

REVIEWS OF TOPICAL PROBLEMS PACS numbers: **05.45.+b, 47.52.+j**

Short-wavelength instability and transition to chaos in distributed systems with additional symmetry

M I Tribel'skiĭ†

Contents

Abstract. A review is presented of the theory of pattern formation in extended, dissipative, highly nonequilibrium systems. Emphasis is placed on systems which, in addition to spatial translations and rotations, have additional continuous symmetry group(s). It is shown that the additional symmetry destabilizes dramatically the ground state of the system thus causing it to make a direct transition from a spatially uniform to a turbulent state in an analogous fashion to the second-order phase transition in quasi-equilibrium systems. Apart from the theoretical analysis, a discussion of experimental data is given.

To the respectful memory of I M Lifshitz.

1. Introduction

I regard it as a great honour to have my paper published in the special issue of *Physics Uspekhi* dedicated to I M Lifshitz on the memorial of his 80th birthday. It is a great responsibility too, bearing in mind the problems with which the present review is concerned. In fact, the field the review belongs to, namely self-organization and transition to chaos, attracted much of I M Lifshitz' attention during the last years of his life. I personally believe that in spite of the fact that he has not published any paper devoted to the subject, this was mainly due to the fact that he had been collecting information in this field that was new to him. This preliminary work would most certainly have been crowned with an article (or a series of papers) laying down a programme of further progress, as had been the case of his studies in polymer physics [1]. It would have been... but it did not take place — this hope was dashed by the premature death of I M Lifshitz.

The objective of the present review is to draw attention of research scientists to selected results of recent studies on pattern formation and transition to chaos in dissipative systems. These results in many respects digress from traditional concepts in this field indicating that the symmetry of the problem is of greater importance than it was believed to be till now.

The review is addressed to a broad audience of both specialists and non-experts. For this reason I shall try to avoid, whenever appropriate, cumbersome mathematical calculations, which are normally inseparable from the theoretical description of the problems and refer the readers to original papers for details. This is supposed to allow to focus on the key aspects of the analysis with special reference to specific qualitative features of the phenomena of interest.

Terminology relevant to the problems discussed is this review has not settled yet, and some authors read their own thoughts into the terms which actually have quite different sense. To avoid misunderstanding, I shall start with explaining the meaning of special terms which the reader will encounter below.

Distributed systems are any objects or phenomena described by equations in partial derivatives (mainly, nonlinear ones). The term *dissipative* implies that the considered systems are non-equilibrium and certain dissipative processes occur. In the case of open systems (where energy is fed from the outside to be further dissipated and eventually removed from the system or released 'to infinity'), 'the energy flux through the system' is supposed to remain constant. This

† The author is also known as M I Tribelsky. The name used here is a transliteration under the BSI/ANSI scheme adopted by this journal.

M I Tribel'skiĭ Graduate School of Mathematical Science,
University of Tokyo
3-8-1 Komaba, Mediro-ku Tokyo 153, Japan
E-mail: tribel@ms.u-tokyo.ac.jp

Received 22 October 1996
Uspekhi Fizicheskikh Nauk **167** (2) 167 – 190 (1997)
Translated by Yu V Morozov; edited by A I Yaremchuk

creates necessary prerequisites for the formation of stationary non-equilibrium states in the system.

The problem of pattern formation and transition to chaos in such systems is very broad and includes a wealth of specific issues pertaining to hydrodynamic flows, the combustion theory, phase transition kinetics, the theory of reaction-diffusion systems of the Belousov – Zhabotinsky type, biology, and even sociology and psychology (see, for instance, Refs [2 – 4] and references therein).

Of the whole variety of issues, I have chosen to consider systems with short-wavelength instability of spatially uniform states and an additional continuous group of symmetry. These terms also need to be specified.

To be concrete let us consider, by way of example, convection in a horizontal fluid layer placed in the field of gravity forces known as the Rayleigh – Benard problem. In the simplest realization of this problem, both lower and upper surfaces of the layer are kept at constant temperatures (T_1 and T_2, respectively, with $T_1 > T_2$), that is the layer is heated from below and cooled from above. Evidently, such a system is open because energy (heat flux) comes in through the lower surface and goes out through the upper one.

It is well-known (see, for instance, Ref. [5]) that as long as $\Delta T \equiv T_1 - T_2$ is below a certain critical value (*convection instability threshold* or, simply, *convection threshold*), the stationary state of a fluid in the absence of any hydrodynamic motion is stable against (infinitely) small perturbations. Hydrodynamic velocity of the fluid in such a stationary state identically vanishes, and its temperature and pressure depend only on one spatial coordinate, z, where z-axis is vertical. Thus, the state of the fluid at the layer plane (plane xy) is *spatially uniform*. Such spatially uniform states of a dissipative system are henceforth termed trivial.

However, the heated fluid at the bottom of the layer tends upwards whereas the cool one atop is always ready to sink. Therefore, the increasing difference between temperatures of the lower and upper surfaces of the layer must eventually result in a hydrodynamic flow, that is in a loss of stability of the stationary spatially uniform state.

It is essential that the continuity equation does not allow such a flow to be realized in the spatially uniform manner: the bottommost layer can not rise as a whole since there would be no room for the fluid above it to go. For this reason, any convective flow inevitably results in a loss of spatial uniformity in the horizontal plane. This, in turn, creates a certain spatial scale at this plane. On the other hand, the thickness of the layer, d, is the only characteristic spatial scale of the problem and is therefore responsible for the characteristic horizontal scale of the flow.

In other words, exceeding convection threshold leads to a *spontaneous symmetry breaking* of the system's state. Mathematically it is expressed as instability of the spatially uniform state to spatially non-uniform perturbations described by multipliers of the form

$$\exp(\gamma t + \mathrm{i}\mathbf{k} \cdot \mathbf{r}) \,, \qquad (1)$$

where $\mathbf{k}$ is the two-dimensional wave vector of the perturbation which lies in the xy plane and $\mathbf{r} = (x, y)$ is the corresponding two-dimensional radius-vector.

The dispersion function $\gamma(\mathbf{k})$ in the Rayleigh – Benard problem is purely real ($\mathrm{Im}\,\gamma \equiv 0$ is *aperiodic instability*). It follows from the previous discussion that $\max\{\gamma(k)\} \equiv \gamma(k_c) = 0$ at the instability threshold, and the critical wave number k_c satisfies the relation $k_c d \sim 2\pi$.

Therefore, the above account of the processes, leading to the spontaneous creation of the characteristic scale of the problem at the horizontal plane, arises from the consideration of instabilities in systems with *constrains* (the upper surface that constrains fluid's motion) and *conservation laws* (of mass, in this case) [4].

In view of the importance of this problem, it is worth to examine it from a different standpoint. Recall that the problem is isotropic in the horizontal plane, and, therefore, γ can not depend on the orientation of $\mathbf{k}$ in this plane; actually, $\gamma(\mathbf{k})$ means $\gamma(k)$, and perturbations (1) may be substituted by those of the form

$$\exp(\gamma t + \mathrm{i}kx) \qquad (2)$$

without loss of generality.

Note further that the problem contains *two* physically different mechanisms of dissipation, i.e., viscosity and heat conduction, whose relative importance varies with changing k. Let us consider two limiting cases, namely long-wavelength ($k \to 0$) and short-wavelength ($k \to \infty$) perturbations. Trajectories of a small element of fluid volume (fluid particle) in case of long-wavelength and short-wavelength perturbations are shown in Figs 1a and 1b, respectively.

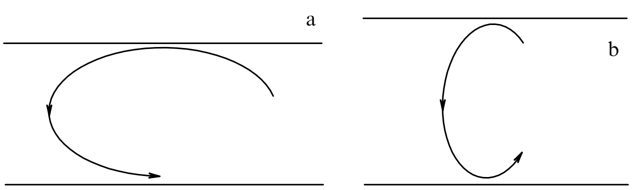


Figure 1. Trajectories of a 'fluid particle' in case of long-wavelength (a) and short-wavelength (b) perturbations of a spatially uniform state. Generally speaking, the trajectories are not closed since the flow is not stationary.

In the case of long-wavelength perturbations, the large characteristic size of the trajectory accounts for small values of the horizontal projection of the temperature gradient associated with the break of spatial uniformity (i.e., with perturbation) because the magnitude of this projection is inversely proportional to the horizontal size of the trajectory (the upward and downward flows correspond to motions of 'hot' and 'cold' fluids, respectively). Therefore for such perturbations thermoconductive dissipation is suppressed whereas viscous dissipation is pronounced due to the large trajectory length; it grows further with decreasing k. For this reason, very long-wavelength perturbations are 'energetically disadvantageous' and can not lead to instability.

On the other hand, for very short-wavelength perturbations (Fig. 1b) viscous dissipation is relatively weak while thermoconductive dissipation is strong due to the large temperature gradient arising from the small horizontal scale of the trajectory. Hence, such perturbations are also disadvantageous.

Competition between these two mechanisms is responsible for selection of the most 'dangerous' (i.e., leading to instability) perturbations whose wavenumbers are neither too large nor too small but have a certain *finite* value.

However, it should be borne in mind that such an explanation for arising of characteristic spatial scale in the problem is actually based on the minimum entropy production principle [2], which holds exclusively for steady states and only under the condition that deviations from equilibrium are small. The application of this principle to strongly non-equilibrium and (or) non-stationary problems may lead sometimes to erroneous and even absurd conclusions.

Thus, the appearance of a spatial scale in the Rayleigh – Benard problem is related to instability against perturbations with certain finite wavenumbers. Such instability will be referred to as *short-wavelength* in contrast to *long-wavelength* instability, where the wave number region corresponding to unstable perturbations begins at $k = 0$.

Formalizing the definition, it may be said that at the short-wavelength instability threshold $\gamma(k) = 0$ (in a more general case of a complex spectrum, $\mathrm{Re}\,\gamma(k) = 0$) at a certain finite $k = k_c$. Moreover, $\gamma(k) < 0$ ($\mathrm{Re}\,\gamma(k) < 0$) everywhere in the vicinity of k_c. When short-wavelength instability threshold is slightly exceeded, the *instability zone* $k_1 < k < k_2$ with $\gamma(k) > 0$ ($\mathrm{Re}\,\gamma(k) > 0$) appears, where $k_{1,2}$ are finite quantities depending on the magnitude of the excess (see Fig. 2).

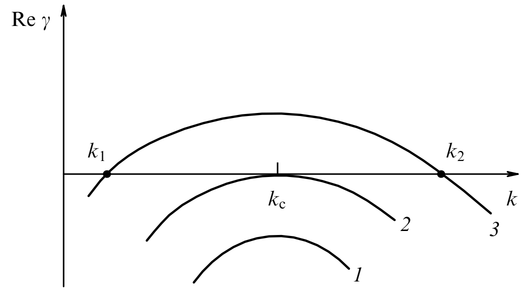


Figure 2. The real part of a dispersion dependence $\gamma(k)$ in case of a short-wavelength instability. Cases *1*, *2*, and *3* correspond to different values of the control parameter: *1* — slightly below the instability threshold, *2* — right at the threshold, *3* — slightly beyond the threshold. k_c is the critical wave number, k_1 and k_2 are the boundaries of the instability zone.

I introduce here the concept of an *instability zone* because further discussion will concern systems infinitely extended in the direction of spatial uniformity of a highly symmetric state, with k being a continuous quantity.

Now let us consider the problem's symmetry which is of paramount importance in the context of this review.

It should be emphasized that the problem's symmetry is the symmetry of equations and complementary boundary conditions of the corresponding boundary-value problem rather than the symmetry of its solution which is quite a different thing. The fact is that the symmetry of all non-trivial solutions discussed in the present review is lower than the initial symmetry of the problem (spontaneous symmetry breaking previously mentioned with reference to the Rayleigh – Benard problem).

Usually, problems of infinite spatial extent in a certain direction are translationally invariant in the same direction, i.e., the corresponding boundary problem remains unaltered by transformations of coordinates like

$$\mathbf{r} \to \mathbf{r} + \mathbf{a}\,. \tag{3}$$

Here $\mathbf{r}$ is a radius-vector (one- or two-dimensional, depending on how the problem is formulated), and $\mathbf{a}$ is an arbitrary constant vector of the same spatial dimension.

In the case of a two-dimensional vector $\mathbf{r}$, we additionally assume the *isotropy* of the problem in the corresponding plane, i.e., *rotational invariance* with respect to an arbitrary axis perpendicular to the plane of spatial uniformity.

Moreover, 'left' and 'right' are considered to be equivalent (left-right parity), i.e., the problem is invariant with respect to a change of the sign of vector $\mathbf{r}$ components (*spatial reflections*).

Finally, the problem normally contains time only in the form of derivatives which accounts for invariance with respect to the time shift

$$t \to t + \mathrm{const}\,. \tag{4}$$

The whole set of the above symmetry transformations will be referred to as the *conventional* symmetry. In certain cases, however, the problem may have an additional continuous one-parametric symmetry group. This implies the existence of a group of transformations which differ from those described in previous paragraphs and cause no change in the examined boundary-value problem. Each transformation in this group is unambiguously fixed by a particular value of one scalar parameter (*uniparametricity*), which may vary continuously in a certain segment, either finite or infinite (*continuity*). It is important for the forthcoming discussion that the segment includes the zero value of the scalar parameter.

Here are several examples of such problems. In the first place, it is worth mentioning different formulations of the problem of motion of phase boundaries during first-order phase transitions and of chemical reaction fronts (e.g., combustion). In the case of steady motion of a plane boundary (front), the problem is translationally invariant in the direction of motion, unless the latter is governed by external factors. Since the translation occurs in the direction *normal* to the boundary plane (i.e., in the direction of *broken* spatial uniformity), such a transformation can not be reduced to (3) and is, actually, an additional symmetry. Problems of this type exhibiting short-wavelength instability of plane boundaries (fronts) have been considered in Refs [6, 7] (only those references are selected which appear to be most pertinent to the details of the problems discussed below).

Another example involves problems with Galilean invariance, they undergo no change upon the transformation

$$v \to v + v_0\,, \qquad \frac{\partial}{\partial t} \to \frac{\partial}{\partial t} + v_0\,\frac{\partial}{\partial x}\,, \tag{5}$$

see Refs [4, 8]. Here v is velocity and v_0 is an arbitrary constant. It seems appropriate to remind that the discussion concerns an additional one-parameter group of symmetry; hence, the system is Galilean invariant with respect to only one component of the velocity vector.

Another important and theoretically most thoroughly developed problem is the *Rayleigh – Benard convection with free-slip or no-stress boundary conditions* [4, 9 – 15]. This issue is worth discussing at greater length.

When a layer of a fluid is enclosed between solid surfaces, it follows from the natural boundary conditions that the velocity of the fluid vanishes at these surfaces: $v = 0$ (*rigid* or *no-slip* boundary conditions). Nevertheless, the majority of theoretical works on the Rayleigh – Benard problem starting from those by Lord Rayleigh himself [9, 10] proceed from the concept of free-slip boundaries, $v_z = 0$, $\partial v_{x,y}/\partial z = 0$. Indeed, these boundary conditions are convenient to use by virtue of the form of z-dependence of eigenfunctions for the stability problem of the trivial state ($\mathbf{v} \equiv 0$). In the case of free-slip boundaries, these eigenfunctions are reduced to ordinary trigonometric functions, which simplifies calculations substantially.

It is worthwhile to note that the boundary conditions $v_z = 0$, $\partial v_{x,y}/\partial z = 0$ admit of persistent motion of a fluid along such boundaries. Really, if the fluid moves as a solid (i.e., the vertical component of its velocity is identically zero, the horizontal components show no dependence on z, and shear strains are absent), the viscous stress tensor corresponding to such a flow vanishes identically and the flow becomes non-dissipative. Evidently, two types of such flow are conceivable. One is translational motion with constant horizontal velocity, the other is rotation of the liquid layer about the vertical axis at a constant angular speed. The possibility to maintain a flow with constant horizontal velocity leads to Galilean invariance of the problem discussed in previous sections. It should be borne in mind, however, that, notwithstanding the assumption of the infinite extension of the fluid layer at the horizontal plane, such a formulation of the problem is, actually, an idealization, and there are always sidewalls under the real conditions which preclude realization of a flow with non-zero mass transfer to 'infinity.' Therefore, Galilean invariance in the convective layer is normally suppressed even though it may be restored in some special cases.

However, if the translational motion of the liquid layer is forbidden, there is nothing to hamper its rotation, which does not lead to mass transfer to infinity: invariance of the problem with respect to such rotations (analogs of Galilean invariance for rotational motion) is the additional symmetry of interest. It should be emphasized once again, to avoid misunderstanding, that invariance in question is the one with respect to rotations with an arbitrary fixed angular velocity contrary to the conventional invariance under rotations through an arbitrary angle.

It is worthy of note that despite experimental realization of the free-slip convection is a difficult problem, the case is not altogether hopeless (see, for instance, Ref. [16] and the discussion of this question in Ref. [4]).

