. The latter include the continuity equation with sources, made up of the components normal to the river of the downhill and subsurface runoffs from the right and left slopes, and equations of motion analogous to (50.4) (with h replaced by $h + h_r$, where h_r is the depth of the river, and with the longitudinal Taylor diffusion described with the aid of $\mathbf{F}$), it also being necessary here to specify the 'equation of state' $\Sigma \approx \Sigma(h_r)$.

It should be noted that in the climate models the indicated hydrological program has not yet been used; rather it has been replaced by certain simplified empirical relations. Going ahead a bit, we should mention, too, that in almost all the climate models the continental ice sheets were not calculated either, only their relief being specified and the presence or absence of a snow cover (this was done during the calculation of both the present climate and the climate at the peak of the Würm glaciation 18,000 years ago).

The first three-dimensional climate model was constructed by Manabe and Bryan (1969). The atmosphere was described in this model just as in the 'wet' model presented in Section 45. In order to describe the hydrology, it was assumed that the evaporation E from the ocean, from a highly moistened land surface (including dense vegetation), and from snow and ice, is proportional to the difference between the humidity of saturated air at a temperature T_s and the humidity at the upper limit of the atmospheric boundary layer, while for a soil humidity σ less than the critical value $\sigma_* \approx \frac{3}{4}\sigma_m$ (where σ_m is the moisture capacity of the soil) it is reduced by a factor of σ_*/σ . For $\sigma \approx \sigma_*$ and a seepage rate $V_z \approx P - E > 0$ it was assumed that $\partial\sigma/\partial t = 0$, while for $\sigma < \sigma_*$ the equation $\partial\sigma/\partial t = P - E$ was used. The evolution of a snow cover was described by equation (46.5), in which q_{wi} was found from (45.6) for the freezing point T_{wi} and $q_s = 0$. The moisture content of snow was taken to be zero (and that of the soil under the snow was assumed to be as above).

The ocean was described as in the model of Bryan and Cox (1967–1969), presented in Section 46 (with exchange coefficients $\mathcal{K}_z = 1.5 \text{ cm}^2/\text{s}$ and $\mathcal{K}_h = 10^8 \text{ cm}^2/\text{s}$, the second of which is apparently too high). The ocean depth was assumed to be constant ($h = 4 \text{ km}$). The pack ice was introduced somewhat differently than in (46.6), its thickness being found from the equation

$$\frac{\partial h_i}{\partial t} + \text{div}(\gamma \bar{\mathbf{V}}_h h_i) = \mathcal{K}_h \Delta h_i + P_s - E - (Q_1 + Q_2) \rho_i^{-1} \mathcal{L}_{wi}^{-1}, \quad (50.5)$$

where $\gamma = 1$ for $h_i < 3 \text{ m}$ and $\gamma = 0$ for $h_i > 3 \text{ m}$; $\bar{\mathbf{V}}_h$ is the mean velocity of the current in the upper layer of the ocean; Q_1 and Q_2 are the upward and downward heat fluxes in the ice. During the first stage of the calculations τ_0 , E , P , and T_w were specified according to the data of the atmospheric block as an average over the last 100 days of its integration for an inert ocean with horizontal heat transfer. It was assumed that $Q_1 = kT_1 (h_i + k/\varepsilon)^{-1}$ and $Q_2 = kT_2 h_i^{-1}$, where temperatures T_1 and T_2 are reckoned from the freezing level of seawater 271.2 K, with $k = 2.1 \text{ W}/(\text{m K})$ and $\varepsilon = 123 \text{ W}/(\text{m}^2 \text{ K})$. In the second stage at the water surface it was assumed that $\mathcal{K}_z \partial T_w / \partial z = q_s$, where q_s was found from (45.6); for ice a procedure was used which guaranteed the conditions $T_1 \leq 273.1 \text{ K}$ and $T_2 = 271.2 \text{ K}$. The salt flux (46.3) was 'parametrized' similarly.

The main innovation of this climate model was the artificial 'synchronization' of the ocean and atmosphere (in order to solve the second of the above-cited problems): the equations of the atmospheric block were integrated (according to a 9-level scheme) in steps of 10 minutes each (over a year), while the oceanic block was integrated (according to a 5-level scheme) in steps of 166.7 minutes each (over 100 years), and then the state of the ocean in each step was 'joined' to the state of the atmosphere in each of its 18th steps, the latter state being smoothed in time (with exponentially decreasing weighting factors into the past).

Calculations based on this model were carried out in an arbitrary spherical region (in Mercator projection) with a longitude size of 120° (with cyclic repetition), in which there are equal amounts of land and sea between latitudes $\pm 66.5^\circ$, but only land at latitudes higher than these (up to $\pm 81.7^\circ$). The results of the

calculations were averaged statistically over the last 100 atmospheric days. An accomplishment of this model was a considerably better reproducibility of various features of the present-day climate, *vis-à-vis* an isolated atmospheric model with an inert ocean.

The meridional transport of heat by ocean currents observed here, amounting to about half of the atmospheric transport, was found to reduce the meridional temperature difference δT . Accordingly, all the quantities depending on δT were reduced as well: the intensity of the zonal circulation, the gradient $\partial u_h / \partial z$ of the wind velocity in the troposphere, the baroclinic instability of the atmosphere, the kinetic energy of the large-scale turbulence, the intensity of the backward cell of the meridional circulation in the middle latitudes and of 'centers of action' of the atmosphere, the meridional flux of momentum, the exchange of angular momentum between the atmosphere and the solid Earth, etc.

Because of the heating of the atmosphere by the ocean, at the east coast of a continent in the middle and high latitudes the cyclogenesis was enhanced and the amount of precipitation increased (which restricted the subtropical desert to the western half of the continent), whereas cold upwelling in the equatorial ocean weakened the tropical rainfall there and, conversely, increased it over the continent. In the north latitudes the transport of warm water by the subarctic gyre produced similar effects at the west coast of the continent. The rain belts of the middle latitudes and the boundaries of the snow cover were shifted poleward somewhat.

The calculation of the fields of T and s at the ocean surface (rather than the specification of these fields, as in previous models) represented an important step forward. A quite satisfactory reproduction of the salinity field over the entire depth of the ocean was achieved, with the halocline in the arctic zone, the maximum of s at the ocean surface in the subtropics, and a weak minimum of s at a depth of about 1 km.

On the whole, the transition from the models of the noninteracting atmosphere and ocean, described in Sections 45 and 46, to a unified model of the atmosphere—ocean—land climate system made it possible to treat a new feature of this system: the effects of the large-scale interaction of the atmosphere and ocean, the most salient of which were mentioned in the three preceding paragraphs and which are apparently responsible for the long-term weather anomalies.

At the same time, we should note that the artificial 'synchronization' of the ocean and atmosphere carried out in the model of Manabe and Bryan proved to be quantitatively unsatisfactory: statistical equilibrium in the thermal state of the ocean during the 100 ocean years calculated was not achieved, and at the end of this period the ocean continued to receive heat from the atmosphere at an average rate of 10 W/m^2 (which corresponds to a mean rate of ocean heating of about 2°C per century). Nevertheless, a similar 'synchronization' was used in a number of later models, to be described briefly below.

Another means of 'joining' the ocean and the atmosphere was used in the model of the Institute of Oceanology of the USSR Academy of Sciences (Zilitinkevich *et al.* 1976). It consisted in dividing the ocean into an upper active layer (of thickness $h_1 = 350 \text{ m}$), synchronously 'joined' with the atmosphere, and a deep part, the

climatic state of which was calculated separately. Thus, this model comprised four blocks: the free atmosphere (A), the atmospheric boundary layer (ABL), the active layer of the ocean (ALO), and the deep ocean (O).

In block A prognostic equations were employed to compute the fields of velocity, temperature, humidity, and surface pressure and to 'parametrize' the horizontal exchange, radiative heat exchange, convection, phase changes of moisture, and precipitation. The ABL block was 'parametrized' according to the methods in Section 36. In the ALO block the temperature T_s of the upper mixed layer and its thickness h were calculated, and the horizontal exchange, the vertical structure, and the vertical velocity were 'parametrized'. In block O the stream function of the barotropic flow ψ and the temperature were calculated, while the 'parametrized' quantities in this block were the horizontal exchange and the vertical structure.

From A the values of $\mathbf{u}_1$, T_1 , s_1 , F^1 are transmitted to the ABL, and the value of p_u to the ALO. From the ABL the values of T_s , q , E , and τ are transmitted to A, the values of q_s and τ to the ALO, and the values of q_s and τ to O. From the ALO the value of T_s is transmitted to the ABL, and from O the values of T and $\mathbf{u}$ are transmitted to the ALO. On land and on a surface of sea ice it was assumed that $q_s = 0$. In regions of perpetual ice with $T_s > 0^\circ\text{C}$ this temperature was replaced by zero and it was assumed that q_s is expended on melting the ice. The air humidity at the ocean surface, and when it rains, on land, was taken to be saturating (but 2 hours after a rainfall on land it dropped to half the saturation value). The albedo of ice or snow was taken to be 0.7, that of vegetation 0.2, and that of a water surface 0.1.

In block A the vertical coordinate was taken to be the normalized pressure $\sigma = p/p_s$, the surface pressure p_s being found from the so-called *tendency equation*:

$$\frac{\partial p_s}{\partial z} + \frac{1}{r_0 \sin \theta} \int_0^1 \left(\frac{\partial p_s u_\theta \sin \theta}{\partial \theta} + \frac{\partial p_s u_\lambda}{\partial \lambda} \right) d\sigma = 0, \quad (50.6)$$

obtained by integrating the continuity equation over the entire thickness of the atmosphere. In the numerical calculations the atmosphere was divided into four layers in all, with boundaries at heights of 1.5, 4.5, and 11 km. However, the amounts of cloudiness were not specified climatically; rather at each of these three levels they were computed using empirical formulas, as linear functions of the corresponding relative humidity s/s_m , and the total cloud amount was taken to be $n = 1 - (1 - n_1)(1 - n_2)(1 - n_3)$.

The radiative heat fluxes were then calculated taking the cloud amounts into account. The short-wave fluxes were found using the coefficients of transmission and absorption in cloudy and cloudless squares of the horizontal grid, and the long-wave fluxes were calculated from the values of the transmission function of water vapor $\mathcal{P}(m)$ (where m is the mass of water vapor). 'Convective adjustment' was used here. The reduction of moisture content due to the fall of precipitation was taken to be $\delta s = s - cs_m$ (if $s \geq cs_m$), where $c = 0.7$ for convective precipitation and $c = 1$ for nonconvective precipitation. The corresponding temperature

rise was $\delta T = \mathcal{L} \delta s / c_p$. Division of δT and δs by the time of 'convective adjustment' (2 hours) yielded the influxes of heat $Q_c + Q_p$ and moisture $P_c - P_p$ (where Q_c and P_c are the convective influxes). The sum of P_p over all the layers gave the precipitation. Finally, the coefficient of horizontal exchange was taken $\mathcal{K}_h = 0.1 \tilde{\epsilon}^{-1/3} l^{4/3} U / U_0$, where $\tilde{\epsilon} = 0.5 \text{ W/g}$, l is the grid interval, U is the wind velocity, and $U_0 = 17 \text{ m/s}$.

In the ALO block the mean currents over the depth were assumed to be the sums of the climatic currents (from block A) and the Ekman drift, calculated from the frictional stress of the wind. The vertical temperature profiles in the active layer of the ocean were assumed to possess similarity, and their parameters T_s and h were determined from the equation of heat transfer and the equation of the budget of turbulent energy, both integrated over the thickness of the ALO. It was assumed that for $T_s < -1.8^\circ \text{C}$ there is ice on the ocean surface.

As the O block the model of ocean circulation devised by Kagan *et al.* (1974) was used. In the upper 2-km layer of the ocean it was assumed that $T = T_B + A(z) \Theta(\theta, \lambda, t)$, where T_B is the fixed abyssal temperature, $A(z)$ is a standard function (with a fixed thickness of the upper mixed layer), and Θ is found from the equation of the mean (over the depth) horizontal heat transfer by drift currents and barotropic currents. The stream function ψ of the barotropic current is found from the vorticity equation, allowing for the wind-caused and near-bottom frictional stresses. The equations of this block were integrated on a latitude-longitude grid with a 5° interval, using the method of stationing for zero initial conditions $\psi = \Theta = 0$ with values of τ_0 and q_s found from contemporary mean annual atmospheric data.

Using this model of the atmosphere—ocean—land system, only a preliminary numerical experiment was carried out: with the actual distribution of continents and oceans on a sphere, but with a horizontal grid interval that was too large, about 1000 km (360 points in each spherical layer, giving only 7560 degrees of freedom). The initial conditions were as follows: zonal distributions of T in A and ALO which were close to the mean annual values; an adiabatic $T(A)$ profile; absence of wind; $p_s = \text{const}$; $s/s_m = 0.5$; and a thickness of the upper mixed layer of 50 m. The integration was carried out in time steps 20 min long in A and 2 hours in ALO, with a natural synchronous 'joining' of these.

Calculations were carried out for 1000 days, allowing for the annual variation of the insolation. A statistically equilibrium quasiperiodic regime was reached in a year. In this regime the mean (over the mass of the troposphere) wind velocity was found to be 17 m/s, for a temperature of 244 K ($T_s = 282 \text{ K}$ with a minimum mean monthly value of 234 K in Antarctica and a maximum of 308 K in North Africa), a mean specific humidity of 1.6‰, $n = 0.47$, a radiative heat flux, assimilable by the Earth's surface, of 470 W/m² (of which 80% was carried away by longwave radiation), and a thickness of the upper mixed layer of 82 m.

The atmospheric pressure over the oceans was found to be lower than over the land (994 mbar, compared with 1040 mbar), but the air temperature was higher (by 5°C), as was the mean specific humidity (2.1‰, compared with 1.2‰), the cloud amount (0.8, compared with 0.4), the evaporation, and the precipitation (6 mm/day, compared with 3.1 mm/day). The amplitudes of the annual variation of

the mean (over the mass) atmospheric temperature were found to be a maximum in the equatorial zone, especially over the oceans of the northern winter (but over the continents these maxima were over the summer tropics); the evaporation was a maximum over the oceans of the middle latitudes in winter; the amount of cloudiness was higher at the equator throughout the year and it was found to have winter maxima in the middle and high latitudes. In winter the upper mixed layer was found to spread over the entire active layer of the ocean; with the start of the spring warmup h decreases to some tens of meters in the middle latitudes and to 120–150 m in the subtropics, while in summer h is a minimum and its values exhibit a very high spatial diversity.

A third means of 'joining' the ocean and atmosphere was proposed by Washington *et al.* (1980) in their model of the seasonal fluctuations of the climate system. In this model the equations of the atmospheric block were integrated (according to an 8-level procedure) for January, April, July and October. The annual and semianual harmonics of the effect of the atmosphere on the ocean were determined, and then these were used to integrate (according to a 4-level procedure) the equations of the oceanic block over a 5-year period. These results were next used for a new integration of the oceanic block, etc.

Of the above three methods for 'joining' the ocean and atmosphere, the most suitable for studying fluctuations of the present-day climate and its reaction to anthropogenic factors is the method of the Oceanology Institute of the USSR Academy of Sciences, and we can expect to see it used widely in the future. On the other hand, the method of Manabe and Bryan has been developed further. It was used, firstly, in their later variant of the model of the mean annual state of the atmosphere–ocean–land system (Manabe *et al.*, 1975; Bryan *et al.*, 1975), secondly, in their model with seasonal fluctuations (Manabe *et al.*, 1979), and, thirdly, in the model of the Siberian Branch of the Soviet Academy (Marchuk *et al.*, 1980).

The 1975 model was constructed for the actual geometry of the continents and oceans on a sphere with a 500-km step of the horizontal grid, 9 levels in the atmosphere, and 12 levels in the ocean. The radiative heat fluxes were calculated for a constant CO_2 concentration and a specified zonal climatic distribution of clouds and ozone, but with a calculated evolution of the water-vapor distribution. Every 1.3 atmospheric years was 'synchronized' with 430 oceanic years.

In the 1979 model the seasonal fluctuations were introduced via the insolation and the ozone distribution and allowing for seasonal variations of the specified zonal distributions of the albedos of the continents and oceans. It was integrated on a grid with a horizontal step in the atmosphere of 500 km during the first three years and 250 km after this, and a step in the ocean of 500 km. The initial state of the atmosphere–ocean–land system was chosen on the basis of a long-term integration of a model with the mean annual insolation (starting from a dry atmosphere with $T = \text{const} = 6.8^\circ\text{C}$ and a homogeneous ocean with $T = \text{const} = 2.5^\circ\text{C}$ and $s = \text{const} = 34.8\%$). The boundary conditions for A at the ocean surface were in this case the mean monthly values of the temperature of the upper mixed layer and the thickness of the sea ice, while for O they were the mean monthly values of the momentum and moisture fluxes, as well as of the

temperature, air humidity, wind velocity, and radiative heat flux at the lower level in the atmosphere.

The calculation was carried out for 4.2 atmospheric years, 'synchronized' with 1200 oceanic years; the relaxation time in the oceanic block was shortened by means of an artificial enhancement of the reaction of the temperature and salinity fields in the deep layers of the ocean (a fivefold enhancement at the greatest depths). In spite of this, after 1200 model oceanic years the temperature, which in the upper layers of the ocean attained statistically equilibrium values, continued to rise in the deep layers (at a rate of about 0.1°C per century). The salinity, on the other hand, especially in the upper layers of the ocean, approached statistically equilibrium values even more gradually (since there is no feedback between it and the hydrologic cycle), and during the 1200 years it only reached in the upper layers values which were too low (less than 34‰).

This model reproduced quite satisfactorily the seasonal fluctuations of a number of characteristics of the atmosphere—ocean—land system (for instance, the heat contents of the atmosphere and ocean), and it improved the reproduction of some characteristics, in comparison with the model with the mean annual insolation (in particular, the thickness of the pack ice in the Arctic found to be a little less than 3 m, and the thickness of the snow cover in the high latitudes on the continents of the Northern Hemisphere, which according to the earlier model accumulated continually). At the same time, for example, the values obtained for the meridional heat fluxes in the ocean and atmosphere, as well as the amplitudes of the annual variation of the surface temperature T_w in the middle latitudes of the Northern Hemisphere, were somewhat too low, while the values obtained for T_w around Antarctica were 6 to 10°C too high.

In the model of the Siberian Branch of the Soviet Academy, for 40-min time steps in A and 2-day steps in O, 'joining' of the atmospheric and oceanic fields was accomplished after every 6 atmospheric steps and every 7 oceanic steps, the data from A being transferred to O smoothed with weighting factors, going to zero exponentially into the past. During the 'joining' operation, the effect of O on A was determined by the field $T_s = T_w$, found from the thermodynamic equation for O allowing for the transfer of heat by currents, neglecting salinity and with a fixed thickness of the upper mixed layer of 100 m. On the basis of T_w the radiative heat transfer through the ocean surface was calculated, as well as the turbulent fluxes of sensible and latent heat in the near-water air layer. Then from the heat-budget equation (45.6) the heat flux into the ocean depths was calculated. The frictional stress of the wind at the ocean surface was calculated in the oceanic block neglecting ocean currents.

Block A included evolution equations for u_θ , u_λ , T , s , and p_s , 'parametrization' of the ALO using techniques like those in Section 36, 'convective adjustment', differing somewhat in the cases of wet and dry convection, and so forth. For this block for the specified climatic-mean January values of the ocean-surface temperature $T_w(\theta, \lambda)$, the distribution of sea ice and land snow cover (assuming the land to be flat), and the insolation, and for initial conditions $u_\theta = u_\lambda = 0$, $p_s = 1018$ mbar, $T = 287 + 60(p/p_s - 1)$, and $s = 0.8s_m(T)$, the mean January atmospheric circulation was calculated (according to a 3-level model with horizontal steps of 6°

for θ and 10° for λ and with 10-min time intervals). The calculation was carried out for 131 days, and the statistical characteristics of the atmospheric circulation were computed for days 100 to 131.

This model reproduced fairly well the level of kinetic energy of the atmosphere K (although, because of the low spatial resolution, the values obtained for the available potential energy Π and the nonzonal parts K' and Π' were a little low): the longitudinal spectra of the zonal and meridional parts of K at all three levels and the values of K' at latitude 45° (with an interval of an inverse-cube law in the K' spectrum), as well as the zonal distributions of T , s , and u_λ (although with low values of the subtropical jet stream in the Southern Hemisphere), the three-cell meridional circulation in the winter hemisphere, the qualitative (but only approximately quantitative) behavior of the meridional fluxes of heat, moisture, and angular momentum, and, finally, the quasistationary ultralong waves in the p_s field.

The oceanic block of the model of the Siberian Branch included a complete system of equations describing the dynamics of an ocean of varying depth $h_w(\theta, \lambda)$ with a realistic equation of state $\rho_w(p, T, s)$, described in Section 46 (where the first published version of this model, that of Marchuk and Zalesnyi, 1974, was mentioned). According to this block, for a fixed temperature of the ocean surface $T_w(\theta, \lambda)$ a calculation was made for a period of 118 years (100 years of heat diffusion and 18 years of interaction between the fields of the temperature and the currents).

The quasiequilibrium states of the atmosphere and ocean, obtained on the basis of the isolated blocks of the atmosphere and ocean for fixed fields of $T_w(\theta, \lambda)$ and the insolation, were used as initial data for integrating the unified model of the atmosphere—ocean—land system. The calculation was carried out for 11 oceanic years ≈ 2 atmospheric months, and the statistical characteristics of the states of the system were determined for the last 20 atmospheric days.

Inclusion of the ocean—atmosphere interaction was found to reduce the zonal values of T_w in the Northern Hemisphere and the tropics and to increase T_w in the Southern Hemisphere; in addition, there was a 5°C lowering of the temperature of the middle troposphere in the tropics, a reduction of the meridional advective and turbulent heat fluxes in the atmosphere, a weakening of the baroclinic instability of the atmosphere, a reduction of the kinetic energy and the other elements of the energy cycle of the atmosphere, a weakening of the subtropical jet stream of the Southern Hemisphere and of the inverse cell of the meridional circulation in the Northern Hemisphere, a deepening of the centers of low pressure p_s over the continents of the Southern Hemisphere, a shift of the zone of tropical rains by 10° southward etc.

On the whole, inclusion of the large-scale ocean-atmosphere interaction led to the same effects as in the 1969 model of Manabe and Bryan, so that the nature of these effects can already be considered to be reliably established in many details. At the same time, we should note that, in view of the absence, in the 1980 model of the Siberian Branch, of seasonal fluctuations of the insolation and heat sinks near Antarctica in the Southern Ocean, there was a trend toward an increase in T_w , which remained even for a fixed value $T_w = 1^\circ\text{C}$ on the coasts of Antarctica in an additional numerical experiment.

Therefore, it can apparently be assumed that the difficulty connected with the longer relaxation times of the ocean (measured in centuries) compared to those of the atmosphere (measured in months) is overcome to some extent in the three-dimensional models of the atmosphere—ocean—land system, albeit with aid of artificial techniques (and for the time being without properly taking into account the eddy 'negative viscosity' in the ocean). The form and evolution of the continental ice sheets, on the other hand, the relaxation times of which amount to tens of millennia, can within these limits be calculated only by ignoring the inverse effect of the evolution of the ice sheets on the statistically equilibrium states of the atmosphere and ocean, that is, for the time being without describing the glacial fluctuations of the climate (to be discussed in the next section).

The continental ice sheets can be considered as *pseudoplastic* bodies, flowing as a result of their own weight, in which for an increase in the deformation rates Φ the viscous stresses τ increase more and more slowly, while the viscosity, as a result, decreases; in addition, the viscosity also drops with an increase in the temperature:

$$\Phi_{ik} = K \tau^{n-1} F(T) \tau_{ik}; \quad F(T) = e^{-\kappa(T_w/T-1)} \quad (50.7)$$

(Shumskii, 1975), where for a pressure of one atmosphere $T_{wi} = 273.15$ K (while with an increase in pressure T_{wi} drops at a rate of 7.52×10^{-8} K/Pa; $\kappa \approx 21.1$; for the values of τ not exceeding 10^5 Pa which are typical of ice sheets, $n \approx 1$ and $K \sim 10^{-14}$ (Pa) $^{-1}$ s $^{-1}$). The ice flow in a sheet can be considered quasi-hydrostatic and nondivergent, and the stresses set up in it are compensated approximately by the pressure forces:

$$(\tau_{xz}, \tau_{yz}) = -g\rho_i(\zeta - z) \left(\frac{\partial \zeta}{\partial x}, \frac{\partial \zeta}{\partial y} \right), \quad (50.8)$$

where $z = \zeta(\mathbf{x}_h, t)$ is the equation of the surface of the ice sheet, and $\rho_i \approx 0.91676$ g/cm ^{3} is the density of ice, which depends very slightly on the pressure and temperature. Taking into account that $(\Phi_{xz}, \Phi_{yz}) = \frac{1}{2}(\partial u/\partial z, \partial v/\partial z)$, from (50.7) and (50.8) integration over z on condition that $u, v \rightarrow 0$ ($z \rightarrow 0$) yields

$$(u, v) = -U \left(\frac{\partial \zeta}{\partial x}, \frac{\partial \zeta}{\partial y} \right); \quad U = 2g\rho_i K \int_0^z (\zeta - z) F(T) dz. \quad (50.9)$$

In these formulas for the velocity field in an ice sheet there are two unknown functions ζ and T . In order to determine the first of these, which specifies the size and shape of the ice sheet, we have to use a kinematic boundary condition at the sheet surface similar to (46.1), but written for $z = \zeta$ rather than $z = 0$ (since here ζ is not small) and with the subtrahend on the right-hand side replaced by the quantity $W = \rho_i^{-1}(P_s - E + E_{wi})$. Now, integrating for this condition the equation of nondivergence of the velocity field over the thickness of the glacier $0 \leq z \leq \zeta$,

we obtain for z the nonlinear nonhomogeneous parabolic equation

$$\frac{\partial \xi}{\partial t} = \text{div}(V \nabla \xi) + W + w_0; \quad V = \int_0^{\xi} U \, dz, \quad (50.10)$$

where w_0 is the vertical velocity at the lower boundary of the ice sheet. Now, comparing the first and second terms on the right-hand side and taking (50.9) into account, we can estimate the height of an ice sheet with a horizontal dimension L as

$$\xi \sim \left(\frac{WL^2}{2g\rho_i K} \right)^{1/4}. \quad (50.11)$$

In order to find the field T in an ice sheet, a quantity which determines the ice viscosity and which enters into formula (50.9) for U , we can use the familiar heat-conduction equation $dT/dt = \chi_i \Delta T + \varepsilon/c_i \rho_i$, where d/dt is the individual derivative, $c_i = 2 \times 10^3 \text{ J/(kg K)}$ is the specific heat of the ice, $\chi_i = 1.1 \times 10^{-6} \text{ m}^2/\text{s}$ is the thermal diffusivity in the ice, and $\varepsilon = \tau_{\alpha\beta} \Phi_{\alpha\beta}$ is the rate of dissipation of kinetic energy into heat.

An ice sheet can be described for some specified climate at its surface if we know the temperature T_ξ and the quantity W . At the bottom for $T < T_{wi}$ we can set $u_0 = v_0 = w_0 = 0$, and the downward flux of heat into the bottom ($-c_i \rho_i \chi_i \partial T / \partial z$) is equal to the geothermal flux q_T . For $T = T_{wi}$ we have

$$w_0 = -\rho_i^{-1} \mathcal{L}_{wi}^{-1} \left(c_i \rho_i \chi_i \frac{\partial T}{\partial z} + q_T \right),$$

it being assumed that even during melting of the ice its flow at the bottom is completely retarded. At the edges of the ice sheet we can set $\xi = 0$.

The above model of a continental ice sheet was formulated by Verbitskii and Chalikov (1980) and used by Verbitskii (1981) to describe the steady-state regime of the present-day ice sheets of Eastern Antarctica, Greenland, and Severnaya Zemlya. In this case the vertical temperature profiles were approximated by quadratic polynomials in z/ξ , and the heat-conduction equation was integrated over the thickness of the glacier.

The results of the calculations showed a good fit with empirical data. For example, for the Eastern Antarctic Ice Sheet, which is $10.5 \times 10^6 \text{ km}^2$ in area, the feeding rate of atmospheric precipitation was found to range from 30 mm/year at the center to 500 mm/year at the margins, the temperature at the bottom of the active layer ranged from -50°C at the center to -20°C at the margins, and $q_T = 0.063 \text{ W/m}^2$. The calculations were carried out on a grid with a horizontal step of 200 km. The volume of this sheet was found to be $24.3 \times 10^6 \text{ km}^3$, which is within the range of the empirical estimates. The velocity of ice flow was of the order of 1 m/year at the center and 10 m/year at the sheet margins. At the bottom

of the glacier a region of melting was detected, with a maximum (up to 3 mm/year) at the sheet center and with increased melting again at the margin, where the rate of dissipation of kinetic energy is a maximum (and comparable with q_T). The volume of the bottom melting obtained was $16.8 \text{ km}^3/\text{year}$, as compared with the empirical estimate of $20 \text{ km}^3/\text{year}$ of Zotikov (1977).

51. MODELING OF GLACIAL PERIODS

The modeling of glacial periods requires a calculation of the evolution (appearance, growth, and decay) of continental ice sheets over hundreds of thousands of years, taking into account the inverse effect of these ice sheets on the state of the atmosphere and ocean. However, since it is practically impossible to integrate over such long periods the equations of the unsteady-state three-dimensional models described in the preceding section, we have to return to simplified models like those in Section 49, after adapting them in a manner appropriate to the problem at hand.

The first numerical model of the glacial periods (involving a comparatively small number of integral parameters, related by simple empirical and semi-empirical formulas and equations) was constructed by two brothers, V. Ya. Sergin and S. Ya. Sergin, in a series of studies which were summarized in their 1978 book. The climatic fluctuations in this model were produced mainly by the negative feedback between the ocean and the continental ice (with an area in the Northern Hemisphere of S_{ci}^n): increased $T_w \rightarrow$ increased E and $P \rightarrow$ increased $S_{ci}^n \rightarrow$ lower T_w , and so forth. However, such feedback would be able to produce self-oscillations only if the relaxation times of the ice sheets were assumed to be as short as the relaxation times of the ocean, which is, of course, unlikely.

Here we will present a more recent quasihydrodynamic model of a climate system with evolving ice sheets, described by Verbitskii and Chalikov (1983) and Verbitskii *et al.* (1986). Since it was assumed that the formation of continental ice sheets depends considerably on the distribution of continents and oceans with respect to the poles and equator, in the model this distribution was specified in a realistic, albeit simplified, manner (Figure 51.1).

