by Brosl Hasslacher

The invention of a totally discrete model for natural phenomena was made by Ulam and von Neumann in the early fifties and was developed to the extent possible at the time. A few years earlier von Neumann had designed the architecture for the first serial digital computers containing stored programs and capable of making internal decisions. These machines are built of electronic logic devices that understand only packets of binary bits. Hierarchies of stored translators arrange them into virtual devices that can do ordinary or radix arithmetic at high speed. By transcribing continuum equations into discrete form, using finite difference techniques and their variants, serial digital computers can solve complex mathematical systems such as partial differential equations. Since most physical systems with large numbers of degrees of freedom can be described by such equations, serial digital machines equipped with large memories have become the standard way to simulate such phenomena.

PART I
BACKGROUND FOR LATTICE GAS AUTOMATA

The lattice gas automaton is an approach to computing fluid dynamics that is still in its infancy. In this three-part article one of the inventors of the model presents its theoretical foundations and its promise as a general approach to solving partial differential equations and to parallel computing. Readers less theoretically inclined might begin by reading "Calculations Using Lattice Gas Techniques" at the end of Part II. This sidebar offers a summary of the model's advantages and limitations and a graphic display of two- and three-dimensional lattice gas simulations.

As the architecture of serial machines developed, it became clear to both Ulam and von Neumann that such machines were not the most natural or powerful way to solve many problems. They were especially influenced by biological examples. Biological systems appear to perform computational tasks using methods that avoid both arithmetical operations and discrete approximations to continuous systems.

Though motivated by the complex information processing of biological systems, Ulam and von Neumann did not study how such systems actually solve tasks. Biological processes have been operating in hostile environments for a long time, finding the most efficient and often devious way to do something, a

way that is also resistant to disturbance by noise. The crucial principles of their operation are hidden by the evolutionary process. Instead, von Neumann chose the task of simulating on a computer the least complex discrete system capable of self-reproduction. It was Ulam who suggested an abstract setting for this problem and many other totally discrete models, namely, the idea of cellular spaces. The reasoning went roughly like this.

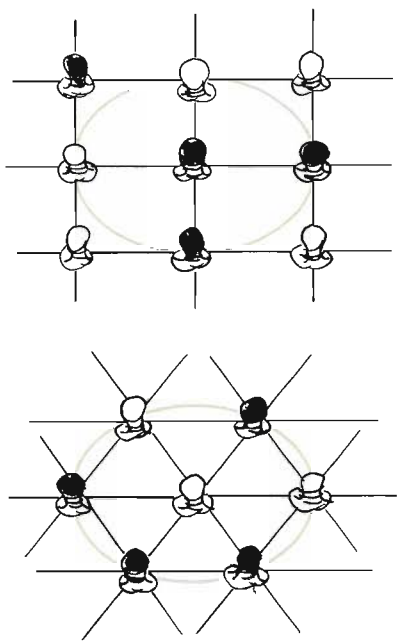
The question is simple: Find a minimal logic structure and devise a dynamics for it that is powerful enough to simulate complex systems. Break this up into a series of sharper and more elementary pictures. We begin by setting up a collection of very simple finite-state machines with, for simplicity, binary values. Connect them so that given a state for each of them, the next state of each machine depends only on its immediate environment. In other words, the state of any machine will depend only on the states of machines in some small neighborhood around it. This builds in the constraint that we only want to consider local dynamics.

We will need rules to define how states combine in a neighborhood to uniquely fix the state of every machine, but these can be quite simple. The natural space on which to put all this is a lattice, with elementary, few-bit, finite-state machines placed at the vertices. The rules for updating this array of small machines can be done concurrently in one clock step, that is, in parallel.

One can imagine such an abstract machine in operation by thinking of a fishnet made of wires. The fishnet has some regular connection geometry, and there are lights at the nodes of the net. Each light can be on or off. Draw a disk around each node of the fishnet, and let it have a 1-node radius. On a square net there are four lights on the edge of each disk, on a triangular net six lights (Fig. 1). The next state of the light at the center of the disk depends on the state of the lights on the edge of the disk and on nothing else. Imagine all the disks in the fishnet asking their neighbors for their state at the same time and switching states according to a definite rule. At the next tick of an abstract clock, the pattern of lights on the fishnet would in general look different. This is what Ulam and von Neumann called a nearest-neighbor-connected cellular space. It is the simplest case of a parallel computing space. You can also see that it can be imaged directly in hardware, so it is also the architecture for a physical parallel computing machine.

CELLULAR SPACES

Fig. 1. Two examples of fishnets made of wires with lights at the nodes. The lights are either on or off. In each example a disk with a radius of 1 node is drawn around one of the lights. The next state of the light at the center depends on the states of the lights on the edge of the disk and on nothing else. Thus these are examples of nearest-neighbor-connected cellular spaces.



We have not shown that such a device can compute. At worst, it is an elaborate light display. Whether or not such a cellular space can compute depends on the definition of computation. The short answer is that special cases of fishnets are provably universal computers in the standard Turing machine sense; that is, they can simulate the architecture of any other sequential machine.

But there are other interpretations of computation that lie closer to the idea of simulation. For any given mathematical situation, we want to find the minimum cellular space that can do a simulation of it: At what degree of complexity does repeated iteration of the space, on which are coded both data and a solution algorithm, possess the power to come close to the solution of a complex problem? This depends on the complexity or degrees of freedom present in the problem.

An extreme case of complexity is physical systems with many degrees of freedom. These systems are ordinarily described by field theories in a continuum for which the equations of motion are highly nonlinear partial differential equations. Fluid dynamics is an example, and we will use it as a theoretical paradigm for many "large" physical systems. Because of the high degree of nonlinearity, analytic solutions to the field equations for such systems are known only in special cases. The standard way to study such models is either to perform experiments or simulate them on computers of the usual digital type.

Suppose a cellular space existed that evolved to a solution of a fluid system with given boundary conditions. Suppose also that we ask for the simplest possible such space that captured at least the qualitative and topological aspects of a solution. Later, one can worry about spaces that agree quantitatively with ordinary simulations. The problem is threefold: Find the least complex set of rules for updating the space; the simplest geometry for a neighborhood; and a method of analysis for the collective modes and time evolution of such a system.

At first sight, modeling the dynamics of large systems by cellular spaces seems far

too difficult to attempt. The general problem of a so-called "inverse compiler"—given a partial differential system, find the rules and interconnection geometry that give a solution—would probably use up a non-polynomial function of computing resources and so be impractical if not impossible. Nevertheless cellular spaces have been actively studied in recent years. Their modern name is cellular automata, and specific instances of them have simulated interesting nonlinear systems. But until recently there was no example of a cellular automaton that simulated a large physical system, even in a rough, qualitative way.

Knowing that special cases of cellular automata are capable of arbitrarily complex behavior is encouraging, but not very useful to a physicist. The important phenomenon in large physical systems is not arbitrarily complex behavior, but the collective motion that develops as the system evolves, typically with a characteristic size of many elementary length scales. The problem is to simulate such phenomena and, by using simulations, to try to understand the origins of collective behavior from as many points of view as possible. Fluid dynamics is filled with examples of collective behavior—shocks, instabilities, vortices, vortex streets, vortex sheets, turbulence, to list a few. Any deterministic cellular-automaton model that attempts to describe non-equilibrium fluid dynamics must contain in it an iterative mechanism for developing collective motion. Knowing this and using some very basic physics, we will construct a cellular automaton with the appropriate geometry and updating rules for fluid behavior. It will also be the simplest such model. The methods we use to do this are very conservative from the viewpoint of recent work on cellular automata, but rather drastic compared to the approaches of standard mathematical physics. Presently there is a large gap between these two viewpoints. The simulation of fluid dynamics by cellular automata shows that there are other complementary and powerful ways to model phenomena that would normally be the exclusive domain of partial differential equations.

The Example of Fluid Dynamics

Fluid dynamics is an especially good large system for a cellular automaton formulation because there are two rich and complementary ways to picture fluid motion. The kinetic picture (many simple atomic elements colliding rapidly with simple interactions) coincides with our intuitive picture of dynamics on a cellular space. Later we will exploit this analogy to construct a discrete model.

The other and older way of approaching flow phenomena is through the partial differential equations that describe collective motions in dissipative fluids—the Navier-Stokes equations. These can be derived without any reference to an underlying atomic picture. The derivation relies on the idea of the continuum; it is simpler to grasp than the kinetic picture and mathematically cleaner. Because the continuum argument leads to the correct functional form of the Navier-Stokes equations, we spend some time describing why it works. The continuum view of fluids will be called "coming down from above," and the microphysical view "coming up from below" (Fig. 2). In the intersection of these two very different descriptions, we can trap the essential elements of a cellular-automaton model that leads to the Navier-Stokes equations. Through this review we wish to show that cellular automaton models are a natural and evolutionary idea and not an invention come upon by accident.

Coming down from Above—The Continuum Description

The notion of a smooth flow of some quantity arises naturally from a continuum description. A flow has physical conservation laws built-in, at least conservation of mass and momentum. With a few additional remarks one can include conservation of energy. The basic strategy for deriving the Euler and Navier-Stokes equations of fluid dynamics is to imbed these conservation laws into statements about special cases of the generalized Stokes theorem. We use the usual Gauss and Stokes theorems, depending on dimension, and apply them to small surfaces and volumes that are still large enough to ignore an underlying microworld. The equations of fluid dynamics are derived with no reference to a ball-bearing picture of an underlying atomic world, but only with a serene reliance on the idea of a smooth flow in a continuum with some of Newton's laws added to connect to the observed world. As a model (for it is not a theory), the Navier-Stokes equations are a good example of how concepts derived from the intuition of daily experience can be remarkably successful in building effective phenomenological models of very complex phenomena. It is useful to go through the continuum derivation of the Euler and Navier-Stokes equations presented in "The Continuum Argument" for several reasons: First, the reasoning is short and clear; second, the concepts introduced such as the momentum flux tensor, will appear pervasively when we pass to discrete theories of fluids; third, we learn how few ingredients are really necessary to build a fluid model and so mark out that which is essential—the role of conservation laws.

It is clear from its derivation that the Euler equation describing inviscid flows is essentially a geometrical equation. The extension to the full Navier-Stokes equations, for flows with dissipation, contains only a minimal reference to an underlying fluid microphysics, through the stress-rate of strain relation in the momentum stress tensor. So we see that continuum reasoning alone leads to nonlinear partial differ-

ential equations for large-scale physical observables that are a phenomenological description of fluid flow. This description is experimentally quite accurate but theoretically incomplete. The coupling constants that determine the strength of the nonlinear terms—that is, the transport coefficients such as viscosity—have a direct physical interpretation in a microworld picture. In the continuum approach however, these must be measured and put in as data from the outside world. If we do not use some microscopic model for the fluid, the transport coefficients cannot be derived from first principles.

Solution Techniques—The Creation of a Microworld. The Navier-Stokes equations are highly nonlinear; this is prototypical of field-theoretical descriptions of large physical systems. The nonlinearity allows analytic solutions only for special cases and, in general, forces one to solve the system by approximation techniques. Invariably these are some form of perturbation methods in whatever small parameters can be devised. Since there is no systematic way of applying perturbation theory to highly nonlinear partial differential systems, the analysis of the Navier-Stokes equations has been, and still remains, a patchwork of ingenious techniques that are designed to cover special parameter regimes and limited geometries.

After an approximation method is chosen, the next step toward a solution is to discretize the approximate equations in a form suitable for use on a digital computer. This discretization is equivalent to introducing an artificial microworld. Its particular form is fixed by mathematical considerations of elegance and efficiency applied to simple arithmetic operations and the particular architecture of available machines. So, even if we adopt the view that the molecular kinetics of a fluid is unimportant for describing the general features of many fluid phenomena, we are nevertheless forced to describe the system by a microworld with a particular microkinetics. The idea of a partial differential equation as a physical model is tied directly to finding an analytic solution and is not particularly suited to machine computation. In a sense, the geometrically motivated continuum picture is only a clever and convenient way of encoding conservation laws into spaces with which we are comfortable.

