RMP Colloquia

This section, offered as an experiment beginning in January 1992, contains short articles intended to describe recent research of interest to a broad audience of physicsts. It will concentrate on research at the frontiers of physics, especially on concepts able to link many different subfields of physics. Responsibility for its contents and readability rests with the Advisory Committee on Colloquia, U. Fano, chair, Robert Cahn, S. Freedman, P. Parker, C. J. Pethick, and D. L. Stein. Prospective authors are encouraged to communicate with Professor Fano or one of the members of this committee.

Approach to chaos through instabilities

Hie-Tae Moon

Physics Department, Korea Advanced Institute of Science and Technology, Daeduk Science-town 305-701, Korea

Breaking of spatial homogeneity and the emergence of spatio-temporal order and disorder are striking phenomena abundant in nature. One possible origin for such ordering and disordering processes and for the transition between them is discussed.

CONTENTS

I. INTRODUCTION

Nature is full of complex phenomena, yet very many of them have long evaded the grasp of the human intellect. The governing equations for such phenomena, are available in many cases, usually in the form of nonlinear partial differential equations. These equations, however, are difficult to work with and have made for slow progress in our understanding of complex phenomena. The past two decades have seen an upsurge of interest in understanding the qualitative contents of these equations, under the name of "chaos."

In the physical sciences, important advances have been further stimulated by the interplay between theory and experiment. From the experimental side, oscillating physicochemical reactions have been the subject of fairly extensive research efforts. They exhibit very rich dynamical behavior including chaos; they also have certain implications for biological systems. Here we consider this phenomenon from a semi-empirical point of view in order to gain insight into the nature of the chaos involved. The study of physicochemical oscillations has wider applications, as we shall see, to many other physical situations. We first briefly outline the experimental facts on these chemical reactions.

A. Physicochemical oscillations: experimental facts

Belousov (1959) reported that a well-stirred homogeneous solution containing bromate ions, malonic acid, sulfuric acid, and a small amount of catalysis, cerous ion, and iron will undergo temporal oscillation. If one chemical component is represented by red and the other by blue, the whole solution turns alternately from red to blue and back again, say, every five minutes. This "chemical clock" behavior was surprising, as it certainly appeared counterintuitive.

Various scientists have since reported the ability of such a solution to form spatial patterns. For example, when the reaction proceeds in a long thin, vertical tube, horizontal bands may appear corresponding to alternating high-concentration regions of the chemicals (Busse, 1969). This is an instance of a spatially inhomogeneous mixture in a single dimension.

If a system remains closed and the raw materials necessary for the reaction are exhausted, its oscillations eventually die out. In the usual open reactors, in the form of long narrow horizontal cylinders, raw materials are constantly supplied and then washed away after use. Chemical reactions in such open systems can present striking features of self-sustained spatio-temporal oscillation, including chaos. Understanding the characteristics of chaotic oscillations has been a central issue during the past decade (Epstein, 1983; Roux *et al.*, 1983).

 0034-6861/93/65(4)/1535(9)/$05.90

B. Theoretical framework

Chemical reaction-diffusion (Nicolis and Prigogine, 1977) and many other physical phenomena are governed by macroscopic equations of the type

$$\mathbf{C}_t - D\mathbf{C}_{rr} = \mathbf{F}(\mathbf{C}) , \tag{1}$$

where the particle density vector $\mathbf{C}$ represents the concentrations $(c_1, \ldots, c_n)$ of n constituents. This equation relates the nonuniformity of a concentration gradient to diffusion through the second-order space derivative. The mixture is assumed to be so dilute that the diffusion coefficients D_{ij} are constant and no couplings occur between fluxes of different components i and j, yielding a diagonal matrix D of constant diffusion coefficients. The function $\mathbf{F}$ describes general nonlinear phenomena. We further assume there that no convection results from the mass flow,[1] so that the function $\mathbf{F}$ describes the overall rate of production of c_i purely from the chemical reactions. In such a case the F_i will be nonlinear polynomial functions of the $\{c_i\}$ reflecting the products of densities relevant to collision frequencies.

Consider a system that admits a steady uniform solution $\mathbf{C}_0$ for which the nonlinear term $\mathbf{F}(\mathbf{C}_0)$ identically vanishes. Now add to $\mathbf{C}_0$ a small perturbation $\mathbf{c}$, so that the system is homogeneous in space but oscillating in time with a characteristic frequency ω_0. The situation may be expressed as $\mathbf{C}=\mathbf{C}_0+\mathbf{c}$, where $c_i = a_0 e^{i\omega_0 t}$, $a_0 \ll |\mathbf{C}_0|$, with ω_0 determined by the system. Any friction in the system may cause the fluctuation to die out. Suppose we now furnish a "vitamin" to the fluctuation, which then starts to grow at a certain critical rate R_0. Here R may denote a catalysis, or temperature, called in general a "control parameter." Thus at $R = R_0$, the linear growth rate ζ of a_0 ($\sim e^{\zeta t}$) changes from negative to positive. But if R continued to grow, the nonlinear term F would become important, bringing the system into the nonlinear regime, which we do not know much about.

Just above the critical threshold, however, we may still understand the behavior by writing a new solution as $c_i = a_0 \psi e^{i\omega_0 t}$, where ψ presumably describes the possible modifications of the homogeneous phase. For definiteness, let ϵ be the order of magnitude of the fluctuations, $a_0 \sim \epsilon$, and we express the post-critical regime in terms of R as $R = R_0(1+O(\epsilon^2))$. The relevant long time and large length scales that follow accordingly are[2] $\tau = \epsilon^2 t$, $\xi = \epsilon \mathbf{r}$. The introduction of the smallness parameter ϵ can then be exploited to single out the dominant term of Eq. (1). The precise form of the equation for the slowly varying amplitude function $\psi(\xi,\tau)$ reduces to (Kuramoto and Tsuzuki 1975)

$$\psi_\tau = \alpha_r \psi + \beta \psi_{\xi\xi} - \gamma |\psi|^2 \psi , \tag{2}$$

where α_r is real, while $\beta = \beta_r + i\beta_i$ and $\gamma = \gamma_r + i\gamma_i$ are in general complex. For example, β_r and β_i are expressed for a two-component system in terms of the diffusion coefficients of the constituents as $\beta_r \propto (D_{11}+D_{22})$ and $\beta_i \propto (D_{11}-D_{22})$. We have thus extracted from the set of rather complicated dynamical equations the only part physically relevant to dynamical phase transitions.