The Pacific, Atlantic, and Indian Oceans were distinguished; they were assumed to have constant angular widths Λ_i and a constant depth h , and to be connected with the Southern Ocean along latitude boundaries $\Phi_1 = -72.5^\circ$ and $\Phi_2 = -42.5^\circ$. The Arctic Ocean was assumed to connect with the Atlantic along latitude boundaries $\Phi_3 = 77.5^\circ$ and $\Phi_4 = 90^\circ$. Each of the three main oceans was divided horizontally by a western boundary layer of constant angular width λ_0 and the open ocean, and vertically by a thermocline of thickness h_0 and the abyssal. The Southern Ocean and the Arctic Ocean were considered to be separate entities. Thus, the World Ocean was divided into 14 'boxes'. The currents in the open ocean, *as an average over the ocean depth h* , were determined with the aid of Sverdrup's relation (46.12) (for $\tau_\lambda \approx \tau = \text{const}$, $\tau_\varphi \approx 0$) and the condition of nondivergence of the total flow (with the condition $U_0 = 0$ at the right-hand shore

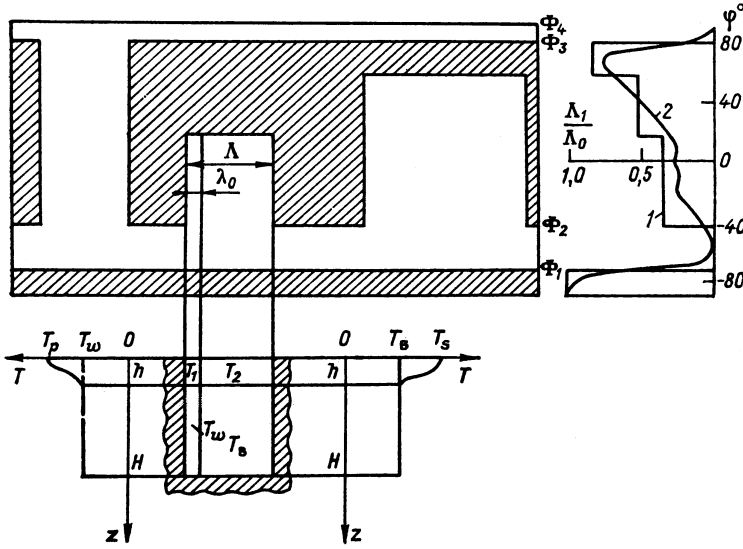


Fig. 51.1. Model oceans (continents shaded). At right: model (1) and empirical (2) latitude distributions of continental fraction Λ_1/Λ_0 . At bottom: vertical section of ocean and typical temperature profiles; T_2 and T_b pertain, respectively, to top and bottom layers of open ocean (angular width $\Lambda - \lambda_0$), and T_1 and T_w to top and bottom layers of western boundary layer (angular width λ_0); H and h are depths of ocean and thermocline; T_s and T_p are surface temperatures of open ocean and western boundary layer; $\Phi_1 = -72.5^\circ$; $\Phi_2 = -42.5^\circ$; $\Phi_3 = 77.5^\circ$; $\Phi_4 = 90^\circ$.

of the ocean):

$$V_0 = \text{curl} \frac{\tau}{\beta h}; \quad U_0 = (\Lambda - \lambda) \frac{\partial}{\partial \varphi} (V_0 \cos \varphi). \quad (51.1)$$

Using the hydrostatic equation $p \approx -g\rho_0 \xi + g \int_{-2}^0 \rho \, dz$ for $\rho \approx \rho_0[1 - \alpha(T - T_0)]$, we arrive at the following formulas for the currents as an average over the thermocline ($i = 1$) and the abyssal ($i = 2$):

$$V_i = V_0 - \frac{P_{0\lambda} - P_{i\lambda}}{f r_0 \cos \varphi} - \frac{h - h_0}{h_0} \frac{\tau}{f h}; \quad U_i = U_0 + \frac{P_{0\varphi} - P_{i\varphi}}{f r_0}; \quad (51.2)$$

$$\frac{1}{r_0 \cos \varphi} \left(\frac{\partial V_1 \cos \varphi}{\partial \varphi} + \frac{\partial U_1}{\partial \lambda} \right) + \frac{W_0}{h_0} = 0 \quad (51.3)$$

(for $i = 2$ we have to replace $(h - h_0)/h_0$ by unity). Here $(P_{i\lambda}, P_{i\varphi})$ are the mean values over the corresponding layers from $-g\alpha \int_{-z}^0 (\partial \tau / \partial \lambda, \partial \tau / \partial \varphi) \, dz$, while W_0 is the vertical velocity at depth h_0 (at the equator it was assumed that $W_0 = 0$). The temperature profile was selected in the universal form $T = T_B + (T_s - T_B)\theta(z)$ in

the thermocline (with a mean value of T_z) and $T = T_B$ deeper than the thermocline. The second terms in (51.2) are then linear combinations of gradients T_2 and T_B . In the western boundary layer it was assumed that $(\bar{V}_{ip}, \bar{W}_p) = -(\bar{V}_i, \bar{W}_0)$ $(\Lambda - \lambda_0)/\lambda_0$, where the overbar signifies averaging over the latitude. The equations of heat transfer in the ocean were averaged over each 'box'. For example, for the thermocline of the open ocean it was found that

$$\begin{aligned} \frac{\partial T_2}{\partial t} + \frac{\partial[\bar{V}_1 T_2 - \mathcal{K} \partial T_2 / r_0 \partial \varphi] \cos \varphi}{r_0 \cos \varphi \partial \varphi} - \frac{U_1(\lambda_0)}{r_0(\Lambda - \lambda_0) \cos \varphi} \cdot \\ \cdot \frac{T_2 + T_1}{2} + \frac{\bar{W}_0}{h_0} \cdot \frac{T_2 + T_B}{2} = \frac{Q - Q_1}{\rho_w c_w h_0}, \end{aligned} \quad (51.4)$$

while for (T_B, T_1, T_w) the value of $\bar{V}_1$ in the second term on the left is replaced by $(\bar{V}_2, \bar{V}_{1p}, \bar{V}_{2p})$; similarly, $\Lambda - \lambda_0$ and T_1 in the third term on the left are replaced by $(\Lambda - \lambda_0, \lambda_{02}, \lambda_0)$ and (T_w, T_2, T_B) ; W_0, h_0 , and T_B in the fourth term are replaced by $(\bar{W}_0, \bar{W}_p, \bar{W}_p)$, $(h_0 - h, h_0, h_0 - h)$ and (T_2, T_w, T_1) ; $Q - Q_1$ and h_0 on the right-hand side are replaced by $(Q_1, Q_p - Q_{1p}, Q_{1p})$ and $(h - h_0, h_0, h - h_0)$, where Q and Q_p are the heat fluxes from the atmosphere determined from the atmospheric block, while Q_1 and Q_{1p} are the heat fluxes into the ocean depths, 'parametrized' as follows:

$$Q_1 = \frac{2\rho_w c_w k(T_2 - T_B)}{h}, \quad Q_{1p} = \frac{2\rho_w c_w k(T_1 - T_w)}{h}, \quad (51.5)$$

where $k = 5 \text{ cm}^2/\text{s}$ for $T_2 \geq T_B$ and $T_1 \geq T_w$ and $k = 5 \times 10^2 \text{ cm}^2/\text{s}$ in the opposite cases. The temperatures of the Southern and Arctic Oceans are found from the equations of the integral heat budget:

$$\begin{aligned} \cos \Phi_2 \int_0^{2\pi} \int_{-h}^0 (VT)_{\Phi_2} dz d\lambda = \frac{r_0}{\rho_w c_w} \int_{\Phi_1}^{\Phi_2} \int_0^{2\pi} Q \cos \varphi d\varphi d\lambda; \\ -\cos \Phi_3 \int_0^{2\pi} \int_{-h}^0 (VT)_{\Phi_3} dz d\lambda = \frac{r_0}{\rho_w c_w} \int_{\Phi_3}^{\Phi_4} Q \cos \varphi d\varphi d\lambda, \end{aligned} \quad (51.6)$$

where for the freezing temperature it was assumed that $Q = 0$. Calculations of the mean annual state of the oceans based on this model for the specified atmosphere yielded quite satisfactory results. The T_s values were within 1°C of the empirical values. The integral heat fluxes in the Pacific, Atlantic, and Indian Oceans were $Q = (1.38; 0.55; 0.72) \times 10^{15} \text{ W}$, and these were balanced by the flux from the Southern Ocean $Q = -2.65 \times 10^{15} \text{ W}$, while the Arctic Ocean was found to be frozen ($Q = 0$). The model gave anticyclonic gyres of water in the thermocline and cyclonic gyres in the lower layer in both hemispheres, with downwelling in the polar regions and upwelling in the middle latitudes and the tropics.

For the atmosphere the model adopted was a mean annual zonal model, averaged over the thickness of the atmosphere. It included equations describing the meridional transfer of heat and moisture:

$$-\frac{1}{r_0 \cos \varphi} \frac{\partial}{\partial \varphi} \left(\mathcal{K}_h \left| \frac{\partial \Theta}{\partial \varphi} \right| \frac{\partial \Theta}{\partial \varphi} \cos \varphi \right) =$$

$$= -\frac{g}{c_p p_s} \left(\frac{p_s}{p_{500}} \right)^{R/c_p} \sum_{i=1}^4 \sum_{k=1}^2 \Lambda_k q_{ik}, \quad (51.7)$$

$$-\frac{1}{r_0 \cos \varphi} \frac{\partial}{\partial \varphi} \left(\mathcal{K}_h \left| \frac{\partial \Theta}{\partial \varphi} \right| \frac{\partial s}{\partial \varphi} \cos \varphi \right) = \frac{g}{p_s} (E - P), \quad (51.8)$$

where Θ is the potential temperature, s is the specific moisture content, Λ_1 and Λ_2 are the fractions occupied by continents and oceans, q_1 and q_{2k} are the influx of short-wave radiation and the budget of long-wave radiation, q_{3k} is the vertical turbulent heat influx, q_{4k} is the latent heat of the phase transformations of moisture (which also contains a term with the mass budget for the ice sheets $\rho_i \mathcal{L}_{wi} (1 - \chi) W$, where χ is the fraction of sea in the sheet boundaries).

All the fluxes of heat and moisture on the right-hand sides of (51.7) and (51.8) were 'parametrized' the same way as in Section 45. T_s on land was determined (assuming zero heat conduction) and Q on the ocean was found from the somewhat simplified boundary condition (45.6). The latter now takes the form $Q = q_{1s} + q_{2s} - q_{3s} - \mathcal{L}E$, where $q_{1s} = (1 - \chi_0) (1 - A_0) (1 - A) S_0$ (here χ_0 is the opacity of the atmosphere and A_0 is its albedo; A is the albedo of the underlying surface, assumed to be on land 0.2 for $\zeta = 0$ and 0.8 for $\zeta > 0$ or $T_s < -10^\circ \text{C}$ and on the ocean 0.05 for $T_s > T_{wi}$ and 0.8 for $T_s \leq T_{wi}$; S_0 is the insolation) and $q_{2s} = \sigma (\nu^4 T^4 - T_s^4)$. The specific humidity at the underlying surface was taken to be saturating at the temperature T_s . The temperature at the boundary of the atmospheric boundary layer was assumed to be $T - \gamma(\Delta z - \zeta)$, where T , determined from the solution Θ of Equation (51.7), is the temperature at the mean level of the atmosphere, Δz is the distance from this level to the undisturbed boundary of the atmospheric boundary layer, and γ is the standard lapse rate. The precipitation P was found from (51.8) for $s \approx s_m(p_{500}/p_s)^2$.

The calculations according to this model of the atmosphere for the specified distributions of S_0 , T_w , A_0 , and ζ were found to reproduce the meridional distributions of T_s and Q quite satisfactorily. Although in the distribution of P no subtropical minimum was obtained, the P values in the high latitudes, which are the most important values *vis-à-vis* the ice sheets, showed a good fit with the empirical data. On the basis of this success, calculations were made of the annual fluctuations of the atmospheric characteristics, for a specified seasonal variation of the distributions of S_0 and T_w ; the results again showed a good fit with the data.

In the above model the continental ice sheets were calculated not according to Equation (50.10) but rather according to a formula ensuing from (50.11): $\zeta^* = BS^{1/4}$, $B = (W^*/2g\rho_i\mathcal{K})^{1/4}$. The latter gives the mean height ζ^* of the ice sheet over the sheet area S . The mass-budget equation of the ice sheet $\partial(\zeta^* S)/\partial t =$

$(1 - \chi) W^*$ S is then easily integrated, and after differentiating with respect to time the formulas

$$S_1 = r_0^2 \int_{-\pi/2}^{\Phi_1^*} \Lambda_1 \cos \varphi \, d\varphi; \quad S_2 = r_0^2 \int_{\Phi_2^*}^{\pi/2} \Lambda_1 \cos \varphi \, d\varphi, \quad (51.9)$$

where S_1 is the area of the southern sheet with a northern latitude boundary Φ_1^* , and S_2 is the area of a northern sheet with a southern boundary Φ_2^* , we obtain

$$\begin{aligned} \Lambda_1(\Phi_1^*) \frac{\partial \Phi_1^*}{\partial t} \cos \Phi_1^* &= \frac{4(1 - \chi) W^*}{5Br_0^{1/2}} \left[\int_{-\pi/2}^{\Phi_1^*} \Lambda_1 \cos \varphi \, d\varphi \right]^{3/4}; \\ -\Lambda_1(\Phi_2^*) \frac{\partial \Phi_2^*}{\partial t} \cos \Phi_2^* &= \frac{4(1 - \chi) W^*}{5Br_0^{1/2}} \left[\int_{\Phi_2^*}^{\pi/2} \Lambda_1 \cos \varphi \, d\varphi \right]^{3/4}. \end{aligned} \quad (51.10)$$

From the atmospheric model together with Equation (51.10), the climates of the Paleocene and Eocene 67 to 37 million years ago were calculated. The departures of T_w from the present-day values were $+13^\circ\text{C}$ in the polar regions and 0°C in the tropics. The climates of the Oligocene and the Early-to-Middle Miocene 37 to 15 million years ago were also calculated ($+5^\circ\text{C}$ and -8°C), as well as those of the Late Miocene and Early Pliocene 15 to 5 million years ago (0°C and 0°C). It was found that during the first of these periods glaciers existed only in a small area in the middle of Antarctica. In connection with the formation of a deep channel between Antarctica and Australia 38 million years ago, a general cooling took place and Antarctica became covered with ice. Then, due to the formation of the deep-water Drake Passage, and the completed development of the West Wind Drift, 12 million years ago, the latitudinal zonality of the climate became more marked, and the Arctic became covered with ice. At 70°N the mean annual temperature of the continents dropped from $+2^\circ\text{C}$ to -16°C .

The equations of the unified model of the atmosphere-ocean-land system were integrated on a five-degree grid, in such a way that in each time step of 1000 years for the glaciers the equations of the oceanic model were integrated in steps of 3 days over 10 years, and the corresponding quasisteady state of the atmosphere was sought. The entire system was integrated over several hundred thousand years. It was found that the climatic state of the system as a whole can be characterized quite completely by four parameters: the latitude boundaries of the ice sheets Φ_1^* and Φ_2^* and the surface temperatures T_A of the Atlantic at 45°N and T_S of the Southern Ocean.

Numerical experiments with different initial values of latitudes Φ_1^* and Φ_2^* and temperatures of all five oceans revealed a nonuniqueness of the climate, and the possibility of five steady states (Φ_1^* , Φ_2^* , T_A , T_S): (I) the present-day climate without

continental glaciations in the Northern Hemisphere, but with an ice-covered Antarctica ($\Phi_1^* = -72.5^\circ$, $\Phi_2^* = 77.5^\circ$), which is reproduced by the model quite satisfactorily in many details; (II) a climate without continental glaciations ($\Phi_1^* = -87.5^\circ$, $\Phi_2^* = 77.5^\circ$); (III) a glacial period ($\Phi_1^* = -72.5^\circ$, $\Phi_2^* = 40^\circ$); (IV) a climate with glaciation only in the Northern Hemisphere ($\Phi_1^* = -87.5^\circ$, $\Phi_2^* = 40^\circ$); (V) a 'white Earth' with ice-covered continents and oceans frozen almost to the bottom ($T_s = 226$ K, $T_a = 201$ K; the thickness of the ice sheets $\zeta^* = 3500$ m, sea level would be 1300 m lower, and the sea ice would be 2300 m thick; just to raise T_s to T_{wi} , it would be necessary to triple the solar constant!). In regions of continental ice the mean annual temperature would be 20°C lower.

The glaciations in each hemisphere turned out to be very stable. They cannot be liquidated (or brought into the 'white Earth' regime) by variations of the external parameters within reasonable limits. The nonglaciated states are, on the other hand, unstable: to change them into a glacial period, an initial extent of the glaciers of no less than 5° is sufficient. In order to produce a primary ice sheet by varying the solar constant, the latter has to be lowered by nearly 20%; under these conditions the glaciation in the Northern Hemisphere increases rapidly, no longer reaching just to 40°N but continuing almost to the equator. In mountainous regions a primary ice sheet can also appear for a slight reduction of the solar constant.

It should be mentioned at this point that the above model of the atmosphere—ocean—land system did not yield oscillatory solutions, corresponding to alternating glacial and nonglacial periods: for the steady-state solutions obtained, slight fluctuations in the insolation can cause only slight variations of the margins of the ice sheets, and no more. The reason for this shortcoming is the fact that the model contains only one high-inertia element: the continental ice sheets, whereas in order to generate oscillatory regimes the system would have to possess at least two related elements with comparable relaxation times.

Following the idea of Pollard (1982), we could take as our second such element the viscoplastic sublithospheric layer of the Earth's upper mantle, the *asthenosphere*. As a result of glacial loading, the blocks of the continental lithosphere resting on the asthenosphere sink, the margins of the ice sheet are below the level of the ocean, the ablation rate increases (to 20 m/year), and this leads to a rapid breakup of the ice sheet. This mechanism may explain the asymmetry in the evolution of the Pleistocene continental ice sheets (growth over tens of millennia and breakup over millennia), revealed by paleogeographers, as well as the postglacial elevation of Fennoscandia and the Hudson region of North America (apparently already by 400 to 600 m, with a prospective further elevation by 200 m until isostatic equilibrium is reached), with the formation at their margins of postglacial seas and lakes.

Let us assume that zonal currents in the asthenosphere are quasihydrostatic and nondivergent, with equilibrium between the gradients of the pressure and the viscosity force. At the bottom of the asthenosphere $z = -h_a$ we impose the conditions of adhesion $v = w = 0$, and at its disturbed surface $z = \zeta_a$ we set $v = 0$, $p = g\sigma_i\zeta$, and $\partial\zeta_a/\partial t = w$ (since $|\zeta_a| \ll h_a$, these conditions can be applied to the undisturbed surface $z = 0$). Then, exactly analogously to the derivation of

Equation (50.10), for ξ_a we obtain

$$\frac{\partial \xi_a}{\partial t} = \frac{1}{\tau_a \cos \varphi} \frac{\partial}{\partial \varphi} \left[\left(\frac{\partial \xi_a}{\partial \varphi} - \frac{\rho_i}{\rho_a} \frac{\partial \xi}{\partial \varphi} \right) \cos \varphi \right];$$

$$\tau_a = \frac{12\mu_a r_0^2}{g\rho_a h_a^3}, \quad (51.11)$$

where μ_a is the viscosity of the asthenosphere. If we assume for it a value of 10^{18} kg/(m s) and if we set $h_a \approx 200$ km, then we get $\tau \sim 30,000$ years, just as for the ice sheets. Note that the steady-state solution of this equation has the form $\xi_a = (\rho_i/\rho_a) \xi \approx \xi/3$, so that for $\xi \sim 3$ km the ice sheets may sink 1 km, which shows fairly good fit with the data presented above concerning the postglacial elevation of Fennoscandia and Canada.

Using (51.11), we must reckon the height of the surface of the ice sheets from the level $h - \xi_a$ of the relief h of the icefree land, 'dropped' to the depth ξ_a of sinking of the asthenosphere surface, at the same time introducing the appropriate corrections into Equation (50.10). In the model of the present section, instead of (50.10) we use (51.10), and a correction is needed only in the calculation of the temperature at the limit of the atmospheric boundary layer, which is now taken to be $T - \gamma (\Delta z - \xi - h + \xi_a)$. This means that during the sinking of the ice sheets it becomes warmer at the sheet surface and the melting of the sheets speeds up. However, the hydroablation is stronger: if in the marginal region of the sheet $\xi_a > h$ and $\rho_w(\xi_0 + \xi_a) > \rho_i \xi$ (where $\xi_0 = \rho_i \delta V / \rho_w S_w$ is the level of the ocean, which varies during the evolution of the ice sheets, so that it also has to be calculated; here δV is the ice-sheet volume and S_w is the area of the ocean), then the ablation in this region is increased by 20 m/year.

For integration of the equations of the model with an asthenosphere, the time step for (51.11) was taken to be 1000 years, just as for the ocean. Mountains 2.5 km high were assumed to be present in the region from 67.5°N to 72.5°N. Figure 51.2 depicts the results of the integration over 500,000 years for the Northern Hemisphere. Self-oscillations are seen to appear in such a system, their largest period being about 120,000 to 140,000 years.

The ice sheet, originating in the mountains at the northern edges of the continents and affecting the atmosphere via the albedo, advances southward to a latitude of 40°; beneath the sheet the asthenosphere sinks with a time lag, whereas ahead of the sheet it is lifted. At a latitude of 40° the sheet edge executes high-frequency oscillations due to $T-P$ feedback. Then the sheet rapidly breaks up, the asthenosphere slowly flows back up, and conditions for a new growth of the sheet are created. However, the asthenosphere is still folded downward, and the sheet advances southward only to a latitude of 62.5°. Then it breaks up, the asthenosphere returns almost to its initial state, and the entire cycle is repeated (although in Figure 51.2 a slight trend in the behavior of the system is also seen, so that it would be desirable to increase the integration time; it would also be desirable to vary the values of τ_a and, if feasible, to introduce seasonal variations into the model).

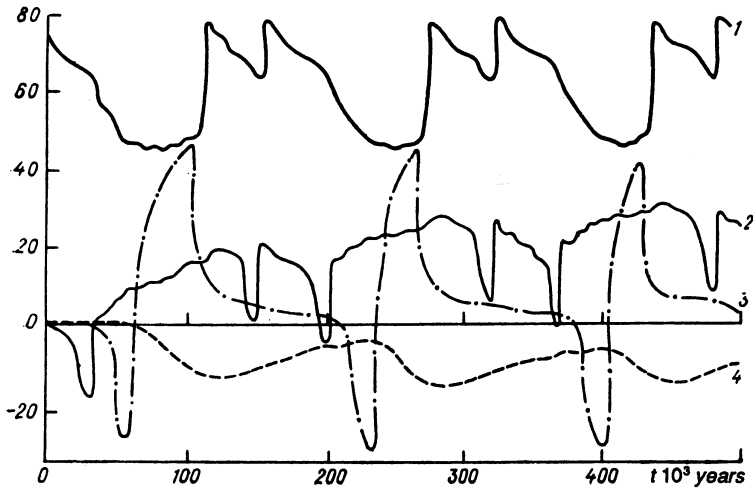


Fig. 51.2. Glacial periods. (1) Latitude (in degrees) of southern boundary of northern ice sheet; (2, 3, 4) plots of ξ_a (in tens of km) at latitudes of 62.5°, 47.5°, and 32.5°.

The conditions for the onset of a glacial period are as follows: (1) the arrangement of the continents must be such that in the atmosphere—ocean—land system (without the asthenosphere) at least two stable states are possible: a glacial state and a nonglacial state (for a geography close to that at present, the northern margins of the continents must be at high latitudes and the Arctic Basin must communicate with the ocean); (2) the relief of the land must be appropriate for the formation of primary ice sheets, including the albedo mechanism: the land must be quite close to the pole or must have quite high mountains (so that astronomical fluctuations of the distribution of insolation can have an effect, since it is such fluctuations which cause cool summers and snowy winters in the regions of origin of glaciers); (3) the relief of the continents in the middle latitudes must be predominantly level.

Such conditions could have occurred repeatedly during the geological history of the Earth. Actually, the geological chronicle contains traces of glacial periods of the Early Proterozoic 2300 to 2100 million years ago, the Late Riphean 850 to 750 million years ago, possibly the Ordovician—Devonian, undoubtedly the Carboniferous—Permian 300 to 250 million years ago, and, finally, for the last 38 million years Antarctica, and during the last 1.5 to 2 million years the Pleistocene glaciations of the northern continents (of which the Greenland Ice Sheet remains today as a vestige).

Fluid Dynamics of Planetary Interiors

52. FLUID DYNAMICS OF THE EARTH'S CORE

Following Monin (1973) and Gavrilin and Monin (1974), let us consider a simplified model of the Earth in three layers: (1) an outer solid spherical layer (mantle and crust) with an outer radius $r_1 \approx 6370$ km (radius of Earth) and an inner radius $r_2 \approx 3470$ km; (2) a liquid outer layer of the core; and (3) a solid spherical core with a radius $r_3 \approx 1470$ km, floating in the middle of the liquid occupying the innermantle space. The quantities pertaining to these three layers will be denoted by indexes 1, 2, 3. The densities of the mantle ρ_1 and the solid core ρ_3 are assumed to be functions of the distance from their centers.

Let us attempt to describe the motions inside the Earth, comprising the rotation of the mantle (that is, the *rotation of the Earth*, observed using astronomical methods), magnetohydrodynamic currents in the liquid layer, and displacements and rotations of the solid core. These motions of the three layers interact due to the hydromagnetic stresses at both boundaries of the liquid layer and due to the variations of the intrinsic gravitational field during displacements of the solid core.

A number of simplifications have been introduced in this model. The rotation-caused oblateness of the Earth is neglected (and thus the tendency toward precession, although the role of the latter, as a source of heat release, should be taken into account). The tidal and convective motions in the mantle and inner core are excluded. The effects of the movements of the atmosphere and ocean are ignored, too, as well as the gravitational effects of external forces of the Solar System. These simplifications do not distort much the class of motions in the Earth's interior which are of interest to us here.

The equations of motion in all three media are taken in the form (5.1), (5.2), except that in the mantle and solid core we omit the stresses $p\delta_{ik} - \sigma'_{ik}$ and we can write

$$\mathbf{v}_1 = \boldsymbol{\omega}_1 \times \mathbf{r}; \quad \mathbf{v}_3 = \dot{\mathbf{x}} + \boldsymbol{\omega}_3 \times (\mathbf{r} - \mathbf{x}), \quad (52.1)$$

where $\mathbf{r}$ is the radius vector in an absolute frame of reference, moving with the center of the mantle; $\mathbf{x}(t)$ is the coordinate of the center of the solid core; $\boldsymbol{\omega}_1(t)$ and $\boldsymbol{\omega}_3(t)$ are the angular velocities of rotation of the mantle and the solid core. The equations of motion must be supplemented by the magnetohydrodynamic equation of evolution of the magnetic field (equation of induction (3.6)). From these equations the local energy equations (5.3)–(5.5) are derived.

By integrating these equations over the volumes of the mantle, liquid layer, and solid core, we obtain 9 equations for the kinetic energy K , magnetic energy M , and internal energy E of each of the three layers. Figure 52.1 gives the corresponding diagram of the energy conversions (with quantitative estimates, which will be derived below). One of the most interesting parts of the diagram gives the rate of energy conversion $K_2 \rightarrow M_2$, equal to $\int \Pi'_{ik} (\partial v_i / \partial x_k) dV_2$ and describing the generation of the geomagnetic field by the hydromagnetic dynamo in the liquid layer.

First let us consider the displacement of the solid core. The equation of motion of its center is obtained by integrating (5.1) over the volume V_3 and using for $\mathbf{v}$ the

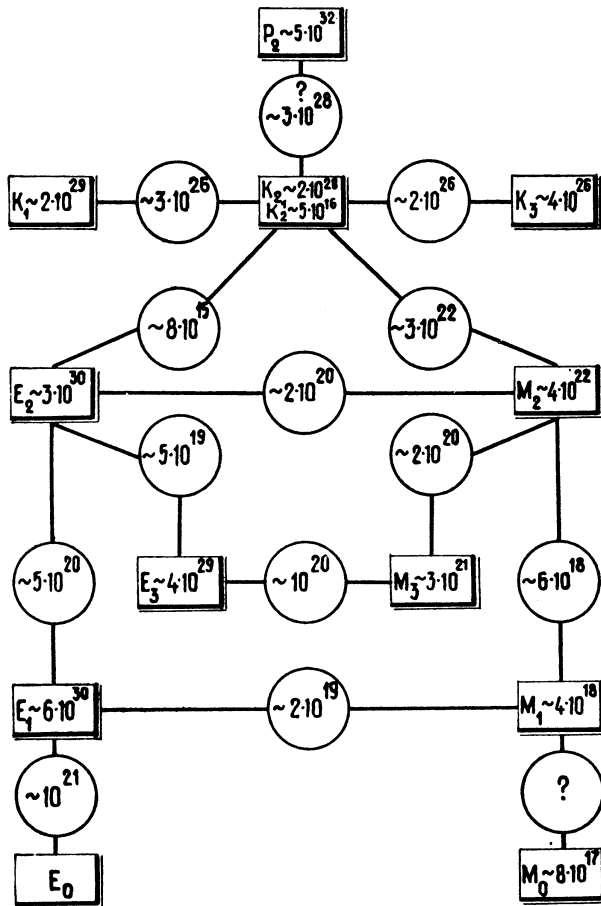


Fig. 52.1. Energy conversions in the Earth's interior. Rectangles enclose energy components in joules; circles enclose rates of energy conversion in joules per century.

second formula (52.1); for the volume integral of $\text{div } \Pi$ we use the Gauss formula

$$m_3 \ddot{\mathbf{x}} = - \int \mathbf{n} \cdot \Pi \, dS_{23} + \int \rho \nabla \Phi_2 \, dV_3, \quad (52.2)$$

where S_{23} is the boundary between the liquid layer and the solid core, and it is taken into account that the gravitational potential Φ_1 , produced by the mass of the mantle, is constant in the innermantle space, while the intrinsic potential Φ_3 of the core mass does not affect its displacement. Potential Φ_2 is calculated for the simplifying assumption $\rho_2 \approx \text{const}$ (actually ρ_2 increases slightly with depth: approximately from 9.9×10^3 to $12.2 \times 10^3 \text{ kg/m}^3$). Then at points $\mathbf{r}$ of the solid core

$$\begin{aligned} \Phi_2(\mathbf{r}) &= G\rho_2 \left[\int_{(2)+(3)} \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} - \int_{(3)} \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{x} - \mathbf{r}'|} \right] \\ &= 2\pi G\rho_2 \left[\left(r_2^2 - \frac{1}{3} r^2 \right) - \left(r_3^2 - \frac{1}{3} |\mathbf{r} - \mathbf{x}|^2 \right) \right] \\ &= \text{const} + \frac{2}{3} \pi G\rho_2 (|\mathbf{r} - \mathbf{x}|^2 - r^2), \end{aligned} \quad (52.3)$$

where $G = 6.67 \times 10^{-11} \text{ m}^3/(\text{kg s}^2)$ is the gravitational constant. Thus, for the second term on the right-hand side of (52.2), we obtain the quantity $-Gm_3\bar{m}_3r_3^{-3}\mathbf{x}$, where $\bar{m}_3 = \frac{4}{3}\pi r_3^3\rho_2$ is the mass which the solid core would have if its density were ρ_2 . Consequently, during displacements of the solid core, the gravitational field of the liquid layer produces a restoring force (proportional to the displacements) and thus a tendency toward an oscillatory nature of these displacements.