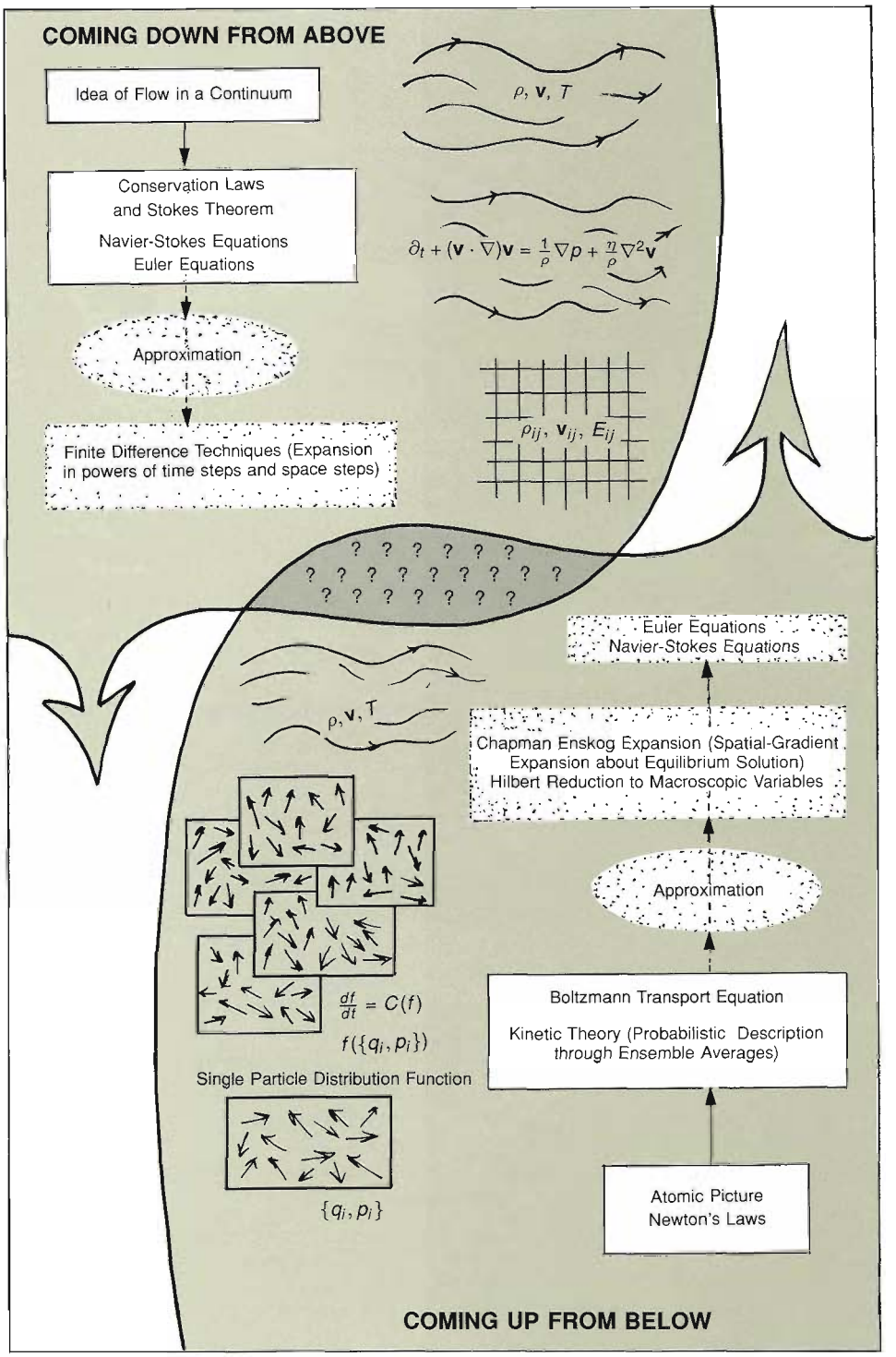


Fig. 2. Both the continuum view of fluids and the atomic picture lead to the Navier-Stokes equations but not without approximations (dashed lines). The text emphasizes how cellular-automaton models embody the essentials of both points of view.

Coming up from Below—The Kinetic Theory Description

Kinetic theory models a fluid by using an atomic picture and imposing Newtonian mechanics on the motions of the atoms. Atomic interactions are controlled by potentials, and the number of atomic elements is assumed to be very large. This attempt at fluid realism has an imme-

continued on page 181

THE CONTINUUM ARGUMENT

Let $\mathbf{v}(\mathbf{x},t)$ be a vector-valued field referred to a fixed origin in space, which we identify with the velocity of a "macroscopic" fluid cell. The cell is not small enough to notice a particle structure for the fluid, but it is small enough to be treated as a mathematical point and still agree with physics.

To derive the properties of a flow defined by the vector field, one now invokes the generalized Stokes theorem:

$$\oint_{\partial\Sigma} \mathbf{A} = \oint_{\Sigma} d\mathbf{A},$$

where Σ is a generalized surface or volume, $\partial\Sigma$ the boundary of Σ, $\mathbf{A}$ an n-differential form and $d\mathbf{A}$ an $(n+1)$-differential form. This very general theorem has two familiar forms: one is the classical Stokes theorem from one to two dimensions,

$$\oint_{\partial\Sigma} \mathbf{A}\cdot\mathbf{d}\ell = \oint_{\Sigma} \nabla\times\mathbf{A}\cdot d\mathbf{S},$$

and the other is the Gauss law from two to three dimensions,

$$\oint_{\partial\Sigma} \mathbf{A}\cdot d\mathbf{S} = \int_{\Sigma} (\nabla\cdot\mathbf{A})dV,$$

where ℓ is a curve, $\mathbf{S}$ is a surface, and V is a volume in three-dimensional Euclidean space $\mathbf{R}^3$.

Conservation Laws and Euler's Equation. First, we deal with the idea of continuity, or conservation of flow. If ρ is the density, or mass per unit volume, then the mass of the fluid in volume V (that is, Σ), is equal to $\int_{\Sigma}\rho dV$. A two-dimensional surface in $\mathbf{R}^3$ has an outward normal vector $\mathbf{n}$ which is defined to be positive. The total mass of fluid flowing out of a volume Σ can be written as

$$\oint_{\partial\Sigma} \rho\mathbf{v}\cdot d\mathbf{S} = \int_{\partial\Sigma} \rho\mathbf{v}\cdot\mathbf{n}dS.$$

Continuity of the flow implies a balance between the flow through the surface and the loss of fluid from the volume. That is, the decrease in mass in the volume must equal the outflow of fluid mass through the surface of the volume, which implies by the Gauss law that

$$\oint_{\partial\Sigma} \rho\mathbf{v}\cdot d\mathbf{S} = -\partial_t \int_{\Sigma} \rho dV = \int_{\Sigma} \nabla\cdot(\rho\mathbf{v})dV.$$

This gives the first evolution equation for a fluid, the continuity, or mass-conservation, equation:

$$\partial_t\rho + \nabla\cdot(\rho\mathbf{v}) = 0. \tag{1}$$

Now we introduce the idea of pressure p as the force exerted by the fluid on a unit surface area of an enclosed volume and use Newton's second law, $\mathbf{F} = ma$. The total force acting on a volume of fluid due to the remainder of the fluid is given by $-\oint_{\partial\Sigma} p d\mathbf{S}$. Using Stokes theorem we can write

$$-\oint_{\partial\Sigma} p d\mathbf{S} = -\oint_{\Sigma} \nabla p dV.$$

The translation of $\mathbf{F} = m\mathbf{a}$ to a continuous medium is

$$-\nabla p = \rho d\mathbf{v}/dt,$$

where $d\mathbf{v}/dt$ is a total derivative. The chain rule on $d\mathbf{v}(\mathbf{x},t)/dt$ gives

$$\rho d\mathbf{v}/dt = \rho\{\partial_t\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{v}\}$$

Substituting this result into the equation for $-\nabla p$ yields Euler's equation for an ideal, dissipation-free fluid:

$$\partial_t\mathbf{v} = -(\mathbf{v}\cdot\nabla)\mathbf{v} - \frac{1}{\rho}\nabla p. \tag{2}$$

THE CONTINUUM ARGUMENT (continued)

One can generalize Euler's equation to a form more useful for a dissipative fluid. For this we look at the flux of momentum through a fluid volume. The momentum of fluid passing through an element dV is $\rho\mathbf{v}$, and its time rate of change expressed in components is

$$\partial_t(\rho v_i) = (\partial_t \rho)v_i + \rho(\partial_t v_i).$$

We can rewrite $\partial_t \rho$ and $\partial_t v_i$ as spatial derivatives by using Eqs. 1 and 2. Then

$$\partial_t(\rho v_i) = -\partial_k \Pi_{ik}, \tag{3}$$

where the momentum flux tensor $\Pi_{ik} \equiv p\delta_{ik} + \rho v_i v_k$.

The meaning of the momentum flux tensor can be seen immediately by integrating Eq. 3 and applying Stokes theorem.

$$\partial_t \int_\Sigma \rho v_i d\Sigma = -\int_\Sigma \partial_k \Pi_{ik} d\Sigma = -\oint_{\partial\Sigma} \Pi_{ik} n_k dS.$$

So

$$\partial_t \int_\Sigma \rho v_i dV = -\oint_{\partial\Sigma} \Pi_{ik} n_k dS,.$$

where the left-hand side is the rate of change of the ith component of momentum ρv_i in the volume and $\Pi_{ik} n_k d\Sigma$ is the ith component of momentum flowing through dS. Therefore, Π_{ik} is the ith component of momentum flowing in the kth direction. This is more easily seen by writing

$$\Pi_{ik} n_k = p\delta_{ik} n_k + \rho v_i v_k n_k = p\mathbf{n} + \rho\mathbf{v}(\mathbf{v}\cdot\mathbf{n}).$$

Equations 1, 2, and 3 are the basic formalism for classical Newtonian ideal fluids (fluids with no dissipation) and are also true for flows in general.

Classical Dissipative Fluids—The Navier-Stokes Equations. The general Euler's equation is $\partial_t(\rho v_i) = -\partial_k \Pi_{ik}$, where Π_{ik} is now the momentum stress tensor. The form of this tensor changes if the fluid is dissipative, for example, if viscous forces convert the energy in the flow into heat. Traditionally, Π_{ik} is modified in the following way. Take $\Pi_{ik} = p\delta_{ik} + \rho v_i v_k$ and introduce an unknown tensor σ'_{ik} that describes the effects of viscious stress. Then rewrite the momentum stress tensor as

$$\Pi_{ik} = p\delta_{ik} + \rho v_i v_k - \rho\sigma'_{ik} \equiv \sigma_{ik} + \rho v_i v_k,$$

where $\sigma_{ik} = p\delta_{ik} - \rho\sigma'_{ik}$ is called the stress tensor and σ'_{ik} the viscosity stress tensor.

The form of σ'_{ik} can be deduced on general grounds. First we assume that the gradient of the velocity changes slowly so σ_{ik} is linear in $\partial_k v_i$. Moreover, σ'_{ik} is zero for $\mathbf{v} = 0$, and under rotation it must vanish since uniform rotation produces no overall transport of momentum. The unique form that has these properties is

$$\sigma'_{ik} = a(\partial_k v_i + \partial_i v_k) + b\delta_{ik}\partial_j v_j,$$

where a and b are unknown coefficients. It is usually written in the form

$$\sigma'_{ik} = \nu(\partial_k v_i - \partial_i v_k - 2/3\delta_{ik}\partial_j v_j) + \zeta\delta_{ik}\partial_j v_j,$$

where ν is the kinematic shear viscosity and ζ is the kinematic bulk viscosity.

For an incompressible fluid (the density is constant so $\rho = \rho_0$) this tensor simplifies, and Euler's equation goes over to the incompressible Navier-Stokes equations:

$$\partial_t\mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\frac{1}{\rho}\nabla p + \mathbf{v}\nabla^2\mathbf{v} \quad \text{and} \quad \nabla\cdot\mathbf{v} = 0.$$

In tensor notation we have

$$\partial_t v_i + (v_j\partial_j)v_i = -\frac{1}{\rho}\partial_i p + \nu\frac{\partial^2}{\partial_k\partial_k}v_i \quad \text{and} \quad \partial_k v_k = 0.$$

In Part II we end the theoretical discussion of the lattice gas by giving the incompressible limit of the lattice gas Navier-Stokes equations. ■

continued from page 178

diate difficulty. We are unable to specify completely the initial state of the system or to follow its microdynamics. It follows that we cannot use a microdynamics that is this detailed. The obvious strategy is to make a smoothened model that reduces the number of degrees of freedom in the system to just a few. This reduction assumes maximum ignorance of the details of the system below some time and distance scale and replaces exact data on events by probabilistic outcomes. Measurements are assumed to be average values of quantities over large ensembles of representative systems. The assumption is that after a sufficiently long time these average observables are a close description of the fluid.

This approach seems very familiar and obvious from elementary courses in statistical mechanics. But it is unclear how to go from a statistical-mechanical description of an atomic system to the prediction of the details of collective motions that come from the evolution of that system. Fidelity to the atomic picture brings with it considerable mathematical difficulties. As we will see below and in "The Hilbert Contraction," the success of the derivation of the Navier-Stokes equations from the kinetic theory picture—that one derives the Navier-Stokes equations with the correct coefficients and not some other macrodynamics—is justified after the fact.