An analog of the Rayleigh – Benard problem for anisotropic fluids (liquid crystals) is electroconvection in thermotropic nematics [17 – 20]. Molecules constituting such a substance have a shape stretched in a certain direction. In the nematic phase, while the position of the center of mass of each molecule is not fixed, the molecules are aligned parallel to one another. Thus, the corresponding substance being a fluid becomes anisotropic. In the electroconvection an AC-voltage V is applied to two parallel (usually transparent†) electrodes confining a nematic layer whose typical width is about 10 – 100 μm. Under conditions described in the above references, a rise in V over a certain threshold value V_c results in the convective motion analogous to that observed in the Rayleigh – Benard convection. In this case, unlike that of thermoconvection, the system normally becomes thermostatic, with the squared voltage playing the role of temperature difference between the upper and lower boundaries of the layer and its frequency corresponding to the Prandtl number‡ in thermoconvection.

Where does the additional symmetry come from in this problem? The thing is that rigid surfaces bounding the nematic layer are normally treated in such a way as to impose the desired orientation of liquid crystal molecules in contact with them. Let us consider the case of homeotropic orientation, where nematic molecules are aligned perpendicular to the surfaces. In the absence of an external electric field, the orientation of the molecules by the surfaces determines molecular orientation throughout the layer.

What effect may an electric field have on such oriented layer? It will be significantly dependent on the sign of dielectric anisotropy of the nematic. I am interested in *negative* anisotropy when molecules of a liquid crystal tend to turn perpendicular to the field lines, that is their orientation by the field is in conflict with the orientation imposed by the boundary conditions.

It is well-known that such a situation is characterized by a certain threshold voltage V_F. At $V < V_F$ the uniform homeotropic distribution of molecules remains constant. At $V > V_F$ the *Freedericksz transition* occurs, and the equilibrium orientation of the molecules becomes oblique relative to the layer plane everywhere except a narrow subsurface layer adjacent to the boundary surfaces [17 – 19].

Therefore, at $V > V_F$ molecule axes have the a non-zero projection on the layer plane, which yields anisotropy in this plane. At the same time, all external factors that affect the system, including the electrical field responsible for this anisotropy, remain spatially uniform and isotropic at the plane of the layer. For this reason the anisotropy axis may randomly take any orientation.

Certainly, all states different only in the direction of the anisotropy axis are physically equivalent. In other words, the system undergoes a spontaneous break of isotropy in the plane of the layer followed by degeneracy with respect to rotations through an arbitrary angle about the axis perpendicular to the layer plane [21].

In typical cases, Refs [22 – 24], the threshold of the Freedericksz transition is lower than that of convection instability. Therefore, the system's state at the threshold of convection is degenerated with respect to the above rotations, which brings about the required additional symmetry.

It should be emphasized that a specific feature of this problem is that symmetry with respect to rotations through an arbitrary angle referred to as conventional in a previous paragraph is actually an additional symmetry due to the spontaneous breaking of the system's isotropy.

It follows from these and many other examples that the number of problems with short-wavelength instability of the spatially uniform state and an additional continuous group of symmetry is not so small. It will be shown below that in such problems evolution of unstable perturbations is qualitatively different from that in analogous systems without the additional symmetry. However, before undertaking the comparison, it needs to be understood how such perturbations behave in the conventional cases. This issue is briefly discussed in the next Section.

2. Short-wavelength instability in conventional systems

2.1 Universality of the Swift – Hohenberg and Ginzburg – Landau equations

Let us turn back to the Rayleigh – Benard problem and consider (at first, qualitatively) what happens with an arbitrary small spatially non-uniform perturbation of a trivial state upon a slight rise over the convection threshold. Expanding such a perturbation in the plane xy in the Fourier

† This allows to study the problem employing optical technique.

‡ Ratio of kinematic viscosity to thermal conductivity.

integral and employing the superposition principle, which remains valid till the Fourier transform amplitude is sufficiently small, one arrives at the conclusion that at the early stage evolution of each spatial harmonic is described by the factor $\exp[\gamma(k)t]$ (increment $\gamma(k)$ is shown in Fig. 2). It means that regardless of the particular form of the initial perturbation, almost all of its spatial harmonics undergo fast exponential fall, except a narrow wave packet centered in the vicinity of $k = k_c$. Therefore, long-time behavior of the problem will be determined by slow evolution of the modes from this narrow packet.

At a small rise over the threshold $\gamma_{max} \equiv \max\{\gamma(k)\}$ is also relatively small (see Fig. 2). Since the characteristic time for the instability development is determined by the inverse increment value (i.e., by $1/\gamma_{max}$), the smallness of γ_{max} is a quantitative measure of the process slowness. In practice, however, it is more convenient to relate the slowness of the instability development to the value of the control parameter responsible for the stability or instability of the system. The role of such a parameter in the Rayleigh – Benard problem is played by the temperature difference, ΔT, between the upper and lower surfaces of the layer or, to be precise, by a certain dimensionless combination of problem constants called the *Rayleigh number*, *Ra*, which is proportional to this difference [4, 5]. Near the instability threshold, Ra_c, it is appropriate to introduce the *normalized* or *reduced control parameter*,

$$\varepsilon \equiv \frac{Ra - Ra_c}{Ra_c} \,, \tag{6}$$

which will be referred to as a *control parameter*, for brevity.

It is clear that γ_{max} is a certain function of ε satisfying the condition $\gamma_{max} = 0$ at $\varepsilon = 0$. Assuming the dependence $\gamma_{max}(\varepsilon)$ to be smooth and expanding it in powers of ε, it is easy to see that at small ε the quantities γ_{max} and ε may be considered proportional to each other.

Let us now turn back to dynamics of instability development in the Rayleigh – Benard problem. As the amplitude of unstable modes grows, the linear approximation becomes inapplicable sooner or later, and *non-linear mode coupling* should be taken into account. To understand the role of this coupling note that the arising convective flow improves heat exchange between the upper and lower surfaces compared to that in a quiescent fluid. The improvement must decrease the temperature difference between the top and the bottom of the layer and hence diminish the primary cause of instability. In other words, the *non-linear mode coupling* promotes stabilization of the flow. For this reason it may be expected that the asymptotic state of the system should correspond to hydrodynamic motion with relatively low characteristic velocities — the lower the smaller ε (weakly non-linear regime), what does agree with experiment [4, 5].

These features of the phenomenon are employed as a basis to develop various versions of perturbation theory, where non-linearity of the problem yields certain correction to the eigenfunctions of the linear stability problem of the trivial (non-convective) state. The important point of such an approach is that the amplitudes of the eigenfunctions are not arbitrary small constants any more — they are fixed at certain small values obtained as a result of the analysis. These studies initiated by the works of Gor'kov [25], Malkus and Veronis [26], and Schluter et al. [27], now constitute a self-contained area of hydrodynamic research (see the review of recent progress in this field in Ref. [4]).

The initial equation in all versions of the theory is the Navier – Stokes equation. The z-dependence of the solution is assumed to be the same as for eigenfunctions of the linear stability problem, the coordinate z is easily excluded and the problem is reduced to a two-dimensional. It should be emphasized that near the threshold of the short-wavelength instability different problems exhibit certain universality. To some extent the situation is similar to that in phase transition phenomena. As well as in the latter case when peculiarity of a given system in the vicinity of a transition point is reduced to a set of critical exponents [28], peculiarities of a given dissipative system near the threshold of the short-wavelength instability are reduced to certain universal non-linear partial differential equations [2 – 4].

The universal nature of these equations can be understood in the framework of the phenomenological approach [4, 29] tracing back to Landau's basic works on the theory of phase transitions [28, 30] and turbulence [31, 32]. Let us consider the problem arising when z-dependence of the solution is already removed, confining ourselves, for the sake of simplicity, to a one-dimensional case, where the dependence of the solution on the coordinate y is suppressed, say due to geometry (convection in a long narrow container with the width close to that of the layer). Let $u(x,t)$ be the z-component of the velocity at the midplane (the plane passing through the midst of the layer), and $U_k(t)$ be its spatial Fourier transform:

$$U_k(t) \equiv \frac{1}{2\pi} \int u(x,t) \exp(ikx)\, dx \,.$$

It is clear that in the linear approximation $U_k(t)$ must satisfy the following equation (here and henceforth, $\gamma(k)$ is taken to be purely real):

$$\frac{dU_k}{dt} = \gamma(k) U_k \,. \tag{7}$$

Furthermore, it is worth to note, following Refs [31, 32], that the right-hand side of Eqn (7) is actually the first term of the expansion in powers of U_k. Adding to this expansion terms of higher order, one obtains that in the most general case the evolution equation must have the following form [33]:

$$\frac{dU_k}{dt} = \gamma(k) U_k - \int \alpha(k, k_1, k_2)\, U_{k_1} U_{k_2} \delta(k - k_1 - k_2) \times dk_1\, dk_2 - \int \beta(k, k_1, k_2, k_3)\, U_{k_1} U_{k_2} U_{k_3} \times \delta(k - k_1 - k_2 - k_3)\, dk_1\, dk_2\, dk_3 - \ldots \tag{8}$$

Here, $\alpha(k, k_1, k_2)$ and $\beta(k, k_1, k_2, k_3)$ are certain real coefficients (*vertices*, or *matrix elements*), while the δ-functions reflect the fact that non-linear mode coupling affects a mode with a given wave vector **k** if and only if the sum of the wave vectors of these modes is **k**.

Formally, integration in Eqn (8) is from plus to minus infinity. However, it is clear from the above qualitative features of the problem that the major contribution to the integrals comes from the vicinities of the point $k = k_c$ and the points $k = nk_c$, $n = 0, \pm 1, \pm 2, \ldots$, which correspond to wave vectors of the mode's spatial satellites. Satellites appear due to non-linear effects responsible for generation of multiple spatial harmonics. It is essential that all satellites lie far in the linear stability region (since$\gamma(nk_c) < 0$, and its absolute value is high at $n \neq 1$). On the other hand, the non-linear effects are weak, and each subsequent term on the right-hand side of Eqn (8) has additional smallness in U. It is therefore natural to suppose that characteristic values of U_k decrease with $|n|$ increasing, and U_k with $k \simeq \pm k_c$ have the highest values.

Thus, the integrals on the right-hand side of Eqn (8) actually split up into sums of integrals taken over small vicinities of points $k = nk_c$. In this case, assuming that coefficients $\alpha(k, k_1, k_2)$ and $\beta(k, k_1, k_2, k_3)$ are smooth functions of wave vectors, expanding them in small deviations of each argument from the corresponding nk_c, and restricting ourselves to the main approximation (i.e., keeping only the first non-trivial terms), we obtain that

$$\alpha(k, k_1, k_2) \simeq \alpha(nk_c, n_1 k_c, n_2 k_c) = \text{const}$$

in each vicinity. Similar considerations give

$$\beta(k, k_1, k_2, k_3) \simeq \beta(nk_c, n_1 k_c, n_2 k_c, n_3 k_c) \simeq \text{const}\,.$$

Certainly, the values of the constants are different for different sets of n.

It has been already mentioned that the modes with $k \simeq k_c$ are of special interest. Let us suppose that $k = k_c + p$ in Eqn (8), with p being a small quantity, and show that in this case the quadratic term on the right-hand side of Eqn (8) is actually of the third order of smallness and may be included into the cubic one. Indeed, if the conjectural hierarchy of distribution of characteristic values of U_k with the number n is taken into account, this term in the main approximation at a given k is divided into two integrals,

$$\begin{aligned}&\int \alpha(k, k_1, k_2)\, U_{k_1} U_{k_2} \delta(k - k_1 - k_2)\, \mathrm{d}k_1\, \mathrm{d}k_2 \\ &\simeq 2\left\{\alpha(k_c, k_c, 0) \int U_{k_c+p_1} U_{p_2} \delta(p - p_1 - p_2)\, \mathrm{d}p_1\, \mathrm{d}p_2 \right. \\ &\left. + \alpha(k_c, 2k_c, -k_c) \int U_{2k_c+p_1} U_{-k_c+p_2} \delta(p - p_1 - p_2)\, \mathrm{d}p_1\, \mathrm{d}p_2 \right\},\end{aligned} \tag{9}$$

(factor 2 on the right-hand side of Eqn (9) appeared due to the symmetry of coefficient α with respect to permutation of 'dummy' variables $k_{1,2}$: $\alpha(k, k_1, k_2) = \alpha(k, k_2, k_1)$).

Let us consider for certainty the first integral. Besides the amplitude U_k of the mode with $k = k_c$ it includes the amplitude U_{p_2} of a long-wavelength mode with small p_2. In turn, the equation for long-wavelength modes' amplitudes in the main approximation has the form

$$\begin{aligned}\frac{\mathrm{d}U_p}{\mathrm{d}t} &= \gamma(p) U_p - 2\alpha(0, k_c, -k_c) \int U_{k_c+p_1} U_{-k_c+p_2} \\ &\quad \times \delta(p - p_1 - p_2)\, \mathrm{d}p_1\, \mathrm{d}p_2\,.\end{aligned} \tag{10}$$

It is worthwhile to note that temporal evolution of the long-wavelength mode with $k = p$ occurs generally due to its non-linear coupling with short-wavelength modes with $k \simeq k_c$. Therefore, the characteristic time of the amplitude $U_p(t)$ changes coincides with that for $U_{k_c}(t)$; in other words, it is of the order $1/\gamma_{\max}$. Hence,

$$\left|\frac{\mathrm{d}U_p}{\mathrm{d}t}\right| \sim \gamma_{\max} |U_p| \ll |\gamma(p) U_p|\,, \tag{11}$$

since $\gamma(p) \approx \gamma(0)$ is a large negative quantity.

Condition (11) provides grounds for neglecting $\mathrm{d}U_p/\mathrm{d}t$ in Eqn (10), compared to $\gamma(p)U_p$. This allows $U_p(t)$ to be expressed in the explicit form. Substitution of this expression into Eqn (9) yields the formula

$$\begin{aligned}&\int \alpha(k_c, k_c, 0) U_{k_c+p_1} U_{p_2} \delta(p - p_1 - p_2)\, \mathrm{d}p_1\, \mathrm{d}p_2 \\ &\quad \simeq \frac{2\alpha(k_c, k_c, 0)\alpha(0, k_c, -k_c)}{\gamma(0)} \int U_{k_c+p_1} U_{k_c+p_2} U_{-k_c+p_3} \\ &\quad \times \delta\big[(k_c + p) - (k_c + p_1) - (k_c + p_2) - (-k_c + p_3)\big] \\ &\quad \times \mathrm{d}p_1\, \mathrm{d}p_2\, \mathrm{d}p_3\,,\end{aligned} \tag{12}$$

which has the same form as that of the main approximation to the cubic term on the right-hand side of Eqn (8)†. The second integral on the right-hand side of Eqn (9) is reduced to the same form by a similar procedure.

Thus, the quadratic term on the right-hand side of Eqn (8) is excluded and the equation in the said approximation takes the form

$$\begin{aligned}\frac{\mathrm{d}U_k}{\mathrm{d}t} &= \gamma(k) U_k - \beta \int U_{k_1} U_{k_2} U_{k_3} \\ &\quad \times \delta(k - k_1 - k_2 - k_3)\, \mathrm{d}k_1\, \mathrm{d}k_2\, \mathrm{d}k_3\,,\end{aligned} \tag{13}$$

where β is a constant which may be derived from $\beta(k_c, k_c, k_c, -k_c)$ if we take into account the symmetry relative to permutations of dummy variables $k_{1,2,3}$ and the contributions due to the exclusion of the quadratic term. Despite actually integration in Eqn (13) is carried over a small vicinity of $k_1 = k_c$, $k_2 = k_c$, and $k_3 = -k_c$, it may be extended from minus to plus infinity because of fast convergence of the integrals.

Now let us discuss the dispersion relation $\gamma(k)$. Note that invariance of the problem with respect to spatial reflections dictates that $\gamma(k)$ must be the an even function of k, since the replacement $x \to -x$ in x-space is equivalent to the replacement $k \to -k$ for Fourier transforms. Moreover, it follows from the above that the most important part in Eqn (13) is the region of k close to k_c. Therefore, the true dependence $\gamma(k)$ may be approximated in such a manner that the approximation is in quantitative agreement with the true dispersion relation within the specified region, while outside of it only qualitative agreement (viz $\gamma(k) < 0$ and should not be small at $|k - k_c| \geqslant k_c$) is sufficient. The simplest approximation of this kind is represented by the expression

$$\gamma(k) = \frac{1}{\tau_0}\left[\varepsilon - \xi_0^4 (k^2 - k_c^2)^2\right], \tag{14}$$

where τ_0 and ξ_0 are constants having the meaning of the characteristic temporal and spatial scales, respectively.