Of the forces represented in the first term on the right-hand side of (52.2), let us single out the hydrostatic pressure $p_0 = \rho_2 \Phi_2(\mathbf{r}) + \text{const}$; the other forces will be called conditionally the resistance forces. The contribution of the hydrostatic pressure $-\int \mathbf{n} p_0 dS_{23}$ is equal to $-\int \rho_2 \nabla \Phi_2 \, dV_3 = G\bar{m}_3^{-2}r_3^{-3}\mathbf{x}$, so that, ignoring for the moment the resistance forces, for the period of the oscillations of the solid core, produced by gravity and the hydrostatic pressure, from (52.2) we obtain

$$\tau_0 = 2\pi \left[\frac{4}{3} \pi G\rho_2 \left(1 - \frac{\bar{m}_3}{m_3} \right) \right]^{-1/2} = \left[\frac{G\rho_2}{3\pi} \left(1 - \frac{\rho_2}{\bar{\rho}_3} \right) \right]^{-1/2}, \quad (52.4)$$

where $\bar{\rho}_3$ is the mean density of the solid core. For example, for $\rho_2 = 10.9 \times 10^3 \text{ kg/m}^3$ and $\bar{\rho}_3 = 12.1 \times 10^3 \text{ kg/m}^3$ we have $\tau_0 \approx 3 \text{ h}$, that is, the period turns out to be very short. The resistance forces increase the period, but they damp the oscillations.

For example, in the case of a Stokes resistance $\mathbf{F} = 6\pi r_3\rho_2\nu_2\dot{\mathbf{x}}$, the period increases by a factor of $[1 - (3r_3\rho_2\nu_2\tau_0/2m_3)^2]^{-1/2}$ and for $\nu_2 > (8\pi/3)(\bar{\rho}_3/\rho_2)r_3^2\tau_0^{-1} \approx 6 \times 10^{-8} \text{ m}^2/\text{s}$ oscillations become impossible. Much lower

estimates of the viscosity ν_2 of the liquid layer are usually assumed, but we will nevertheless assume that $\mathbf{x} = 0$ in subsequent calculations.

With this condition let us derive the equation of rotation of the solid core, having multiplied (5.1) on the left vectorially by $\mathbf{r}$ and then integrating over V_3 . The left-hand side then becomes $I_3 \dot{\boldsymbol{\omega}}_3$, where $I_3 = (8\pi/3) \int_0^{r_3} \rho r^4 dr \approx 1.4 \times 10^{35} \text{ kg m}^2$ is the moment of inertia of the core. Now, assuming the gravitational field to be symmetric, we have $\Phi = \Phi(r)$ and $\mathbf{r} \times \nabla \Phi = 0$. For $\int \mathbf{r} \times \text{div } \Pi dV$ let us use the expression $\int \mathbf{r} \times (\mathbf{n} \cdot \Pi) dS + 2 \int \boldsymbol{\omega} dV$, analogous to the Gauss formula, where $\omega_i = \frac{1}{2} \varepsilon_{ijk} \Pi_{kj}$ in our case goes to zero, because of the symmetry of tensor Π . On S_{23} we have $\mathbf{n} = \mathbf{r}/r$ and, using (5.2), we obtain

$$\begin{aligned} \mathbf{n} \cdot \Pi = \rho_2 \nu_r \mathbf{v} + \left[p - \rho_2 \left(\nu_1 - \frac{2}{3} \nu \right) \text{div } \mathbf{v} + \frac{H^2}{8\pi} \right] \frac{\mathbf{r}}{r} - \\ - \rho_2 \nu_2 \left(\frac{\partial \mathbf{v}}{\partial r} - \frac{\mathbf{v}}{r} + \nabla v_r + \frac{v_r}{r} \mathbf{r} \right) - \frac{H_r \mathbf{H}}{8\pi} \end{aligned} \quad (52.5)$$

Then, taking into account that on S_{23} it must be true that $\mathbf{v} = \boldsymbol{\omega}_3 \times \mathbf{r}$, so that $v_r = 0$ and $\partial \mathbf{v} / \partial r - \mathbf{v} / r = \partial (\mathbf{v} - \boldsymbol{\omega}_3 \times \mathbf{r}) / \partial r$, we get finally

$$I_3 \dot{\boldsymbol{\omega}}_3 = \int \mathbf{r} \times \left[\rho_2 \nu_2 \frac{\partial}{\partial r} (\mathbf{v} - \boldsymbol{\omega}_3 \times \mathbf{r}) + \frac{H_r \mathbf{H}}{4\pi} \right] dS_{23}. \quad (52.6)$$

The equation of rotation of the mantle is derived similarly. After vector multiplication of (5.1) by $\mathbf{r}$ and integration over V_1 , the left-hand side becomes $I_1 \dot{\boldsymbol{\omega}}_1$, where $I_1 = (8\pi/3) \int_{r_1}^{r_2} \rho r dr \approx 7.25 \times 10^{37} \text{ kg m}^2$ is the moment of inertia of the mantle. The right-hand side becomes equal to the difference of the integrals of $\mathbf{r} \times (\mathbf{r} \cdot \Pi / r)$ over the surfaces $r = r_2$ and $r = r_1$. The hydrodynamic stresses at the surface of the solid Earth, that is, the flux of mechanical momentum from the atmosphere and ocean, as discussed above, will here be neglected (it produces, for instance, the annual oscillations of $\boldsymbol{\omega}_1$, but they have to be calculated separately). The magnetic field outside the Earth is a potential field, so that the integral over the surface $r = r_1$ of $\mathbf{r} \times (\mathbf{r} \cdot \Pi' / r)$, equal to the integral of $\mathbf{r} \times \text{div } \Pi'_p$ over the entire volume of the Earth, vanishes, since $\text{div } \Pi'_p = 0$. As a result, similarly to (52.6) we obtain

$$I_1 \dot{\boldsymbol{\omega}}_1 = - \int \mathbf{r} \times \left[\rho_2 \nu_2 \frac{\partial}{\partial r} (\mathbf{v} - \boldsymbol{\omega}_1 \times \mathbf{r}) + \frac{H_r \mathbf{H}}{4\pi} \right] dS_{12}. \quad (52.7)$$

The right-hand side of this equation should account for the observed fluctuations of $\boldsymbol{\omega}_1$ with periods of decades, which according to Westin (1953) correlates with the fluctuations of the velocity of the western drift of the nondipole component of the geomagnetic field. For example, for 1910–1930 an acceleration of $\dot{\boldsymbol{\omega}}_1 \approx 7 \times$

10^{-21} s^{-2} was detected, and if we specify the right-hand side of (52.7) in the form $4\pi r_2^3 \rho_2 u_*^2$, then on S_{12} a 'friction velocity' $u_* \approx 0.03 \text{ cm/s}$ corresponds to this acceleration. For the Earth as a whole, excluding external forces, the angular momentum is conserved:

$$I_1 \omega_1 + \int \mathbf{r} \times \rho \mathbf{v} dV_2 + I_3 \omega_3 = \text{const.} \quad (52.8)$$

Since $I_1 : I_2 : I_3 \approx 725 : 86 : 1.4$, it is clear that moderate variations in the rotation of the liquid layer will lead to slight variations in the rotation of the mantle but major variations in the rotation of the solid core.

In order to estimate the energy components of the solid core, we will integrate over its volume equations (5.3)–(5.5), on the right-hand sides of which here only the terms with v_m and Q are different from zero:

$$\begin{aligned} \dot{K}_3 &= \int (\omega_3 \times \mathbf{r}) \left[\rho_2 v_2 \frac{\partial}{\partial r} (\mathbf{v} - \omega_3 \times \mathbf{r}) + \frac{H_3 \mathbf{H}}{4\pi} \right] dS_{23}; \\ \dot{M}_3 &= \frac{1}{4\pi} \int v_m (\mathbf{H} \times \text{curl } \mathbf{H})_r dS_{23} + \frac{1}{4\pi} \int v_m |\text{curl } \mathbf{H}|^2 dV_3; \\ \dot{E}_3 &= - \int q_r dS_{23} + \frac{1}{4\pi} \int v_m |\text{curl } \mathbf{H}|^2 dV_3 + \int Q dV_3. \end{aligned} \quad (52.9)$$

Let us evaluate the terms, using the characteristics of the steady-state model of the geomagnetic dynamo devised by Bullard and Gellman (1954): $u \sim 2.5 \times 10^{-4} \text{ m/s}$, $H_T \sim 500 \text{ gauss}$ (*toroidal field*; some investigators assign considerably lower values to it), $H_r \sim 4 \times 10^{-4} \text{ tesla}$, $H^2/8\pi \sim 2.3 \times 10^2 \text{ J/m}^3$, $\frac{1}{2} \rho u^2 \sim 3 \times 10^{-4} \text{ J/m}^3$ (that is, the kinetic-energy density of the relative motions is six orders of magnitude less than for the magnetic field). For $\omega_3 \sim \omega_1 = 7.29 \times 10^{-5} \text{ s}^{-1}$ we get $K_3 = \frac{1}{2} I_3 \omega_3^2 \sim 3.7 \times 10^{26} \text{ J}$, $M_3 \sim (H^2/8\pi) V_3 \sim 3 \times 10^{21} \text{ J}$, and $E_3 \sim c_3 m_3 T_3 \sim 3.7 \times 10^{29} \text{ J}$ (where $c_3 \sim 4.6 \times 10^2 \text{ J/(kg K)}$, $T_3 \sim 5 \times 10^3 \text{ K}$, and $m_3 \approx 1.6 \times 10^{23} \text{ kg}$).

Let us assume for v_2 near S_{23} the estimate of Zharkov (1962) of $10^7 \text{ m}^2/\text{s}$ (although much lower estimates are often assumed). The contribution to $\dot{K}_3$ of the viscous stresses will then be $4\pi \omega_3 r_3^2 \rho_2 v_2 u \sim 5 \times 10^{13} \text{ kW}$, while the contribution to $\dot{K}_3$ of the magnetic stresses will be of the order of $\omega_3 r_3^2 H_T H_r \sim 4.6 \times 10^{13} \text{ kW}$. The time of the variations of ω_3 is found to be of the order of $K/\dot{K} \sim 250 \text{ years}$, that is, it is very short, and the rotation of the solid core can be assumed to be adapted to the regime of the hydromagnetic processes in the liquid layer.

If we assume $v_m \sim 2.5 \text{ m}^2/\text{s}$, then we get $\dot{M}_3 \sim \frac{1}{4\pi} v_m r_3 H_T^2 \sim 3 \times 10^7 \text{ kW}$, so that the time of the ohmic dissipation is $M_3/\dot{M}_3 \sim 3000 \text{ years}$. To remove the Joule heat from the solid core we would require a heat flux at S_{23} with a density $q \sim \dot{M}_3/4\pi r_3^2 \sim 10^{-3} \text{ J/(m}^2\text{s)}$, to which for an adiabatic temperature gradient in the liquid layer $|\partial T/\partial r| \approx 1.4 \times 10^{-4} \text{ K/m}$ corresponds a thermal conductivity of the order of $1.4 \times 10^{-5} \text{ m}^2/\text{s}$.

For the mantle the equations analogous to (52.9) have the form

$$\begin{aligned}\dot{K}_1 &= - \int (\boldsymbol{\omega}_1 \times \mathbf{r}) \left[\rho_2 \nu_2 \frac{\partial}{\partial r} (\mathbf{v} - \boldsymbol{\omega}_1 \times \mathbf{r}) + \frac{H_r \mathbf{H}}{4\pi} \right] dS_{12}; \\ \dot{M}_1 &= -\dot{M}_0 - \frac{1}{4\pi} \int \nu_m (\mathbf{H} \times \text{curl } \mathbf{H})_r dS_{12} - \frac{1}{4\pi} \int \nu_m |\text{curl } \mathbf{H}|^2 dV_1; \\ \dot{E}_1 &= - \int q_r dS_{01} + \int q_r dS_{12} + \frac{1}{4\pi} \int \nu_m |\text{curl } \mathbf{H}|^2 dV_1 + \int Q dV_1,\end{aligned}\quad (52.10)$$

where M_0 is the energy of the potential magnetic field $\mathbf{H} = \nabla W$ outside the Earth (determined completely by the field H_r at S_{01} by the solution of the Neumann problem for the Laplace equation $\Delta W = 0$). For a mean density of the mantle $\rho_1 \approx 4.37 \times 10^3 \text{ kg/m}^3$ its mass $m_1 = 4.2 \times 10^{24} \text{ kg}$, volume $V_1 = 9.2 \times 10^{20} \text{ m}^3$, outer surface area $S_{01} = 5.1 \times 10^{14} \text{ m}^2$, inner surface area $S_{12} = 1.5 \times 10^{14} \text{ m}^2$, and kinetic energy $K_1 = \frac{1}{2} I_1 \omega_1^2 = 1.9 \times 10^{29} \text{ J}$. For a dipole external field with a potential $W = h_1^0 r_1^3 r^{-2} \cos \theta$, $h_1^0 = -0.3 \times 10^{-4} \text{ tesla}$, we obtain $M_0 = \frac{1}{2} (h_1^0)^2 r_1^3 \approx 7.8 \times 10^{17} \text{ J}$ and $M_1 = M_0 [(r_1/r_2)^2 - 1] \approx 4.3 \times 10^{18} \text{ J}$. For $c_1 = 7.1 \times 10^2 \text{ J/(kg K)}$ and $T_1 \sim 2 \times 10^3 \text{ K}$ we have $E_1 \sim c_1 m_1 T_1 \sim 6 \times 10^{30} \text{ J}$.

The contribution of the magnetic stresses to $\dot{K}_1$ are approximately the same as to $\dot{K}_3$, except that the effect of the viscous stresses is much less, since ν_2 near S_{12} is much lower than near S_{23} . In order to obtain as a result of the viscous stresses the above value $u_* \approx 3 \times 10^{-4} \text{ m/s}$ at S_{12} , we have to assume $\nu_2 \sim u_*^2 \delta / u$ there, and for a boundary-layer thickness $\delta \sim 10^6 \text{ m}$ we obtain $\nu_2 \sim 3 \times 10^2 \text{ m}^2/\text{s}$. Ohmic dissipation $\dot{M}_1 \sim \nu_m \delta_1 H_T^2$ takes place only in the lower layer of the mantle $\delta_1 \sim 10^7 \text{ cm}$ thick, above which the electrical conductivity is negligibly small; in this layer ν_m is apparently three orders of magnitude higher than in the liquid layer, but H_T is attenuated rapidly there.

For the liquid layer, on condition that $\text{div } \mathbf{v}_2 = 0$, the equations analogous to (52.9) take the form

$$\begin{aligned}\dot{K}_2 &= -\dot{K}_1 - \dot{K}_3 + \int \rho_2 \mathbf{v} \cdot \nabla \Phi dV_2 - \mathcal{D} - \frac{1}{4\pi} \int \mathbf{H} \cdot (\mathbf{H} \cdot \nabla) \mathbf{v} dV_2; \\ \dot{M}_2 &= \frac{1}{4\pi} \int \nu_m (\mathbf{H} \times \text{rot } \mathbf{H})_r dS_{12} - \frac{1}{4\pi} \int \nu_m (\mathbf{H} \times \text{curl } \mathbf{H})_r dS_{23} \\ &\quad - \mathcal{D}_m + \frac{1}{4\pi} \int \mathbf{H} \cdot (\mathbf{H} \cdot \nabla) \mathbf{v} dV_2; \\ \dot{E} &= - \int q_r dS_{12} + \int q_r dS_{23} + \int Q dV_2 + \mathcal{D} + \mathcal{D}_m,\end{aligned}\quad (52.11)$$

where $\mathcal{D}$ is the viscous dissipation of kinetic energy, and $\mathcal{D}_m$ is the ohmic dissipation of magnetic energy in the liquid layer. The volume of the liquid layer

$V_2 = 1.6 \times 10^{20} \text{ m}^3$, and for a mean density $\bar{\rho}_2 = 10.9 \times 10^3 \text{ kg/m}^3$ it will have a mass $m_2 = 1.7 \times 10^{24} \text{ kg}$ and a kinetic energy of rotation $\bar{K}_2 = \frac{1}{2} I_2 \omega_1^2 \sim 2.3 \times 10^{28} \text{ J}$. On the basis of the foregoing estimates, the kinetic energy of the relative motions is found to be $K'_2 \sim 4.8 \times 10^{16} \text{ J}$, the magnetic energy $M_2 \sim 3.7 \times 10^{22} \text{ J}$, and, for $c_2 \sim c_3 \sim 4.6 \times 10^2 \text{ J/(kg K)}$ and $T_2 \sim 4.3 \times 10^3 \text{ K}$, the internal energy $E_2 \sim c_2 m_2 T_2 \sim 3.4 \times 10^{30} \text{ J}$.

The initial cause of the relative motions in the liquid layer (which in turn produce the geomagnetic dynamo) may well be the liberation of the potential energy $\dot{K}_2 \sim m_2 g' v_r$, where g' is the effective buoyancy acceleration. For instance, for $g' \sim 10^{-1} \text{ m/s}^2$ and $v_r \sim 10^{-4} \text{ m/s}$ we would have $K_2 \sim 10^{16} \text{ kW}$, which would, with a little left over, be enough to compensate all the energy losses. The rate of the mutual conversions of K_2 and M_2 (the last term in the two first equations (52.11)), that is, the rate of generation of the geomagnetic field $\dot{M}_2 \sim \frac{1}{3} r^2 u H_T^2$, is found to be of the order of 10^{10} kW . The Joule losses, heat fluxes, and thermal conductivity were estimated as in the case of the solid core.

All the above quantitative estimates were presented in Figure 52.1 (but many of these are to be considered just as an example, since as yet no reliable values have been obtained even for the order of magnitude of such quantities as u , H_T , and v_r). The following section will present a qualitative discussion of the hydromagnetic processes going on in the liquid layer of the Earth.

53. HYDROMAGNETIC DYNAMOS

The mechanisms whereby a magnetic field is generated by currents of electrically conducting fluids are called *hydromagnetic dynamos*. According to present-day ideas, the existence of a magnetic field at the Sun (consisting of a highly electroconductive plasma) and the Earth (having a molten layer in its core) can be explained in terms of dynamo mechanisms. The same is true of Mercury (apparently for the same reason), Jupiter (containing fluid metallized hydrogen), probably Saturn (for the same reason), Uranus (with its anomalously large 55° inclination of the magnetic axis to the rotation axis), magnetic stars, and the Galaxy as a whole.

Hydromagnetic dynamos are described by the equations of magnetohydrodynamics, that is, the equations of motion (3.2), including the Lorentz force (3.3), together with the induction equation (3.6) and the condition $\text{div } \mathbf{H} = 0$. Let us mention some special features of these equations: (1) because of the effect of the magnetic field on the motion, their solutions include waves of a new type, namely magnetohydrodynamic waves (Alfvén, 1942), the role of which for dynamos should be clarified; (2) for problems of the generation of magnetic fields of planets and stars, spherical geometry is typical; (3) for dynamos it is important that the motion not possess 'reflective symmetry', that is, that the *density of helicity* $\mathbf{u} \cdot \text{curl } \mathbf{u}$ be different from zero; (4) the theory of dynamos should explain the tendency of the global magnetic fields to exhibit polarity reversals as observed on the Sun and at the Earth.

Let us begin with the waves. First, for simplicity, let us consider a nonrotating ($\omega = 0$) homogeneous fluid ($\rho = \text{const}$) without viscosity and ohmic losses

($\nu = \nu_m = 0$), taking into account only the Lorentz force, out of all the external forces; the notation $\mathbf{h} = (4\pi\rho)^{-1/2} \mathbf{H}$ will be used here. Then it is easy to see that the equations of magnetohydrodynamics have solutions $\mathbf{u} = \pm \mathbf{h}_1$, $\mathbf{h} = \mathbf{h}_0 + \mathbf{h}_1$, where $\mathbf{h}_0 = \text{const}$ and $\mathbf{h}_1$ is an arbitrary function of $\mathbf{x} \pm \mathbf{h}_0 t$. If it is localized, then these solutions describe stationary waves, propagating with the Alfvén velocity $\pm \mathbf{h}_0$ (for $h_1 \ll h_0$ they are representable as superpositions of plane harmonic transverse waves with two independent plane polarizations and with frequencies $\sigma = \mathbf{k} \cdot \mathbf{h}_0$).

In these waves the densities of kinetic and magnetic energy $\rho u^2/2$ and $\rho h_1^2/2 = H_1^2/8\pi$ are identical. The integrals $\frac{1}{2} \int |\mathbf{u} \pm \mathbf{h}_1|^2 dV$ do not vary with time; consequently, the same is true of their halfsum (the total energy of the waves) and halfdifference $\int \mathbf{u} \cdot \mathbf{h}_1 dV$ (the so-called *mutual helicity* of fields $\mathbf{u}$ and $\mathbf{h}_1$, serving as a measure of the coupling of the vortex lines of field $\mathbf{u}$ with the force lines of field $\mathbf{h}_1$). These waves apparently do not participate in the hydromagnetic dynamos and during the description of the latter can be ‘filtered out’ by retaining in the projections of induction equation (3.6) only one derivative with respect to time entering into (4.7) (Monin, 1971).

Now let us introduce rotation, but at the same time considering only the linearized equations

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + 2\boldsymbol{\omega} \times \mathbf{u} &= -\nabla P_1 + (\mathbf{h}_0 \cdot \nabla) \mathbf{h}_1 + \nu \Delta \mathbf{u}; \\ \frac{\partial \mathbf{h}_1}{\partial t} &= (\mathbf{h}_0 \cdot \nabla) \mathbf{u} + \nu_m \Delta \mathbf{h}_1 \end{aligned} \quad (53.1)$$

(where P_1 includes both the hydrodynamic pressure and the magnetic pressure). For plane waves (transverse waves, in view of the solenoidality) we obtain the relation $\text{curl } \mathbf{u} = \pm k \mathbf{u}$ (where k is the wave number), signifying that the wave possesses circular polarization ($|\mathbf{u}| = \text{const}$, and vector $\mathbf{u}$ turns along the direction $\mathbf{k}$). The helicity $\langle \mathbf{u} \cdot \text{curl } \mathbf{u} \rangle = \pm k |\mathbf{u}|^2 e^{-2|\sigma|t}$ is here a maximum. Such waves may produce dynamos. In the absence of a magnetic field and viscosity, only inertial oscillations would remain in this case. The onset of a magnetic field distorts them somewhat, but, and this is the main thing, a new class of waves is created, known as *magnetostrophic waves*, in which $2\boldsymbol{\omega} \times \mathbf{u} \approx -\nabla P_1 + (\mathbf{h}_0 \cdot \nabla) \mathbf{h}_1$. The dispersion relation for these waves has the form

$$\sigma = -i\nu_m k^2 \pm \frac{1}{2} k (\mathbf{k} \cdot \mathbf{h}_0)^2 (\mathbf{k} \cdot \boldsymbol{\omega})^{-1}. \quad (53.2)$$

Now let us add the buoyancy force and let us consider the so-called MBC waves (Braginskii, 1964), in which the magnetic, buoyancy, and Coriolis forces approximately compensate one another:

$$2\boldsymbol{\omega} \times \mathbf{u} \approx -\nabla P_1 + (\mathbf{h}_0 \cdot \nabla) \mathbf{h}_1 - \alpha T_1 \mathbf{g}, \quad (53.3)$$

where T_1 is the temperature perturbation, satisfying the linearized equation $\partial T_1 / \partial t - \Gamma w \approx \alpha \Delta T_1$, in which $(-\Gamma) < 0$ is the temperature gradient in a quiescent unstably stratified medium. The constant field $\mathbf{h}_0$ will here be assumed to be horizontal. Let us consider horizontally propagating and vertically standing plane

waves in a flat layer $|z| \leq z_0$ (at the boundaries of which $w = \partial u / \partial z = \partial v / \partial z = T_1 = h_{1z} = \partial h_{1x} / \partial z = \partial h_{1y} / \partial z = 0$), in which $u, v, P_1, h_{1x}, h_{1y}$ depend on z as $\cos nz$, while w, T_1, h_{1z} depend on z as $\sin nz$, where nz_0/π is an integer. For the frequencies σ of these waves we obtain the cubic equation

$$(-i\sigma + \chi k^2)[1 + Y^2(-i\sigma + \nu_m k^2)^2] - Z(-i\sigma + \nu_m k^2) = 0;$$

$$Y = \frac{2\omega n}{h_0^2 k_y^2 k}, \quad Z = \alpha \Gamma g \frac{(k_x^2 + k_y^2)}{h_0^2 k_x^2 k^2}, \quad (53.4)$$

where the modes of interest are unstable modes with $\text{Im } \sigma > 0$ (for $\text{Re } \sigma = 0$ they are nonoscillating). The y axis can conveniently be directed toward the east, and then $k_y = Nm_1$, $N = 0, \pm 1, \pm 2, \dots$. Approximation (53.3) is valid only for low-frequency modes $|\sigma| \ll \omega$. Dissipative factors ν_m and χ are important in this case, and it can be shown that for $q = \chi/\nu_m < 2$ or $Q = 2\omega\nu_m/h_0^2 \geq 3.273$ the first of the unstable modes to be excited are the nonoscillating modes, whereas in the opposite case in the (q, Q) plane there exists a region in which oscillating modes are excited first. Note, finally, the helicity of the unstable modes:

$$\langle \mathbf{u} \cdot \text{curl } \mathbf{u} \rangle = \frac{1}{2} \text{Re}[i(k_x v - k_y u)\omega^*] e^{2\sigma_1 t} \sin 2nz, \quad (53.5)$$

where the averaging is over the horizontal planes, and on the right-hand side u, v, w are the complex amplitudes of the waves.

Next let us consider the specifics of the spherical geometry, in which the solenoidal vector fields $\mathbf{A}$ can be conveniently represented as the sum $\mathbf{A}_T + \mathbf{A}_P$ of a *toroidal component* and a *poloidal component*, defined by the formulas:

$$\mathbf{A}_T = \text{curl}(T\mathbf{r}) = -\mathbf{r} \times \nabla T; \quad T = -\Delta_h^{-1}(\mathbf{r} \cdot \text{curl } \mathbf{A}),$$

$$\mathbf{A}_P = \text{curl curl}(P\mathbf{r}) = -\text{curl}(\mathbf{r} \times \nabla P); \quad P = -\Delta_h^{-1}(\mathbf{r} \cdot \mathbf{A}), \quad (53.6)$$

where $T(\mathbf{r})$ and $P(\mathbf{r})$ are the corresponding scalar fields, and $\Delta_h \approx (r \times \nabla)^2$ is the momentum operator, that is, a two-dimensional Laplace operator on a sphere of unit radius (its eigenvalues are $-n(n+1)$), with the corresponding eigenfunctions $S_n = \sum_{m=0}^n A_n^m P_n^m(\cos \theta) e^{im\lambda}$ with arbitrary complex constants A_n^m . Vector lines $\mathbf{A}_T$ lie on spheres $r = \text{const}$, while $\mathbf{A}_P$, on the other hand, has a nonzero radial component. If field $\mathbf{A}$ is axisymmetric (that is, invariant relative to rotations about the axis), then T and P do not depend on λ , and we get

$$\mathbf{A}_T = (0, 0, A_\lambda), \quad A_\lambda = -\frac{\partial T}{\partial \theta},$$

$$\mathbf{A}_P = \left(-\frac{1}{r^2 \sin \theta} \frac{\partial \chi}{\partial \theta}, \frac{1}{r \sin \theta} \frac{\partial \chi}{\partial r}, 0 \right), \quad \chi = -r \sin \theta \frac{\partial P}{\partial \theta};$$

both fields $\mathbf{A}_T$ and $\mathbf{A}_P$ here have zero helicity.

Using the representation (53.6) for the magnetic field $\mathbf{A} = \mathbf{H}$, we can rewrite

induction equation (3.6) as

$$\frac{\partial T}{\partial t} + \frac{1}{r} \mathbf{u}_h \cdot \nabla_h T - \frac{\nu_m}{r^2} \Delta_h T = \frac{1}{r^2} \left(\mathbf{u}_h \times \nabla_h \frac{\partial r P}{\partial r} \right)_r; \quad (53.7)$$

$$\begin{aligned} \frac{\partial \Delta_h P}{\partial t} + \operatorname{div} \left(\mathbf{u}_h \Delta_h P + u_r \nabla_h \frac{\partial r P}{\partial r} \right) - \\ - \nu_m \Delta \Delta_h P = (\nabla_h T \times \nabla_h u_r)_r. \end{aligned} \quad (53.8)$$

The right-hand side of the first of these equations describes the generation (or absorption) of a toroidal field by a poloidal field, which is possible because of the horizontal currents. The right-hand side of the second equation describes the generation (or absorption) of a poloidal field by a toroidal field, which is possible because of the vertical currents. If fields T and u_r have a common axis of symmetry (do not depend on the longitude), then this right-hand side goes to zero. Then field P will become damped in time due to the ohmic losses, while without excitation by field P , field T will also become damped, that is a dynamo turns out to be impossible. This is the theorem of the ‘antidynamo’, stated by Cowling (1934).

The same result is obtained for $u_r \equiv 0$, which yields the theorem of Bullard and Gellman (1954) stating that a dynamo with a purely toroidal velocity field is impossible. These investigators developed a method of solving the induction equation by expansion in spherical harmonics, and they constructed for $\nu_m = 0$ an approximate stationary model $\mathbf{H}(\mathbf{r})$ of spherical harmonics up to second order, for a specified $\mathbf{u}(\mathbf{r})$ in the form of a zonal current $\mathbf{u}_0(\mathbf{r})$ and several convective cells $\mathbf{u}_1(\mathbf{r})$. The typical ratio u_0/u_1 , playing the role of an eigenvalue, was found to be positive, as required for a dynamo to exist. For the Earth $H_T \sim 0.05$ tesla $\gg H_p$ (note, however, that the field $\mathbf{u}(\mathbf{r})$ specified in this work was later corrected by Gubbins, 1973).

A similar theorem of an ‘antidynamo’ for a plane two-dimensional current (for which $\mathbf{k} \cdot \mathbf{u} = 0$, where $\mathbf{k}$ is a fixed vector) was stated by Zel’dovich (1956) (who also considered the initial enhancement and subsequent attenuation of a magnetic field and the ‘diamagnetic’ ejection of magnetic force lines from an area made turbulent). Consequently, the theorems of an “antidynamo” indicate that motions capable of generating and sustaining a magnetic field must be three-dimensional and unsymmetrical.

