



Geometry, Stability and Symmetry

A. B. GIVENTAL*

*Institute of Chemical Physics,
Soviet Academy of Sciences, Moscow, U.S.S.R.*

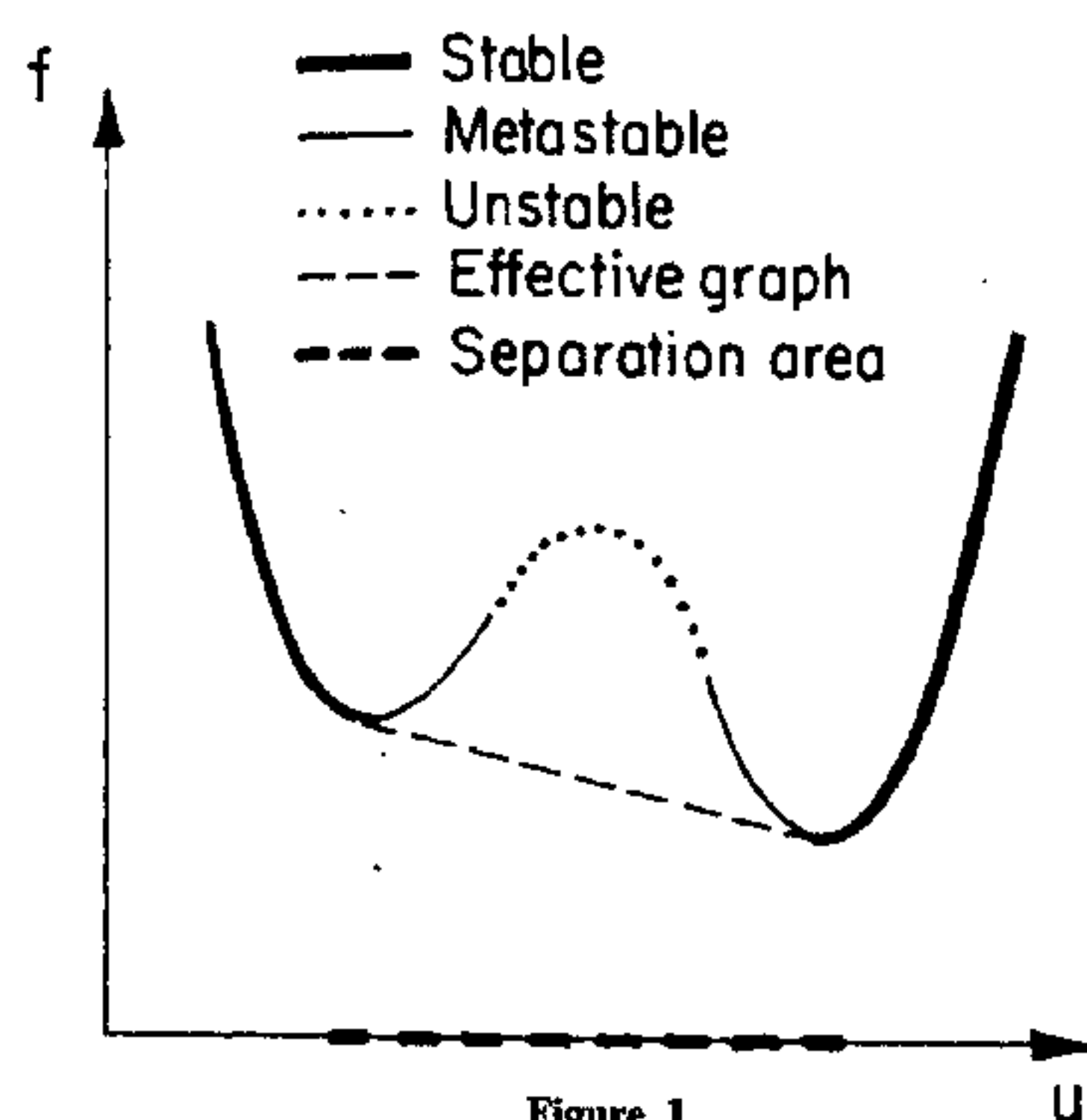
The purpose of these notes is to illuminate some problems and results on stability in thermodynamics, kinetics and dynamics and to introduce the readers to corresponding mathematical tools. Physical phenomena presented here are restricted by the author's interests and knowledge: phase separation, Rayleigh-Bénard convection, non-linear elasticity, short-wave optics and diffraction. Dissipative structures, intermediate asymptotics, solitons and catastrophes appear there as convenient theoretical models. Geometry brings clarity to general considerations.

We suggest that the reader follows further the validity (or invalidity?) of a few metascientific conceptions:

- (i) Stability : everything that is observed in natural or numerical experiments is stable.
- (ii) "Feynman principle" : the same equations have the same solutions.
- (iii) Symmetry implicit and explicit – does everything hidden become evident?
- (iv) Unity of nature – is this its inner property or that of our viewpoint?

THERMODYNAMICS OF PHASE SEPARATION

The phase separation theory deals with an energy function of concentrations $u = (u_1, \dots, u_n)$ defined for pure states of the mixture. Pure states may coexist in mixed ones. Then the average



* Dr. Givental, was unfortunately not able to participate in the symposium at Bangalore in December 1988. Nevertheless, he sent in his paper which is reproduced here – the editors are deeply indebted to him.

energy and average concentrations of components depend on those parts linearly. The mixture tends to the minimum of the energy. This rule leads to the (meta)stability criterion (figure 1). A pure state is metastable (= stable with respect to a small separation) if and only if the graph of f is convex up at the point. An effective average energy function graph is the convex hull of the initial one. Pure states where effective and initial graphs coincide are stable (with respect to any separation).

The next question is how the phase separation happens in space and time. It leads us to diffusion kinetics.

DIFFUSION KINETICS

One may use the Van-der-Waals' gradient approach and introduce the following energy functional

$$\mathcal{F}[u] = \int_{\text{space}} [f(u) + u_x^2/2] dx. \tag{1}$$

Now u is a concentration field varying in the space. Time evolution of the field satisfies a kinetic equation called in the context the Cahn-Hilliard equation

$$\dot{u} = (\delta\mathcal{F} / \delta u)_{,xx} = -u_{xxxx} + (\text{lower terms}). \tag{2}$$

The evolution under the equation is purely dissipative: average concentrations are preserved and energy \mathcal{F} never increases. Therefore time-periodical solutions are impossible. But space-periodical stationary solutions are possible and do exist. For a binary mixture and one-dimensional space (u and x – scalars) they all correspond to segments shown on figure 2a and depend on two continuous parameters – spatial period T and average concentration U . For fixed T , U there is only a finite number of stationary states which differ by a discrete parameter ν – the number of their semiperiods on the interval T (figures 2b – d).