Insertion of relation (14) into Eqn (13) supplemented by inverse Fourier transform (with $k \to -\mathrm{i}\partial/\partial x$) leads to an equation which after scale transformation

$$\begin{aligned}&U \to \frac{(k_c \xi_0)^2}{\sqrt{\beta}}\, U\,, \qquad t \to \frac{\tau_0}{(\xi_0 k_c)^4}\, t\,, \\ &x \to \frac{x}{k_c}\,, \qquad \varepsilon \to (k_c \xi_0)^4 \varepsilon\end{aligned} \tag{15}$$

† This procedure is actually the exclusion of *slaved modes* from the problem according to Haken's approach [3]. Note also, it is clearly seen from Eqn (10) that the conjectured hierarchy of the characteristic values of the amplitudes does hold.

is reduced to the following universal form:

$$\frac{\partial u}{\partial t} = \left[\varepsilon - \left(1 + \frac{\partial^2}{\partial x^2}\right)^2\right] u - u^3 \,. \tag{16}$$

Here, old notations are used for the new dimensionless quantities, since it cannot cause misunderstanding. It is worth noting also that Eqn (16) was derived under an implicit assumption that the constant β is positive. When $\beta > 0$ the non-linear term on the right-hand side of Eqn (13) has stabilizing effect. When $\beta < 0$ the same term destabilizes the problem, and higher non-linearities need to be taken into account to describe the problem adequately, as it was pointed out already by Landau [31].

Eqn (16) known as the Swift – Hohenberg equation was first derived (from the Navier – Stokes equation in a way different from that described above) in the two-dimensional form $\partial^2/\partial x^2 \to \Delta_2$, where Δ_2 is the two-dimensional Laplacian, to study the role of fluctuations in the Rayleigh – Benard convection [34]. The importance of this equation for the description of pattern formation in distributed systems was perceived later (see, for instance, Ref. [4] for more details).

At the same time, Eqn (16) admits of further simplification. To do this, the fast-oscillating part of the solution should be singled out in an explicit form; in other words, $u(x,t)$ must be represented as

$$u(x,t) = \frac{1}{2}\left[\psi(x,t)\exp(\mathrm{i}x) + \psi^*(x,t)\exp(-\mathrm{i}x)\right] , \tag{17}$$

where $\psi(x,t)$ is a slowly varying function, an *envelope*. It ought to be remembered that in new dimensionless variables $k_{\mathrm{c}} = 1$ because now

$$\gamma(k) = \varepsilon - (k^2 - 1)^2 \tag{18}$$

[see Eqn (16)]. Substituting k in the form $1 + p$, where $p \lesssim \sqrt{\varepsilon}$, it is possible to rewrite Eqn (18) as

$$\gamma(k) \simeq \varepsilon - 4p^2 \,, \tag{19}$$

where dropped terms are of order $p^3 \sim \varepsilon^{3/2}$.

On the other hand, representing $U(x,t)$ and $\psi(x,t)$ in the form of the Fourier integrals, it is easy to see that

$$\mathrm{i}p\left[\psi \exp(\mathrm{i}x)\right]_p \to \exp(\mathrm{i}x)\frac{\partial \psi}{\partial x} \,. \tag{20}$$

Therefore, the dispersion relation (19) in the x-representation corresponds to the operator†

$$\varepsilon + 4\frac{\partial^2}{\partial x^2} \,, \tag{21}$$

where differentiation acts *only* on ψ and does not act on $\exp(\mathrm{i}x)$, in accordance with Eqn (20).

At last, raising the right-hand side of Eqn (17) to the third power and omitting higher spatial harmonics, Eqn (16) can be reduced to the form

$$\frac{\partial \psi}{\partial t} = 4\frac{\partial^2 \psi}{\partial x^2} + \varepsilon\psi - \frac{3}{4}|\psi|^2\psi \,.$$

† It is possible to obtain the same result by a much more complicated procedure, substituting Eqn (17) into (16), taking into account that $\partial\psi/\partial x \sim \sqrt{\varepsilon}\psi \ll \psi$ because ψ is slow, and retaining only the main approximation in spatial derivatives.

The scale transformation

$$x = \frac{2}{\sqrt{\varepsilon}}X \,, \qquad t = \frac{T}{\varepsilon} \,, \qquad \psi = 2\sqrt{\frac{\varepsilon}{3}}\,\Psi \tag{22}$$

yields the final form of this equation,

$$\Psi_T = \Psi_{XX} + \left(1 - |\Psi|^2\right)\Psi \,, \tag{23}$$

where indices denote respective differentiations.

Eqn (23) is nothing but the well-known Ginzburg – Landau equation. It is noteworthy that this equation written in the form (23) contains neither small nor large parameters, all its quantities being of order of unity. Hence, characteristic values of the initial variables can be estimated straightforwardly from the scale transformation (22): $\psi \sim \sqrt{\varepsilon}$, the characteristic spatial scale of $\psi(x,t)$ variations is of order $1/\sqrt{\varepsilon}$, and the characteristic temporal scale $\sim 1/\varepsilon$.

It should be emphasized before proceeding further that Eqn (23) is invariant with respect to the transformation $\Psi \to \Psi\exp(\mathrm{i}\varphi_0)$, where φ_0 is an arbitrary constant phase (rotational invariance in the complex plane). This invariance generates a certain conserved quantity maintained to be equivalent to the angular momentum in the problem of a classical particle dynamics in an axially symmetric potential. The presence of an additional conservation law plays, in turn, an important role in some specific problems, e.g., strict selection of the wavenumber of a spatially periodic pattern in a domain wall [35, 36].

It is essential, however, that the symmetry in question is not an additional symmetry of the *initial* problem. It appears upon transition from Eqn (16) to Eqn (23) *due* to approximations made during this transition. Such a symmetry is called *accidental* and disappears in higher orders of the theory of perturbations. Indeed, some terms omitted upon transition from Eqn (16) to Eqn (23) (e.g., those of the form of $\psi^3\exp(3\mathrm{i}x)$) break this symmetry *together with the conservation law* mentioned in the previous section. The symmetry breaking leads to certain qualitatively new effects related to higher orders of perturbation theory [36].

In fact, these issues are beyond the framework of the present review, but I could not help taking this opportunity and drawing attention of readers to the insidiousness of accidental symmetry. The overwhelming majority of equations used in theoretical physics are approximate and can be supplemented with small (compared to the examined approximation) terms. If these small terms reduce the symmetry of a given problem, it is always the cause for expectation of qualitatively new effects which may appear in association with such terms, no matter how small they might be.

Let us turn back to the Swift – Hohenberg equation and discuss how a problem changes if short-wavelength instability is considered in two spatial dimensions. The starting point in this case is again the evolution equation in the form (8) , where scalar k must be now replaced by corresponding two-dimensional vectors. This results in two qualitatively new effects.

First, the quadratic term is of the second (rather than the third, as before) order of smallness, as it should be, and can not be reduced to the cubic one. When the problem contains no special reason accounting for the numerical smallness of coefficient α, it is not difficult to show (see, for example, Ref. [33]) that explosive instability associated with the unlimited growth of the amplitude of a hexagonal cellular pattern occurs at $\varepsilon > 0$.

However, it should be noted that the quadratic term in Eqn (8) breaks the invariance of this expression with respect to change in sign of U_k. Meanwhile, such an invariance in certain problems follows from their symmetry. Specifically, in the case of convection in a fluid layer placed between two solid surfaces, the net force acting on an upflowing fluid particle, overheated by δT compared to the surrounding fluid, is of the same magnitude (although opposite in sign) as that for an analogous sinking particle overcooled by δT and located symmetrically with respect to the midplane. For this reason, the profile of the x-component of the fluid velocity must be invariant under reflections in the midplane, resulting in symmetry of the problem with respect to $U_k \to -U_k$, so that $\alpha(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \equiv 0$ vanishes identically.

When more subtle effects (e.g., temperature dependence of viscosity) are taken into consideration, the symmetry $U_k \to -U_k$ is broken. In such cases, the coefficient $\alpha(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2)$ does not vanish identically but remains small, in conformity with the smallness of the mentioned effects. At small α, the quadratic and cubic terms in the two-dimensional version of Eqn (8) may be of the same order and both must be taken into account.

Another qualitative difference between one and two-dimensional versions of Eqn (8) is that the quantity $\beta(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ is no longer constant even in the main approximation in deviations of vectors $\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3$ from k_c — it becomes a function of the rhomboid cone formed by the four vectors:

$$\beta(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)_{k=k_1=k_2=k_3=k_4=k_c} = \beta(\theta)\,.$$

Thus, in the two-dimensional case the Swift – Hohenberg equation is no longer a universal equation describing weak non-linear dynamics of a broad class of problems with short-wavelength instability. The more general Eqn (8) is reduced to the two-dimensional version of the Swift – Hohenberg equation if and only if $\alpha(k, \mathbf{k}_1, \mathbf{k}_2) \equiv 0$ and $\beta(\theta) = \text{const}$. Nevertheless, the two-dimensional Swift – Hohenberg equation and its various generalizations provide good and rather simple models whose study is beneficial for understanding certain general aspects of pattern formation (see Ref. [4] for a detailed discussion of these issues).

2.2 Spatially periodic solutions and their stability. Eckhaus criterion

Now, it is appropriate to present a summary of major findings pertinent to spatially periodic solutions of Eqns (16), (23) and their stability (readers interested in details are referred to Refs [4, 37 – 43]).

It is easy to see that Eqn (23) has a family of steady spatially periodic solutions of the form

$$\Psi = \sqrt{1 - K^2}\exp\left[\mathrm{i}(KX + \varphi_0)\right], \tag{24}$$

which formally exist at any $K \leqslant 1$. Let us discuss stability of these solutions for small perturbations. To this effect, it is convenient to represent function Ψ as $\Psi = R\exp(\mathrm{i}\Phi)$ and the Ginzburg-Landau equation in the form

$$R_T = R_{XX} + (1 - R^2 - \Phi_X^2)R\,, \tag{25}$$

$$\Phi_T = \frac{1}{R^2}(R^2\Phi_X)_X\,. \tag{26}$$

Then solution (24) is transformed into

$$R = \sqrt{1 - K^2}\,, \qquad \Phi = KX + \varphi_0\,. \tag{27}$$

Linearizing the system (25), (26) about this solution and representing the perturbed solution in the form

$$\widetilde{R} = R + r\exp(\sigma T + \mathrm{i}QX)\,, \tag{28}$$

$$\widetilde{\Phi} = \Phi + \varphi\exp(\sigma T + \mathrm{i}QX)\,, \tag{29}$$

where r and φ are infinitesimal constants†, one arrives at a quadratic dispersion equation for $\sigma(K, Q)$ obtained by the standard procedure, whose solutions are

$$\sigma_{1,2} = K^2 - 1 - Q^2 \pm \sqrt{(1 - K^2)^2 + 4K^2Q^2}\,. \tag{30}$$

It is worthwhile to note that the dispersion relation (30) always corresponds to purely real values of σ. It means that small perturbations of the steady solution evolve monotonically in time. Furthermore, sign 'minus' in front of the radical in Eqn (30) corresponds to negative values of σ at any K and Q, that is one of the branches of the spectrum described by the relation (30) (hereinafter denoted as σ_2, for brevity) is always stable. The second branch σ_1 (sign 'plus' in front of the radical) corresponds to stable perturbations ($\sigma_1 \leqslant 0$ at any Q values) only if

$$K^2 < K_\mathrm{E}^2 \equiv \frac{1}{3}\,. \tag{31}$$

At $K^2 > 1/3$, the spectrum contains an instability region: $\sigma(K, Q)$ becomes positive at

$$-\sqrt{2(3K^2 - 1)} < Q < \sqrt{2(3K^2 - 1)}\,. \tag{32}$$

It is essential that the range (32) always includes the point $Q = 0$. On the other hand, for the examined branch of the spectrum $\sigma = 0$ at $Q = 0$ and any K. It will be shown below that nullification of $\sigma(K, Q)$ at $Q = 0$ is not accidental — it constitutes the basic attribute of the problem related to translational invariance of Eqn (23). The corresponding neutrally stable mode is called the Goldstone mode and plays a very important role in the stability problem.

At small Q, it is easy to obtain from the dispersion relation (30) that

$$\sigma_1 = \frac{3K^2 - 1}{1 - K^2}\,Q^2 + O(Q^4)\,, \tag{33}$$

i.e., instability is associated with change in sign of the coefficient in front of Q^2 in the expansion of $\sigma(K, Q)$ in powers of Q^2. The branch of the spectrum describing unstable modes detaches from the point $Q = 0$ (the *Goldstone branch*), and the wave numbers Q corresponding to unstable modes are concentrated at $0 < 3K^2 - 1 \ll 1$ in a small vicinity of zero. In other words, instability is related to the *long-wavelength* modulation of spatially periodic patterns.

Therefore, formally existing steady spatially periodic solutions of the Ginzburg – Landau equation appear to lose stability when their wavenumber K becomes too large. It is possible to obtain the stability criterion (31) in a somewhat different form employing the expression (27), which relates K to the amplitude of the examined spatially periodic pattern.

† Phase perturbation (29) corresponds to perturbation Ψ of the form $R\exp\left[\Phi + \varphi\exp(\sigma T + \mathrm{i}QX)\right]$ which looks rather unusual. However, taking into account the smallness of φ, this expression may be rewritten as $R\exp(\Phi)\exp\left[\varphi\exp(\sigma T + \mathrm{i}QX)\right] \simeq R\exp(\Phi)[1 + \varphi\exp(\sigma T + \mathrm{i}QX)]$, which reduces the perturbation to the standard form.

Namely, spatially periodic solutions are stable if their amplitude $R(K)$ satisfies the condition

$$\frac{R^2(K)}{R^2_{\max}} > \frac{2}{3}, \quad (34)$$

where $R_{\max} \equiv \max_k\{R(K)\} = R(K_c)$. It should be recalled that K_c corresponds to the maximum of instability growth rate of perturbations of the trivial state. The trivial state for the Ginzburg–Landau equation is $\Psi \equiv 0, \gamma(K) = 1 - K^2$ and $K_c = 0$.

The advantage of the stability criterion in the form (34) is its scale invariance, i.e., independence of its form of the units of measure. Therefore, this criterion is the *universal law* which equally holds for the Swift–Hohenberg equation (16) and *any other equation* with short-wavelength instability which can be reduced to Eqns (16), (23) at a small rise over the threshold by making use of the previously described procedure. It should be emphasized that 2/3 in the right-hand side of Eqn (34) is the *world constant*, the same as, say, π. Conditions equivalent to Eqns (31), (34) were first obtained in Refs [44, 45] and are known as the Eckhaus criterion.

Finally, there is one more variant of the Eckhaus criterion convenient for the purpose of comparing theory and experiment. Let us construct in the plane $k\varepsilon$ a curve describing *neutral stability* of a trivial state, in accordance with the condition $\gamma(k) = 0$. It follows from Eqn (14) that at $|k - k_c| \ll k_c$ this curve is the parabola

$$\varepsilon = 4\xi_0^4 k_c^2 (k - k_c)^2 . \quad (35)$$

On the other hand, the neutral stability condition $\gamma(K) = 0$ for the Ginzburg–Landau equation corresponds to the condition $K^2 = 1$. Comparison with Eqn (35) yields the scale correspondence

$$K^2 = \frac{4\xi_0^4 k_c}{\varepsilon} (k - k_c)^2 . \quad (36)$$

Then, the Eckhaus criterion in the form (31) in dimensional variables corresponds to the condition

$$\varepsilon > 3 \cdot 4\xi_0^4 k_c^2 (k - k_c)^2 , \quad (37)$$

and defines one more parabola *inside* the former one, given by Eqn (35), which is steeper than that. This second parabola separates stable spatially periodic solutions from unstable ones.

Such curves separating stable and unstable spatially periodic states in the space of problem's parameters (conventionally in k_ε plane) are usually referred to as *stability balloons* or *Busse balloons*, named after F. Busse to emphasize a valuable contribution of this author to the problem of stability of spatially periodic patterns in the Rayleigh–Benard convection (a review and bibliography of these Busse's works can be found in Refs [4, 46, 47]).

It is worthy of note that the universal nature of the Eckhaus criterion is essentially related to the reduction of the evolution equation in the general form (8) to the Ginzburg–Landau equation, which is true only to the *main* approximation in small $\sqrt{\varepsilon}$. Terms of higher order in ε generate corrections to the conditions (34), (37), whose particular form is not universal and depends on particular form of the governing equations. It is important, however, that all such corrections are small (see, for instance, Ref. [48]).

To conclude, here is a summary of major qualitative results for conventional systems which were obtained above:

(1) When ε values are sufficiently small, the problem always has a certain family of *stable* stationary spatially periodic solutions.

(2) The mentioned family determines in the $k\varepsilon$ plane a finite Busse balloon whose width (i.e., the distance between the right and left boundaries at a fixed ε) is of order of $\sqrt{\varepsilon}$.

(3) Relaxation of stable perturbations of spatially periodic solutions (inside the Busse balloon) and the growth of unstable ones occur *monotonically* in time ($\operatorname{Im}\sigma \equiv 0$).

(4) Unstable modes outside the Busse balloon are always associated with the Goldstone branch of the spectrum.