Once ψ is determined from Eq. (2), the general solution, expressing $\psi = |\psi| e^{i\Theta}$, can be written up to the order of ϵ, as

$$c_1(x,t;\xi,\tau) = 2\epsilon |\psi| \cos(\omega_0 t + \Theta) , \tag{3}$$

$$c_2(x,t;\xi,\tau) = 2\epsilon |a| |\psi| \cos(\omega_0 t + \Theta + \varphi) , \tag{4}$$

where $|a|$ is generally a complex constant and $\tan\varphi = \mathrm{Im}a / \mathrm{Re}a$. Thus the different variables become interlinked through ψ as they grow. On the other hand, Eq. (2) admits a spatially homogeneous but temporally oscillating solution,

$$\psi(\xi,\tau) = A_0 e^{-i\gamma_i A_0^2 \tau}, \quad A_0 \equiv \sqrt{\alpha_r/\gamma_r} . \tag{5}$$

This solution, when incorporated into Eq. (3), shows how the initially infinitesimal $\mathbf{c}$ relaxes to temporal oscillation with a relatively large amplitude in the post-critical regime, thus explaining the spatially uniform bulk oscillatory "chemical clock" behavior observed by Belousov (1959).

This homogeneous state, however, can be unstable to particular types of fluctuating spatial inhomogeneity whose amplitudes may grow into a nonlinear pattern. In this Colloquium, we limit ourselves to discussing reactions in a homogeneous phase, where breaking of spatial homogeneity and the emergence of spatio-temporal order

---

[1] If there is a mass flow $\mathbf{v}$, its time rate $d\mathbf{v}/dt = \partial\mathbf{v}/\partial t + (\mathbf{v}\cdot\nabla)\mathbf{v}$ naturally involves the quadratic nonlinear convective term $(\mathbf{v}\cdot\nabla)\mathbf{v}$, which complicates the reaction phenomenon. Here mass transport is restricted to be possible only through the diffusion mechanism of the second-order space derivative term. In experiments, a porous gel medium is used as a substrate to remove convection.

[2] For this post-critical regime of $R = R_0(1+O(\epsilon^2))$, the linear growth rate must be $\zeta \sim \Delta R \sim O(\epsilon^2)$. This suggests that fluctuations will grow in time on the scale $\tau = \epsilon^2 t = O(1)$. In many physical systems, only a single characteristic mode becomes unstable at $R = R_0$, while the neighboring modes become unstable as R increases above $R = R_0$, yielding $\partial R/\partial k|_{k_0} = 0$ at $R = R_0$. The linear growth rate ζ of the neighboring mode with a wave number $k = k_0(1+O(\epsilon))$ then becomes $\zeta \sim \Delta R \sim \partial^2 R/\partial k^2|_{k_0} (\Delta k)^2 \sim \epsilon^2$, consistent without initial setting of the post-critical regime as $R = R_0(1+O(\epsilon^2))$. The effect of wave-number disturbance becomes important when $(\Delta k/k_0)x (\sim \epsilon x)$ becomes $O(1)$, which naturally introduces a relevant large length scale $\xi = \epsilon x = O(1)$.

are most striking. We thus aim here to describe the dynamical phase transitions in terms of a universal function ψ.

II. THE UNIVERSAL FUNCTION FOR POST-CRITICAL PHENOMENA

Equation (2) is a universal evolution equation pertaining to the post-critical regime for many physical systems, whose macroscopic governing equations are of the type (Newell, 1974).

$$L(\partial/\partial t, \partial/\partial x; R)\phi(x,t) = M\phi N\phi \ , \tag{6}$$

where L, M, N are linear operators and R is a control parameter. Significantly, the Navier-Stokes equations of motion for fluids fall into this group. The acceleration of the fluids, $d\mathbf{v}/dt = \partial\mathbf{v}/\partial t + (\mathbf{v}\cdot\nabla)\mathbf{v}$, naturally involves the quadratically nonlinear convective term, and its control parameter R is the Reynolds number, namely, the product of the system's mean velocity and of its characteristic length divided by the molecular viscosity. Consequently Eq. (2) appears in various phemomena involving hydrodynamics.

Various authors have derived Eq. (2) in different contexts. We note among others that the equation was originally considered by Ginzburg and Landau in the context of superconductive phase transitions (Lifshitz and Pitaevskii, 1980), and independently by Newell and Whitehead (1967) in the context of Bénard convection[3] and by Kuramoto and Tsuzuki (1975) in chemical reaction-diffusion.

The dynamics of Eq. (2), however, depends quite sensitively on its coefficients. When all of these are real positive, as in the phase transitions of superconductivity and in Bénard convection, energy is constantly supplied at large length scales through the α_τ term. Physically, this supply can be represented by boundary conditions, for instance, by keeping two points of the system at different temperatures. The energy is then dissipated at small length scales through the β_r term for example, in the form of heat flow. In the absence of space-dependent material properties, the nonlinear damping term γ_r would alone damp the energy inflow. The situation here is highly dissipative, flattening out any spatial inhomogeneity.

Another limit arises when the coefficients are all imaginary, effectively representing an energy-conserving Hamiltonian situation. Equation (2) then reduces to the widely known nonlinear Schrödinger equation describing, among others, the electric field modulation in a medium whose refractive index depends on the amplitude (Kelly, 1965), as well as deep-water gravity waves (Benjamin and Feir, 1967). Here the β_i term provides the wave dispersion ("spreading") effect while γ_i yields the steepening or focusing effect. Combination of these two terms can produce large-scale wavelike structures in the medium.