Kinetic Theory and the Boltzmann Transport Equation. Complete information on the statistical description of a fluid or gas at, or near, thermal equilibrium is assumed to be contained in the one-particle phase-space distribution function $f(t,\mathbf{r},\Gamma)$ for the atomic constituents of the system. The variables t and $\mathbf{r}$ are the time and space coordinates of the atoms and Γ stands for all other phase-space coordinates (for example, momenta). In this rapid overview of kinetic transport theory, we will not dwell on the many and difficult questions raised by this description but keep to a level of precision consistent with a general understanding of the basic ideas.

The distribution function f is basically a weighting function that is used to define the mean values of physical observables. The relation

$$N(t,\mathbf{r}) \equiv \int f(t,\mathbf{r},\Gamma)\,d\Gamma \qquad (1)$$

defines the density function $N(t,\mathbf{r})$ for the particles in the system over all space. Therefore NdV is the mean number of particles in the volume dV. Here dV is a physical volume $\propto L^3$ whose characteristic length L is much larger than l_m, the mean free path of a particle, and much smaller than L_g, some global length, such as the edge of a container for the whole gas. Thus $l_m \ll L \ll L_g$.

The basic equation of kinetic theory is the evolution equation for $f(t,\mathbf{r},\Gamma)$ in the presence of gas collisions. Imagine first that the system has no collisions. Conservation of phase-space volumes, or Liouville's theorem, tells us that

$$\frac{df}{dt} = 0, \qquad (2)$$

where d/dt is a total derivative. In an isolated system with no external fields, we can expand the total derivative as

$$\frac{df}{dt} = \partial_t f + \mathbf{v}\cdot\nabla f \equiv \partial_t f + \mathrm{v}_i \partial_i f. \qquad (3)$$

(We use the convention that repeated indices are summed over.) Equation 3 defines the free-streaming operator, which represents the local change in f per unit time caused by the independent motion of particles alone.

Now imagine a simple isolated gas with collisions. If $C(f)$ is a function that models the rate of change of the distribution function f caused by collisions, then $C(f)\,dVd\Gamma$ is the rate of change per unit time of the number of molecules in the phase-space volume element $dVd\Gamma$. The Liouville statement now is modified to become the transport equation:

$$\frac{df}{dt} = C(f), \qquad (4)$$

where $C(f)$ is in general a highly nonlinear function of f.

Boltzmann first gave a simple approximation for the collision operator, which can be thought of as a gain-minus-loss $(\mathcal{G}-\mathcal{L})$ operator. A straightforward physical argument defining its general structure is presented below and is due to Landau.

The Boltzmann Form of the Collision Term. Let the particles in a two-body collision process have incoming distribution functions g_1 and g_2 and outgoing distribution functions $\bar{g}_1$ and $\bar{g}_2$. Fixing attention on particle 1, assume that before colliding it occupies a phase-space region $d\Gamma_1$, and after collision it occupies $d\bar{\Gamma}_1$; similarly, particle 2 occupies $d\Gamma_2$ before colliding and $d\bar{\Gamma}_2$ afterwards. If particle 1 undergoes a collision, $d\bar{\Gamma}_1$ will not in general be in $d\Gamma_1$, and particle 1 is said to be lost from $d\Gamma_1$. From these considerations we can compute the functional structure of the general loss term for a binary collision.

The probability of loss will be proportional to the product of four terms: (1) the number of particles of type 1 already in the volume, namely g_1; (2) the number of type-2 particles that enter the volume from some phase-space range $d\Gamma_2$, namely, $g_2 d\Gamma_2$; (3) the total volume of allowed outgoing phase space, $d\bar{\Gamma}_1 d\bar{\Gamma}_2$; and finally (4) a probability for the collision process $P_g\{\Gamma\}$. Now we sum over all possible allowed volumes of phase space. So the total number of losses $\mathcal{L}$ in the volume dV and from $d\Gamma$ due to binary collision processes is

$$\mathcal{L} = dV\,d\Gamma \int P_g\{\Gamma\} g_1 g_2\, d\Gamma_2 d\bar{\Gamma}_1 d\bar{\Gamma}_2.$$

Similarly, particle gain into the phase

space volume $d\Gamma$ can only come from reversed channel processes $\bar{g}_1, \bar{g}_2 \rightarrow g_1, g_2$, with fixed Γ_1, and summed over all of $\bar{\Gamma}_1$, $\bar{\Gamma}_2$, and Γ_2, so

$$\mathcal{G} = dV\, d\Gamma \int P_g\{\Gamma\}\bar{g}_1\bar{g}_2\, d\Gamma_2 d\bar{\Gamma}_1 d\bar{\Gamma}_2.$$

The Boltzmann form for $C(f)$ is the net flow into the region, which is $\mathcal{G} - \mathcal{L}$. Using this form, we get the Boltzmann transport equation, a highly nonlinear integro-differential equation:

$$\frac{df}{dt} = \partial_t f + \mathrm{v}_i \partial_i f = \mathcal{G} - \mathcal{L}. \qquad (5)$$

In Part II we will use the same reasoning to construct the Boltzmann equation for the discrete lattice gas. The explicit form of the lattice gas collision operator is much simpler than in standard kinetic models.

Note that the Boltzmann form for the $(\mathcal{G} - \mathcal{L})$ collision term implicitly assumes only two-body collisions. It also assumes the collisions are pairwise statistically independent events occurring at a single point with detailed, or at most semi-detailed, balance symmetry for collision probabilities.

Solutions to the Boltzmann Transport Equation. Even though the Boltzmann equation is intractable in general, by using entropy arguments (Boltzmann's H theorem), the following can be stated about possible functional forms for f, the one-particle distribution function. If the system is uniform in space, any form for f will relax monotonically to the global Maxwell-Boltzmann form:

$$f_{\text{global}} \sim \rho e^{-E(\rho,\mathrm{v})/T},$$

in which the macroscopic variables ρ, v, and T (density, macrovelocity, and temperature) are independent of position, or global. In the non-equilibrium case, with a soft space dependence, any distribution function will relax monotonically in velocity space to a *local* Maxwell-Boltzmann form. This means that ρ, v, and T will depend on space as well as time. These local distribution functions are solutions to the Boltzmann transport equation. For the non-uniform case, one gets a picture of the full solution as an ensemble of local Maxwell-Boltzmann distributions covering the description space of the fluid, with some gluing conditions providing the consistency of the patching.

RANGE OF THE BOLTZMANN TRANSPORT EQUATION

The rigorous range of physical parameters in which the Boltzmann transport equation is mathematically meaningful is

$$N \rightarrow \infty \text{ such that } mN \rightarrow K$$

$$N \rightarrow \infty,\ \sigma \rightarrow 0 \text{ such that } (N\sigma^2) \rightarrow l_m^{-1}$$

where N is the number of particles,
m is the mass of each particle,
σ is the range of the force or the effective interaction ball,
l_m is the mean free path,
K is a constant.

These conditions imply a dilute gas, binary collisions, and slowly varying spatial dependence (that is, slow space gradients). As an additional axiom we require that there be no long-range forces in the sense of photon excitations, etc.

Recovering Macrodynamics–The Euler Equations. If we assume a simple fluid and neglect all dissipative processes (viscosity, heat transfer, etc.), we can quickly derive the Euler equations (presented in "The Continuum Argument") from the Boltzmann transport equation. But first we need the notion of average quantities and some observations about collisions in a dissipation-free system.

As before, let $\rho(t,r) = \int f(t,\mathbf{r},\Gamma)\, d\Gamma$ be the density field of the gas. Then a mean gas velocity $\mathbf{v} = \frac{1}{\rho}\int \mathbf{v}' f(t,\mathbf{r},\Gamma)\, d\Gamma$, where $\mathbf{v}'$ is a microvelocity. We will use $\mathbf{v}$ as a macroscopic variable that characterizes cells whose length L in any direction is much, much greater than the mean free path in the gas, l_m; that is, $L \gg l_m$.

Since, by assumption, collisions preserve conservation laws exactly, the moments of $C(f)$, in particular the integrals $\int C(f)\, d\Gamma$ and $\int \mathbf{v}\, C(f)\, d\Gamma$, are equal to zero (similarly for any conserved quantity). We use this fact by integrating the Boltzmann equation in two ways: $\int$ (B.E.) $d\Gamma$ and $\int \mathbf{v}$(B.E.) $d\Gamma$ (where B.E. stands for the Boltzmann equation). The first integral gives the continuity equation:

$$\partial_t \rho + \partial_i(\rho \mathrm{v}_i) = 0. \qquad (6)$$

The second integral gives the momentum tensor equation:

$$\partial_t(\rho \mathrm{v}_i) + \partial_k \Pi_{ik} = 0, \qquad (7)$$

where the momentum flux tensor Π_{ik} is given by

$$\Pi_{ik} \equiv \int \mathrm{v}_i \mathrm{v}_k f\, d\Gamma.$$

In order to derive the Euler equation for ideal gases with the usual form for

the momentum flux tensor, we need to assume that each region in the gas has a local Maxwell-Boltzmann distribution. With this assumption one can show that the momentum flux tensor in Eq. 7 has the following form:

$$\Pi_{ik} = \rho v_i v_k + \delta_{ik} p ,$$

where p is the pressure. This form of Π_{ik} gives the same Euler equation that we found by general continuum arguments. (We will see in Part II that the form of Π_{ik} for the totally discrete fluid is not so simple but depends upon the geometry of the underlying lattice. Again by assuming a form for the local distribution function (the appropriate form will turn out to be Fermi-Dirac rather than Boltzmann), Π_{ik} will reduce to a form that gives the lattice Euler equation.)

Recovering the Navier-Stokes Equation. The derivation of the Navier-Stokes equation from the kinetic theory picture is more involved and requires us to face the full Boltzmann equation. Hilbert accomplished this through a beautiful argument that relies on a spatial-gradient perturbation expansion around some single-particle distribution function f_L assumed to be given at t_0. In "The Hilbert Contraction" we discuss the main outline of his argument emphasizing the assumptions involved and their limitations. Here we will summarize his argument. Hilbert was able to show that the evolution of f for times $t > t_0$ is given in terms of its initial data at t_0 by the first three moments of f, namely the familiar macroscopic variables ρ (density), $\mathbf{v}$ (mean velocity), and T (temperature). In other words, he was able to contract this many-degree-of-freedom system down to a low-dimensional descriptive space whose variables are the same as those used in the usual hydrodynamical description. The beauty of Hilbert's proof is that it is constructive. It explicitly displays a recursive closed tower of constraint relations on the moments of f that come directly from the

THE HILBERT CONTRACTION

The Boltzmann equation is a microscopic equation for colliding-gas evolution valid in a very tight regime. It is first order in time and so requires a complete description of the one-particle distribution function at one time, say $t = 0$, after which its functional form is completely fixed by the Boltzmann transport equation.

Describing the one-particle distribution function completely is a hopeless procedure, since the amount of information is too large. However, one wants to recover hydrodynamics, which is essentially a partial differential equation for a macroscopic description of the fluid at long times and distances compared to molecular scales. So there must exist a contraction mechanism that reduces the number of degrees of freedom required to describe the solution to the Boltzmann transport equation at such long times and distances. It is not obvious how that can happen, but Hilbert gave a proof that is central to understanding that it must happen and in a rather surprising way. We will call this process the Hilbert contraction. All analyses of the Boltzmann equation are based on this contraction. We would like to give it in detail because it is a beautiful argument, but space forbids this, so we outline how Hilbert reasoned.

Since we don't know what else to do when faced with such a highly nonlinear system, we construct a perturbation expansion in a small variable around some distribution function f, assumed to be given to us at t_0. Under some very mild assumptions, and assuming the existence of such a general perturbation expansion in some parameter δ, Hilbert was able to show that the evolution of f for $t > t_0$ is given in terms of its initial data at t_0 by the first three moments of f, namely $\rho, \mathbf{v}$, and T. The system has contracted down to a low-dimensional descriptive manifold whose coordinates are the same variables used by the hydrodynamic description. The beauty of Hilbert's proof is that it is constructive. It explicitly displays a recursive closed tower of constraint relations on the moments of f that come directly from the Boltzmann equation. The proof also shows that such a contracted description is unique—a very powerful result.

It must be pointed out that Hilbert's construction is on the time-evolved solution to the Boltzmann transport equation, not on the equation itself, which still requires a complete specification of f. It amounts to a hard mathematical statement on an effective field-theory description for times much greater than elementary collision times, but with space gradients still smooth enough to entertain a serious gradient perturbation expansion. As such, it says nothing about the turbulent regime, for example, where all these assumptions fail.