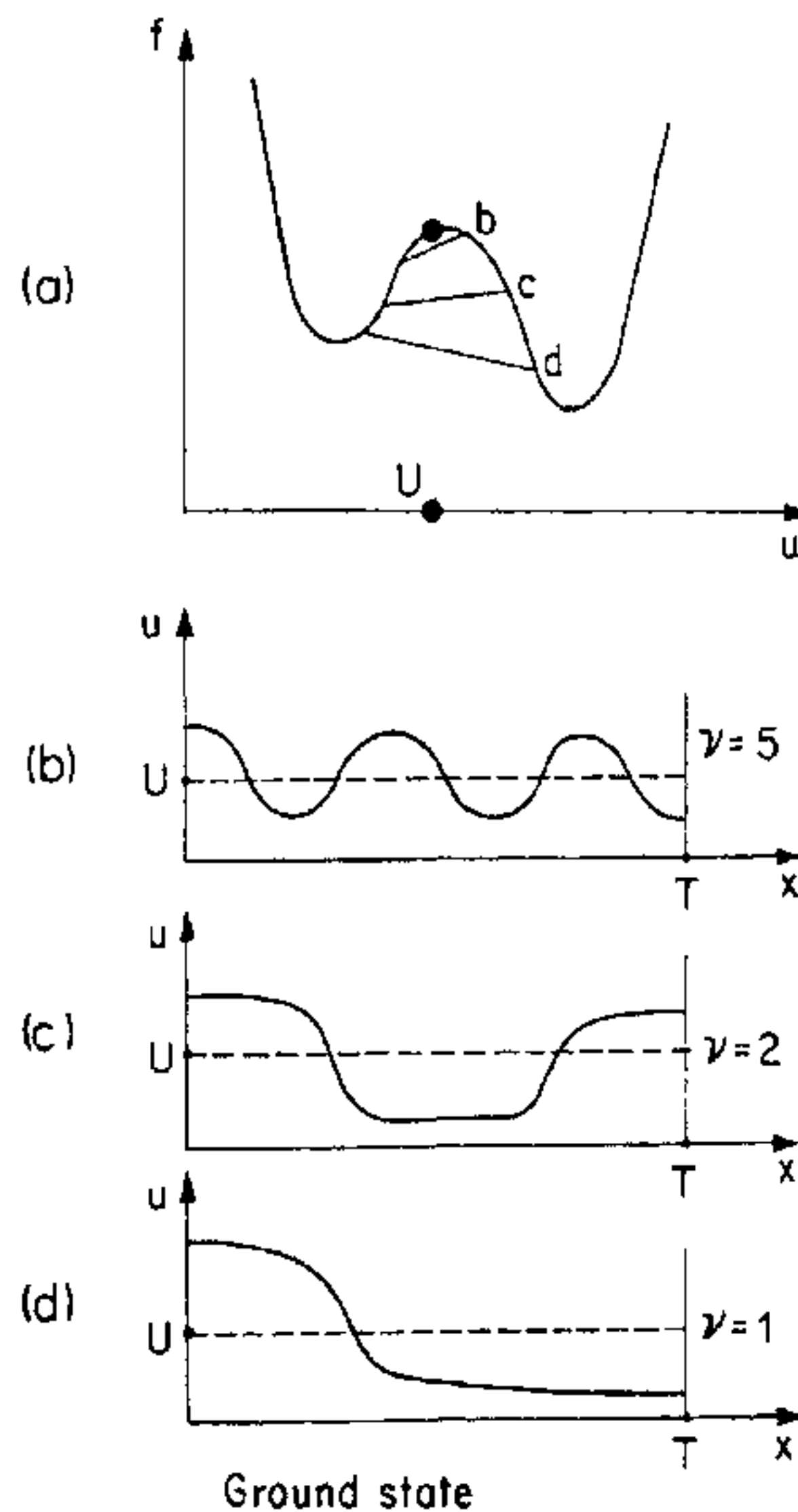


Figure 2.

It is intuitively clear that a separation decay will tend to the ground state with $\nu = 1$ (figure 2d). But a numerical experiment (Mitlin-Manevich-Erukhimovich) shows that the decay of unstable stationary state $u(x) \equiv U$ leads to a periodical stationary regime (as in figure 2b for example) living unlimitedly long. In contradiction to this, one can prove:

Theorem. States with $\nu > 1$ are unstable, they are saddle points of \mathcal{F} in the functional space of concentration fields with fixed T and U .

Moreover a similar result is valid for periodical stationary states of multicomponent mixtures. Moreover similar results can be found for the space of dimension more than one. These results lead us to the instability principle: stable phase separation states must have the simplest structure.

To bring sense to this statement we note that the structure of the whole is determined by the structure of its parts. So it would be better to begin with the parts and then reconstruct the whole. We will give now two statements built by this rule. The first of them is a conjecture, the second a theorem.

(i) Let a stationary state in the area with a flat boundary be given (figure 3a). We may extend it onto an area double the size by reflection in the flat boundary. Then this stationary state in the whole area is unstable.

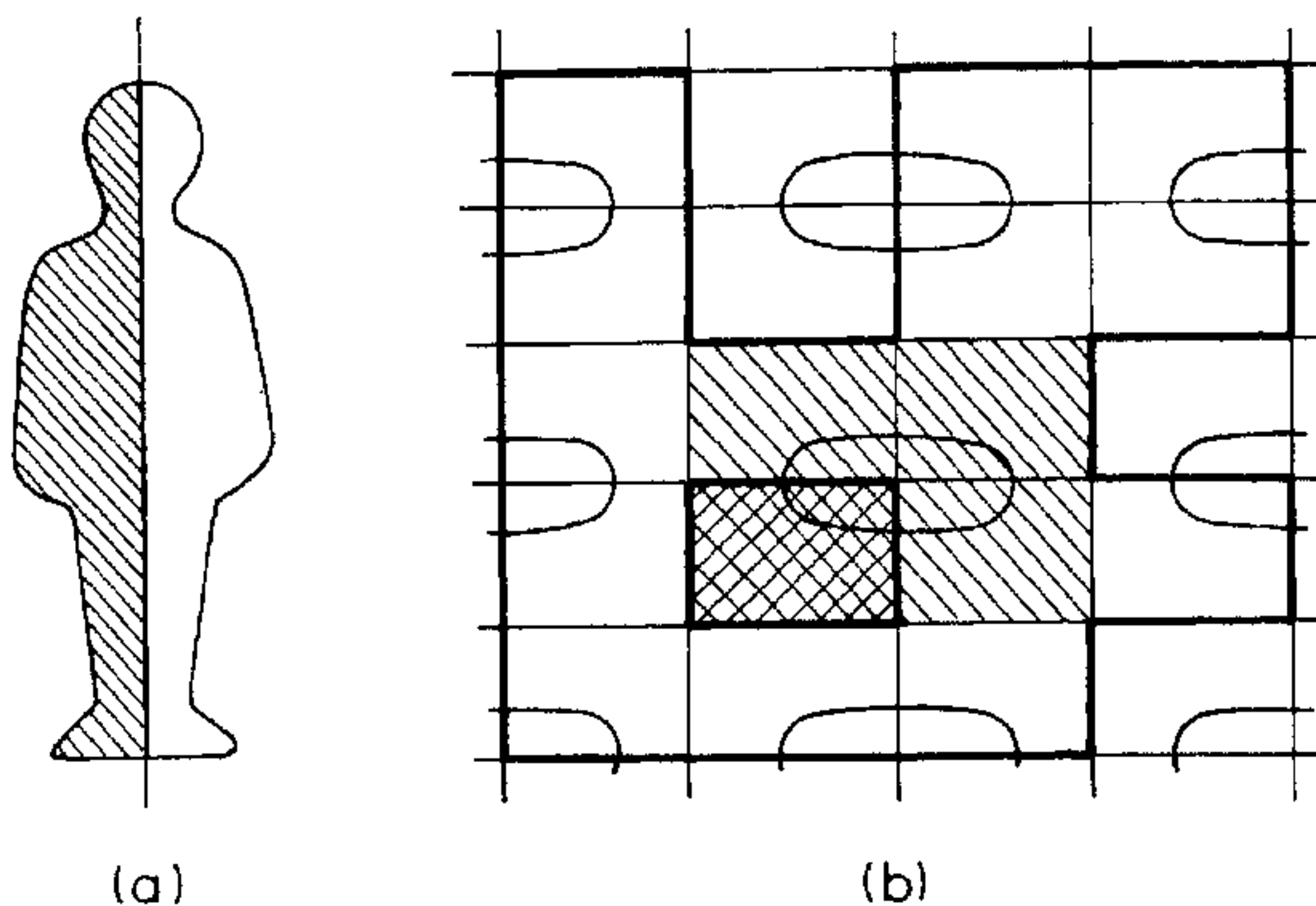


Figure 3.

(ii) Let a stationary state in the rectangle be given (figure 3b). We may extend it by reflection onto nearest rectangles of the lattice and then periodically on the whole space. We get a periodical symmetric stationary state in any area consisting of rectangles. If the number of rectangles is more than one then this state is unstable.

Returning to binary mixtures in one-dimensional space we describe now stability criterion for states with $\nu = 1$. Calculating energy for such states we get a function $F(T, U)$.

Theorem. The state is stable if and only if F is convex up at the point corresponding to the state.

This result returns us to thermodynamics but with a new “extensive” parameter – the size of the space area.

INTERMEDIATE ASYMPTOTICS

Appearances of unstable states in numerical experiments make us think about these states as intermediate asymptotics and to describe the process as wandering to the ground state via unstable ones. Further numerical experiments (Mitlin-Manevich) – on significantly greater time scales than before – justify this program.