When the coefficients are complex, any asymptotic behavior hinges on the overall balance of energy gains and losses as well as on steepening and smoothening effects. This situation arises variously in fluid as well as in plasma physics. In chemical reaction-diffusion systems, in which $\alpha_r, \beta_r, \gamma_r$ are all positive, new raw material is constantly supplied through the α_r term, through the boundary condition of steady matter inflow. The molecules are constantly diffused at small length scales (β_r term) and then washed away, effectively preventing the formation of small-scale structures.

Even with the same hydrodynamic origin, the coefficients depend on the boundary conditions as well as on the different elements of molecular diffusivity. For example, β_i vanishes in Bénard convection (Newell and Whitehead, 1967), where thermal diffusion is the only mechanism. On the other hand, for thermo-haline double diffusive systems, relevant to ocean convection in which heat and salt are the two components with different diffusion rates, we again recover $\beta_i \neq 0$ (Bretherton and Spiegel, 1983). In this context, we note that, for two-component chemical systems, β_r and β_i are explicitly expressed in terms of the diffusion coefficients of each element as $\beta_r \propto D_1 + D_2$, $\beta_i \propto (D_1 - D_2)$. The important roles played by these terms will be pointed out later. We merely mention here, that equality of D_1 and D_2 leads to a single diffusive system with $\beta_i = 0$, as in Bénard convection

A. Modulational instability

The spatially homogeneous state of Eq. (5) becomes in fact unstable to a particular type of fluctuating spatial inhomogeneity, whose origin can be addressed as follows. Suppose we have a nearly monochromatic wave packet with a narrow bandwidth of the order of ϵ. In a linear situation, the modes within the packet simply spread out, moving with their own individual phase velocities. In most practical situations, however, they are not independent from each other, but undergo nonlinear interactions among themselves. For simplicity, we approximate the wave packet as consisting of just three wave-number components: k_0, k_u, k_l, where k_0 is the central mode and k_u, k_l are neighboring upper and lower sidebands with $k_u - k_0 = O(\epsilon)$. We represent the three waves as $a_0 e^{i(k_0 x - \omega_0 t)}$, $a_u e^{i(k_u x - \omega_u t)}$, $a_l e^{i(k_l x - \omega_l t)}$, where the amplitude a_0 is small, the amplitudes a_u, a_l are even smaller, and the wave number and frequencies satisfy $k_l + k_u = 2k_0, \omega_l + \omega_u = 2\omega_0$. We assume that the original

[3]Bénard convection is another example of the instability of a stationary state's giving rise to a phenomenon of spatial pattern formation. In this system the fluid is contained between horizontal, thermally conducting plates heated from below. When the imposed temperature gradient reaches a threshold value, the stationary state becomes unstable to perturbation, forming convective rolls. Here ψ denotes the strength of the rolls.

system has quadratic nonlinearity as in Eq. (6), causing the second harmonic of the central component, $a_0^2 e^{2i(k_0 x-\omega_0 t)}$, to interact with the upper sideband, producing

$$a_0^2 a_u e^{i[(2k_0-k_u)x-(2\omega_0-\omega_u)t]}=a_0^2 a_u e^{i(k_l x-\omega_l t)}, \quad (7)$$

whose parameters match those of the lower sideband.[4] Similarly, the interaction between the second harmonic and the lower sideband produces a term proportional to the upper sideband. The presence of both the upper and lower sidebands thus combines with the second harmonic, bringing about a mutual reinforcement. This resonant mechanism has been known as the "modulation instability" (Benjamin and Feir, 1967). Indeed, this instability provides the physical foundation for many cooperative processes in plasma and laser physics and in chemical reaction-diffusion systems.

B. Linear stability analysis

The long-time evolution of the resonance mechanism described above is represented conveniently in terms of the universal amplitude function ψ. For convenience, we normalize the equation through the transformations

$$\Psi=\sqrt{\gamma_r/\alpha_r}\,\psi, \quad T=\alpha_r\tau, \quad X=\sqrt{\alpha_r/\beta_r}\,\xi .$$

Setting now $c_d\equiv\beta_i/\beta_r$, $c_n\equiv\gamma_i/\gamma_r$, we see that the new state variable $\Psi(X,T)$ satisfies the following dimensionless equation:

$$\Psi_T=\Psi+(1+ic_d)\Psi_{XX}-(1+ic_n)|\Psi|^2\Psi . \quad (8)$$

The spatially homogeneous solution of Eq. (5) reduces then to $\Psi_e(T)=e^{-ic_n T}$. At $T=0$, we assume it to include upper and lower sidebands of the type (7),

$$\Psi(X,0)=\Psi_s(0)+\delta(e^{iqX}+e^{-iqX}), \quad \delta\ll 1 , \quad (9)$$

where q is the sideband wave number corresponding to the large length scale X; $q\sim(k_u-k_0)/\epsilon=O(1)$.

When we set $\delta\sim e^{\sigma t}$, Eq. (8) implies that σ becomes positive real for $0<q<q_m$ with a threshold wave number q_m expressed in terms of c_d and c_n provided $1+c_n c_d<0$. The fluctuation thus builds up within the range $0<q<q_m$, whereas the initial state becomes unstable. Figure 1 plots the linear growth rate σ as a function of the sideband wave number q for $c_d=5$, $c_n=-4$, leading to the instability threshold $q_m=1.209$ beyond the peak growth rate $\sigma=2.4$ at $q=0.8$. The choice of $c_d=5$, $c_n=-4$ is well suited for a two-component chemical system, but, for the sake of generality, we monitor the situation in terms of the stability diagram. In this study, we deal with the phenomena taking place in the hatched unstable region.