Let us also write an expression for the ‘charge of the magnetic field’ (4.7) with the aid of relation (53.6):

$$r\rho H_* = (\nabla_h \eta \times \nabla_h T)_r + \frac{1}{r} \nabla_h \eta \cdot \nabla_h \frac{\partial r P}{\partial r} - \frac{\partial \eta}{\partial r} \Delta_h P \quad (53.9)$$

(formulas (53.7)–(53.9) were published by Gavrilin and Monin, 1972). Thus it is

evident that if the entropy field $\eta(\mathbf{r})$ differs little from the spherically symmetrical field $\eta_0(r)$, then the toroidal magnetic field can introduce into H_* a contribution which is not small compared to the contribution of the poloidal field, provided that it is stronger than the latter by a factor of at least $\eta_0|\eta - \eta_0|^{-1}$. According to the estimates of Bullard and Gellman, such conditions apparently exist in the Earth's core.

Let us consider, too, the *helicity* $I = \int \mathbf{A} \cdot \text{curl } \mathbf{A} \, dV$ of the vector field $\mathbf{A}$, the density of which $\mathbf{A} \cdot \text{curl } \mathbf{A}$ is the product of the polar vector times the axial vector, that is, a pseudoscalar which changes sign during reflections $\mathbf{r} \rightarrow -\mathbf{r}$ (so that for fields with a reflective symmetry the helicity density is identically equal to zero). Topologically, the helicity is a measure of the structural complexity (engagement or entanglement) of the vortex tubes of field $\mathbf{A}$. For example, if $\text{curl } \mathbf{A}$ is different from zero only inside two thin tubes with axes C_1 and C_2 , would around each other N times, then the helicity will be $\pm N\Phi_1\Phi_2$, where $\Phi_i = \int_{S_i} \mathbf{r} \cdot \text{curl } \mathbf{A} \, dS = \oint_{C_i} \mathbf{A} \cdot d\mathbf{r}$ are the vorticity fluxes through the sections of tubes S_i . If the Clebsch representation $\mathbf{A} = \psi \nabla \varphi + \nabla \chi$ exists (so that $\text{curl } \mathbf{A} = \nabla \psi \times \nabla \varphi$), then the helicity equals zero. It can be shown that, using the notation $\mathbf{H} = \text{curl } \mathbf{A}$, from the induction equation we obtain the relation

$$\frac{dI}{dt} = \nu_m \int (\mathbf{H} \cdot \Delta \mathbf{A} + \mathbf{A} \cdot \Delta \mathbf{H}) dV \quad (53.10)$$

(Moffatt, 1969), where the integration can be over any volume at the surface of which $\mathbf{n} \cdot \mathbf{H} = 0$. Therefore, because of diffusion the topological structure of the magnetic field varies with time, and it is impossible to track individual force lines. If, on the other hand, $\nu_m = 0$, then the helicity of the magnetic field does not vary in time. However, we will soon see that for a dynamo the helicity of the velocity field is more important.

Let us represent the velocity field and the magnetic field as $\mathbf{u}_0 + \mathbf{u}'$ and $\mathbf{H}_0 + \mathbf{H}'$, where a zero subscript signifies mean values (over the ensemble of realizations of random fields, or else over a large, for fluctuations, but small, for the mean fields, interval of space or time), and a prime indicates fluctuations, that is, deviations from the mean values. The averaged induction equation and the difference between the complete and averaged equations will be

$$\frac{\partial \mathbf{H}_0}{\partial t} - \text{curl}(\mathbf{u}_0 \times \mathbf{H}_0) - \nu_m \Delta \mathbf{H}_0 = \text{curl } \mathbf{F}; \quad \mathbf{F} = \langle \mathbf{u}' \times \mathbf{H}' \rangle; \quad (53.11)$$

$$\frac{\partial \mathbf{H}'}{\partial t} - \text{curl}(\mathbf{u}_0 \times \mathbf{H}') - \text{curl } \mathbf{G} - \nu_m \Delta \mathbf{H}' = \text{curl}(\mathbf{u}' \times \mathbf{H}_0);$$

$$\mathbf{G} = \mathbf{u}' \times \mathbf{H}' - \mathbf{F}. \quad (53.12)$$

Equation (53.11) is an analog of the Reynolds equations in turbulence theory, and its right-hand side $\text{curl } \mathbf{F}$, playing the role of a source for the mean field $\mathbf{H}_0$, is an analog of the Reynolds stresses. However, in contrast to turbulence theory, in the *electrodynamics of the mean field* the mean electromotive force $\mathbf{F}$ in a number

of cases can be calculated precisely (expressed in terms of the velocity field), since Equation (53.12) for $\mathbf{H}'$ is linear; after this, for $\mathbf{H}_0$ we can seek simple symmetrical solutions, because the 'antidynamo' theorems do not apply to Equation (53.11). This idea, dating back to Parker (1955), was further developed by the Potsdam group of Steenback, Krause, and Rädler (1966 and later works).

Because of the linear dependence of $\mathbf{H}'$ and thus of $\mathbf{F}$, on $\mathbf{H}_0$ we can represent field $\mathbf{F}$ as a rapidly converging series:

$$F_i = \alpha_{ij} H_{0j} + \beta_{ijk} \frac{\partial^2 H_{0j}}{\partial x_k} + \gamma_{ijkl} \frac{\partial^2 H_{0j}}{\partial x_k \partial x_l} + \dots, \quad (53.13)$$

where coefficients α_{ij} , β_{ijk} , γ_{ijkl} , ... are pseudotensors, completely determined by field $\mathbf{u}_0$, by the statistical properties of field $\mathbf{u}'$, and by the value of ν_m . The antisymmetrical part of pseudotensor α_{ij} has the form $(-\varepsilon_{ijk} a_k)$ and it makes a contribution $\mathbf{a} \times \mathbf{H}_0$ to field $\mathbf{F}$, that is, it leads in (53.11) to replacement of $\mathbf{u}_0$ by $\mathbf{u}_0 + \mathbf{a}$. If field $\mathbf{u}'$ is statistically uniform and *pseudoisotropic* (that is, invariant with respect to rotations, but not necessarily with respect to reflections, of the frame of reference), then $\alpha_{ij} = \alpha \delta_{ij}$, where α is a pseudoscalar, so that $\mathbf{u}'$ produces an electromotive force $\mathbf{F} = \alpha \mathbf{H}_0$ (and current $\sigma \mathbf{F}$, parallel to $\mathbf{H}_0$, is, in contrast to the usual induction current $\sigma \mathbf{u}_0 \times \mathbf{H}_0$, perpendicular to $\mathbf{H}_0$). Steenback and Krause (1966) called the appearance of this electromotive force the α effect (it had been discovered previously by Parker, 1955). In the pseudoisotropic case $\beta_{ijk} = \beta \varepsilon_{ijk}$, where β is simply a scalar. If it is constant, then its contribution to curl $\mathbf{F}$ will be $\beta \Delta \mathbf{H}_0$, meaning that it leads to a replacement of ν_m by $\nu_m + \beta$ (where usually β is positive, but the exotic possibility still exists that perhaps $\nu_m + \beta < 0$; according to Kraichnan, 1976, it can be realized as a result of fluctuations of the helicity of the field $\mathbf{u}'$).

For a specific calculation of coefficients α_{ij} and β_{ijk} (the effects related to $\mathbf{u}_0$ will be considered separately), let us confine ourselves to nondivergent statistically stationary and uniform fields $\mathbf{u}'$ with a spectral tensor $\Phi_{ij}(\mathbf{k}, \sigma)$ and accordingly with a helicity spectrum

$$F(k, \sigma) = i \int_{\Omega_k} \varepsilon_{lmn} k_m \Phi_{ln}(\mathbf{k}, \sigma) d\Omega \quad (53.14)$$

(where Ω_k is a sphere of radius k in $\mathbf{k}$ space). Then, in the 'quasilinear' approximation we can neglect in Equation (53.12) the curl $\mathbf{G}$ term (according to Vainshtein, 1972, this approximation corresponds to a Markov random process), so that this equation is chosen in the form $\partial \mathbf{H}' / \partial t - \nu_m \Delta \mathbf{H}' = (\mathbf{H}_0 \cdot \nabla) \mathbf{u}'$. Solving this by the spectral method and calculating $\mathbf{F}$, we then obtain

$$\begin{aligned} \alpha_{ij} &= i \nu_m \varepsilon_{ikl} \int k^2 k_j \Phi_{kl}(\mathbf{k}, \sigma) (\sigma^2 + \nu_m^2 k^4)^{-1} d\mathbf{k} d\sigma; \\ \alpha &= \frac{1}{3} \alpha_{ii} = -\frac{\nu_m}{3} \int k^2 F(k, \sigma) (\sigma^2 + \nu_m^2 k^4)^{-1} dk d\sigma. \end{aligned} \quad (53.15)$$

The first of these formulas is essentially a superposition of the values of $\alpha_{ij} = \alpha^{(3)} \delta_{i3} \delta_{j3}$ for individual 'helical waves' $\mathbf{u}' = \text{Re } \mathbf{u}_1 \exp[i(\mathbf{k} \cdot \mathbf{x} - \sigma t)]$, $\mathbf{u}_1 = u_1 (-i, 1, 0)$, $\mathbf{k} = (0, 0, k)$. The second, on the other hand, shows that α is a weighted integral of the helicity spectrum with opposite sign (obviously, too, in order to produce the α effect, diffusion of the magnetic field must take place, creating a phase shift between $\mathbf{H}'$ and $\mathbf{u}'$). Coefficient β_{ijk} is calculated similarly (for a field $\mathbf{H}_0$ which is linear in $\mathbf{x}$).

The first variant of the electrodynamics of a mean field (only with a specifics of the spherical geometry oriented toward the geomagnetic dynamo) was constructed somewhat earlier by Braginskii (1964). He used averaging over the longitude and sought an axisymmetric field $\mathbf{H}_0$, assuming $\mathbf{u} = \mathbf{u}_{0T} + \varepsilon^2 \mathbf{u}_{0P} + \varepsilon \mathbf{u}'$ and $\mathbf{H} = \mathbf{H}_{0T} + \varepsilon^2 \mathbf{H}_{0P} + \varepsilon \mathbf{H}'$, where $\varepsilon = \text{Re } m^{-1/2}$ is a small parameter. In this case it was found that the toroidal field $\mathbf{H}_{0T}$ is generated by the differential rotation $\mathbf{u}_{0T}$, while the poloidal field $\mathbf{H}_{0P}$ is generated by the α effect. Such a mechanism of generation of the average field is called an $\alpha\omega$ dynamo.

If, on the other hand, the α effect dominates in the equations for both $\mathbf{H}_{0T}$ and $\mathbf{H}_{0P}$, then the generation mechanism is called an α^2 dynamo. Let us give an example of such a dynamo with anisotropic α_{ij} . If we assume that $\mathbf{u}_0 = 0$ and $F_i = \alpha_{ij} H_{0j}$, then Equations (53.11) have solutions in the form of plane waves $\mathbf{H}_0 = \mathbf{H}_1 \exp(i\mathbf{k}_1 \cdot \mathbf{x} + pt)$, $\mathbf{k}_1 \cdot \mathbf{H}_1 = 0$, and for these solutions they have the form $(p + \nu_m k_1^2) H_{1i} = i\varepsilon_{ijk} \alpha_{km} k_{ij} H_{1m}$, where the determinant of the coefficients of H_{1i} must go to zero. This yields a cubic equation for p , the roots of which in the principal axes of pseudotensor α_{ij} have the form

$$p_1 = -\nu_m k_1^2; \quad p_{2,3} = -\nu_m k_1^2 \pm Q; \\ Q^2 = \alpha^{(2)} \alpha^{(3)} k_{11}^2 + \alpha^{(3)} \alpha^{(1)} k_{12}^2 + \alpha^{(1)} \alpha^{(2)} k_{13}^2, \quad (53.16)$$

An exponential increase of $\mathbf{H}_0$ is possible for $\text{Re } p > 0$, that is, for $Q^2 > \nu_m^2 k_1^4$. In the axisymmetric case ($\alpha^{(1)} = \alpha^{(2)}$), for $\alpha^{(1)} \alpha^{(3)} > 0$ and sufficiently small $\mathbf{k}_1$, intensification of field $\mathbf{H}_0$ takes place for any direction $\mathbf{k}_1$, but for $\alpha^{(1)} \alpha^{(3)} < 0$ it is possible only for small k_1 and for values of $\mathbf{k}_1$ in the cone $Q^2 > 0$. Such an anisotropic intensification is possible, in particular, for a spatially periodic field $\mathbf{u}'$. For example, in 'helical waves' $\mathbf{u}' = u_1 (\sin kz, \cos kz, 0)$ we obtain $\alpha_{ij} = \alpha \delta_{i3} \delta_{j3}$, $\alpha = -u_1^2 / \nu_m k_3$, but in a field $\mathbf{u}'$ invariant with respect to a group of rotations of a cube (Childress, 1970), for example, $\mathbf{u}' = u_1 (\sin kz + \cos ky, \sin kx + \cos kz, \sin ky + \cos kx)$, we nevertheless obtain isotropy $\alpha_{ij} = \alpha \delta_{ij}$, $\alpha = -u_1^2 / \nu_m k$.

Thus, helicity of the velocity field $\mathbf{u}'$ because of the α effect results in an intensification of the magnetic field $\mathbf{H}_0$. Vainshtein and Zel'dovich (1972) in the case of turbulence, and Moffatt (1972) for random waves, showed that, conversely, an increase in the magnetic field $\mathbf{H}_0$ due to the effect of the Lorentz force tends to weaken the α effect, thereby limiting the intensification of H_0 to some equilibrium value H_c .

Finally, let us consider reversals of the polarity of global magnetic fields. Such reversals are most evident on the Sun. Powerful magnetic fields of the order of 0.01 to 0.4 tesla are observed (according to the splitting of the spectral lines with their opposite angular polarization due to the Zeeman effect) in sunspots. The

magnetic flux through a typical spot is of the order of 10^{13} webers. Spots are usually detected in groups, of which more than 90% form bipolar magnetic regions, more than 8% form unipolar magnetic regions, and less than 1% form multipolar magnetic regions. In each bipolar magnetic region there is a leading (relative to the Sun's rotation) spot p and, a little further from the equator, one or several rear spots f of opposite magnetic polarity. All the spots p in one hemisphere and spots f in the other hemisphere have the same polarity (Hale and Nicholson, 1938).

Outside the spots the magnetic-field strength ranges from 10^{-4} tesla to 20×10^{-4} tesla. In the latitude zone $|\varphi| < 40^\circ$ bipolar magnetic areas predominate, and at each latitude regions of positive and negative polarity alternate, with a predominant latitude wave number $m = 6$; their axes are so inclined to the latitude circles that their western ends are closer to the equator than their eastern ends (these regions apparently correspond to Rossby waves, transporting angular momentum toward the equator). In regions of $|\varphi| > 40^\circ$ unipolar magnetic regions with the same polarity as the spots p prevail (which points to the presence of a poloidal field).

The heliomagnetic field cyclically reverses polarity, every 11.2 years on the average. During its cycle, the Wolf number (Wolf, 1848) $W = N_1 + 10N_2$, where N_1 is the number of spots and N_2 is the number of spot groups, varies from a minimum of 0 to 11 to a maximum of 49 to 190 and then drops to a new minimum. Figure 53.1 shows the fluctuations of W for the years 1610 to 1974 (Maunder, 1890, noted that practically no sunspots were observed from 1645 to 1715; Eddy, 1977, detected, according to ^{14}C fluctuations in annual tree rings during the last 5000 years, six such minima of the solar activity).

The cycle begins with the appearance of spots of new polarity in both hemispheres at the far edges of the zones of the 'royal latitudes' 52° to 5° (in most cases 30° to 8°); subsequent spots appear, as a rule, closer and closer to the equator (at the time of sunspot maximum near latitudes 15°); the last spots of this polarity appear at the near-equatorial edges of the 'royal latitudes' (Carrington, 1858, and Spörer, 1894, traced this regularity back to 1621). The isolines of the sunspot number have the shape of a butterfly on a latitude-time plot (Maunder, 1922); see Figure 53.2. They show that the first spots of the new polarity appear at high latitudes earlier (3 years earlier, on the average) than the last spots of the old polarity at low latitudes (so that the minima of the $Q(t)$ curve become smoothed out), but the 'butterflies' do not intersect.

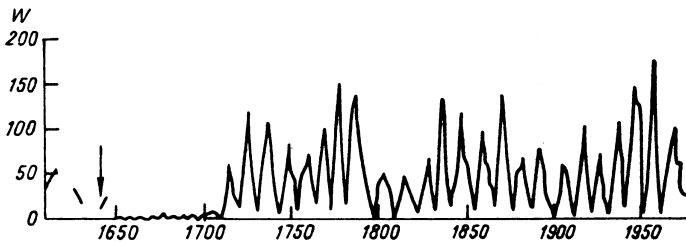


Fig. 53. 1. Oscillations of Wolf number.

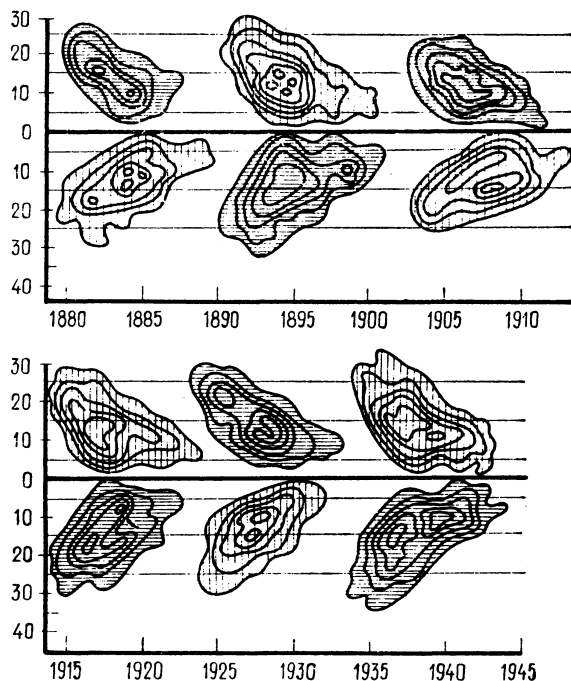


Fig. 53.2. Maunder butterflies for 1874 to 1913. Vertical shading denotes one polarity, and horizontal shading denotes opposite polarity.

Babcock (1961) considered the bipolar magnetic regions to occur because some portion of the force tube of the toroidal magnetic field (having a slight inclination to the latitude circle) rises up out of the subphotospheric convective layer. Buoyancy may be produced by a local intensification of the field (due to nonuniformity of the stretching of the tube by the differential rotation), which because of the tendency toward equalization of the total pressure $p + H^2/8\pi$ must lead to a reduction of p via the expulsion of some of the gas from the region of increased H , and thus to a reduction of ρ and the appearance of a buoyancy force (Parker, 1955). Partial suppression of convection by a strong magnetic field (hindering the motion of plasma across the magnetic force lines) leads to weakening of the convective heat transfer, cooling of the photospheric gas, and the appearance of relatively cool spots.

The heliomagnetic field is apparently generated by the mechanism of the $\alpha\omega$ dynamo: a toroidal field is produced from the poloidal by differential rotation of the Sun's convective zone. According to Babcock's model, three years before the beginning of the next cycle, there is produced on the Sun a poloidal magnetic field with a strength at the equator $H_0 \sim 5 \times 10^{-4}$ tesla and a total magnetic field of the order of 8×10^{13} webers, the force lines of which in the zone $|\varphi| \gtrsim 55^\circ$ pass along the meridians in the convective layer under the photosphere, whereas for $|\varphi| \lesssim 55^\circ$ they come out from under the photosphere and meet high above it.

The differential rotation $\omega \approx 14.28^\circ - 2.77^\circ \sin^2 \varphi \text{ day}^{-1}$ stretches the segments of the force lines in the convective layer, causing them in $(n + 3)$ years to shift in longitude by $\delta_\lambda = 17.6(n + 3) \sin^2 \varphi$ radians and to become spirals, making angles ψ to the meridian, where $\tan \psi = \partial \delta_\lambda / \partial \varphi$, and having an intensity $H = H_0 \sec \varphi \sec \psi$. After 3 years ($n = 0$) the field at latitudes $|\varphi| \sim 30^\circ$ reaches the critical value, on the average over the magnetized layer $H_c \sim 264 \times 10^{-4}$ tesla, while in some magnetic tubes it is several times higher, upon which the stretched parts of the tubes rise, forming spots, and the stretching stops. At lower latitudes the stretching continues longer. After $(n + 3)$ years the value H_c is reached at latitudes determined by the equation $(n + 3) \sin \varphi = 1.5$.

If the initial submerged poloidal field consisted, say, of 8 tubes about $2R_\odot$ long, then after 3 years they will be stretched out into spirals with 5 turns in each hemisphere, with a total length of about $500R_\odot$. After a cycle their total length amounts to $10^5 R_\odot$. If each segment of a stretched tube of length $R_\odot$ comprises 3 or 4 bipolar magnetic regions, then all of these together will form about 3000 regions after a cycle. This shows a good fit with observational data.

A poloidal field of new polarity is produced from the toroidal by α effect. According to Babcock, the loops of magnetic force lines produced over a bipolar magnetic region already make a contribution to the new poloidal field, since their p parts lie closer to the equator than their f parts. Rising into the corona, they are pinched off and carried away by the solar wind. Presumably, the ends left behind then experience crossclosure and the lower branch of the new poloidal field becomes submerged beneath the photosphere.

After we have defined in some way the function $\alpha(r, \theta)$ and the differential rotation $\omega(r, \theta)$ (or the poloidal field of the velocity $\mathbf{u}_{0p}(r, \theta)$), we can look for periodic solutions of the equations of electrodynamics of the mean field (53.11) and then construct the 'Maunder butterflies', as isolines of function $H_{0\lambda}(r, \theta, t)$ for some fixed r . Such calculations have been carried out by a number of investigators, including Steenback and Krause (1969), Roberts (1972), Ivanova and Ruzmaikin (1976, 1977) (in their 1977 model Ivanova and Ruzmaikin also took into account the nonlinear effect of the decrease in α with increasing H_0 , which apparently is in principle promising for the explanation of such long-period variations of solar activity as the above-mentioned 'Maunder minima' of 1645–1715). These calculations showed that for a reasonable choice of the parameters (including $\nu_m \approx 10^8 \text{ m}^2/\text{s}$) we can arrive at the correct length of the solar cycle. Similar calculations of $\alpha\omega$ dynamos for the giant planets should yield some important predictions.

Polarity reversals, albeit with much longer periods, are also characteristic of the geomagnetic field. At the Earth's surface and beyond it, this field is a potential field: $\mathbf{H} = -\nabla\Psi$, $\Psi = \sum_{n=1}^{\infty} \Psi_n$ ($n = 1$ corresponds to a dipole), where in spherical coordinates

$$\Psi_n = -r_1 \left(\frac{r_1}{r} \right)^{n+1} S_n(\theta, \lambda);$$

$$S_n = \sum_{m=0}^n (g_n^m \cos m\lambda + h_n^m \sin m\lambda) P_n^m(\cos \theta). \quad (53.17)$$

At the Earth's surface $r = r_1$ the dipole component is the main one. The north geomagnetic pole (toward which the northern end of a compass needle is attracted) lies near Greenland (73°N , 100°W), and the south geomagnetic pole lies in the Australian sector of Antarctica (68°S , 143°E); the values of H there are, respectively, 0.6×10^{-4} and 0.7×10^{-4} tesla. At the magnetic equator H varies from 0.42×10^{-4} tesla in Indonesia to 0.25×10^{-4} tesla on the east coast of South America. Table 53.1 gives the values of the most recent geomagnetic elements g_n^m and h_n^m in gamma units ($1 \gamma = 10^{-9}$ tesla) and the rates of their variation with time $\dot{h}_n^m$ and $\dot{g}_n^m$ in γ/year in the 1975 epoch. Conversion to the surface of the core $r = r_2$ shows that there $\mathbf{H}$ has a multipolar nature.

In the present epoch the dipole is weakening at a rate of about 5% per century. The nondipolar moment is drifting westward at a rate of $(0.18 \pm 0.015)^\circ$ longitude per year (Bullard *et al.*, 1950). This may be an indication that the rotation of the core lags behind that of the mantle by about 20 km/year ≈ 0.7 mm/s. Westin (1953) noted a correlation between the fluctuations in the rate of westward drift and the fluctuations in the velocity ω_1 of the Earth's rotation.

From the magnetization of lavas of different ages, Cox *et al.* (1963–1968) were able to detect reversals of the polarity of the geomagnetic field. During the last 4.5 million years there have been four *epochs* of alternating polarity sign (called, going backward from the present, the Brunhes, Matuyama, Gauss, and Gilbert epochs), interrupted by brief intervals of opposite polarity, known as *events* (Figure 53.3). On the basis of banded magnetic anomalies of alternating signs of the intensity of the order of 500 γ on the ocean floor, which are oriented parallel to the axes of the midoceanic ridges and are symmetrical about these axes, a calendar of polarity reversals was extended 162 million years into the past (Lamont scale); an epoch from 110 to 85 million years was detected in which there were no polarity reversals. On the basis of the magnetization of continental rocks, Khramov and Sholpo (1967) proposed a tentative scale of the polarity epochs for the entire Phanerozoic (including epochs without polarity reversals, for instance, 280 to 230 million years ago). The presence of a geomagnetic field has been verified back to at least 3 billion years ago.

TABLE 53.1
Most recent geomagnetic elements

n	m	g	h	$\dot{g}$	$\dot{h}$
1	0	-30,103.6	—	26.8	—
	1	-2016.5	5682.6	10.0	-10.1
2	0	-1906.7	—	-25.0	—
	1	3009.9	-2064.7	0.3	-2.8
3	2	1633.0	-58.1	5.5	-18.9
	0	1278.2	—	-3.8	—
	1	-2142.0	-329.8	-10.5	7.2
	2	1254.7	265.9	-4.7	2.8
	3	831.0	-227.0	-4.7	-6.4

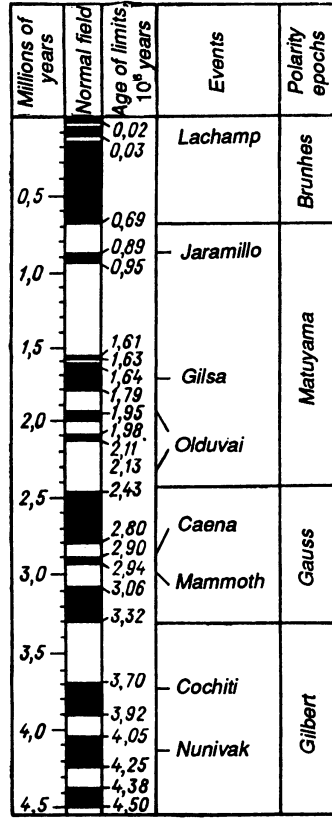


Fig. 53.3. Polarity reversals of geomagnetic field.

In order to calculate the geomagnetic field, it may well be that the same model of an $\alpha\omega$ dynamo is suitable, since the rotation of the liquid layer of the Earth's core may be differential. This is a result of the transfer of angular momentum in the meridional plane, say, by Rossby waves or, more precisely, by MBC waves. We should mention, too, the possibility of the formation of helical waves during flow around irregularities of the inner surface of the mantle; the existence of such waves was suggested when Hide and Meilin (1970, 1971) detected a correlation between the nondipole geomagnetic potential $\Psi - \Psi_1$ and the latitude-shifted gravitational anomalies. With regard to such calculations, it is important to know whether these waves are related to internal gravity waves or, conversely, to convective cells, that is, whether the stratification of the liquid layer is stable or unstable.

Stability may be brought about by the liberation of heat in the upper part of the liquid layer due to gravitational energy of the growth of the core (an increase in its radius r_2 ; see Section 54 below) or due to a difference in the rates of precession of the mantle and the core, these rates being proportional to their dynamic ellip-

ticities $\varepsilon = (C - A)/C$, where C and A are their principal moments of inertia (actually, $\varepsilon_c \approx (3/4) \varepsilon_m$; the precession hypothesis was analyzed by Malkus, 1963, 1968, and Loper, 1975). The presumption of Higgins and Kennedy (1971, 1973) of stability because the temperature in the core is close to the subadiabatically stratified melting point is hardly credible; it is more likely that the liquid layer is superheated. Instability may be produced by heat release from below during growth of the inner solid core.

Mirabel' *et al.* (1980) computed a model of the geomagnetic dynamo with generation of a toroidal field due to differential rotation and a poloidal field due to the radial component of the quasigeostrophic motions (with a horizontal scale L and a low Kibel' number $Ki = u_0/2\omega L$). The magnetohydrodynamic fields were taken to be:

$$\mathbf{u} = u_0[\mathbf{k} \times \nabla(\psi + Ki \psi_1) + Ki \nabla \sigma + \delta Ki w \mathbf{k}];$$

$$\mathbf{H} \approx H_0(\mathbf{k} \times \nabla \chi + \nabla \gamma + \delta H_1 \mathbf{k}); \quad p = p_0 \left(1 + \frac{g_0 \rho_0 \mathcal{D}}{p_0} Fr Ki p_1 \right); \quad (53.18)$$

$$\rho = \rho_0 \left(1 + \frac{r_1}{\mathcal{D} r} Fr Ki \rho_1 \right); \quad T = T_0 \left(1 + \frac{r_1}{\alpha T_0 \mathcal{D} r} Fr Ki T_1 \right),$$

where r_1 and $\mathcal{D}$ are the outer radius and thickness of the liquid layer; r and ∇ are the dimensionless radial coordinate and operator of the horizontal gradient, both measured by scale $\mathcal{D}$; $\mathbf{k}$ is the unit vector of the vertical; $\delta = \mathcal{D}/L$; $Fr = 4\omega^2 L^2/g_0$ $\mathcal{D}$ is the rotational Froude number; α is the coefficient of thermal expansion.

