International Journal of Theoretical Physics, Vol. 33, No. 11, 1994

Postinstability Models in Dynamics

Michail Zak[1]

Received April 6, 1994

This paper is devoted to the concept of instability in dynamical systems with the main emphasis on orbital, Hadamard, and Reynolds instabilities. It demonstrates that the requirement about differentiability in dynamics in some cases is not consistent with the physical nature of motions, and may lead to unrealistic solutions. Special attention is paid to the fact that instability is not an invariant of motion: it depends upon frames of reference, the metric of configuration space, and classes of functions selected for mathematical models of physical phenomena. This leads to the possibility of elimination of certain types of instabilities (in particular, those which lead to chaos and turbulence) by enlarging the class of functions using the Reynolds-type transformation in combination with the stabilization principle: the additional terms (the so-called Reynolds stresses) are found from the conditions that they suppress the original instability. Based upon these ideas, a new approach to chaos and turbulence as well as a new mathematical formalism for nonlinear dynamics are discussed.

1. INTRODUCTION

In recent years an increasing amount of interest has been addressed to the fact that, in many different domains of science (physics, chemistry, biology, engineering), systems with a similar strange behavior are frequently encountered. These systems display irregular and unpredictable time evolution, and are called chaotic. But chaotic motions are not the only motions in dynamics which are unpredictable. Much earlier, about 100 years ago, O. Reynolds studied, experimentally and theoretically, turbulent motions in fluids. Despite the many efforts, the problem of prediction of turbulent motions is still unsolved. Later another type of instability which is associated with a failure of hyperbolicity in distributed systems was discovered by J. Hadamard. In all these cases the postinstability behavior of the solutions to the original models is characterized by supersensitivity

[1]Jet Propulsion Laboratory, Caltech, Pasadena, California 91105.

to initial conditions, and for that reason, it cannot be predicted, since the initial conditions are never known exactly. In this paper we will discuss a possibility to develop a unified approach to prediction of postinstability behavior in dynamics.

1.1. Mathematical Formulations and Dynamical Invariants

Dynamics describes the motion of a system, i.e., the time evolution of its parameters. The time variable t can be discrete or continuous. In discrete-time dynamical systems, the rate of change of a parameter x is defined only for discrete values of t. These systems can be represented as the iteration of a function:

$$\mathbf{x}_{t+1} = \mathbf{v}(\mathbf{x}_t, t), \qquad t = 0, 1, 2, \ldots \tag{1}$$

i.e., as difference equations.

In continuous-time dynamical systems the rate of change of x is defined for all values of t; such systems can be modeled by ordinary differential equations.

$$\frac{d\mathbf{x}}{dt} = \dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, t) \tag{2}$$

or by partial differential equations:

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, \mathbf{x}', \mathbf{x}'', \ldots, t), \qquad \mathbf{x}' = \frac{\partial \mathbf{x}}{\partial \mathbf{s}}, \qquad \mathbf{x}'' = \frac{\partial^2 \mathbf{x}}{\partial \mathbf{s}^2}, \ldots \tag{3}$$

if the rate of change, in addition, depends upon distributions of $\mathbf{x}$ over space coordinates $\mathbf{s}$. In equations (1)–(3), $\mathbf{x}$ represents the state of the dynamical system.

Continuous-time dynamical system theory has adopted basic mathematical assumptions of the theory of differential equations, such as differentiability of the parameters (with respect to time and space) "as many times as necessary," the boundedness of the velocity gradients $\partial\dot{\mathbf{x}}/\partial\mathbf{x}$ (the Lipschitz conditions), etc. Under these assumptions, the existence, uniqueness, and stability of solutions describing the behavior of dynamical systems has been studied. However, the dynamical systems cannot be identified with the mathematical models, i.e., with the differential equations. Indeed, dynamical systems are characterized by scalars, vectors, or tensors which are invariant with respect to coordinate transformations. Hence, equation (2) or (3) models a dynamical system only if it preserves these invariants after any (smooth) coordinate transformation. For instance, any model of a mechanical system must be derivable from variational principles which are expressed via the mechanical invariants (kinetic and potential energy, dissipation functions, etc.). In other words, the difference between dynamical systems and the corresponding differential equations is similar to the difference between a matrix as an object of algebra and a second-rank tensor as an object of geometry: The same tensor can be modeled by different matrices, depend on choices of coordinates; however, all these matrices must have the same eigenvalues. Continuing this analogy, it can be expected that the parameters **x** in equations (2) and (3) can be decomposed (at least, in principle) into "invariant" and "noninvariant" components, in the same way in which a matrix A can be decomposed into invariant (diagonal $\tilde{A}$) and coordinate-dependent $(\theta, \theta)^1$ components:

$$A = \theta \tilde{A} \theta^{-1} \tag{4}$$

1.2. Ignorable Coordinates and Orbital Instability

In mechanical systems, "noninvariant" components of **x** can be associated with ignorable (or cyclic) coordinates which do not enter the Lagrangian function explicitly, and therefore do not affect the energy of the system. For nonconservative systems, in addition to that, the generalized forces corresponding to these coordinates are zero. In terms of Lagrange equations, this property is expressed as the conservation of generalized ignorable impulses P (Gantmacher, 1970):

$$\frac{\partial L}{\partial q_\alpha} = 0, \qquad Q_\alpha = 0, \qquad \text{i.e.,} \quad \frac{\partial L}{\partial \dot{q}_\alpha} = P_\alpha = \text{const}, \qquad \alpha = 1, 2, \ldots, m \tag{5}$$

unlike the equations for the position coordinates, which, in general, do not preserve the position impulses P:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = Q_k, \qquad k = 1, 2, \ldots, n \tag{6}$$

Here L is the Lagrangian, q_α and q_k are ignorable and position coordinates, respectively, and Q_k are nonpotential components of generalized forces.

In order to illustrate the difference between position and ignorable coordinates, consider the following dynamical system:

$$\dot{r} = \sin r, \qquad \dot{\theta} = \omega = \text{const} \tag{7}$$

where r and θ are polar coordinates.

It has periodic attractors:

$$r = \frac{\pi k}{2}, \quad k = 0, 1, \ldots, \qquad \theta = \theta_0 + \omega t \tag{8}$$

Returning to (7), one can easily identify r and θ as position and ignorable coordinates, respectively. Indeed, the Lagrangian and generalized forces for this dynamical system are

$$L = \frac{1}{2}(\dot{r}^2 + \dot{\Theta}^2), \qquad Q_r = \dot{r}\cos r, \qquad Q_\theta = 0 \tag{9}$$

and therefore

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_r} = Q_r \neq 0, \qquad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\theta} = Q_\theta = 0 \tag{10}$$

It is important to emphasize that the position coordinate r is stable at the attractors, while the ignorable coordinate Θ is at the boundary of stability: any small error in Θ will increase linearly (but not exponentially) in time.

Indifference of the energy of a dynamical system to an unlimited growth of ignorable coordinates raises the following question: do there exist such states where all the position coordinates are stable, but some of ignorable coordinates are unstable? Numerical experiments give positive answers to this question. These states are associated with chaotic behavior. Unlike periodic attractors, here any small error in initial values of ignorable coordinates increases exponentially (but not linearly) with time, so that two motion trajectories which initially were indistinguishable (because of finite scale of observation) diverge exponentially, and therefore the behavior of the dynamical system becomes unpredictable. But is such a "multivaluedness" of trajectories consistent with the basic mathematical assumptions about motions of dynamical systems? This problem will be discussed in the next sections in connection with predictability in classical dynamics.

1.3. Distributed Systems and Failure of Differentiability

There are two types of distributed systems—hyperbolic and parabolic—which can model dynamical behaviors. (Elliptic equations are ill-posed for time evolution processes.) Distributed dynamical systems can exhibit more sophisticated behaviors, such as turbulence (whose relation to chaos is still disputed), Hadamard's instability (Zak, 1982*a–c*), which is associated with failure of hyperbolicity, and transition to ellipticity, formation of cumulative effects (Zak, 1970, 1983), etc.

Actually, all these phenomena are associated with spatial effects in distributed dynamical systems resulting from additional mathematical restrictions requiring differentiability of dynamical parameters with respect to spatial coordinates. But are these restrictions always consistent with the physical nature of motions? The following example shows that such restrictions may lead to unrealistic solutions.

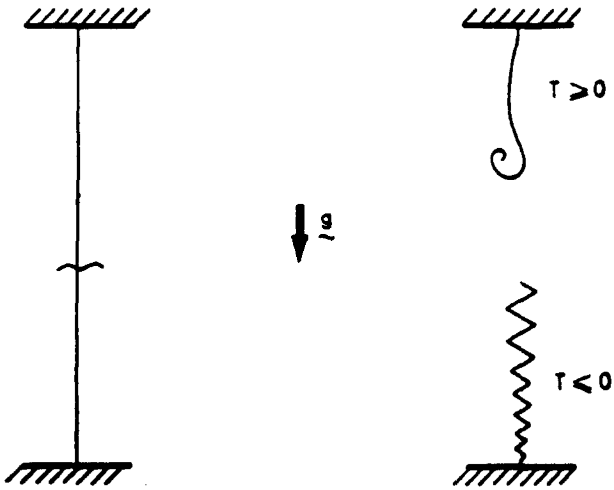

Fig. 1

Consider an ideal filament stretched in the vertical direction, as shown in Fig. 1. Let us cut it at the middle point and observe the behavior of the upper and lower parts. The lower part will fold up in a "thick point," losing differentiability of its configuration. The upper part will preserve differentiability of its configuration in an open interval (which does not include the free end), but at the end small initial disturbances will accumulate and become infinitely large (snap of a whip). Both of these effects are lost in the dynamical model based upon differentiability of the dynamical parameters (for the lower part of the filament) and upon the Lipschitz condition at the free end (for the upper part of the filament (Zak, 1970)).

1.4. Open Problems

As illustrated below, the evolution of ignorable coordinates may be fundamentally different from the evolution of nonignorable (or position) coordinates. For instance, the growth of position coordinates is limited by the boundedness of the system energy, and consequently their instability cannot persist: the system must find an alternative stable state. In contrast, the instability of ignorable coordinates (which is called an orbital instability) can persist all the time without having an alternative stable state. In particular, the indifference of the energy to changes of ignorable parameters is responsible for such phenomena as turbulence, chaos, failure of differentiability and uniqueness of solutions. In turn, the occurrence of these phenomena questions the basic mathematical assumptions about the class of functions in which the dynamical systems are described.

The existence of two different types of parameters in dynamical systems raises some other questions: can instability of ignorable coordinates develop independently of the behavior of the position coordinates? Is instability of ignorable coordinates an invariant of the frame of reference, or of the class of functions in which motions are studied? Can instability of ignorable coordinate be eliminated by change of motion representation?

The answers to these questions, as well as new representations of chaos and turbulence, will be discussed in this paper.

2. INSTABILITY IN DYNAMICS

2.1. Basic Concepts

Most dynamical processes are so complex that a universal theory which would capture all the details during all the time periods is unthinkable. That is why the art of mathematical modeling is to extract only the fundamental aspects of the process and to neglect its insignificant features, without losing the core of information. But "insignificant features" is not a simple concept. In many cases even vanishingly small forces can cause large changes in the dynamical system parameters, and such situations are intuitively associated with the concept of instability. Obviously the destabilizing forces cannot be considered as "insignificant features" and therefore they cannot be ignored. But since they may be humanly indistinguishable, in the very beginning, there is no way to incorporate them into the model. This simply means that the model is not adequate for quantitative description of the corresponding dynamical process: it must be changed or modified. However, the instability delivers important qualitative information: it manifests the boundaries of applicability of the original model.

We will distinguish short- and long-term instabilities. Short-term instability occurs when the system has alternative stable states. For dissipative systems such states can be represented by static or periodic attractors. In the very beginning of the postinstability transition period, the unstable motion cannot be traced quantitatively, but it becomes more and more deterministic as it approaches the attractor. Hence, a short-term instability does not necessarily require a model modification. Usually this type of instability is associated with bounded deviation of position coordinates whose changes affect the energy of the system. Indeed, if the growth of a position coordinate persists, the energy of the system would become unbounded.

The long-term instability occurs when the system does not have an alternative stable state. Such a type of instability can be associated only with ignorable coordinates since these coordinates do not affect the energy of the system. The long-term instability will be the main subject of this paper.

2.2. Orbital Instability

2.2.1. Ignorable Coordinates

As mentioned in the Introduction [see (5)], the coordinate g_α is called ignorable if it does not enter the Lagrangian function L or nonconservative generalized forces Q:

$$\frac{\partial L}{\partial q_\alpha} = 0, \qquad Q_\alpha = 0 \tag{11}$$

Therefore,

$$\frac{\partial L}{\partial \dot{q}_\alpha} = P_\alpha = \text{const} \tag{12}$$

i.e., the generalized ignorable impulse P_α is constant.

As follows from equation (12), there exist such states of dynamical systems (called stationary motions) that all the position coordinates retain a constant value while the ignorable coordinates vary in accordance with a linear law. For example, the regular precession of a heavy symmetric gyroscope is a stationary motion characterized by

$$\Theta = \text{const}, \qquad \dot{\psi} = \text{const}, \qquad \dot{\phi} = \text{const} \tag{13}$$

where the angle of precession ψ and the angle of pure rotation ϕ are ignorable coordinates, while the angle of nutation Θ—the angle formed by the axis of the gyroscope and the vertical—is a position coordinate.

Obviously, stationary motions are not stable with respect to ignorable velocities: a small change in $\dot{q}_\alpha$ at $t = 0$ yields, as time progresses, an arbitrarily large change in the ignorable coordinates themselves. However, since this change increases linearly (but not exponentially), the motion is still considered as predictable. In particular, the Lyapunov exponents for stationary motions are zero:

$$\sigma = \lim_{d(0) \to 0, t \to \infty} \left(\frac{1}{t}\right) \ln \frac{d(0)t}{d(0)} = 0 \tag{14}$$

However, in the case of nonstationary motion, the ignorable coordinate can exhibit more sophisticated behavior. In order to demonstrate this, let us consider the inertial motion of a particle M of unit mass on a smooth pseudosphere S having a constant negative curvature (Fig. 2):

$$G_0 = \text{const} < 0 \tag{15}$$

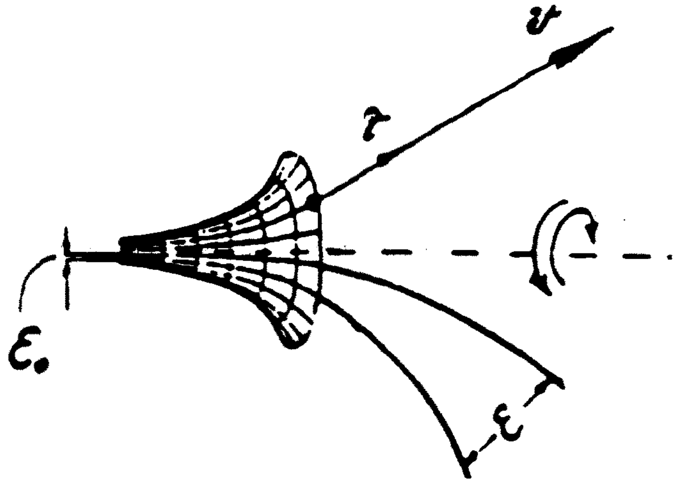

Fig. 2

Remembering that trajectories of inertial motions must be geodesics of S, we will compare two different trajectories assuming that initially they are parallel and that the distance between them ϵ_0 is very small.

As shown in differential geometry, the distance between such geodesics will exponentially increase:

$$\epsilon = \epsilon_0 \exp[(-G_0)^{1/2} t], \qquad G_0 < 0 \tag{16}$$

Hence, no matter how small the initial distance ϵ_0, the current distance ϵ tends to infinity.

Let us assume now that the accuracy to which the initial conditions are known is characterized by L. This means that any two trajectories cannot be distinguished if the distance between them is less than l, i.e., if

$$\epsilon < l \tag{17}$$

The period during which the inequality (17) holds has the order

$$\Delta t \sim \frac{1}{|-G_0|^{1/2}} \ln \frac{l}{\epsilon_0} \tag{18}$$

However, for

$$t \gg \Delta t \tag{19}$$

these two trajectories diverge such that they can be distinguished and must be considered as two different trajectories. Moreover, the distance between them tends to infinity even if ϵ_0 is small (but not infinitesimal). That is why the motion, once recorded, cannot be reproduced again (unless the initial conditions are known exactly), and consequently it attains stochastic features. The Lyapunov exponent for this motion is positive and constant:

$$\sigma = \lim_{t \to \infty, d(0) \to 0} \left(\frac{1}{t}\right) \ln \frac{\epsilon_0 \exp[(-G_0)^{1/2} t]}{\epsilon_0} = (-G_0)^{1/2} = \text{const} > 0 \quad (20)$$

Let us introduce a system of coordinates at the surface S: the coordinate q_1 along the geodesic meridians and the coordinate q_2 along the parallels. In differential geometry such a system is called semigeodesic. The square of the distance between adjacent point on the pseudosphere is

$$ds^2 = g_{11}\, dq_1^2 + 2g_{12}\, dq_1\, dq_2 + g_{22}\, dq_2^2 \quad (21)$$

where

$$g_{11} = 1, \qquad g_{12} = 0, \qquad g_{22} = -\frac{1}{G_0} \exp[-2(-G)^{1/2} q_1] \quad (22)$$

The Lagrangian for the inertial motion of the particle M on the pseudosphere is expressed via the coordinates and their temporal derivates as

$$L = g_{ij} \dot{q}_i \dot{q}_j = \dot{q}_1^2 - \frac{1}{G_0} \exp[-2(-G)^{1/2} q_1] \dot{q}_2^2 \quad (23)$$

and, consequently,

$$\frac{\partial L}{\partial q_2} = 0 \quad (24)$$

while

$$\frac{\partial L}{\partial q_1} \neq 0 \qquad \text{if} \quad \dot{q}_2 \neq 0 \quad (25)$$

Hence, q_1 and q_2 play the roles of position and ignorable coordinates, respectively.

Therefore, an inertial motion of a particle on a pseudosphere is stable with respect to the position coordinate q_1, but it is unstable with respect to the ignorable coordinate. However, in contrast to the stationary motions considered above, here the instability is characterized by exponential growth of the ignorable coordinate, and that is why the motion becomes unpredictable. It can be shown that such a motion becomes stochastic (Arnold, 1988).

Instability with respect to ignorable coordinates can be associated with orbital instability. Indeed, turning to the last example, one can represent the particle velocity v as the product

$$\mathbf{v} = |v| \tau \quad (26)$$

In the course of the instability, the velocity magnitude $|\mathbf{v}|$ and consequently the total energy remain unchanged, while all the changes affect only τ, i.e., the direction of motion. In other words, orbital instability leads to redistribution of the total energy between the coordinates, and it is characterized by positive Lyapunov exponents.