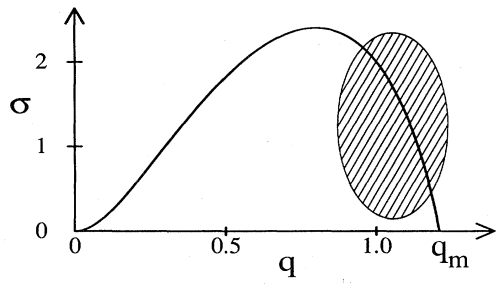

FIG. 1. Stability diagram as a function of sideband wave number q. Here we consider the shaded unstable region, which ranges from the instability threshold $q_m=1.21$ to $q=0.9$ ($\sigma=2.3$), practically the most unstable region. The maximum growth rate is $\sigma=2.4$ at $q=0.8$.

III. BEYOND THE THRESHOLD OF INSTABILITY

A. The emergence of spatio-temporal order

Figure 2 summarizes the evolution near the instability threshold (Moon, 1991). For $1.1<q<q_m=1.209$, a steady state with permanent spatial form evolves. As q drops below 1.1, a new state develops exhibiting recurrent pulses, called "ticks," further into this state a tick is followed by a "tock," namely by a pulse suddenly shifted to neighboring sites. A tick is followed by a single tock with clocklike precision, displaying the phenomenon of spatial oscillation.

In the context of chemical reaction-diffusion, c_n is now normalized as $c_d\propto(D_{11}-D_{22})/(D_{11}+D_{22})$ for a two-component system. As mentioned earlier, a complex c_d term yields a spreading dispersive effect corresponding to a restoring force akin to the tension on a string. Two components of different molecular diffusivities thus yield

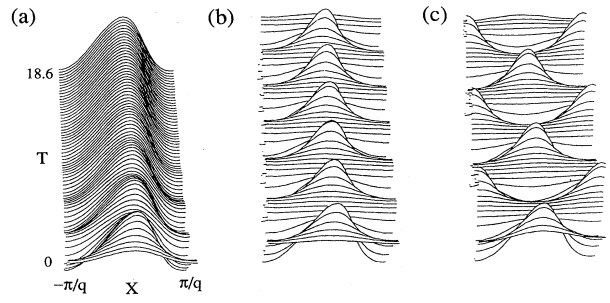

FIG. 2. Emergence of spatio-temporal order in a plot of $|\Psi(X,T)|$: (a) A steady spatial structure at $1.1\le q\le q_m$; (b) a recurring sequence of pulses, which we call "ticks," $1.0<q<1.1$; (c) a state exhibiting spatial oscillations; a tick is followed here by a pulse space-shifted orthogonally (out of phase), which we call a "tock," $0.985\le q\le 1.0$.

[4]Notice the coefficient in Eq. (7), $a_0^2 a_u$, which is cubic in character. Even though the governing nonlinearity of the original system is quadratic, the first nontrivial nonlinearity in the post-critical regime in generically cubic.

a pair of modes that work cooperatively to block the creation of small-scale structures but also compete for the role of the restoring force. This restoring force (originating from diffusion) and the nonlinear driving force (from the chemical process) contribute jointly to the generation of macroscopic patterns.

The first two patterns of Fig. 2 show the formation of a high-concentration range of the X coordinate. However, they pursue different temporal dynamics. The first pattern is steady,[5] while the second pattern exhibits an oscillating population of molecules, showing a *local* "chemical clock" behavior. In the context of deep-water gravity waves, the behavior of Fig. 2(b) has been referred to as the FPU (Fermi-Pasta-Ulam) recurrence phenomenon[6] (Yuen and Ferguson, 1978). Finally, the band of high concentration can flip-flop, exhibiting spatial oscillation between two neighboring sites. These problems have not found a satisfactory answer until recently.

B. Double-well potential in the medium

We now look for the origin of the spatio-temporal order. Let us view the behavior of the concentration as a motion on a string. A linear sinusoidal vertical oscillation of the string at a local point is completely understood by studying the displacement and its velocity, which parallel the motion of a particle moving near the bottom of a parabolic well. Correspondingly, one may define (Moon, 1990)

$$A(T)=|\Psi(0,T)|-1, \quad B(T)=dA(T)/dT , \qquad (10)$$

where $A(T)$, taken at the origin, denotes the concentration measured from the unperturbed homogeneous reference state and $B(T)$ is its rate of change in time. The unperturbed homogeneous state, for example, is represented by the origin of coordinates in this two-dimensional *phase* space. The three phase paths, corresponding respectively to the three patterns of Fig. 2, are drawn and labeled a, b, c in Fig. 3.

Notice that the phase-space orbits thus obtained can also represent the motion of a particle in the double-well potential shown in the inset, assuming A now denotes the distance of the particle and B its velocity. The locally bounded motion of energy $E=E_b$ corresponds to the orbit b, while the locally unbounded motion ($E=E_c$) yields the large orbit c. The ridge potential ($E=E_0$) separates a locally bounded motion from a broader-ranging motion, thus affording a long-range order. The summit of the potential barrier corresponds to the saddle point denoted by "+" in the phase space. The phase path along the ridge potential, called a "separatrix," passes through this singular point, showing the double-well potential structure of the field $|\Psi(X,T)|$.

C. An application to optical communication

The wave patterns of Fig. 2(b) and Fig. 2(c) still exist in the Hamiltonian limit, $\sigma_r=\beta_r=\gamma_r=0$, of Eq. (2) (Moon, 1990). The double-well potential structure still occurs, but does not allow the steady state of Fig. 2(a). In optical communication through glass fibers, we encounter the same equation except that the roles of X and T are interchanged, so that the horizontal X axis in Fig. 2 now refers to time (pulse width), while the vertical axis indicates the distance traveled along the fiber.