The ten unknown functions ψ , ψ_1 , σ , w , χ , γ , H_1 , p_1 , ρ_1 , and T_1 are related by the equations of motion (with the Lorentz force), continuity, induction, non-divergence, state, and heat flux. From these in the quasigeostrophic approximation we can derive equations for the potential vorticity, the vortical and radial components of the magnetic field, the zero H -divergence ($\partial H_1/\partial r + \Delta \gamma = 0$), and heat flux; the latter equations contain only functions ψ , w , χ , γ , and H_1 . We will present only the heat-flux equation:

$$\alpha T_0 \omega = \varepsilon \left\{ -\mu \frac{\partial}{\partial t} \frac{\partial \psi}{\partial r} + Pe^{-1} \left[\Delta \left(\mu \frac{\partial \psi}{\partial r} \right) + \delta^{-2} \mu \frac{\partial^3 \psi}{\partial r^3} \right] + Q \right\}, \quad (53.19)$$

where $\mu = \cos \theta$ (here $\varepsilon = L^2/L_R^2$ and $L_R = \mathcal{D}N/2\omega$ is the Rossby deformation radius), Pe is the Péclet number, and Q is the heat flux in units of $2 Ki^2 Fr \alpha T_0 \omega r^{-1}$. The quantities $\partial \psi/\partial r$, w , $\partial \chi/\partial r$, H_1 were assumed to go to zero at the boundaries of the liquid layer. The liquid layer was divided into four sublayers with boundaries from top to bottom 0, 1, 2, 3, 4, giving six equations in $\psi = \frac{1}{2}(\psi_1 + \psi_3)$, $\xi = \frac{1}{2}(\psi_3 - \psi_1)$, $\chi = \frac{1}{2}(\chi_1 + \chi_3)$; $\eta = \frac{1}{2}(\chi_3 - \chi_1)$; $\gamma = \gamma_3 = -\gamma_1$, and $w = w_2$ (the last quantity is eliminated with the aid of (53.19). These unknowns, and Q as

well, were expanded in series in spherical harmonics, where for Q the harmonics $P_0, P_1, P_2, P_1^1 \cos \lambda, P_1^1 \sin \lambda$ were chosen, for ψ and ξ the harmonics P_1, P_2 , and for χ, η, γ the harmonics $P_1, P_1^1 \cos \lambda, P_1^1 \sin \lambda$, 13 unknowns in all.

Assuming that at the initial time $\psi = \eta = \xi_1 = 0$, we were able to distinguish the least possible seven-component system. Then, we set $L^2 = L_R^2$ and $H_0^2 = 4\pi\rho_0 u_0^2$ and introduced the normalization:

$$y = \frac{24}{5Q_1} \xi_2; \quad z_1 = \frac{\sqrt{2}}{3} \text{Ki}^{1/2} \chi_1; \quad z_2 = \frac{\tau^{-2}}{6\sqrt{2}} Q_1 \text{Ki}^{1/2} \gamma_1;$$

$$(x_1, x_2) = \frac{1}{3} \text{Ki}^{1/2} (\chi_{11}^1 \mp \chi_{21}^1); \quad (x_{11}, x_{22}) = \frac{\tau^{-2}}{24} Q_1 \text{Ki}^{1/2} (\gamma_{11}^1 \mp \gamma_{21}^1); \quad (53.20)$$

where $\tau = r_1/\mathcal{D} - \frac{1}{2} \approx 1.235$. Using these variables, we wrote the seven-component system as:

$$\dot{y} = 3 - AC; \quad (\dot{z}_1, \dot{x}_1, \dot{x}_2) + \mu_1(z_1, x_1, x_2) = y(-z_2, x_{11}, x_{22});$$

$$(\dot{z}_2, \dot{x}_{11}, \dot{x}_{22}) + \mu_2(z_2, x_{11}, x_{22} - B\mu_2^{-1}z_1) = (-2Bx_1, Cx_1, -Cx_2);$$

$$A = \frac{128}{5} \tau^2 Q_1^{-2} \text{Ki}^{-1}; \quad B = \frac{1}{4\sqrt{2}} \tau^{-2} Q_1 \text{Ki} w_0;$$

$$C = x_1 x_{11} + x_2 x_{22} - z_1 z_2 + \frac{2}{A}; \quad w_0 = w_1^1 = w_1^{11};$$

$$\mu_1 = 2\tau^{-2} \text{Re}_m^{-1}; \quad \mu_2 = \mu_1 + 8\delta^{-2} \text{Re}_m^{-1}. \quad (53.21)$$

This system has two steady-state solutions with the same $y = \mu_1 \mu_2 A/3$, differing in sign by the values of z_1 and by the other variables proportional to z_1 . It was integrated numerically for $\text{Ki} = 10^{-4}$ and 10^{-6} , $\text{Re}_m = 10^2$ and 10^3 , $Q_1 = 1$ and 2 , respectively, $E = (15Q_1/16w_0)^2 = 0.5$ and initial conditions corresponding to one of the steady-state solutions. After an extended period of accumulation of small errors, the system entered an oscillatory regime with polarity reversals; for $\text{Ki} = 10^{-4}$ the reversals were almost regular with a cycle length of the order of 10^4 years, and for $\text{Ki} = 10^{-6}$ they were irregular with a cycle length of the order of 10^6 years and a nonaxisymmetric poloidal field (see Figure 53.4).

54. EVOLUTION OF PLANETARY INTERIORS

The evolution of the internal planetary structure is, of course, the fundamental part of the evolution of each planet. The internal structure here signifies the distribution (over the depth) of the chemical composition of the planetary material, its phase state (gaseous, liquid, or solid, and, if solid, the types of crystal lattice), and the physical characteristics: pressure, density, temperature, elasticity, viscosity, electrical conductivity, etc. Let us consider these distributions first using the Earth as an

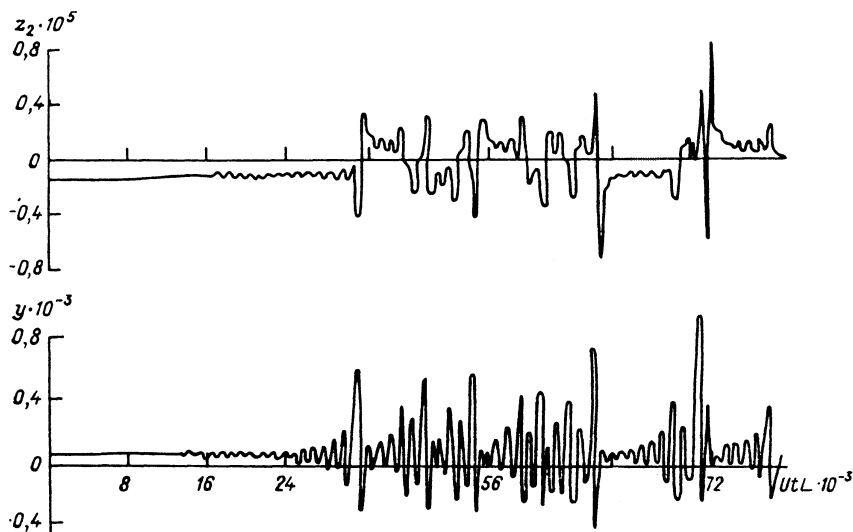


Fig. 53.4. Poloidal field $z_2(t)$ and differential rotation $y(t)$ in seven-component model of geomagnetic field for $K_i = 10^{-6}$.

example and then going on (with some necessary modifications of the results) to the other planets of the terrestrial group. Finally, we will discuss briefly the evolution of the giant planets.

The calculations of the internal structures of planets are based, first of all, on a representation of the assumed approximate *hydrostatic equilibrium* of their interiors, that is, with the simplification of spherical symmetry, on the equations

$$\begin{aligned} \frac{\partial p(r)}{\partial r} &= -g(r)\rho(r); & g(r) &= \frac{Gm(r)}{r^2}; \\ m(r) &= 4\pi \int_0^r \rho(r)r^2 dr, \end{aligned} \quad (54.1)$$

which relate the distributions of pressure $p(r)$ and density $\rho(r)$ (in (54.1) $g(r)$ is the distribution of the free-fall acceleration). However, the equation of state of the planetary material, that is, the dependence of ρ on p , should include at least one more thermodynamic parameter. This parameter may be the temperature, provided we assume, as is quite often likely, that in the planetary interior convective mixing takes place, at least layer-by-layer, and that distribution $T(r)$ is adiabatic (the adiabatic curves being calculated using the methods of theoretical physics).

The calculated model of the internal structure must provide correct values of the astronomically measured planetary radius R , planetary mass M , and moment of inertia $I = 8\pi/3 \int_0^R \rho r^4 dr$, as well as of the available gravitational moments of

higher orders (at these levels of accuracy it may be advisable to change from spherically symmetrical models to axisymmetric models).

In the case of the Earth, hypotheses concerning $T(r)$ are necessary only in order to ascertain the temperature itself, and Equations (54.1) can instead be supplemented by the relations

$$\frac{\partial p}{\partial \rho} = c_p^2 - \frac{4}{3} c_s^2;$$

$$c_p = \left(\frac{K}{\rho} + \frac{4}{3} \frac{\mu}{\rho} \right)^{1/2}; \quad c_s = \left(\frac{\mu}{\rho} \right)^{1/2}, \quad (54.2)$$

where $c_p(r)$ and $c_s(r)$ are the distributions of the velocities of longitudinal and transverse elastic waves, measured using seismic methods (K and μ are the moduli of compression and shear). The results of some such measurements (Bullen, 1966) have been plotted in Figure 54.1. These reveal that the Earth's interior is layered, the most clearly distinguishable layers being the thick solid upper envelope (the mantle) $B-C-D'-D''$ and the core $E-F-G$. At the surface D'/E separating these, which lies at a depth of 2920 km, the velocity c_p drops sharply (approximately from 13.6 to 8.1 km/s), while c_s decreases from 7.4 km/s to zero (the latter signifies that at least the upper layer of the core E is in a liquid, that is, molten, state; other geophysical evidence points to this as well).

From Figure 54.1 it is seen that below the transition layer F (4980–5120 km) the velocity c_p increases steadily (approximately from 10.4 to 11.4 km/s). Data on reflections of longitudinal waves from the surface of the inner core F'/G and on the passage of seismic waves from the Alaskan earthquake of 1964 showed that transverse waves also pass through the inner core G , albeit with quite low velocities $c_s \approx 3.6$ km/s. This means that the inner core is nevertheless solid (apparently, it is softened and close to molten).

Some investigators have used Equations (54.1), (54.2) to calculate models of $\rho(r)$ and $p(r)$. For example, according to the *Zemlya-2* model of Zharkov *et al.*

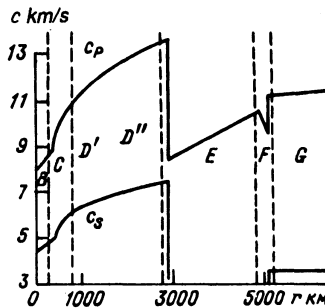


Fig. 54.1. Distributions $c_p(r)$ and $c_s(r)$ inside Earth, according to Bullen. B is upper mantle; C is middle mantle or Golitsyn layer; D' is lower mantle; D'' is transition layer; E is outer layer of core; F is transition layer; G is inner core.

(1971), at a depth of 2920 km for a pressure $p_0 = 1.386$ Mbar the density has a discontinuity $\rho_0^- - \rho_0^+ = 5.56-10.08$ g/cm³, while at the Earth's center $p_0 = 3.657$ Mbar and $\rho_c = 12.68$ g/cm³.

The difference between the mantle and the core is apparently a primarily chemical difference: roughly speaking, the mantle consists of silicates, while the core is iron (which at a fixed pressure has a higher density, but a lower melting point, than silicates). The hypothesis that the core consists of metallized silicates (Lodochnikov, 1939; Ramsay, 1949) is evidently not verified by experiments on the metallization of substances under *static* p , T conditions of the Earth's core or, certainly, by experiments on the *impact* compression of substances and by theoretical estimates (it is even doubtful that the depths of metallization for different silicates are the same).

It has been established that the Earth's crust consists of low-melting aluminosilicates. In the continental crust (mass 2.25×10^{22} kg) the mean composition is: 58% SiO₂, 15% Al₂O₃, 8% FeO + Fe₂O₃, 6% CaO, 4% MgO, 4% Na₂O, 2–2.5% K₂O, etc. In the oceanic crust (mass 0.6×10^{22} kg) the mean composition under a 0.5-km layer of loose sediments, in the 1.7-km 'second layer' of so-called tholeiitic basalts (pillowlike on the top and dike-like down below), is: 50.1% SiO₂, 16.7% Al₂O₃, 8.8% FeO + Fe₂O₃, 11.6% CaO, 7.8% MgO, 2.8% Na₂O, 0.19% K₂O, etc. The 5-km 'third layer' of the oceanic crust apparently consists of plutonic forms (gabbroids) of approximately the same makeup.

The *lherzolite* model of the subcrustal upper mantle was constructed on the basis of rock samples taken from the rift zones of the midoceanic ridges and from diamond-bearing kimberlite pipes: five parts high-melting magnesium silicate peridotite (harpurite) and one part tholeiitic basalt. The composition is: 45.7% SiO₂, 3.7% Al₂O₃, 5.1% Fe₂O₃ + 3.6% FeO, 2.3% CaO, 38.4% MgO, 0.3% Na₂O, 0.1% K₂O, etc. Assuming that the entire mantle is in a state of convective mixing, we can extend the *lherzolite* model to the mantle as a whole.

The hypothesis that the Earth's core is predominantly iron goes back to Wiechert (1897). According to experiments on the impact compression of substances, the core must contain at least 80–90% iron, although pure iron or an alloy of it with nickel would have a density and melting point which are too high. Therefore, in the core (at least in layer *E*) there must exist small amounts of lighter substances.

Taking into account that the most ancient rocks are considerably richer in iron, primarily in the form of FeO, but are poor in the light substances Si, Al, Mg, Ca, C, and S, Sorokhtin (1971) assumed that the main light addition to the iron in the outer core is the element oxygen, which is widely distributed throughout the crust and mantle. On the basis of data on the impact compression of Fe and FeO, he decided on the composition of these Fe₂O, which has a density suitable for the outer core and, apparently, at the corresponding pressures, is stable, being formed, in particular, during the disintegration of fayalite $\text{Fe}_2\text{SiO}_4 \rightarrow \text{Fe}_2\text{O} + \text{SiO}_2 + \text{O}$.

However, for the pressures of about 3 Mbar at the surface of the inner core, in iron a new electron-phase transition can take place, bringing it closer to the properties of nickel, with the reaction $\text{Fe}_2\text{O} \rightarrow 2\text{Fe} + \text{O}$. Then the inner core consists of Fe or an Fe–Ni alloy. Growth of the inner core may produce a source of energy for convection in the liquid layer, generating the geomagnetic field. In this section, though, assuming that the liquid layer constitutes about 30% of the

mass of the entire Earth, while the inner core constitutes only 1.2% (or 2.5% of the mass of the entire core), we will ignore the latter and assume that the 'core' material is just made up of Fe_2O .

Now we can add together the above-indicated compositions of the crust, mantle, and core, with weights equal to the relative masses of these envelopes, to obtain for the Earth as a whole the following composition: 31.2% SiO_2 , 2.6% Al_2O_3 , 23.6% FeO , 13.1% Fe , 1.6% CaO , 26.2% MgO , 0.2% Na_2O , 0.1% K_2O , etc. On the basis of cosmogonic considerations it can apparently be assumed that the primordial Earth was homogeneous, and that the rocks at its surface were homogeneous as well; they differed radically from the present-day rocks.

The main process of evolution of the planetary interiors (at any rate, for the terrestrial planets) is gravitational differentiation of substances of different weight, whereby heavier substances (that is, those with a higher density at a given depth) move downward, while lighter substances rise. As a consequence, the planet becomes stratified into envelopes, the deeper envelopes being made up of the heavier materials.

In the simplest model, the two-phase model, the planet is assumed to consist of two substances: a heavier 'core' substance with a mean concentration c_0 and a lighter 'mantle' substance with a mean concentration $c_1 = 1 - c_0$. The evolution of the planetary interior then consists simply in the separation and gradual growth of the core. The evolutionary stage is denoted by the fraction x of the core material which has become differentiated and passed into the core, which thus at this stage will have a mass $c_0 x M$, where M is the mass of the planet. At this stage of the evolution the mantle has a mass $(1 - c_0 x)M$ and it contains the entire mass $(1 - c_0)M$ of mantle material plus the mass $c_0(1 - x)M$ of the core material still remaining in it, so that the mean concentrations of these in the mantle are, respectively, $(1 - c_0)(1 - c_0 x)^{-1}$ and $c_0(1 - x)(1 - c_0 x)^{-1}$.

For instance, the lherzolite of the present-day mantle of the Earth contains 6.5% Fe , and thus $c_0(1 - x)(1 - c_0 x)^{-1} = 7.52\%$ core material Fe_2O . The relative mass of the present core is $c_0 x = 32.18\%$. This yields a concentration of core material in the Earth as a whole $c_0 = 37.3\%$, and the fraction of this material which has already passed into the core is $x = 86.3\%$. Consequently, the gravitational differentiation of core material in the Earth is already 86.3% completed.

Let us calculate approximately some details of this process. If $c(A)$ is mantle concentration of substance A and $c_0(A)$ is its initial concentration (in the primordial Earth), and if we assume that the core material Fe_2O can form in the mantle from free iron Fe , and from its oxides FeO and Fe_2O_3 , then taking into account their molecular weights we have

$$c(\text{Fe}_2\text{O}) = \frac{128}{112} c(\text{Fe}) + \frac{128}{144} c(\text{FeO}) + \frac{128}{160} c(\text{Fe}_2\text{O}_3). \quad (54.3)$$

In the primordial Earth there was not enough oxygen, and no ferric oxide existed, that is, then

$$c_0(\text{Fe}_2\text{O}) = \frac{128}{112} c_0(\text{Fe}) + \frac{128}{144} c_0(\text{FeO}). \quad (54.4)$$

Above we cited the values $c_0(\text{Fe}) \approx 13.1\%$ and $c_0(\text{FeO}) \approx 23.6\%$, for which $c_0 = c_0(\text{Fe}_2\text{O}) \approx 36.7\%$ (the value $c_0 \approx 37.3\%$ given previously was obtained taking into account the additional admixtures present in the inner core). During the first stage, as long as the mantle still contained free iron, the oxide Fe_2O was formed via the reduction of FeO according to the reaction $\text{FeO} + \text{Fe} \rightarrow \text{Fe}_2\text{O}$ (during the decomposition $2\text{FeO} \rightarrow \text{Fe}_2\text{O} + \text{O}$ oxygen most probably went to oxidize Fe to Fe_2O , which in the sum gave the previous reaction; the decomposition reaction $4\text{FeO} \rightarrow \text{Fe}_2\text{O} + \text{Fe}_2\text{O}_3$ in the presence of free Fe is much less likely, and it will be neglected here). The budgets for the mantle materials in this reaction are:

$$\begin{aligned} c_0 - c(\text{Fe}_2\text{O})(1 - c_0x) &= \frac{128}{72} [c_0(\text{FeO}) - c(\text{FeO})(1 - c_0x)] = \\ &= \frac{128}{56} [c_0(\text{Fe}) - c(\text{Fe})(1 - c_0x)]. \end{aligned} \quad (54.5)$$

From this we get

$$\begin{aligned} c(\text{Fe}) &= (1 - c_0x)^{-1} \left[c_0(\text{Fe}) - \frac{56}{128} c_0x \right]; \\ c(\text{FeO}) &= (1 - c_0x)^{-1} \left[c_0(\text{FeO}) - \frac{72}{128} c_0x \right]. \end{aligned} \quad (54.6)$$

According to the first of these formulas, $c(\text{Fe})$ goes to zero, that is, free iron disappears from the mantle, when $x = x_1 = (128/56)c_0(\text{Fe})/c_0 \approx 0.816$ (on the chronological scale to be given below, this corresponds to the beginning of the Phanerozoic, about 600 million years ago). The concentrations for $x = x_1$ will be denoted by a subscript 1, so that $c_1 = c_1(\text{Fe}_2\text{O}) \approx 9.8\%$ and $c_1(\text{FeO}) \approx 11\%$. In the second stage, when $x > x_1$, core material is produced via the reaction $4\text{FeO} \rightarrow \text{Fe}_2\text{O} + \text{Fe}_2\text{O}_3$ (the reactions $2\text{FeO} \rightarrow \text{Fe}_2\text{O} + \text{O}$ and $\text{Fe}_2\text{O}_3 \rightarrow \text{Fe}_2\text{O} + \text{O}_2$ lead to this reaction, since the oxygen will most probably go to oxidize FeO to Fe_2O_3). This gives the following mass budgets:

$$\begin{aligned} c_1(1 - c_0x_1) - c(\text{Fe}_2\text{O})(1 - c_0x) &= \frac{128}{160} c(\text{Fe}_2\text{O}_3)(1 - c_0x) = \\ &= \frac{128}{288} [c_1(\text{FeO})(1 - c_0x_1) - c(\text{FeO})(1 - c_0x)]. \end{aligned} \quad (54.7)$$

Then we have

$$\begin{aligned} c(\text{FeO}) &= (1 - c_0x)^{-1} \left[c_1(\text{FeO})(1 - c_0x_1) - \frac{288}{128} c_0(x - x_1) \right]; \\ c(\text{Fe}_2\text{O}_3) &= \frac{160}{128} c_0(x - x_1)(1 - c_0x)^{-1}. \end{aligned} \quad (54.8)$$

According to the first of these formulas, $c(\text{FeO})$ goes to zero, that is, ferrous iron disappears from the mantle, when $x = x_2 = (128/288)c_1(\text{FeO}) (1 - c_0x_1)/c_0 \approx 0.908$ (about one billion years in the future). The concentrations for $x = x_2$ will be denoted by a subscript 2, so that $c_1 = c_2(\text{Fe}_2\text{O}) \approx 5\%$ and $c_2(\text{Fe}_2\text{O}_3) \approx 6.4\%$. In the third stage, when $x > x_2$, core material is produced via the reaction $\text{Fe}_2\text{O}_3 \rightarrow \text{Fe}_2\text{O} + \text{O}_2$ with a mass budget

$$c_2(1 - c_0x_2) - c(\text{Fe}_2\text{O})(1 - c_0x) = \frac{128}{160} [c_2(\text{Fe}_2\text{O}_3)(1 - c_0x_2) - c(\text{Fe}_2\text{O}_3)(1 - c_0x)], \quad (54.9)$$

which yields

$$c(\text{Fe}_2\text{O}_3) = (1 - c_0x)^{-1} \left[c_2(\text{Fe}_2\text{O}_3)(1 - c_0x_2) - \frac{160}{128} c_0(x - x_2) \right]. \quad (54.10)$$

It follows from (54.3) that this expression must be equal to $(160/128)c(\text{Fe}_2\text{O})$, where $c(\text{Fe}_2\text{O}) = c_0 (1 - x) (1 - c_0x)^{-1}$. Taking the second formula of (54.8) into account, we then obtain $2x_2 = 1 + x_1$.

The mantle concentrations of relatively immobile substances, which do not leave the mantle to go into the core or the upper envelopes of the Earth, or which are so widespread in the mantle that their relative losses remain low (e.g., SiO_2 , MgO , and MnO), increase in time according to the law

$$c(A) = c_0(A)(1 - c_0x)^{-1}. \quad (54.11)$$

In particular, as demonstrated above, the concentration of silica SiO_2 increased from 31.2% in the primordial Earth to 45.7% in the present-day mantle lherzolite (which naturally brought about qualitative changes in the mantle rocks). The mantle concentrations of substances leaving the mantle for the upper envelopes must decrease with time more rapidly than according to law (54.11). To calculate these, it is convenient to characterize the stage of planetary evolution not by x , but rather by the fraction $y = nM_m/n_1M_{m1}$ of completed convective mass exchange in the mantle, where $M_m = (1 - c_0x)M$ is the mass of the mantle, n is the fraction of this mass which has already become differentiated, and subscript 1 corresponds to the value $x = 1$ (Sorokhtin, 1972); the quantity y can be assumed to be proportional to the mass of the *continental* crust which has already formed. The dimensionless rate of convective mass exchange dy/dt can be called the *tectonic activity* of the planet. The rate of convective mass exchange $d(nM_m)/dt$ can be assumed to be proportional to the rate of increase of core mass and inversely proportional to the mass of the mantle, that is, equal to $(M/M_m) d(c_0xM)/dt$. Thus

for the conditions $y = 0$ ($x = x_0$) and $y = 1$ ($x = 1$) we have

$$y = \left(\ln \frac{1 - c_0 x_0}{1 - c_0} \right)^{-1} \ln \frac{1 - c_0 x_0}{1 - c_0 x}. \quad (54.12)$$

For $x_0 = 0$, a value of $y = 0.833$ corresponds to the present-day $x = 0.863$. Now, if we write the rate of decrease of the mantle mass of movable material A in the form $\dot{M}(A) = -\kappa(A)M(A)\dot{y}$, where $\kappa(A)$ is its *mobility*, then instead of (54.11) we get

$$c(A) = c_0(A) \frac{1 - c_0 x_0}{1 - c_0 x} e^{-\kappa(A)y}. \quad (54.13)$$

Monin and Sorokhtin calculated the variations with time of the mantle concentrations of a number of substances; these are shown in Figure 54.2, on a chronological scale to be discussed below. As an important example, we consider the content in the Earth of the long-lived radioisotopes ^{40}K , ^{238}U , ^{235}U , and ^{232}Th , assuming these to be *lithophilic*, that is, having an affinity to silicates, so that they are absent in the core. At present ^{40}K constitutes 0.0119% of the natural potassium; 88% of this ^{40}K becomes ^{40}Ca via beta decay, with a constant decay rate $\lambda_\beta \approx 0.472 (10^9 \text{ years})^{-1}$, while the other 12% of the ^{40}K becomes ^{40}Ar via K capture, with a constant rate $\lambda_\kappa = 0.0585 (10^9 \text{ years})^{-1}$, and ^{40}Ar accounts for 99.6% of the natural Ar. The constant decay rates of the other three isotopes are: $\lambda_{238} = 0.15369$, $\lambda_{235} = 0.97216$, and $\lambda_{232} = 0.048813 (10^9 \text{ years})^{-1}$.

Now there are 6.525×10^{16} kg of ^{40}Ar in the atmosphere and 0.084×10^{16} kg in the hydrosphere. According to Ronov and Yaroshevskii (1967), in the continental crust (mass of 2.25×10^{22} kg) 2% K is present, with $\text{K/U} = 1.3 \times 10^4$, $\text{K/Th} = 3.3 \times 10^3$, and $^{238}\text{U}/^{235}\text{U} = 137.7$. This gives in the continental crust 5.49×10^{16} kg ^{40}K (plus 0.0066×10^{16} kg in seawater), 3.46×10^{16} U, and 1.36×10^{17} kg Th. Let us use the equation of evolution of ^{40}K in the continental crust and degassing of ^{40}Ar from the mantle and crust:

$$\begin{aligned} {}^{40}\text{K}_{\text{cr}} &= {}^{40}\text{K}_0 e^{-\lambda_\beta t - \lambda_\kappa t} (1 - e^{-\kappa y}); \\ {}^{40}\text{Ar}_{\text{atm}} &= \frac{{}^{40}\text{K}_0 \lambda_\kappa}{\lambda_\kappa + \lambda_\beta} (1 - e^{-\lambda_\beta t - \lambda_\kappa t}) (1 - e^{-\kappa y}) (1 + e^{-\kappa y} - e^{-\kappa_1 y}), \end{aligned} \quad (54.14)$$

where $\kappa_1 \approx 0.976$ is the mobility of the Ar degassed from the crust as its bedrock is transformed into sedimentary rock, found from the condition $dm_{\text{cr}} = -k_1 m_{\text{cr}} dy$ for the present-day value $\dot{y} \approx 0.25 (10^9 \text{ years})^{-1}$. For $t = 4.65 \times 10^9$ years and $y = 0.833$, from (54.14) we get $\kappa(\text{K}) = 0.8$, an initial amount of ^{40}K in the Earth ${}^{40}\text{K}_0 = 1.315 \times 10^{18}$ kg, and a present-day amount of 1.131×10^{17} kg.

On the basis of the isotope ratios in lunar basalts and recent basalts, Sorokhtin obtained the values $\kappa(\text{U}) = 2.17$ and $\kappa(\text{Th}) = 2.85$. These yielded the following

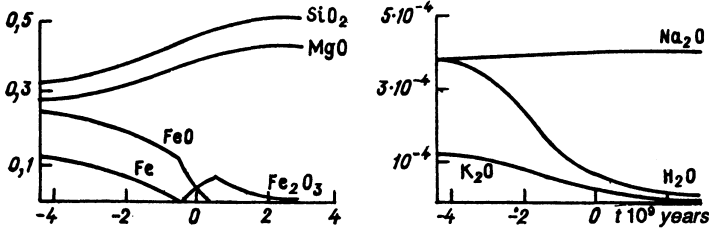


Fig. 54.2. Variations with time of mantle concentrations of various substances.

present-day amounts in the Earth: $^{238}\text{U} = 4.11 \times 10^{16}$ kg, $^{235}\text{U} = 3 \times 10^{14}$ kg, and $^{232}\text{Th} = 1.5 \times 10^{17}$ kg, as contrasted with their initial values: $^{238}\text{U}_0 = 8.4 \times 10^{16}$ kg, $^{235}\text{U}_0 = 2.76 \times 10^{16}$ kg, and $^{232}\text{Th}_0 = 1.88 \times 10^{17}$ kg. Assuming rates of heat liberation 0.94×10^{-4} W/kg for ^{238}U , 5.7×10^{-4} W/kg for ^{235}U , 0.26×10^{-4} for ^{232}Th , and 0.29×10^{-4} W/kg for ^{40}K , we arrive at a present-day radiogenic heat liberation in the Earth of 1.13×10^{13} W and an initial value (not taking short-lived isotopes into account) of 6.56×10^{13} W, that is, six times greater than the present value. If the specific heat content of the matter is 1.3×10^3 J/(kg K), then the gradual heating of the Earth is of the order of $0.3^\circ\text{C}/10^6$ years. Integration over time indicates that, during the course of the entire existence of the Earth, long-lived radioisotopes have released into it 0.41×10^{31} joules of heat.

As a second important example, we cite the calculation of the growth of the Earth's hydrosphere by Monin and Sorokhtin (1982). These investigators deduced that in the first stage, during the Early Katararchean, 4.6 to 3.8 billion years ago, there was as yet very little free water on the Earth's surface. The mass m_w of water in the oceanic crust increased in accordance with the decrease in the water mass M_w of the mantle, which took place at a rate $\dot{M}_w = -\kappa_w M_w \dot{y}$. Therefore,

$$m_w = M_{w0}(1 - e^{-\kappa y}) = 23.6(1 - e^{-4.05y}), \quad (54.15)$$

where M_{w0} is the total mass of water in the Earth (the numerical values will be explained below). During the second stage, in the Late Katararchean and Archean, 3.8 to 2.6 billion years ago, the oceans had already appeared but the water level had not yet reached the tops of the midoceanic ridges, and only the first (sedimentary) and second (basaltic) layers of the oceanic crust were water-saturated, the third layer still being anhydrous. If we assume that the portion s of the Earth's surface occupied by the continental crust increases according to the law $s = s_2 y/y_2$, where the subscript 2 denotes values 2.6 billion years ago, then at this stage

$$m_w = \frac{m_{w2}}{1 - s_2} \left(1 - S_2 \frac{y}{y_2} \right) = 1.35 - 0.54y. \quad (54.16)$$

If the third stage, during the first half of the Early Proterozoic, 2.6 to 2.2 billion years ago, the third layer of the oceanic crust also gradually became saturated with water, and law (54.15) again was valid. During the fourth stage, which covers all

the subsequent time, the hydrosphere has again grown according to a law analogous to (54.16), but with different parameters:

$$m_w = \frac{m_{w1}}{1 - s_1} \left(1 - s_1 \frac{y}{y_1} \right) = 5 - 2y, \quad (54.17)$$

where the subscript 1 denotes the values at present. Everywhere, measuring m_w in units of 10^{20} kg, we assume here that $m_{w1} = 3.34$ and $s = \frac{1}{3}$ (and, as above, $y_1 = 0.833$). On the basis of the present area of the oceanic crust, a water content of 0.9 in its first and second layers, and the value of y_2 (see below), we arrive at the coefficients in (54.16). Then, knowing m_{w1} , y_1 , m_{w2} , and y_2 , we can find the coefficients in (54.15). The rate of liberation of water during the formation of the continental crust can be computed as $\dot{z}_w = m_w \dot{y} / m_{w1} \dot{y}_1$. Integration over time then yields the water mass in the continental crust $m_{wc} = m_{wc1} z_w / z_{w1} \approx m_{wc1} y / y_1$, where according to present estimates $m_{wc1} = 4.86$. Now, finally, we can calculate the mass of the growing hydrosphere:

$$M_{\text{hydro}} = M_{w0} (1 - e^{-ky}) - m_w - m_{wc}. \quad (54.18)$$

The results of the calculations are shown in Figure 54.3.