2.2.2. *Orbital Instability of Inertial Motions*

The results described above were related to inertial motions of a particle on a smooth surface. However, they can be easily generalized to motions of any finite-degree-of-freedom mechanical system by using the concept of configuration space. Indeed, if the mechanical system has N generalized coordinate q^i $(i = 1, 2, \ldots, N)$ and is characterized by the kinetic energy

$$W = \alpha_{ij}\dot{q}^i\dot{q}^j \tag{27}$$

then the configuration space can be introduced as an N-dimensional space with the following metric tensor:

$$g_{ij} = a_{ij} \tag{28}$$

while the motion of the system is represented by the motion of the unit-mass particle in this configuration space.

In order to continue the analogy to the motion of the particle on a surface in actual space we will consider only two-dimensional subspaces of the N-dimensional configuration space, without loss of generality. Indeed, a motion which is unstable in any such subspace has to be qualified as unstable in the entire configuration space.

Now the Gaussian curvature of a two-dimensional configuration subspace (q^1, q^2) follows from the Gauss formula:

$$\begin{aligned} G = {} & \frac{1}{a_{11}a_{22} - a_{12}^2}\left(\frac{\partial^2 a_{12}}{\partial q^1 \partial q^2} - \frac{1}{2}\frac{\partial^2 a_{11}}{\partial q^2 \partial q^2} - \frac{1}{2}\frac{\partial^2 a_{22}}{\partial q^1 \partial q^1}\right) \\ & - \Gamma_{12}^{\gamma}\Gamma_{12}^{\delta}a_{\gamma\delta} - \Gamma_{11}^{\alpha}\Gamma_{22}^{\beta}a_{\alpha\beta} \end{aligned} \tag{29}$$

where the connection coefficients Γ_{sk}^{l} are expressed via the Christoffel symbols:

$$\Gamma_{sk}^{l} = \frac{1}{2}a^{lp}\left(\frac{\partial a_{sp}}{\partial q^k} + \frac{\partial a_{kp}}{\partial q^s} - \frac{\partial a_{sk}}{\partial q^p}\right) \tag{30}$$

while

$$a^{\alpha\beta}a_{\beta\gamma} = a_{\gamma}^{\alpha} = \begin{cases} 0 & \text{if} \quad \alpha \neq \gamma \\ 1 & \text{if} \quad \alpha = \gamma \end{cases} \tag{31}$$

Thus, the Gaussian curvature of these subspaces depends only on the coefficients a_{ij}, i.e., it is fully determined by the kinematical structure of the system [see equation (27)]. In case of inertial motions, the trajectories of the representative particle must be geodesics of the configuration space. Indeed, as follows from (26);

$$\frac{d\tau}{dt} = \frac{d\tau}{ds}\dot{s} = 0 \qquad \text{if} \quad \dot{v} = 0 \qquad \text{and} \quad |v| = |\dot{s}| = \text{const} \neq 0 \tag{32}$$

where s is the arc coordinate along the particle trajectory:

$$ds = a_{ij}\, dq^i\, dq^j \tag{33}$$

But then

$$\frac{d\tau}{ds} = 0 \tag{34}$$

which is the condition that the trajectory is geodesic.

If the Gaussian curvature (29), which is uniquely defined by the parameters of the dynamical system a_{ij}, is negative,

$$G < 0 \tag{35}$$

then the trajectories of inertial motions of the system that originated at close, but different points of the configuration space diverge exponentially from each other, and the motion becomes unpredictable and stochastic. Some examples of orbital instability in inertial motions are discussed by Zak (1985*b*).

2.2.3. Orbital Instability of Potential Motions

Turning back to the motion of the particle M on a smooth pseudosphere (Fig. 2), let us depart from inertial motions and introduce a force F acting on this particle. For noninertial motions ($F \neq 0$) the trajectories of the particle will not be geodesics, while the rate of their deviation from geodesics is characterized by the geodesic curvature χ. It is obvious that this curvature must depend on the forces F:

$$\chi = \chi(F) \tag{36}$$

Synge (1926) showed that if the force F is potential,

$$F = -\nabla\Pi \tag{37}$$

where Π is the potential energy, then the condition (35) is replaced by

$$G_0 + 3\chi^2 + \frac{1}{W}\left(\frac{\partial^2\Pi}{\partial q^i \partial q^j} - \Gamma^k_{ij}\frac{\partial\Pi}{\partial q^k}\right)n^i n^j < 0; \qquad i, j = 1, 2 \tag{38}$$

Here Γ_{ij}^k are defined by equations (30), and n^i are the contravariant components of the unit normal $\mathbf{n}$ to the trajectory.

The geodesic curvature χ in (38) can be expressed via the potential force $\mathbf{F}$:

$$\chi = \frac{\mathbf{F} \cdot \mathbf{n}}{2W} = -\frac{\nabla\Pi \cdot \mathbf{n}}{2W} \tag{39}$$

As follows from (38) and (39), the condition (38) reduces to (35) if $\mathbf{F} = 0$.

Suppose, for example, that the elastic force

$$F = -\alpha^2 \epsilon, \qquad \alpha^2 = \text{const} \tag{40}$$

proportional to the normal deviation ϵ from the geodesic trajectory is applied to the particle M moving on the smooth pseudosphere. If the initial velocity is directed along one of the meridians (which are all geodesics), the unperturbed motion will be inertial, and its trajectory will coincide with this meridian since there $\epsilon = 0$, and therefore $F = 0$. In order to verify the orbital stability of this motion, let us turn to the criterion (38). Since

$$\chi = 0 \qquad \text{and} \qquad \frac{\partial \Pi}{\partial q^k} = F^k = 0 \tag{41}$$

for the unperturbed motion, one obtains the condition for orbital stability:

$$G_0 + \frac{\alpha^2}{2W} > 0, \qquad \text{i.e.,} \qquad \alpha^2 < -2WG, \quad G < 0 \tag{42}$$

where

$$W = \frac{1}{2} m v_0^2 \tag{43}$$

As in the case of inertial motions, the inequality

$$\alpha^2 < -2WG_0 \tag{44}$$

leads to unpredictable (stochastic) motions which are characterized by

$$\sigma = \left(G_0 - \frac{\alpha^2}{2W}\right)^{1/2} = \text{const} > 0 \tag{45}$$

For pure inertial motions ($\alpha = 0$), equation (45) reduces to equation (15).

After the discovery of chaotic attractors, the stochastic motions which are generated by the instability and are characterized by positive Lyapunov exponents are called chaotic. Hence, the inequalities (35) and (50) can be associated with criteria of chaos: if the left-hand part in (50) is bounded away from zero by a negative number $-B$ in all the configuration space where the motion can occur, then the motion will be chaotic, and its positive Lyapunov exponent will be

$$\sigma \geq B^2 \tag{46}$$

Unfortunately, this criterion is too "strong" to be of practical significance: it is sufficient, but not necessary. Indeed, this criterion assumes that not only global, but also the local Lyapunov exponents are positive at any point of the configuration space. At the same time, for many chaotic motions, local Lyapunov exponents in certain domains of the configuration space are all negative or zero, although some of the global exponents are still positive.

2.2.4. General Case

Following J. L. Synge, the results for the orbital instability of inertial and potential motions for a system of material points can be generalized to arbitrary motions.

Since the motion of a system of material points in the configuration space with the metric (28) is represented by a unit-mass point, the momentum equation follows from Newton's second law:

$$\ddot{q}^r + \Gamma^r_{mn}\dot{q}^n = Q^r \tag{47}$$

where Q^r is the force applied to the point. Let q^r be the coordinates of the representative point M moving along an undisturbed natural trajectory C, and $(q^r + \eta^r)$ the coordinates of the corresponding (simultaneous) point M^* of the disturbed natural trajectory C, while η^r is an infinitesimal disturbance vector. The condition for stability of the motion is that the magnitude of the disturbance vector should remain permanently small.

Introducing a unit disturbance vector μ^r codirectional with η^r, so that

$$\eta^r = \eta\mu^r, \qquad a_{mn}\mu^m\mu^n = 1 \tag{48}$$

where η is the disturbance vector magnitude, and substituting $q^r + \eta^r$ into equation (47), one can obtain an invariant differential equation with respect to the scalar η (Synge, 1926):

$$\ddot{\eta} + A\eta = 0 \tag{49}$$

where the scalar A is uniquely defined by the metric coefficients a_{ij} and the forces Q_i, namely

$$A = G_{mnsl}\mu^m\dot{q}^n\mu^s\dot{q}^l - \tilde{\mu}^2 - Q_{rs}\mu^r\mu^s \tag{50}$$

in which G_{mnsl} is the curvature tensor of the configuration space expressed in the covariant form

$$G_{msnl} = \frac{\partial \Gamma^m_{nl}}{\partial q^s} - \frac{\partial \Gamma^m_{ns}}{\partial q^l} + (\Gamma^u_{nl}\Gamma^v_{ms} - \Gamma^u_{ns}\Gamma^v_{ml})\alpha^{uv}, \qquad a^u a_{vp} = \delta^p_u \tag{51}$$

Here Γ^r_{mn} are the Christoffel symbols defined by equation (30), and

$$Q_{rs} = \left(\frac{\partial Q^r}{\partial q^l} + \Gamma^r_{ln} Q^n\right) a_{ls}, \qquad \tilde{\mu}^r = \Gamma^r_{mn}\mu^m \dot{q}^n, \qquad \tilde{\mu} = |\tilde{\mu}^r| \tag{52}$$

while the metric tensor of the configuration space is given by equation (27).

Equation (49) leads to a sufficient condition for a dynamical system given in the form (47) to be exponentially unstable. If the Riemannian curvature of the manifold of configurations corresponding to every two-space element $x^m x^n$ containing the direction of the given trajectory is bounded away from zero by a constant negative value, and $Q_{mn} x^m x^n$ is bounded away from zero by a constant positive value in all the domains of the configuration space where the motion can occur, then the motion will be exponentially unstable; since this instability persists, the motion will attain stochastic features (as in the case of the inertial or potential motion of a particle on a smooth pseudosphere), and therefore it will become chaotic. Actually the condition (38), which was formulated earlier without a proof, follows directly from equation (49).

Obviously, the persistency of the instability in equation (49) can occur only due to a contribution of the exponential growth of the ignorable coordinates to the total magnitude of the disturbance vector η. For instance, in the case of inertial motion of the particle M on a smooth pseudosphere, the disturbance vector can be represented by the components ϵ and ν, which are codirectional and normal to the unperturbed (geodesic) trajectory. The component ν corresponds to the ignorable coordinate, and its evolution is described by equation (49), which reduces to

$$\ddot{\nu} + 2W_0 G_0 \nu = 0 \tag{53}$$

The exponential instability of ν when $G_0 = \text{const} < 0$ leads to chaos. At the same time, the position coordinate ϵ is eliminated from equation (53) and it can be found from the energy conservation:

$$\epsilon = \epsilon_0 \frac{\dot{\epsilon}}{\dot{S}_0} \tag{54}$$

where $\dot{S}_0$ and ϵ_0 are the initial conditions at $t = 0$ for the motion velocity along the trajectory and the position coordinate of the disturbance vector, respectively. In spite of some limitations of the results described above (the conditions for chaos are sufficient, but not necessary; the forces Q^r depend only upon coordinates, but not upon velocities), they nevertheless elucidate the physical origin of orbital instability, chaos, and consequently, of unpredictability of motions in classical dynamics.

2.3. Hadamard's Instability

2.3.1. General Remarks

The results presented in the previous section can be applied to distributive systems after a discretization technique which reduces them to finite-dimensional systems. For instance, as noticed by Arnold (1988), an inviscid stationary flow with a smooth velocity field

$$v_x = A \sin Z + C \cos Y, \qquad v_y = B \sin X + A \cos Z, \\ v_z = C \sin Y + B \cos X \tag{55}$$

has chaotic trajectories $X(t)$, $Y(t)$, $Z(t)$ of fluid particles (Lagrangian turbulence) due to negative curvature of the configuration space which is obtained as a finite-dimensional approximation of a continuum. However, there are some special types of instability in distributed systems which can be lost in the course of the discretization, and we focus on them in this section.

As noticed in the previous section, the long-term instability which may lead to chaos is associated with the orbital instability, i.e., with the instability of ignorable coordinates. However, in distributed systems described by partial differential equations, there is another possibility for long-term instability which is associated with the decrease of scale of motions, i.e., with the growth of spatial derivates of the system parameters. In mathematical terms this means a failure of differentiability of the solutions to the corresponding governing equations. However, an unlimited growth of spatial derivatives must be consistent with the boundedness of energy. Indeed, the stresses in continuous media depend not upon displacements or velocities, but upon their gradients, i.e., upon their space derivatives. Hence, we have to find such situations when an unlimited growth of these derivatives does not lead to unbounded stresses.

Turning to the geometry of displacements and their gradients in continua, let us introduce the displacement vector

$$\mathbf{u} = \mathbf{r} - \mathbf{r}_0 \tag{56}$$

where $\mathbf{r}_0$ and $\mathbf{r}$ are the radius vectors of the same particle before and after deformation, respectively. In elastic bodies, the stress tensor depends upon the displacement gradient $\nabla \mathbf{u}$ via the strain tensor ϵ:

$$\epsilon = \frac{1}{2}[\nabla \mathbf{u} + (\nabla \mathbf{u})^T + \mathbf{u} \cdot (\nabla \mathbf{u})^T] = \frac{1}{2}[\nabla \mathbf{r} \cdot (\nabla \mathbf{r})^T - \mathring{g}] \tag{57}$$

where $\mathring{g}$ is the unit (the initial state) tensor, while the current state metric tensor is defined as

$$g = 2\epsilon + \mathring{g} \tag{58}$$

The tensor gradient $\nabla\mathbf{r}$ in (57) can be decomposed as

$$\nabla\mathbf{r} = CB \tag{59}$$

where C is a symmetric tensor:

$$C = +[\nabla\mathbf{r} \cdot (\nabla\mathbf{r})]^{1/2} \tag{60}$$

and B is an orthogonal tensor:

$$B = +[\nabla\mathbf{r} \cdot (\nabla\mathbf{r})^T]^{1/2} \cdot (\nabla\mathbf{r})^T = (B^{-1}), \qquad \det B = 1 \tag{61}$$

As follows from (60), the strain tensor is

$$\epsilon = \frac{1}{2}(c^2 - \mathring{g}) \tag{62}$$

and consequently the stress tensor depends only upon the symmetric part of the tensor gradient $\nabla\mathbf{r}$, and does not depend upon its orthogonal component B, which corresponds to rigid rotations of elementary volumes. However, indirectly an unlimited growth of these rotations can lead to unbounded stresses in three-dimensional elastic bodies. Indeed as follows from the identity

$$\nabla \times \nabla\mathbf{r} = 0 \tag{63}$$

the components of the tensor gradient $\nabla\mathbf{r}$ must satisfy six additional constraints, which are called the compatibility equations. Loosely speaking, they follow from the requirement that after deformations, the continuum should not have any "holes" or "cracks." In geometrical terms, (62) represents the fact that after deformations, the actual space remains Euclidean, i.e., the curvature tensor is zero:

$$R = 0 \tag{64}$$

However, six constraints imposed upon the tensor gradient $\nabla\mathbf{r}$ by (63) or (64) are also not independent. Indeed, according to another identity

$$\nabla \cdot \nabla \times \nabla\mathbf{r} \equiv 0 \tag{65}$$

which holds even if

$$\nabla \times \nabla\mathbf{r} \neq 0 \tag{66}$$

and which is equivalent to three scalar equations, only three of the six constraints (63) are truly independent. In geometrical terms equation (65) can be associated with the Bianchi identities (Fluge, 1962).

Thus, nine components of the vector gradient $\nabla \mathbf{r}$ must satisfy three independent compatibility equations, and therefore if all six components of the stress tensor ϵ are given, then the remaining three components of $\nabla \mathbf{r}$ and consequently all the rigid rotations will be uniquely defined. This means that in isotropic three-dimensional elastic bodies, an unlimited decrease of scale of motions would lead to unbounded stresses, which is physically impossible.

Let us turn to one-dimensional continua (filaments). In this case, rigid rotations define the external geometry of the model (the rotations about the binormal to the filament correspond to the first curvature, and the rotations about the tangent to the filament correspond to the second curvature, or twist; Fig. 3), and they do not depend upon the elongations of the curve which define the stress. Indeed, let us introduce the filament equation in the form

$$\mathbf{r} = \mathbf{r}(\psi, t), \qquad \left|\frac{\partial \mathbf{r}}{\partial \psi}\right| = 1 \tag{67}$$

where ψ plays the role of an Eulerian coordinate. Then the motions associated with changes of the internal geometry and therefore the stresses are described by the function

$$\psi = \psi(s, t) \tag{68}$$

where s is the Lagrangian coordinate of an individual particle.

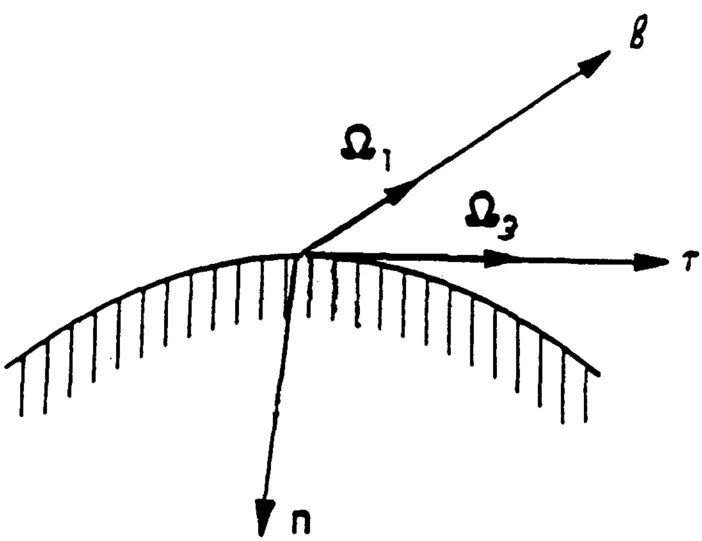

Fig. 3

At the same time, the curvatures of the filament configurations can be expressed as

$$\Omega_3 = \left|\frac{\partial^2 \mathbf{r}}{\partial \psi^2}\right|, \qquad \Omega_1 = \frac{(\partial \mathbf{r}/\partial \psi \times \partial^2 \mathbf{r}/\partial \psi^2) \cdot \partial^3 \mathbf{r}/\partial \psi^3}{|\partial^2 \mathbf{r}/\partial \psi^2|} \tag{69}$$

Consequently, both curvatures are independent of the internal geometry characterized by (68), and in particular, upon the stress defined by the derivative $\partial\psi/\partial s$.