(5) Stability criterion for spatially periodic solutions in the main (in small $\sqrt{\varepsilon}$) approximation can be obtained from the main approximation to the evolution equation; small corrections to this equation yield small corrections to the stability criterion.

In the case of short-wavelength instability in two-dimensional systems the problem is more complicated: there appear new types of patterns (spatially periodic cellular patterns [4] and even quasiperiodic patterns [4, 49, 50–53]) and new types of instability. At the same time, the above qualitative features of the stability problem for steady patterns as a rule hold for two-dimensional problems also.

The situation changes dramatically in the case of problems with additional symmetry. The presence of an additional continuous symmetry group regardless of spatial dimension of the problem leads to a change in all the five features listed above. Let us try to understand what causes such changes.

3. Short-wavelength instability in systems with an additional continuous group of symmetry

3.1 Symmetry and Goldstone modes

It should be emphasized from the very beginning that any continuous symmetry group as a rule leads to appearance of the corresponding Goldstone branch(s) in the spectrum of perturbations of steady solutions. Indeed, let $u(\mathbf{r})$ be a certain steady solution of the problem and $\widehat{T}_a$ the operator of a continuous one-parameter symmetry group, which transforms the solution $u(\mathbf{r})$ into $u_a(\mathbf{r})$:

$$\widehat{T}_a u(\mathbf{r}) = u_a(\mathbf{r}) . \quad (38)$$

Here a is the parameter which specifies this transformation. Specifically, for the spatial translation group we have

$$\widehat{T}_a u(x) = u(x + a) , \quad (39)$$

for the group of rotations in a complex plane

$$\widehat{T}_a u(\mathbf{r}) = u(\mathbf{r}) \exp(\mathrm{i}a) , \quad (40)$$

etc.

Certainly, $u_a(\mathbf{r})$ is also a stationary solution to the problem, that is it identically satisfies the corresponding equations and boundary conditions. Now, let us consider the case of an infinitesimal a, when

$$u_a(\mathbf{r}) = u(\mathbf{r}) + a\left[\frac{\partial}{\partial a}\, u_a(\mathbf{r})\right]_{a=0} \quad (41)$$

(we took into account that $u_a(\mathbf{r})|_{a=0} \equiv u(r)$). On the other hand, the second term on the right-hand side of Eqn (41) may be regarded as a perturbation of the stationary solution $u(\mathbf{r})$. This perturbation transforms $u(\mathbf{r})$ into another stationary state $u_a(\mathbf{r})$. Such a perturbation must be neutrally stable. In other words, $(\partial u_a/\partial a)_{a=0}$ must be an eigenfunction of the corresponding stability problem for $u(\mathbf{r})$ whose eigenvalue equals zero, i.e., a Goldstone mode†.

Therefore, the number of Goldstone modes in the problem of stability of a given solution $u(\mathbf{r})$ as a rule coincides with the number of continuous scalar quantities which parametrize all possible symmetry transformations, i.e., the entire continuous symmetry group of the problem of interest. Passage 'as a rule' means that it is not the case always. Indeed, there are exceptions to this rule including the two major ones below.

Firstly, an individual solution subjected to some symmetry transformations may remain unchanged, i.e., these transformations applied to this solution do not generate any new *equivalent* solution, transforming the solution into itself, so that $u_a(\mathbf{r}) \equiv u(\mathbf{r})$. In this case, $(\partial u_a/\partial a)_{a=0}$ is identically equal to zero and no Goldstone mode arises. Examples include spatial translations, rotations, etc., of the trivial solution $u \equiv 0$, or translation along the axis y in a two-dimensional problem applied to an y-independent quasi-one-dimensional solution $u(x)$.

Another important exception may occur in a problem with several different symmetry transformations (the entire symmetry group G is the direct product of a certain number of lower-symmetry subgroups: $G = G_1 \times G_2 \times \ldots \times G_n$). In this case action of *different* subgroups on a given particular solution may be described by *the same operator* $\widehat{T}_a$ due to accidental degeneracy of the solution. It is the case, e.g., when solution (27) to the Ginzburg – Landau equation is subjected to spatial translations ($X \to X + a$), or rotations in a complex plane [$\Psi \to \Psi \exp(\mathrm{i}\varphi_0)$]: the results are the same if $\varphi_0 \equiv aK$. For this reason the accidental symmetry of the problem with respect to rotations in the complex plane does not cause any complication in the stability analysis of solutions (24). The stability problem has a *single* Goldstone mode [see Eqn (33)] despite the presence of *two* continuous one-parameter groups of symmetry.

Why do I pay so much attention to Goldstone modes? The fact is that these modes are not so much important by themselves as they are due to their giving rise to Goldstone branches of the spectrum, which determine bands of slowly evolving perturbations. Presence of these *bands* must in turn be taken into account explicitly in both derivation of the corresponding evolution equation (Goldstone branches in the stability problem for the trivial state) and analysis of stability of its solutions.

For example, in the above discussion of short-wavelength instability in systems with conventional symmetry, the problem of the trivial state stability had no Goldstone modes at all, while the problem of stability of weakly non-linear spatially periodic solutions had only one Goldstone mode related to translational symmetry of these solutions. Accordingly, in the derivation of the evolution equation, slowly evolving modes were only those with the wave numbers $k \approx k_c$ associated with short-wavelength instability. All other modes followed them *adiabatically*. As a result, the evolution equation was reduced to the Swift-Hohenberg equation (16) and the problem of stability of spatially periodic solutions to the Eckhaus criterion (34). Let us trace changes resulting from the appearance in the problem of Goldstone's modes associated with its additional symmetry.

† The operator $(\partial \widehat{T}_a/\partial a)_{a=0}$ which produces this mode from the solution $u(\mathbf{r})$ is called an infinitesimal generator of the corresponding symmetry group.

3.2. (Quasi)one-dimensional spatially periodic solutions and their stability

As it has been noted in the Introduction the Rayleigh – Benard problem with free-slip boundaries is one of the best-known problems of short-wavelength instability in systems with additional symmetry [4, 9 – 15]. The additional symmetry (associated with steady rotation of the layer as a whole) is responsible for appearance of a Goldstone branch in the perturbation spectrum of the trivial (non-convective) state. The branch describes slow relaxation of vortex modes with the vertical component of the vorticity vector

$$\boldsymbol{\Omega} \equiv [\nabla \times \mathbf{v}] = (0, 0, \Omega_z)\,, \qquad (42)$$

where $\mathbf{v}$ is the three-dimensional velocity vector. With this in mind, the authors of Refs [9 – 15] derived a system of evolution equations for two slowly evolving fields coupled by *quadratic* non-linearity: the field $u(x, y, t)$, where u is the v_z value at the midplane (as before), and the field Ω_z. The starting point was the system of the Navier – Stokes equations in the so-called Oberbeck – Boussinesq approximation leading to the symmetry of the problem with respect to change of sign of u. It is known (see, for instance, Ref. [4]) that the only weakly non-linear stationary spatially periodic solution to the problem in this approximation is formed by quasi-one-dimensional rolls, where u depends on a single spatial variable and has the form‡

$$u = R(k)\cos(kx + \varphi_0) + O(\varepsilon)\,, \qquad (43)$$

see Eqns (17), (22), (24).

Such a solution corresponding to zero vorticity ($\Omega_z = 0$) was obtained in Refs [9 – 15]. However, the key aspect is the stability analysis of this solution. The results of this analysis reported in Ref. [14] have nothing in common with the Eckhaus criterion and can be summarized as follows:

(1) There is an anomalous contraction of the Busse balloon to the width of order of ε or its eventual disappearance (depending on the Prandtl number), i.e., instability of all stationary spatially periodic solutions.

(2) In the presence of a finite Busse balloon there is oscillatory relaxation of perturbations of stable solutions inside the balloon, oscillatory growth of unstable perturbations outside one of the balloon's boundaries and their monotonic growth beyond the other.

However, the most unexpected result of the study [14] is that the stability criteria and some types of the most 'dangerous' (i.e., the first to cause instability) perturbations differ significantly from the criteria and perturbations obtained in *the same problem* by other authors [12, 13]. It should be emphasized that studies [12, 13] and [14] are based on *different* but equivalent versions of *the perturbation theory*, and the above discrepancy is not related to a trivial mistake in calculations.

‡ In this case, more complicated two-dimensional solutions $u(x, y)$ describing cellular patterns do not satisfy the evolution equations.

This paradox is resolved by Bernoff [54] who shows that the stability problem in the case under consideration has the following peculiarity: certain small (in $\sqrt{\varepsilon}$) corrections to the evolution equations always ignored in the traditional approach (see Section 1) may undergo rescaling during calculations in the given problem, so that the contribution of these terms into the final stability criterion is of the same order in $\sqrt{\varepsilon}$ as that for the main approximation. In other words, the problem becomes very sensitive to truncation of the series of perturbation theory. Approximation to the first non-trivial terms of this series used in Refs [12, 13] by analogy with the traditional approach is responsible for inadequate stability conditions. However, the approaches adopted in Refs [12, 13] and [14] are in excellent agreement with Ref. [54] if higher orders of the theory are taken into account.

This result of Ref. [54] naturally generates a number of questions regarding the minimally acceptable accuracy of the series of perturbation theory and the most convenient specific realizations of this theory for the analysis of stability of the solutions for a given problem.

To answer these questions as well as to make sure that the highlighted peculiarities of the stability problem should be regarded as the common property of all problems with additional symmetry, rather than a unique feature of the free-slip convection, let us consider, following Refs [55–57], one of the simplest realizations of the problem, involving the equation

$$\frac{\partial v}{\partial t} + \frac{\partial^2}{\partial x^2}\left[\varepsilon - \left(1 + \frac{\partial^2}{\partial x^2}\right)^2\right] v + v\frac{\partial v}{\partial x} = 0\,. \tag{44}$$

Eqn (44) was suggested by Nikolaevsky [58, 59] to describe seismic waves in visco-elastic media. The equation is written in dimensionless variables and v has the meaning of the displacement velocity. It is important that quadratic non-linearity $v\,\partial v/\partial x$ in Eqn (44) is due to the transition to a 'running' variable and is therefore typical of a broad class of problems involving front and/or interface motion [4]. As usual, the problem is formulated in infinite space: $-\infty < x < \infty$.

The analysis of stability of the trivial solution $v \equiv 0$ with respect to spatially periodic perturbations of the form (2) yields the dispersion relation

$$\gamma(k) = k^2\left[\varepsilon - (k^2 - 1)^2\right]\,, \tag{45}$$

which differs from Eqn (14) by presence of the Goldstone mode: $\gamma(k) = 0$ at $k = 0$. Let us show that this mode is related to an actual additional symmetry of the problem and is not accidental. With this end in view we consider the equation

$$\frac{\partial u}{\partial t} + \frac{\partial^2}{\partial x^2}\left[\varepsilon - \left(1 + \frac{\partial^2}{\partial x^2}\right)^2\right] u + \left(\frac{\partial u}{\partial x}\right)^2 = 0\,. \tag{46}$$

It is evident that differentiation with respect to x and redesignation $2\partial u/\partial x \equiv v$ reduce Eqn (46) to Eqn (44), i.e., these equations are equivalent in terms of formation of spatially periodic patterns. However, Eqn (46) unlike (44) explicitly exhibits additional symmetry with respect to the transformation

$$u \to u + \text{const}\,. \tag{47}$$

Despite apparent similarity of the problem (44) and the Swift–Hohenberg equation, their properties are significantly different, and this was understood very early [60, 61]. A detailed analysis of Eqn (44) and a number of its generalizations can be found in Ref. [61]. However, this analysis based on reduction of Eqn (44) to equations of the Ginzburg–Landau type for slowly varying envelopes is based on *main* approximation to these equations, that is, it contains the same mistake as that in Refs [12, 13]. For this reason, such a consideration is on the whole inadequate despite the correct description of certain qualitative features of the problem. Specifically, the presence of the finite Busse balloon for spatially periodic solutions of Eqn (44) reported in Ref. [61] is in conflict with the real situation; in fact, all such solutions turn out to be unstable [56].

To be sure, let us firstly obtain spatially periodic solutions to Eqn (46) in explicit form†. Representing $u(x, t)$ as a Fourier series

$$u(x,t) = \sum_{n=-\infty}^{\infty} U_{nk}(t)\exp(\mathrm{i}nkx)\,, \qquad U^*_{nk} = U_{-nk}\,, \tag{48}$$

and substituting Eqn (48) into (46), we have the following sequence of equations for amplitudes U_{nk}:

$$\frac{\mathrm{d}U_{nk}}{\mathrm{d}t} = \gamma(nk)U_{nk} + k^2\sum_{l=-\infty}^{\infty} l(n-l)\,U_{lk}U_{(n-l)k}\,. \tag{49}$$

It follows from the condition $\gamma(nk) = 0$ at $n = 0$ and the concrete form of non-linearity in Eqn (46) (the squared *gradient* of u) that the amplitude $U_0(t)$ is absent on the right-hand side of Eqn (49) and can always be completely excluded from the problem by means of transformation $\tilde{u}(x,t) = u(x,t) - U_0(t)$. For this reason later we will not pay any attention to this amplitude.

Let us consider the case of small positive ε and find the boundaries $k_{1,2}$ of the range $(0 < k_1 \leqslant k_2)$ of neutral stability in agreement with their standard definition, $\gamma(k_{1,2}) = 0$. It is easy to see that if the series (48) is truncated to a finite number of terms, i.e., if $|n| < N$, where N is any integer higher than or equal to two, for any k from the segment $k_1 \leqslant k \leqslant k_2$ there is a stationary solution of the form

$$U_{nk} = \sum_{m=0}^{N} U^{(m)}_{nk}\,, \qquad U^{(m)}_{nk} = O(\varepsilon^{(n+2m)/2})\,. \tag{50}$$

For example, for $N = 3$,

$$|U_k|^2 = -\frac{\gamma(k)\gamma(2k)}{4k^4}\left[1 + O(\varepsilon) + O(\varepsilon^2)\right]\,, \tag{51}$$

$$U_{\pm 2k} = -\frac{k^2 U^2_{\pm k}}{\gamma(2k)}\left[1 + O(\varepsilon)\right]\,, \tag{52}$$

$$U_{\pm 3k} = O(\varepsilon^{3/2}) \tag{53}$$

[at $|k-1| \sim \sqrt{\varepsilon}$ $\gamma(k) \sim \varepsilon$, and $\gamma(2k) \sim 1$, see Eqn (45)].

Therefore, to the main order in $\sqrt{\varepsilon}$, the amplitude of the stationary solution for the problem being examined is $\sim \sqrt{\varepsilon}$ [see Eqn (51)], that is of the same order as that for the Swift-Hohenberg equation [see Eqns (17), (22), (24)]. In other

† Due to pronounced additional symmetry of Eqn (46) mentioned above, this equation is more convenient for the subsequent analysis than Eqn (44).

words, the presence of an additional symmetry in Eqn (46) has no effect on the order of magnitude of its stationary solution amplitude. The reason for it is quite obvious. Equations for stationary amplitudes can include only amplitudes coupled resonantly with U_k. Of the whole spectrum of long-wavelength modes related to the additional symmetry, only the Goldstone mode itself satisfies this condition. On the other hand, $k = 0$ for the Goldstone mode, i.e., the mode is x-independent. Such a mode can not by itself play any role in formation of spatially periodic patterns and can always be excluded from the consideration either by transforming the order parameter u, as above, or by its differentiation with respect to x.

The stability problem constitutes quite a different case. Here, *any* spatially periodic perturbation can be related to the mode $U_k \exp(ikx)$, including those belonging to the long-wavelength Goldstone branch of the spectrum with small but non-zero wave vector. Interaction between such a perturbation and a stationary solution results in the long-wavelength spatial modulation of the latter, and it is impossible to avoid this modulation by trivial transformations of the order parameter. In other words, despite the additional symmetry does not affect the structure of spatially periodic solutions of the problem, it must result in dramatic changes in their stability problem. These changes and their nature are revealed in the next Section.

3.3 Symmetry and dispersion equation in case of mixed ε-scales

First of all let us try to understand the cause of the above-mentioned scale-mixing in the expansion of operators of evolution equations and their solutions in powers of $\sqrt{\varepsilon}$ (ε-scale-mixing, Refs [55, 56]) and to reveal what determines the minimally acceptable accuracy of calculations in such cases.

The formal answer to the latter question is trivial: the accuracy of intermediate calculations should be gradually increased until the stability criterion for stationary solutions stops changing at the leading order in $\sqrt{\varepsilon}$. However the real situation is not so simple as this. The calculation of each consecutive correction in $\sqrt{\varepsilon}$ requires more and more efforts because the number of terms of the order $(\sqrt{\varepsilon})^n$ sharply increases with n growing. On the other hand, only a few of these small corrections undergo rescaling in course of calculations and are responsible for a non-zero contribution to the main order of the stability criterion whereas all others may be dropped without serious detriment to the problem [56]. It is therefore crucial to be able to estimate the order of magnitude of the contribution made by each correction to the final stability conditions without calculating this contribution in the explicit form. This problem was solved in Ref. [56] due to an appropriate choice of the form of the perturbation theory in powers of $\sqrt{\varepsilon}$. This is a special question and its detailed discussion is beyond the scope of the present review. Therefore, only key aspects of such analysis are touched upon below. The reader is referred to Ref. [56] for more details.