To realize it we need to know the lifetime of unstable states. It is an open question on instability increments. An exact formulation of the problem concerns to do with the family of stationary states

in the infinite one-dimensional space with fixed average concentration U and growing spatial period T . The question is: what is the asymptotic behaviour of the largest eigenvalue of the linearized equation (2) when $T \rightarrow \infty$ and what does the eigenfunction look like? The answer is found only for stationary states that are close to the constant U not just in average but pointwise (Dobrynin-Givental'). This problem is nontrivial only for U close to the inflection point of energy function graph (figure 1) – the boundary between meta- and instability. Otherwise the relaxation time tends to that of the constant unstable solution $u(x) \equiv U$. For the nontrivial case the relaxation time tends to infinity as $(\text{const}) T^4$. The coefficient can be found explicitly. The eigenfunction has a period $2T$. This means that decay of the stationary states in question will pass preferably in the direction of the period's doubling.

These results are based on a surprising mathematical relation between phase separation and quantum scattering theory. It begins with the stupid question: what are the periodical functions u for which our eigenvalue problem can be reduced to that for Schrödinger's equation with potential u ? The answer is:

- (i) If and only if u is a solution of stationary Cahn-Hilliard equation with polynomial energy function f of degree three!
- (ii) Such solutions are two-zone potentials and spectral problems for them are exactly solvable.

The further idea is to approximate any energy function near inflection point by a cubic polynomial.

The general mathematical causes of these surprising coincidences are hidden inside the theory of Korteweg-de-Vries equation but are not completely clear.

CHEMICAL KINETICS

The instability of periodical solutions of equation (2) is closely related to its purely dissipative character. Being linearized near a pure state $u(x) \equiv U$ equation (2) remains spatially homogeneous. So

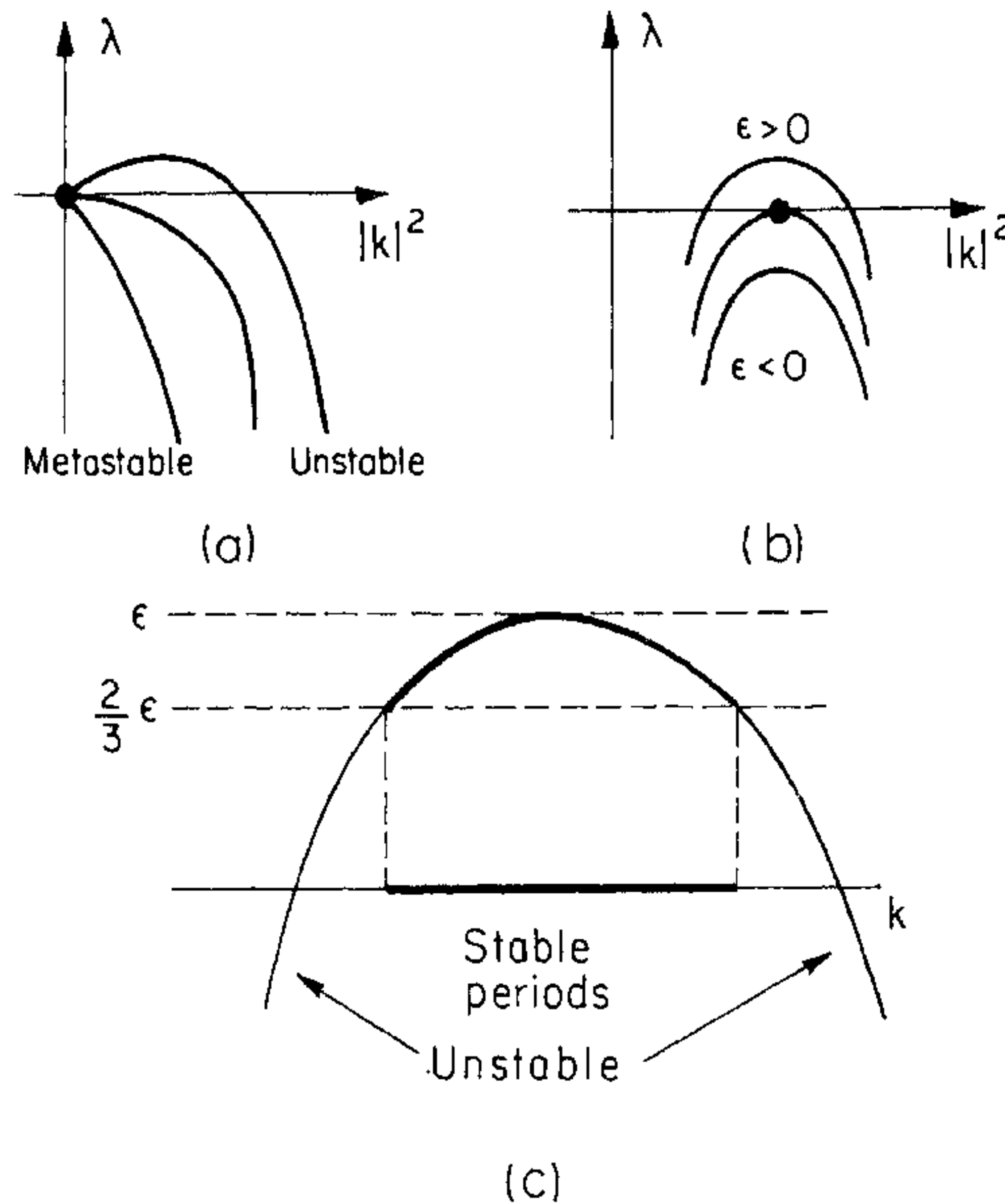


Figure 4.

eigenfunctions of a linearized equation are exponents $\exp(ikx)$. It remains isotropic. So the eigenvalue λ depends on $|k|^2$ only. This dependence is called a dispersion relation and looks like figure 4a. Amplitudes of eigenfunctions vary in time as $\exp[\lambda(k)t]$. Thus pure states lose their stability through long-wave perturbations.

Quite another loss of stability – through middle-wave perturbations (figure 4b) – can take place for equations of “diffusion + reaction” type where average concentrations are not preserved. This actually happens in the model (and really representative) equation

$$\dot{u} = -u_{xxxx} - 2u_{xx} - (1 - \epsilon)u - u^3. \quad (3)$$

It is known (a strict proof seems to be due to Arnol'd) that periodical stationary solutions exist for all “unstable” spatial frequencies ($\epsilon > 0$, $\lambda(k) > 0$). The question about their stability is answered by the Echhaus criterion displayed in figure 4c: stable periodical stationary states do exist and their spatial frequencies are determined asymptotically for $\epsilon \rightarrow 0$ by the universal constant $2/3$.

RAYLEIGH-BÉNARD CONVECTION

A similar situation takes place in Rayleigh-Bénard convection theory. A layer of viscous fluid heated from below becomes unstable for some critical value of the heat flow and convective motions arise. Under small overcriticality the convective flows form stationary periodical structures of regular symmetry: rolls, rhombia or hexagons filling the plane. Mathematically an investigation of the phenomenon can be reduced to that of an equation of type (3) but in two spatial variables. The thermoconductivity and hydrodynamics replace the diffusion and reaction. The unknown u itself has no evident sense but determines all physical observables. This was studied and a two-dimensional analogy of the universal asymptotic Echhaus criterion was found (Malomed-Tribel'sky). To illuminate the nature of the problem and the style of results we shall restrict ourselves to the case of periodical perturbations with fixed period lattice having regular hexagonal symmetry (really the lattice can vary and perturbations can be unperiodic). We suppose additionally the existence of some energy functional \mathcal{F} minimized under kinetics.

The hierarchy of symmetry groups containing the lattice translations and contained in the motion group of the plane is presented in figure 5a. Stationary states are critical points of functional \mathcal{F} . Stable states are local minima of \mathcal{F} . \mathcal{F} is invariant under the whole group of motions of the plane. But its critical points may belong to any symmetry type.

\mathcal{F} may depend on parameters. Two of them – overcriticality ϵ and gravitational asymmetry α – are essential physically. Coexistence of stable states with given symmetry types is shown in figure 5b. Such pictures on parameter space are called bifurcation diagrams. When parameters vary the state of the system changes continuously or by jumps and the resulting state may depend on the initial state and on the direction of the process (hysteresis). How it happens can be understood from figure 5c where the diagram of catastrophes – the dependence of energy in stationary states on parameters – is pictured.