This means that unlimited growth of the curvature may not cause stress at all, and therefore, instability in the form of unlimited decrease of scale of motions is possible (Fig. 1).

The situation becomes more complicated in two-dimensional continua (films, membranes). Here the internal geometry is defined by two-dimensional versions of equations (56)–(63), while the external geometry is described by the coefficients of the second fundamental form:

$$b_{ij} = \frac{\partial^2 \mathbf{r}}{\partial \psi^i \, \partial \psi^j} \cdot \mathbf{n}, \qquad i, j = 1, 2 \tag{70}$$

where ψ^i are coordinates on the surface, and $\mathbf{n}$ is the unit normal to the surface.

However, these coefficients are not independent: they are coupled with the strains by three compatibility equations:

$$b_{11}b_{22} - b_{12}^2 = \Gamma^v_{12}\Gamma^\delta_{12}g_{v\delta} - \Gamma^\alpha_{11}\Gamma^\beta_{22}g_{\alpha\beta} - \frac{1}{2}\frac{\partial^2 g_{11}}{\partial \psi^{(2)2}} + \frac{\partial^2 g_{12}}{\partial \psi^{(1)}\partial \psi^{(2)}} - \frac{1}{2}\frac{\partial^2 g_{22}}{\partial \psi^{(1)2}} \qquad (v, \delta, \alpha, \beta = 1, 2) \tag{71}$$

$$\frac{\partial b_{il}}{\partial \psi^{(2)}} - \frac{\partial b_{i2}}{\partial \psi^{(1)}} = \Gamma^1_{i2}b^i_{11} - \Gamma^2_{il}b_{22} + (\Gamma^2_{i2} - \Gamma^1_{il})b_{12} \qquad (i = 1, 2) \tag{72}$$

where

$$g_{11} = 1 + 2\epsilon_{11}, \qquad g_{12} = 2\epsilon_{12}, \qquad g_{22} = \mathring{g}_{22} + 2\epsilon_{22} \tag{73}$$

The two-dimensional Christoffel symbols are

$$\Gamma^n_{ij} = \frac{1}{2}g^{nl}\left(\frac{\partial g_{li}}{\partial \psi^{(j)}} + \frac{\partial g_{lj}}{\partial \psi^{(i)}} - \frac{\partial g_{ij}}{\partial \psi^{(l)}}\right) \qquad (n, i, j = 1, 2) \tag{74}$$

while

$$\|g_{ij}\| = \|g^{ij}\|^{-1} \tag{75}$$

Hence, in general, three coefficients b_{ij} are defined by the strains ϵ_{ij} from the three equations (71) and (72), and, consequently, change in b_{ij} affects the strains ϵ_{ij}.

Nevertheless, there are situations when the unlimited growth of curvature may not affect the stress at all. In order to describe this case, recall that on a surface with negative or zero Gaussian curvature

$$G = \frac{b_{11}b_{22} - b_{12}}{g_{11}g_{22} - g_{12}^2} \tag{76}$$

there exists a family of asymptotic lines where the second fundamental form is equal to zero:

$$b_{11} \tan^2 \phi + 2b_{12} \tan \phi + b_{22} = 0 \tag{77}$$

while the angle ϕ between an asymptotic line and the coordinate line ψ_1 is found as

$$\tan \phi = -\frac{b_{12}}{b_{11}} \pm \frac{(b_{11}^2 - b_{11}b_{22})^{1/2}}{b_{11}} \qquad (b_{11} \neq 0) \tag{78}$$

Selecting the coordinate ψ_1 as an asymptotic line, one obtains

$$\tan \phi = 0 \tag{79}$$

and, as follows from equations (77) and (78),

$$b_{22} = 0 \tag{80}$$

Now it is obvious that along the asymptotic line the curvature b_{11} can be selected arbitrarily without affecting the parameters of the internal geometry g_{ij}, and consequently the stress. Indeed, since $b_{22} = 0$, b_{11} is eliminated from (71). In addition, as follows from (72),

$$\frac{\partial b_{11}}{\partial \psi^{(2)}} - \frac{\partial b_{12}}{\partial \psi^{(1)}} = \Gamma_{12}^1 b_{11} + (\Gamma_{11}^2 - \Gamma_{11}^1) b_{12} \tag{81}$$

$$\frac{\partial b_{12}}{\partial \psi^{(2)}} = \Gamma_{22}^1 b_{11} + (\Gamma_{22}^2 - \Gamma_{21}^1) b_{12} \tag{82}$$

The derivative $\partial b_{11}/\partial \psi_1$ is not defined, i.e., that asymptotic line ϕ_1 coincides with the characteristic of the partial differential equations (71) and (82).

This means that the curvature b_{11} can be chosen arbitrarily along the asymptotic lines of the surface without affecting any parameters of the film, including stresses. In other words, an unlimited growth of the curvature b_{11} may be consistent with the unboundedness of stresses and it can be associated with the formation of wrinkles along the asymptotic lines.

So far we have been concerned with elastic continua. Turning to fluids, one should recall that their stresses depend only upon the velocities, but not upon the displacements. That is why an unlimited growth of any component of the displacement vector (56) or of the tensor gradient (59) is consistent with the unboundedness of stresses, and it can be associated with the Lagrangian turbulence.

In terms of velocities, the situation is different. In order to demonstrate that, recall that in a viscous fluid the stress tensor depends upon the velocity gradient $\nabla \mathbf{v}$ via the time derivative of the strain tensor (57). The velocity gradient $\nabla \mathbf{v}$ has the same type of structure as the vector gradient $\nabla \mathbf{r}$: it can be decomposed into a symmetric tensor of the rate of strain

$$\dot{\epsilon} = \frac{1}{2}[\nabla \mathbf{v} + (\nabla \mathbf{v})^T] \tag{83}$$

and an antisymmetric tensor

$$\omega = \frac{1}{2}[\nabla \mathbf{v} - (\nabla \mathbf{v})^T] \tag{84}$$

which is equivalent to the vector of vortex

$$\omega = \frac{1}{2}\,\text{Curl}\,\mathbf{v} = \frac{1}{2}\nabla \times \mathbf{v} \tag{85}$$

while

$$\nabla \mathbf{v} = \dot{\epsilon} + \omega \tag{86}$$

Since

$$\nabla \times \nabla \mathbf{v} \equiv 0 \quad \text{and} \quad \nabla \cdot \nabla \times \nabla \mathbf{v} \equiv 0 \quad \text{(even if)} \quad \nabla \times \nabla \mathbf{v} \neq 0 \tag{87}$$

one comes to the same conclusion as in the case of the vector gradient $\nabla \mathbf{r}$ [see (63) and (65)]: nine components of the tensors $\dot{s}$ and ω are coupled by three compatibility equations. Hence, six components of the rate of strain tensor $\dot{\epsilon}$ uniquely define the velocity gradient, and for that reason an unlimited growth of the vortices in viscous fluids would lead to unlimited growth of stresses.

The situation becomes different in inviscid fluids, where stress is defined only by a scalar—the divergence $\nabla \times \mathbf{v}$. But since any velocity field can be uniquely defined based upon two independent components of its gradient $\nabla \mathbf{v}$, which are the divergence $\nabla \cdot \mathbf{v}$ and the vorticity $\nabla \times \mathbf{v}$, one concludes that an unlimited growth of vorticity in an inviscid fluid may not lead to unbounded stresses. This conclusion can be loosely applied to motions of viscous fluids characterized by high Reynolds number when viscous stresses are ignorable in comparison to the inertia forces. In this case an "unlimited" growth of vortices can be associated with turbulence.

Thus, in this section we have analyzed a possibility "in principle" of an unlimited decrease of scale of motions in continua from the viewpoint of a consistency of this type of instability with the boundedness of stresses and energy. This means that if such an instability exists, it can be found only in one- or two-dimensional elastic models or in fluid motions with high Reynolds number.

2.3.2. Failure of Hyperbolicity in Distributed Systems

Mathematical models of continua are based on the assumption that the functions describing their states can be differentiated "as many times as necessary" at any point exclusive of some special surfaces of discontinuities simulating shock waves or coinciding with the characteristics of the governing equations. In other words, these functions must be at least piecewise differentiable. From the physical viewpoint this means that any point as a center of mass of an infinitesimal volume represents all the properties of this volume. Obviously, the assumption about the smoothness of the functions allows us to use the mathematical technique of differentiable equations.

However, this artificial mathematical limitation follows neither from the principles of mechanics nor from the definition of a continuum. The price paid for such a mathematical convenience is instability (in the class of smooth functions) of the solutions to the corresponding governing equations in some regions of the parameters. This instability is characterized by unlimited decrease of the scale of the motions, in the course of which the derivatives of the corresponding functions tend to infinity although the functions themselves remain finite. In other words, the solution tends to "go out" from the class of differentiable functions.

Most of the instability phenomena leading to unlimited decreasing of the scale of continua motions are associated with the failure of hyperbolicity of the corresponding governing equations, i.e., with the appearance of imaginary characteristic speeds (Zak, 1982*b*,*c*).

In order to illustrate this, we will start with the governing equations of motion of elastic bodies in the following form:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left[\frac{\partial \Pi}{\partial(\partial u_i/\partial x_j)} \right] + F_i, \qquad i = 1, 2, 3 \tag{88}$$

where u_i are the displacements, Π is the potential energy of strains, ρ is the density, F_i are the external forces, and x_i are the material coordinates, posing the initial value problem

$$u_i^0 = u_i \bigg|_{t=0} = \begin{cases} (1/\lambda_0^2) \sin \lambda_0 x_1 & \text{if } \; |x_1| \le x_0 \\ 0 & \text{if } \; |x_i| > x_0, \quad i = 1, 2, 3 \end{cases} \tag{89}$$

$$\left(\frac{\partial u}{\partial t}\right)_t = 0, \qquad i = 1, 2, 3 \tag{90}$$

where the parameter λ_0 can be made as large as necessary, i.e.,

$$\lambda_0 \to \infty \tag{91}$$

The region of the initial disturbance can be arbitrarily shrunk, i.e.

$$|x_0| \to 0 \tag{92}$$

Consequently, the initial disturbances u_i and their first derivatives $\partial u_i/\partial x_i$ can be made as small as necessary. This means that for the corresponding infinitesimal period of time Δt_0 the equations (88) can be linearized and the solution subject to the initial conditions (89) can be sought in the form $u_j = (x_i, t)$, i.e.,

$$\rho_0 \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^{3} a_{ij} \frac{\partial^2 u_j}{\partial x_1^2}, \qquad \text{while} \quad \left.\frac{\partial u_j}{\partial x_j}\right|_{j \neq 1} \equiv 0 \tag{93}$$

where

$$a_{ij} = \left.\frac{\partial^2 \Pi}{\partial(\partial u_i/\partial x_1)\partial(\partial u_j/\partial x_1)}\right|_{\partial u_i/\partial x_i, \partial u_j/\partial x_1 = 0} \tag{94}$$

Let us assume that one of the eigenvalues of the matrix a_{ij} is negative:

$$\lambda_1 < 0 \tag{95}$$

Then the solution to the equation (93) will contain the term

$$\frac{1}{\lambda_0^2} \exp\left[\lambda_0 \left(\frac{|\lambda_1|}{\rho_0}\right)^{1/2} \Delta t\right] \sin \lambda_0 x \tag{96}$$

which tends to infinity if $\lambda_0 \to \infty$ within an arbitrary short period of time Δt_0 and within an infinitesimal volume around the point x_i. Hence, one arrives at the following situation:

$$|u_i| \to \infty \tag{97}$$

in spite of the fact that

$$\left.|u_i|\right|_{t=0} \to 0 \tag{98}$$

However, strictly speaking, because of utilization of the governing equation (88) in a linearized form, the conditions (98) must be weakened:

$$|u_i| \not\to 0 \qquad \text{if} \quad \left.|u_i|\right|_{t=0} \to 0 \tag{99}$$

The formula (99) shows that the appearance of negative eigenvalues of the matrix (94), and consequently imaginary characteristic roots of the governing equation (88) (failure of its hyperbolicity), leads to the violation of a continuous dependence between the initial and transient disturbances during an arbitrarily short period of time and within an arbitrarily selected volume. This type of instability was first observed by J. Hadamard in connection with the ill-posedness of the Cauchy problem for the Laplace equation. Further results with applications to the instability of a string, film, and free surfaces of elastic bodies were reported by Zak (1982*b*,*c*).

The result formulated above was obtained under specially selected initial conditions (89), but it can be generalized to include any initial conditions. Indeed, for equations (93) let the initial conditions be arbitrarily defined by

$$|u_i|_{t=0} = u_i^{00} \tag{100}$$

and the corresponding solution is

$$u_i = f_i(x, t) \tag{101}$$

By altering the initial conditions to

$$u(0, t) = u_i^0 + u_i^{00} \tag{102}$$

where u_i is defined in (89), we observe from the preceding argument by superposition that vanishingly small change in the initial conditions would lead to unboundedly large solutions.

To obtain a geometrical interpretation of the above-described instability, let us turn to expression (97) of the solution and note that if the second derivatives $\partial^2 u_i/\partial t^2$, $\partial u_i/\partial x_i^2$ are of order λ_0, then the first derivatives $\partial u_i/\partial t$, $\partial u_i/\partial x_i$ are of order 1, and u_i are of order $1/\lambda_0$. Hence, the period of time Δt_0 can be selected in such a way that the second derivatives will be as large as necessary, but the first derivatives and u_i are still sufficiently small. Taking into account that the original governing equation (88) is quasilinear with respect to the second derivatives and therefore the linearization does not impose any restrictions on their values, one can conclude that the linearized equation (93) is valid for the solution during the above-mentioned period of time Δt_0. Turning to the formula (97), one can now interpret the solution by the function having an infinitesimal amplitude and changing its signs with an infinite frequency ($\nu = \lambda_0 \to \infty$). The first derivatives of this function $\partial_i/\partial t$, ∂_i/x_i can be small and change their signs by finite jumps (with the same infinite frequency ν), so that the second derivatives $\partial^2 u_i/\partial t^2$, $\partial^2 u_i/\partial x_1^2$ at the points of such jumps are infinite. Thus, within an arbitrarily small volume there is located an arbitrarily large number of points at which the strains have jumps. From the mathematical point of view, the function describing such a field of displacements u_i is considered as a continuous but nondifferentiable function. This function can be simulated, for instance, by a function with a multivalued derivative.

2.3.3. The Criteria of Hadamard's Instability

Let us fix an arbitrary point M and an arbitrary direction x_i at this point in an elastic body. According to the above-formulated result, the instability at the point M in the x direction results from the negative eigenvalues of the matrix:

$$a_{ij} = \left\{ \frac{\partial^2 \Pi}{\partial(\partial u_i/\partial x_1)\partial(\partial u_j/\partial x_1)} \right\} \Bigg|_{\partial u_i/\partial x_1, \partial u_j/\partial x_1 = 0} \tag{103}$$

Assuming that the unperturbed state at this point is characterized by the initial stresses

$$T^0_{1j} \neq 0, \qquad T^0_{2j} = 0, \qquad T^0_{3j} = 0, \qquad j = 1, 2, 3 \tag{104}$$

but zero strains

$$\left(\frac{\partial u_i}{\partial x_1}\right)_0 = 0, \qquad \text{i.e.,} \qquad \epsilon^0_{11} = 0, \quad \gamma^0_{12} = 0, \quad \gamma^0_{13} = 0 \tag{105}$$

let us utilize the following for a variation of the specific potential energy from the initial stresses defined by (104):

$$\delta\Pi = T_{11}\delta\epsilon_{11} + T_{12}\delta\gamma_{12} + T_{13}\delta\gamma_{13} = \cdots = T_{ij}\delta\epsilon_{ij} \tag{106}$$

Taking into account that

$$\epsilon_{11} = \frac{\partial u_1}{\partial x_1} + \frac{1}{2}\left[\left(\frac{\partial u_1}{\partial x_1}\right)^2 + \left(\frac{\partial u_2}{\partial x_1}\right)^2 + \left(\frac{\partial u_3}{\partial x_1}\right)^2\right]$$

$$\gamma_{1i} = \frac{\partial u_i}{\partial x_1} + \cdots \qquad (i = 1, 2, 3)$$

i.e.,

$$\delta\epsilon_{11} = \left(1 + \frac{\partial u_1}{\partial x_1}\right)\delta\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_1}\delta\frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_1}\delta\frac{\partial u_3}{\partial x_1}$$

$$\delta\gamma_{1i} = \delta\frac{\partial u_i}{\partial x_1} \qquad (i = 1, 2, 3)$$

$$T_{1i}\big|_{\partial u_1/\partial x_1} = 0 \tag{107}$$

one obtains for $\partial u_i/\partial x_1 \to 0$

$$a_{11} = T^0_{11} + \frac{\partial T_{11}}{\partial\epsilon_{11}} \tag{108}$$

$$a_{22} = T^0_{11} + \frac{\partial T_{12}}{\partial\gamma_{12}} = T^0_{11} + \frac{1}{2}\frac{\partial T_{12}}{\partial\epsilon_{12}} \tag{109}$$

$$a_{33} = T_{11}^0 + \frac{\partial T_{13}}{\partial \gamma_{13}} = T_{11}^0 + \frac{1}{2}\frac{\partial T_{13}}{\partial \epsilon_{13}} \tag{110}$$

$$a_{12} = a_{21} = a_{13} = a_{31} = a_{23} = a_{32} = 0 \tag{111}$$

where the stresses T_{ij} are related to the local Cartesian coordinates x_1, x_2, x_3 at the point M.

Now the eigenvalues of the matrix (106) can be written in the form

$$\lambda_i = a_{ii}$$

i.e.,

$$\lambda_1 = T_{11}^0 + \frac{\partial T_{11}}{\partial \epsilon_{11}} \tag{112}$$

$$\lambda_2 = T_{11}^0 + \frac{\partial T_{12}}{\partial \gamma_{12}} = T_{11}^0 + \frac{1}{2}\frac{\partial T_{12}}{\partial \epsilon_{12}} \tag{113}$$

$$\lambda_3 = T_{11}^0 + \frac{\partial T_{13}}{\partial \gamma_{13}} = T_{11}^0 + \frac{1}{2}\frac{\partial T_{13}}{\partial \epsilon_{13}} \tag{114}$$

Hence, the criteria of instability are

$$T_{11}^0 < -\frac{\partial T_{11}}{\partial \epsilon_{11}} \tag{115}$$

$$T_{11}^0 < -\frac{\partial T_{12}}{\partial \gamma_{12}} \tag{116}$$

$$T_{11}^0 < -\frac{\partial T_{13}}{\partial \gamma_{13}} \tag{117}$$

Each inequality leads to the failure of differentiability of the corresponding component of strains: ϵ_{11}, ϵ_{12}, or ϵ_{13}, while the potential energy $\Pi(\epsilon_{1i})$ has a local maximum.