A perturbation of the stationary solution of Eqn (46) is represented as

$$\delta u = \exp(\sigma t) \sum_n V_{nk+q} \exp\left[\mathrm{i}(nk+q)x\right] , \qquad (54)$$

where $k_1 \leqslant k \leqslant k_2$ and q is the wave vector of the perturbation. Substitution of expression (54) into Eqn (46) linearized near the solution (48), (50) results in a dispersion equation, which has a usual form of a certain determinant equal to zero†. It is easy to see that the diagonal elements of such a determinant have the from $\sigma - \gamma(nk+q)$.

It is relevant to note that the above analysis of stability of stationary solutions for the Swift – Hohenberg equation has shown that characteristic values of σ for unstable modes turn out to be of the order ε and characteristic values $q \sim \sqrt{\varepsilon}$, see Eqns (17), (22), (30), (32), (36). It is therefore believed that similar relations must hold in the present case also. Certainly, this hypothesis needs to be verified as soon as the final answer (i.e., the dispersion dependence $\sigma(k, q)$) is available.

Bearing in mind what has been postulated in previous sections, the quantity $\gamma(nk+q)$ appears to be of the order σ at $n = 0, \pm 1$, and of order of unity in all other cases (it should be remembered that the condition $k_1 \leqslant k \leqslant k_2$ implies $k = 1 + \varkappa$, where $\varkappa \sim \sqrt{\varepsilon}$). For this reason, the product of elements of the leading diagonal of the determinant at any large but finite order proves to be of the order of ε^3, which defines both the accuracy of the determinant's evaluation and the order of the dispersion equation, σ^3, at the main approximation in $\sqrt{\varepsilon}$.

However, the fact is that the minimal (in $\sqrt{\varepsilon}$) order of terms arising from the evaluation of the determinant turns out to be actually $\varepsilon^{5/2}$ rather than ε^3, and it is related to *non-diagonal* elements of the row corresponding to the projection of the evolution equation for the eigenvector of the problem (54) on the long-wavelength mode $\sim \exp(\mathrm{i}qx)$, i.e., a mode belonging to a new (compared to the Swift – Hohenberg problem) Goldstone branch of the spectrum. Therefore, to obtain the dispersion equation with the specified accuracy $\sim \varepsilon^3$ one should take into account *corrections* to elements of the determinant with the relative smallness up to order $\sqrt{\varepsilon}$. The necessity to consider these corrections is the actual cause of ε-scale-mixing. Thus, the ε-scale-mixing follows directly from the fact of the existence of an additional Goldstone mode in the perturbation spectrum, that is from the symmetry of the problem.

No serious difficulties are encountered in connection with performing calculations necessary to take into account the above corrections, even though they are very tiresome. They lead to a dispersion equation whose coefficients depend on subtle peculiarities of behavior of $\gamma(k)$ in the vicinity of the maximum corresponding to short-wavelength instability (such as $\gamma'|_{k=1} = O(\varepsilon)$ and $\gamma'''|_{k=1}$, where the prime indicates differentiation with respect to k), and also on the value of derivative $\gamma'|_{k=2}$. It can be inferred that the spectrum of the stability problem 'feels,' already in the main approximation in $\sqrt{\varepsilon}$, the deviation of the wave number maximizing $\gamma(k)$ from $k_c = 1$ (defined by $\gamma'|_{k=1}$), asymmetry of this maximum described by $\gamma'''|_{k=1}$, and the slope of the dispersion curve far in the linear stability region ($\gamma'|_{k=2}$). Taken together, these make the problem essentially non-local.

If the stability problem is generalized and is represented in the integral form (8), it shows dependence not only on the aforementioned subtle characteristics of the dispersion relation $\gamma(k)$, which determines *linear* evolution of amplitudes U_k, but also on fine details of *non-linear* mode coupling that arise from the expansion of vertices in powers of deviation $k - k_c$ [56]. Certainly, in this case contrary to Eqn (8), quadratic non-linearity is not reduced to cubic because long-wavelength

† Summation in Eqn (54) must be stopped at a certain, sufficiently large number of terms $|n| \leqslant M$, similar to how it is done in building up a stationary solution of Eqns (49).

modes are now independent degrees of freedom and no longer adiabatically follow modes with $k \simeq k_c$.

It has been noted before that the dispersion equation describing the dependence $\sigma(k, q)$ is in the general case very complicated (see Ref. [56]). Therefore, it is not presented here. It is however remarkable that the *structure* of this equation is sufficiently universal to be obtained without calculations since it follows from the symmetry of the problem [55].

Indeed, we know that the dependence $\sigma(k, q)$ in the main (in $\sqrt{\varepsilon}$) approximation is defined by a cubic equation, which means that the required dispersion equation has the form

$$\sigma^3 + a_2(k,q)\sigma^2 + a_1(k,q)\sigma + a_0 = 0\,. \tag{55}$$

This is actually due to the fact that the perturbation spectrum (54) contains three spatial harmonics with wave numbers $-k+q$, q and $k+q$, respectively, and they correspond to small γ. Note, that since there are only two such harmonics ($-k+q$ and $k+q$) for the Swift–Hohenberg equation, the corresponding dispersion equation for the stability problem in the main approximation is quadratic in σ [see the discussion of Eqns (25)–(29)].

Furthermore, according to Eqns (51), (52), the two amplitudes, U_k and U_{2k}, of the examined steady solution of Eqn (46) may be regarded as real quantities without loss of generality. This feature of the solution means that in the x-representation $u(x)$ is an even function: $u(x) = u(-x)$. Finally, linearization of Eqn (49) about the solution (51), (52) at real U_k, U_{2k} leads to the equation for amplitudes V_{nk+q} with *real* coefficients; hence the coefficients $a_{0,1,2}(k,q)$ in Eqn (55) are real too. Moreover, the problem of $u(x, y)$ evolution must also be invariant with respect to the transformation $x \to -x$, since $u(x)$ is even and the initial non-linear problem which defines evolution of δu is symmetric with respect to spatial reflections. It has been noted earlier that the reflection symmetry in the x-representation corresponds, for Fourier transforms, to invariance of the problem with respect to the transformation $q \to -q$, which means that the coefficients $a_{0,1,2}(k,q)$ are *even* in q.

Note also, the dependence of these coefficients on q originates from differentiation of Eqn (54) with respect to x (which yields only *integer powers* of q) and evaluation of the corresponding determinant. Therefore, we conclude that $a_{0,1,2}$ must be rational functions of q^2 and may be expanded in the Taylor series at $q = 0$ (since *all* the coefficients $a_{0,1,2}$ must remain finite at $q = 0$).

The last thing to be taken into account is the presence of two Goldstone modes in the stability problem, which correspond to two continuous symmetry groups: spatial translations ($x \to x + \text{const}$), and order parameter translations ($u \to u + \text{const}$). Due to this, Eqn (55) must have a *double root* $\sigma_{1,2} = 0$ at $q = 0$.

It follows from the above that in the limit of long-wavelength perturbations ($q \to 0$), Eqn (55) must have the form

$$\sigma^3 + b_2(k)\sigma^2 + b_1(k)q^2\sigma + b_0(k)q^2 = 0\,. \tag{56}$$

Building up its solution as a power series in q, it is easy to find that two Goldstone branches of the perturbation spectrum are described by the relation

$$\sigma_{1,2} = \pm q\sqrt{-\frac{b_0}{b_2}} + \frac{q^2}{2b_2}\left(\frac{b_0}{b_2} - b_1\right) + O(q^3)\,. \tag{57}$$

Eqn (57) immediately leads to the conclusion that stability of the problem is essentially dependent on the sign of the ratio $b_0(k)/b_2(k)$. When the ratio is negative, sign '+' in front of the radical in Eqn (57) corresponds to an unstable branch of the spectrum, with σ for this branch being *purely real* (monotonic development of instability).

Conversely, if $b_0(k)/b_2(k) > 0$, the first term on the right-hand side of Eqn (57) is purely imaginary and stability of the solution depends on the *correction* to this term.

In this case, stability conditions are reduced to the following obvious inequalities

$$\frac{b_0(k)}{b_2(k)} > 0\,, \tag{58}$$

$$\frac{1}{2b_2(k)}\left[\frac{b_0(k)}{b_2(k)} - b_1(k)\right] < 0\,, \tag{59}$$

which must be satisfied simultaneously. Evidently, these two conditions define the two boundaries of the Busse balloon. Eqns (57)–(59) imply oscillatory relaxation of perturbations inside the balloon. If the condition (58) fails to be fulfilled, unstable perturbations monotonically grow in time whereas they show oscillatory growth when condition (59) is violated (cf. properties of the perturbation spectrum in the case of convection with free-slip boundaries discussed in the beginning of Subsection 3.2).

Certainly, conditions (58) and (59) may be incompatible. In such a case, a Busse balloon does not exist at all, i.e., all steady spatially periodic solutions of the problem are unstable. The particular expressions for coefficients $b_{0,1,2}$ obtained in Ref. [56] for the solutions (51)–(53) result in the conclusion that the solutions are stable against long-wavelength perturbations in the narrow domain

$$\frac{91}{144} < \frac{\varkappa}{\varepsilon} < \frac{11}{12}\,, \tag{60}$$

where $\varkappa$ is the introduced earlier deviation of the wave number k of the solution (51)–(53) from unity: $\varkappa = k - 1$. Note a sharp contraction of the stability region compared to the Eckhaus criterion [see Eqn (37)].

However, stability against long-wavelength perturbations does not mean stability with respect to arbitrary perturbations. Generally speaking, a band of unstable perturbations can be separated from Goldstone modes by a finite gap. A detailed analysis of the stability problem regardless of the limitations imposed by the condition $q \to 0$ shows that it is true in the examined case: instability persists in the region defined by the inequality (60), but turns out to be associated with perturbations corresponding to finite values of q [56]. It should be emphasized that such a property of the perturbation spectrum constitutes another qualitative distinction of the problem with additional symmetry from the Swift–Hohenberg equation in which instability is always associated with the Goldstone branch of the spectrum [see Eqn (32)].

Note also that the above estimations of $\sigma \sim \varepsilon$ and $q \sim \sqrt{\varepsilon}$ are valid only at $\varkappa \sim \varepsilon$. On the other hand, solutions (50) exist at any value of k from the interval $k_1 \leqslant k \leqslant k_2$ to which $\varkappa \sim \sqrt{\varepsilon}$ corresponds. For such $\varkappa$ values the dependence of σ on ε proves to be even more complicated and, generally speaking, is not reduced to the proportionality of σ to a certain fixed ε [56]. In the context of the stability problem values $\varkappa \sim \sqrt{\varepsilon}$ are of no special interest because all such solutions are unstable already with respect to long-wavelength perturbations [see Eqn (60)]. However, the above complicated dependence of σ on ε indicates that, unlike the case where the Swift – Hohenberg equation may be reduced to the Ginzburg – Landau equation in the form (23), the problem can not be reformulated in such a way as to allow ε to be excluded by a simple scale transformation of variables [57, 61]. This in turn provides grounds to expect complicated spatiotemporal evolution of the solution associated with interplay of different characteristic scales at any positive ε, as small as desired [55, 57]. Indeed such evolution takes place and will be discussed in the next Section.

Ending this Section it seems relevant to note the following. Although all steady spatially periodic solutions of Eqn (46) turn out to be unstable, it may be regarded as a specific property of this particular equation – other equations with an additional continuous *one-parameter* group of symmetry may still possess stable solutions of such a type. Let us discuss how the general situation changes when the problem has *two* additional one-parametric groups of symmetry, apart from the conventional translational symmetry. In this case, following the same line of reasoning as before, it is easy to conclude that in the relevant pattern stability problem the dispersion equation has a *triple* zero $\sigma_{1,2,3} = 0$ at $q = 0$, which corresponds to *three* Goldstone branches of the spectrum. Then, in the long-wavelength limit, this equation has the from

$$q^2 c_0(k) + q^2 c_1(k)\sigma + q^2 c_2(k)\sigma^2 + \sigma^3 + \ldots = 0\,, \qquad (61)$$

where dots denote higher powers of σ. At small σ, Eqn (61) has the form

$$\sigma_{1,2,3} = (-q^2 c_0)^{1/3} \exp\left(\mathrm{i}\frac{2\pi m}{3}\right), \qquad m = 0, 1, 2\,. \qquad (62)$$

Evidently, the real part of at least one of the roots in Eqn (62) is positive at any sign of $c_0(k)$, which means instability of all stationary spatially periodic solutions in the problem with such symmetry.

However, it needs to be emphasized that in spite of the apparent simplicity of these arguments they are applicable only within a strictly defined region and any digression may be fraught with serious mistakes. An example of such incorrect generalization is provided by the stability problem of a *two-dimensional* pattern of square cells with only the conventional (but two-dimensional, i.e., two-parameter) group of symmetry with respect to spatial translations. At first sight, it appears that since two independent transformations of this group, i.e., translations along directions x and y, give rise to two independent Goldstone modes, this problem is reduced to that discussed earlier in this paper, and the criterion of stability with respect to long-wavelength perturbations must be defined by inequalities (58), (59).

Nevertheless, the structure of the dispersion equation is altogether different from that defined by expression (56), because nullification of the free term in the dispersion equation leading to the appearance of Goldstone modes occurs at both $q_x \to 0$, $q_y \neq 0$ and $q_y \to 0$, $q_x \neq 0$, where $q_{x,y}$ are the corresponding projections of the perturbation wave *vector*. Thus, the equation describing Goldstone branches of the spectrum in the long-wavelength limit has to have the following form†:

$$q_x^2 q_y^2 g_0 + [q_x^2 g_{1x} + q_y^2 g_{1y}]\sigma + g_2\sigma^2 + \ldots = 0\,. \qquad (63)$$

Here coefficients $g_{0,1x,1y,2}$ are independent of the perturbation wave vector. Certainly, properties of the dispersion relation defined by Eqn (63) differ dramatically from those of $\sigma(k, q)$, which follow from Eqn (55) and have been discussed earlier.

4. Soft turbulent modes and 'continuous' transition to chaos

A natural question that arises from instability of *all* stationary spatially periodic solutions of Eqn (46) is 'What is to come of this ?' In other words, what is the asymptotic state into which the unstable trivial solution $u \equiv 0$ evolves at $t \to \infty$? Since analytical study of such a problem is hardly possible, to answer this question it is natural to employ numerical methods. It is worth noting, however, that the problem (46), similar to the majority of problems with additional symmetry, is very difficult for numerical integration. Besides trivial reasons, such as approximation to high-order derivatives by finite differences, these difficulties are closely connected with intrinsic properties of the problems, e.g., with its symmetry. The following two of these 'internal factors' need to be mentioned in the first place. To begin with, the aforementioned scale mixing and the impossibility to reformulate the problem scaling out small ε result in simultaneous presence in the problem of ε-independent characteristic small scales of order of unity and characteristic large scales diverging at $\varepsilon \to 0$. A proper description of such a problem at small scales requires integration in small spatio-temporal steps. At the same time, integration for the description of the *asymptotic* state of the system must cover time intervals substantially wider than $1/\varepsilon$, which imposes strict constraints on the stability of the differential scheme and leads to the enormous expenditure of machine time which is the greater the smaller ε.

Secondly, long-wavelength modes play a very important role in the problem being examined as was noted more than once in the previous Sections. However, in any computer simulation the problem on an infinite straight line $-\infty < x < \infty$ is replaced by a problem on a finite segment $0 \leqslant x \leqslant L$ (usually with periodic boundary conditions at the ends), which cuts the spectrum at wave numbers of order $2\pi/L$ ($2\pi/L$ is the *exact* spectrum boundary from below for the periodic boundary conditions). It is essential that the value of $2\pi/L$ must be compared with $1/\varepsilon$ [see Eqn (60)], rather than with unity or $1/\sqrt{\varepsilon}$ (as is the case of the Swift – Hohenberg equation [see Eqn (22)]. Therefore, the L value at small ε must be very large, which also greatly contributes to the expenditure of machine time.

One of the earliest numerical studies on the short-wavelength instability in problems such as (44), (46) has been reported in Ref. [65]. But this paper considered the generalization of Eqn (44) rather than the equation itself, with the linear differential operator supplemented with the third derivative in x. This difference is crucial, for it does not only account for the complex dispersion dependence $\gamma(k)$, but also breaks symmetry of the problem with respect to spatial reflection ($x \to -x$). Ref. [65] discusses evolution of a small initial perturbation in the form of white noise. Evolution leads to the appearance of a spatial pattern interpreted by the authors as stationary one. However, their results admit of another interpretation. Dependences $v(x)$ reported in Ref. [65] as examples of stationary asymptotic states have clear-cut

† A detailed discussion of the stability problem for cellular patterns in dissipative systems can be found in Refs [49, 62, 63], and for Hamiltonian systems in Ref. [64].

long-wavelength modulation which may be regarded as the onset of long-wavelength instability. Unfortunately, difficulties encountered by the authors made it impossible to advance farther in time than a few inverse ε^{-1}, and this is not enough to judge the asymptotic states for the reasons given above. It appears from Ref. [65] that the value of L in this simulation conformed to the condition $k_c L/(2\pi) \sim 10$, which is not quite satisfactory either.