It was proposed that the tick pulses of Fig. 2(b) would provide an extra-high-bit-rate optical communication; experiments did demonstrate that one could produce such pulses by applying a sideband modulation with a high repetition at the fiber input of an incoming monochromatic light wave (Tai *et al.*, 1986). One problem of long-distance transmission, however, lies in the gradual power loss along the distance traveled. Currently, long-

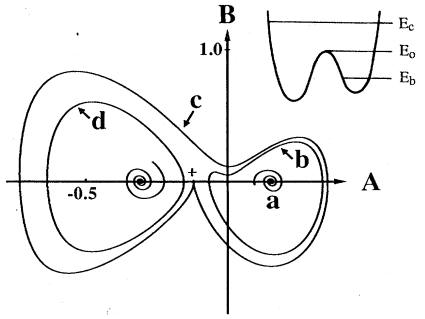

FIG. 3. Phase paths in the two-dimensional phase space spanned by $A(T)=|\psi(0,T)|-1$ and $B(T)=dA(T)/dT$. The paths denoted by a,b,c correspond, respectively, to the three patterns of Fig. 2. The orbit d on the left is obtained when $A(T)$ is measured at a site, displaced to the left by a half wavelength. Notice that the same trajectories are followed by a particle moving in the double-well potential shown in the inset. The locally bounded motion with energy $E=E_b$ corresponds to the orbit b, while the locally unbounded motion ($E=E_c$) yields the large orbit c. The ridge potential ($E=E_0$) separates the locally bounded motion from the unbounded motion, thus affording a long-range order. The summit of the potential barrier corresponds to the saddle point denoted by + in the phase space. The phase path of the ridge potential passes through this singular point. The concentration therefore evolves in the double-well potential field.

[5]But it requires a constant supply of new raw materials to retain its shape because molecules constantly diffuse away in regions of density gradient.

[6]An important assumption of classical mechanics has been that "small nonlinearities" lead to equipartition of energy. Fermi, Pasta, and Ulam tested this assumption by studying the vibration of 64 masses connected by nonlinear strings. The results were surprising. No tendency toward thermalization was observed. Energy originally put in the lowest-frequency mode returned almost entirely to that mode after a while. A fairly extensive literature has since developed to understand what has become known as the FPU problem.

distance transmission is considered feasible with the use of optical amplification. It was soon realized, however, that repeated amplification degrades the signal-to-noise ratio in proportion to the number of amplifications, thus presenting a severe limitation to exploiting such pulses (Hasegawa and Tai, 1989).

The pulse dynamics in the fiber are governed by the dimensionless equation[7]

$$\Psi_X = [-\eta + \zeta \sin^{2n}(\pi X/L)]\Psi + i\Psi_{TT} + i|\Psi|^2\Psi \ , \qquad (11)$$

where η denotes the gradual power loss and the ζ term represents the periodic power amplification. The specific form $\zeta \sin^{2n}(\pi X/L)$ is proposed here to represent the amplifiers spaced at intervals L. $L=10$, $n=2$ are chosen. At the input of the fiber the incoming light is modulated as $\Psi(0,T)=1+0.01\cos(\omega T)$ with the repetition frequency $\omega=1$; this is the frequency at which linearity is most unstable. For weak loss and amplification, represented by $\eta=0.005$ and $\zeta=0.002$, Fig. 4 displays a possible signal distribution along the fiber, where the three different stages are denoted by I, II, III, respectively. In the first stage, we observe tick pulses only. Ideally, one would hope to see only tick pulses throughout to achieve good quality of communication. At the second stage, however, a tick is suddenly time-shifted orthogonally. In the double-well potential picture of Fig. 3, these pulses arise by crossing the ridge potential from the right to the left well and staying in the left well, a stage characterized by the phase orbit "d" in Fig. 3. The third stage indicates another crossing of the ridge potential from the left well to an unbounded higher-energy state, where a tick is followed by a tock. The result would be just catastrophic if such crossings occurred irregularly. Clearly care is required not to upset the ridge potential in exploiting tick pulses for optical communication.

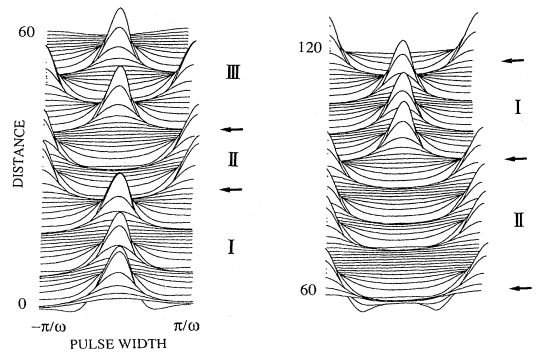

FIG. 4. Wave propagation in an optical fiber with weak power-loss and amplification. Only "tick pulses" occur in the first stage denoted by I. At the second stage II, all ticks are time-shifted orthogonality, according to the double-well potential model of Fig. 3, passing from the right well into the left, where each tick remains. This tick propagation corresponds to the phase path d in Fig. 3. The third stage is the result of crossing from the left well to the locally unbounded higher-energy state, where a tick is followed only by a tock. Each crossing is indicated by an arrow. Note that the ridge potential, which afforded the long-range order, is now a source of disorder in the presence of weak perturbations.

IV. APPROACH TO CHAOS

Figure 4 illustrates how the ridge potential, which affords a long-range order, can also become self-destructing in the presence of weak perturbations of the system. Indeed, the ridge potential is often a harbinger of chaos in dissipative systems. We first discuss the chaos found at $q=0.92$ in the initial condition of Eq. (9).