Recall that all the above-formulated results are related to an arbitrary point M_0 and arbitrary selected direction x_1, with the unidirectional initial stress T_{11}^0.

In the general case when all the components of the initial stresses are nonzero,

$$T_{ij}^0 \neq 0 \tag{118}$$

one can decompose them into spherical and deviatoric parts,

$$T_{ij}^0 = \frac{1}{3} T_0 E + T_{ij}^{0*}, \qquad T_{ij}^{0*} = \text{Dev}\, T_{ij}^0 \tag{119}$$

where E is the unit tensor, and

$$T_0 = \frac{1}{3}\sum_{i=1}^{3} T_{ij}, \qquad \sum_{i=1}^{3} T_{ij}^* = 0 \tag{120}$$

Now equation (106) can be rewritten in the following form:

$$\delta\Pi = T_0\delta\epsilon_0 + T_{11}^*\delta\epsilon_{11} + T_{12}^*\delta\epsilon_{12} + \cdots = T_0\delta\epsilon_0 + T_{ij}^*\delta\epsilon_{ij} \tag{121}$$

where ϵ_0 is the spherical part of the strain tensor:

$$\epsilon_0 = \frac{1}{3}\sum_{i=1}^{3} \epsilon_{ij} \tag{122}$$

and instead of equations (108)–(110), one obtains

$$\begin{aligned} a_{11} &= T_{11}^0 + \frac{\partial T_{11}^*}{\partial \epsilon_{11}} \\ a_{22} &= T_{11}^0 + \frac{1}{2}\frac{\partial T_{12}^*}{\partial \epsilon_{12}} \\ a_{33} &= T_{11}^0 + \frac{1}{2}\frac{\partial T_{13}^*}{\partial \epsilon_{13}} \end{aligned} \tag{123}$$

Consequently, the sufficient conditions of the instability in some directions at the fixed point for an isotropic elastic material for which the derivatives $\partial T_{ij}/\partial\epsilon_{ij}$ do not depend on a selected direction x_1 are

$$\hat{T}_{11}^0 < -\frac{\partial T_{11}}{\partial \epsilon_{11}} \tag{124}$$

$$\hat{T}_{11}^0 < -\frac{1}{2}\frac{\partial T_{12}}{\partial \epsilon_{12}} \tag{125}$$

$$\hat{T}_{11}^0 < -\frac{1}{2}\frac{\partial T_{13}}{\partial \epsilon_{13}} \tag{126}$$

where $\hat{T}_{11}^0$ is one of the principal deviatoric stresses.

The instability emerges in any direction if these inequalities are valid for all the principle deviatoric stresses $\hat{T}_{ii}^0$:

$$\hat{T}_{ii}^0 < -\frac{1}{2}\frac{\partial T_{ij}}{\partial \epsilon_{ij}} \qquad (i \neq j) \tag{127}$$

because usually

$$\frac{\partial T_{ii}}{\partial \epsilon_{ii}} > \frac{\partial T_{ij}}{\partial \epsilon_{ij}} \qquad (i \neq j) \tag{128}$$

For a Hook's material the criteria of the instability are expressed in terms of Young's modulus E and the Poisson ratio ν,

$$\frac{\partial T_{ii}}{\partial \epsilon_{ii}} = \frac{E(1-\nu)}{(\nu+1)(1-2\nu)} \quad \frac{\partial T_{ij}}{\partial \epsilon_{ij}} = \frac{E}{\nu+1} = 2G \tag{129}$$

if the initial stress tensor is spherical ($T_{ij}^0 = p$), where E, G, and μ are the Young's and shear moduli and ν is the Poisson ratio.

2.3.4. Boundaries of Applicability of the Classical Models of Distributed Systems

All the results discussed above were based on the formal analysis of mathematical models of elastic materials, and their practical usefulness has to be demonstrated. The most obvious and visualizable application of these results can be found in the area of one- and two-dimensional models such as strings, membranes, etc., whose states are defined not only by internal geometry (strains), but also by external geometry (shape). As shown in Section 2.3.1, in this model, unlimited decrease of the scale of motions may be consistent with the boundedness of stresses and energy. The problem of the shape instability there occurs as a result of any local compression and manifests itself in wrinkling, in the course of which the shape loses its smoothness.

Examples. (a) For one-dimensional continua such as an ideally flexible inextensible string, two types of the characteristic speeds are obtained (Zak, 1968):

$$\lambda_{1,2} = \pm\left(\frac{T}{\rho}\right)^{1/2} \tag{130}$$

$$\lambda_{3,4} = \pm\left(\frac{T}{\rho} + \frac{F_n}{\Omega}\right)^{1/2} \tag{131}$$

where T is the tension, ρ is the linear density, Ω is the first curvature of the string's shape, and F_n is the normal component of the external tracking force. These characteristic speeds correspond to discontinuities of the curvature and twist of the string, respectively (Fig. 4).

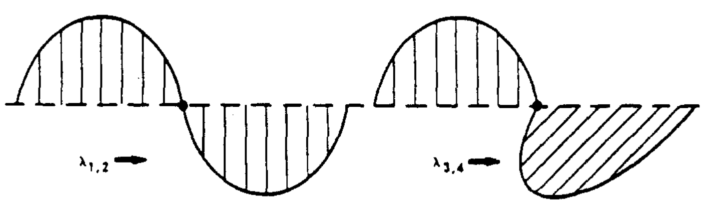

Fig. 4

These conditions of the instability of the string's shape following from the failure of hyperbolicity are given in the form

$$T < 0 \tag{132}$$

$$T < \frac{F_n}{\Omega} \tag{133}$$

The inequality (131) expresses the well-known fact that a compressed string is unstable (the loss of stability of the first curvature; Fig. 1). The shape of such a string cannot be described by differentiable functions, and, theoretically, that string can be rolled up in a point. The inequality (132) shows that even a stretched string can be unstable if subjected to the corresponding tracking force (the loss of the stability of the twist).

These results are generalized to a one-dimensional, ideally flexible pipe within which an ideal fluid flows (Zak, 1982*b,c*):

$$\begin{aligned} \lambda^1_{1,2} &= \frac{\rho^1}{\rho+\rho^1}u \pm \left[\frac{T}{\rho+\rho^1} - \frac{\rho\rho^1}{(\rho+\rho^1)^2}u^2\right]^{1/2} \\ \lambda^1_{3,4} &= \frac{\rho^1}{\rho+\rho^1}u \pm \left(\frac{T}{\rho+\rho^1} + \frac{F_n}{(\rho+\rho^1)\Omega} - \frac{\rho\rho^1}{(\rho+\rho^1)^2}u^2\right)^{1/2} \end{aligned} \tag{134}$$

where T is the tension referred to the entire pipe's cross section, ρ^1 is the linear density of the fluid, and u is the velocity of the fluid.

Then, the conditions of the failure of hyperbolicity are given by

$$T < \frac{\rho\rho^1}{(\rho+\rho^1)}u^2 \tag{135}$$

$$T < \frac{F_n}{\Omega} + \frac{\rho\rho^1}{(\rho+\rho^1)^2}u^2 \tag{136}$$

This means that a flow within the pipe destabilizes its shape. In order to illustrate the last results, let us consider a vertical, ideally flexible, inextensible pipe with a free lower end suspended in the gravity field. Assuming that the flow within this pipe has constant velocity u_0, let us define the area of the instability (Fig. 5). The tension T referred to the entire pipe cross section is given by

$$T = \rho g\zeta(l-x) \tag{137}$$

where l is the length of the pipe, x is the coordinate along the length of the pipe, and ζ is the ratio of the area of the cross section occupied by the pipe's walls relative to the entire cross-sectional area. Substituting (137) into (135), one obtains the unstable area of the pipe:

$$l \geq x \geq l - \frac{u_0^2}{g\zeta(1+\epsilon)}, \qquad \epsilon = \frac{\rho}{\rho^1} \tag{138}$$

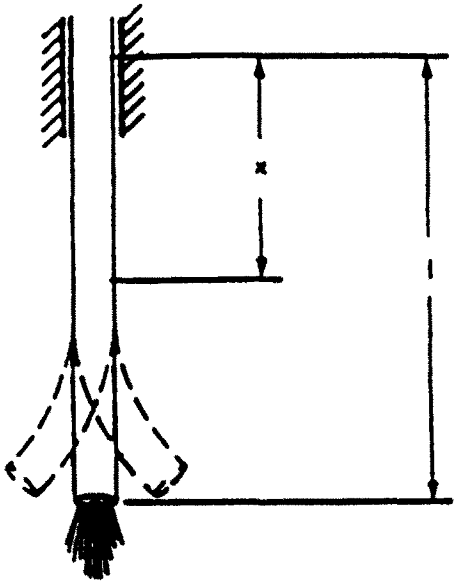

Fig. 5

Hence, for the ideal flexible pipe, the free end is always unstable. (Such a phenomenon is well known from experiments.) In the limit case $u_0 \to 0$, when the pipe can be considered as a string, the unstable area is concentrated around the free end. As shown by Zak (1970, 1983), such an instability manifests itself in an accumulation of energy at the free end (snap of a whip).

(b) For two-dimensional continua, such as membranes, films, and nets, as shown by Zak (1979), the characteristic speed corresponding to discontinuities of the shape (i.e., the coefficients of the second fundamental form) is given by

$$\lambda_{1,2} = \pm\left(\frac{T_n}{\rho}\right)^{1/2} \tag{139}$$

where T_n is the tension normal to the front of the wave of a discontinuity.

Hence, the failure of hyperbolicity emerges in the region where at least one of the principal stresses is negative. Such a failure manifests itself in the formation of wrinkles. The wrinkles can be observed, for instance, in the course of shearing, twisting, or bending of a membrane (Fig. 6). If both of the principal stresses are negative, then even the lines of wrinkles lose their smoothness, and a membrane can be rolled up in a point.

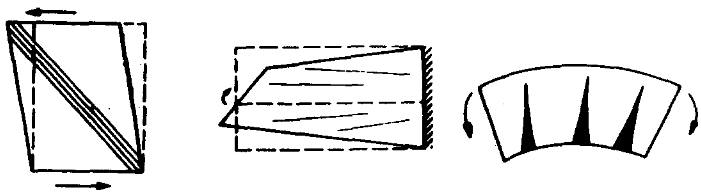

Fig. 6

Recall that in contrast to one-dimensional continua, where the shape parameters (curvature and twist) can be changed independently of the elongations, in two-dimensional continua there are some limitations imposed on the changes of the shape in the form of the equations of compatibility with the changes of strains [the Gauss equations (71) and (72)]. As follows from equations (78)–(82), at the points of negative Gaussian curvature, there are two directions of possible shape wave propagations (Fig. 7b). At the point of zero Gaussian curvature, there is only one such direction (Fig. 7c). Finally, at the points of positive Gaussian curvature, the shape discontinuities are impossible (Fig. 7d).

Thus the instability of the shape defined in terms of the coefficients of the second fundamental form b_{ij} is possible only if a compression occurs in

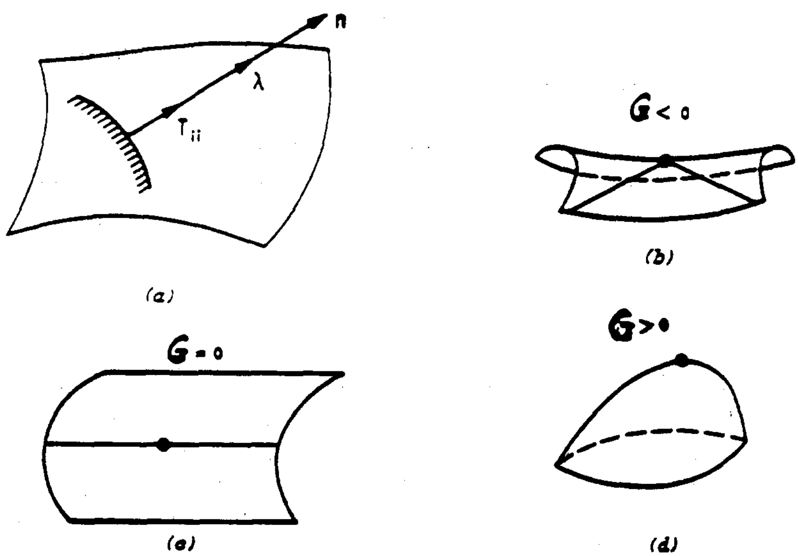

Fig. 7

the direction normal to the asymptotic line of the surface at the corresponding point.

Slightly different criteria of the Hadamard instability can be obtained for liquid films considered as two-dimensional continua (Zak, 1985*a*).

The Hadamard instability for three-dimension continua was studied by Zak (1982*a*–*c*). In this paper we will focus our attention on the instability of a surface separating an elastic body and ideal fluid (Fig. 8).

As shown by Zak (1982*b*), the characteristic speed of waves transporting discontinuities of the surface shape is expressed as

$$\mu = \frac{\rho_2}{\rho_1 + \rho_2} v \pm \left[\frac{1}{\rho_1 + \rho_2} T_{nn} + \frac{E}{2(1+\nu)} - \frac{\rho_1 \rho_2}{(\rho_1 + \rho_2)^2} v^2 \right]^{1/2} \tag{140}$$

in which ρ_1, E, ν, and T_{nn} characterize the density, Young's modulus, Poisson ratio, and the stress normal to the front of the propagating wave of the elastic body, and ρ_2 and v characterize the density and the velocity of the fluid.

Hence, the Hadamard instability occurs if:

$$T_{nn} < \frac{\rho_1 \rho_2}{(\rho_1 + \rho_2)^2} v^2 - \frac{E}{2(1+\nu)} \tag{141}$$

As a particular case of equation (140), one can arrive at the Hadamard instability of surface of a tangential jump in velocity in an inviscid fluid (Fig. 9)

$$\lambda = \frac{1}{2} \{(u_2 - u_1) \pm [-(u_2 - u_1)^2]^{1/2}\} = \frac{1}{2}(u_2 - u_1)(1 \pm i) \tag{142}$$

This is a well-known result stating that tangential jumps of velocities in

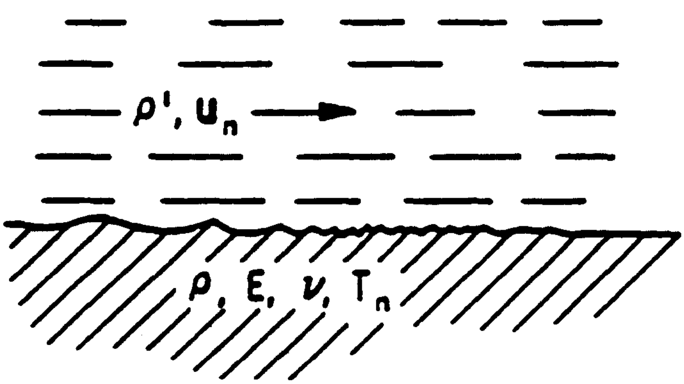

Fig. 8

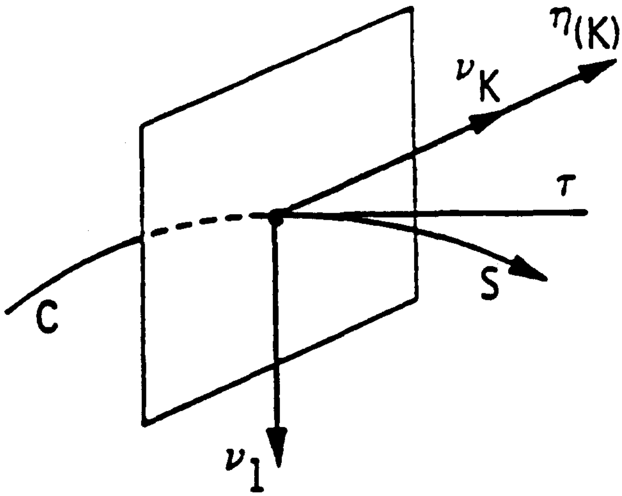

Fig. 9

inviscid fluids are always unstable. (In fluid mechanics this phenomenon is called the Kelvin–Helmholtz instability.)

2.4. Cumulative Effects

2.4.1. Degenerating Hyperbolic Equations

A cumulative effect can be introduced as a preinstability state which is associated with the change of type of governing equation from hyperbolic to parabolic when at least one of the characteristic speeds becomes zero. Actually this represents the boundary for the Hadamard instability, and depending on how the motion approaches this boundary, it may remain stable or unstable. The simplest example of this type of situation is the governing equation for a vertical, ideally flexible, inextensible string with a free lower end suspended in a gravity field (Fig. 8). Projecting this equation into the horizontal direction, one arrives at the governing equations for small transverse motion of the string:

$$\frac{\partial^2 x}{\partial t^2}+\frac{T}{\rho}\frac{\partial^2 x}{\partial \psi^2}=0 \tag{143}$$

with the characteristic speeds

$$\mu = \pm\left(\frac{T}{\rho}\right)^{1/2} \tag{144}$$

Since the tension of the string T vanishes at the free end,

$$T = 0 \qquad \text{at } S = l \tag{145}$$

where l is the length of the string, the characteristic speeds (144) vanish, too, at $S = l$, and therefore equation (143) degenerate into parabolic type at the very end of the string.

As a second example, consider a one-dimensional model of the shear wave propagation in a soil column of height H:

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(G \frac{\partial u}{\partial x}\right) \tag{146}$$

where ρ is the density, u is the horizontal displacement, G is the shear modules, t is time, and x is the vertical coordinate with the origin at the surface.

Ignoring the small shear stresses at the surface, one can take the shear modulus in the following form:

$$G = 0.5\rho g x \tag{147}$$

Since

$$G = 0 \qquad \text{at} \quad x = 0 \tag{148}$$

equation (146) degenerates into parabolic type at the soil surface.

For the sake of concreteness, we will investigate the solution to equation (146) subject to the initial and boundary conditions

$$u(x, 0) = \varphi(x), \qquad \frac{\partial u}{\partial x}(x, 0) = \psi(x), \qquad U(H, t) = \alpha(t), \qquad \frac{\partial u}{\partial x}(0, t) = 0 \tag{149}$$

Thus, it is assumed that the soil column is fixed at $x = H$ and there is no shear stress at the surface, i.e., at $x = 0$.

One should notice that for simplicity in this model all the damping and creep effects are ignored.