A general method of catastrophe theory is to reduce any object – function, family etc. – to some universally normal form. Where a physicist neglects higher terms, a mathematician “kills” them by transformation of variables. The remaining terms and parameters are essential and form the normal forms. Its analysis leads to pictures similar to figure 5.

Indeed our normal form depends also on some third essential parameter and is unstable to its variations. But this instability is analytical and invariants of the pictures vary continuously. The situation does not change qualitatively until the third parameter value crosses some critical point. After the crossing figure 5b is replaced by 5d. This gives a theoretical prediction: one can observe stable structures with triangular symmetry (Golubitsky 1983, Malomed – Tribel'sky – Givental' 1986).

For Rayleigh-Bénard convection the third parameter value never crosses the critical point, for physical reasons, and the picture 5b works in accordance with experiments. It would be interesting to find out where the triangular structures can appear and to observe them.

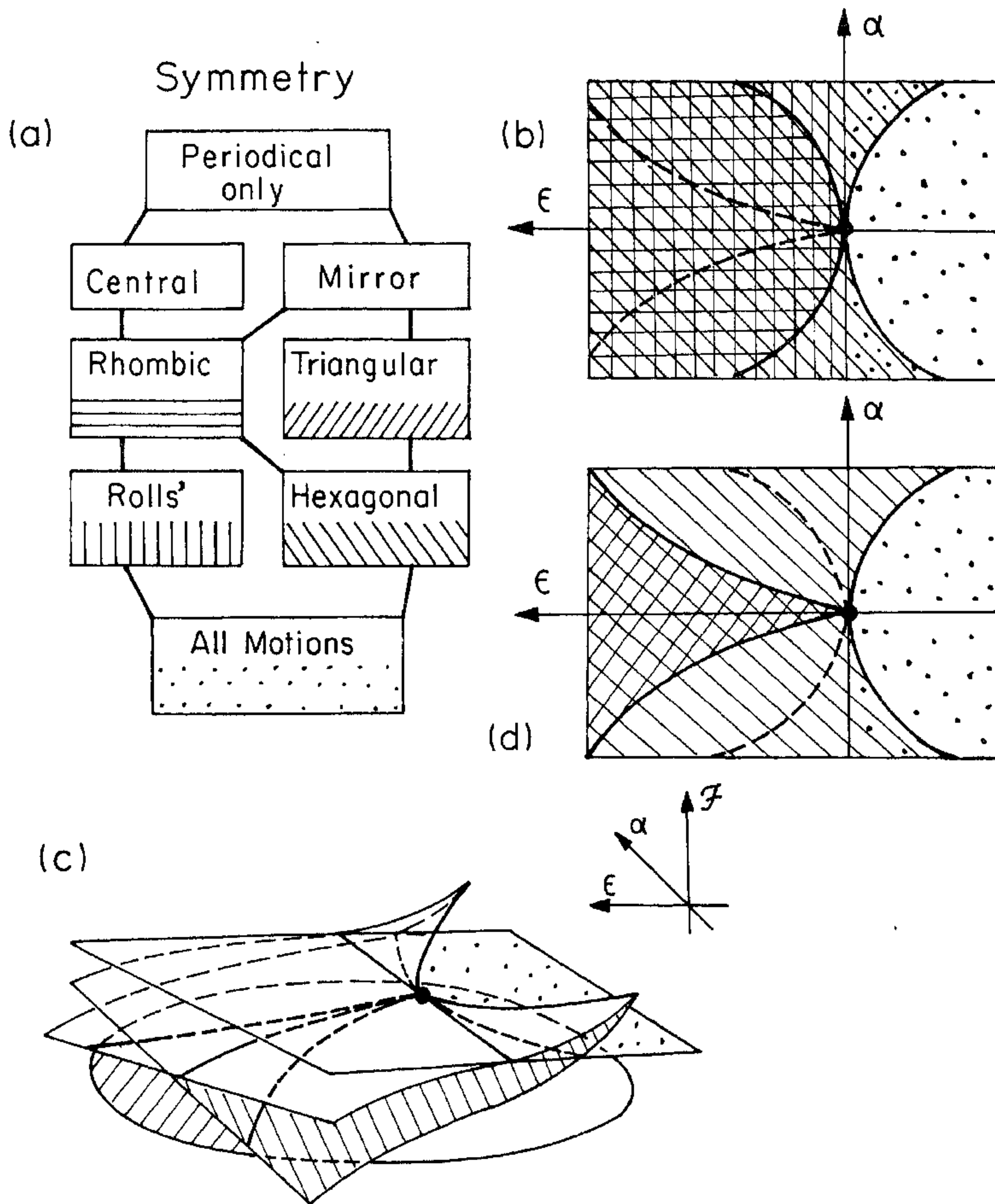


Figure 5.

STABILITY OF SOLITONS

Equation (2) is unstable when an additional time derivative is added to the left hand side: it changes its character from dissipative to a conservative one. The resulting equation describes non-linear elastic waves in one-dimensional media with dispersion. It appears for example under continuous discrete atom chain with pair potential of neighbours f and harmonic long-range interaction. The unknown u is a strain field. Under change $v_x = u$ the equation becomes a Hamiltonian system with kinetic energy $\int (v^2/2)dx$ and potential energy

$$u[v] = \int [v_{xx}^2/2 + f(v_x)] dx. \quad (4)$$

The corresponding dynamical equation takes the form

$$\ddot{v} = - [(v_x)_{xx} - f'(v_x)]_x. \quad (5)$$

Among solutions of (5) there are solitons of strain field. By definition a soliton is a running solitary wave. From the statistical mechanics viewpoint solitons have a negligible value among all solutions: observable wave processes are "rarely solitary". But their existence in the model shows that such processes should exist and were often used in phenomenological theories infinitely. (Even such notions as "a solitary wave packet" appeared to overcome the contradiction with statistical mechanics.)

According to the general stability conception a soliton phenomenological theory is valid only if the soliton solutions used are stable. In our case solitons depend on two parameters – strain value ϵ at infinity and soliton velocity c . Let us define "soliton square"

$$F(c^2) = \int_{-\infty}^{\infty} (u - \epsilon)^2 dx.$$

The soliton is stable if and only if

$$F(c^2) + 2c^2 \frac{d}{dc^2} F(c^2) < 0.$$

In particular standing solitons are unstable. Analogous stability criteria can be found for other classes of dynamical equations containing Korteweg-de-Vries or Sine-Gordon equation with an arbitrary form of non-linearity. But there are open questions. So a stability criterion is unknown for opposite sign before the quadratic term in (4).

GEOMETRICAL OPTICS

The wave-particle duality shows itself in short wave theory as an equivalence of ray and front approaches in geometrical optics. According to the Huygens principle one may take any instant front as the initial one and investigate a light propagation considering parallel fronts or the ray system orthogonal to these fronts.

Let the initial front be a circle on the plane. Then the ray system has focus at the centre. This situation is unstable: if the initial front varies then the focus vanishes. Instead the whole focal locus – caustics – appears (an aberration). It is shown in figure 6a together with ray and front systems.

Geometry of diffraction on an obstacle leads to another example. If rays must not cross some curve – boundary of the obstacle – then they move some time along the boundary and then leave it in a tangent direction. Figure 6b shows rays and fronts near an inflection point on the obstacle boundary. Here the ray distance is a sum of boundary segment length and straight one. Fronts are shown together with their "analytical continuation". It means that time on straight rays runs forward in positive direction from tangency point and backward in negative one.

Both pictures were already contained in the first course on Mathematical Analysis (l'Hospitale \approx 1800) but their nature was understood just recently. They are managed by regular polyhedra – by the tetrahedron (Arnol'd, 1972) and the icosahedron (Scherbak, 1983) respectively.

Consider a surface in 3-space (figure 6c,d) formed by instant fronts lifted to different heights – the graph of time function. The surfaces are the discriminant varieties of the polyhedra symmetry groups.