In standard physics texts one can read all kinds of plausibility arguments as to why this contraction process should exist, but they lack force, for, by arguing tightly, one can make the conclusion go the other way. This is why the Hilbert contraction is important. It is really a powerful and mathematically unexpected result about a highly nonlinear integro-differential equation of very special form. Beyond Hilbert's theorem and within the Boltzmann transport picture, we can say nothing more about the contraction of descriptions.

The construction of towers of moment constraints, coupled to a perturbation expansion that Hilbert developed for his proof of contraction, was used in a somewhat different form by Chapman and Enskog. Their main purpose was to devise a perturbation expansion with side constraints in such a way as to pick off the values of the coupling constants—which are called transport coefficients in standard terminology—for increasingly more sophisticated forms of macrodynamical equations.

One makes the usual kinetic assump-

THE HILBERT CONTRACTION (continued)

tions: The gas reaches local equilibrium in a collision time or so; the one-particle distribution function has a local Maxwell-Boltzmann form (or whatever form is appropriate), call it f_L; a second time scale is assumed where space gradients are still small, but collective modes develop at large distances and at times much greater than molecular collision times. Then one assumes a general functional perturbation expansion exists of the form

$$f = f_L(1 + \xi^{(1)} + \xi^{(2)} + \cdots),$$

which turns out to be explicitly a spatial gradient expansion:

$$f = f_L(1 + c_1(\mathrm{v})(\lambda\nabla) + c_2(\mathrm{v})(\lambda\nabla)^2 + \cdots)$$

where λ is the mean free path in the system and v is the macrovelocity.

The perturbation expansion is set up so that at nth order, the correction to f_L obeys an integral equation of the form $f_L C(\xi^{(n)}) = L_n$, where C is the Boltzmann collision operator and L_n is an operator that depends only on lower order spatial derivatives. This generates a recursive tower of relations $\xi^{(n)}$ whose solubility conditions at order n are the $(n-1)$th-order hydrodynamical equations.

For example, assume

$$f = f_L\left(1 + \xi^{(1)}\right);$$

that is, we keep only 1st order in ξ. Then in the Boltzmann collision term keep consistently only order $\xi^{(1)}$ and in the streaming operator put $\xi^{(0)} \equiv f_L$. So we get

$$(\partial_t + \mathrm{v}_\alpha\partial_\alpha + a_\alpha\frac{\partial}{\partial \mathrm{v}_\alpha})f_L = f_L C(\xi^{(1)}),$$

which is of the form

$$f_L C(\xi^{(1)}) = L_1.$$

The solubility conditions for this are that L_1 must be orthogonal to the five zero eigenmodes of $C(\xi^{(1)}) = 0$ (the solutions are 1, $\mathbf{v}$, and $\mathbf{v}^2$). These solubility conditions are the Euler equation for ρ, v, and T and the ideal gas equation of state. In this way one derives a sequence of hydrodynamical equations with explicit forms for the transport coefficients. Order 0 gives the Euler equation, order 1 gives the Navier-Stokes equations, order 2 and greater give the generalized hydrodynamical equations, which have some validity only in special situations. The expansion is an asymptotic functional expansion, so going beyond Navier-Stokes takes one away from ordinary fluids rather than closer to them. Solving explicitly for the various $\xi^{(n)}$ gives a way to evaluate the transport quantities (viscosity, etc.).

There are many other ways to do the same thing—multiple time expansions, dispersion methods, etc. We have developed everything so far within the conceptual frame of the Boltzmann transport equation. Within that framework the problem of deriving macrodynamical equations and associated transport coefficients reduces to tedious but straightforward linear algebra that has absorbed the best efforts of excellent technical people since the turn of the century. It is a problem best suited to a computer but only recently have algebraic processors of sufficient power been available.

This asymptotic perturbation expansion is a way to compute measurable quantities from microdynamical properties, but the physical insight one gains from doing it is small. The other methods mentioned, especially correlation-function techniques, are much more revealing. All of these comments and approaches carry over directly to the discrete case of the lattice gas. Nothing conceptually new arises in the totally discrete case, but explicit calculations are a great deal easier. ■

Boltzmann equation. The zero-order relation gives the Euler equations and the second-order relation gives the Navier-Stokes equations. However, Hilbert's method is an asymptotic functional expansion, so that the higher order terms take one away from ordinary fluids rather than closer to them. Nevertheless, solving explicitly for the terms in the functional expansion provides a way of evaluating transport coefficients such as viscosity. (See the "Hilbert Contraction" for more discussion.)

Summary of the Kinetic Theory Picture. Our review of the kinetic theory description of fluids introduced a number of important concepts: the idea of local thermal equilibrium; the characterization of an equilibrium state by a few macroscopic observables; the Boltzmann transport equation for systems of many identical objects (with ordinary statistics) in collision; and the fact that a solution to the Boltzmann transport equation is an ensemble of equilibrium states. In "The Hilbert Contraction" we introduced the linear approximation to the Boltzmann equation with which one can derive the Navier-Stokes equations for systems not too far (in an appropriate sense) from equilibrium in terms of these same macroscopic observables (density, pressure, temperature, etc.). We then outlined a method for calculating the coupling constants in the Navier-Stokes system—that is, the strengths of the nonlinear terms—as a function of any particular microdynamics.

This review was intended to give a flavor for the chain of reasoning involved. We will use this chain again in the totally discrete lattice world. However, just as important as understanding the kinetic theory viewpoint is keeping in mind its limitations. In particular, notice that perturbation theory was the main tool used for going from the exact Boltzmann transport equation to the Navier-Stokes equations. We did not discover more pow-

erful techniques for finding solutions to the Navier-Stokes equations than we had before. To go from the Boltzmann to the Navier-Stokes description, we made many smoothness assumptions in various probabilistic disguises; in other words, we recreated an approximation to the continuum. It is true one could compute (at least for relatively simple systems) the transport coefficients, but in a sense these coefficients are a property of microkinetics, not macrodynamics.

We are at a point where we can ask some questions about the emergence of macrodynamics from microscopic physics. It is clear by now that microscopic conservation laws, those of mass, momentum, and energy are crucial in fixing the form of large-scale dynamics. These are in a sense sacred. But one can question the importance of the description of individual collisions. How detailed must micromechanics be to generate the qualitative behavior predicted by the Navier-Stokes equations? Can it be done with simple collisions and very few classes of them? There exists a whole collection of equations whose functional form is very nearly that of the Navier-Stokes equations. What microworlds generate these? Do we have to be exactly at the Navier-Stokes equations to generate the qualitative behavior and numerical values that we derive from the Navier-Stokes equations or from real fluid experiments? Is it possible to design a collection of synthetic microworlds that could be considered local-interaction board games, all having Navier-Stokes macrodynamics? In other words, does the detailed microphysics of fluids get washed out of the macrodynamical picture under very rapid iteration of the deterministic system? If the microgame is simple enough to update it deterministically on a parallel machine, is the density of states required to see everything we see in ordinary Navier-Stokes simulations much smaller than the density of atoms in real physical fluids? If so, these synthetic microworlds become a potentially powerful analytic tool.

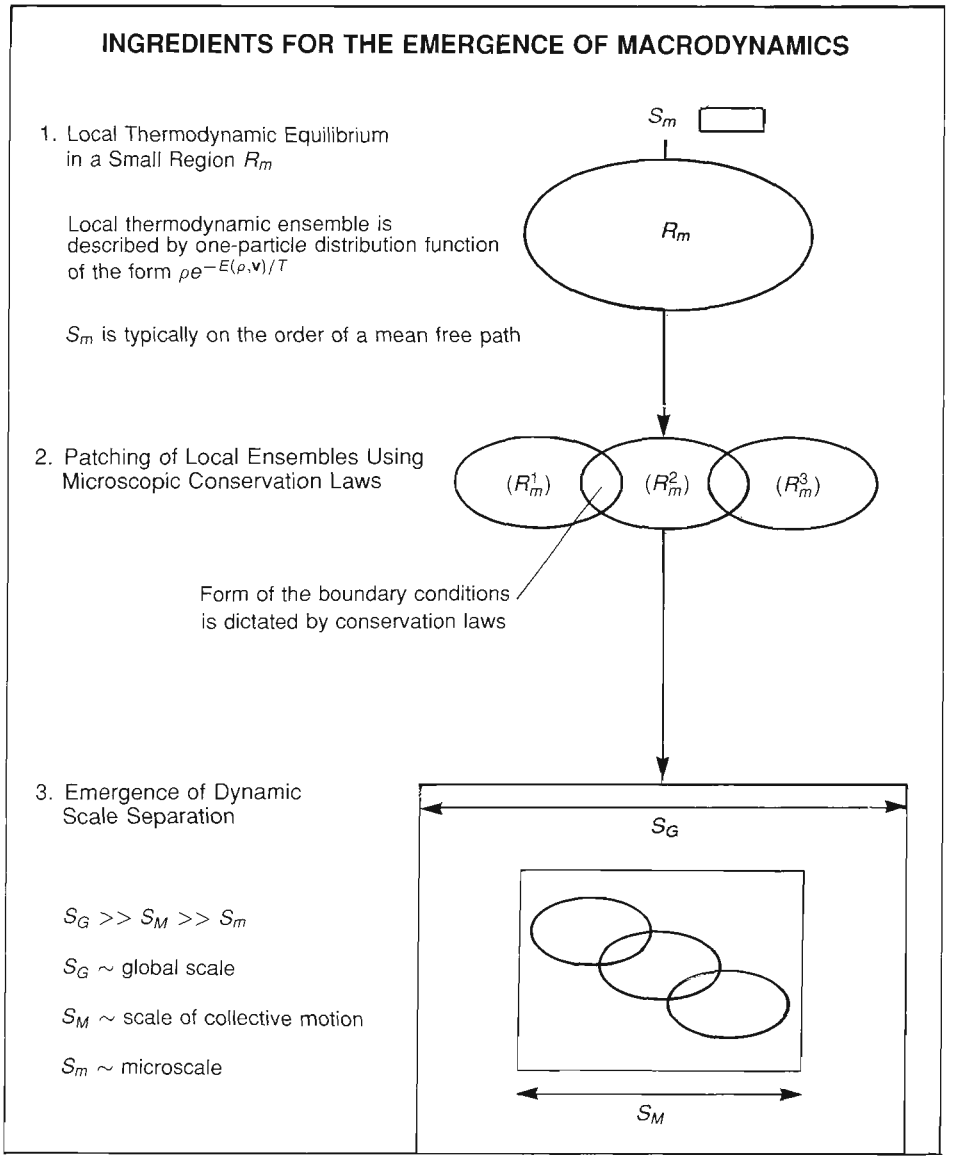


Fig. 3. Three ingredients are needed for the emergence of macrodynamics: local thermodynamic equilibrium, conservation laws, and scale separation between microkinetics and collective motion.