2.4.2. Uniqueness of the Solution

Let us assume that there exist two solutions of the problem under consideration, $u'(x, t)$ and $u''(x, t)$, and let us examine the difference

$$u^*(x, t) = u'(x, t) - u''(x, t) \tag{150}$$

The function $u^*(x, t)$ satisfies equation (146) with additional homogeneous conditions:

$$\rho \frac{\partial^2 u^*}{\partial t^2} = \frac{\partial}{\partial x}\left(G \frac{\partial u^*}{\partial x}\right) \tag{151}$$

$$u^*(x,0)=0, \qquad \frac{\partial u^*}{\partial t}(x,0)=0 \tag{152}$$

$$u^*(H,t)=0, \qquad \frac{\partial u^*}{\partial x}(0,t)=0 \tag{153}$$

For the total energy, one gets

$$E(t)=E(0)=\frac{1}{2}\int_0^H \left\{G\left(\frac{\partial u^*}{\partial x}\right)^2+\rho\left(\frac{\partial u^*}{\partial t}\right)^2\right\}\bigg|_{t=0} dx=0 \tag{154}$$

If the solution is sought in the open interval

$$0<x\leq H \tag{155}$$

which does not include the surface point $x=0$, then the uniqueness of the solution is obvious.

However, this proof cannot be applied to the closed interval

$$0\leq x\leq H \tag{156}$$

which includes the surface point $x=0$. Indeed, in this case, according to equation (148),

$$G=0 \qquad \text{at} \quad x=0 \tag{157}$$

and any arbitrarily selected derivative $\partial u^*/\partial x$ at $x=0$ will satisfy the equality in (154).

Thus, for the closed interval (156), the uniqueness of the solution can be guaranteed only in the class of functions having continuous derivative $\partial u^*/\partial x$, otherwise an infinite number of different solutions can be offered to satisfy equations (146) with the conditions (145) and (149) in the interval (156). As will be shown in the following, the artificial mathematical restriction about the continuity of the derivative $\partial u^*/\partial x$ excludes such important physical phenomena as cumulative effects.

From the mathematical point of view, the singularity at the point $x=0$ is associated with the fact that the original equation is hyperbolic in the open interval (155), but degenerates into a parabolic equation at the point $x=0$. The physical meaning of this singularity will be discussed in the following section.

2.4.3. Stability of the Solution

Starting with the conditions (148), let us assume that

$$\varphi(x)\begin{cases} >0 & \text{for} \quad 0<x_1^*<x<x_2^*<H \\ =0 & \text{for} \quad x<x_1 \quad \text{and} \quad x>x_2 \end{cases} \tag{158}$$

$$\psi(x)|_{0 \le x \le H} = 0 \tag{159}$$

$$\alpha(t) = 0, \qquad t \ge 0 \tag{160}$$

i.e., we consider an initial disturbance in a local interval $[x_1^*, x_2^*]$ contained within the interval $[0, H]$.

From the differential equation of the characteristics, one finds the equations of the characteristics passing through x_1 and x_2:

$$t' = \int_{x_1}^{x_2^*} \frac{d\epsilon}{[G(\epsilon)/\varphi]^{1/2}} \qquad t'' = \int_{x_2}^{x_2^*} \frac{d\epsilon}{[G(\epsilon)/\varphi]^{1/2}} \tag{161}$$

$$0 < x_1 < x_2 < H$$

Here x_1 and x_2 are the coordinates of the leading and trailing fronts of the discontinuity wave of derivatives $\partial^2 u/\partial t^2$ and $\partial^2 u/\partial x^2$, where

$$x_1 = x_1^*, \qquad x_2 = x_2^* \qquad \text{at} \quad t = 0 \tag{162}$$

A singular solution coincident for both characteristics holds for

$$x_1 = x_2 = 0$$

because

$$\left.\frac{dx_1}{dt}\right|_{x_1 = 0} = \left.\frac{dx_2}{dt}\right|_{x_2 = 0} = 0 \tag{163}$$

Two cases may arise: (A) The improper integral

$$\int_x^{x_0} \frac{d\eta}{[G(\eta)/\rho]^{1/2}} < \infty \qquad \text{for} \quad x \to 0 \tag{164}$$

i.e., converges, which then means that coincidence of the characteristics occurs for finite $t = t^*$. Then

$$\left|\frac{\partial u}{\partial t}\right| \to \infty \qquad \text{for} \quad t \to t^* < \infty \tag{165}$$

From the mathematical viewpoint this instability predicts an accumulation of the shear strain energy at the soil surface $x = 0$. At the same time, it illustrates the ambiguity in the solution which has been remarked in the investigation of equation (15).

(B) If the improper integral (164) diverges, then the characteristics (161) coincide at $t^* \to \infty$ and the accumulation effect does not occur.

For the particular case of soil where the shear modulus is given in the form (147) the integral (164) converges and the time t^* defining the moment of the formation of the shear strain energy accumulation at the soil surface is

$$t^* = 2\left(\frac{2x_2^*}{g}\right)^{1/2}$$

In the general case when the shear modulus is a more complicated function of the elevation the cumulative effect occurs if

$$G > \epsilon_1 x^{2+\epsilon_2}$$

where ϵ_1, ϵ_2 are arbitrarily small positive constants, because then the integral (164) converges.

2.4.4. Snap of a Whip

The results presented above can be applied to equation (163) describing transverse oscillations of a vertical, ideally flexible, inextensible string with a free lower end suspended in a gravity field.

The tension of the string due to gravity is given by

$$T = \gamma(l - x) \tag{166}$$

where γ and l are the specific weight and length of the string, respectively.

Referring to the formula (144), one concludes that the characteristic speed of transverse displacements tends to zero at the free end:

$$T|_{x=l} = 0, \qquad \tilde{\gamma}_{3,4} \to 0 \qquad \text{if} \quad x \to l \tag{167}$$

In other words, for small transverse displacements of the string, the governing equation is of hyperbolic type only in the open interval, excluding the end:

$$0 \le x < l \tag{168}$$

As shown in the previous section, in this open interval there exists a unique stable solution. However, in the closed interval, including the end

$$0 \le x \le l \tag{169}$$

the solution is not unique and there are unstable solutions if the improper integral

$$\int_0^x \frac{d\xi}{[T(\xi)/\rho]^{1/2}} \tag{170}$$

converges for $x \to l$.

This result has a very clear physical interpretation: Suppose that an isolated transverse wave of small amplitude was generated at the point of suspension (Fig. 1b). The speed of propagation of the leading front of the transverse wave will be smaller than the speed of the trailing front because the tension decreases from the point of suspension to the free end [see equations (144) and (166)]. Hence, the length of the above wave will be decreasing and in some cases [see (170)] will tend to zero. Then according to the law of conservation of energy, the specific kinetic energy per unit of length will tend to infinity, producing a snap (snap of a whip).

It can be easily verified by substituting (166) in (170) that for the string in the gravity field the integral (170) converges, i.e., instability in the form of snap occurs.

The same type of instability as a result of accumulation of energy near the boundary of the failure of hyperbolicity can exist in two- and three-dimensional models in the domains where the inequalities (116)–(118) are close enough to the corresponding equalities (Zak, 1982*c*).

2.4.5. Failure of Lipschitz Conditions

The cumulative effects are accompanied by a very interesting mathematical phenomenon: failure of Lipschitz conditions for the differential equations of characteristics:

$$\frac{ds}{ds} = \dot{s} = \lambda(s) \tag{171}$$

Indeed, if the characteristic speed follows from equation (143) or equation (146), i.e.,

$$\lambda = \pm\left(\frac{T}{S}\right)^{1/2} \quad \text{or} \quad \lambda = \pm\left(\frac{G}{S}\right)^{1/2}, \quad \text{respectively} \tag{172}$$

then

$$\left|\frac{\partial \lambda}{\partial s}\right| = \left|\frac{1}{2\lambda}\frac{\partial \lambda^2}{\partial s}\right| \to 0 \quad \text{at} \quad S \to S_0 \tag{173}$$

if

$$\left|\frac{\partial \lambda^2}{\partial S}\right| > 0 \quad \text{at} \quad S \to S_0$$

As follows from equations (147) and (166),

$$\left|\frac{\partial \lambda^2}{\partial s}\right| = 0.5g > 0 \quad \text{and} \quad \left|\frac{\partial \lambda^2}{\partial s}\right| = g > 0, \quad \text{respectively} \tag{174}$$

Hence, the loss of the uniqueness of the solution to equations (143) and (146) can be formally associated with the failure of the Lipschitz condition at the point where the characteristics coincide.

In general, failure of the Lipschitz conditions in dynamics was analyzed by Zak (1988, 1992, 1993*a*,*b*).

2.5. Comments on Other Types of Instability in Dynamics

As follows from the previous section, the Hadamard instability occurs in idealized models such as elastic bodies or ideal fluids where the energy dissipation can be ignored. The main property of this type of instability is that the solution becomes unbounded during a finite time interval ($t < \infty$). However, there are many other types of instability (especially in fluid dynamics) which also lead to unlimited decrease of the scale of motion, although they are not so "strong" as the Hadamard instability: the solution becomes unbounded only at $t \to \infty$. Since all of these Lyapunov-type instabilities are well represented in the literature, we will give here only a brief description of them.

Thermal instability arises when a fluid is heated from below. When the temperature difference across the fluid layer is great enough, the stability effects of viscosity and thermal conductivity are overcome by the destabilizing buoyancy, and the instability occurs in the form of a thermal convection.

Centrificial instability occurs in a fluid owing to the dynamical effects of rotation or of streamline curvature. For instance, as shown by Rayleigh, an inviscid flow between two rotating coaxial cylinders is unstable if the angular momentum $|r^2\Omega|$ decreases anywhere inside the interval $r_1 < r < r_2$, where Ω is the angular velocity of rotation of the fluid and r_1 and r_2 are the radii of the coaxial cylinders.

It can be demonstrated that in general, centrificial instability arises from adverse distributions of angular momentum.

Rayleigh–Taylor instability derives from the character of the equilibrium of an incompressible heavy fluid of variable density. For instance, it can be shown that in the case of a variable density of exponential distribution

$$\rho = \rho_0 e^{\beta z}, \qquad \beta = \text{const} \tag{175}$$

where z is the vertical coordinate, the equilibrium is unstable if

$$\beta > 0 \tag{176}$$

i.e., if the heavier layers are above the lighter layers.

Reynolds instability results from an imbalance between the inertial and viscous forces. It occurs when the Reynolds number R exceeds certain critical values which depend upon the type of flow and its boundary conditions. For a particular case of inviscid shear flow ($R \to \infty$) with parallel streamlines, Rayleigh showed that a necessary condition for instability is that the basic velocity profile should have an inflection point.

3. STABILIZATION PRINCIPLE

3.1. Instability as Inconsistency Between Models and Reality

3.1.1. General Remarks

It has been demonstrated in the previous section that there are some domains of dynamical parameters where the motion cannot be predicted because of instability of the solutions to the corresponding governing equations. How can this be interpreted? Does it mean that Newton's laws are not adequate? Or is there something wrong with our mathematical models? In order to answer these questions, we will discuss some general aspects of the concept of instability, and, in particular, the degree to which it is an invariant of motion. We will demonstrate that instability is an attribute of a mathematical model rather than a physical phenomenon, that it depends upon the frame of reference, the class of functions in which the motion is described, and the way in which the distances between the basic and perturbed solutions is defined.

3.1.2. Instability Dependence upon Metrics of Configuration Space

Let us turn to orbital instability discussed in Section 2.2. The metric of configuration space where the finite-degree-of-freedom dynamical system with N generalized coordinates q^i $(i = 1, 2, \ldots, N)$ is represented by a unit-mass particle was defined by equations (27) and (28). Now there are at least two possible ways to define the distance between the basic and disturbed trajectories. Following Synge (1926), we will consider the distance in both kinematic and kinematicostatic senses. In the first case the corresponding points on the trajectories are those for which time t has the same value. In the second case the correspondence between points on the basic trajectory C and a disturbed trajectory C^* is established by the condition that P (a point on C) should be the foot of the geodesic perpendicular let fall from P^* (a point on C^*) on C, i.e., here every point of the disturbed curve is adjacent to the undisturbed curve (regardless of the position of the moving particle at the instant t). As shown by Synge, both definitions of stability are invariant with respect to coordinate transformations, and in both cases the stability implies that the corresponding distance between the curves C and C^* remains permanently small.

It is obvious that stability in the kinematic sense implies stability in the kinematicostatic sense, but the converse is not true. Indeed, consider the motion of a particle of unit mass on a plane under the influence of a force system derivable from a potential:

$$\Pi = -x + \frac{1}{2} y^2 \tag{177}$$

Writing down the equations of motion and solving them we get

$$x = \frac{1}{2}t^2 + At + B \tag{178}$$

$$y = c\sin(t + \alpha) \tag{179}$$

where A, B, C, and D are constants of integration.

Let the undisturbed motion be

$$x = \frac{1}{2}t^2 + t \tag{180}$$

$$y = 0 \tag{181}$$

The motion is clearly unstable in the kinematic sense. However, from the viewpoint of stability in the kinematicostatic sense, the distance between corresponding points is

$$PP^* = y = C\sin(t + D) \tag{182}$$

This remains permanently small if C is small. Hence, there is stability in the kinematicostatic sense.

Thus, the same motion can be stable in one sense and unstable in another, depending upon the way in which the distance between the trajectories is defined.

It should be noticed that in both cases, the metric of configuration space was the same [see equations (27) and (28)]. However, as shown by Synge (1926), for conservative systems, one can introduce a configuration space with another metric:

$$g_{mn} = (E - \Pi)\alpha_{mn} \tag{183}$$

where α_{mn} are expressed by equation (27), and E is the total energy.

The system of motion trajectories here consists of all the geodesics of the manifold. The correspondence between points on the trajectories is fixed by the condition that the arc O^*P^* should be equal to the arc OP, where O and O^* are arbitrarily selected origins on the basic trajectory and any disturbed one, respectively.

As shown by Synge, the problem of stability here (which is called stability in the action sense) is that of the convergence of geodesics in Riemannian space. If two geodesics pass through adjacent points in nearly parallel directions, the distance between points on the geodesics equidistant from the respective initial points is either permanently small or not. If not, there is instability. It appears that stability in the action sense may not be equivalent to stability in the kinematicostatic sense for distances which change the total energy E.

Turning to example (177), let us take the initial point O at the origin of coordinates and the initial point O^* on the y axis. Then the disturbance being infinitesimal, the (action) distance between corresponding points is

$$P^* = (E - \Pi)^{1/2} y = 2^{-1/2}(t + 1)C \sin(t + D) \tag{184}$$

Hence, the motion is unstable in the action sense.

3.1.3. Instability Dependence upon the Frame of Reference

Dynamical instability depends not only upon the metric in which the distances between trajectories are defined, but also upon the frame of reference in which the motion is described. Such a dependence was already noticed above [see equations (55)]. In this section we will introduce and discuss an example which illustrates the dependence of criteria of hydrodynamic stability and the onset of turbulence upon the frame of reference.

The linear theory of hydrodynamic stability is based upon the Eulerian representation of fluid motions in which the frame of reference is chosen *a priori*. Strictly speaking, such a representation provides criteria of stability for the velocity field rather than the fluid motion. The difference between these two types of stability was illustrated by Arnold (1988), who introduced flows with stable velocity fields and unstable trajectories (Lagrangian turbulence). If the classical (Eulerian) turbulence is associated with the instability of streamlines, then it is reasonable to study this instability in a streamline frame of reference in which streamlines form a family of initially unknown Eulerian coordinates, while the remaining two Lagrangian coordinates are found from the compatibility conditions. Such a frame of reference is completely defined by the motion, and therefore it contains a minimum of arbitrarily chosen parameters.

First of all, we will show that criteria of stability in this frame of reference do not necessarily coincide with the classical criteria which are derived from the Orr–Sommerfeld equation. For this purpose, we will introduce a small disturbance velocity field for incompressible plane flow in Cartesian coordinates x, y:

$$V_x = \phi'(y)e^{i(\alpha x - \beta t)}, \qquad V_y = -i\alpha\phi(y)e^{i(\alpha x - \beta t)}, \qquad \alpha, \beta = \text{const} \tag{185}$$

where the prime denotes differentiation.

The angle θ between streamlines and the x direction is

$$\theta = \frac{i\alpha\phi}{\overset{0}{V}} e^{i(\alpha x - \beta t)} \tag{186}$$

in which $\overset{0}{V}(y)$ is the velocity profile of the basic flow. The orthogonal streamline coordinates ξ, ζ are found from the system

$$\frac{\partial x}{\partial \xi} = H_1 \cos\theta, \qquad \frac{\partial x}{\partial \zeta} = -H_2 \sin\theta, \qquad \frac{\partial y}{\partial \xi} = H_1 \sin\theta, \qquad \frac{\partial y}{\partial \zeta} = H_2 \cos\theta \tag{187}$$

where H_1 and H_2 are the Lamé coefficients defined by the compatibility conditions ($\partial^2 x/\partial\xi\,\partial\zeta = \partial^2 x/\partial\xi\,\partial\xi$, etc.)

$$\left(\frac{\partial\theta}{\partial x} + \frac{\partial\theta}{\partial y}\tan\theta\right)H_1 = \tan\theta\frac{\partial H_1}{\partial x} - \frac{\partial H_1}{\partial y} \tag{188}$$

and

$$\left(-\frac{\partial\theta}{\partial y} + \frac{\partial\theta}{\partial x}\tan\theta\right)H_2 = \tan\theta\frac{\partial H_2}{\partial y} - \frac{\partial H_2}{\partial x} \tag{189}$$

It follows from equations (186)–(189) that the coordinate transformation

$$x = x(\xi, \zeta, t), \qquad y = y(\xi, \zeta, t) \tag{190}$$

in general will depend on time. Hence, for the stream function one obtains

$$\psi = \phi(y)e^{i(\alpha x - \beta t)} = \phi[y(\xi, \zeta, t]e^{i[\alpha x(\xi, \zeta, t) - \alpha t]} \tag{191}$$

i.e.,

$$\left.\frac{\partial\psi}{\partial t}\right|^{x,y=\text{const}} \neq \left.\frac{\partial\psi}{\partial t}\right|_{\xi,\zeta=\text{const}} \tag{192}$$

In other words, the stability criteria in frames x, y and ξ, ζ are not necessarily the same.

This preliminary conclusion provides motivation to analyze criteria of hydrodynamic stability in streamline coordinates.

Confining our investigation to a plane incompressible inviscid flow, we derive the momentum equations in streamline coordinates from the Lagrange equation

$$\frac{d}{dt}\frac{\partial w}{\partial \xi} - \frac{\partial w}{\partial \xi} = \frac{1}{\rho}\frac{\partial p}{\partial \xi}, \qquad -\frac{\partial w}{\partial \zeta} = -\frac{1}{\rho}\frac{\partial p}{\partial \zeta} \tag{193}$$

in which the kinetic energy is

$$W = \frac{1}{2}H_1^2\xi^2 \tag{194}$$

and the velocity is

$$V = V_1 = H_1\xi, \qquad V_2 = H_2\xi = 0$$

while p and ρ are pressure and density, respectively.