Consequently, we see that the two-phase model enabled a very simple calculation of a whole set of important characteristics of the Earth's evolution. Let us examine, within the framework of this model, two more problems: the calculation of the evolution of distributions $p(r)$ and $\rho(r)$, and the establishment of the relationship between the evolution parameter x and the time t . In order to solve the first of these, we must know the equations of state $\rho_0(p, T)$ of the core material and $\rho_1(p, T)$ of the mantle material. Having in mind the planets of the terrestrial group, whose densities vary over a relatively small interval from 2 to 15 g/cm³, their temperature dependences being quite weak, Monin approximated the

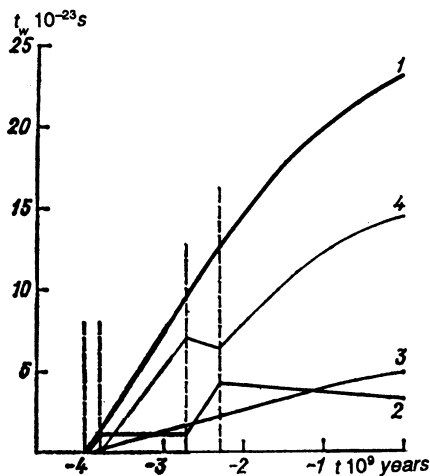


Fig. 54.3. Changes in time of water masses: degassed from mantle (1), in oceanic crust (2), in continental crust (3), and in hydrosphere (4).

equations of state as parabolas:

$$\rho_0 = \alpha_0 \rho_1(p); \quad \rho_1 = \left(\rho_s^2 + \frac{p}{2\pi GL} \right)^{1/2}, \quad (54.19)$$

where the parameters $\alpha_0 > 1$, ρ_s , and L , and also the two coefficients for 'joining' the solutions of Equations (54.1) at the surface of the core (so as to ensure continuity of p and g there), have to be selected according to the present mass of the Earth $M = 5.98 \times 10^{24}$ kg, radius $R = 6371$ km, dimensionless moment of inertia $I_1 = I/MR^2 = 0.3308$, core radius $R_0 = 3451$ km, and relative core mass $c_0 x = 0.3218$. The equation of state of the mantle material here has the form $\rho = \alpha_0 \beta \rho_1(p)$, where $\beta = (1 - c_0 x)[c_0(1 - x) + \alpha_0(1 - c_0)]^{-1}$, and Equations (54.1) are solved analytically: in the core $\rho_0 = A\xi^{-1}\sin \xi$, while in the mantle $\rho_1 = \xi^{-1}(B \sin \xi + C \cos \xi)$, where $\xi = \alpha_0 r/L$, and coefficients A , B , and C are determined from the conditions of joining and the requirement that $\rho_1 = \rho_s$ at $r = R$. The calculation gave the following values for the parameters of model (54.19): $\alpha_0 = 1.68$; $\rho_s = 3.42 \times 10^3$ kg/m³; $L = 4017$ km.

The structure of the Earth calculated for these parameters at different stages x of its gravitational differentiation is characterized by the following data. The Earth's radius R during a time from $x = 0$ to $x = 1$ almost does not vary (the calculated decrease of 25 km probably lies within the accuracy limits of the model). The radius of the core R_0 increases slightly more rapidly than $x^{1/3}$ (reaching a limit of 3635 km), so that the mean density of the core decreases slightly (from 11.65×10^3 kg/m³ for $x = 0.2$ to 11.09×10^3 kg/m³ for $x = 1$). The density at the Earth's center ρ_c increases from 11.34 to the present-day 13.86 and then to the limit value 14.08×10^3 kg/m³, while the pressure p_c goes from 2.30 to 3.82, and then 3.97 Mbar. With the growth of the core, there is a corresponding decrease in the amount of (core) material above it, so that there is a drop in the pressure p_0 at the boundary of the core (from 2.04 Mbar for $x = 0.2$ to 1.25 Mbar for $x = 1$) and also in the corresponding densities ρ_0^+ (7.43×10^3 to 5.5×10^3 kg/m³) and ρ_0^- (10.86×10^3 to 9.25×10^3 kg/m³), but their discontinuity $\rho_0^- - \rho_0^+$ increases. The rise in the concentration of mass toward the center of gravity reduces both the absolute moment of inertia I (from 9.00×10^{37} kg m² for $x = 0$ to 7.88×10^{37} kg m² for $x = 1$) and the relative moment I_1 (from 0.3681 to 0.3250). The potential energy of the Earth in the field of Newtonian attraction of all its masses is computed as follows:

$$\begin{aligned} U &= -G \int_0^M \frac{m(r) dm(r)}{r} = - \int_0^M g(r) r dm(r) \\ &= -4\pi \int_0^R g(r) \rho(r) r^3 dr = 4\pi \int_0^{p_c} r^3 dp(r) = \\ &\quad -12\pi \int_0^R p(r) r^2 dr. \end{aligned} \quad (54.20)$$

Its decrease $E(x) = U(0) - U(x)$ is a powerful source of heat release inside the Earth. A calculation shows that it increases from 0.32×10^{31} J for $x = 0.2$ to 1.61×10^{31} J at present, with a future limit of 1.86×10^{31} J (note that the total energy release throughout the entire past history of the Earth is four times greater than the heat release of long-lived radioisotopes). The spatial density of the influx of this energy (accurate to terms of the nature of the three-dimensional divergence) is $e(r, t) = 3[p(r, t) - p(r, 0)]$, so that the rate of energy release (the conversion of gravitational energy into internal energy, and the division of the latter into thermal and elastic energy, will not be considered here) per unit volume is $\dot{e} = 3(\partial p / \partial x)\dot{x}$. This rate is positive in the core and a maximum at its center; in the mantle, on the other hand, which is expanding slightly, potential energy is expended. A computation of the rate of energy release in spherical layers of unit thickness $4\pi r^2 \dot{e}$ showed that most of the energy release takes place in the core near its upper boundary.

In order to complete our estimates of the components of the Earth's heat budget, it now remains just to calculate the rate ε_T of heat release in the Earth due to the dissipation of tidal energy. To do this, we will take into account only lunar tides (since the moment of the tide-producing forces of the Sun is many times less than that of the Moon: $(m_1 r_s^3 / m_s r_1^3)^2$ times less, where subscript 1 pertains to the Moon and subscript s to the Sun; now this ratio is about 5). In addition, such a calculation is simplified greatly by the fact that the approximation of a circular equatorial orbit of the Moon is suitable for it (Ruskol, 1975). Then the work of the moment of the tide-producing forces of the Moon per unit time ε_T is given by the formula

$$\varepsilon_T = \frac{3}{2} \frac{k_2 G m_1^2 R^5}{Q r_1^6} (\omega_0 - \omega_1), \quad (54.21)$$

where R is the radius of the Earth, $m_1 = 7.35 \times 10^{22}$ kg is the mass of the moon, r_1 is the distance from the Moon, ω_0 and ω_1 are the angular velocities of the Earth's rotation and the Moon's orbital motion, $k_2 \approx 0.29$ is the so-called Love number, and $Q \approx (2\delta)^{-1}$ is the *quality factor* of the Earth (δ is the lag angle of the tides). The coefficient of $\omega_0 - \omega_1$ on the right-hand side is here equal to dM_1/dt , where $M_1 = m_0 m_1 (m_0 + m_1)^{-1} r_1^2 \omega_1$ is the angular momentum of the Moon's orbital motion. The quantities ω_0 , ω_1 , and r_1 are related to one another by the law of conservation of angular momentum and by the third law of Kepler:

$$C\omega_0 + M_1 = \text{const}; \quad C = I + \frac{2k_s R^5}{9G} \omega_0^2;$$

$$\omega^2 r_1^3 = \mu = \text{const} \approx 4.036 \times 10^{14} \text{ m}^3/\text{s}^2 \quad (54.22)$$

where $k_s \approx 0.947$ is the so-called secular Love number. From these and from the above-mentioned expression for dM_1/dt in the time intervals during which the

values of Q can be assumed quasiconstant, we obtain

$$r_1^{1.3/2}(t_2) - r_1^{1.3/2}(t_1) = \frac{39}{2} \frac{k_2 m_1 \mu^{1/2} R^5}{m_0 Q} (t_2 - t_1). \quad (54.23)$$

Monin *et al.* (1986) divided the history of the Earth into five such intervals: (1) the Early Katararchean 4.65 to 4.0 billion years ago, until the Earth heated up and when there were not yet any oceans, so that $Q \sim 1500$, as for the present-day lithosphere for $r_1(4.6) = 2 \times 10^7$ m (near the Roche limit, at which tidal forces break up the satellite), from (54.23) $r_1(4.0) = 1.576 \times 10^8$ m; (2) the Late Katararchean and Early Archean 4.0 to 3.2 billion years ago, when due to the heat release from the primordial core the Earth's interior warmed up and Q dropped sharply; (3) the Middle to Late Archean 3.2 to 2.6 billion years ago, when the Earth cooled somewhat and Q increased; (4) the Proterozoic 2.6 to 0.6 billion years ago, when Q still increased a little; (5) the Phanerozoic 0.6 to 0 billion years ago, when due to an abundance of shallow seas Q decreased significantly.

For interval (5) from the diurnal microlamination of corals of the Middle Devonian, 0.38 billion years old, it was established that there were 400 days in a year, that is, $\omega_0(0.38) = 7.964 \times 10^{-5} \text{ s}^{-1}$, so that from (54.22) we get $r_1(0.38) = 3.844 \times 10^8$ m and from (54.23) $Q \approx 11.8$ (now, however, is a geocratic epoch and $Q \approx 13$). Consequently, we have $r_1(0.6) = 3.569 \times 10^8$ m. For interval (4) from the lamination of stromatolites 1.8 billion years old it was established that there were 445 days in a year, that is, $\omega_0(1.8) = 8.86 \times 10^{-5} \text{ s}^{-1}$, so that from (54.22) we have $r_1(1.8) = 3.451 \times 10^8$ m, and taking into account $r_1(0.6)$ from (54.23) we get $Q \approx 75$ and $r_1(2.6) = 3.361 \times 10^8$ m. On the basis of $r_1(2.6)$ and $r_1(4)$ for the interval (3) + (4) from (54.23) we obtain on the average $Q = 25.6$; then, dividing Q^{-1} over intervals (3) and (4) in proportion to their lengths, we obtain for (4) $Q = 17.8$ and for (3) $Q = 61$.

The plot of ε_T calculated on the basis of these values of Q is shown in Figure 54.4, curve 1. It shows that at the initial time $\varepsilon_T \approx 1.3 \times 10^{12}$ MW, which is 30,000 times higher than the present rate of total heat loss by the Earth, equal to approximately 4.5×10^7 MW. However, in 10 million years ε_T dropped by a factor of 1800, and by a time 4.0 billion years ago it had decreased to 1.2×10^7 MW. At this limit because of the drop in Q , due to the overheating during the heat release of the primordial core, a new peak $\varepsilon_T \approx 1.04 \times 10^9$ MW appeared, but here, too, ε_T dropped rapidly. The present value $\varepsilon_T = 2.76 \times 10^6$ MW is very low. Curves 2 and 3 in Figure 54.4 show the rates of radiogenic and gravitational energy release (evidently, the latter is the main mechanism).

The calculation of the evolution of the Earth's interior based on the two-phase model can be extended to the other terrestrial planets if we assume that they all consist of the same mantle material and core material (having the same equations of state (54.19) and parameters α_0 , ρ_s , and L), except in different amounts, and thus in a different ratio to c_0 . To apply the model, for each planet of specified radius R and mass M , we must use the formulas of the two-phase model to

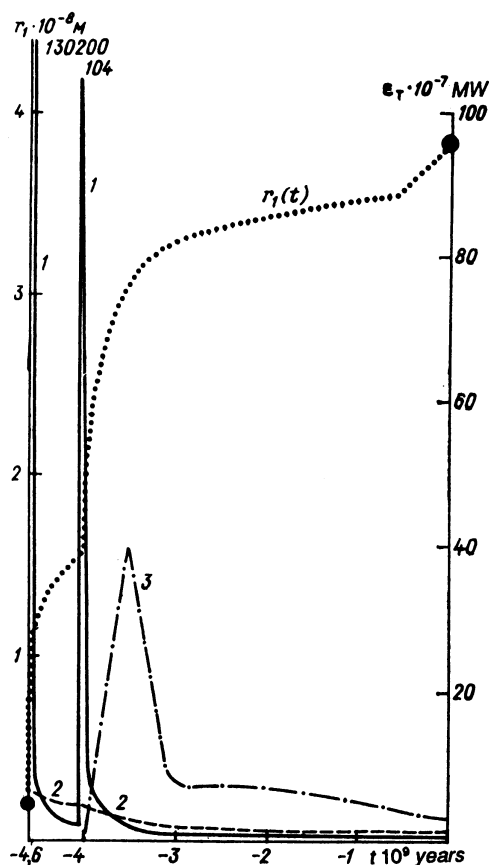


Fig. 54.4. Rates of energy release ε_T in Earth: (1) tidal, (2) radiogenic, (3) gravitational; r_1 is distance from moon.

determine the concentration c_0 and the present-day value x^* of the evolution parameter.

This can be done successfully for Mercury, Venus, and Mars, but not for the Moon, whose mean density $3.34 \times 10^3 \text{ kg/m}^3$ is less than the minimum density in our model $\rho_s = 3.42 \times 10^3 \text{ kg/m}^3$ (probably because the Moon has a very large crust, so that a third material, a light 'crustal' material, with a density, say, $\alpha_2 \rho_1(p)$, where $\alpha_2 < 1$, has to be introduced into the model). Therefore, instead of the Moon, let us calculate the evolution of some hypothetical Quasimoon, with the Moon's mass and with an iron concentration of 13%, obtained for the Moon by estimating its mean magnetic susceptibility from data on the Moon's perturbation of the Earth's magnetic wake, and let us find its radius (the radius of the Quasimoon turns out to be 68 km less than the Moon's radius, and the mean density is $3.68 \times 10^3 \text{ kg/m}^3$; however, the evolution of the Quasimoon provides

us with a qualitative evaluation of the evolution of the Moon). The results of the calculations are presented in Table 54.1 and Figure 54.5(a).

The table shows the present-day radii of the cores R_0^* , the initial and limit central densities ρ_{c0} , ρ_{c1} in 10^3 kg/m^3 and pressures p_{c0} , p_{c1} in megabars, the present-day dimensionless moments of inertia I_0^* , and the limit values of the gravitational energy $E(x = 1)$ in 10^{25} MJ . We note that the radii R of all the planets are decreasing monotonically, but by no more than 0.4%. The concentration c_0 of heavy matter decreases with increasing distance from the Sun; the slight departure of the Earth from this law is partly compensated by the low value of c_0 for the Moon. The value of $E(x)$ drops sharply with decreasing size of the planet: for Venus it is only a quarter of the value for the Earth, but for Mars it is 1/100, for Mercury 1/200, and for the Quasimoon 1/8000 of the value for our planet. The radiogenic heat is approximately proportional to the mass, so that its relative role increases for the small planets.

Now let us go on to establish the chronological scale $x(t)$. In order to do this, Sorokhtin (1971) assumed that the rate of increase of the core mass $c_0 M \dot{x}$ is proportional to its surface area $4\pi R_0^2$ and the concentration of core material in the mantle $c = c_0 (1 - x) (1 - c_0 x)^{-1}$, or, more precisely, to its difference $c - c_1^*$ from the saturation concentration c_1^* of the core material in the mantle solid solutions at the core surface:

$$c_0 M \dot{x} = \mathcal{K} \cdot 4\pi R_0^2 (c - c_1^*), \quad (54.24)$$

where $\mathcal{K}$ is a proportionality factor. The latter factor has to be found in accordance with the assumed representation of the micromechanisms of distribution of mantle and core materials. At first a simplified representation was assumed, whereby the distribution took place as a surface reaction, equivalent to melting, so that $\mathcal{K}$ can differ only by some numerical factor from the rate constant of this reaction. Then, in (54.24) we substituted $R_0 = (c_0 M x / \frac{4}{3} \pi \rho_c)^{1/3}$, where ρ_c is the mean core density, which, as mentioned above, decreases slightly (quasilinearly) with increasing x , that is, $\rho_c = \rho_\infty (1 + \delta - \delta x)$, where $\delta \sim 10^{-2}$. Factor $\mathcal{K}$ may in general be a function of the temperature and pressure at the core surface. The

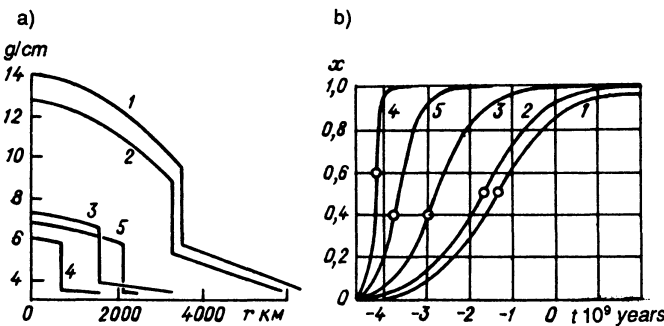


Fig. 54.5. Terrestrial planets: (a) present structure $\rho(r)$; (b) evolution $x(t)$; (1) Earth, (2) Venus, (3) Mars, (4) Mercury, (5) Quasimoon.

TABLE 54.1
Evolution of terrestrial planets

Planet	$R^*(\text{km})$	$M/M_\oplus$	c_0	x^*	$R_0^*(\text{km})$	$\rho_{c,0}$
Earth	6371	1	0.373	0.863	3451	11.34
Venus	6050	0.815	0.358	0.92	3342	10.37
Mars	3394	0.1077	0.183	1	1615	6.65
Mercury	2439	0.0543	0.792	1	2144	6.60
Quasimoon	1738	0.0123	0.149	1	759	5.96

Planet	$\rho_{c,1}$	$\rho_{c,0}$	$\rho_{c,1}$	I_1^*	E_1	k_0
Earth	14.08	2.30	3.97	0.3308	1.86	0.891
Venus	12.64	1.790	3.050	0.3337	1.252	0.994
Mars	6.99	0.267	0.377	0.3712	0.167	1.689
Mercury	6.76	0.252	0.303	0.3639	0.00748	5.748
Quasimoon	6.02	0.058	0.075	0.3778	0.000228	3.283

value of c in (54.24) should, strictly speaking, be taken at the core surface, rather than as an average over the mantle; this difference may create a transition layer D'' . The saturation concentration of core material in the mantle solid solutions is given by the usual thermodynamic formula

$$c^* = c_0^* \exp \left(-p \frac{\delta V}{RT} \right), \quad (54.25)$$

where c_0^* is the value for $p = 0$, δV is the variation of the molar volume of the solution during the dissolution of one mole of core material, and R is here the gas constant. The value of c^* decreases with depth and at the surface of the Earth's core has a value of about 0.03 (see below). If we ignore in the above-mentioned simplified representation all the minor corrections, then Equation (54.24) can be put into the form $\dot{x} = k_0 x^{2/3} c$, where $k_0 = 4\pi R_0^2(1) \mathcal{K}/M$. It is integrated in elementary functions, and from the value $x = 0.863$ for the Earth at present for $t = 4.6$ billion years we obtain a universal constant $\mathcal{K} = 33.4 \text{ kg}/(\text{m}^2 \text{ year})$. After this, for each planet we find the quantity k_0 determining its rate of evolution. These values, in $(10^9 \text{ years})^{-1}$, are given in Table 54.1. The evolution curves $x(t)$ for the various planets are shown in Figure 54.5(b).

The higher the value of k_0 for a planet, the greater will be its evolutionary age at a fixed time. Since $k_0 \sim R_0/M$, the rate of evolution will be lower for the more massive planets, while for a fixed planetary mass it will be less, the smaller the core of the planet, that is, the lower the amount of core material in it. Figure 54.5(b) shows that the most iron-rich planet Mercury evolved the most rapidly (the highest rate of evolution $\max \dot{x}$ was attained at $t_m = 0.47$ billion years, and the end of evolution $x = 0.99$ corresponds to about $t_\infty = 0.7$ billion years). Then the small Moon solidified ($t_m = 0.86$; $t_\infty = 2.13$), and, only recently, the more massive Mars

($t_m = 1.65$; $t_\infty = 4.15$); Venus is still evolving ($x^* = 0.92$; $t_m = 2.94$; $t_\infty = 5.93$), and the Earth is evolutionally even younger ($x^* = 0.863$; $t_m = 3.25$; this time, 1.4 billion years ago, corresponds to the so-called Gotskaya tectonomagmatic epoch, dividing the Early and Middle Riphean and considered to be the beginning of the *Neogaea*; $t_\infty = 6.45$, so that on this scale the Earth will continue a slowing evolution for another 1.8 billion years).

Note, however, that the considered chronological scale $x(t)$ obtained from (54.24) for the simplifying assumption $\mathcal{K} = \text{const}$ provides only a qualitatively plausible picture of the evolution of the planets, which is quantitatively in need of correction. For example, the evolution rate $\dot{x}_1 = \dot{x}(t_1) = 0.15 (10^9 \text{ years})^{-1}$ obtained for the Earth at the present time $t = t_1$ gives an excessively high rate $(\varepsilon_G)_1 = (\dot{x}_1/x_1) E(x_1) \approx 9.3 \times 10^7 \text{ MW}$ (for $x_1 = 0.863$ and $E = 1.6 \times 10^{25} \text{ MJ}$), three times higher than the most likely thermal-equilibrium value $\varepsilon_G = \varepsilon_F - \varepsilon_R - \varepsilon_T = (4.5 - 1.13 - 0.276) 10^7 \text{ MW} = 3.094 \times 10^7 \text{ MW}$.

In order to 'slow down' the present-day evolution, it is necessary to 'speed up' the earlier evolution, that is, according to Monin and Sorokhtin (1981), to change from an insufficiently intensive *surface* reaction of separation of core material from mantle material with $\mathcal{K} = \text{const}$ to a more intensive *volume* reaction with $\mathcal{K} = \mathcal{K}_1 g_0$, where $g_0 = GM_0/R_0^2$ is the free-fall acceleration at the surface of the core, $M_0 = c_0 x M$ is its mass, and R_0 is its radius. Such a volume differentiation of the iron oxides and silicates should take place, in view of the dissociation of their solid solutions at high pressures and, say, depths greater than 2000 km (where a considerable reduction of the quality factor of the mantle is observed), the diffusion of iron oxides from silicate grains into the intragranular spaces (on condition that $c > c^*$), and finally, the downward percolation of dissolved core material under the influence of gravity. Equation (54.24) for $\mathcal{K} = \mathcal{K}_1 g_0$ and its solution are:

$$x = \frac{x}{\tau} \left(c_0 \frac{1-x}{1-c_0 x} - c_1^* \right); \quad \tau = (4\pi \mathcal{K}_1 G)^{-1}; \quad (54.26)$$

$$\frac{t-t_0}{\tau} = \frac{1}{1-c_1^*} \left[\ln \frac{a-x}{a-x_0} - \frac{1}{ac_0} \ln \frac{x_0(a-x)}{x(a-x_0)} \right];$$

$$a = \frac{c_0 - c_1^*}{c_0(1-c_1^*)} \quad (54.27)$$

(Monin and Sorokhtin, 1982). In order to calculate a , we take into account that at zero pressure the miscibility of the iron oxides with the silicates is complete, so that in fayalite Fe_2SiO_4 we have $c_0^*(\text{Fe}_2\text{O}) = 0.63$. In the ilmenite of the present mantle, with the appropriate scaling, $c(\text{Fe}_2\text{O}) = 0.073$. Assuming at a depth of 2000 km in the mantle $c^*(\text{Fe}_2\text{O}) = 0.073$ and taking there an adiabatic temperature 2450 K and a pressure $p = 0.87 - 0.13 = 0.74 \text{ Mbar}$ (where 0.87 is the actual pressure, and 0.13 is the pressure of the first phase transition in iron), we

obtain $\delta V = 0.59 \text{ cm}^3/\text{mole}$. Considering that the ratio $\delta V/V$ varies little with depth, at the bottom of the mantle ($p = 1.35 \text{ Mbar}$, $T = 2650 \text{ K}$) we obtain $c_1^*(\text{Fe}_2\text{O}) = 0.027$. According to (54.4) $c_0 = 0.367$, so that finally $a \approx 0.946$.

Function $x(t)$ according to (54.27) differs significantly from the solution for $\mathcal{K} = \text{const}$, by the presence of a singularity at $x_0 = x(t_0) \rightarrow 0$, which appears because in Equation (54.26) the free-fall acceleration g_0 was introduced, and it goes to zero at the Earth's center. Therefore, gravitational differentiation cannot begin from this point, and to initiate it there must exist a nucleus of the core with a relative mass $x_0 > 0$, at the surface of which $g_0 > 0$.

An early core nucleus could have been formed by the melting of the center of the Protoearth, due to the heat liberation of short-lived radioisotopes. However, it is more certain that a recent core nucleus was formed as a result of the appearance in one of the upper layers of the Earth of a melt of nuclear substance, which then moved toward the Earth's center as drops or streamlets. The most suitable layer for this is the one in which, at the end of the Earth's formation, there existed a temperature maximum (according to the cosmogonic calculations of Safronov, 1978, $T_{\text{max}} = 2200 \text{ K}$ at a depth of about 1500 km), and the melt was caused by two things: the gradual heating up of the Earth due to the heat liberation of long-lived radioisotopes (then distributed approximately uniformly throughout the interior of the Earth) and the dissipation of tidal energy.

Because of the release of gravitational energy, the recent nucleus of the core was probably created violently and rapidly. Calculations of the structure of the primordial homogeneous Earth according to (54.1) and (54.19) give $p_c = 2.30 \text{ Mbar}$, below the second electronphase transition in iron, so that the core nucleus was a melt of Fe_2O (with a density of $11.34 \times 10^3 \text{ kg/m}^3$ and an adiabatic temperature of 2900 K). The subsequent evolution of the Earth's interior took place more calmly, in accordance with formula (54.27).

This formula contains two parameters, x_0 and τ . The second of these can be found from (54.26), having specified the present (that is, for $x = 0.863$) value $\dot{x}_1 = (x \varepsilon_G / E)_1$, which at the above-mentioned thermal equilibrium $(\varepsilon_G)_1 \approx 3.094 \times 10^7 \text{ MW}$ and at Sorokhtin's estimate of $1.23 \times 10^{25} \text{ MJ}$ for the thermal part of E_1 , is $\dot{x}_1 \approx 0.065 (10^9 \text{ years})^{-1}$. This gives $\tau \approx 640$ million years. Then parameter x_0 can be found from (54.27), once we specify the age of the core nucleus $t_1 - t_0$. Assuming $t_1 - t_0 = 3.8$ billion years (the age of the most ancient continental magmatic and sedimentary rocks), we get $x_0 \approx 0.35$, that is, for these initial data the nucleus of the core was found to be very large. Now we can determine from (54.12) the maximum tectonic activity $\max \dot{y}$, reached at $x = a(2 - ac)^{-1}$, which has an age $t_1 - t_m \approx 2.6$ billion years (this corresponds to the so-called Belomorskaya tectonomagmatic epoch, separating the Archean from the Proterozoic). Figure 54.6 shows the corresponding $\dot{x}(t)$ and $\dot{y}(t)$ curves.

To conclude this section, let us consider briefly the evolution of the giant planets. Present-day ideas about their internal structure, which were discussed in Section 43 using Jupiter as an example, picture these planets as having fluid hot adiabatic interiors, existing in a state of convection. Table 54.2 presents a number of their characteristics. The solar constant I is given in units of 10^{-3} times the value at the Earth. The mean radii R and masses M (in fractions of the values for

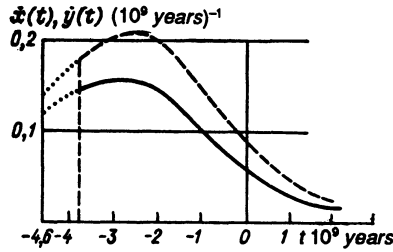


Fig. 54.6. Rate of Earth's evolution assuming volume differentiation of planetary interior. Solid curve $\dot{x}(t)$, dashed curve $\dot{y}(t)$.

the Earth) are known quite precisely. The mean densities $\bar{\rho}$ clearly demonstrate the difference between these planets and the terrestrial planets: the giant planets are in a fluid state, J and S being hydrogen-helium planets (the ratio He/H for J is apparently close to that of the Sun, while in the upper layers of S it is possibly somewhat less than the solar value) and U and N apparently consisting mainly of 'ices': of water, methane, ammonia, etc. The rotation period τ is needed to account for the oblateness of the figures of these planets. The observed effective surface temperatures T were found to be considerably higher than the equilibrium temperatures T_r , calculated from the solar constant and the albedo according to formula (43.5); this points to the existence of an internal heat flux q_i .

The further data in Table 54.2 were obtained in calculations according to some of the models of Zharkov and Trubitsyn (1980). These include the relative radius R_m/R of metallization of hydrogen and the selected (according to the mean density of the planet) values of the mass M_c/M and radius R_c/R (as well as the pressure p_c at the boundary) of the core, consisting of a 'heavy component' and 'ices'. Also given are the pressure p_c , the temperature T_c (in 10^3 K), and the density ρ_c at the center of the planet. Finally, the table also gives estimates of the internal heat flux q_i in 10^{-1} W/m³.