The symmetry group of the regular polyhedra acts in 3-space and is generated by reflections in its symmetry planes. An orbit space of the action is 3-space again and the discriminant variety by definition is the collection of nonregular orbits formed by the symmetry planes. For example the tetrahedron group acts in 3-space by permutations of its vertices in the same way as permutations of roots of polynomials (6)

$$x^4 + Ax^2 + Bx + C = (x - x_0)(x - x_1)(x - x_2)(x - x_3), \quad x_0 + x_1 + x_2 + x_3 = 0 \quad (6)$$

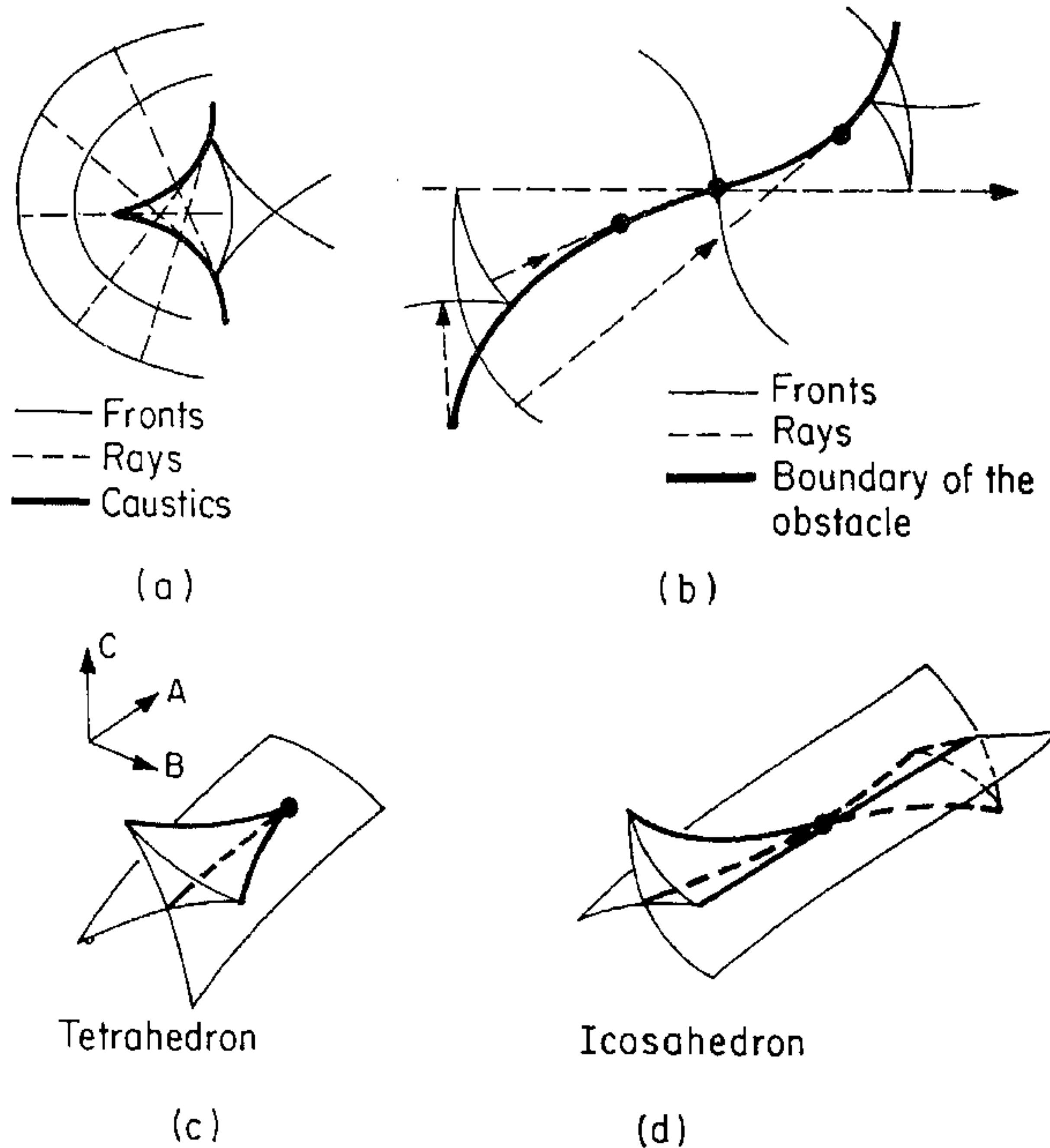


Figure 6.

act on the root space itself. So the tetrahedron discriminant coincides with the space of polynomials (6) having multiple root ($x_i = x_j$ for some i, j) in the (A, B, C) - space of all polynomials.

The discriminants are normal forms of time function graphs only. They coincide with the graphs up to smooth exchange of variables in surrounding space near the origin. So the catastrophe theory method is applied here again.

A similar approach can be used anywhere the rays or fronts propagation meets (see for example Arnol'd - Zel'dovich - Shandarin's works on large-scale structure of the Universe).

CATASTROPHE THEORY AND CONTACT GEOMETRY

It seems the contact geometry first appeared in physical context in Gibbs' "Graphical methods in the thermodynamics of fluids" (1873), (The Scientific Papers, Vol. 1, (1906), pp. 1-32).

We consider the following quantities: v —volume, p —pressure, t —temperature (absolute), ϵ —energy, η —entropy of given liquid in some state, and also W —work produced by liquid under a transition from one state to another and H —heat received by liquid under this. These quantities satisfy relations expressed by the following differential equations: . . . $d\epsilon = dH - dW$, $dW = p dv$, $dH = t d\eta$. Eliminating dW and dH yields

$$d\epsilon = t d\eta - p dv. \tag{7}$$

The quantities v, p, t, ϵ and η are defined if some state of the liquid is given, so one should call them functions of the state. The state of the liquid being understood in the fluid thermodynamics

sense admits two independent variations so there exist three finite equations between five quantities v, p, t, ϵ and η which are generally different for distinct substances but never contradict the differential equation (7).

In modern terms the 5-dimensional phase space of thermodynamics carries the contact structure (7) and states of the given substance form a Legendrian surface in it.

Contact geometry is well adapted to a phenomenological description of phase transitions. So figure 7 completely describes the Landau phase transition theory near Van-der-Waals' critical point. The projection of Legendrian surface into (ϵ, η, v) -space gives the internal energy graph. If we want to replace extensive variables v, η by intensive ones p, t and to preserve the contact structure under this then we automatically will come to Gibbs' free energy $G = \epsilon - t\eta + pv$:

$$dG = -\eta dt + v dp.$$

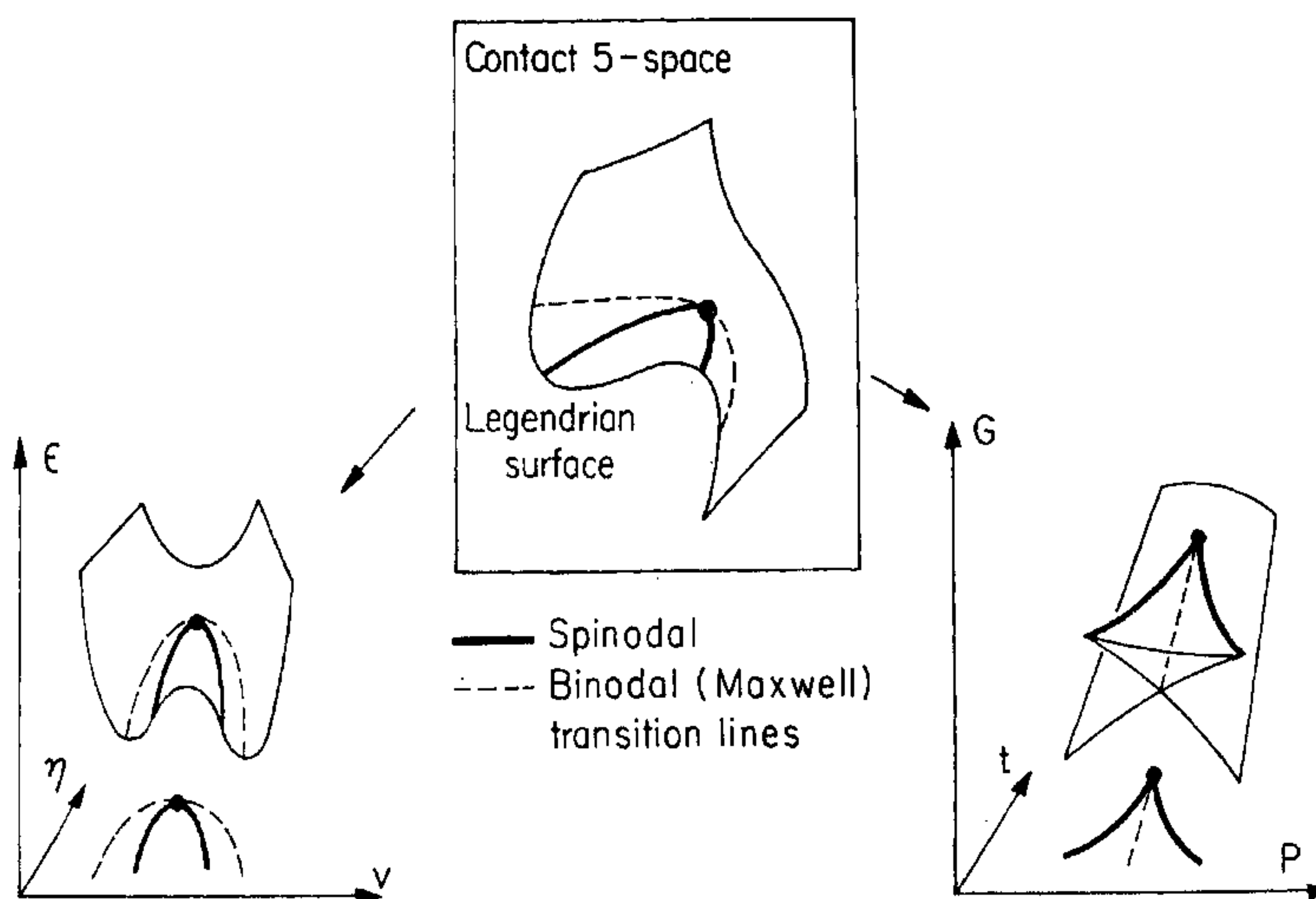


Figure 7.

The projection of the same Legendrian surface into (G, t, p) -space is the free energy graph and can be used as a catastrophe diagram for investigation of hysteresis phenomenon under phase transitions.

The appearance of the tetrahedron discriminant here is not just a chance. The point is that the contact geometry is well adapted to description of geometrical optics also. Velocities of rays form an n -dimensional Legendrian submanifold in $(2n + 1)$ -space of all possible velocities p , spatial coordinates q and time variable t with contact structure

$$dt = p_1 dq_1 + \dots + p_n dq_n. \tag{8}$$

The time function being single-valued on the Lagrangian submanifold is multi-valued being considered on q -space only. Its graph is a projection of the Legendrian submanifold into space-time.

Legendrian submanifolds appearing in geometrical optics are not always smooth as in figure 7. They may have singularities as in the case of diffraction (figure 8) for example. The catastrophe theory method of classification of stable time function graphs for Legendrian submanifolds with the simplest singularities leads to a surprising result. If n -dimensional Lagrangian submanifolds are

smooth or have singularities along $(n - 1)$ -dimensional smooth "edge" as in figure 8a (and look like a singular plane curve in transversal direction) then the classification in question coincides with that of symmetry groups of regular polyhedra in multidimensional spaces (Arnol'd 1972, 1978; Scherbak 1983, 1988; Givental' 1988). Graphs of corresponding time functions look like discriminants of these symmetry groups!

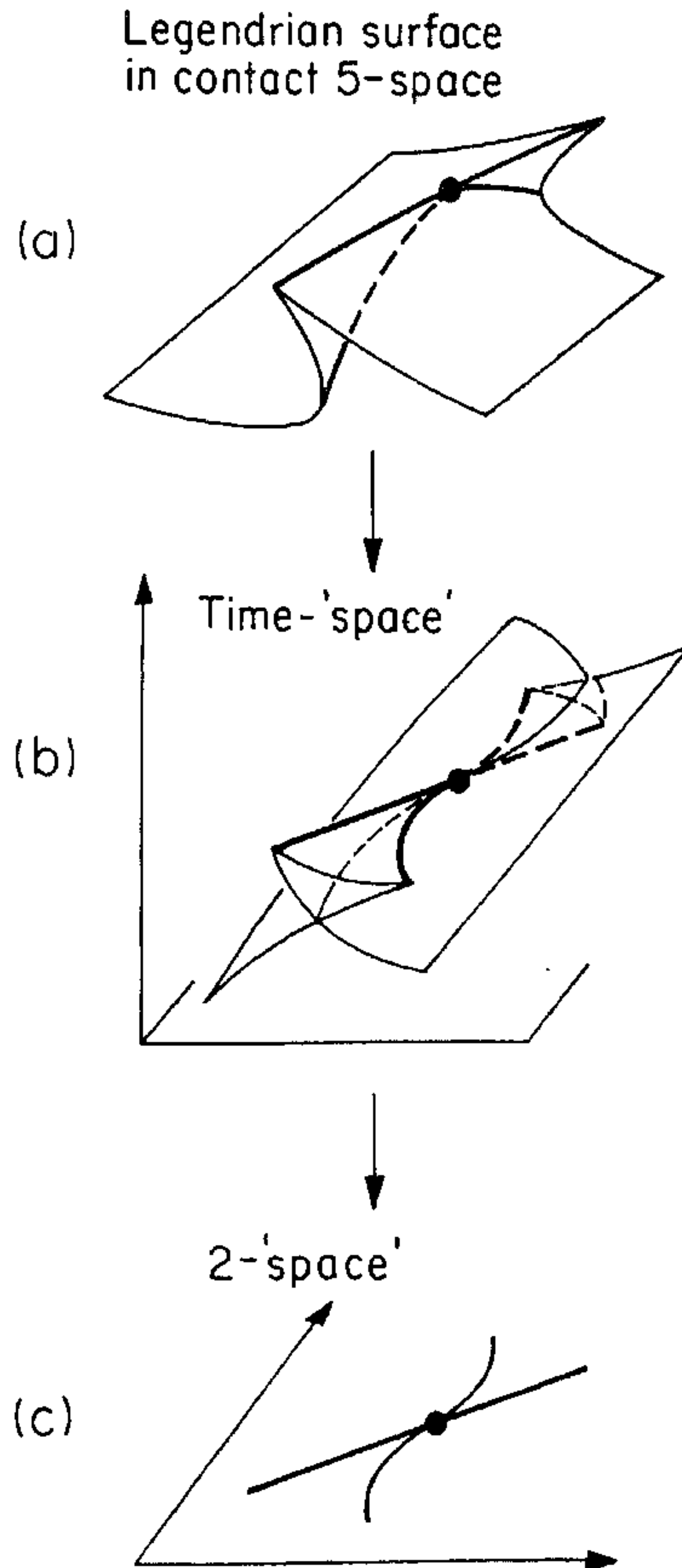


Figure 8.

SYMPLECTIC GEOMETRY AND STURM THEORY

The Hamiltonian method in conservative dynamics (of moving particles, rotating solids and so on) begins from the Extremal Action Principle. One should consider an action functional

$$S = \int [(p_1 dq_1 + \dots + p_n dq_n) - H(p, q, t) dt]. \tag{9}$$

The integration is carried out along a path $(p(t), q(t))$ in generalized momenta-configurations phase

space. Extremals of the action functional are motion trajectories in the phase space of the mechanical system with Hamiltonian function H . They satisfy the Hamilton equations $\dot{p} = -H_q$, $\dot{q} = H_p$.

Let us introduce the symplectic structure in the phase space

$$dp_1 \wedge dq_1 + \dots + dp_n \wedge dq_n. \quad (10)$$

This expression behaves like a scalar product but an anti-symmetric one and defines, by integration, the oriented area of a surface in the phase space. This means that the area may be positive or negative and should be equal to zero for closed surfaces.

The first term in (9) can be understood geometrically as oriented area of a surface bounded by the path. This leads to the following variational principle for autonomous Hamiltonian system. Its trajectories are extremals of the oriented area functional inside the class of paths situated at the same level of the Hamiltonian function. This variational principle is a prototype of all ones in mathematical physics. It allows us to consider Hamiltonian systems with phase spaces, more complicated than coordinate ones: we need a symplectic structure and a Hamiltonian function only. For example an ordinary sphere, provided with an area element, becomes the phase space of the classical spin system and this statement acquires an exact sense in geometrical quantization theory (see Kirillov).