Our approach in building a cellular space is to move away from the idea of a fluid state and focus instead on the idea of the macrodynamics of a many-element system. In abstract terms, we want to devise the simplest deterministic local game made of a collection of few-bit, finite-state machines that has the Navier-Stokes equations as its macrodynamical description. From our brief look at kinetic transport theory, we can abstract the essential features of such a game (Fig. 3). The many-element system must be capable of supporting a notion of local thermodynamic equilibrium and must also include local microscopic conservation laws. The state of a real fluid can be imagined as a col-

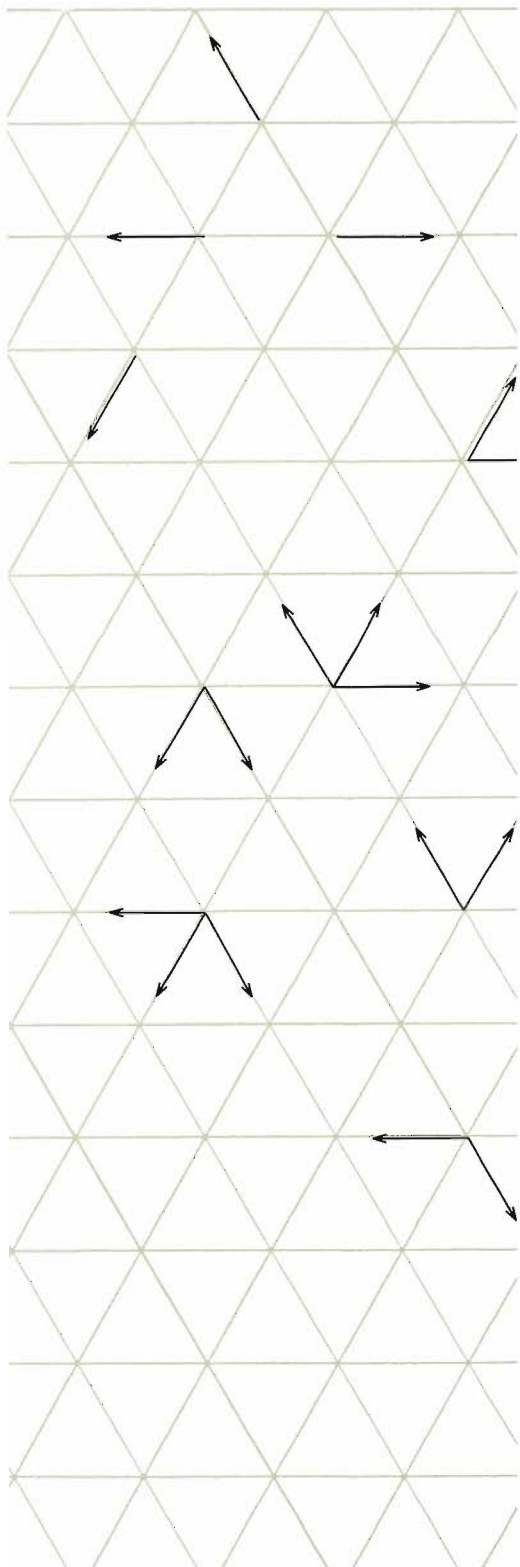
lection of equilibrium distribution functions whose macroscopic parameters are unconstrained. These distribution functions have a Maxwell-Boltzmann form, $e^{-E(\rho,v)/T}$. If these distribution functions are made to deviate slightly from equilibrium, then local conservation laws impose consistency conditions among their parameters, which become constrained variables. These consistency conditions are the macrodynamical equations necessary to put a consistent equilibrium function description onto the many-element system. In physical fluids they are the Navier-Stokes equations. This is the general setup that will guide us in creating a lattice model.

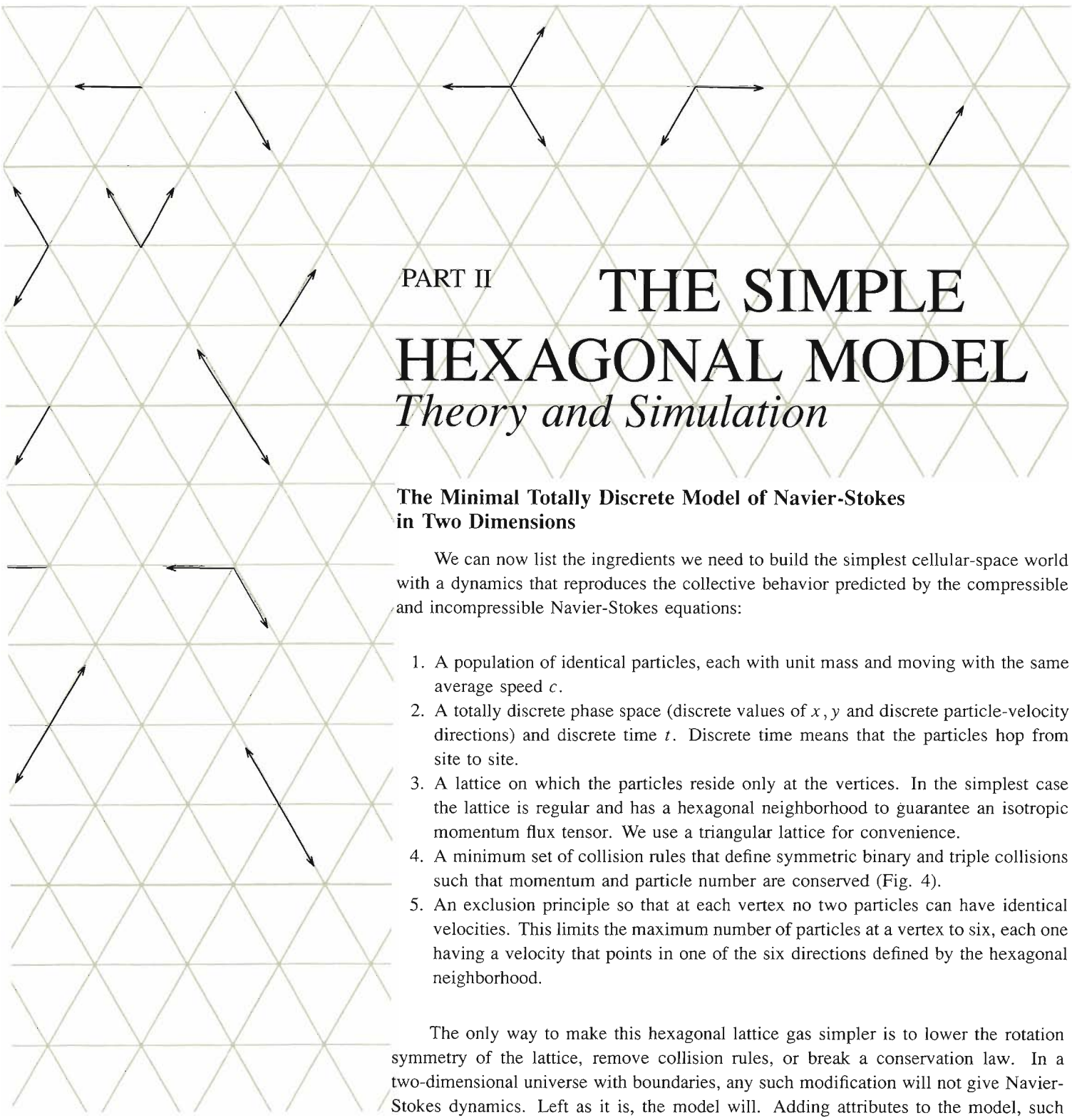
Evolution of Discrete Fluid Models

Continuous Network Models. The Navier-Stokes equations, however derived, are analytically intractable, except in a few special cases for especially clean geometries. Fortunately, one can avoid them altogether for many problems, such as shocks in certain geometries. The strategy is to rephrase the problem in a very simple phase space and solve the Boltzmann transport equation directly. If a single type of particle is constrained to move continuously only along a regular grid, the Boltzmann equation is so tightly constrained that it has simple analytic solutions. In the early 1960s Broadwell and others applied this simplified method of analysis to the dynamics of shock problems. Their numerical results agreed closely with much more elaborate computer modeling from the Navier-Stokes equations. However, there was no real insight into why such a calculation in such a simplified microworld should give such accurate answers. The accuracy of the limited phase-space approach was considered an anomaly.

Discrete Skeletal Models. The next development in discrete fluid theory was a discrete modification of the continuous-speed network models of the Broadwell class. By forming a loose analogy to the structure of the Ising model (spins on a lattice), Hardy, de Pazzis, and Pomeau created the first minimalist fluid model on a two-dimensional square lattice. It was a simple, binary-valued, nearest-neighbor gas with a single species of molecule, limited to binary collisions. The new feature was a totally discrete velocity and state space for the gas. Particles hopped from one site to the next without a notion of continuous movement between sites. Particles were confined to the vertices of the network, and the velocity vector of each particle could point in only one of four directions. Since there was no natural way to deal with bound states, these authors imposed the arbitrary rule that the maximum number of particles occupying any vertex be four.

This simple model possessed remarkable properties including local thermodynamic equilibrium and the emergence of a scale separation; that is, the typical collective motion scale L is much greater than the microscopic mean free path l_m; $L \gg l_m$. However, the macrodynamics that emerged was not that of the Navier-Stokes equations but a more complex one with unphysical features. The square model was the first example of rich dynamics emerging wholly on a cellular space. It had all the right ingredients except one: isotropy under the rotation group of the lattice. The momentum flux tensor must reduce to a scalar for isotropy, but this is impossible with a square lattice. In two dimensions the neighborhood that has the minimal required symmetry and tiles the plane is a hexagonal neighborhood. In Part II we will present the simple hexagonal model, analyze it mathematically, and describe the simulations of fluid phenomena that have been done so far.





PART II

THE SIMPLE HEXAGONAL MODEL

Theory and Simulation

The Minimal Totally Discrete Model of Navier-Stokes in Two Dimensions

We can now list the ingredients we need to build the simplest cellular-space world with a dynamics that reproduces the collective behavior predicted by the compressible and incompressible Navier-Stokes equations:

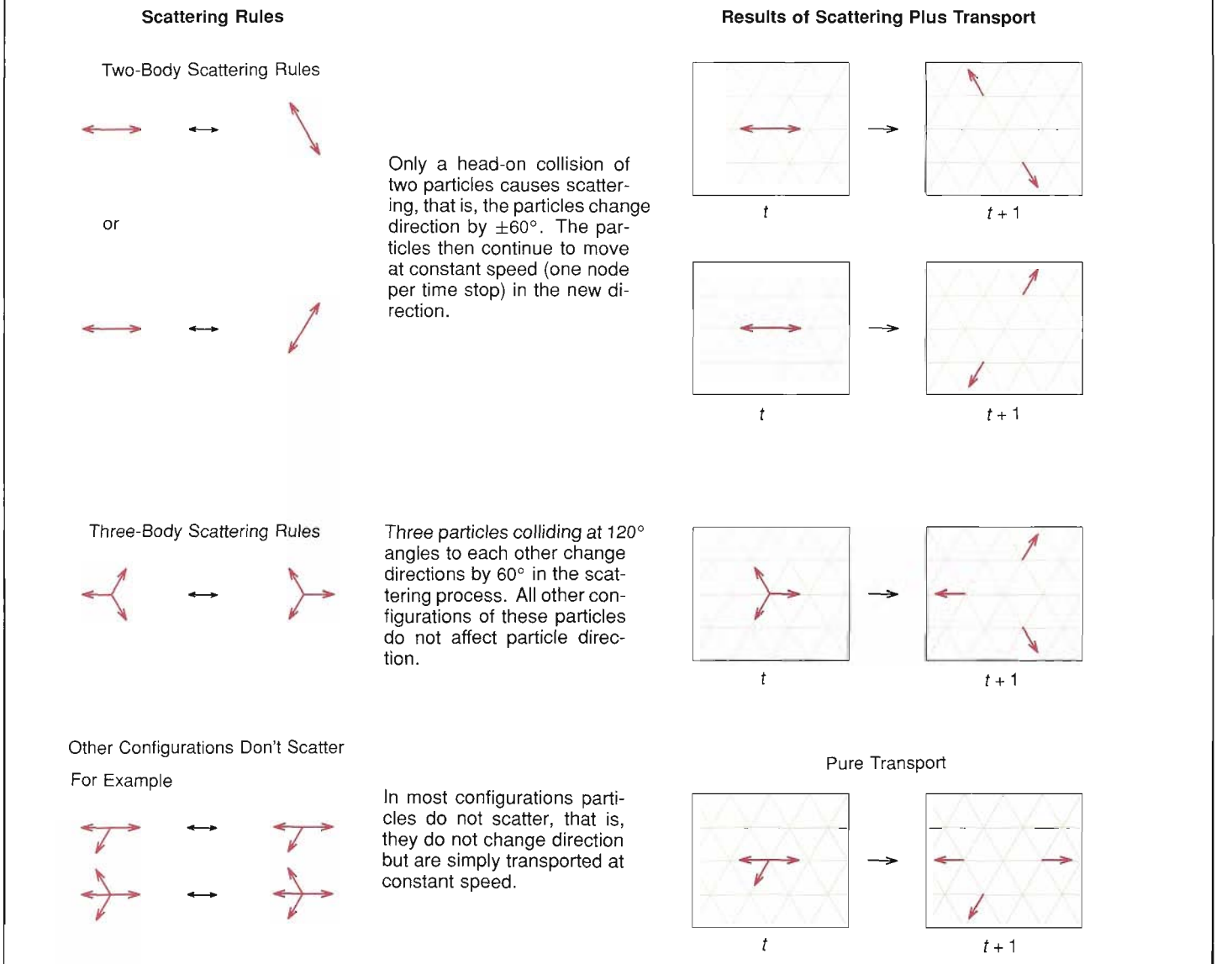
1. A population of identical particles, each with unit mass and moving with the same average speed c.
2. A totally discrete phase space (discrete values of x, y and discrete particle-velocity directions) and discrete time t. Discrete time means that the particles hop from site to site.
3. A lattice on which the particles reside only at the vertices. In the simplest case the lattice is regular and has a hexagonal neighborhood to guarantee an isotropic momentum flux tensor. We use a triangular lattice for convenience.
4. A minimum set of collision rules that define symmetric binary and triple collisions such that momentum and particle number are conserved (Fig. 4).
5. An exclusion principle so that at each vertex no two particles can have identical velocities. This limits the maximum number of particles at a vertex to six, each one having a velocity that points in one of the six directions defined by the hexagonal neighborhood.