The evolution of the giant planets consists in a contraction of these bodies under the influence of their intrinsic gravitation. The protoplanetary stage was the most dramatic stage of all. For instance, according to the calculations of Bodenheimer *et al.* (1980), a protoplanet with the mass of Jupiter (74% hydrogen and 24% helium) had an initial radius 4530 times greater than the present value of R and the following characteristics typical of the circumsolar nebula: a mean density of 1.47×10^{-8} kg/m³; a central density of 6.3×10^{-8} kg/m³; a mean temperature of 45 K, a central temperature of 51 K, and an irradiance $L = 10^{-4.1}$ of the solar irradiance $L_\odot$. After 4.5×10^5 years it contracted to a radius 168 times greater than the present radius with $\rho_c = 2 \times 10^{-5}$, $T_c = 2100$ K, and $L = 10^{-6} L_\odot$, where down to $T_c = 1600$ K it was convective, after which a radiative-equilibrium core gradually formed (35% in mass).

Next, as a result of the dissociation of molecular hydrogen, gravitational collapse took place, which after 0.25 year produced a hydrostatic core with $\rho_c = 0.2 \times 10^3$ kg/m³ and $T_c = 2 \times 10^4$ K and an outward-propagating shock wave of accretion of external material with a radius of $3R$ (at which molecular hydrogen

TABLE 54.2
Structural data for giant planets

Planet	$10^{-3}I/I_{\oplus}$	$M/M_{\oplus}$	$R(\text{km})$	$\rho \times 10^{-3}$ (kg/m^3)	τ h min s	T (K)	T_r (K)
Jupiter	37	317.8	70000	1.314	9 50 30	130 ± 4	109.4
Saturn	11	95.2	57820	0.703	10 14	95 ± 5	76 ± 4
Uranus	2.7	14.6	25800	1.25	17 12	58 ± 2	57
Neptune	1.1	17.2	24500	1.67	17 48	55.5 ± 2	46

Planet	R_m/R	M_s/M	R_s/R	p_s (Mbar)	p_c (Mbar)	$10^{-3} T_c$ (K)	$\bar{\rho} \times 10^{-3}$ (kg/m^3)	$10q_i$ (W/m^3)
Jupiter	0.76	3–4%	0.15	45	78	24	16.5	8100 ± 1500
Saturn	0.46	20–25%	0.27	10	48	22	13.5	2800 ± 1400
Uranus		0.94	0.78	0.12	5.8	11	4.43	< 180
Neptune		0.97	0.85	0.07	7.4	12	4.85	285 ± 80

becomes heated and dissociates). Almost the entire planet enters into a hydrostatically equilibrium convective state in 10^7 years, its radius reaching $1.3R$, $\rho_c = 2.73 \times 10^3 \text{ kg/m}^3$, $T_c = 3.5 \times 10^4 \text{ K}$, and $L = 10^{-5.65} L_{\odot}$.

The further evolution (over 4.5 billion years) consists in a quasihydrostatic contraction and cooling, for which radiation $4\pi R^2 \sigma (T^4 - T_r^4)$ takes place via a reduction of the sum of the gravitational energy ($-U$) (defined by formula (54.20)) and the internal energy. According to existing estimates, the present value of q_i for Jupiter provides a contraction $\dot{R} \approx 1 \text{ mm/year}$.

55. MANTLE CONVECTION

Convection is motion which arises in a medium with an unstable density stratification $\rho_0(r)$, as a result of the action of buoyancy forces $\rho'g$ on volumes with density fluctuations ρ' . What happens is that volumes lighter than the surrounding fluid ($\rho' < 0$) rise, while heavier volumes ($\rho' > 0$) sink (see (11.1)–(11.6) and (17.3)–(17.5)).

Thermal convection (for which $\rho' \sim (-T')$) is produced, due to heating from below, in the outer layer of the Sun (thickness about $\frac{1}{3}R_{\odot}$), in the interiors of the giant planets, in the atmosphere of Venus, at times (and in the intertropical zone always) in the lower layers of the Earth's atmosphere, and also, due to cooling from above, in some regions of the World Ocean.

Convection may also arise in the mantle of the Earth, which, contrary to what was stated in Section 52, for long relaxation times (more than 10^4 years) should be assumed not to be solid, but rather *viscoelastic* or even *viscoplastic*. A viscoplastic body experiences plastic deformations when the stresses upon it exceed some *yield threshold*, the deformation rate being proportional to the amount the stress

exceeds this threshold. For the Earth the equation of state of the mantle material can be written as $\rho = \rho(p, T, c)$, where c is the concentration of heavy core material, and the density fluctuations are

$$\frac{\rho'}{\rho_0} \approx -\alpha T' + \beta c', \quad (55.1)$$

where α and β are positive coefficients. If here we take into account only the temperature fluctuations T' , then the convection can be treated as thermal; if only the concentration fluctuations c' are taken into account, then we are dealing with *concentration convection*.

Many investigators, considering just the thermal convection, have limited its development to the upper mantle, to depths of 650–700 km, where an increase in the velocities of seismic waves is observed, in connection with the transition of quartz to the more densely packed crystal lattice of stishovite. This limitation originated, firstly, from the fear that the surfaces of the phase transitions for convection are impermeable and, secondly, from the fact that the Rayleigh number $Ra = \rho g a h^3 \delta\Theta / k \eta$ in the lower mantle (which has a layer thickness $h \sim 2500$ km, $\rho \sim 5 \times 10^3$ kg/m³, $\alpha \sim 10^{-5}$ K⁻¹, a thermal diffusivity $k \sim 10^{-6}$ m²/s, and, what is most important, an enormous viscosity $\eta \sim 10^{25}$ Pa s) can reach the lower (for thermal convection) critical value of the order of 2×10^3 only for unrealistically great differences in potential temperature $\delta\Theta \sim 2600$ K. In the upper mantle, on the other hand, this difficulty is removed, since there the viscosity is several orders of magnitude lower. However, convection in a thin layer yields values of the geothermal heat flux and velocities of lithospheric plates which are an order of magnitude too low, and it can hardly be expected to take the form of the cells of global horizontal dimensions which are typical of the main lithospheric plates.

There are also some more important arguments showing that the concentration factor has to be taken into account. For example, radiogenic heat liberation does not constitute an adequate source of thermal convection. Actually, radioactive isotopes are lithophylic and a significant portion of them exit from the mantle into the Earth's crust, where their heat release (that is, heating from above) does not promote, but rather hinders, the mantle convection. In the mantle, however, according to the data of the preceding section, there is now only 52%, or 5.8×10^{16} kg, K⁴⁰; 16%, or 6.8×10^{15} kg, U²³⁸; 16%, or 4.9×10^{13} kg, U²³⁵; and 9.3%, or 1.4×10^{16} kg, Th²³², enough to create a radiogenic heat release at a rate of 2.7×10^6 MW, only about 6% of the present heat loss by the Earth. Even less heat (three times less) is supplied to the mantle by the dissipation of tidal energy.

Consequently, the main source of thermal convection must be the heat liberation accompanying the gravitational differentiation of matter. This, incidentally, also takes place mainly just beneath the mantle, in the upper layers of the core, so that the convection produced by it has to penetrate the entire mantle from bottom to top. However, a concentration factor also enters in during gravitational differentiation, since the volumes losing part of their heavier material become lighter and should move upward.

If the stratification is not too highly unstable, the convection motions are subdued even *laminar* motions, traveling horizontally as rolls or cells, sometimes as hexagonal Bénard cells (see (17.6) and (17.7)). In the case of weak instability, *closed* cells are formed, at the centers of which the movement is upward and at the periphery downward (these are sometimes figuratively called Abel bonfires). For greater instability, *open* cells are formed, with downward motions at the centers and upward motions at the periphery (Cain bonfires).

In a spherical layer like the Earth's mantle the simplest form of convection turns out to be a single global cell with opposite poles of uplift and subsidence of material. Then the mantle currents carry the lithospheric plates floating on the asthenosphere into the region of the subsidence pole. If this kind of convection persists for a long enough time, then all the continents become massed together into the subsidence region, that is, a single supercontinent is created.

The direction $\mathbf{m} = \boldsymbol{\omega}/\omega$ of the Earth's rotation axis should remain unchanged (in the space of the fixed stars), due to the conservation of angular momentum on the time scales of the convective cycles (of the order of 10^2 million years). On the other hand, the Earth should so turn about $\mathbf{m}$ that the maximum of the uplifted continental masses lies at the rotation equator (since the rheology allows the formation of an equatorial bulge, it will obviously also allow it to move across the body of the Earth). Such a rotation of the body of the Earth can be portrayed as a motion of the rotation pole over this body.

Therefore, for single-cell convection the center of the supercontinent (the convection pole) must lie at the rotation equator, this is, at a distance of about 90° from the rotation pole. According to paleogeographic reconstructions, this situation existed 300 to 220 million years ago, at the time of the supercontinent Pangaea (Wegener, 1912); the South Pole then lay at the edge of Pangaea, near the tips of South America, Southern Africa, and Western Antarctica. Note, too, that the surfaces of Mars and the Moon have a single-cell structure (Monin, 1982).

The next most complicated form of mantle convection is two-cell convection with closed cells, in which there are two opposite poles of uplift and an equator of subsidence of material (the convection equator may be curved, like the seam on a tennis ball glued together out of two pieces of material). In this case the mantle currents must carry the lithospheric plates toward the subsidence equator and distribute the continents along it, so that the rotation equator will coincide with the subsidence equator. If during the transition from single-cell convection to two-cell convection with closed cells (or vice versa) the convection axis in the body of the Earth remains unchanged, then the rotation pole must be shifted thereby by 90° .

For two-cell convection with open cells there will be an equator of uplift and two poles of subsidence of mantle material, in the vicinities of which the continents will become massed together. The rotation equator must then pass through the subsidence poles (as in the case of single-cell convection). The present movement of the lithospheric plates of the Earth conforms well to this type of convection (in the 'tennis-ball' configuration). The part of the uplift equator is played by the chain of midoceanic ridges (almost closed through the Arctic but complicated by

branching through the Indian Ocean into the Red Sea and the East African Rift). The part of one of the subsidence regions is played by the Western-Pacific region of subduction zones (complicated by branching in the Himalayan—Mediterranean belt), while that of the other is played by the Eastern-Pacific subduction zones in Central and South America (with branchings in the Caribbean and toward the Hawaiian Islands).

The two-cell structure of the Earth also manifests itself in the geophysical fields (Monin, 1986), namely in the deviations (up to ± 100 m) of the geoid from an axisymmetric reference ellipsoid, which are negative in two regions (Asia-Antarctica and North America) and positive in the belt separating these, as well as in the nondipole component H' of the geomagnetic field (Asian and American regions with the same sign of H'_z , separated by a band with the opposite sign of H'_z). The two-cell structure of the surface of Venus should also be mentioned (two approximately antipodal supercontinents: Ishtar and Aphrodite).

It has been suggested that the single-cell and two-cell forms of mantle convection alternate; see Figure 55.1 (Monin and Sorokhtin, 1977). In the single-cell form 55.1(a), a stagnation zone St is formed, in which there is a tendency toward subsidence (subsidence equator Su in intermediate form 55.1(b)). This leads to a two-cell form with closed cells 55.1(c). In it two stagnation zones St already appear. With the breakthrough of the lighter material which has accumulated under the St zones in zones Su (Figure 55.1(d)), along one of the existing uplift axes, the single-cell mode of convection is restored. If, on the other hand, the breakthrough occurs in the zone of the previous subsidence equator, then a two-cell mode with open cells (Figure 55.1(e)) appears, and after it decays via the intermediate form 55.1(f) we return to the initial single-cell form.

This hypothesis can account for the alternation of tectonomagmatic epochs in the Earth's history, with a possible formation of supercontinents about 2600, 1650 (the Megagaea of Stille), 950, and 300 (the Pangaea of Wegener) million years ago.

In order to calculate the mantle convection within the framework of the two-phase model discussed in Section 54, we assume $\rho = (\Sigma_1^2 c_n \rho_n^{-1})^{-1}$ and $\mathbf{u} = \Sigma_1^2 c_n \mathbf{u}_n$, where $n = 1$ corresponds to core material and $n = 2$ corresponds to mantle material. According to Barenblatt (1953), the horizontal components of the current velocities of both materials can be assumed to be the same, while the vertical components differ by an amount w (the 'hydraulic size' of the particles of the heavy substance), so that

$$u_{1r} = u_r - w(1 - c); \quad u_{2r} = u_r + wc. \quad (55.2)$$

From the laws of conservation of mass of the two materials we obtain the continuity equation for the mixture, which can be approximated as $\text{div } \mathbf{u} = 0$, and the equation of convective diffusion of the heavy substance:

$$\rho \frac{dc}{dt} = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \rho w c (1 - c) - \text{div } \mathbf{i}, \quad (55.3)$$

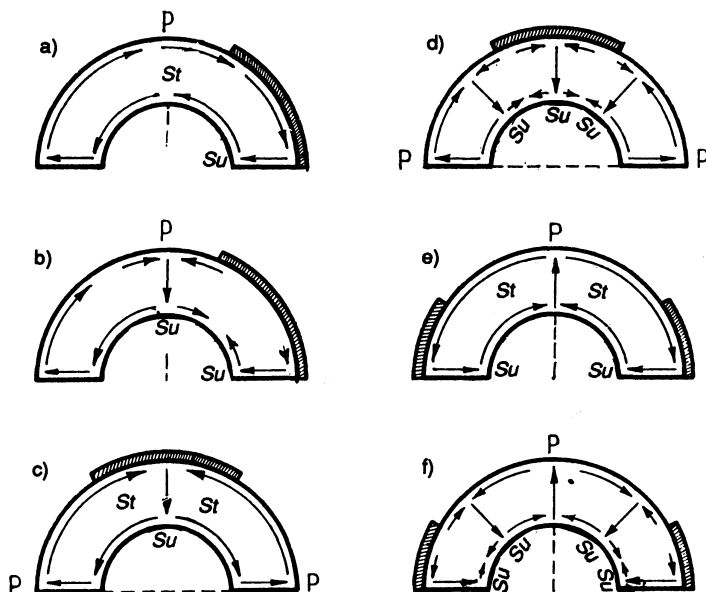


Fig. 55.1. Alternation of forms of mantle convection.

where i is the density of the diffusive mass flux of the heavy substance. In the equations of motion the use of formulas (55.2) entails the addition to the free-fall acceleration g of the sedimentation acceleration

$$g' = \frac{1}{r^2 \rho} \frac{\partial}{\partial r} r^2 \rho w^2 c (1 - c). \quad (55.4)$$

Moreover, the viscosity coefficient acquires the Einstein correction factor $1 + \frac{5}{2} c \rho / \rho_1$. The equations of motion can be taken in the Boussinesq approximation, and because of the slowness of the mantle convection and the enormous viscosity we can ignore the inertial forces in comparison with the viscosity forces. Then the equation for the vorticity simplifies to

$$\nabla r \times \nabla (g \rho' + \rho_0 g') + \text{curl } \mathbf{F} = 0, \quad (55.5)$$

where $\mathbf{F}$ is the viscosity force. In the axisymmetric case, with the aid of the condition of nondivergence of the velocity in the meridional convection plane, we can introduce the stream function ψ , setting

$$u_r = \frac{1}{r \sin \theta} \frac{\partial \psi \sin \theta}{\partial \theta}; \quad u_\theta = -\frac{1}{r} \frac{\partial \psi r}{\partial r}. \quad (55.6)$$

Here it is sufficient to consider only the longitudinal projection of Equation (55.5),

which becomes

$$\begin{aligned} \frac{g}{r} \frac{\partial \rho'}{\partial \theta} + \frac{\rho_0}{r} \frac{\partial g'}{\partial \theta} = \frac{1}{r} \frac{\partial}{\partial \theta} \left(\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \eta \Phi_{rr} + \right. \\ \left. + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} \eta \Phi_{r\theta} \sin \theta - \eta \frac{\Phi_{\theta\theta} + \Phi_{\lambda\lambda}}{r} \right) - \frac{1}{r} \frac{\partial}{\partial r} \left(\frac{1}{r^2} \frac{\partial}{\partial r} r^3 \eta \Phi_{\theta r} + \right. \\ \left. + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \eta \Phi_{\theta\theta} \sin \theta - \eta \Phi_{\lambda\lambda} \operatorname{ctg} \theta \right). \end{aligned} \quad (55.7)$$

where Φ_{ik} is the tensor of the deformation rates, expressible using ordinary formulas in terms of velocity components (55.6), and thus in terms of ψ . This equation should be integrated along with the equation of state (55.1) for $\beta = \rho_0(\rho_1 - \rho_2)\rho_1^{-1}\rho_2^{-1}$, the equation of convective diffusion (55.3) (with the boundary condition of absorption of core material $c = 0$ or for a specified flow of this material at the lower boundary of the mantle), and the heat-transfer equation. The preliminary numerical experiments (without the thermal factor and, unfortunately, for the time being only for low c values) represented, as a rule, 3 or 4 cells with a quasiperiodic transfer of energy between them, so that most of the time one or two cells predominated. The convection velocities obtained ranged up to tens of cm/year.

However, in these preliminary experiments no successful calculation was made of the evolution of mantle convection with a growth in time of the Earth's core: the evolution slowed down prematurely, and the growth of the core ceased. Therefore, assuming a volume separation of substances ($\mathcal{K} = \mathcal{K}_{1g_0}$) instead of a surface melting out of iron from silicates at the core surface ($\mathcal{K} = \text{const}$ in (54.24)), and assuming seepage of a heavy fluid through a porous medium instead of sedimentation of heavy 'particles' with an artificially selected hydraulic size w , Monin *et al.* (1987) solved equations (55.3) and (55.7) for $w = 0$ (and without an Einstein correction to the viscosity), but with the following boundary condition for the concentration of the heavy substance at the core—mantle interface:

$$\mathcal{D} \frac{\partial c}{\partial r} = V(c - c^*), \quad (55.3')$$

where $\mathcal{D} \approx 0.01 \text{ cm}^2/\text{s}$ is the coefficient of diffusion of the heavy substance, $c^* \approx 0.027$ is its saturation concentration in the solid solutions, and V is the effective velocity of its movement from the mantle to the core, which, in the numerical experiments, varied between 0.01 and 0.5 cm/year. The equations were integrated numerically in an axisymmetric model on a uniform grid of 1037 points in a meridional section of the mantle, with time step of about 1/4 million years in

$N = 10,000$ steps with an initial state of rest. A number of numerical experiments were carried out for different values of parameters r_1 (core radius), $\nu(r)$ (mantle viscosity), and V .

In all the experiments, after $N = 500$ to 1100 steps a maximum energy of convection was reached in a four-cell form (see the $N = 600$ plot in Figure 55.2, which depicts the streamlines $\psi = \text{const}$ of the convective circulations in a meridional section of the mantle, rectified on the figure, whereas the calculation pertained to a spherical layer; circulations in different directions are represented by dashed and solid lines). Then, up to $N = 1700$ –2500 the energy dropped to a minimum, retaining the four-cell form of convection, and weakly damped relaxation self-oscillations with periods of the order of 150 to 200 million years began. It should be noted that the value $V = 0.01$ was obviously low (for it, the mean value of c decreased during 10,000 steps by only 0.005), while the value $V = 0.5$ was apparently rather high (for it self-oscillations were observed only up to $N = 3000$, after which the convection energy decreased monotonically).

The most interesting results were obtained in the experiment with $r_1 = 3000$ km, $c_0 = 0.100$ with a random disturbance δc of amplitude 0.001 (the Earth 1 billion years ago) for $\nu = \text{const} = 10^{22}$ cm²/s and $V = 0.05$ cm/year. These results are shown in Figure 55.2. Here from the moment of maximum energy $N = 600$ in field c a zone of uplift of lightened material was formed, which gradually moved toward the left. After a time $N = 2000$ –2100 (again an energy maximum) a second uplift zone formed to the right of the first one, and for a short time the convection became six-cell (with very weak polar cells). By $N = 2500$ –2600 these uplift zones had faded and merged, and the four-cell form returned (due to the weakness of the polar cells, this was practically two-cell) with an energy maximum at $N = 2800$ –2900.

Then the uplift zone once again shifted to the left, at $N = 5300$ –5400 a second uplift zone again formed to the right of it, and at $N = 5400$ –5700 the convection became five-cell (the left polar cell disappeared; $N = 5500$ was a new energy maximum). By $N = 5800$ –5900 the uplift zones had faded and merged, creating four-cell convection again (almost two-cell); at $N = 6600$ –7000 the left polar cell disappeared.

The uplift zone shifted to the left, and by the new energy maximum at $N = 6800$ –6900 the convection became three-cell; then the weak left polar cell again appeared, and the convection once more became four-cell. A gradual leftward shift and weakening of the uplift zone during the period $N = 7100$ –8000 led this time to a qualitatively different result: the second uplift zone to the right of the first now did not form, but instead the right-hand tropical cell spread to the left pole (inclusively), keeping at the same time its descending branch in the right circumpolar zone (where the polar cell became so weak that the convection became practically one-cell). For $N = 8000$ –8100 here a new energy maximum was reached, and for $N = 8200$ –8300 for the first time the convection pole turned out to be a single region of uplift of lightened material.

At $N = 8300$ in the convective cell around the pole two cores formed, and at $N = 8400$ between these in the lower half of the mantle a small opposite cell and a region of uplift of material appeared. By $N = 8500$ these had spread along the

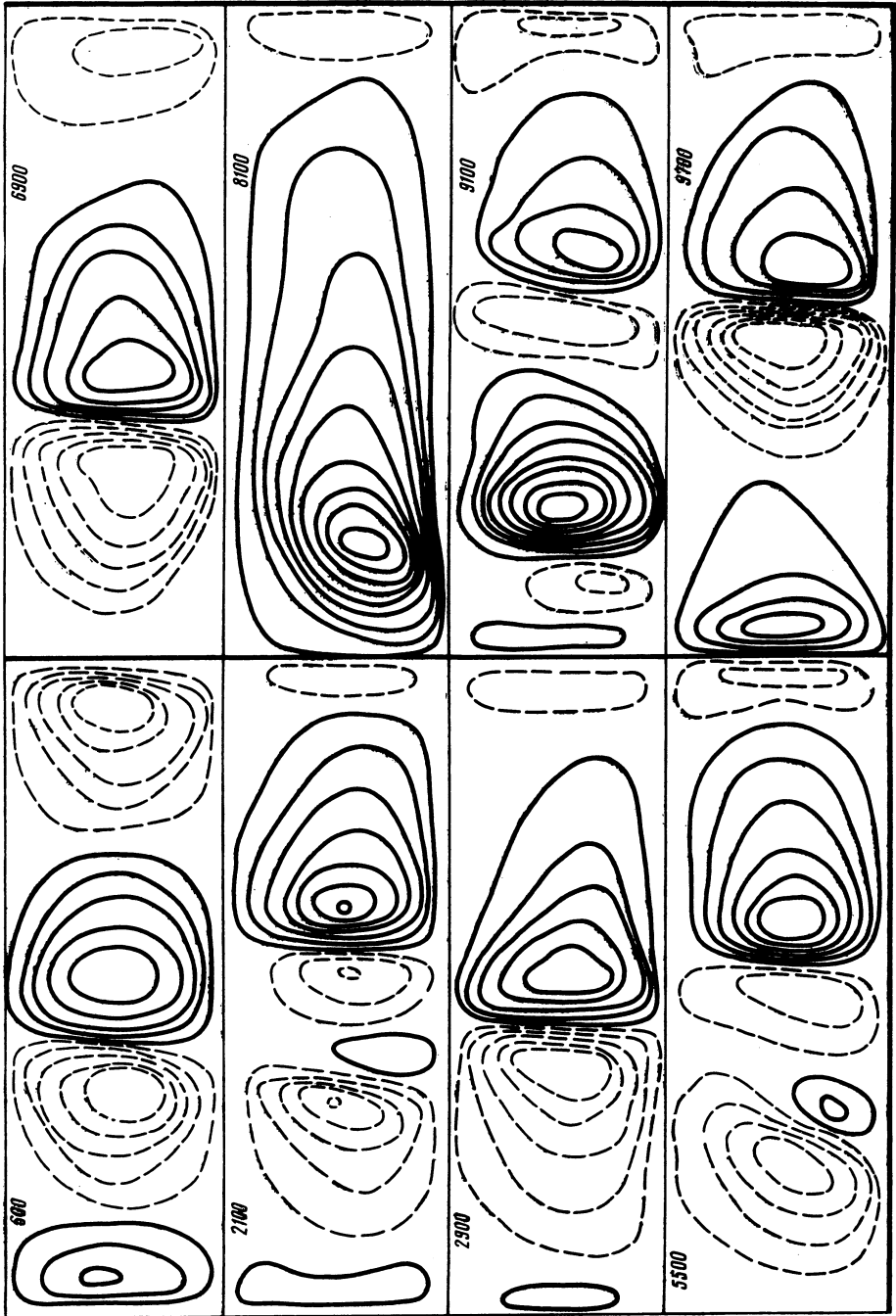


Fig. 52.2.

vertical throughout almost the entire thickness of the mantle, but relative to this circumpolar disturbance the one-cell form proved to be stable, and by $N = 8600$ – 8800 the disturbance had disappeared. At $N = 8800$ two cores again appeared in the cell, and at $N = 8900$ between these around the equator an opposite cell and a zone of uplift formed, that is, there was a return to four-cell convection. The one-cell form proved to be unstable with respect to such an equatorial disturbance: at $N = 9000$ in the left-hand circumpolar cell there appeared a second zone of uplift, and at $N = 9100$ – 9200 the convection became six-cell (again with an energy maximum). In all, the one-cell form persisted during a period from $N = 8100$ to 8800 , that is, for about 200 million years.

The second uplift zone then weakened and disappeared, and at $N = 9300$ – 9800 the usual four-cell form occurred (with an energy maximum at $N = 9700$). Then, at $N = 9900$ – $10,000$, a second uplift zone again appeared in the left hemisphere, the six-cell form occurred, and the energy increased.

On the whole, it could be assumed that this experiment verifies qualitatively the convection patterns shown in Figure 55.1 (but it showed that unstable disturbances of the concentration field are zones of uplift of lighter material formed at the bottom, rather than zones of subsidence of relatively heavy material formed at the cell centers). It should be noted, however, that the forms of convection depend considerably on the parameters of the model.

Mantle convection is assumed to be the only candidate for the position of the force driving the tectonics of the lithospheric plates. One of the consequences of this tectonics (and thus, ultimately, of the mantle convection as well) is movement of the poles. The latter can be calculated with the aid of the Liouville equation $\dot{\mathbf{M}} + \boldsymbol{\omega} \times \mathbf{M} = \mathbf{K}$ for the angular momentum of the Earth $\mathbf{M}$ in a frame of reference moving with the Earth, where $\mathbf{K}$ is the moment of the external forces. In a geocentric Cartesian coordinate system $OX_1X_2X_3$ we obtain $M_i = I_{ij} \omega_j + M'_i$, where $\mathbf{M}'$ is the angular momentum of the internal motions inside the Earth (which, like $\mathbf{K}$, can be neglected here) and I_{ij} is the inertia tensor, defined by the formula

$$I_{ij} = \int (x^2 \delta_{ij} - x_i x_j) \rho \, dV = \frac{2A + C}{3} \delta_{ij} + \frac{C - A}{1 + \mu} \left(m_i m_j - \frac{1}{3} \delta_{ij} \right) + I'_{ij}, \quad (55.8)$$

where C and A are the principal moments of inertia, the second term gives the contribution of the deformation of the Earth due to its rotation (equatorial bulge) within the framework of rigid-elastic rheology, $\mu = \frac{19}{2} \mu_0 / \bar{\rho} g R \approx 2.3$ is the dimensionless (μ_0 being dimensional) 'effective tidal' rigidity, while I'_{ij} is the contribution of the nonrotational deformations, produced by the mass distribution between the oceans and continents and, consequently, determined by the plate

tectonics:

$$\begin{aligned}
 I'_{ij} &= \int \left(\delta_{ij} - \frac{r_i r_j}{r^2} \right) [\rho_{oc}(r) (1 - \chi_{con}) - \rho_{con}(r) \chi_{con}] r^4 dr d\Omega = \\
 &= -N \int n_i n_j (\chi_{con} - \bar{\chi}_{con}) d\Omega(\mathbf{n}); \quad N = \int (\rho_{con} - \rho_{oc}) r^4 dr, \quad (55.9)
 \end{aligned}$$

where $\rho_{con}(r)$ and $\rho_{oc}(r)$ are the radial density distributions on the continental and oceanic areas; $\mathbf{n}$ is a variable of integration (a point on a sphere of unit radius) in the absolute ('ancient') coordinate system, whose origin is the *initial* position of the pole; $\chi_{con}[\mathbf{n}, t; \mathbf{m}(t)]$ is a characteristic function of the continents at time t (equal to 1 on them and 0 outside of them), specified empirically (according to paleogeographical reconstructions) in relative coordinates whose origin is the paleopole $\mathbf{m}(t)$ and then converted into coordinates $\mathbf{n}$ via the change $\mathbf{m}(t) \rightarrow \mathbf{m}(0)$.

In order to make migration of the poles possible, we invoke the viscoelastic rheological model of Maxwell, in which the deformation rates [strain rates] are linear combinations of the rates of change of the stresses with a coefficient μ_0^{-1} and the stresses themselves with a coefficient η_0^{-1} , where $\eta_0 = \tau\mu_0$ is the viscosity, and τ is the time of viscous relaxation of the stresses. Then the stresses are proportional to the strains, with a coefficient $\mu_0 \tau \mathcal{D} (1 + \tau \mathcal{D})^{-1}$, where $\mathcal{D} = d/dt$.

Since time τ is much shorter than the typical time needed for a pole to shift a considerable distance, operator $\tau \mathcal{D}$ can be assumed to be small. Therefore, the conversion to this rheology consists simply in replacing μ by $\mu \tau \mathcal{D}$ and $(1 + \mu)^{-1}$ by $1 - \mu \tau \mathcal{D}$ in (55.8). Then with the aid of (55.8) we obtain $\mathcal{D} I_{ij} m_j \approx C \dot{m}_i$ (Monin, 1972) and the Liouville equation can be approximated as

$$C \dot{m}_i - \varepsilon_{ijk} (C - A) \mu \tau \omega m_j \dot{m}_k + \varepsilon_{ijk} \omega m_j I'_{kl} m_l = 0. \quad (55.10)$$

Since the dimensionless viscosity $\eta = \mu \tau \omega (C - A)/2C$ is large (for $\eta_0 \sim 10^{23}$ Pa s we get $\eta \sim 10^5$), the first term in Equation (55.10) can be neglected, so that in view of (55.9) this equation reduces to

$$\dot{\mathbf{m}} = -\mathcal{K} \int \mathbf{n}(\mathbf{m} \cdot \mathbf{n}) (\chi_{con} - \bar{\chi}_{con}) d\Omega(\mathbf{n}); \quad \mathcal{K} = \frac{\omega N}{2\eta C}. \quad (55.11)$$

This equation contains a single parameter $\mathcal{K}$, absorbing the unreliably estimated values of N and η . Preliminary calculations based on Phanerozoic paleogeographic reconstructions with idealization of the paleostable continental blocks as material points (with weighting factors proportional to their areas) revealed that excursions of the pole by the 90° amount predicted above are obtained for $\mathcal{K} = 0.133$. For the estimate of $N = 1.78 \times 10^{32}$ kg m² obtained by Munk (1958), this corresponds to an effective mantle viscosity $\eta_0 = 5.9 \times 10^{22}$ Pa s.