Now we are going to explain how the instability theorems in diffusion kinetics should be proved. Until the space is one-dimensional it should be considered mathematically as time. Thus stationary dissipative states turn into movements of some non-linear autonomous Hamiltonian system!

An analogy with geometrical optics helps further. A particle, forced to remain at a curved surface, moves along a geodesic – a locally shortest line on the surface. But geodesics starting at a given point become non-minimal after the first focal point. For example meridians are geodesics on a globe, and the South pole is focal to the North one (of course this situation is unstable and on the geoid the focal point turns into focal locus – caustics – looking as in figure 9a). A pencil of neighbour geodesics is shown in figure 9b. By analogy a stationary dissipative state has a non-minimal energy (instead of geodesics' length) because of its periodical character (figure 9c).

Formally the problem can be reduced to determination of the number of negative eigenvalues of some non-autonomous linear Hamiltonian system, describing extremals of the functional (compare with (4)):

$$\int_a^b [A(t)\ddot{u}^2 + B(t)\dot{u}^2 + C(t)u^2] dt. \quad (11)$$

If $A \equiv 0$ then (11) yields the Sturm-Liouville (or stationary Schrödinger) equation. Then the number of zeroes of its solutions predicts the number of negative eigenvalues. This theory is well known by the name "Sturm theory". But it is not so well known that the same theory is valid for functionals (11) containing second or even higher derivations. Namely let us consider the space of functions on the interval $[a, b]$ satisfying "the first boundary condition"

$$u = \dot{u} = 0 \quad (12)$$

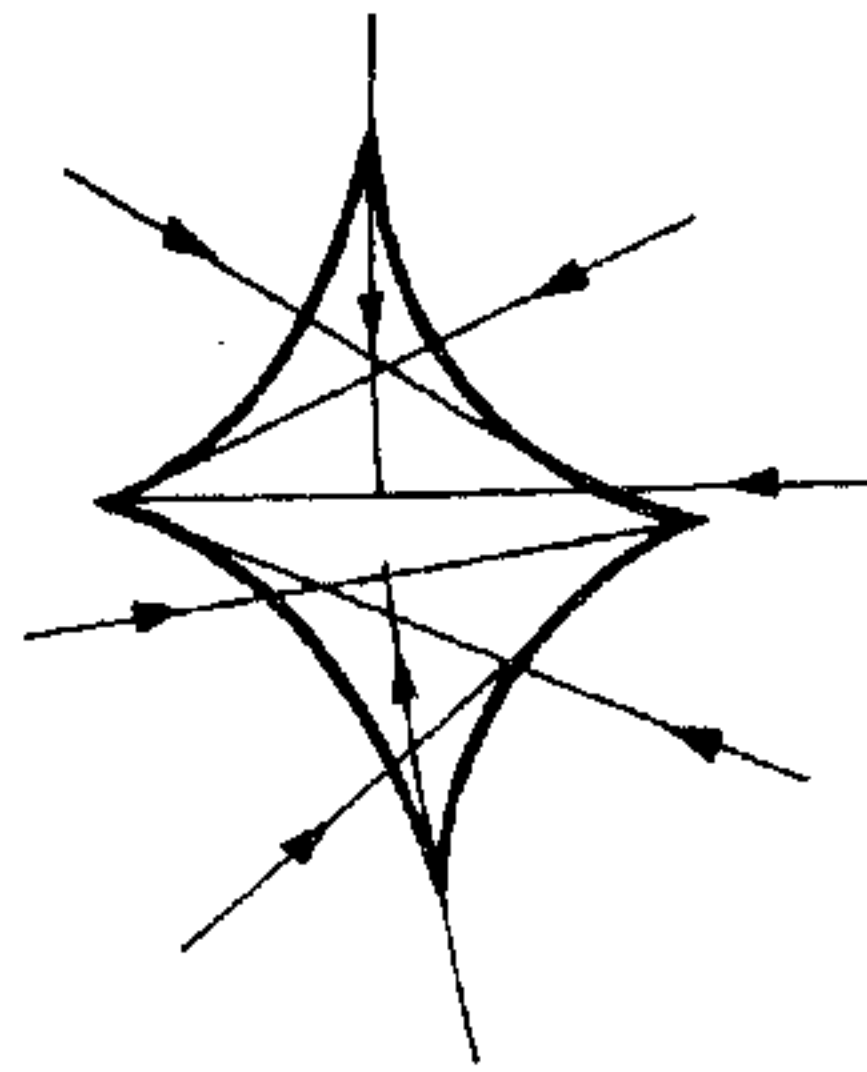
and call a point $a < t < b$ a focal one if there exists a solution of the corresponding linear Hamiltonian system, satisfying (12) at t .

Theorem. If A is positive then the number of negative squares of the functional (11) on our space of functions is equal to the number of the focal points.

A similar result is valid in a rather more general situation: the functional may depend on higher derivatives, may contain mixed terms and u may be a vector, so that A, B, C are symmetric matrices. The only need is that: the highest term A must be positive definite.

This generalized Sturm theory is also used in soliton stability results.

For an explanation of generalized Sturm theorems from the symplectic geometry viewpoint and for references see Arnol'd, 1984.



The Antarctic

(a)

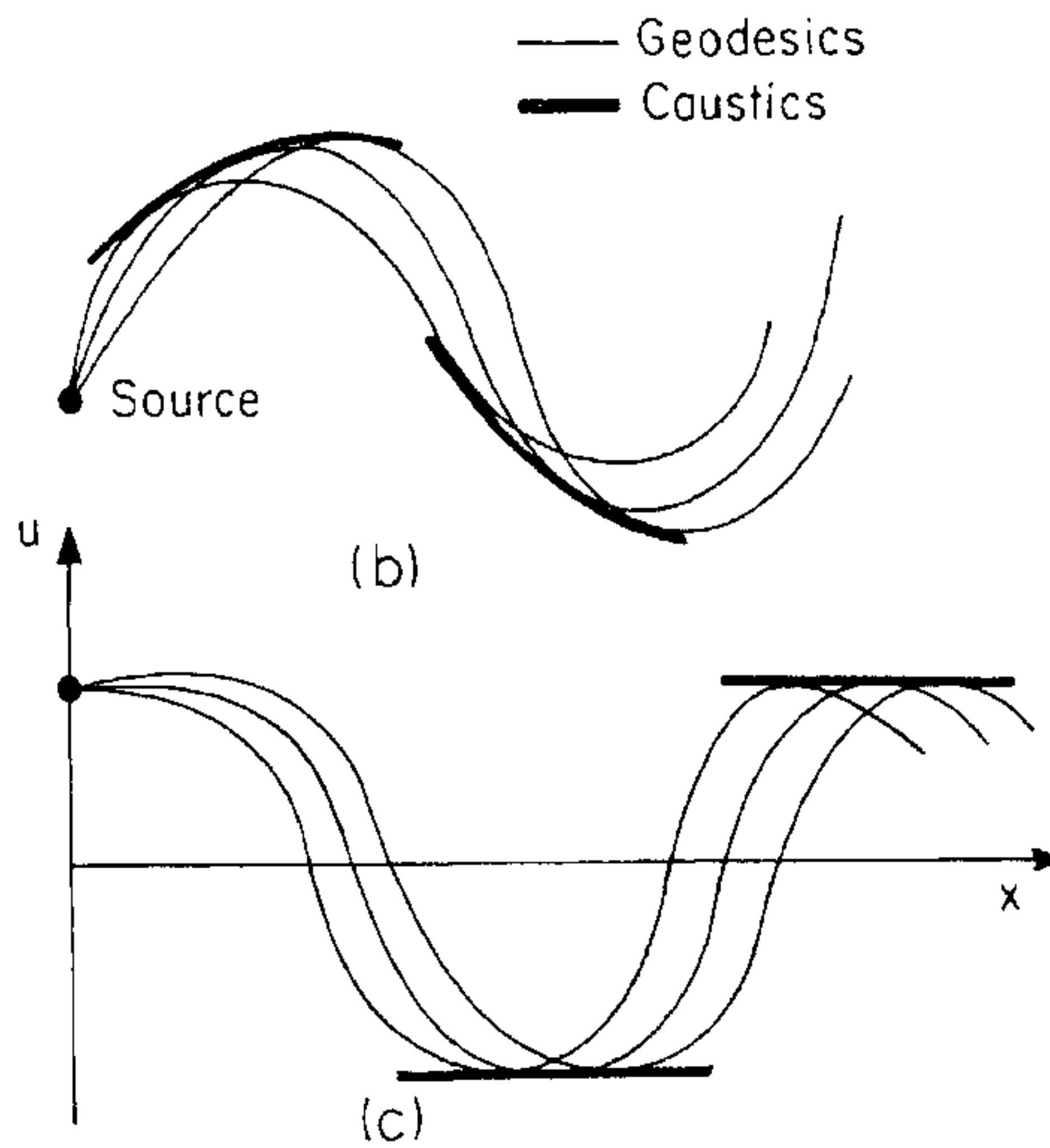


Figure 9.