The only way to make this hexagonal lattice gas simpler is to lower the rotation symmetry of the lattice, remove collision rules, or break a conservation law. In a two-dimensional universe with boundaries, any such modification will not give Navier-Stokes dynamics. Left as it is, the model will. Adding attributes to the model, such

as different types of particles, different speeds, enlarged neighborhoods, or weighted collision rules, will give Navier-Stokes behavior with different equations of state and different adjustable parameters such as the Reynolds number (see the discussion in Part III). The hexagonal model defined by the five ingredients listed above is the simplest model that gives Navier-Stokes behavior in a sharply defined parameter regime.

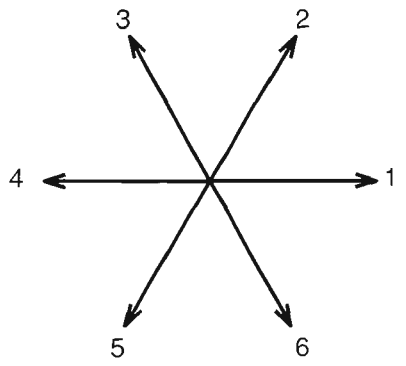
At this point it is instructive to look at the complete table of allowed states for the model (Fig. 5). The states and collision rules can be expressed by Boolean logic

Fig 4. SCATTERING RULES FOR SIMPLE HEXAGONAL MODEL



operations with the two allowed values taken as 0 and 1. From this organization scheme we see that the hexagonal lattice gas can be seen as a Boolean parallel computer. In fact, a large parallel machine can be constructed to implement part or all of the state table locally with Boolean operations alone. Our simulations were done this way and provide the first example of the programming of a cellular-automaton, or cellular-space, machine that evolves the dynamics of a many-degrees-of-freedom, nonlinear physical system.

Fig. 5. All possible states of the hexagonal lattice gas are shown in the state table. Each state can be expressed in 6-bit notation (a combination of 3 right bits and 3 left bits). For example, the empty state is written (000,000) and the maximally occupied state shown in the lower right hand corner of the table is written (111, 111). Collision states for the simplest hexagonal model are shown in red and shaded in gray. The scattering rules for these states are written beside the table. All other states do not result in scattering. The extended hexagonal model includes scattering rules for four-body states (shaded in gray). The extended model lowers the viscosity of the lattice gas.



PARTICLE DIRECTIONS IN THE HEXAGONAL MODEL

Fig. 6. The velocity vector of each particle can point in one of six possible directions. All particles have the same speed c.

Theoretical Analysis of the Discrete Lattice Gas

Before presenting the results of simulations with the lattice-gas automaton, we will analyze its behavior theoretically. The setup we work on is a regular triangular grid with hexagonal neighborhood. The natural explicit coordinate system for a single-speed, six-directional world (Fig. 6) is the set of unit vectors:

$$\hat{\mathbf{i}}_\beta = \left\{\cos\left(\frac{2\pi\beta}{6}\right),\ \sin\left(\frac{2\pi\beta}{6}\right)\right\}, \quad \beta = 1,\cdots,6. \tag{8}$$

One never requires this much detail except to work out explicit tensor structures and scalar products particular to the hexagonal model case, but the index conventions are important to avoid disorientation. From now on the Greek indices $\alpha,\beta,\cdots$ label lattice direction indices; $\hat{\mathbf{i}},\hat{\mathbf{j}},\hat{\mathbf{k}},\ldots$ are lattice unit vectors and i,j,k label space indices $(x_1,x_2,\ldots)$; on a square lattice we have $\mathbf{r} = (x_1,x_2) = (x,y)$.

The first thing we will look at is pure transport on the lattice with no collisions. Because the basic space is a discrete lattice with a fundamental lattice spacing, rather than a continuum, a shadow of the lattice is induced into the coupling constant of the theory, namely the viscosity. This lattice effect is not obvious, but we will make it so by looking at transport on the lattice in detail. As a corollary we will derive the usual Euler equations for the "macroscopic" flow of the lattice gas.

To do a quick analysis on lattice models we lift the restriction of a deterministic gas and pass to a probabilistic description familiar from kinetic theory; then we can use familiar stochastic and kinetic theory tools outlined in Part I of this article. In going from a continuous to a discrete probabilistic formalism we introduce the lattice form of the single-particle distribution function by making the identifications

$$f(\mathbf{r},\Gamma,t) \rightarrow f_\beta(\mathbf{r},t)$$

$$\int f(\mathbf{r},\Gamma,t)d\Gamma \rightarrow \Sigma_\beta f_\beta(\mathbf{r},t) \equiv \rho,.$$

and

$$\rho\mathbf{v} \rightarrow \Sigma_\beta \hat{\mathbf{i}}_\beta f_\beta$$

To begin we write the master equation for f_β in the absence of collisions. The master equation expresses conservation of probability. For simplicity we write it for a square lattice with the following conventions: $n_\beta(\mathbf{r}+\hat{\mathbf{i}}_\beta,t)$ = number of particles in the direction β at the node $\mathbf{r}+\hat{\mathbf{i}}_\beta$ at time t. The master equation for the system, neglecting collisions and written in a continuum notation for convenience, is

$$f_\beta(\mathbf{r}+h,t+k) - f_\beta(\mathbf{r},t) = 0, \text{ with } h = \hat{\mathbf{i}}_\beta d_x, k = d_t,$$

where $d_t, d_x \ll 1$.

If we expand the first term in the master equation out to $O^2(h,k)$ using the Taylor series expansion $f(x_0+h,y_0+k) = \sum_{\lambda=0}^{m-1} \frac{1}{\lambda!}(h\partial_x + k\partial_y)^\lambda f(x_0,y_0) + R_m$, we obtain

$$0 = d_t\partial_t f_\beta + d_x\hat{\mathbf{i}}_\beta\cdot\nabla f_\beta + \frac{1}{2}d_t^2\partial_t^2 f_\beta + \frac{1}{2}d_x^2(\hat{\mathbf{i}}_\beta\cdot\nabla)^2 f_\beta + d_x d_t(\hat{\mathbf{i}}_\beta\cdot\nabla)\partial_t f_\beta.$$

To lowest order in h and k, we have

$$\partial_t f_\beta + \hat{\mathbf{i}}_\beta \cdot \nabla f_\beta = 0, \tag{9}$$

which has the standard form of the kinetic theory transport equation in the absence of collisions. If we include collisions, the full Boltzmann transport equations schematically become

$$\partial_t f_\beta + \hat{\mathbf{i}}_\beta \cdot \nabla f_\beta = C_\beta(f) \tag{10}$$

where $C_\beta(f)$ is the collision operator on the lattice. The form of the lattice collision operator will tell us a great deal about how the model works, but for the moment we just look at the general structure of the "macroscopic" equations for the lattice gas to the lowest order in the lattice expansion parameters.

As in standard kinetic theory, the usual zero integrals of the motion hold, since the lattice model is assumed to have some kind of detailed balance (that is, microscopic reversibility of reaction pathways). Accordingly $\sum_\beta C_\beta(f) = 0$ and $\sum_\beta \hat{\mathbf{i}}_\beta C_\beta(f) = 0$ for a skeletal gas. Following the kinetic theory procedure, we write the continuity and momentum equations that follow from these conditions as:

$$\partial_t \rho + \partial_i(\rho \mathrm{v}_i) = 0 \tag{11}$$

and

$$\partial_t(\rho \mathrm{v}_i) + \partial_j \Pi_{ij} = 0, \tag{12}$$

where the tensor Π_{ij} is defined as

$$\Pi_{ij} = \sum_\beta (\hat{i}_\beta)_i (\hat{i}_\beta)_j f_\beta. \tag{13}$$

So far we have kept only the leading terms of the Taylor series expansion in the scaling factors that relate to the discreteness of the lattice. It's easy to show that keeping quadratic terms in this lattice-size expansion leaves the continuity equation invariant but alters the momentum equation by introducing a free-streaming correction to the measured viscosity. This rather elegant way of viewing this correction was first developed by D. Levermore. The correction comes from breaking the form of a Galilean covariant derivative and is a geometrical effect. Specifically, to second order in the lattice size expansion, the momentum equation does not decompose simply into factors of these covariant derivatives but instead the expansion introduces a nonvanishing covariant-breaking term:

$$\text{Noncovariant term} = \sum_\beta \{(\hat{i}_\beta)_i \partial_i \partial_t + \left((\hat{i}_\beta)_i \partial_i\right)^2\} \hat{\mathbf{i}}_\beta f_\beta. \tag{14}$$

This term is of the same order as those terms that contribute to the viscosity. Later we will show how to use the Chapman-Enskog expansion to compute an explicit form for the lattice-gas viscosity.

The Chapman-Enskog Expansion and the Direct Expansion. The form of Π_{ij} depends on the form of f, the solution to the full lattice Boltzmann transport equation.

By Hilbert's construction we know that an efficient expansion can be developed in terms of the collision invariants of the model up to powers of terms linear in the gradient of the macroscopic velocity. In whatever perturbation expansion of f we choose, the coefficients in the expansion are fixed by solving for them under the Lagrange multiplier constraints of mass and momentum conservation: $\rho = \sum_\beta f_\beta$ and $(\rho \mathbf{v}) = \sum_\beta \hat{\mathbf{i}}_\beta f_\beta$. In the simple hexagonal model there is no explicit mechanism provided for storing energy in internal state space, so there is no independent energy equation.

For the lattice case, the Chapman-Enskog version of Hilbert's expansion reduces to an expansion in all available scalar products using the vectors $\hat{\mathbf{i}}_\beta, \mathbf{v}$ and the vector operator $\overrightarrow{\partial}$. The expansion is made around the global equilibrium solution for $\mathbf{v} = 0$, which we will call $N_{\text{eq}}^{\mathbf{v}=0}$ and terms are kept up to those linear in $\overrightarrow{\partial}$. The relevant scalar products are

$$(\hat{\mathbf{i}}_\beta \cdot \hat{\mathbf{i}}_\beta),\ (\hat{\mathbf{i}}_\beta \cdot \mathbf{v}),\ (\hat{\mathbf{i}}_\beta \cdot \overrightarrow{\partial})(\hat{\mathbf{i}}_\beta \cdot \mathbf{v}),\ (\overrightarrow{\partial} \cdot \mathbf{v}),\ (\hat{\mathbf{i}}_\beta \cdot \mathbf{v})^2,\ (\mathbf{v} \cdot \mathbf{v}) + O(\mathbf{v}^3).$$

The systematic expansion becomes

$$f_\beta = N_{\text{eq}}^{\mathbf{v}=0} \left\{ 1 + \alpha \hat{\mathbf{i}}_\beta \cdot \mathbf{v} + \beta \left[(\hat{\mathbf{i}}_\beta \cdot \mathbf{v})^2 - \frac{1}{2}|\mathbf{v}|^2 \right] + \beta_1 \left[(\hat{\mathbf{i}}_\beta \overrightarrow{\partial})(\hat{\mathbf{i}}_\beta \mathbf{v}) - \frac{1}{2} \overrightarrow{\partial} \cdot \mathbf{v} \right] + O(\mathbf{v}^3) \ldots \right\}. \tag{15}$$

In the usual kinetic theory approach the coefficients α and β can be found by neglecting collisions and β_1, the gradient term, can be determined only by an explicit solution to the full Boltzmann equation including collision terms. In this way one obtains the viscosity in terms of β_1. For the discrete lattice, however, both β and β_1 depend on the explicit form of the solution to the full Boltzmann equation with collisions. We also need that form to recover the correction to the raw viscosity that, as mentioned in the last section, comes from pure translation effects on the lattice.

Given that we have to use the full solution to the Boltzmann transport equation almost immediately, we now derive its structure, find the general and equilibrium solution, and then use a direct expansion to fix both β and β_1. In the process we will recover the Euler equations for inviscid flow and the Navier-Stokes equations for the flow with dissipation.

The Lattice Collision Operator and the Solution to the Lattice Boltzmann Transport Equation. We will write down the discrete form of the Boltzmann equation, especially noting the collision operator, for a number of reasons. First, writing the explicit form of the collision kernel builds up an intuition of how the heart of the model works; second, we can show in a few lines that the Fermi-Dirac distribution satisfies the lattice gas Boltzmann equation; third, knowing this, we can quickly compute the lattice form of the Euler equations; fourth, we can see that many properties of the lattice-gas model are independent of the types of collisions involved and come only from the form of the Fermi-Dirac distribution.