The momentum equations read

$$H_1 \frac{\partial V}{\partial t} + V\left(\frac{\partial H_1}{\partial t} + \frac{\partial V}{\partial \xi}\right) = -\frac{1}{\rho}\frac{\partial p}{\partial \xi} \tag{195}$$

and

$$-\frac{V^2}{H_1 H_2}\frac{\partial H_1}{\partial \xi} = \frac{1}{\rho}\frac{\partial p}{\partial \zeta} \tag{196}$$

The continuity equation follows from the condition

$$\operatorname{div} V = 0, \qquad \text{i.e.,} \qquad \frac{\partial}{\partial \xi}(VH_2) = 0 \tag{197}$$

Equations (187)–(189) are completed by the compatibility (Lamé) equation

$$\frac{\partial}{\partial \xi}\left(\frac{1}{H_1}\frac{\partial H_2}{\partial \xi}\right) + \frac{\partial}{\partial \zeta}\left(\frac{1}{H_2}\frac{\partial H_1}{\partial \zeta}\right) = 0 \tag{198}$$

Linearizing these equations with respect to an unperturbed shear flow.

$$\mathring{V} = \mathring{V}(y) \tag{199}$$

and taking into account that for this flow the streamline coordinates coincide with the Cartesian coordinates

$$\xi = x, \qquad \zeta = y, \qquad \tilde{H}_1 = \tilde{H}_2 = 1 \tag{200}$$

one obtains after eliminating the pressure

$$\frac{\partial^2 \tilde{V}}{\partial t\, \partial \zeta}\frac{\partial^2 \tilde{H}_1}{\partial t\, \partial \zeta} + \mathring{V}(\zeta)\frac{\partial^2 \tilde{V}}{\partial \xi\, \partial \zeta} + \mathring{V}^2(\zeta)\frac{\partial^2 \tilde{H}_1}{\partial \xi\, \partial \zeta} = 0 \tag{201}$$

$$\frac{\partial \tilde{V}}{\partial \xi} + \mathring{V}(\zeta)\frac{\partial \tilde{H}_2}{\partial \xi} = 0 \tag{202}$$

$$\frac{\partial^2 \tilde{H}_2}{\partial \xi^2} + \frac{\partial^2 \tilde{H}_1}{\partial \zeta^2} = 0 \tag{203}$$

where $\tilde{V}$, $\tilde{H}_1$, and $\tilde{H}_2$ are small perturbations of V, H_1, and H_2, respectively.

If the solution for $\tilde{V}$ is assumed to be of the form

$$\tilde{V} = \mathring{V}(\zeta)\phi'(\zeta)e^{i(\alpha\xi - \beta t)}, \qquad \alpha, \beta = \text{const} \tag{204}$$

then it follows from equations (202) and (203) that

$$\tilde{H}_2 = -\phi'(\zeta)e^{i(\alpha\xi - \beta t)}, \qquad -\frac{\partial \tilde{H}_1}{\partial \zeta} = -\alpha^2\phi(\zeta)e^{i(\alpha\xi - \beta t)} \tag{205}$$

Substituting the values (204) and (205) into equation (201), one arrives at the governing equation for $\phi(\xi)$:

$$\phi'' - \frac{c\mathring{V}'(\zeta)}{\mathring{V}(\zeta)[\mathring{V}(\zeta) - c]}\phi' - \alpha^2\phi = 0, \qquad c = \frac{\beta}{\alpha} \tag{206}$$

which is different from the Orr–Sommerfeld equation.

If the basic flow $\mathring{V}(y)$ is bounded by rigid walls

$$y = y_1, \qquad y = y_2 \tag{207}$$

then the streamlines at $\eta = y_1$ and $\zeta = y_2$ must coincide with these walls, i.e.,

$$\frac{\partial \tau}{\partial \xi} = -\frac{1}{H_2}\frac{\partial H_1}{\partial \zeta}\mathbf{n} = 0 \qquad \text{at} \quad y = y_1 \quad \text{and} \quad y = y_2 \tag{208}$$

in which τ and $\mathbf{n}$ are the unit tangent and the unit normal vectors to the streamlines, respectively.

Hence

$$\frac{\partial H_1}{\partial \zeta} = 0 \qquad \text{at} \quad y = y_1 \quad \text{and} \quad y = y_2 \tag{209}$$

and therefore, with references to equations (205),

$$\phi(\zeta_1 = y_1) = 0, \qquad \phi(\zeta_2 = y_2) = 0 \tag{210}$$

These equations express the boundary conditions for equations (206).

In order to show that the stability criteria in streamline coordinates are different from those given by the Orr–Sommerfeld equations, let us select a special velocity profile $\mathring{V}(y)$ such that the coefficient of ϕ' in equation (206) reduces to a constant. Obviously, such a profile must satisfy the first-order differential equation

$$\frac{\mathring{V}'}{\mathring{V}(\mathring{V} - c)} = \gamma = \text{const}, \qquad \operatorname{Im} \gamma = 0 \tag{211}$$

and consequently

$$\mathring{V} = \frac{c}{1 - e^{c\gamma y}} \tag{212}$$

while equations (206) for this profile reduces to

$$\phi'' - c\gamma\phi' - \alpha^2\phi = 0 \tag{213}$$

Its general solution is

$$\phi = C_1 \exp(\lambda_1 y) + C_2 \exp(\lambda_2 y) \tag{214}$$

where

$$\lambda_{1,2} = \frac{cy}{2} + \left(\frac{c^2y^2}{4} + \alpha^2\right)^{1/2} \tag{215}$$

Substitution of the boundary conditions (210) into equation (214) leads to a system of homogeneous equations:

$$\begin{aligned} C_1 \exp(\lambda_1 y_1) + C_2 \exp(\lambda_2 y_1) = 0 \\ C_1 \exp(\lambda_1 y_2) + C_2 \exp(\lambda_2 y_2) = 0 \end{aligned} \tag{216}$$

and for a nontrivial solution

$$\begin{aligned} &\det\begin{pmatrix} \exp(\lambda_1 y_1) & \exp(\lambda_2 y_1) \\ \exp(\lambda_1 y_2) & \exp(\lambda_2 y_2) \end{pmatrix} \\ &\quad = \exp(\lambda_1 y_1 + \lambda_2 y_2) - \exp(\lambda_1 y_2 + \lambda_2 y_1) = 0 \end{aligned} \tag{217}$$

i.e., $\lambda_1 = \lambda_2$; or, with reference to equation (215),

$$c = \pm i2\alpha/y = \pm c_0 i \tag{218}$$

Since α and γ are real, c is imaginary, and therefore solutions (205) are unstable for any y_1 and y_2.

Now we will show that the Orr–Sommerfeld equation predicts stability for the same profile. Indeed, substituting c from equation (218) into equation (217) and separating the real part of the velocity profile, one obtains

$$\mathrm{Re}\, \dot{V} = \pm \mathrm{cotan}(2\alpha y) \tag{219}$$

This profile has only one inflection point [at $y = \pi/(4\alpha)$]. Consequently, according to the point-of-inflection criterion proved by Tollmien, any profile of the form (219) which does include the inflection point, i.e.,

$$0 \le y_1 \le y \le y_2 < \frac{\pi}{4\alpha} \tag{220}$$

is stable.

It is important to emphasize that these two different results regarding the same velocity profile are not mutually exclusive: the first is related to the stability of the fluid motion referred to streamline coordinates, while the second is related to the stability of the velocity field. But which of these approaches is actually related to the onset of turbulence? The dynamics of fluid motion, and in particular the stability of streamlines, are directly related to the onset of turbulence inasmuch as the stability of particle trajectories is directly related to the onset of Lagrangian turbulence. At the same time, the stability of velocity fields is indirectly related to the onset of turbulence. That is why the linearized version of the classical theory of stability cannot explain the instability of plane Couette flows. In this connection it is worth noting that by an appropriate selection of α, y_1, and y_2 in equations (219) and (220), the velocity profile (219) can be made as close as necessary to a straight line, thereby predicting the instability of any flow which is arbitrarily close to the Couette flow.

3.1.4. Instability Dependence upon the Class of Functions

The properties of solutions to differential equations such as existence, uniqueness, and stability have a mathematical meaning only if they are referred to a certain class of functions. For instance, as shown above in equations (143) and (146), we have a unique stable solution in an open interval (155) in the class of bounded functions, while in a closed interval (156), the uniqueness and stability are not guaranteed. Most of the results concerning the properties of solutions to differential equations require differentiability (up to a certain order) of the functions describing the solutions. However, the mathematical restrictions imposed upon the class of functions which guarantee the existence of a unique and stable solution do not necessarily lead to the best representation of the corresponding physical phenomenon. Indeed, turning again to equations (143) and (145), one notices that the unique and stable solution (155) does not describe accumulation effect (a snap of a whip) which is well pronounced in experiments. At the same time, an unstable solution in a closed interval (156) gives a qualitative description of this effect. Hence, pure mathematical restrictions imposed upon the solutions are not always consistent with the physical nature of motions. In this context, the long-term instability in classical dynamics discussed in Section 2 can be interpreted as a discrepancy between these mathematical restrictions and physical reality. This means that unpredictability in classical dynamics is the price paid for mathematical "convenience" in dealing with dynamical models. Therefore, the concept of unpredictability in dynamics should be put as unpredictability in a selected class of functions, or in a selected metric of configuration space, or in a selected frame of reference.

Now the following problem can be posed. How does one select an appropriate mathematical representation of a physical phenomenon? The answer to this question will be discussed below.

3.2. Dynamics in Rapidly Oscillating Frame of Reference

As shown in the previous sections, the instability and therefore the occurrence of chaos or turbulence in the description of mechanical motions means only that these motions cannot be properly described by smooth functions if the scale of observations is limited. These arguments can be linked to Gödel's (1931) incompleteness theorem and Richardson's (1968) proof that the theory of elementary functions in classical analysis is undecidable.

But since instability is not an invariant of motions, the following question can be posed: it is possible to find such a new (enlarged) class of functions, or a new metric of configuration space, or a new frame of reference in order to eliminate instability? Actually such a possibility would lead to different representative parameters describing the same motion in such a way that small uncertainties in external forces cause small changes of these parameters. For example, in turbulent and chaotic motions, mean velocities, Reynolds stresses, and power spectra represent "stable" parameters, although classical governing equations neither are explicitly expressed via these parameters nor uniquely define them.

The first step toward the enlarging of the class of functions for modeling turbulence was made by Reynolds (1895), who decomposed the velocity field into the mean and pulsating components, and actually introduced a multivalued velocity field. However, this decomposition brought new unknowns without additional governing equations, and that created a "closure" problem. In 1986 Zak showed that the Reynolds equations can be obtained by referring the Navier–Stokes equations to a rapidly oscillating frame of reference, while the Reynolds stresses represent the contribution of inertia forces. From this viewpoint the "closure" has the same status as a "proof" of Euclid's parallel postulate, since the motion of the frame of reference can be chosen arbitrarily. In other words, the "closure" of the Reynolds equations represents a case of undecidability in classical mechanics. However, based upon the interpretation of the Reynolds stresses as inertia forces, it is reasonable to choose the motion of the frame of reference such that the inertia forces eliminate the original instability. In other words, the enlarged class of functions should be selected such that the solution of the original problem in that class of functions will not possess an exponential sensitivity to changes in initial conditions. This stabilization principle has been formulated and applied to chaotic and turbulent motions by Zak (1982, 1985*a*, 1986*a,b*, 1990). As shown there, the motions which are chaotic (or turbulent) in the original frame of reference can be represented as a sum of the "mean" motion and rapid fluctuations, while both components are uniquely defined. It is worth emphasizing that the amplitude of velocity fluctuation is proportional to the degree of the original instability, and therefore the rapid fluctuations can be associated with the measure of the uncertainty in the description of the motion. It should be noticed that both "mean" and "fluctuation" components representing the originally chaotic motion are stable, i.e., they are not sensitive to changes of initial conditions, and are fully reproducible.

Let us refer the original equation to the motion of a noninertial frame of reference which rapidly oscillates with respect to the original inertial frame of reference. Then the absolute velocity $\dot{q}$ can be decomposed into the relative velocity $\dot{q}_1$ and the transport velocity $\dot{q}_2 = 2\dot{q}_{2(0)}$:

$$\dot{q} = \dot{q}_1 + 2\dot{q}_{2(0)} \cos \omega \to \infty \tag{221}$$

while $\dot{q}_1$ and $\dot{q}_2^0$ are "slow" functions of time in the sense that

$$\omega \gg \frac{1}{\tau} \tag{222}$$

where τ is the time scale upon which the changes $\dot{q}_1$ and $\dot{q}_{2(0)}$ can be ignored.

Then for the mean $\bar{q}$

$$q \cong q_1 \quad \text{since} \quad \int_0^{t \gg \tau} \dot{q}_{2(0)} \cos \omega t \, dt \simeq \frac{1}{\omega} \dot{q}_{2(0)} \sin \omega t \to 0 \qquad \text{if} \quad \omega \to \infty \tag{223}$$

In other words, rapidly oscillating velocity practically does not change the displacements.

Taking into account that

$$\begin{gathered} \frac{\omega}{2\pi} \int_0^{2\pi/\omega} \dot{q}_1 \, dt \simeq \dot{q}_1 \\ \int_0^{2\pi/\omega} \dot{q}_{2(0)} \sin \omega t \, dt = 0 \\ \int_0^{2\pi/\omega} \dot{q}_{2(0)}^2 \cos^2 \omega t \, dt = \frac{1}{2} \dot{q}_{2(0)}^2 \end{gathered} \tag{224}$$

one can transform the system

$$\dot{x}^i = a_j^i + b_{jm}^i x^i x^m, \qquad i = 1, 2, \ldots, n \tag{225}$$

into the following form:

$$\dot{\bar{x}}_i = a_j^i \bar{x}^j + b_{jm}^i \bar{x}^j \bar{x}^m + b_{mj}^i \overline{x^j x^m}, \qquad i = 1, 2, \ldots, n \tag{226}$$

where $\bar{x}^i$ and $\overline{x^i x^j}$ are means and double correlations of x^i as random variables, respectively.

As will be shown below, the transition from (225) to (226) is identical to the Reynolds transformation: i.e., applied to the Navier–Stokes equations, it leads to the Reynolds equations, and therefore the last term in (226) (which is a contribution of inertial forces due to fast oscillations of the frame of reference) can be identified with the Reynolds stresses. From a mathematical viewpoint, this transformation is interpretable as enlarging the class of smooth functions to multivalued ones. Indeed, as follows from (222), for any arbitrarily small interval Δt, there always exists such a large frequency $\omega > \Delta t/2\pi$ that within this interval the velocity $\dot{q}$ runs through all its values, and actually the velocity field becomes multivalued.

Clearly equations (226) result from time averaging. In case of applicability of the ergodic hypothesis, the same equations can be obtained from ensemble averaging. However, formally the averaging procedure can be introduced axiomatically based upon the Reynolds conditions:

$$\overline{a+b} = \bar{a} + \bar{b}, \qquad \overline{ka} = k\bar{a}, \qquad \bar{k} = k \qquad (k = \text{const})$$
$$\overline{\frac{\partial a}{\partial l}} = \frac{\partial \bar{a}}{\partial l}, \qquad \overline{\bar{a}b} = \bar{a}\bar{b}$$

This leads to the identity

$$\overline{ab} = \bar{a}\bar{b} + \overline{a'b'}$$

where $a = \bar{a} + a'$ and $b = \bar{b} + b'$.

Let us consider a mechanical system with N degrees of freedom and the kinetic energy

$$W = a_{ij}\dot{q}^i\dot{q}^j \tag{227}$$

where q^i and $\dot{q}^i$ are the generalized coordinates and velocities, respectively, and introduce an N-dimensional (abstract) space with the metric

$$ds^2 = a_{sk}\, dg^s\, dq^k, \qquad \dot{s}^2 = 2W \tag{228}$$

Then the equations of motion

$$q^i = q^i(t) \tag{229}$$

satisfy the following differential equation:

$$\ddot{q}^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{q}^\beta\dot{q}^\gamma = Q^\alpha \tag{230}$$

where Q^α is the force vector and $\Gamma^\alpha_{\beta\gamma}$ are the Christoffel symbols:

$$\begin{aligned} \Gamma^e_{sk} &= \frac{1}{2}\alpha^{ip}\left[\frac{\partial a_{sp}}{\partial q^k} + \frac{\partial a_{kp}}{\partial q^s} - \frac{\partial a_{sk}}{\partial q^p}\right] \\ a^{\alpha\beta}a_{\beta\gamma} &= \delta^\alpha_\gamma = \begin{cases} 0 & \text{if} \quad \alpha \neq \gamma \\ 1 & \text{if} \quad \alpha = \gamma \end{cases} \end{aligned} \tag{231}$$

Equation (230) can be interpreted as a parametric equation of the trajectory C of a representative point M with the contravariant coordinates q^α. The unit tangent vector $\tau = v_0$ to this trajectory is defined as

$$\tau^\alpha = v_0^\alpha = \frac{dq^\alpha}{ds} = \frac{1}{(2W)^{1/2}}\dot{q}^\alpha, \qquad \alpha_{mn}v_0^m v_0^n = 1 \tag{232}$$

while the unit normals $\nu_1, \nu_2, \ldots, \nu_{N-1}$ are given by the Frenet equations:

$$\frac{d\nu_p^i}{ds} + \Gamma_{kq}^i \nu_p^q \frac{dq^k}{ds} = -\chi_p \nu_{p-1}^i + \chi_{p+1} \nu_{p+1}^i \tag{233}$$

where $\chi_1, \chi_2, \ldots, \nu_{N-1}$ are the curvatures of the trajectory and S is the arc coordinate along this trajectory.

The principal normal ν_1 is coplanar with the tangent ν_0 and the force vector Q. The remaining curvatures as well as the directions of the rest normals are defined by equation (233); see Fig. 9.

For simplicity we will confine ourselves to the particular case when

$$\Gamma_{sk}^l = \text{const} \tag{234}$$

Substituting the decomposition (221) into equation (230), one obtains

$$\ddot{q}_1^\alpha + \Gamma_{\beta\delta}^\alpha \dot{q}_1^\beta \dot{q}_1^\gamma + \Gamma_{\beta\delta}^\alpha \dot{q}_{2(0)}^\beta \dot{q}_{2(0)}^\alpha = Q^\alpha \tag{235}$$

Here the terms

$$Q_{(i)}^\alpha = -\Gamma_{\beta\gamma}^\alpha \dot{q}_{2(0)}^\beta \dot{q}_{2(0)}^\gamma \tag{236}$$

represent the inertia forces caused by the transport motion of the frame of reference.