CARTAN FORMALISM AND FIELD THEORY

If physical space in our kinetic equations is not one-dimensional then stationary states satisfy some field equation in the space. Thus to apply space-time analogy one needs Hamiltonian theory "with multidimensional time". Such a theory does exist, and is called Hamilton-Cartan formalism and leads to equations

$$\overrightarrow{\text{grad}} q = H_{\vec{p}}, \text{div } \vec{p} = -H_q. \tag{13}$$

The formalism starts from action functional (for example - in 3-time)

$$\begin{aligned} S &= \int p \wedge dq - H(p, q, t) dt, \\ p &= p_1 dt_2 \wedge dt_3 + p_2 dt_3 \wedge dt_1 + p_3 dt_1 \wedge dt_2 \\ dt &= dt_1 \wedge dt_2 \wedge dt_3, \quad q = q(t_1, t_2, t_3). \end{aligned} \tag{14}$$

The integration is carried out over some time domain T . Extremals of (14) satisfy (13) in T .

This approach allows us to formulate some "conic" generalization of Sturm theory (figure 10a). Considering "the first boundary problem" for functions in the domain T we call a level $t_1 = t$ focal if our linear boundary problem has non-trivial solution in the domain $a \leq t_1 \leq t$. Then the number of negative squares of the quadratic functional, whose extremals we are interested in, is equal to the number of focal levels between a and b if our functional is positive definite in the highest derivatives in other words – if our boundary problem is elliptic).

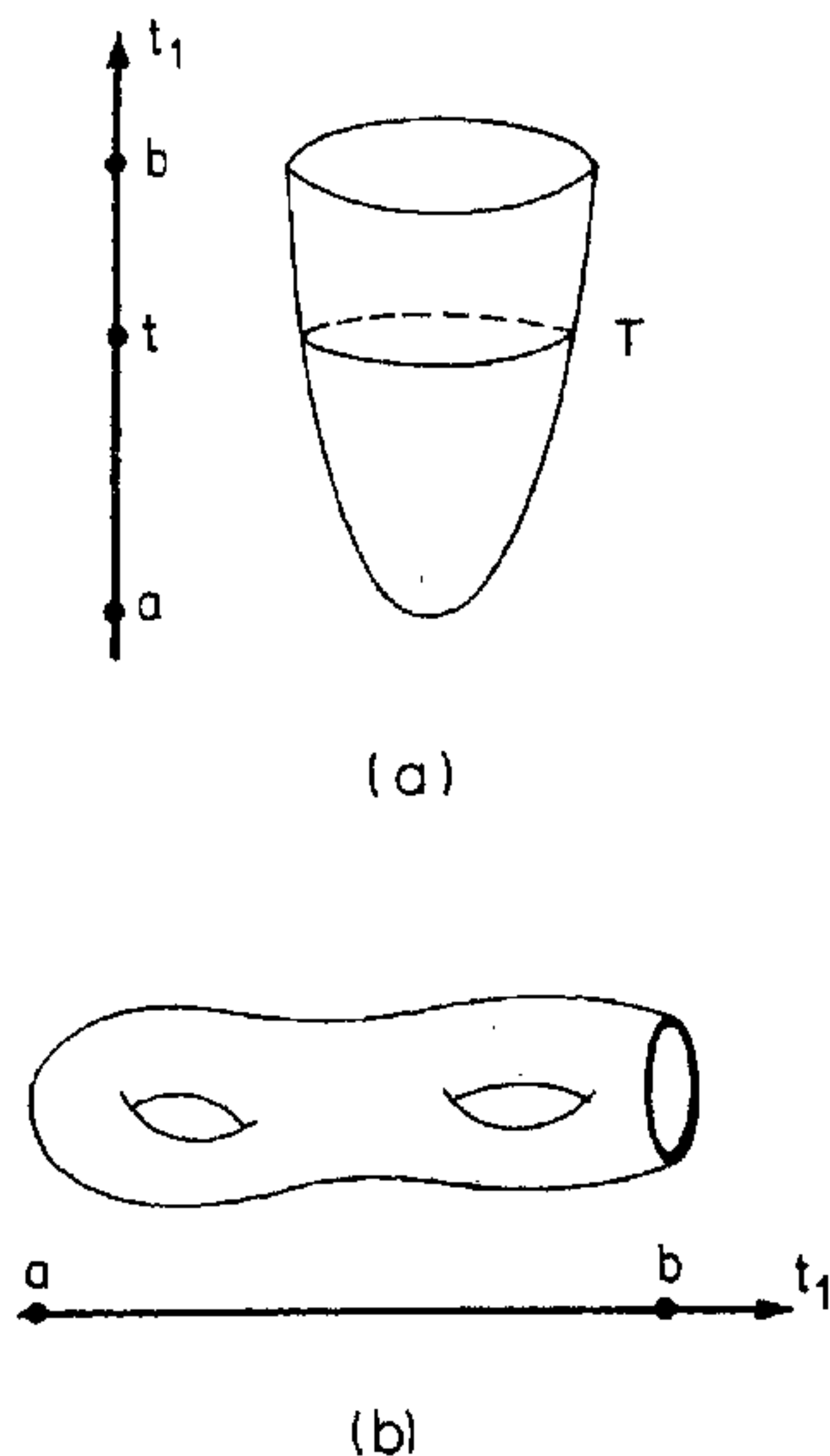


Figure 10.

In conclusion let us formulate two problems.

- (i) Extend "Sturm theory" to the topologically non-trivial situation where the time domain T is replaced by a manifold (figure 10b) and the elliptic quadratic functional is defined on sections of a vector bundle.
- (ii) Our formulation of Hamilton-Cartan theory is local in "time" T . Give a global one in which both "time" and "phase space" could be manifolds and exist separately.

LITERATURE

1. Arnol'd, V. I., Normal forms for functions near degenerate critical points, the Weyl groups A_k , D_k , E_k and Lagrangian manifolds, *Funct. Anal. Appl.* 1972, **6**, 254.
2. Arnol'd, V. I., Normal forms for functions on a manifold with boundary the Coxeter groups B_n , C_n , F_4 and evoluta singularities, *Russ. Math. Surv.* 1978, **31**, 2.
3. Arnol'd, V. I., *Mathematical methods of classical mechanics*, Moscow, Nauka, 1974.
4. Arnol'd, V. I. and Givental', A. B., Symplectic geometry, *Encyclopaedia of Math. Sc.*, 1988, **4**, VINITI-Springer.
5. Arnol'd, V. I., Symplectic geometry and Sturm theory, *Funct. Anal. Appl.* 1984, **18**, 4.
6. Cahn, J. W. and Hilliard, J. E., Free energy of a nonuniform system I, *J. Chem. Phys.*, 1958, **28**, 258.
7. Givental', A. B., Unstability of stationary structures under diffusion kinetics, *J. Theor. Math. Phys.* 1988, **76**, 462.
8. Givental', A. B. and Dobrynin, A. V., Phase separation and 2-zone potentials, *Usp. Mat. Nauk*, 1988, **41**, 3, (in Russian only).
9. Givental', A. B., Criterion of soliton stability (Presented to *J. Theor. Math. Phys.*).
10. Givental', A. B., Singular Lagrangian varieties and their Lagrangian mappings, *Contemp. Probl. of Math.*, 1988, **33**, VINITI, (in Russian).
1. Kirillov, A. A., Geometrical quantization, *Encyclopaedia of Math. Sc.*, 1988, **4**, VINITI-Springer.