Collision operators for lattice gases with continuous speeds were derived by Broadwell, Harris, and other early workers on continuum lattice-gas systems. For totally discrete lattice gases with an exclusion principle, we must be careful to apply this principle correctly. It is similar to the case of quasi-particles in quantum Fermi liquids.

The construction reduces to following definitions of collision operators introduced in the section on classical kinetic theory and counting properly.

Taking any hexagonal neighborhood, let i be one of the six directions and use the convention $i, i+1, i+2, \ldots \equiv \hat{\mathbf{i}}_{\beta=i}, \hat{\mathbf{i}}_{\beta=i+1}, \hat{\mathbf{i}}_{\beta=i+2}, \ldots$ for convenience. (Later we will return to our original notation.) First consider binary collisions alone, and assume detailed balance, which implies microscopic reversibility of a collision at each vertex. One need not use detailed balance, but other balancing schemes are algebraically tedious and conceptually similar extensions of this basic case. Given a vertex at $(\mathbf{r}, t)$ we compute the gain and loss of particles into a neighborhood along a fixed direction, say i. This is, by definition, the collision kernel for binary processes. First compute the number of particles thrown in a collision into a phase-space region along the direction i. Let $n_i(\mathbf{r}, t)$ be the probability that a particle is at the node $(\mathbf{r}, t)$ and has a velocity in the ith direction.

If a particle scatters into a vector direction i, it must have come from binary processes along directions ($i+1$ and $i+4$) or ($i+2$ and $i+5$) (see the two-body scattering rules in Fig. 4). Interpreted as probabilities for the two events to happen, the probability for gain in the i direction due to binary processes alone is

$$P_i^{\text{binary}} = n_{i+1} n_{i+4} \bar{n}_i \bar{n}_{i+2} \bar{n}_{i+3} \bar{n}_{i+5} + n_{i+2} n_{i+5} \bar{n}_i \bar{n}_{i+1} \bar{n}_{i+3} \bar{n}_{i+4},$$

where $\bar{n}_k \equiv (1 - n_k)$. The $\bar{n}_k$'s impose the exclusion rule in the output channel, namely, that a particle cannot scatter there if one is already present.

Loss of a particle from direction i can occur only by the binary collision $(i+3, i)$, and this can happen for each of the two choices of gain collisions separately. So we have $(-2 n_i n_{i+3} \bar{n}_{i+1} \bar{n}_{i+2} \bar{n}_{i+4} \bar{n}_{i+5})$ as the probability for loss in the i direction due to binary collisions alone. Note that these products can be compactly expressed as $\hat{n}_i \hat{n}_{i+3} \Pi_{i=0}^{5} (1 - n_i)$ where $\hat{n}_i \equiv \frac{n_i}{1-n_i}$.

The three-body gain-loss term can be written down by inspection in the same way as the binary term. The complete two- and three-body collision term for the ith direction, in compact notation, is

$$C_i^{2+3} = [(\hat{n}_{i+1}\hat{n}_{i+4} + \hat{n}_{i+2}\hat{n}_{i+5} - 2\hat{n}_i\hat{n}_{i+3}) + (\hat{n}_{i+1}\hat{n}_{i+3}\hat{n}_{i+5} - \hat{n}_i\hat{n}_{i+2}\hat{n}_{i+4})] \, \Pi_{i=0}^{5}(1 - n_i).$$

For extensive calculations more compact notations are easily devised, but this one clearly brings out the essential idea in constructing arbitrary collision schema. With some minor modifications this form for the collision operator can be reinterpreted as a master equation for a transition process, which is useful as a starting point for a detailed microkinetic analysis by stochastic methods.

Given the $C(f)$ for two- and three-body collisions in the above compact form, and given detailed balance, we show that $C(f_\beta) = 0$ for the Fermi-Dirac distribution. The proof is simple and well known from quantum Fermi-liquid theory where the same functional form for the collision operator appears but with a different interpretation.

If n is a Fermi-Dirac distribution, it has the form $(1 + e^E)^{-1} = n(E)$ where E is expanded in collision invariants, in this case particle number and momentum. Then note that $\frac{n}{1-n} \equiv \hat{n} = e^{-E}$, the form of the Maxwell-Boltzmann distribution. This is also the form of the collision kernel, and the exponential terms just contain the sum of momenta in the collision. Since this sum is conserved, each collision term (binary,

triple, etc.) vanishes separately, because of the exclusion principle. So the solution is a Fermi-Dirac distribution. This proof also shows that as long as conservation laws of any kind are embodied in the collision term, each type of collision is separately zero under the Fermi-Dirac distribution. Accordingly, the Fermi-Dirac solution is universal across collision types. This implies that one cannot alter the character of the Fermi-Dirac distribution in the lattice gas by adding collision types that respect collision invariants.

Since f_β is now assumed to be a Fermi-Dirac distribution, we take it as

$$f_\beta = (1 + e^E)^{-1},$$

with

$$E = \alpha(\rho, \mathbf{v}) + \overrightarrow{\beta}(\rho, \mathbf{v}) \cdot \hat{\mathbf{i}}_\beta.$$

(Here we have returned to our original conventions for $\hat{\mathbf{i}}_\beta$.) The equilibrium value for f_β at $\mathbf{v} = 0$, namely $N_{\text{eq}}^{\mathbf{v}=0}$, is $\frac{\rho}{6}$ where ρ is the density. Expanding the Fermi-Dirac form for f_β about this equilibrium value gives us

$$f_\beta = \frac{\rho}{6}\left\{1 + \alpha(\rho)(\hat{\mathbf{i}}_\beta \cdot \mathbf{v}) + \beta(\rho)\left[(\hat{\mathbf{i}}_\beta \cdot \mathbf{v})^2 - \frac{1}{2}|\mathbf{v}|^2\right] + \cdots\right\}, \tag{16}$$

the same form as the Chapman-Enskog expansion (Eq. 15). To fix α and β we use number and momentum conservation as constraints, so that f_β becomes

$$f_\beta = \frac{\rho}{6}\left\{1 + 2(\hat{\mathbf{i}}_\beta \cdot \mathbf{v}) + 4g(\rho)\left[(\hat{\mathbf{i}}_\beta \cdot \mathbf{v})^2 - \frac{1}{2}|\mathbf{v}|^2\right] + \cdots\right\},$$

where we have taken the particle speed as 1 ($c = 1$). The coefficient $g(\rho)$ is

$$g(\rho) = \frac{3 - \rho}{6 - \rho}.$$

If we substitute this result for f_β in the momentum tensor (Eq. 13) and do the sum over β, the particle directions, we have

$$\Pi_{ik} = \frac{\rho}{2}\left(1 - g(\rho)\mathbf{v}^2\right)\delta_{ik} + \rho g(\rho)\mathbf{v}_i\mathbf{v}_k.$$

The lattice Euler equation (Eq. 12) thus becomes

$$\partial_t(\rho \mathbf{v}_i) + \partial_j\left[\rho g(\rho)\mathbf{v}_i\mathbf{v}_j + \cdots\right] = -\partial_i p. \tag{17}$$

In the usual Euler equation $g(\rho) = 1$. Here $g(\rho)$ is the lattice correction to the convective term due to the explicit lattice breaking of Galilean invariance. The equation of state for Eq. 17 is

$$p = \frac{\rho}{2}\left(1 - g(\rho)\frac{\mathbf{v}^2}{2}\right).$$

For general single-speed models with particle speed c and b velocity vectors in D dimensions, the result above generalizes to

$$f_\beta = \frac{\rho}{b} \text{when} \quad \mathbf{v} = \mathbf{0}$$

and

$$g(\rho) = \frac{D}{D+2}\frac{b-2\rho}{b-\rho}.$$

These forms depend only on the structure of tensor products of $\hat{i}_\alpha$ in D dimensions.

When we discuss the full Navier-Stokes equations, we will show how to absorb the $g(\rho)$ Galilean-invariance-breaking term in Eq. 17 into a rescaling of variables.

Isotropy and The Momentum Tensor. We will go on to discuss viscosity and the lattice form of the Navier-Stokes equation, but first we comment briefly on how the structure of the momentum tensor depends on the geometry of the lattice. Those interested in all the details can find them discussed from several viewpoints in Frisch, d'Humières, Hasslacher, Lallemand, and Pomeau 1987.

By definition $\Pi_{ij} \equiv \sum_\beta (\hat{i}_\beta)_i (\hat{i}_\beta)_j f_\beta$, where f_β is determined by the Chapman-Enskog, or direct, expansion (Eq. 15). Isotropy implies invariance under rotations and reflections; tensors that are isotropic are proportional to a scalar. Define the tensors $E^{(n)} = \sum_\beta (\hat{i}_\beta)_{i_1} \ldots (\hat{i}_\beta)_{i_n}$. For $E^{(n)}$ with regular b-sided polygons, we can derive conditions on b for $E^{(n)}$ to be isotropic. These conditions are

$$\left(E^{(2)}|b>2\right),\ \left(E^{(3)}|b\geq 2, b\neq 3\right),\ \left(E^{(4)}|b>2, b\neq 4\right),\ \left(E^{(5)}|b\geq 2, b\neq 3,5\right),\ldots.$$

For $b = 4$, the case of the HPP (Hardy, de Pazzis, and Pomeau) square lattice, $E^{(4)}$ is not isotropic. For $b = 6$, the hexagonal neighborhood case, all tensors up to $n = 5$ are isotropic.

Using the Chapman-Enskog expansion for f_β and the notation above for tensors, Π_{ij} has the following tensor structure.

$$\Pi_{ij} \approx N_{\text{eq}}^{\mathbf{v}=\mathbf{0}} \left(E_{ij}^{(2)} + \alpha E_{ijk}^{(3)} \mathrm{v}_k + \beta \left[E_{ijkl}^{(4)} \mathrm{v}_k \mathrm{v}_l, E_{ij}^{(2)} \mathrm{v}_k \mathrm{v}_k \right] + \beta_1 \left[E_{ijkl}^{(4)} \partial_k \mathrm{v}_l, E_{ij}^{(2)} \partial_k \mathrm{v}_k \right] \right)$$

where we are following the discussion of Wolfram. The momentum stress tensor must be isotropic up to $E^{(4)}$ in order that the leading terms in the momentum equation (corresponding to the convective and viscous terms in the Navier-Stokes equation) be isotropic. For the square model, the original discrete-lattice model, we have nonisotropy manifested in two places through the momentum flux tensor.

$$\Pi_{11} = \rho g(\rho)(\mathrm{v}_1^2 - \mathrm{v}_2^2) + \frac{\rho}{2} + O(\mathrm{v}^4),$$

$$\Pi_{22} = \rho g(\rho)(\mathrm{v}_2^2 - \mathrm{v}_1^2) + \frac{\rho}{2} + O(\mathrm{v}^4),$$

$$\Pi_{12} = \Pi_{21} = 0,$$

where

$$g(\rho) = \frac{2-\rho}{4-\rho}.$$

See Frisch et al. for further discussion.

The nonisotropy implies that we do not get a Navier-Stokes type equation for the square lattice. For the hexagonal model, $\beta = 6$, isotropy is maintained through order $E^{(4)}$. By using general considerations on tensor structures for polygons and polyhedra in D-dimensional space, one can quickly arrive at probable models for Navier-Stokes dynamics in any dimension. The starting point is that isotropy, or the lack of it, in both convective and viscous terms (the Euler and the Navier-Stokes equations), is controlled completely by the geometry of the underlying lattice. This crucial point was missed by all earlier workers on lattice models who thought that the geometry of the underlying lattice was irrelevant.

Viscosity for Lattice Gas Models. In "The Continuum Argument" we saw that the general form of the compressible Navier-Stokes equation with bulk viscosity $\zeta = 0$ is

$$\partial_t(\rho \mathrm{v}_i) + \partial_j(\rho(\mathrm{v}_i \mathrm{v}_j) = -\partial_i p + \partial_j \left(\nu\rho(\partial_j \mathrm{v}_i + \partial_i \mathrm{v}_j - \frac{2}{3}\delta_{ij}\partial_k \mathrm{v}_k)\right),$$

where ν is the kniematic shear viscosity. To derive this form for the discrete model, one must solve for Π_{ij} using both the Chapman-Enskog approximation for f_β and the momentum-conservation equation. We noted earlier that the momentum equation contained corrections as powers of the lattice spacing but chose to ignore these at first pass. However, if we use the full Taylor expansion developed in the lattice-size scaling, we find that the contribution to the viscous term of the momentum equation is $-\frac{1}{8}\rho\nabla^2\mathbf{v}$. Note that the correction to the viscosity is a constant (see Eq. 19) that depends only on the lattice and dimension and is independent of the scattering-rule set. This extra noncovariant-derivative contribution to the viscosity must be subtracted from the bare viscosity calculated from the normal perturbation expansion to get the renormalized viscosity, which is the one actually measured in the lattice gas. In other words, the bare coupling constant of the lattice gas model gets "dressed" by this constant amount, owing to the discrete vacuum that the particle must pass through, to become the physical lattice-gas viscosity.