Applying a velocity decomposition similar to (221)

$$\mathbf{v} = \bar{\mathbf{v}} + 2\tilde{\mathbf{v}} \cos wt, \qquad \omega \to \infty \tag{237}$$

to the momentum equation for a continuum in Eulerian representation

$$\rho\left(\frac{\partial \bar{\mathbf{v}}}{\partial t} + \bar{\mathbf{v}}\, \nabla \bar{\mathbf{v}}\right) = \nabla \cdot \sigma \tag{238}$$

where σ is the stress tensor, one obtains

$$\rho\left(\frac{\partial \bar{\mathbf{v}}}{\partial t} + \bar{\mathbf{v}}\, \nabla \bar{\mathbf{v}}\right) = \nabla \cdot (\sigma + \tilde{\sigma}) \tag{239}$$

in which $\tilde{\sigma}$ is the Reynolds stress tensor with the components

$$\tilde{\sigma}_{ij} = -\rho \overline{\tilde{v}_i \tilde{v}_j} \tag{240}$$

In terms of the Reynolds equations, $\bar{\mathbf{v}}$ and $\tilde{\mathbf{v}}$ represent the mean velocity and the amplitude of fast velocity fluctuations, respectively.

The most significant advantage of the Reynolds-type equations (226), (235), and (239) is that they are explicitly expressed via the physically reproducible parameters $\overline{x^i}$, $\overline{x^i x^j}$ which describe for instance, a mean velocity profile in turbulent motions, or a power spectrum of chaotic attractors. However, as a price for that, these equations require a closure since the number of unknowns is larger than the number of equations. Actually the closure problem has existed for almost 100 years, since the Reynolds equations were derived. In the next sections, based upon the stabilization principle introduced by Zak (1985*a*, 1986*a*,*b*, 1990), this problem will be discussed.

3.3. Stabilization Principle and the Closure Problem

3.3.1. General Remarks

Revisiting the dynamical systems (226), (235), and (239) which describe motions in the class of multivalued functions, one notes that these systems are not complete, in the sense that the number of unknowns is larger than the number of equations. In particular, the vector which expresses the bulk contribution of the "microscale" motions to the averaged motion represents excessive unknowns. Such an incompleteness creates a closure problem. This problem first was identified in connection with the Reynolds equation describing turbulent motions. The problem of turbulence arose almost 100 years ago as a result of the discrepancy between theoretical fluid mechanics and experiments. However, in spite of considerable research activity, there is no general approach to the prediction of turbulence based upon theoretical models. Most effort has been directed toward finding a "physical" law which would couple the Reynolds stresses with the rate of strain of the average motion, and thereby would represent additional equations required for the closure of the Reynolds equations. For instance, Prandtl introduced the mixing length assumption

$$\tau = \rho l^2 \left|\frac{\partial u}{\partial y}\right| \frac{\partial u}{\partial y} \tag{241}$$

for the two-dimensional version of the Reynolds equation:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = \frac{1}{\rho}\frac{\partial \tau}{\partial y}, \qquad \tau = -\rho\overline{u_r v_r} \tag{242}$$

Here u, v, u_r, and v_r are the mean and fluctuation velocity projections on the Cartesian coordinates x and y, respectively, τ is the shear component of the Reynolds stress, and l is a so-called mixing length which is supposed to be found from experiments.

By exploiting the closure (241), Prandtl solved several problems of the two-dimensional theory of turbulence: he found the mean velocity profile of an axisymmetric turbulent flow in a pipe, described the smoothing out of velocity discontinuity, etc., while all of his solutions were sufficiently close to experimental results.

However, the same closure (241) failed to provide satisfactory solutions in many other cases, which means that the closure (241) cannot be considered as a "physical" law. But does any "physical" law of the type (241) exist in principle? And is such a law necessary for the closure? Indeed, as shown in the previous section, the Reynolds stresses can be interpreted as a contribution of the inertia forces of a rapidly oscillating frame of reference, while this frame of reference can be chosen arbitrarily! However, such an interpretation leads to another question: it is possible to find such a frame of reference which provides the "best" representation of the motion? Obviously, in this representation the motion must be stable, and therefore the restoration of stability of the originally unstable motion can be chosen as the main criterion for selection of the frame of reference, and therefore of the Reynolds stresses. From the mathematical viewpoint, this means that if the original motion is unstable in the class of smooth functions, this instability can be eliminated by enlarging the class of functions. From that viewpoint, Prandtl's closure (241) can be treated as a feedback which stabilizes an originally unstable laminar flow. Indeed, turning, for instance, to a plane Poisson flow with a parabolic velocity profile, one arrives at its instability if the Reynolds number is larger than $R_{\mathrm{cr}} \cong 5772$. Experiments show that a new steady turbulent profile is no longer parabolic: it is very flat near the center and is very steep near the walls. The same profile follows from the Prandtl solution based upon the closure (241). But since this profile can be experimentally observed, it must be stable, and this stabilization is carried out by the "feedback" (241).

3.3.2. Formulation of the Stabilization Principle

Based upon remarks made in the previous section, we will now formulate the following stabilization principle. Consider a dynamical model which in some domain of its parameters becomes unstable in the class of differentiable functions, i.e., its instability leads to an unbounded growth of ignorable variables. As noticed earlier, this means that the corresponding physical phenomena cannot be adequately described in the class of differentiable functions, and the original model must be modified. The modification of the model should be based upon the enlarging the original class of functions in such a way that the instability is eliminated. This mathematical formulation can be complemented and specified by the following physical reasonings: The application of the Reynolds-averaging conditions to any nonlinear dynamical model leads to another nonlinear system which differs from the original one by additional variables—the Reynolds "stress" [see equations (226), (235), and (239)]. In a symbolic form, the transformation from the Newtonian (N) to the Reynolds (R) dynamics can be represented as

$$\mathrm{R} \cdot x = \mathrm{N} \cdot x + \sigma_{\mathrm{R}}$$

If the original dynamical system

$$\mathrm{N} \cdot x = 0$$

is unstable, but the Reynolds-averaged system

$$\mathrm{N} \cdot x + \sigma_{\mathrm{R}} = 0$$

is stable, obviously the stabilization is performed by the Reynolds stresses σ_{R}: driven by the mechanism of instability of the original model, they grow until the instability is suppressed down to a neutral stability. As will be shown below, the last condition may uniquely define σ_{R} as well as all the averaged parameters of the dynamical system. Mathematical justification of the neutral stability of Reynolds-averaged models will be given in Section 4.3 [see equations (275) and (276)].

Experimental verification of neutral stability of free turbulent jets was reported by Lessen and Poillet (1976).

In the next sections the stabilization principle will be applied for the prediction of the postinstability behavior of fluids (turbulent motions) and of finite-dimensional dynamical systems (chaos).

3.4. Application of the Stabilization Principle to Predictions of Chaotic Motions

The strategy for the application of the stabilization principle to predict chaotic motions for inertial, potential, and dissipative systems will be presented in this section.

3.4.1. Inertial Motions

In order to clarify the main idea of the approach, let us turn to the inertial motion of a particle M of unit mass in a smooth pseudosphere S having a constant negative curvature (15). As shown there, the orbital instability and therefore the chaotic behavior of the particle M can be eliminated by the elastic force (40),

$$F = -\alpha^2 \epsilon, \qquad \alpha^2 = \text{const} > -2WG, \qquad G < 0 \tag{243}$$

which is proportional to the normal deviation ϵ from the geodesic trajectory which is applied to the particle M. But such a force can appear as an inertial force if the motion of the particle M is referred to an appropriate noninertial system of coordinates.

Indeed, so far this motion has been referred to an inertial system of coordinates q_1, q_2, where q_1 is the coordinate along the geodesic meridians and q_2 is the coordinate along the parallels. Let us introduce now a frame of reference which rotates about the axis of symmetry of the pseudosphere with a rapidly oscillating transport velocity:

$$\dot{\epsilon} = 2\epsilon_0 \cos \omega t, \qquad \omega \to \infty \tag{244}$$

so that the components of the resultant velocity along the meridians and parallels are, respectively,

$$v_1 = \dot{q}_1, \qquad v_2 = \dot{q}_2 + 2\dot{\epsilon}_0 \cos \omega t \tag{245}$$

Since equation (245) has the same structure as equation (221), the Lagrangian of the motion of the particle M relative to the new (noninertial) frame of reference can be written in the form [see equation (22)]

$$L^* = \dot{q}_1^2 - \frac{1}{G_0}\{\exp[-2(-G_0)^{1/2}q_1]\}(\dot{q}_2^2 + \dot{\epsilon}_0^2) \tag{246}$$

The last term in equation (246) represents the contribution of the inertia forces in the new frame of reference.

So far the transport velocity $\dot{\epsilon}_0$ has not been specified, and therefore the Lagrangian (252) has the same element of arbitrariness as the governing equations (235) describing chaotic motions. Now, based upon the stabilization principle, we are going to specify the transport motion in such a way that the original orbital instability of the inertial motion of the particle M is eliminated. Turning to the condition (42), one obtains

$$\frac{\partial^2 L}{\partial \epsilon^2} \geq -2WG_0 \tag{247}$$

where $W = \frac{1}{2}mv_0^2$ is the kinetic energy of the particle. This condition can be satisfied if the transport velocity $\dot{\epsilon}_0$ is coupled with the normal deviation ϵ as follows:

$$-\frac{1}{G_0}\{\exp[-2(-G_0)^{1/2}q_1]\}\dot{\epsilon}_0^2 = -WG_0\epsilon^2 \tag{248}$$

As follows from equation (45), in this limit case the Lyapunov exponent of the relative motion in the new (noninertial) frame of reference will be zero:

$$\sigma = \left(-G_0 - \frac{\alpha^2}{W}\right)^{1/2} = 0, \qquad \alpha^2 = \frac{\partial^2 L}{\partial \epsilon^2} \tag{249}$$

and the trajectories of perturbed motions do not diverge. The normal deviation from the trajectory of the relative motion (in case of zero perturbed velocity $\dot{\epsilon}_0$) can be written in the following form:

$$\epsilon = \epsilon_0 = \text{const}, \qquad \epsilon_0 = \epsilon(t = 0) \tag{250}$$

which means that in the new frame of reference an initial error ϵ_0 does not grow—it remains constant. The relative motion along the trajectory is described by the differential equation following from the Lagrangian (246), which takes the following form [after substituting equation (248)]:

$$L = \dot{q}_1^2 - \frac{1}{G_0}\{\exp[-2(-G_0)^{1/2}q_1]\}\dot{q}_2 - WG_0\epsilon^2 \tag{251}$$

i.e.,

$$\ddot{q}_1 - \frac{2(-G_0)^{1/2}}{G_0}\{\exp[-2(-G_0)^{1/2}q_1]\}\dot{q}_2 = 0 \tag{252}$$

But the original (unperturbed) motion was directed along the meridians, i.e., $\dot{q}_2 \equiv 0$. Consequently,

$$\ddot{q}_1 = 0, \qquad \dot{q}_1 = v_0 = \text{const} \tag{253}$$

i.e., the relative motion along the trajectory remains unchanged.

Returning to the original (inertial) system, one obtains the resultant velocity by summing the relative and transport velocities:

$$v_\tau = v_0 \tag{254}$$

$$v_\epsilon = -mv_0^2 G_0 \epsilon_0 \cos \omega t \qquad (\omega \to \infty) \tag{255}$$

in which v_τ and v_ϵ are the velocity components parallel and normal to the undisturbed (geodesic) trajectory, respectively.

The equations of the disturbed motion in the original frame of reference are

$$\sigma = v_0 t \tag{256}$$

$$\epsilon = \epsilon_0 + \left(\frac{1}{\omega} mv_0^2 G_0 \epsilon_0 \sin \omega t\right) \qquad (\omega \to \infty) \tag{257}$$

in which σ is the coordinate along the undisturbed (geodesic) trajectory.

It follows from equations (254)–(257) that the motion in the original frame of reference is stable in the sense that the current deviations of displacements and velocities do not exceed their initial values. However, the displacement–time function (257) is not differentiable, because its derivative (255) is multivalued. Indeed, for any arbitrarily small interval Δt there always exists a large frequency $\omega > \Delta t/2\pi$ such that within this interval the velocity (255) runs through all its values. In other words, one arrives at stability in the class of nondifferentiable functions. (The mathematical meaning of this result will be discussed below.)

Thus, chaotic motion of a particle on a smooth pseudosphere is represented by the "mean" motion (256) along the undisturbed geodesic trajectory [with the constant velocity (244)] and the fluctuation motion (257) normal to this trajectory. The "amplitude" of these fluctuations is vanishingly small, but the velocity "amplitude" is finite. It is worth emphasizing that this amplitude is proportional to the Gaussian curvature of the surface S, i.e., to the degree of the orbital instability. Therefore, it can be associated with the measure of the uncertainty in the description of the motion.

It is worth mentioning that both "mean" and "fluctuation" components representing the originally chaotic motion are stable. That is why they are not sensitive to initial uncertainties and are fully reproducible. In other words, such a representation of the originally chaotic motion is deterministic.

One should notice that the condition $\omega \to \infty$ is a mathematical idealization. Practically, ω is finite

$$\omega \gg 1/T \tag{258}$$

where T is a time scale over which changes of the parameters of the motion are negligible. The concepts of differentiability and multivaluedness have to be understood in the same sense. Indeed, the multivaluedness of the functions (261) and (262) means that the time interval between two different values of these functions is smaller than the scale of observation T of the examined motion, and therefore these values can be associated with "almost" the same argument.

As discussed above, the concept of stability is related to a certain class of functions or a type of space: the same solution can be stable in one space and unstable in another, depending on the definition of the "distance" between two solutions. Indeed, if the distance between the solutions in (263) is defined as

$$\rho = \sum_{k=0}^{n} \max\left|\epsilon_2^{(k)}(t) - \epsilon_1^{(k)}(t)\right| \tag{259}$$

then the solution (257) is stable for $n = 0, 1$, but it is unstable for $n = 2, 3, \ldots$, since its derivatives $\epsilon^{(2)}, \epsilon^{(3)}, \ldots$, are unbounded. In other words, the concepts of stability as well as chaos are attributes of a mathematical model rather than of a physical phenomenon.

Hence, from a formal mathematical point of view, the occurrence of chaos in the description of mechanical motions means only that these motions cannot be properly described by smooth functions if the scale of observation is finite.

One can notice that the application of the stabilization principle to the representation of chaotic motions in Lagrangian dynamics can be linked to a control problem. Indeed, we are introducing additional rapidly fluctuating forces (coming from noninertial motions of the frame of reference) which are coupled with the parameters of motion in such a way that the original instability is eliminated.

In the particular case of an inertial motion of a particle M on a pseudosphere, the rate of divergence of the trajectories was constant [see equation (15)], which means that local and global Lyapunov exponents are the same. That is why by eliminating the positive local Lyapunov exponent we "automatically" eliminate the global one. In the general case, the situation is more complex: the local Lyapunov exponents depend upon the position of the system, and by eliminating all the local positive Lyapunov exponents, one overstabilizes the motion. Indeed, nonpositive global Lyapunov exponents can exist even if the local ones are positive in some domain of space where the motion can occur. As we will see later, the elimination of global Lyapunov experiments is a much harder problem, and that is why in many practical situations we confine ourselves to the easier problem of elimination of local exponents, i.e., with the overstabilized representations.

3.4.2. Potential Motions

Based upon equation (235), for potential motions, the governing equations can be written in the following form:

$$\ddot{q}^{\alpha} + \Gamma^{\alpha}_{\beta\delta}\dot{q}^{\beta}\dot{q}^{\delta} = -\frac{\partial \Pi}{\partial q^{\alpha}} + a^{\alpha}_{(i)} \tag{260}$$

$$\frac{\partial \Pi}{\partial q^{\alpha}} = -Q^{\alpha} \tag{261}$$

where Π is the potential energy of the dynamical system and $Q^{\alpha}_{(i)}$ are the inertia forces [or the "Reynolds stresses" caused by the rapidly oscillating transport motion of the frame of reference; see equation (236)].

For simplicity, we will confine ourselves to a two-dimensional dynamical system, assuming that $\alpha = 1, 2$.

Following the same strategy as those applied to inertial motions, let us couple the inertia forces with the parameters of the dynamical system in such a way that the original orbital instability (if it occurs) is eliminated. For that purpose, first we will represent these forces in the form

$$a^{\alpha}_{(i)} = -\frac{\partial \Pi_{(i)}}{\partial q^{\alpha}} \tag{262}$$

where $\Pi_{(i)}$ is a fictitious potential energy equivalent to the kinetic energy of the fluctuations. Then, turning to the criteria of local orbital stability (38), one finds this potential energy $\Pi_{(i)}$ and consequently the inertia forces $a^{\alpha}_{(i)}$

from the condition that original local orbital instability is eliminated:

$$G+3\left[\frac{\nabla(\Pi+\Pi_{(i)}\cdot\mathbf{n}}{2W}\right]^2 + \frac{1}{2W}\left[\frac{\partial^2(\Pi+\Pi_{(i)})}{\partial q^i\,\partial q^j} - \Gamma^k_{ij}\frac{\partial(\Pi+\Pi_{(i)})}{\partial q^k}\right]n^i n^j = 0, \qquad i,j=1,2 \tag{263}$$

Here W, G, and Γ^k_{ij} are defined by the parameters of the dynamical system (260) via equations (27), (29), and (30), respectively, and n_i are the contravariant components of the unit normal n to the trajectory of the basic function.

Equation (263) contains only one unknown $\Pi_{(i)}$, which can be found from it, and will define the inertia forces, or the "Reynolds stresses" (262).

It should be noticed that unlike the case of the inertial motion of a particle on a pseudosphere, here the Gaussian curvature G as well as the gradients of the potential energy Π are not constants, and consequently the local Lyapunov exponents may be different from the global ones. This means that the condition (263) eliminates local positive exponents, and therefore the solution to equations (260) and (263) represents an overstabilized motion. Obviously, elimination of only global positive Lyapunov exponents would lead to solutions with less uncertainties, while some of local exponents in certain domains of the phase space may even remain positive. However, the strategy for elimination of global positive exponents is more sophisticated, and it can be implemented only numerically.

It is worth noting that equation (269) is simplified to

$$G+\frac{1}{2W}\left[\frac{\partial^2(\Pi+\Pi_{(i)})}{\partial q^i\,\partial q^j}\right]n^i n^j = 0, \tag{264}$$

if the basic motion is characterized by zero potential forces

$$\frac{\partial\Pi}{\partial q^i}=0 \tag{265}$$

This may occur, for instance, when the dynamical system is in a relative equilibrium with respect to a moving frame.

Examples of the application of the stabilization principle to elastic systems and to ideal fluids are given by Zak (1987), and Zak (1986*a*), respectively.