The numerical study of Eqn (46) presented in Refs [55–57] yielded more relevant data. The simulation was performed using a CRAY C-90 supercomputer and the differential scheme specially designed to integrate equation (46) at extremely small ε values [66], which enabled the authors to advance in time as far as $t \sim 10^3/\varepsilon$. The scheme was thoroughly tested by comparing the results of simulations with the corresponding analytical expressions obtained in Ref. [56] by the methods of the theory of perturbations.

The simulation provided the following results. Regardless of smallness of ε (the smallest examined ε was 10^{-4} at L large enough (typical values of $k_c L/(2\pi)$ at $\varepsilon = 10^{-4}$ were about 10^3) evolution of small initial perturbations invariably resulted in a chaotic regime with very unusual properties. This regime is characterized by excitation of a large number of modes (apparently, up to continuum at $L \to \infty$) concentrated in small vicinities of the zero wave number (Goldstone branch) and $k = k_c$ (short-wavelength instability-associated branch) supplemented by their spatial satellites (see Fig. 3). At the same time, the amplitudes of these modes, being of order of $\sqrt{\varepsilon}$, undergo chaotic time-dependent variations (Figs 4 and 5). Such dynamics exhibits all features of developed chaos including exponential separation of adjacent trajectories in the phase space and exponential decay of autocorrelation functions [57]. It is noteworthy that characteristic values of amplitudes of long-wavelength modes (which in the linear problem correspond to the slowly varying but *stable* branch of the spectrum with $\gamma(k) < 0$ in

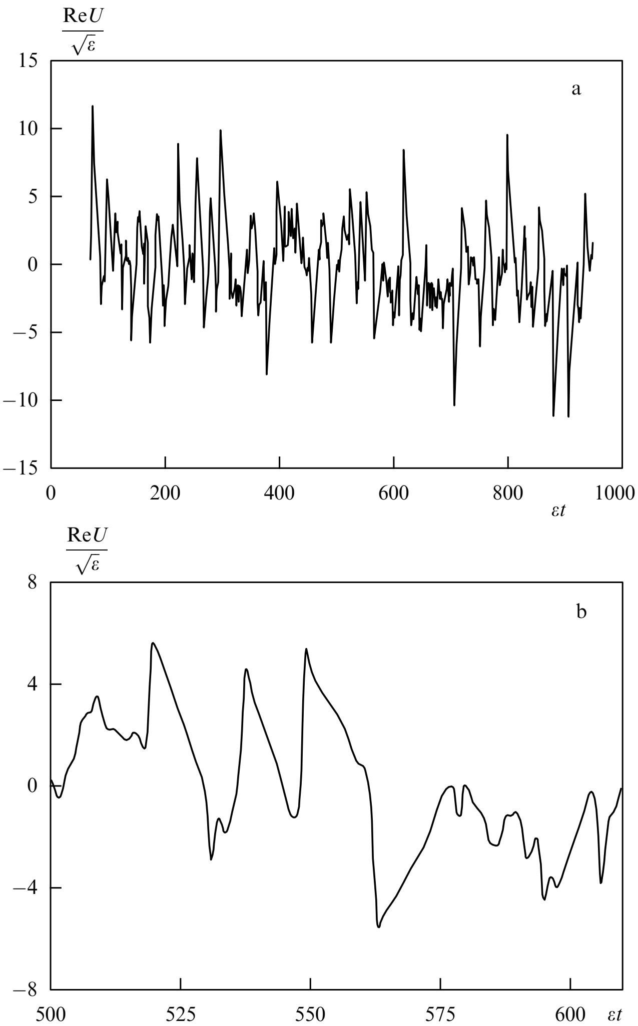


Figure 4. Temporal evolution of the real part of the amplitude of a mode with $k = 3.125 \times 10^{-3}$ corresponding to Fig. 3, (a), and a selected fragment of this curve suggesting the 'fine structure', (b), (from Ref. [57]).

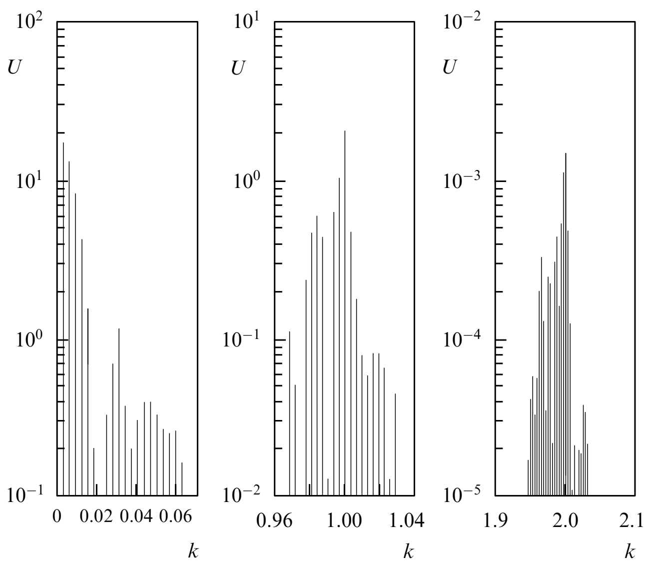


Figure 3. Typical dependence of the amplitude of spatial harmonics in the solution of Eqn (46) on their wave number at a fixed moment after the asymptotic state is achieved. Numerical integration with periodic boundary conditions was performed at $\varepsilon = 10^{-4}$. The distance Δk between two adjacent modes related to spectrum discretization due to finiteness of L is 3.125×10^{-3}. Each of the three parts of the spectrum is presented in its own scale (from Ref. [55]).

the vicinity of $k = 0$) turn out to be greater than those for the modes concentrated in the vicinity of $k = 1$ which are directly related to short-wavelength instability: $\gamma(k) > 0$ at $k = 1$ (Fig. 3).

A conspicuous feature of the problem is that chaotic evolution of each mode has a specific 'fine structure' with a certain characteristic time $\tau(k)$ which actually determines the time of decay of the autocorrelation function (correlation time) (Figs 4b, 5b). At a fixed ε, the value of $\tau(k)$ increases with decreasing k (cf. Figs 4b and 5b). It is however essential that all these characteristic times are not of order of unity as might have been expected, but of the order of $1/\varepsilon$; in other words, they *diverge* at $\varepsilon \to 0$.

All this indicates that at $\varepsilon = 0$ the system undergoes bifurcation which leads to *direct* transition of the 'quiescent state' ($u \equiv 0$) to chaotic dynamics. In this process, the corresponding turbulent† modes are *soft*, i.e., the 'rest-to-chaos' transition is continuous. Of course, the word *continuous* in this context does not imply that asymptotic behavior of the system *gradually* (?) becomes chaotic in a certain

† In this review, the term 'turbulence' is used in the broad sense to describe any spatio-temporal chaos.

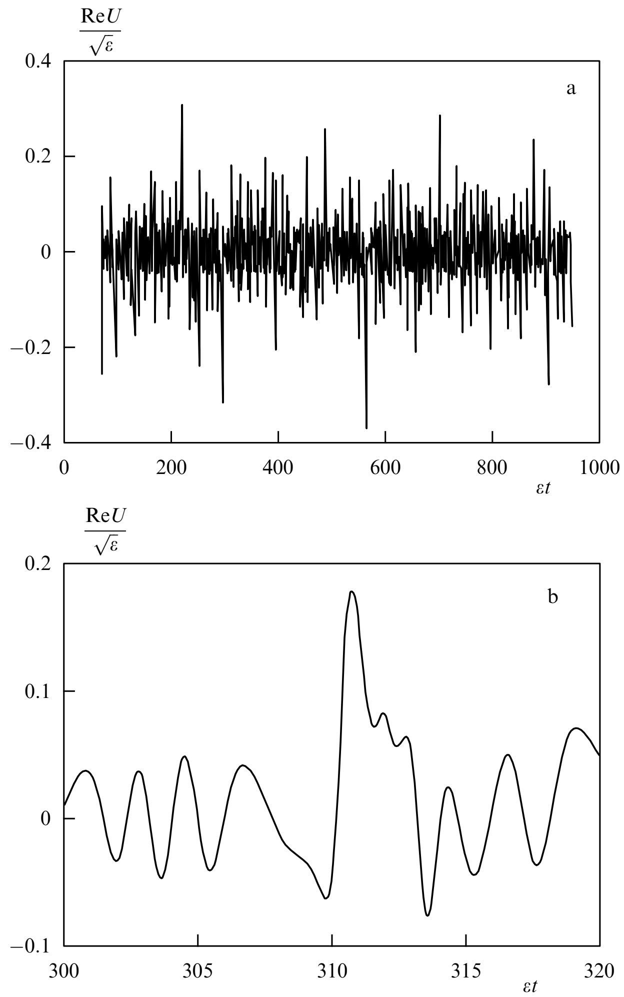


Figure 5. Temporal evolution of the real part of the amplitude of a mode with $k = 1$ corresponding to Fig. 3, (a), and a selected fragment of this curve suggesting the 'fine structure', (b), (from Ref. [57]).

mysterious way. This would be absurd. Chaos is either present or absent. It appears jumpwise at $\varepsilon = 0$. But *quantitative parameters* of this chaos, such as the characteristic amplitudes of turbulent modes or characteristic correlation time (as well as characteristic time of transition to the asymptotic regime), merge to the quiescent state in a continuous manner through vanishing of the amplitudes and divergence of the characteristic time (*critical slowing down*) at $\varepsilon = 0$. In this context, the word 'continuous' has the same sense as that set by Landau when he applied this attribute to second order phase transitions.

An additional evidence that turbulent modes in this problem are soft is absence of any hysteresis in the 'adiabatic' scanning of ε from small positive towards small negative values and in the opposite direction (see Ref. [57]).

The simulations also show that the problem is highly sensitive to cutting off the long-wavelength part of the spectrum due to the finiteness of L. If L is not large enough, the asymptotic state corresponds to strongly non-linear, but periodic oscillations of amplitudes of the modes, rather than to chaos. For example, such oscillations at $\varepsilon = 10^{-4}$ were possible to observe even at $k_c L/(2\pi) = 50$. This gives reason to believe that phase space of the problem contains limiting cycles (cycle) stable in certain directions and unstable in others. When L is sufficiently large, chaotic dynamics corresponds to random wandering of the phase trajectory which is in fact an attraction along stable directions followed by repulsion along unstable ones. In such a case reduction of the phase space caused by decrease of L can stabilize the limiting cycle(s), giving rise to the periodic dynamics mentioned above [57].

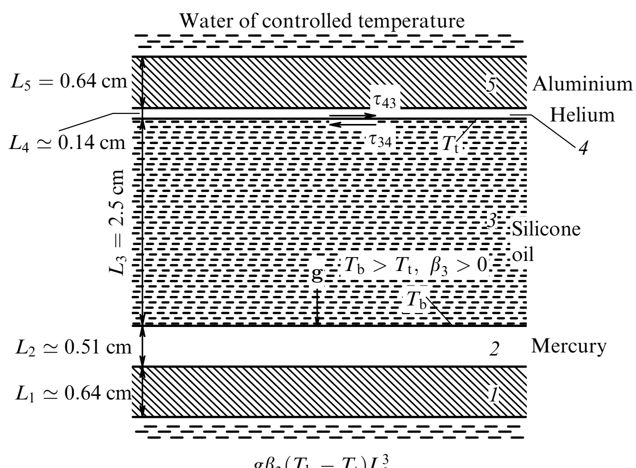
5. Experiment. Turbulence with zero critical Reynolds number

'Symmetry, Goldstone modes, numerical simulation. All this is very good, but what does it have to do with real physics? All kind of equations may be thought of but few are likely to withstand experimental validation.'

These are thoughts which are sure to occur to the reader who has worked through the text as far as this Section. In fact, what can be learnt from the experiment?

Ref. [16] was one of the first attempts to realize experimentally a system with short-wavelength instability and additional symmetry for the case of Rayleigh – Benard convection in a layer with free-slip boundaries. The free-slip boundary conditions were achieved by virtue of the original design of the apparatus in which a layer of silicone oil was separated from the rigid walls by a layer of liquid mercury from below and a helium layer (gaseous phase) from above (Fig. 6). Due to the drastic difference between coefficients of kinematic viscosity of helium and mercury on the one hand, and silicone oil on the other hand (the corresponding ratios are 2.1×10^{-6} and 1.6×10^{-4}, respectively), such a system provides a good model of free-slip boundaries for the test oil layer. Indeed, the critical Rayleigh number corresponding to the onset of convective motion measured in Ref. [16] is in excellent agreement with that predicted by Lord Rayleigh himself for convection with free-slip boundaries [9, 10].

Unfortunately, the study [16] had been carried out long before Siggia and Zippelius took notice of the role of vertical vorticity in the destabilization of stationary patterns during convection with free-slip boundaries [11]. Although the authors of Ref. [16] mentioned that they made all measurements after the asymptotic regime of convection was established, they did not describe this regime in detail, so it was not



$$Ra = \frac{g\beta_3 (T_b - T_t) L_3^3}{\nu_3 \alpha_3}$$

Figure 6. Cross-sectional scheme of the test apparatus to study thermal convection with the free-slip boundary conditions (from Ref. [16]).

clear whether the regime corresponded to steady spatially periodic or to chaotic states.

Another example of systems with the desired additional symmetry is provided by electroconvection in a layer of a homeotropically oriented nematic which has been extensively used as a model in recent experimental studies [22–24, 67, 68].

These experiments were almost simultaneously initiated by two different groups which will be henceforth referred to for brevity as European [22, 23] and Japanese [24, 67, 68]. The two groups designed their experiments in a similar way. A reference nematic, *p*-methoxybenzilidene-*p*′-*n*-butylaniline (MBBA), with negative dielectric anisotropy was used, which allowed the onset of convective motion to be preceded by the Freedericksz transition. Convection was studied at small values of the control parameter ε derived from the condition

$$\varepsilon = \frac{V^2 - V_c^2}{V_c^2} ,$$

where V_c is the critical voltage applied to the sample and responsible (at a given oscillation frequency f) for convection instability. Degeneracy associated with the Freedericksz transition was removed (if required) by a magnetic field applied at the layer plane to create a singled out direction in this plane. The results obtained by the two groups are in agreement. Minor differences unessential in the context of this review can be accounted for by different values of the material constants of the nematic and other experimental parameters. However, the Japanese group, as opposed to the European one, reported not only the *observed* patterns, but also important *quantitative* information about their time spectra, which is crucial for the comparison with the theories described in previous Sections. Therefore, further discussion will be focused on the results obtained by the Japanese authors.

The most important details of the experimental design are as follows. The MBBA layer was 50 μm in width and had lateral size of 1×1 cm^2. The homeotropic orientation of the nematic was achieved by the treatment of the glass plate surfaces between which the layer was enclosed with a surfactant, *n*-*n*′-dimethyl-*n*-octadecyl-3-aminopropyl-trimethoxysilyl chloride (DMOAP). Nematic conductance was $\sigma_{\|} = 3.30\times\ \times 10^{-7}\,\Omega\,\mathrm{m}^{-1}$ and $\sigma_{\perp} = 2.34 \times 10^{-7}\ \Omega\,\mathrm{m}^{-1}$ and was monitored by the addition of 0.012 weight percent of tetra-n-butylammonium bromide (TBBA). Dielectric constants were $\epsilon_{\|} = 4.21$ and $\epsilon_{\perp} = 4.70$. A sample was placed in a thermostat at $30 \pm 0.02\,^{\circ}$C. The most systematic study was carried out at the frequency of the applied voltage equal to 500 Hz (the corresponding $V_c = 8.34$). The threshold Freedericksz transition V_F was 3.92 V. As soon as convective motion was apparent, spatial modulations of the director (a unit vector $\mathbf{n}$ indicating the direction of the predominant molecular orientation) led to corresponding variations of refractive index which facilitated visualization of the convection patterns. All images were recorded on magnetic tapes and disks for further computer-assisted analysis.

The experiment was carried out as follows. Firstly, the voltage applied was jumpwise increased from zero to 6.00 V (higher than V_F but lower than V_c) to trigger the Freedericksz transition in the non-convective state. The sample was then left at this state for 50 min to ensure that the angle between the director and the vertical axis (associated with the Freedericksz transition) had time to relax to its equilibrium value. Next, the voltage was again raised stepwise and then fixed at a certain value above the convection instability threshold. Measurements were made after a 50-min pause from the moment of this voltage jump to avoid influence of transient processes. Upon completion of measurements (that normally took about 5 min), the voltage was removed and the sample left for 50 min for reset to erase 'memories' of the experiment. Then the procedure was repeated.