A. Deterministic chaos

Figure 5(a) displays the chaotic wave field at $q=0.92$, exhibiting irregular spatio-temporal oscillations. How

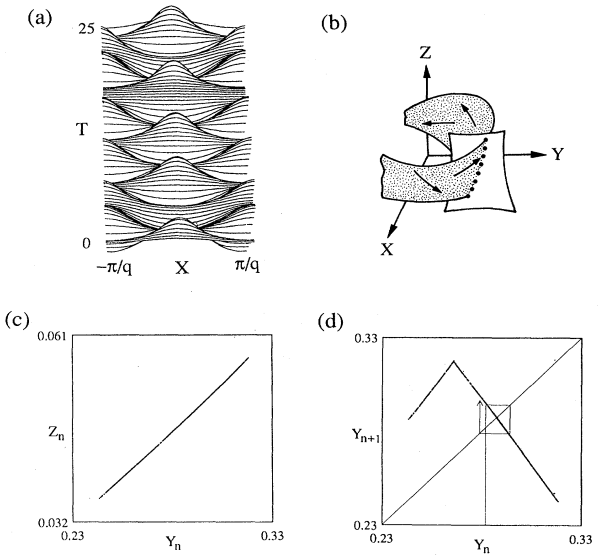

FIG. 5. Successive stages of evolution: (a) Emergence of disorder at $q=0.92$. The spatial pattern is transformed into the Fourier 3D phase space spanned by $X\equiv|a_0|$, $Y\equiv|a_q|$, $Z\equiv|a_{2_q}|$, where $|a_k|$ is the amplitude of the Fourier component with wave number k. (b) the phase-space trajectory is drawn schematically. A Poincaré surface of section $X\approx\langle X\rangle$, where $\langle X\rangle$ is the time average of X, is placed across the direction of the flow, generating a sequence of points $[Y(t_n),Z(t_n)]$, where t_n is the time of the nth piercing of the section. The linear arrangement of the set of crossing points indicates that the trajectory lies on a strip. (c) a plot of the set of points shown in Fig. 5(b). (d) by plotting Y_{n+1} vs Y_n, one effectively reduces the 3D dynamics into the 1D dynamics $Y_{n+1}=G(Y_n)$. A well-defined functional form G appears. Arrows are inserted to show iteration of G. For convenience, a line is drawn at a 45° angle.

[7]The dimensionless variables are related to the physical variables as $\Psi\propto\phi/\epsilon$, $T\propto\epsilon(t-x/v_g)$, $X=\epsilon^2x/\lambda$, where ϕ is the electric-field amplitude of the pulse envelope, t is the time, x is the distance of transmission, λ is the wavelength of the carrier wave, and v_g is the group velocity.

can this evolution be truly chaotic, when it arises from a purely deterministic system? To answer this question, we need to look at the system's dynamics in the three-dimensional phase space spanned by $X \equiv |a_0|$, $Y \equiv |a_q|$, $Z \equiv |a_{2q}|$, where $|a_k|$ is the amplitude of the Fourier mode with wave number k. The axes represent the energy of the steady mode, of the upper sideband, and of its second harmonic term, respectively. Notice that the wave-field amplitude $|\Psi(X,T)|$ displayed in Fig. 5(a) is the result of cooperative processes among the various Fourier components. When the motions were simple enough, as shown in Fig. 2, a two-dimensional phase space sufficed to describe the dynamics. Here the situation is more complex, requiring at least a three-dimensional phase space.

The phase-space trajectory corresponding to the evolution of Fig. 5(a) is drawn *schematically* in Fig. 5(b). A popular method of analyzing trajectories in a 3D phase space is to place a Poincaré surface of section across the direction of the flow as shown in Fig. 5(b), thereby generating a succession of points at which the trajectory traverses the plane. We chose the plane $X \approx \langle X \rangle$, where $\langle X \rangle$ is the time average of X, and obtained a sequence of points $[Y(t_n), Z(t_n)]$, where t_n is the time of the nth piercing of the section. Figure 5(b) presents the schematic drawing of the Poincaré surface section and also the set of crossing points on it. Notice that the points are aligned on a finite segment of line. With the notation $Y(t_n) \equiv Y_n$, Fig. 5(c) displays an enlarged picture of the cross section schematically drawn in Fig. 5(b), defining a two-dimensional discrete function, $(Y_{n+1}, Z_{n+1}) = P(Y_n, Z_n)$, $n = 1,2,\ldots$. Notice how the functional form of P is a truly finite line. The linear structure in the enlarged cross section implies that the trajectory in the 3D phase space contracts onto a surface in the form of the finite strip at the place where the section is inserted[8] in the schematic drawing of the trajectory in Fig. 5(b). At this point, we may further reduce the dynamics into a one-dimensional representation, plotting, say, Y_{n+1} versus Y_n, as drawn in Fig. 5(d). We find that the function G in $Y_{n+1} = G(Y_n)$ has a definite functional form, which is not derived theoretically but dug out of a mass of seemingly erratic data. As we shall see, this form affords an understanding of the whole dynamical history of $|\Psi(X,T)|$.

The reduction of very complex original systems to 1D maps, first reported by Lorenz (1963), has now become a basic tool of the laboratory (Roux *et al.*, 1983). An easier way to iterate a discrete one-dimensional function is the following: Starting from an initial Y_0, (1) move vertically to the graph of $G(Y)$, then (2) move horizontally to the 45° line of Fig. 5(d), and repeat steps 1 and 2. This process is illustrated by the arrows in Fig. 5(d).

The first piece of information from the functional form G is that its slope $|G'(Y)|$ has magnitude greater than one everywhere (except at its singular maximum), generating disorder as follows. Starting from any arbitrarily close pair of points, $Y_i, Y_i + \delta, \delta \ll 1$, an iteration gives the next points at Y_{i+1} and $Y_{i+1} + G'(Y_i)\delta$. These two points are now separated by a greater distance, since $|G'(Y)|$ is greater than unity, and the separation grows exponentially as the iteration progresses. All sequences are therefore unstable to small modifications. In other words, the dynamics now depends sensitively on initial conditions, a property foreign to periodic states.