Selected References*

- Agafonova, E. G., Galerkin, L. I., and Monin, A. S.: 1972, 'On the origin of thermohaline circulation in the ocean', *Okeanologiya* **12** (6), 950—955.
- Alfvén, H.: 1942, *Cosmic Electrodynamics*.
- Andreev, O. A., Kagan, B. A., and Oganessian, L. A.: 1976, 'On the global circulation in a two-layer ocean', *Okeanologiya* **16** (1), 5—14.
- Arnold, V. I.: 1965, 'Sur la topologie des écoulements stationnaires des fluides', *C. R. Acad. Sci. Paris* **261**, 17—21.
- Barenblatt, G. I.: 1978, 'Dynamics of turbulence spots and intrusion in a stably stratified fluid', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **14** (2), 195—206.
- Batchelor, G. K.: 1969, 'Computation of the energy spectrum in homogeneous two-dimensional turbulence', *Phys. Fluids* **12** (11), 233—239.
- Bjerknes, V.: 1937, 'Application of line integral theorems to the hydrodynamics of terrestrial and cosmic vortices', *Astrophys. Norw.* **2**, 263—339.
- Blinova, E. N.: 1947, 'On the annual mean temperature distribution in the Earth's atmosphere taking into account continents and oceans', *Izv. Akad. Nauk SSSR, Ser. Geogr. Geofiz.* **XI** (1), 3—13.
- Braginsky, S. I.: 1964, 'Self-excitation of a magnetic field during the motion of a highly conducting fluid', *Sov. Phys. JETP* **20**, 726—735.
- Bretherton, F. P.: 1966, 'Critical layer instability in baroclinic flows', *Quart. J. Roy. Meteor. Soc.* **92**, 325—334.
- Bryan, K. and Cox, M. D.: 1968, 'A nonlinear model of an ocean driven by wind and differential heating', Parts I and II, *J. Atmos. Sci.* **25**, 945—978.
- Bryan, K. and Lewis, L. J.: 1979, 'A water mass model of the World Ocean', *J. Geophys. Res.* **84** (5).
- Bryan, K., Manabe, S., and Pacanowski, R. C.: 1975, 'A global ocean-atmosphere climate model', Part II, 'The oceanic circulation', *J. Phys. Oceanogr.* **5** (1), 30—46.
- Budyko, M. I.: 1968, 'On the origin of ice ages', *Meteorologiya i Gidrologiya*, No. 2, 3—12.
- Bullard, E. C. and Gellman, H.: 1954, 'Homogeneous dynamos and terrestrial magnetism', *Phil. Trans. Roy. Soc.* **A247**, 213—278.
- Charney, J. G. and Stern, M.: 1962, 'On the stability of internal baroclinic jets in a rotating atmosphere', *J. Atmos. Sci.* **19**, 159—172.
- Childress, S.: 1970, 'New solutions of the kinematic dynamo problem', *J. Math. Phys.* **11**, 3063—3076.
- Cowling, T. G.: 1934, 'The magnetic field of sunspots', *Mon. Nat. Roy. Astron. Soc.* **94**, 39—48.
- Cox, M. D.: 1970, 'A mathematical model of the Indian Ocean', *Deep-Sea Res.* **17**(1), 47—75.
- Defant, A.: 1921, 'Die Zirkulation der Atmosphäre in den gemässigten Breiten der Erde', *Geogr. Ann.* **3**, Stockholm.
- Dikii, L. A.: 1969, *Theory of Oscillations of the Earth's Atmosphere* [in Russian], Gidrometeoizdat, Leningrad.

* The author supplied only a short list of recommended books (28 in all). For the convenience of the reader many of the other works referred to in the text have been identified and added to this list by the translator.

- Dikii, L. A.: 1976, *Hydrodynamic Stability and the Dynamics of the Atmosphere* [in Russian], Gidrometeoizdat, Leningrad.
- Eckart, C. H.: 1960, *Hydrodynamics of Oceans and Atmospheres*, Pergamon.
- Eckmann, J. P.: 1981, 'Roads to turbulence in dissipative dynamical systems', *Rev. Mod. Phys.* **53**, 643–654.
- Ertel, E.: 1942, 'Ein neuer hydrodynamischer Wirbelsatz', *Meteorol. Zeitsch.* **59**, 277.
- Faegre, A.: 1972, 'An intransitive model of the earth-atmosphere-ocean system', *J. Appl. Meteorol.* **11**, 4–6.
- Feigenbaum, M. J.: 1978, 'Quantitative universality for a class of nonlinear transformations', *J. Stat. Phys.* **19**, 25–52.
- Fjørtoft, R.: 1950, 'Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex', *Geophys. Publ.* **17** (6), 52.
- Gavrilin, B. L. and Monin, A. S.: 1969, 'On primitive and balanced hydrodynamic equations of meteorological fields', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **5**, 763–774.
- Gavrilin, B. L., Mirabel, A. P., and Monin, A. S.: 1972, 'On the energy spectrum of synoptic processes', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **8** (5), 483–493.
- Gledzer, E. B., Dolzhanskii, F. V., and Obukhov, A. M.: 1981, *Systems of the Fluid-Dynamical Type and Their Application* [in Russian], Fizmatgiz, Moscow.
- Golitsyn, G. S.: 1968, 'Estimates of some general circulation parameters in the planetary atmospheres of the terrestrial group', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **4** (11), 1131–1138.
- Golitsyn, G. S.: 1970, 'Similarity theory for large-scale motions of planetary atmospheres', *Dokl. Akad. Nauk SSSR* **190** (2), 323–326.
- Golitsyn, G. S.: 1973, *An Introduction to the Dynamics of Planetary Atmospheres* [in Russian], Gidrometeoizdat, Leningrad.
- Gollub, J. P. and Benson, S. V.: 1980, 'Many routes to turbulent convection', *J. Fluid Mech.* **100**, 449–470.
- Greenspan, H. P.: 1968, *The Theory of Rotating Fluids*, Cambridge Univ. Press, New York.
- Gubbins, D.: 1973, 'Numerical solutions of the kinematic dynamo problem', *Phil. Trans. Roy. Soc. A* **274**, 493–521.
- Hadley, G.: 1735, 'Concerning the cause of the general tradewinds', *Phil. Trans. Roy. Soc. London* **39**, 58.
- Hale, G. E. and Nicholson, S. B.: 1938, *Magnetic Observations of Sunspots I, II*, Washington.
- Holland, W. R. and Lin, L. B.: 1975, 'On the generation of mesoscale eddies and their contribution to the oceanic circulation', *J. Phys. Ocean* **5** (4), 642–669.
- Hough, S. S.: 1898, 'On the application of harmonic analysis to the dynamical theory of the tides', Part II, 'On the general integration of Laplace's dynamical equations', *Phil. Trans. Roy. Soc. London A* **191**, 139–185.
- Ilin, A. M., Kamenkovich, V. M., Zhugrina, T. G., and Silkina, M. M.: 1969, 'On the computation of total flows in the World Ocean (steady-state problem)', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **5** (11), 1160–1172.
- Jeffreys, H.: 1926, 'On the dynamics of geostrophic winds', *Q. J. Roy. Met. Soc.* **52**.
- Kagan, B. A., Laikhtman, D. P., Oganasyan, L. A., and Pyaskovskii, R. V.: 1972, 'A numerical experiment on the seasonal variability of the global circulation in a barotropic ocean', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **8** (10), 1052–1071.
- Kagan, B. A., Laikhtman, D. P., Oganasyan, L. A., and Pyaskovskii, R. V.: 1974, 'A two-dimensional thermal model of the World Ocean', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **10**, 1118–1122.
- Kamenkovich, V. M.: 1958, 'On some simplifications of the dynamic equations of steady currents in a baroclinic sea', *Dokl. Akad. Nauk SSSR* **119** (6), 1134–1137.
- Kamenkovich, V. M.: 1977, *Fundamentals of Ocean Dynamics*, Elsevier, Amsterdam.
- Kamenkovich, V. M., Koshlyakov, M. N., and Monin, A. S.: 1982, *Synoptic Eddies in the Ocean* [in Russian], Gidrometeoizdat, Leningrad.
- Kibel', I. A.: 1943, 'The temperature distribution in the Earth's atmosphere', *Dokl. Akad. Nauk SSSR* **39** (1), 18–22.

- Kirichkov, S. E.: 1978, 'Numerical experiments with a zonal atmospheric circulation model', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **14** (7), 691–702.
- Kochergin, V. P. and Shcherbakov, A. V.: 1977, 'Numerical modeling of a baroclinic World Ocean', *Meteorologiya i Gidrologiya* **7**, 56–63.
- Kochin, N. E. (ed.): 1936, *Dynamic Meteorology*, Centr. Admin. United Hydrometeor. Service, USSR.
- Kochin, N. E.: 1961, *Vector Analysis and the Beginning of Tensor Analysis* [in Russian], Izd. Akad. Nauk SSSR, Moscow.
- Kraichnan, R. H.: 1967, 'Inertial ranges in two-dimensional turbulence', *Phys. Fluids* **10** (7), 1417–1423.
- Kraichnan, R. H.: 1976, 'Diffusion of weak magnetic fields by turbulence', *J. Fluid Mech.* **75**, 657–676.
- Landau, L. D. and Lifshits, E. M.: 1953, *Continuum Mechanics* [in Russian], GITTL, Moscow.
- Longuet-Higgins, M. S. and Gill, A. E.: 1967, 'Resonant interactions between planetary waves', *Proc. Roy. Soc. (London)* **A299**, 120–140, 141–145.
- Lorenz, E. N.: 1960, 'Maximum simplification of the dynamical equations', *Tellus* **12** (3), 243–254.
- Lorenz, E. N.: 1963, 'Deterministic nonperiodic flow', *J. Atmos. Sci.* **20**, 130–141.
- Manabe, S. and Bryan, K.: 1969, 'Climate calculations with a combined ocean-atmosphere model', *J. Atmos. Sci.* **26**, 786–789.
- Manabe, S. and Holloway, J. L.: 1975, 'The seasonal variation of the hydrologic cycle as simulated by a global model of the atmosphere', *J. Geophys. Res.* **80**, 1617–1649.
- Manabe, S. and Strickler, R.: 1964, 'Thermal equilibrium of the atmosphere with a convective adjustment', *J. Atmos. Sci.* **21** (4), 361–385.
- Manabe, S., Bryan, K., and Spelman, M. J.: 1975, 'A global ocean-atmosphere climate model: Part I. The atmospheric circulation', *J. Phys. Oceanogr.* **5** (1), 3–29.
- Manabe, S., Bryan, K., and Spelman, M. J.: 1979, 'A global ocean-atmosphere climate model with seasonal variation for future studies of climate sensitivity', *Dynamics of Atmos. and Ocean* **3**, 393–420.
- Manabe, S., Smagorinsky, J., and Strickler, R. F.: 1965, 'Simulated climatology of a general circulation model with a hydrologic cycle', *Mon. Wea. Rev.* **93** (12), 769–798.
- Marchuk, G. I. and Zalesnyi: 1974, 'A numerical model of the large-scale circulation in the World Ocean', in *Numerical Methods of Calculating Ocean Currents* [in Russian], Computation Center of the Siberian Division of the USSR Academy of Sciences, Novosibirsk, pp. 3–20.
- Maunder, E. W.: 1890, *M. N.* **50**, No. 6.
- Miropol'ski, Yu. Z.: 1981, *The Dynamics of Internal Gravity Waves in the Ocean* [in Russian], Gidrometeoizdat, Leningrad.
- Moffatt, H. K.: 1969, 'The degree of knottedness of tangled vortex lines', *J. Fluid Mech.* **35**, 117–129.
- Moffatt, H. K.: 1978, *Magnetic Field Generation in Electrically Conducting Fluids*, Cambridge Univ. Press.
- Monin, A. S.: 1958, 'On precalculation of the zonal atmospheric circulation', *Izv. Akad. Nauk SSSR, Ser. Geofiz.* **19**, 1250–1253.
- Monin, A. S.: 1965, 'On the structure of the atmospheric boundary layer', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **1** (3), 258–265.
- Monin, A. S.: 1969, *Weather Forecasting as a Problem in Physics* [in Russian], Nauka, Moscow.
- Monin, A. S.: 1970, 'On turbulent mass flows in the ocean', *Dokl. Akad. Nauk SSSR* **190** (2), 240–248.
- Monin, A. S.: 1972, *Earth's Rotation and Climate* [in Russian], Gidrometeoizdat, Leningrad [English Translation, 1976. Radok-Radhakrishna, Delhi, India].
- Monin, A. S.: 1977, *The History of the Earth* [in Russian], Nauka, Leningrad.
- Monin, A. S.: 1980, *The Solar Cycle* [in Russian], Gidrometeoizdat, Leningrad.
- Monin, A. S.: 1986, *An Introduction to the Theory of Climate*, Reidel (Kluwer), Dordrecht.
- Monin, A. S. and Krasitskii, V. P.: 1985, *Phenomena on the Ocean Surface* [in Russian], Gidrometeoizdat, Leningrad.

- Monin, A. S. and Ozmidov, R. V.: 1981, *Ocean Turbulence* [in Russian], Gidrometeoizdat, Leningrad.
- Monin, A. S. and Shishkov, Yu. A.: 1979, *The History of Climate* [in Russian], Gidrometeoizdat, Leningrad.
- Monin, A. S. and Simuni, L. M.: 1980, 'A model of the differential rotation of the Sun', *Dokl. Akad. Nauk SSSR* **250** (6), 1348–1351.
- Monin, A. S. and Sorokhtin, O. G.: 1977, *Dokl. Akad. Nauk SSSR* **234** (2), 413–416.
- Monin, A. S. and Yaglom, A. M.: 1965 (Part 1) and 1967 (Part 2), *Statistical Fluid Dynamics* [in Russian], Nauka, Moscow [English Translation, 1971, MIT Press].
- Monin, A. S., Kamenkovich, V. M., and Kort, V. G.: 1974, *The Variability of the World Ocean* [in Russian], Gidrometeoizdat, Leningrad.
- Monin, A. S., Seidov, D. G., Sorokhtin, O. G., and Sorokhtin, Yu. O.: 1987, *Dokl. Akad. Nauk SSSR* **294** (1), 58–63.
- Munk, W. H.: 1950, 'On the wind-driven ocean circulation', *J. Meteorol.* **7**, 79–83.
- Needler, G. T.: 1971, 'Thermocline models with arbitrary barotropic flow', *Deep-Sea Res.* **18** (9), 895–904.
- North, G. R.: 1975, *Numerical Models of Ocean Circulation*, Nation. Acad. Sci., Wash. No. 20418.
- Obukhov, A. M.: 1962, 'On the dynamics of a stratified fluid', *Dokl. Akad. Nauk SSSR* **145** (6), 1239–1242.
- Oceanology* — *Physics of the Ocean*, 1978, Vol. 1, *Hydrophysics of the Ocean*, Vol. 2, *Hydrodynamics of the Ocean* [in Russian], Nauka, Moscow.
- Oceanology* — *Geophysics of the Ocean*, 1979, Vol. 2, *Geodynamics* [in Russian], Nauka, Moscow.
- Ogura, Y.: 1952, 'The theory of turbulent diffusion in the atmosphere', *J. Meteor. Soc. Japan* **30** (2), 53–58.
- Oort, A. H.: 1964a, 'On estimates of the atmospheric energy cycle', *Mon. Wea. Rev.* **92** (11), 483–493.
- Oort, A. H.: 1964b, 'On the energetics of the mean and eddy circulation in the lower stratosphere', *Tellus* **16** (4).
- Oort, A. H. and Peixoto, J. P.: 1974, 'The annual cycle of the energetics of the atmosphere on a planetary scale', *J. Geophys. Res.* **79** (18), 2705–2719.
- Oort, A. H. and Rasmusson, E. M.: 1971, 'Atmospheric circulation statistics', *NOAA Prof. Paper* **5**.
- Parker, E. N.: 1955, 'Hydrodynamic dynamo models', *Astrophys. J.* **122**, 293–314.
- Pedlosky, J.: 1965, 'A study of the time-dependent ocean circulation', *J. Atmos. Sci.* **22**, 267–272.
- Pedlosky, J.: 1979, *Geophysical Fluid Dynamics*, Springer, New York/Heidelberg/Berlin.
- Phillips, N. A.: 1956, 'The general circulation of the atmosphere: a numerical experiment', *Quart. J. Roy. Met. Soc.* **82** (358), 123–164.
- Pomeau, Y. and Manneville, P.: 1980, 'Intermittent transition to turbulence in dissipative dynamical systems', *Commun. Math. Phys.* **74**, 189–197.
- Rhines, P.: 1970, 'Edge- bottom- and Rossby waves in a rotating stratified fluid', *Geophys. Fluid Dyn.* **1**, 273–302.
- Rhines, P.: 1975, 'Waves and turbulence on a beta-plane', *J. Fluid Mech.* **69** (3), 417–444.
- Roberts, P. H.: 1972, 'Kinematic dynamo models', *Phil. Trans. Roy. Soc.* **A272**, 663–698.
- Rossby, C.-G.: 1939, 'Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacement of the semipermanent centers of action', *J. Mar. Res.* **2**, 38–55.
- Rossby, C.-G.: 1940, 'Planetary flow patterns in the atmosphere', *Quart. J. Roy. Met. Soc.*, Suppl. **66**, 68–87.
- Rossby, C.-G.: 1941, 'The scientific basis of modern meteorology', *Climate and Man. Yearbook of Agriculture*, US Govt. Printing Office, Washington, D.C., pp. 599–655.
- Rossby, C.-G.: 1947, 'On the distribution of angular velocity in gaseous envelopes under the influence of large-scale horizontal mixing processes', *Bull. Amer. Meteor. Soc.* **28**, 53–68.
- Ruelle, D. and Takens, F.: 1971, 'On the nature of turbulence', *Commun. Math. Phys.* **20**, 167–192 and **23**, 343–344.
- Seidov, D. G.: 1978, 'A numerical scheme for the study of synoptic eddies in the ocean', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **14** (7), 757–767.
- Seidov, D. G.: 1980, 'Synoptic eddies in the ocean. A numerical experiment', *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana* **16** (1), 73–87.

- Sellers, W. D.: 1969, 'A global climatic model based on the energy balance of the Earth-atmosphere system', *J. Appl. Met.* **8** (3), 392–400.
- Sergin, V. Ya. and Sergin, S. Ya.: 1978, *A Systematic Analysis of the Relation between Large Climatic Oscillations and the Earth's Glaciation* [in Russian], Gidrometeoizdat, Leningrad.
- Shtokman, V. B.: 1946, 'The vertical propagation of heat waves in the sea and indirect methods for determining the thermal conductivity', *Trudy Inst. Okeanolog. Akad. Nauk SSSR* **1**, 3–46.
- Steenback, M. F., Krause, F., and Rädler, K.-H.: 1966, 'A calculation of the mean electromotive force in an electrically conducting fluid in turbulent motion, under the influence of Coriolis forces', *Z. Naturforsch.* **21A**, 369–376.
- Stommel, H.: 1948, 'The westward intensification of wind-driven ocean currents', *Trans. Amer. Geophys. Union* **29** (2), 202–206.
- Sverdrup, H. U.: 1947, 'Wind-driven currents in a baroclinic ocean, with application to the equatorial currents of the eastern Pacific', *Proc. Nat. Acad. Sci. USA* **33** (11), 318–326.
- Taylor, G. I.: 1923, 'Experiments on the motion of solid bodies in rotating fluids', *Proc. Roy. Soc. A* **104**, 213–218.
- Vainshtein, S. I. and Zeldovich, Ya. B.: 1972, 'Origin of magnetic fields in astrophysics', *Usp. Fiz. (SSSR)* **106**, 431–457 (Engl. trans. *Sov. Phys. Usp.* **15**, 159–172).
- Verbitskii, M. Ya.: 1981, 'Numerical modeling of the evolution of the cover glaciation', *Dokl. Akad. Nauk SSSR* **256** (6), 1333–1337.
- Verbitskii, M. Ya. and Chalikov, D. V.: 1980, 'A thermodynamic model of a large ice sheet', *Dokl. Akad. Nauk SSSR* **250** (1), 59.
- Washington, W. M., Semtner, A. J., Meehl, G. A., Knight, D. J., and Mayer, T. A.: 1980, 'A general circulation experiment with a coupled atmosphere, ocean, and sea ice model', *J. Phys. Oceanogr.* **10** (12), 1887–1908.
- Welander, P.: 1959, 'On the vertically integrated mass transport in the oceans', in *The Atmosphere and the Sea in Motion*, B. Bolin (ed.), Rockefeller Inst. Press/Oxford Univ. Press, pp. 95–101.
- Williams, G. P. and Davis, D. R.: 1965, 'A mean-motion model of the general circulation', *Quart. J. Roy. Met. Soc.* **91** (390), 471–489.
- Zharkov, V. N. and Trubitsyn, V. P.: 1980, *Physics of the Planetary Interiors* [in Russian], Fizmatgiz, Moscow.
- Zharkov, V. N., Trubitsyn, V. P., and Samsonenko, L. V.: 1971, *Physics of the Earth and Planets* [in Russian], Fizmatgiz, Moscow.
- Zilinkevich, S. S., Monin, A. S., Turikov, V. G., and Chalikov, D. V.: 1976, 'Numerical modeling of the combined atmospheric and oceanic circulation', *Dokl. Akad. Nauk SSSR* **230** (3), 556–559.
- Zotikov, I. A.: 1977, *The Thermal Regime of the Antarctic Ice Sheet* [in Russian], Gidrometeoizdat, Leningrad.

Index

- Acoustic waves 50, 64
Active layer 312–314
Adaptation of fields 36, 51–55
 convective 295–296
Adiabatic approximation 165
Adiabatic process 7, 33ff, 50
Airy functions 173–174, 178, 199
Anisotropization 268
Antarctica 317, 328–330, 338–340
Antidynamo 351–353
Asthenosphere 339–341
Atmosphere, height of 9, 38
 models of 15
Attractors 81, 99–100, 114

Baric tendency 55
Baroclinic instability 73–81, 269–272, 294
Baroclinic medium 4, 11, 231
Baroclinic waves 39, 53, 64, 237–239, 259, 268
Barotropic instability 73
Barotropic medium 4, 11, 55, 230, 254
Barotropic waves 39, 53, 64, 235–238, 247, 257–264
Bénard cells 111
Bernoulli equation 119, 137–139
Bifurcations 81–85, 113–115
Blini 218–219
Boundary layers 222–228
Boussinesq approximation 11, 37, 41, 65, 165, 182, 188, 384
Brunt-Väisälä frequency 8, 33, 36, 172, 193, 220
Buoyancy parameter 209–211, 216
Buoyancy scale 216

Cantor sets 98
Capillary waves 119ff
Centers of action 280
Circulation index 240
Climate 312ff
Climate models 315ff
Cnoidal wave 142–144, 183

Concentration convection 381–385
Convective interval 216
Convective rollers 111–113, 116, 285
Coriolis force 10, 50–51, 199, 204, 209, 222, 237, 279–280
Coriolis parameter 15, 43, 73, 165, 203, 237
Costras 382
Couette flow 88–89, 92, 108–109
Crossover phenomenon 29–31

Darcy's law 323
Departing wave 39
Double diffusion 222
Drag crisis 93

Earth, interior of 342ff, 363–372
Eddy-resolving models 281, 306–311
Eddy waves 166
Eigencurves 56–64
Eigenfrequencies 37, 56, 167
Eikonal 124
Ekman layers 193, 203, 205, 222–224, 228, 296, 302
Ekman number 301–304
Enstrophy 31–32, 231–236, 266–267
Equivalent depths 238
Ergodicity 95, 180
Evolution equations 10, 206

Fjørtoft condition 79, 81
Flow visualization 90–93
FPU recovery 148, 154
Friction velocity 162, 203, 224
Fridman-Keller equations 204, 296
Frozen magnetic field 12, 14

Galerkin approximation 23–24, 68, 203–204
Garrett-Munk model 197–198
General circulation 279ff
Geoid 7
Geomagnetic field 342–348
Geophysical fluid dynamics, defined xii

- Geostrophic motions 35, 175, 269
 Geostrophic turbulence 228–236
 Geostrophic wind 44, 73–75
 Glacial periods 334–341
 Gravity waves 38–40, 119ff, 165–172, 199
 Gyres 301, 307–310, 336
 Gyroscopic rigidity 54
 Gyroscopic waves 37, 39, 64, 199–200
- Hamilton equations 155–160, 188–194
 Hamilton system 31–32, 261–269
 Helicity 348–353
 Helmholtz equation 24–25
 Helmholtz instability 70
 Hydraulic approximation 323
 Hydrodynamic dynamos 348–361
 Hydrodynamic noise 40
 Hydrodynamic systems 22–32
 Hydrostatic approximation 44, 50
 Hydrostatic equation 8–9, 44, 290, 335
 Hydrostatic equilibrium 51, 362
- Ice ages 334–341
 Ice sheets 332–341
 Induction equation 11
 Inertial oscillations 199–201
 Insolation 329
 Instability 65ff
 Internal waves 49–50, 64, 165–201
 Irregular waves 131
 Isentropic process 7
 Isothermal atmosphere 43–51
- Jet streams 280–282
 Jupiter 236, 241, 282ff
- Kármán constant 210
 Kármán street 93
 KdV equation 131, 143–147, 184–185
 Kibel' number 54, 73, 305
 Kinetic equation 157
 Kolmogorov spectra 158–160, 193, 216
- Labile energy 21
 Lagrangian description 128
 Lagrangian invariants 6, 7, 14, 17, 246–248
 Laplace tidal equation 55–64, 155, 240
 Legendre equation 61
 Lin-Vasov theorem 86–87
 Linear invariants 34, 43–44
 Linear theory 119–123, 165–172, 237–241
 Lorentz force 10, 14, 26, 354
 Lorenz flow 112–116
 Lyapunov indices 99, 105, 116
 Lyapunov stability 68–71
- Mach rotation number 282, 288
 Macrostability parameter 22
 Magnetostrophic waves 349
 Manifolds 82
 Mantle convection 380–389
 Mars 279ff
 Maunder “butterflies” 355–357
 Microstructure 185–188
 Mixing length 205
 Monodromy operator 84
 Moon, interior of 372–374
- Ocean
 currents 334–338
 depth 56
 Orr-Sommerfeld equation 85–89, 92
- Parametric resonance 29
 Phase flow 68, 81–85
 Phase space 68
 Planets 361ff. *See also* individual planets
 Poiseuille flow 89–93
 Polytypic processes 50
 Potential approximation 119
 Potential density 7, 180–184, 203
 Potential temperature 15, 337
 Potential vorticity 13–17, 44, 48, 76, 229–230
 Potential waves 130–138
 Pseudoentropy 16
 Pycnocline 168–169, 184, 193, 219
- Quasigeostrophic approximation 54, 73–76, 230, 271, 295–297, 307, 360
 Quasilinear theory 161–164
- Radiation equilibrium 319
 Radiation stresses 125–127
 Ray theory 123–127, 172–180, 239
 Rayleigh condition 79–81
 Rayleigh equation 67–69, 72–73, 79, 86–92, 161
 Refraction velocity 172
 Reversal point 174–180
 Reynolds equations 204–205, 228, 352
 Reynolds number 81
 critical 83–85
 Reynolds stress 171, 204, 352
 Richardson number 72, 169–171, 203–210, 225
 Rollers 111–116
 Rossby deformation radius 15, 46, 230
 Rossby waves 41–43, 64, 203, 230, 235ff, 280

- Salt fingers 222
- Scattering data 145
- Self-focusing 184
- Semiempirical theories 205–207, 217, 296
- SHU model 205
- Solar constant 290
- Solitons 143–155, 236, 251–261
- Spectra 194–198
- Spectral densities 214
- Stability changes 93, 109–111
- Stationary waves 128–155, 180–185, 246–261
- Stochasticity 93–100
 - scenarios 100–108
- Stokes lines 88–90
- Stratification 7–13, 65, 199–218
- Structure functions 208–209
- Strange attractors 81, 95, 99–101
- Sturm-Liouville problem 167
- Sunspots 354–357
- Superreflection 171
- Surface tension 20
- Surface waves 38
- Symmetrized systems 31–32
- Synoptic eddies 280, 296, 307–311

- Taylor flow 82, 108–116
- Tendency equation 327
- Thermal convection 82
- Thermobaric energy 34

- Thermocline 336
- Thermodynamically active admixtures 3, 7, 13
- Thermodynamic potential 3
- Thermodynamic process 6
- Tidal forces 10, 37, 41–42, 194, 313
- Topographic waves 242–246, 269–271
- Trade winds 279
- Trapped modes 245–246
- T-S curves 8–9
- Tsunamis 193
- Turbulence 95, 202ff
 - geostrophic 228–236
 - scale 206

- Venus 241, 279ff
- Viscosity, disappearing 86
 - negative 268, 280–285, 306–310, 332
- Vorticity, potential 13–17, 44, 48, 76, 229–230

- Wave action 157–160, 173, 176
- Wave breaking 159, 162
- Wave channel 176
- Wave train 172
- Wave traps 178–180
- Weak nonlinearity 156, 190, 193
- Weather forecasting 241
- Wind waves 160–164
- WKB method 174–178
- Wolf number 355

ENVIRONMENTAL FLUID MECHANICS

Managing Editor:

G.T. Csanady, *Woods Hole Oceanographic Institution, Woods Hole, Massachusetts*

Editorial Board:

A.J. Davenport, *University of Western Ontario, London, Ontario*

B.B. Hicks, *Atmospheric Turbulence and Diffusion Laboratory, Oak Ridge, Tennessee*

G.R. Hilst, *Electric Power Research Institute, Palo Alto, California*

R.E. Munn, *University of Toronto, Ontario*

J.D. Smith, *University of Washington, Seattle, Washington*

Publications:

1. W. Brutsaert: *Evaporation into the Atmosphere*. Theory, History, and Applications. 1982; Reprint 1988 ISBN 90-277-1247-6
2. G.T. Csanady: *Circulation in the Coastal Ocean*. 1982 ISBN 90-277-1400-2
3. A.S. Monin and R.V. Ozmidov: *Turbulence in the Ocean*. 1985 ISBN 90-277-1735-4
4. S. Panchev: *Dynamic Meteorology*. 1985 ISBN 90-277-1744-3
5. V.M. Kamenkovich, M.N. Koshlyakov and A.S. Monin (eds.): *Synoptic Eddies in the Ocean*. 1986 ISBN 90-277-1925-X
6. A.S. Monin: *Theoretical Geophysical Fluid Dynamics*. 1990 ISBN 0-7923-0426-8