Convection patterns observed in this experiment were *chaotic* both in space and time at all values of ε including minimal experimentally available. The patterns underwent slow evolution in time and showed no tendency to reach any stationary state. A typical pattern at a fixed moment of time is shown in Fig. 7a. Interestingly to note that the pattern retains characteristic 'micro'-scale (short-range order) related to the critical wave number of short-wavelength instability while the development of chaos is associated with long-wavelength modulations of the 'micro'-relief (absence of long-range order).

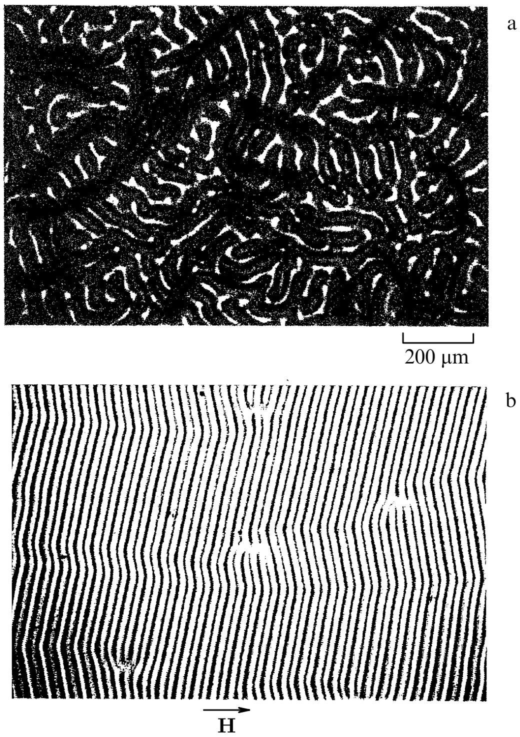


Figure 7. Typical turbulent pattern near the electroconvection threshold in a homeotropically aligned liquid crystal at a fixed moment $\varepsilon = 0.1$, $V = 8.75$ V, (a), and its stabilization, (b), by a magnetic field applied to the layer plane, $H = 1600$ Gauss, $V = 10.07$ V. Taking into account variations of V_c caused by the magnetic field, this corresponds to $\varepsilon = 0.1$. See text for details (from Ref. [67]).

What should be expected if a similar experiment is carried out in the presence of a magnetic field applied to the layer plane? Such a field generates anisotropy in this plane, i.e., reduces the symmetry of the problem to conventional. This, in turn, must lead to the appearance of spatially periodic patterns provided the theoretical background is relevant. It has been experimentally confirmed that application of the magnetic field indeed resulted in the stabilization of chaos and the appearance of long-range order in the system (Fig. 7b). The pattern shown in this figure (the so-called *oblique rolls*) is typical of electroconvection at onset in case of the *planar* orientation of the director (the director is parallel to the layer plane, i.e., the problem is originally anisotropic in this plane) [17 – 20]. It is essential that stabilizing effect of the magnetic field on chaotic dynamics is *reversible*. That is, removal of the field restores additional symmetry of the problem and leads to spontaneous destruction of stationary spatially periodic patterns; as a result, the system is brought back to chaotic dynamics shown in Fig. 7a. The latter confirms that spatio-temporal chaos is not a protracted trivial transitional relaxation to the spatially uniform state of the initial (spatially non-uniform at large distances) azimuthal orientation of the director projection upon the layer plane, but is rather an intrinsic property of the problem [23, 24, 67, 68].

The following approach was used to obtain temporal characteristics of the observed turbulence. An arbitrary line of the image served as the x axis. The brightness of different points at this line was loaded into computer in the form of a space-time profile $\xi(x,t)$ (Fig. 8). The local autocorrelation function,

$$\phi_x(t) = \left\langle \left[\xi(x, t+t') - \left\langle \xi(x,t') \right\rangle\right] \left[\xi(x,t') - \left\langle \xi(x,t') \right\rangle\right] \right\rangle , \tag{64}$$

was calculated for each point, where angular brackets denote averaging over time t' $(0 \leqslant t' \leqslant T$ here, T is the total observation time).

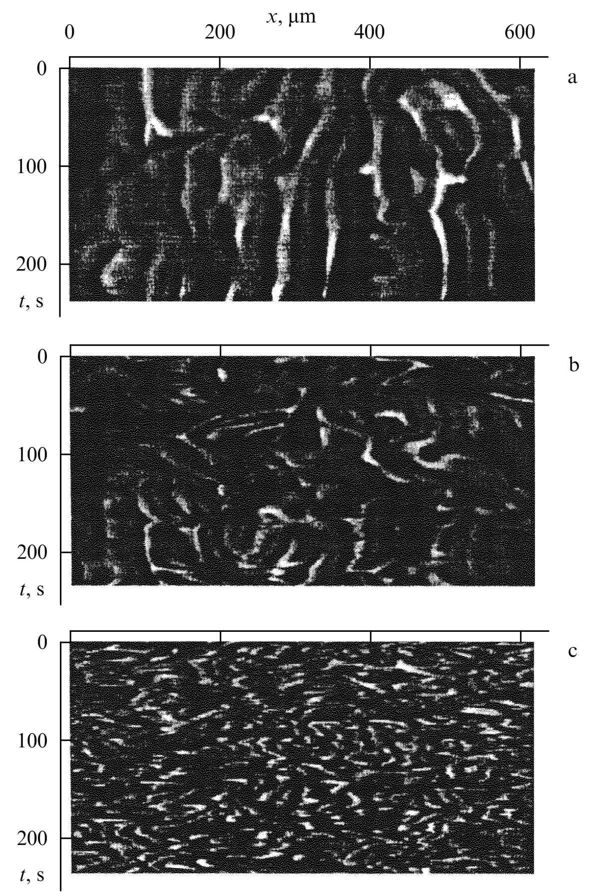


Figure 8. Space-time intensity profiles of an arbitrary line of the image. Spatial resolution is 1.27 μm pixel$^{-1}\times$ 512 pixels. Time resolution 1.0 s pixel$^{-1}\times$256 pixels; $\varepsilon = 0.010$, (a); $\varepsilon = 0.24$, (b); $\varepsilon = 0.69$, (c) (from Ref. [67]).

In the end, the resulting set $\phi_x(t)$ was averaged over x. The autocorrelation functions $\phi(t)$ obtained by this approach at different ε values are shown in Fig. 9. It can be seen that autocorrelation functions rapidly decay regardless of ε values, which is typical of a developed chaos.

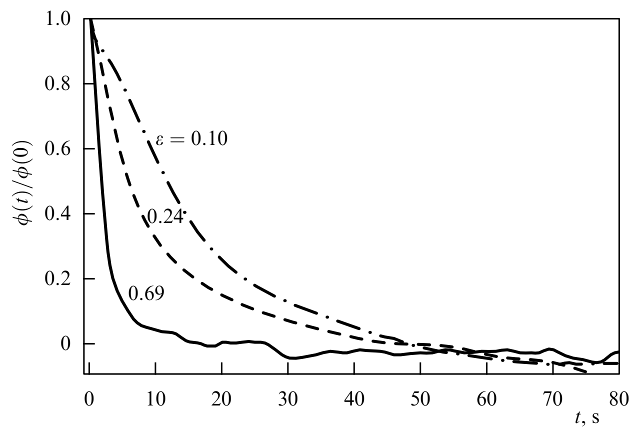


Figure 9. Autocorrelation functions corresponding to space-time profiles in Fig. 8. See text for details (from Ref. [68]).

Characteristic correlation time τ was estimated by the approximate representation of experimentally obtained functions $\phi(t)$ in the form

$$\phi(t) = \phi(0) \exp\left(-\frac{t}{\tau}\right) , \tag{65}$$

where the value of $\phi(0)$ was found in the experiment and τ was the only adjusting parameter determined by the least square method. The obtained dependence $\phi(0)$ is shown in Fig. 10. Experimental points fit well to the straight line $1/\tau = \mathrm{const}\,\varepsilon$.

Another important characteristic of a random quantity is its spectral density. The spectral density of a quasistationary random process is not an independent quantity and may be expressed in terms of Fourier components of the autocorrelation function [28]. Therefore, comparison of these two quantities provides a good criterion for checking up all propositions underlying the present discussion. At the same time, the expression relating spectral density of a random quantity to its autocorrelation function was obtained in [28]

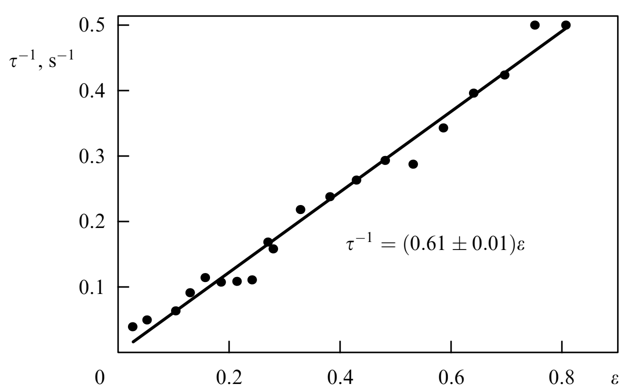


Figure 10. Plot of correlation time vs ε. See text for details (from Ref. [68]).

essentially on the grounds on the ergodicity hypothesis, whose application to fluctuations in non-equilibrium systems requires a special substantiation. Moreover, Ref. [28] considers the case of $T \to \infty$, whereas the transition from finite to infinite observation time in the given problem is not trivial. For this reason, it appears appropriate to describe the derivation of this relation taking no account of ergodicity of the process and relevant at finite values of T (see Appendix). With the assumed form of $\phi(t)$ [see Eqn (65)] and ε-dependence of the correlation time $\tau = (\Lambda\varepsilon)^{-1}$, where $\Lambda = \mathrm{const}$, the formula (A.7) leads to the following expression:

$$\frac{\varepsilon S_\omega}{T\phi(0)} = \frac{2\Lambda}{\Lambda^2 + \Omega^2} \,. \qquad (66)$$

Here $\Omega \equiv \omega/\varepsilon$, and S_ω is the absolute value of Fourier coefficients for spectral expansion of time-dependence of signal intensity averaged over x in the same manner as that in the case of local autocorrelation functions (subscript n in ω_n in Eqn (66) is omitted, for brevity).

It is worthwhile to note that the right-hand side of Eqn (66) is a universal ε-independent function and that $\phi(0)$ and Λ in Eqn (66) were deduced from examining the autocorrelation functions, while ε and T are given by the experimental conditions; therefore, the expression (66) does not contain any adjusting parameters.

Spectral density of signal brightness measured experimentally at different ε is compared to the expression (66) in Fig. 11. Evidently, the measured values and Eqn (66) are in excellent agreement.

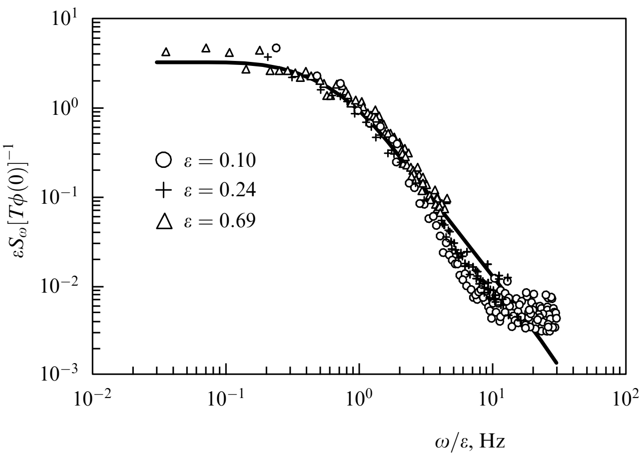


Figure 11. Power spectrum of brightness of the turbulent pattern images at different ε. Solid line — the right-hand side of dependence (66).

To summarize, experiments confirmed all qualitative inferences and specific features of the transition to chaos in systems with additional symmetry as well as characteristic scale relations following from the theoretical analysis of the phenomenon in previous Sections.

Unfortunately, neither theoretical nor quantitative experimental data on *spatial* properties of chaos are currently available. Nevertheless, on the grounds of the qualitative description of the patterns in [22–24, 67, 68] it may be concluded that the observed chaos is of *spatiotemporal* nature and arises at $\varepsilon = 0$, i.e., at the onset of any hydrodynamic motion. In other words, what is observed is a *turbulence with the critical Reynolds number equal to zero*.

In the end, a great deal of progress has been achieved in development of the quantitative weakly non-linear (valid at small ε) theory of electroconvection in a layer of a homeotropically aligned nematic [69, 70]. Numerical integration of the equations performed at the values of parameters close to the experimental ones [22, 23] leads to spatio-temporal chaos similar to the patterns observed in Refs [22, 23]; processing of these data also yields ε-dependence of correlation time of the form $\tau \propto 1/\varepsilon$. However, the equations derived and studied in Refs [69, 70] are very complicated, and their discussion is beyond the scope of the present review, being of interest for narrow experts rather than for a broad audience†.

6. Conclusion

Usually, Conclusion is intended to sum up the results and discuss prospects of further studies. In our case, this is hardly reasonable. Any summary would be premature since neither theory nor experiment provided enough data for the purpose. The same is true of prospects; in fact, any study should be encouraged if it promises a breakthrough in the area in question. At present, it appears safe to assert only that, firstly, the behavior of different systems with short-wavelength instability and additional symmetry has much in common and, secondly, this behavior is so much unlike the conventional one that such systems ought to be undoubtedly regarded as a special class of pattern-forming systems. Their study would not only enlarge our knowledge about the origin of order and chaos, but also stimulate revision of many deep-rooted ideas.

In conclusion, I would like to make a few general remarks on both the similarity and the difference between the problems reviewed in this paper and those concerning thermodynamics and kinetics of phase transitions.

To begin with, there is an analogy between these two problems, which inevitably rises in mind and was noticed by many authors (see Refs [2–4] and references therein); also, it was many times emphasized in the present review (the role of symmetry, expansion in powers of the order parameter, etc.). However, this analogy is usually 'strongly anisotropic', i.e., directed to the transfer of ideas from the nature and well-known field of phase transitions to the younger and therefore less-developed area of pattern formation in dissipative

† It should be emphasized that the case of soft-mode turbulence in electroconvection of a homeotropically aligned layer is a little bit more complicated than it could seem from the above discussion. The point is that there are two different types of perturbations of a spatially uniform state, viz (i) perturbations whose wave vector $\mathbf{k}$ is orthogonal to the director projection on the midplane $\mathbf{n}_\parallel$ — the so called normal rolls, and, (ii), all the rest mutual orientations of $\mathbf{k}$ and $\mathbf{n}_\parallel$ — the so called oblique rolls. In case (i) at small ε the governing equations admit weakly non-linear steady solutions of the normal-roll-type [21, 69, 70]. All these solutions are unstable under very general conditions [21, 69, 70], so the situation is indeed similar to that of Refs [55–57]. As for the case (ii), the corresponding steady solutions do not exist since the torque on $\mathbf{n}_\parallel$ exited at the oblique orientation cannot be compensated [21, 69, 70]. On the other hand, in the discussed experiments [24, 67, 68] the most unstable perturbations of the spatially uniform state belong to the case (ii), so strictly speaking, the analogy with the problem discussed in Refs [55–57] may be irrelevant. Nevertheless, preliminary experiments mentioned in Ref. [68] show that change of the type of the most unstable modes from (ii) to (i) caused by an appropriate change of the frequency of the AC-voltage does not make much difference in the pattern dynamics. (*Note added in English proof.*)

systems. I would like to change the sign of this anisotropy and discuss which properties of dissipative systems with additional symmetry are likely to emerge in the field of phase transitions and how they can be recognized.

One of the most important attributes of dissipative systems with additional symmetry is destabilization of stationary spatially periodic patterns resulting from their coupling with long-wavelength modes of the Goldstone branch of the spectrum. It is essential that destabilization may under certain conditions be responsible for the complete disappearance of stable spatially periodic solutions of the problem. Another important aspect is mixing of ε-scales, which makes it inadequate to consider the problem in the lowest order of perturbation theory.

Formation of stationary spatially periodic patterns close to onset of short-wavelength instability obviously corresponds in phase transition phenomena to the transition with a spontaneous *reduction* of symmetry from a certain highly symmetric (e.g., isotropic) phase I to a less symmetric phase II. In the framework of the analogy being examined if such a phase transition problem has an additional symmetry, the interaction between long-wavelength macroscopic fluctuations detaching from the corresponding Goldstone modes and phase II must result in its destabilization up to complete suppression of the phase transition in a certain range of values of the problem parameters.