The second message is that the function G has a single maximum, falling at both ends. This tells us that the trajectory stays in a finite region where it traverses the Poincaré section, despite the fact that any two nearby trajectories stretch out exponentially from each other. It thus implies global stability of the strip of contracted surface, whose cross section is shown in Fig. 5(c). Since any two nearby trajectories are stretched exponentially away from each other, stability of the strip requires that the trajectory be folded back onto itself somewhere on its way to the strip. No two trajectories, however, may intersect, since one initial condition (the point of intersection) would then give rise to different trajectories, contradicting the deterministic nature of the description.[9] This difficulty can be resolved only if the surface actually consists of infinite layers of sheets spaced arbitrarily close, so that any pair of almost crossing trajectories can stay on different layers. This is why we needed at least a 3D phase space. Conversely, the functional form G results from this multilayer structure. In conclusion, the one-dimensional dynamics based on G implies that the phase-space trajectory corresponding to the evolution shown in Fig. 5(a) represents a deterministic chaotic motion within a finite but very complicated structure.[10]

B. Cascading instabilities

The question follows naturally, what makes chaos and how? Let us start with the periodic state of Fig. 2(c) observed for a finite range of q, $0.985 \leq q \leq 1.0$. For a q value smaller than 0.985, the state becomes unstable, and a new state emerges. To trace the steps of this transition we shall once again examine how a 1D map arises from 3D trajectories. This time we choose the surface of section that contains the maxima of Y, that is, the surface

[8]Dissipative systems have the property that an evolving ensemble of states occupies a region of phase space whose volume decreases with time, eventually having zero volume. A surface has a zero volume.

[9]Except at a singular saddle point where two orbits can merge or emanate. Actually, the nondifferentiable edge shape maximum is the manifestation of the presence of the ridge potential.

[10]Such a region has been named a chaotic or strange attractor (Ruelle and Takens, 1971).

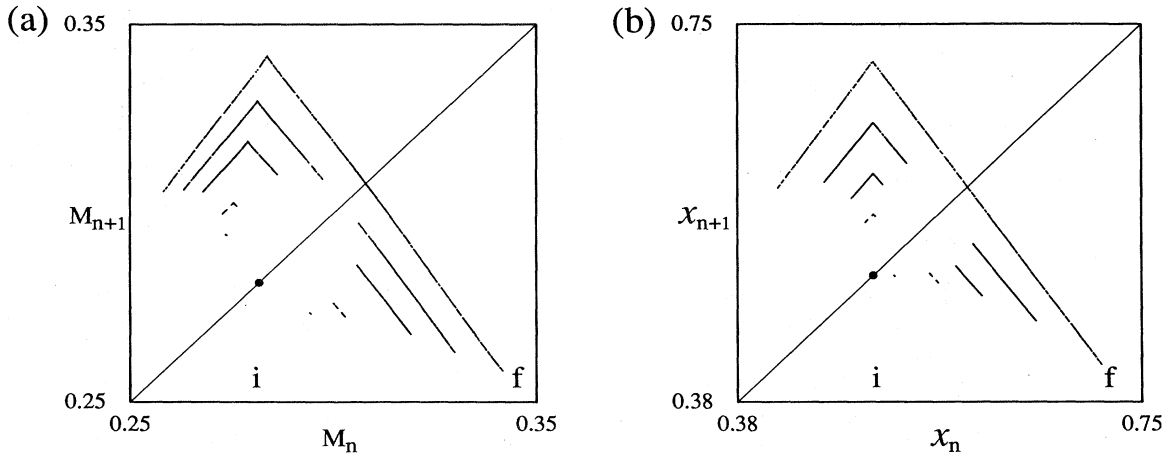

FIG. 6. Further aspects of evolution: (a) Approach to chaos as a function of the sideband wave number q. The point state denoted by i corresponds to the dynamic state shown in Fig. 2(c). At this reduced level, a "point" corresponds to a mode-locked state, which is destroyed as the value q changes further. The final state denoted by f corresponds to the fully chaotic state of Fig. 5(a); (b) the approach to chaos in the "tent map" $x_{n+1}=\lambda(1-2|x_n-1/2|)$, $n=1,2,3\ldots$. The parameter λ is varied from $1/2$ (state i) to $1/\sqrt{2}$ (state f). The transition patterns (a) and (b) are identical, implying that the tent map approach to chaos is universal.

defined by $dY/dt=0$.[11] From this plane of the section, we again obtain a one-dimensional dynamics by plotting M_{n+1} versus M_n, $n=1,2,\ldots$, where M_n is the nth maximum of Y.

For the periodic state of Fig. 2(c), all M_n coincide because the corresponding trajectory in the 3D phase space is a closed circular loop, which necessarily crosses the section at one point. This initial state is denoted by i in Fig. 6(a). The dynamics of Figs. 2(a) and 2(b) also display a single point. Thus, upon further reduction, the details of the dynamics of the three states of Fig. 2 are washed away, and only one common feature is left in the form of a point. Since the axes X, Y, Z represent the three predominant waves, their common dynamic feature lies apparently in their *mode locking*.

This mode-locking status is destroyed when the value of the q is decreased. The state i splits into two points at the resolution of the figure, and then into a state represented by four small bands, further into a state with two bigger bands, and finally into one mapped into a single big band. The final state denoted by f corresponds to the chaotic state of Fig. 5(a). Notice that the final map structure is the same as that obtained from the $X\approx\langle X\rangle$ plane of Fig. 5(d). No single state is ever stable within the range $0.920\le q<0.985$, exhibiting cascading instabilities until it becomes fully chaotic at $q=0.92$.

C. Universal nature of the approach to chaos

Astonishingly, the transition thus described turns out to match that of the following algebraically defined "tent map" (Moon, 1993):

$$x_{n+1}=\lambda(1-2|x_n-1/2|)\ , \tag{12}$$

defined for $0\le x\le 1$, with a nonlinearity parameter λ. We plot the transition sequence of the tent map in Fig. 6(b) for comparison with Fig. 6(a). The parameter λ has been varied from $\lambda=1/2$ to $1/\sqrt{2}$. The initial state denoted by i corresponds to the parameter value $\lambda=1/2$. This state becomes unstable and changes continuously as λ increases, and the final state denoted by f corresponds to $\lambda=1/\sqrt{2}$. The transitions to chaos in both cases follow paths that are identical in all respects.