3.4.3. General Case

When motions of a dynamical system are not potential, in many cases it is more convenient to represent equation (235) in the form of a system of first-order differential equations. For simplicity, we confine ourselves to dynamical systems which can be represented in the following form:

$$\dot{x}^i = a^i_j x^j + b^i_{jm} x^j x^m, \qquad i = 1, 2, \ldots, n \tag{266}$$

Applying the transformation (221) to the variables x^i, one arrives at the following Reynolds-type equation [which is equivalent to equation (235)]:

$$\dot{\bar{x}}^i = a^i_j \bar{x}^j + b^i_{jm} \bar{x}^j \bar{x}^m + b^i_{jm} \overline{x^j x^m}, \qquad i = 1, 2, \ldots, n \tag{267}$$

with the additional terms $b^i_{jm} \overline{x^j x^m}$ representing the Reynolds stresses.

a. The Closure Problem. As in the previous cases, because of additional unknowns $\overline{x^j x^m}$ in (267), the closure problem arises. Analogously, we will seek additional coupling between the mean motion and the fluctuations:

$$\overline{\dot{x^j x^m}} = a^{jm}_l \bar{x}^l \bar{x}^l + a^{jm}_{ln} \bar{x}^l \bar{x}^n \cdots \tag{268}$$

based upon the stabilization principle, the application of which will be clarified below.

First, we recall that the solution to equations (266) are chaotic, and consequently, some of the Lyapunov exponents of equation (266) are positive:

$$\lambda^+_m > 0, \qquad m = 1, 2, \ldots, S \tag{269}$$

Second, we are looking for a decomposition in which the mean motion is periodic, rather than chaotic. Hence, the fluctuations should be coupled with the mean motion such that all positive Lyapunov exponents become zero, while the rest of the exponents are unchanged. Indeed, in this case the mean motion is a regular motion which is the "closest" to the original chaotic motion. Since the Lyapunov exponents for the system (267), (268) depend on the "feedback" coefficients a^{jm}_l, a^{jm}_{ln}, etc., the closure can now be formulated as follows:

$$\begin{aligned}
&\lambda^+_i (a^{jm}_l, a^{jm}_{ln}, \ldots) = 0, \qquad i = 1, 2, \ldots, S_+ \\
&\lambda^0_i (a^{jm}_l, a^{jm}_{ln}, \ldots) = \lambda^0_i (0, 0, \ldots), \qquad i = 1, 2, \ldots, S_0 \\
&\lambda^-_i (a^{jm}_l, a^{jm}_{ln}, \ldots) = \lambda^-_i (0, 0, \ldots), \qquad i = 1, 2, \ldots, S_-
\end{aligned} \tag{270}$$

in which λ^+, λ^0, and λ^- are positive, zero, and negative Lyapunov exponents, respectively. Obviously, those coefficients a^{jm}_l which do not appear in (270) must be zero.

Thus, the system (267), (268), (270) is closed. It defines the regular mean motion and fluctuations which represent the original chaotic motion. Since all the Lyapunov exponents for this system are not positive, the solution is stable and predictable in the sense that small changes in the initial conditions cause small changes in both the mean motion and the fluctuations.

In the next subsection the application of this approach to the Lorenz strange attractor is illustrated.

b. Higher-Order Approximations. The Reynolds decomposition of the variables x^i in equation (266) generates not only pair correlations $\overline{x^i x^i}$, but also correlations of higher orders, such as triple correlations $\overline{x^i x^j x^k}$, quadruple correlations $\overline{x^i x^j x^k x^m}$, etc.

Indeed, multiplying equations (266) by x^k and averaging and combining the results, one obtains the governing equation for the pair correlations $\overline{x^i x^k}$:

$$\dot{\overline{x^i x^k}} = a^i_j \overline{x^j x^k} + a^k_j x\overline{x^j x^i} + b^i_{jm}(\overline{x^k x^j x^m} + \overline{x^k x^j x^m} + \overline{x^k x^m x^j} + \overline{x^j x^m x^k})$$
$$+ b^i_{jm}(\overline{x^i x^j x^m} + \overline{x^i x^j x^m} + \overline{x^i x^m x^j} + \overline{x^j x^m x^i}) \tag{271}$$

which contains nine additional triple correlations $\overline{x^i x^j x^k}$.

Similar equations for the triple correlations will contain all the quadruple correlations, etc. In general, one arrives at an infinite hierarchy of equations which are open, since any first N equations relate $(N + 1)$ correlations.

From this viewpoint all the closures discussed above can be considered as first-order approximations which define only the mean components of the chaotic motions. In order to define both the mean motion and the double correlations, one should consider the Reynolds equation (267) together with equation (271). In this case the evolution of the double correlations is already prescribed by equation (271), and consequently the stabilizing feedback must now couple the triple correlations with the mean and pair correlation components:

$$\overline{x^i x^j x^m} = F(\overline{x^k}, \overline{x^l x^m}) \tag{272}$$

The system (267), (271), (272) will define periodic mean and pair correlation components. It is possible that the mean components may be different from those found before (in the same way in which the second-order approximation may be different from the first-order one).

The higher-order approximations can be introduced using the same procedure.

c. Computational Strategy. As follows from the above, the closure, i.e., the stabilizing feedback between the Reynolds stresses and the mean components of the motion, can be written in explicit form only if the criteria for the onset of chaos are formulated explicitly. Since such a situation is the exception rather than the rule, we develop below a computational strategy which allows one to find the closure regardless of the complexity of the original equations.

We will demonstrate this strategy using equation (266). The same strategy will be suitable for the Navier–Stokes equations, since after an appropriate discretization technique they reduce to the form (266).

Turning to equation (267), which follow from equation (266) as a result of the Reynolds decomposition, let us linearize it with respect to the original ("laminar") state of x_0^i:

$$\dot{\bar{x}} = (a_j^i + 2b_{jm}^i \bar{x}_0^m)\bar{x}^j, \qquad \text{with} \quad \overline{x^i x^j} = 0 \quad \text{at} \quad \bar{x}^i = \bar{x}_0^i \tag{273}$$

Introducing small "laminar" disturbances in the form

$$\bar{x}^i = \bar{x}_*^i \exp(\lambda_0 t) \tag{274}$$

one arrives at a truncated analog of the Orr–Sommerfeld equations:

$$\lambda^0 \delta_j^i = (a_j^i + 2b_{jm}^i \bar{x}_0^m)\bar{x}^j \tag{275}$$

where the local eigenvalues of equation (267)

$$\lambda^0 = \lambda_1^0, \lambda_2^0, \ldots, \lambda_n^0 \tag{276}$$

are the roots of the characteristic equation

$$\det(\lambda^0 \delta_j^i - a_j^i - 2b_{jm}^i \bar{x}_0^m) = 0 \tag{277}$$

Applying the same procedure to the second-order Reynolds equation (271), one obtains instead of equation (272),

$$\frac{1}{2}\lambda^0 \delta_k^i = (a_j^i + 2b_{jm}^i \bar{x}_0^m)\overline{x^j x^k} \tag{278}$$

and therefore the local eigenvalues of equation (271) are twice as large as those for equation (267), i.e., instead of equation (274)

$$\overline{x^i x^k} = \sim\exp(2\lambda_0 t) \tag{279}$$

If the original "laminar" state $\bar{x}_0^i$ is unstable, i.e., there are λ_i^0 with positive real parts in equation (276),

$$\operatorname{Re} \lambda_i^0 > 0 \tag{280}$$

then the pair correlations (279) will grow much faster than the mean motion disturbances (277), and one can assume that these correlations will be large enough to stabilize equation (267) while the mean motion will remain sufficiently close to its original state $\bar{x}_0^i$. This property makes possible the following computational strategy.

Let us seek a closure to equation (267) in the neighborhood of the original laminar state $\bar{x}_0^i$ in the form

$$b_m^i \overline{x^j x^m} = C_m^i \bar{x}^j \tag{281}$$

Substituting equation (281) into equation (267) and linearizing with respect to the original "laminar" state $\bar{x}_0^i$, one obtains

$$\dot{\bar{x}}^i = (a_j^i + 2b_{jm}^i \bar{x}_0^m + C_j^i)\bar{x}^j \tag{282}$$

while the eigenvalues for this equation follow from

$$\det(\lambda\delta_j^i - a_j^i - 2b_{jm}^i \bar{x}_0^m - C_j^i) = 0 \tag{283}$$

The sought coefficients C_j^i must be selected such that

$$\operatorname{Re}\lambda_i = \frac{1}{2}(\operatorname{Re}\overset{0}{\lambda}_i - |\operatorname{Re}\overset{0}{\lambda}_i|) \tag{284}$$

Indeed, in this case all the positive real parts of the local eigenvalues causing the instability of the "laminar flow" become zero, while the rest of these eigenvalues remain unchanged.

In order to find C_j^i from condition (284), we diagonalize the matrix

$$a_j^i + 2b_{jm}^i \bar{x}_0^m = \{F_{ij}\} \tag{285}$$

such that

$$\Theta^{-1}F\Theta = [\lambda_i, \ldots, \lambda_n] \tag{286}$$

then the matrix of the sought coefficients

$$C_j^i = \{C_{ij}\} \tag{287}$$

is found to be

$$\{C_{ij}\} = \frac{1}{2}\Theta[C_1, C_2, \ldots, C_n]\Theta^{-1} \tag{288}$$

in which

$$C_i = -\frac{1}{2}(\operatorname{Re}\lambda_i + |\operatorname{Re}\lambda_i|) \tag{289}$$

Substituting equation (248) into equation (282), one obtains a linearized governing equation for the turbulent or chaotic motion at the very beginning of the transition from the laminar motion. Selecting a small time step Δt_1, one can find the next state $\bar{x}_1^i$:

$$\bar{x}_1^i = \bar{x}_0 + \dot{\bar{x}}_0 \Delta t_1 \tag{290}$$

Repeating this procedure for x_1^i, Δt_2, x_2^i, Δt, etc., one arrives at the evolution of the turbulence, or chaos. The process ends when the solution approaches a regular (static or periodic) attractor whose existence is assumed.

A numerical implementation of this strategy can be based upon a direct suppression of the exponential growth of errors in initial conditions by means of an appropriate selection of the Reynolds stresses. As an example of application of such a strategy, we will illustrate the prediction of the probabilistic structure of the Lorenz attractor by using the stabilization principle.

Applying the Reynolds transformation to the Lorenz attractor

$$\begin{aligned} \dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz \end{aligned} \tag{291}$$

one obtains

$$\begin{aligned} \dot{\bar{x}} &= -\sigma\bar{x} + \sigma\bar{y} \\ \dot{\bar{y}} &= r\bar{x} - \bar{y} - \bar{x}\,\bar{z} - \overline{xz} \\ \dot{\bar{z}} &= -b\bar{z} + \bar{x}\,\bar{y} + \overline{xy} \end{aligned} \tag{292}$$

where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the mean values of x, y, and z, while $\overline{xz}$ and $\overline{xy}$ are double correlations representing the Reynolds "stresses."

As extra variables, these double correlations must be found from the condition that they suppress the positive Lyapunov exponent to zero. In this case, both the mean and the double correlation components of the motion will be represented by periodic attractors, i.e., in a fully deterministic way.

Numerical implementation of this strategy performed for $\sigma = 10$, $r = 28$, and $b = 8/3$ leads to the following results. Figure 10 represents the original chaotic attractor as a solution to equation (291). In Fig. 11, this attractor is decomposed into two deterministic (periodic) motions: the mean motion (Fig. 11a) and the double correlations, i.e., the Reynolds stresses (Figs. 11 and 11c). In order to find all the double correlations, one should exploit the system for triple correlations, which can be obtained in a straightforward way from equation (291). In this system all the triple correlations, as extra variables, must be found from the stabilization principle in a similar way. By continuing this process, one can find the probabilistic structure of the solution to the Lorenz equations (291) to required accuracy.

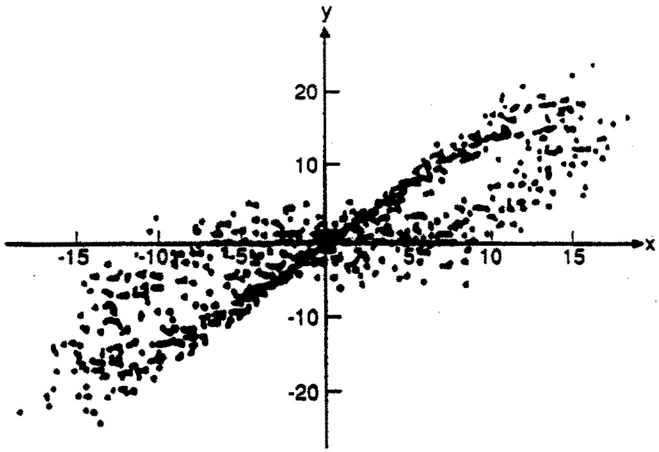

Fig. 10. Plot of x vs. y for one million points sampled at 1000 points.

It should be stressed that the solution to equations (292) plotted in Fig. 11 is stable (in the new class of functions which includes "multivalued" fluctuations): small changes in initial conditions will lead to small changes in the solution.

One should recall that although equations (292) are different from the original Lorenz equations (291), they describe the same physical phenomenon in a specially selected rapidly oscillating frame of reference.

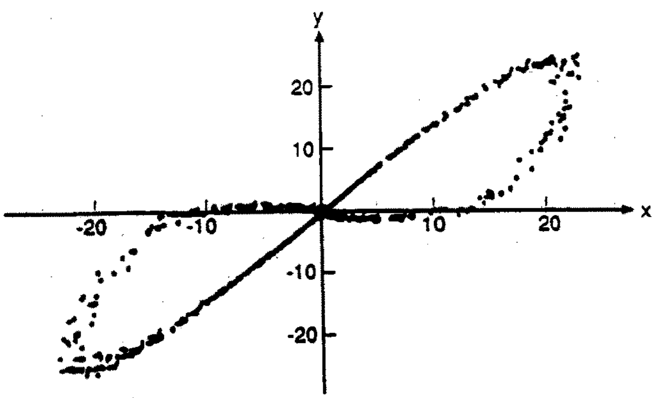

Fig. 11(a).

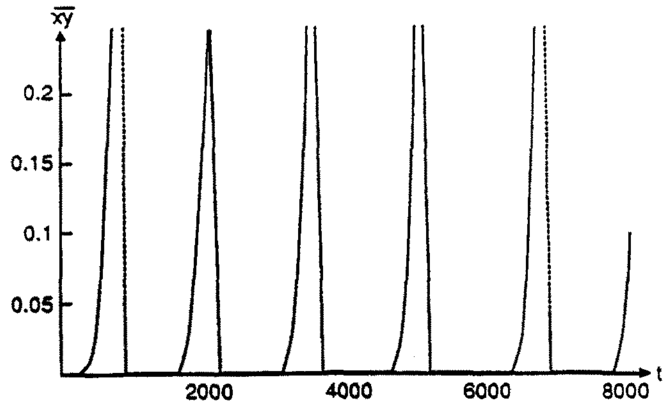

Fig. 11(b).

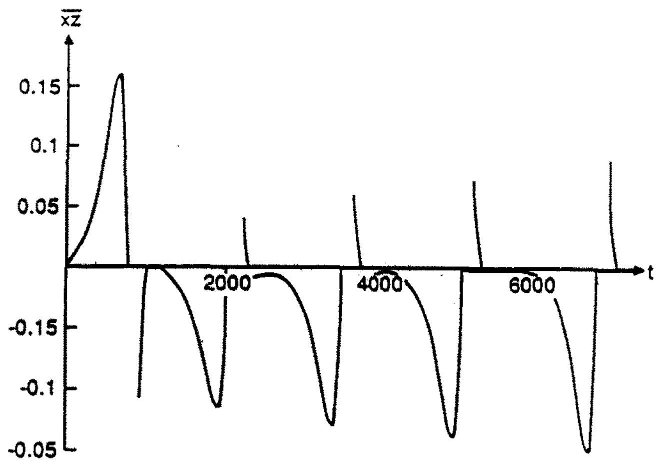

Fig. 11(c)

Fig. 11. (a) Plot of x vs. y for one million points sampled at 1000 points. (b) $\overline{xy}$: Double correlations. Plot over time (8000 points). (c) $\overline{xz}$: Double correlations. Plot over time (8000 points).

ACKNOWLEDGMENTS

This research was carried out at the Center for Space Microelectronic Technology, Jet Propulsion Laboratory, California Institute of Technology. Support for the work came from agencies of the U.S. Department of Defense, including the Innovative Science and Technology Office of the

Strategic Defense Initiative organization, through an agreement with the National Aeronautics and Space Administration. The effort was also supported by Dr. R. E. Meyers, U.S. Army Atmospheric Laboratory.

REFERENCES

Arnold, V. (1988). *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, p. 331.

Da Costa, N., and Doria, F. (1991). *International Journal of Theoretical Physics*, **30**(8), 1041–1073.

Fluge, N. (1962). *Handbook on Engineering Mechanics*, McGraw-Hill, New York, p. 33-9.

Gantmacher, F. (1970). *Lectures in Analytical Mechanics*, MIR, Moscow, pp. 80, 81.

Gödel, K. (1931). *Monatshefte fur Mathematik und Physik*, **38**, 173.

Lessen, M., and Poillet, (1976). *Physics of Fluids*, **19**(7).

Reynolds, O. (1895). *Philosophical Transactions of the Royal Society*, **1895**, 186.

Richardson, D. (1968). *Journal of Symbolic Logics*, **33**, 514.

Synge, J. L. (1926). *Philosophical Transactions of the Royal Society of London A*, **226**, 31–106.

Zak, M. (1968). *Solid Mechanics Journal* (Moscow), **3**(3), 68–53 [in Russian] (English transl., pp. 32–35).

Zak, M. (1970). *Prikladnaya Matematika i Mekhanika*, **39**, 1048–1052.

Zak, M. (1982*a*). *Acta Mechanica*, **43**, 97–117.

Zak, M. (1982*b*). *Solid Mechanics Archives*, **7**, 467–503.

Zak, M. (1982*c*). *Journal of Elasticity*, **12**(2), 219–229.

Zak, M. (1983). *ASME, Journal of Applied Mechanics*, **50**, 227–228.

Zak, M. (1985*a*). *Physics Letters A*, **107A**(3), 125–128.

Zak, M. (1985*b*). *International Journal of Nonlinear Mechanics*, **20**(4), 297–308.

Zak, M. (1986*a*). *Physics*, **18D**, 486–487.

Zak, M. (1986*b*). *Physics Letters A*, **118**(3), 139–143.

Zak, M. (1987). *International Journal of Math Modelling*, **9**(8), 599–612.

Zak, M. (1988). *Physics Letters A*, **133**(1,2), 18–22.

Zak, M. (1990). *Mathematical and Computer Modelling*, **13**(1), 33–37.

Zak, M. (1993). *International Journal of Theoretical Physics*, **32**(1), 159–189.