A possible example of such a situation is constituted by the isotropic phase – cholesteric phase or nematic – smectic transition in a liquid crystal placed in a capillary. One of the types of long-wavelength fluctuations in such system is associated with excitation of hydrodynamic modes. On the other hand, it is well-known that capillary flow of a cholesteric or smectic liquid crystal with a sufficiently low speed resembles 'permeation' of molecules through helicoid (cholesteric) or lamellar (smectic) patterns of liquid-crystal order whose position is fixed due to interaction with the capillary walls (the so-called Helfrich mechanism [17, 18, 71]). For this reason, long-wavelength Helfrich modes in such a problem are analogous to the modes associated with vertical vorticity in the Rayleigh – Benard problem with free-slip boundaries; both cause fluid displacement relative to the small-scale pattern (liquid-crystal in the case under consideration and convective in the Rayleigh – Benard problem), which arises in this fluid [72].

Application of concepts and methods discussed in this review to investigation into stability of the cholesteric (smectic) phase with respect to 'Helfrich fluctuations' would provide an interesting example of the 'feedback' effect of the problems of the theory of dissipative pattern formation on the problems of the phase transitions.

On the other hand, it should be remembered that any analogy is only similarity but not identity. Therefore, the presence of certain properties in the phase transition problem only gives *reason* to expect similar properties in an analogous problem for dissipative patterns, but in no way warrants their existence. The role of additive (thermal) fluctuations in both problems is a good illustration of how dangerous the indiscreet use of such analogies may be.

The crucial role of fluctuations in a close neighborhood of phase transition points is widely known [28]. Sometimes fluctuations may even change order of phase transition from the second to the first [73]. It may seem that negation of fluctuations invalidates the conclusions of this review concerning behavior of dissipative systems in a close vicinity of the short-wavelength instability threshold. Nevertheless, such negation is fully justified. The fact is that patterns formed in dissipative systems are invariably of *macroscopic* scale unlike the situation in phase transitions, where spontaneous symmetry breaking is associated with the *atomic* scale ordering. This explains why the corresponding dimensionless parameter responsible for the width of the 'fluctuation' region (Ginzburg number) in dissipative systems always includes the characteristic atomic to macroscopic size ratio. As a rule, this causes such shrinking of the fluctuation region, that the question of fluctuation *effect* on system's dynamics becomes senseless [34]. It should be emphasized, to avoid misunderstanding, that I am speaking about the effect of fluctuations on the dynamics of short-wavelength instability, but not of the fluctuations themselves. The problem of examining fluctuations seems quite reasonable and was elegantly solved in Ref. [74] with the aid of short-wavelength instability as a sort of amplifier for thermal fluctuations.

Acknowledgements. I am indebted very much to M Mimura and T Ochiai for being invited to the University of Tokyo, where this study was carried out, and for the provision of excellent conditions to do research. I am grateful to S Kai for stimulating discussions of many problems scrutinized in the present review and to L Kramer for the opportunity to read his paper [70] prior to publication.

7. Appendix

In order to derive the expression relating the spectrum of a random quantity to its autocorrelation function, let us consider a real random quantity $\zeta(t)$ observed for some time $T(-T/2 \leqslant t \leqslant T/2)$ and satisfying the condition $\langle\zeta\rangle = 0$. Let us further define $\zeta(t)$ at the entire time-axis $(-\infty < t < \infty)$ periodically extending this function beyond the boundaries of the interval $-T/2 \leqslant t \leqslant T/2$ $(\zeta(t+T) = \zeta(t))$ and then expanding it into a Fourier series, which is convenient to write down in the form

$$\zeta(t) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \zeta_{\omega_n} \exp(\mathrm{i}\omega_n t)\,,$$

$$\omega_n = \frac{2\pi n}{T}\,, \quad \xi_{\omega_n} = \xi^*_{\omega_{-n}} = \int_{-T/2}^{T/2} \zeta(t)\exp(-\mathrm{i}\omega_n t)\,\mathrm{d}t\,. \quad \text{(A.1)}$$

Note that in this case ω_n satisfies the evident identity $\omega_n + \omega_m = \omega_{n+m}$. Then, in accordance with the definition of autocorrelation function, we have

$$\phi(t) = \big\langle \zeta(t+t')\zeta(t') \big\rangle$$
$$= \frac{1}{T^2}\sum_{n,m} \zeta_{\omega_n}\zeta_{\omega_m} \exp(\mathrm{i}\omega_n t)\big\langle \exp(\mathrm{i}\omega_{n+m} t')\big\rangle\,. \qquad \text{(A.2)}$$

The last expression is due to averaging over t' and the only t'-dependent quantity under the summation sign is $\exp(\mathrm{i}\omega_{n+m}t')$. On the other hand,

$$\big\langle \exp(\mathrm{i}\omega_{n+m}t')\big\rangle = \frac{1}{T}\int_{-T/2}^{T/2} \exp(\mathrm{i}\omega_{n+m}t')\,\mathrm{d}t' = \delta_{-nm}\,, \quad \text{(A.3)}$$

where δ_{-nm} is the Kronecker symbol. Substituting Eqn (A.3) into (A.2) and taking into account that $\xi_{\omega_{-n}} = \xi^*_{\omega_n}$, we have

$$\phi(t) = \frac{1}{T^2} \sum_n |\zeta_{\omega_n}|^2 \exp(\mathrm{i}\omega_n t) \,, \tag{A.4}$$

i.e.,

$$\phi_{\omega_n} = \frac{1}{T} |\zeta_{\omega_n}|^2 \,, \tag{A.5}$$

where ϕ_{ω_n} are coefficients of expansion of $\phi(t)$ into a Fourier series. Also, since ϕ_{ω_n} are real, we see that $\phi(t) = \phi(-t)$, as expected [28].

For an exponentially decaying autocorrelation function (65), direct integration at $T \gg \tau$ yields (according to general rules, in the exponent $|t|$ must be substituted for t if $t < 0$, see Ref. [28])

$$\phi_{\omega_n} \simeq \frac{2\tau\phi(0)}{1 + \omega_n^2 \tau^2} \,. \tag{A.6}$$

Comparison of Eqns (A.5) and (A.6) leads to the conclusion that the following relation holds for the autocorrelation function (65):

$$\frac{1}{T} S_{\omega_n} = \frac{2\tau\phi(0)}{1 + \omega_n^2 \tau^2} \,. \tag{A.7}$$

Here the notation $S_{\omega_n} = |\zeta_{\omega_n}|^2$ is introduced and by definition $\phi(0) \equiv \langle \zeta^2 \rangle$, [see Eqn (A.2)].

There are two remarks with respect to the relations (A.5), (A.7) and the method of their derivation. First, it may seem that deriving (A.5) we did employ besides the assumption of ergodicity (the problem of ergodic fluctuations in this case is totally immaterial since a *single* random quantity was considered) also the assumption of quasistationarity of the process, i.e., independence of statistical characteristics from a particular instant of the beginning of observations. However, it is not the case: *periodic extension* of the random function defined on a finite time-interval to the entire axis automatically guarantees the quasistationarity (the integral of the periodic function over its period is evidently independent of the position of the initial integration point).

The second remark concerns transition to the limit $T \to \infty$. Introducing $\Delta\omega \equiv 2\pi/T$, writing Eqn (A.1) as the integral sum, and passing to the limit, we find

$$\zeta(t) = \frac{1}{2\pi} \Sigma \zeta_{\omega_w} \exp(\mathrm{i}\omega_n t) \Delta\omega \underset{T\to\infty}{\longrightarrow} \frac{1}{2\pi} \int_{-\infty}^{\infty} \zeta_\omega \exp(\mathrm{i}\omega t) \, \mathrm{d}\omega \,,$$

where the formal limit ζ_{ω_n} at $T \to \infty$ is denoted as ζ_ω [see Eqn (A.1)]. Furthermore, denoting the limit ϕ_{ω_n} at $T \to \infty$ as ϕ_ω, we find that the expression (A.7) turns into the familiar relation given in [18]. It is however essential that the two limits, ζ_ω and ϕ_ω, generally speaking, can not be finite at the same time. Indeed, if ϕ_ω is assumed to be finite, it follows from Eqn (A.5) that $\zeta_{\omega_n} \propto \sqrt{T}$ and diverges at $T \to \infty$, which is actually due to the fact that $\zeta(t)$ does not tend to zero at $t \to \infty$.

References

1. Lifshitz I M *Zh. Eksp. Teor. Fiz.* **55** 2408 (1968) [*Sov. Phys. JETP* **28** 1280 (1969)]
2. Nicolis G, Prigogine I *Self-organization in Nonequilibrium Systems, from Dissipative Structures to Order through Fluctuations* (New York: Wiley, 1977)
3. Haken H *Advanced Synergetics* (Berlin: Springer-Verlag, 1983)
4. Cross M C, Hohenberg P C *Rev. Mod. Phys.* **65** 851 (1993)
5. Chandrasekhar S *Hydrodynamic and Hydrodynamic Stability* (Oxford: Clarendon Press, 1961)
6. Anisimov S I, Tribel'skiĭ M I, Epel'baum Ya G *Zh. Eksp. Teor. Fiz.* **78** 1597 (1980) [*Sov. Phys. JETP* **51** 802 (1980)]
7. Bernoff A J et al. *SIAM J. Appl. Math.* **55** 485 (1995)
8. Coullet P, Fauve S *Phys. Rev. Lett.* **55** 2857 (1985)
9. Lord Rayleigh *Proc. R. Soc. London A* **93** 148 (1917)
10. Lord Rayleigh *Phil. Mag.* **32** 529 (1916)
11. Siggia E D, Zippelius A *Phys. Rev. Lett.* **47** 835 (1981)
12. Zippelius A, Siggia E D *Phys. Rev. A* **26** 1788 (1982)
13. Zippelius A, Siggia E D *Phys. Fluids* **26** 2905 (1983)
14. Busse F H, Bolton E W *J. Fluid Mech.* **146** 115 (1984)
15. Bolton E W, Busse F H *J. Fluid Mech.* **150** 487 (1985)
16. Goldstein R J, Graham D J *Phys. Fluids* **12** 1133 (1969)
17. Chandrasekhar S *Liquid Crystals* (Cambridge: Cambridge University Press, 1977)
18. De Gennes P G *The Physics of Liquid Crystals* (Oxford: Clarendon Press, 1982)
19. Blinov L M *Electro-Optical and Magneto-Optical Properties of Liquid Crystals* (New York: Wiley, 1983)
20. Kai S *Forma* **7** 189 (1992)
21. Hertrich A et al. *J. de Phys. II* **2** 1915 (1992)
22. Richter H, Buka A, Rehberg I *Phys. Pev. E* **51** 5886 (1995)
23. Richter H et al. *Europhys. Lett.* **30** 37 (1995)
24. Tribelsky M I et al. *Bussei Kenkyu* (*Kyoto*) **66** 592 (1966); *Proc. of the 1st Tohwa Univ. Int. Meeting on Stat. Phys., Fukuoka, Japan, Nov. 1995*
25. Gor'kov L P *Zh. Eksp. Teor. Fiz.* **33** 402 (1957) [*Sov. Phys. JETP* **6** 311 (1958)]
26. Malkus W V R,Veronis G *J. Fluid Mech.* **4** 225 (1958)
27. Schluter A, Lortz D, Busse F H *J. Fluid Mech.* **23** 129 (1965)
28. Landau L D, Lifshits E M *Statisticheskaya Fizika* Chast' 1 (Statistical Physics) (Moscow: Nauka, 1976) [Translated into English 2 Vols, 3rd ed. (Oxford: Pergamon Press, 1980)]
29. Tribel'skiĭ M I *Izv. Akad. Nauk SSSR Ser. Fiz.* **53** 722 (1989) [*Bull. Acad. Sci. USSR Phys. Ser.* **53** 107 (1989)]
30. Landau L D *Sobranie Trudov* T. 1 (Collected Works Vol. 1) (Moscow: Nauka, 1969) p. 234 [Collected Papers of L D Landau (Oxford: Pergamon, 1965)]
31. Landau L D *Sobranie Trudov* T. 1 (Collected Works Vol. 1) (Moscow: Nauka, 1969) p. 447 [Collected Papers of L D Landau (Oxford: Pergamon, 1965)]
32. Landau L D, Lifshits E M *Gidrodinamika* (Hydrodynamics) (Moscow: Nauka, 1988) [Translated into English (Oxford: Pergamon Press, 1990)]
33. Kuznetsov E A, Spektor M D *PMTF* **21** (2) 76 (1980) [*J. Appl. Mech. Tech. Phys.* **21** 220 (1980)]
34. Swift J B, Hohenberg P C *Phys. Rev. A* **15** 319 (1977)
35. Malomed B A, Nepomnyashchy A A, Tribelsky M I *Phys. Rev. A* **42** 7244 (1990)
36. Nepomnyashchy A A, Tribelsky M I, Velarde M G *Phys. Rev. E* **50** 1194 (1994)
37. Langer J S, Ambegaokar V *Phys. Rev.* **164** 498 (1967)
38. McCumber D E, Halperin B I *Phys. Rev. B* **1** 1054 (1970)
39. Newell A C, Whitehead J A *J. Fluid Mech.* **38** 279 (1969)
40. Stuart J T, DiPrima R C *Proc. R. Soc. London A* **362** 27 (1978)
41. Kramer L, Zimmermann W *Physica D* **16** 221 (1985)
42. Kramer L, Schober H R, Zimmermann W *Physica D* **31** 212 (1988)
43. Tribelsky M I, Kai S, Yamazaki H *Prog. Theor. Phys.* **86** 963 (1991)
44. Eckhaus W *J. Mecanique* **2** 153 (1963)
45. Eckhaus W *Studies in Nonlinear Stability Theory* (Springer Tracts in Natural Phylosophy, Vol. 6) (Berlin: Springer-Verlag, 1965)
46. Busse F H *Rep. Prog. Phys.* **41** 1929 (1978)
47. Busse F H, in *Hydrodynamic Instabilities and the Transition to Turbulence* (Eds H L Swinney, J P Gollub) (Berlin: Springer-Verlag, 1981) p. 97 [Translated into Russian (Moscow: Mir, 1984) p. 124]
48. Malomed B A, Staroselsky I E, Konstantinov A B *Physica D* **34** 270 (1989)
49. Malomed B A, Nepomnyashchiĭ A A, Tribel'skiĭ M I *Zh. Eksp. Teor. Fiz.* **96** 684 (1989) [*Sov. Phys. JETP* **69** 388 (1989)]
50. Edwards W S, Fauve S *J. Fluid Mech.* **278** 123 (1994)
51. Ishimasa T, Nissen H -U, Fukano Y *Phys. Rev. Lett.* **55** 511 (1985)
52. Bendersky L *Phys. Rev. Lett.* **55** 1461 (1985)

53. Wang N, Chen H, Kuo K H *Phys. Rev. Lett.* **59** 1010 (1987)
54. Bernoff A J *Europ. J. Appl. Math.* **5** 267 (1994)
55. Tribelsky M I *Int. J. Brf. Chaos* (*Proc. of Int. Workshop on Nonlinear Dynamics and Chaos, Pohang, Korea, July, 1995*) (in press)
56. Tribelsky M I, Velarde M G *Phys. Rev. E* (in press)
57. Tribelsky M I, Tsuboi K *Phys. Rev. Lett.* **76** 1631 (1996)
58. Nikolaevskiĭ V N *Dokl. Akad. Nauk SSSR* **307** 570 (1989)
59. Nikolaevskiĭ V N, in *Recent Advances in Eng. Sci.* (Lecture Notes in Eng. 39) (Eds S L Koh, C G Speciale) (Berlin: Springer-Verlag, 1989) p. 210
60. Malomed B A, Mitlin V S, Nikolaevskiĭ V N *Prikl. Mat. Mekh.* **55** 795 (1991)
61. Malomed B A *Phys. Rev. A* **45** 1009 (1992)
62. Busse F H, Riahi N *J. Fluid Mech.* **96** 243 (1980)
63. Malomed B A, Tribel'sky M I *Zh. Eksp. Teor. Fiz.* **92** 539 (1987) [*Sov. Phys. JETP* **65** 305 (1987)]
64. Kuznetsov E A, Spektor M D *Zh. Eksp. Teor. Fiz.* **71** 262 (1976) [*Sov. Phys. JETP* **44** 136 (1976)]
65. Beresnev I A, Nikolaevskiy V N *Physica D* **66** 1 (1993)
66. Tsuboi K, Tribelsky M I *Transactions of the Japan Society for Industrial and Applied Mathematics* (in Japanese) (in press)
67. Hidaka Y et al. *Technology Reports of Kyushu Univ.* (in Japanese) **69** 411 (1996)
68. Hidaka Y et al. *Phys. Rev. Lett.* (in press)
69. Rossberg A G et al. *Phys. Rev Lett.* **76** 4729 (1966)
70. Rossberg A G, Kramer L *Phys. Scripta* (in press)
71. Helfrich W *Phys. Rev. Lett.* **23** 372 (1969)
72. Dubois-Violette E, Guazzelli E, Prost J *Philos. Mag. A* **48** 727 (1983)
73. Brazovskiĭ S A *Zh. Eksp. Teor. Fiz.* **68** 175 (1975) [*Sov. Phys. JETP* **41** 85 (1975)]
74. Rehberg I et al. *Phys. Rev. Lett.* **67** 596 (1991)