The system seems to forget all of a sudden its original equations, following the transition pattern dictated by the totally different tent map. This result seems to imply that the details specific to any system may be irrelevant at the onset of chaos. On this subject, one can give the following general argument. Given a physical system whose dynamics can reduce to one-dimensional dynamics in the form of $x_{n+1}=F(x_n)$, we know that F should have a maximum in order to confine its trajectories within a finite region of phase space. One may then classify the functional form F into two groups depending on the differentiability of the maximum. The smooth-maximum group may be represented by the quadratic map

$$x_{n+1}=4\lambda x_n(1-x_n)\ , \tag{13}$$

defined on $0\le x\le 1$ with a nonlinearity control parameter λ, $0\le\lambda\le 1$. This map exhibits the well-known period-doubling transition route to chaos (May, 1976; Feigenbaum, 1978). In the past two decades, it has been observed repeatedly that any system reducible to a smooth functional form invariably follows the approach to chaos of the quadratic map. This property has been termed "universality" (Feigenbaum, 1978). The present

[11]This approach has been popular for selecting a surface section. The section chosen in Fig. 5(b) was chosen for purposes of simple illustration.

study appears to broaden the class of universality, raising the possibility that the tent map approach to chaos is universal.

V. CONCLUSIONS

We have discussed the phenomenon of spatial symmetry breaking, of the emergence of order and disorder when a dynamical system enters into a nonlinear stage. In such a critical regime, the original set of nonlinear equations reduces to a much simpler form. In the reduction processes, many properties specific to individual systems are bound to be eliminated. The final form therefore is believed to hold a common dynamical essence shared by diverse physical systems. For this reason, Eq. (2) becomes important in optics, fluid dynamics, plasma physics, and chemical reaction-diffusion. Specifically, a simple change in the interpretation of Ψ might permit the description of diverse phenomena such as optical pulses and physicochemical patterns.

The concentration variables expressed as $c_i = c_i(x,t,\xi,\tau)$ signify the presence of two different scales, in space and in time. This fact cannot be overemphasized, because the presence of two different scales can be a source of great confusion epistemologically. Furthermore, the span of ranges is very broad. For example, the small and large time scales of a pulse in an optical fiber are, respectively, of the order of 10^{-2} and 1 ps, while in chemical reaction-diffusion they are of the order of a minute and of several hours.

The "chemical clock" behavior, whether bulk or local, the Fermi-Pasta-Ulam recurrence phenomenon, and the spatial oscillations of molecular populations are all described within the context of our universal function Ψ. In addition, they are not independent of each other but globally coupled through the double-well potential structure.

The ridge potential of the double-well potential in a continuous medium is the source of the long-range self-organization. In dissipative systems, however, it can also become a source of disorder at the onset to chaos. It thus exhibits a duality that seems contradictory by underlying both long-range order and chaos.

Finally, we point out that the dynamics described by Ψ further reduces to one-dimensional dynamics at the onset to chaos. For the reasons stated above, this reduced form is believed to hold the dynamic essence of diverse systems. In fact, it shows that the approach to chaos follows all of a sudden a common channel set by the simple tent map, implying the universality of the approach to chaos.

ACKNOWLEDGMENTS

The author wishes to thank M. V. Goldman, P. Huerre, J. I. Kim, M. J. Kim, L. G. Redekopp, and J. Toomre, for many valuable discussions. He would like especially to thank U. Fano, S. Freedman, and D. Stein, for both criticism and encouragement. This work was supported by the Korea Science and Engineering Foundation.

REFERENCES

Belousov, B. P., in *Sb. Ref. Radiat. Med.* (Collection of abstracts on radiation medicine) (Medgiz Moscow, 1959), p. 145.
Benjamin, T. B., and J. E. Feir, 1967, J. Fluid Mech. **27**, 417.
Bretherton, C. S., and E. A. Spiegel, 1983, Phys. Lett. A **96**, 152.
Busse, H. G., 1969, J. Phys. Chem. **73**, 750.
Epstein, I. R., 1983, Physica D **7**, 47.
Feigenbaum, M. J., 1978, J. Stat. Phys. **19**, 25.
Fermi, E., J. Pasta, and S. Ulam, 1965, in *Collected Papers of Enrico Fermi,* Vol. 2, edited by E. Segré (University of Chicago), p. 978.
Hasegawa, A., and K. Tai, 1989, Opt. Lett. **14**, 512.
Kelly, P. L., 1965, Phys. Rev. Lett. **15**, 1005.
Kuramoto, Y., and T. Tsuzuki, 1975, Prog. Theor. Phys. **54**, 687.
Lifshitz, E. M., and L. P. Pitaevskii, 1980, *Statistical Physics,* Vol. 2 (Pergamon, Oxford).
Lorenz, E. N., 1963, J. Atmos. Sci. **20**, 130.
May, R. M., 1976, Nature **261**, 459.
Moon, H. T., 1990, Phys. Rev. Lett. **64**, 412.
Moon, H. T., 1991, Phys. Fluids A **3**, 2709.
Moon, H. T., 1993, Phys. Rev. E **47**, 772.
Newell, A. C., 1974, Lect. Appl. Math. **15**, 157.
Newell, A. C., and J. A. Whitehead, 1969, J. Fluid Mech. **38**, 279.
Nicolis, G., and I. Prigogine, 1977, *Self-Organization in Nonequilibrium Systems* (Wiley, New York).
Roux, J.-C., R. H. Simoyi, and H. L. Swinney, 1983, Physica D **8**, 257.
Ruelle, D., and F. Takens, 1971, Commun. Math. Phys. **20**, 167.
Tai, K., A. Hasegawa, and A. Tomita, 1986, Phys. Rev. Lett., **59**, 135.
Yuen, H. C., and W. E. Ferguson, 1978, Phys. Fluids **21**, 1275.