Viscosity is a coupling constant and can be found by any method that can isolate the β_1 term in the Chapman-Enskog expansion. The simplest methods involve solving for the eigenvalues and right eigenvectors of the linearized collision operator, which is a tedious exercise in linear algebra. Using the results of such a calculation, we can write the Navier-Stokes form of the momentum equation in which the viscosity $\nu(\rho)$ appears explicitly:

$$\partial_t(\rho \mathrm{v}_i) + \partial_j \Pi_{ij} = \partial_j S_{ij}, \tag{18}$$

where the momentum tensor Π_{ij} and the viscosity stress tensor are

$$\Pi_{ij} = c_s^2\rho(1 - g(\rho)\frac{\mathrm{v}^2}{c^2})\delta_{ij} + \rho g(\rho)\mathrm{v}_i\mathrm{v}_j$$

and

$$S_{ij} = \nu(\rho)\left(\partial_i(\rho \mathrm{v}_j) + \partial_j(\rho \mathrm{v}_i) - \frac{2}{D}\delta_{ij}\partial_k(\rho \mathrm{v}_k)\right).$$

The coefficient c_s^2 is given by

$$c_s^2 = \frac{c^2}{D},$$

and c_s can be identified as the speed of sound. For the simple hexagonal model $c_s = 1/\sqrt{2}$, and the viscosity is given by

$$\nu = \frac{1}{12}\frac{1}{d(1-d)^3} - \frac{1}{8}, \tag{19}$$

where $d = \frac{\rho_0}{6}$, that is, the mass density per cell. (The $-\frac{1}{8}$ in the viscosity was mentioned above as the noncovariant correction due to the finite lattice size.)

The Incompressible Limit. Many features of low Mach number ($M = v/c_s \ll 1$) flows in an ordinary gas can be described by the incompressible Navier-Stokes equations:

$$\partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \nu \nabla^2 \mathbf{v} \tag{20}$$

and

$$\nabla \cdot \mathbf{v} = 0.$$

We end this theoretical analysis by showing under what conditions we recover these equations for lattice gases. One way is to freeze the density everywhere except in the pressure term of the momentum equation (Eq. 18). Then, in the low-velocity limit, we can write the lattice Navier-Stokes equations as

$$\rho_0 \partial_t \mathbf{v} + \rho_0 g(\rho_0)\mathbf{v} \cdot \nabla \mathbf{v} = -c_s^2 \nabla \rho' + \rho_0 \nu(\rho_0)\nabla^2 \mathbf{v}, \tag{21}$$

and

$$\nabla \cdot \mathbf{v} = 0$$

where $\rho = \rho_o + \rho'$ and we allow density fluctuations in the pressure term only. As it stands, Eq. 21 is not Galilean invariant. To make it so, we must scale away the $g(\rho_0)$ term in a consistent way. We rescale time and viscosity as follows:

$$t \rightarrow \frac{t}{g(\rho_0)} \quad \text{and} \quad \nu \rightarrow g(\rho_0)\nu.$$

To be more precise, we do an ϵ expansion of the momentum equation, where ϵ^{-1} is the same order as the global lattice size L_g (see Frisch et al. for details), and rescale the variables as follows:

$$\mathbf{r} = \epsilon^{-1}\mathbf{r}_1, \qquad t = \frac{1}{g(\rho_0)}\epsilon^{-2}T,$$

$$\mathbf{v} = \epsilon \mathbf{V}, \qquad \rho' = \frac{\rho_0 g(\rho_0)}{c_s^2}\epsilon^2 P',$$

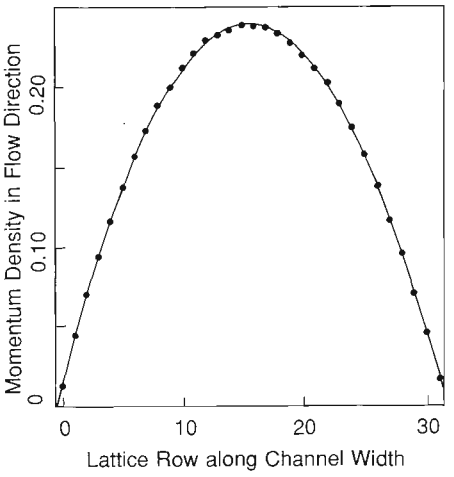
and

$$\nu = g(\rho_0)\nu'.$$

where ϵ^{-1} is on the order of the global lattice size L_g. (Note that this rescaling

SIMULATED VELOCITY PROFILE

Fig. 7. The predicted velocity profile was obtained in a low-velocity lattice gas simulation of two- dimensional flow in a channel with viscous boundaries (Kadanoff, McNamara, and Zanetti 1987).



of variables keeps the Reynolds number fixed.) Now all the relevant terms in the momentum equation are of $O(\epsilon^3)$ and higher order terms are $O(\epsilon^4)$ or smaller. So to leading order (where ∇_1 means $\frac{\partial}{\partial \mathbf{r}_1}$) we get

$$\partial_T \mathbf{V} + \mathbf{V} \cdot \nabla_1 \mathbf{V} = -\nabla_1 P' + \nu' \nabla_1^2 \mathbf{V} \text{ and } \nabla_1 \cdot \mathbf{V} = 0.$$

Thus we recover the incompressible Navier-Stokes equations. To obtain this result, we have done a fixed-Reynolds-number, large-scale, low-Mach-number expansion and Galilean invariance has been restored, at least formally, by a time rescaling.

Simulations of Fluid Dynamics with the Hexagonal Lattice Gas Automaton

In the last two years several groups in the United States and France have done simulations of fluid-dynamical phenomena using the hexagonal lattice-gas automaton. The purpose of these simulations was twofold: first, to check the internal consistency of the automaton, and second, to determine, by both qualitative and quantitative measures, whether the model behaves the same or nearly the same as the known analytic and numerical solutions of the Navier-Stokes equations.

The classes of experiments done can be grouped roughly as free flows, flow instabilities, flows past objects, and flows in channels or pipes. These simulations were run in a range of Reynolds numbers between 100 and 700 (and for relatively low mean flow velocities, so that the fluid is nearly incompressible). We first checked to see whether the automaton developed various classic instabilities when triggered by two types of mechanisms, external perturbations and internal noise. The two classic instabilities studied were the Kelvin-Helmholtz instability of two opposing shear flows and the Rayleigh-Taylor instability. We describe the Kelvin-Helmholtz instability in some detail.

In the Kelvin-Helmholtz instability one is looking for the development of a final-state vortex structure of appropriate vortex polarity. From an initial state of two opposing flows undergoing shear, the detailed development of the instability depends on the initial perturbation of the flows. Left unperturbed, except by internal noise in the automaton, at first the two opposing flows develop velocity fields that signal the development of a boundary layer, then sets of vortices develop in these boundary layers, and finally vortex interactions occur that trigger a large-scale instability and the development of large-vortex final states. The same pattern appears in standard two-dimensional numerical simulations of the Navier-Stokes equations near the incompressible regime. No pathological non-Navier-Stokes behavior was observed. These results extend over the entire range of Reynolds numbers (100–700) run with the simple hexagonal model. It is notable that the Kelvin-Helmholtz instability is self-starting due to the automaton internal noise, and the instability proceeds rapidly.

The Rayleigh-Taylor instability was simulated by a French group in a slightly compressible fluid range, where it behaves like a Navier-Stokes fluid with no anomalies.

These global topological tests check whether automaton dynamics captures the correct overall structure of fluids. In general, whenever the automaton is run in the Navier-Stokes range, it produces the expected global topological behavior and correct

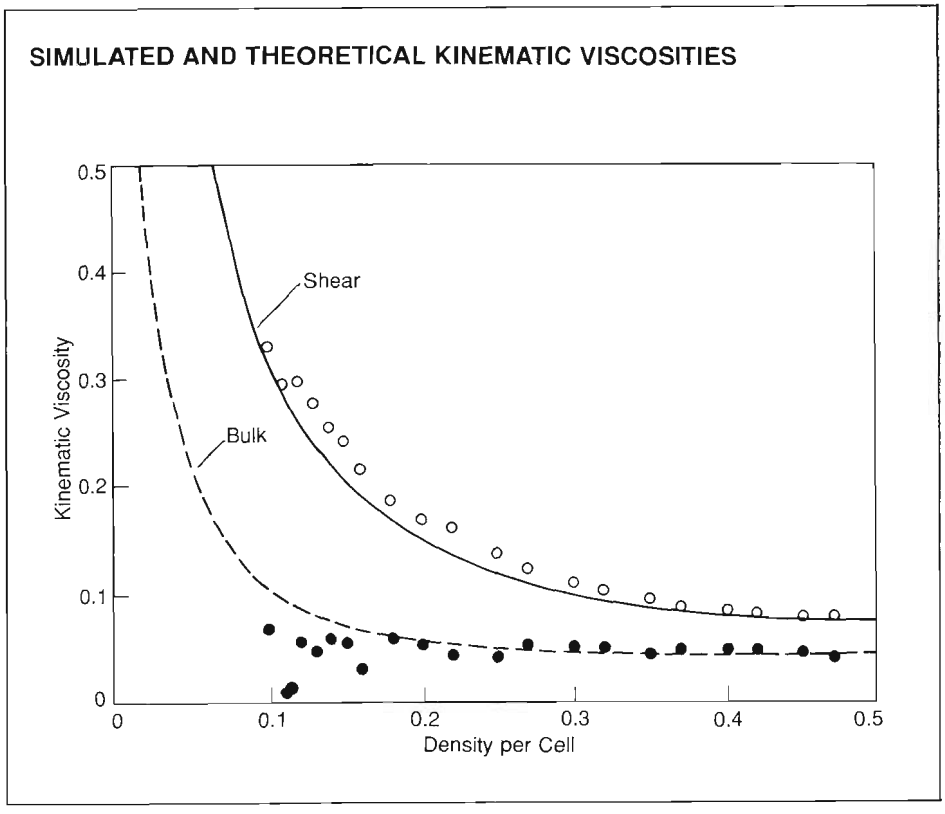


Fig. 8. Theoretical shear (solid line) and bulk (dashed line) reduced viscosities as a function of reduced density compared with the results of hexagonal lattice gas simulation with rest particles and all possible collisions (d'Humiéres and Lallemand 1987).

functional forms for various fluid dynamical laws. The question of quantitative accuracy of various known constants is harder to answer, and we will take it up in detail later.

The next broad class of flows studied are flows past objects. Here, we look for distinctive qualitative behavior characteristic of a fluid or gas obeying Navier-Stokes dynamics. The geometries studied, through a wide range of Reynolds numbers, were flows past flat plates placed normal to the flow, flows past plates inclined at various angles to the flow, and flows past cylinders, 60-degree wedges, and typical airfoils. The expected scenario changes as a function of increasing Reynolds number: recirculating flow behind obstacles should develop into vortices, growing couples of vortices should eventually break off to form von Karman streets with periodic oscillation of the von Karman tails; finally, and as the Reynolds number increases, the periodic oscillations should become aperiodic, and the complex phenomena characteristic of turbulent flow should appear. The lattice gas exhibits all these phenomena with no non-Navier-Stokes anomalies in the range of lattice-gas parameters that characterize near incompressibility.

The next topic is quantitative self-consistency. We used the Boltzmann transport approximation for the discrete model to calculate viscosities for the simple hexagonal automaton as well as models with additional scattering rules and rest particles. We then checked these analytic predictions against the viscosities deduced from two kinds of

simulations. We ran plane-parallel Poiseuille flow in a channel, saw that it developed the expected parabolic velocity profile (Fig. 7) and then deduced the viscosity characteristic of this type of flow. We also ran an initially flat velocity distribution and deduced a viscosity from the observed velocity decay. These two simulations agree with each other to within a few percent and agree with the analytic predictions from the Boltzmann transport calculation to within 10 percent. Viscosity was also measured by observing the decay of sound waves of various frequencies (Fig. 8). The level of agreement between simulation and the computed Boltzmann viscosity is generic: we see a systematic error of approximately 10 percent. Monte Carlo calculations of viscosities computed from microscopic correlation functions improve agreement with simulations to at least 3 percent and indicate that the Boltzmann description is not as accurate an analytic tool for the automaton as are microscopic correlation techniques. One would call this type of viscosity disagreement a Boltzmann-induced error. Other consistency checks between the automaton simulation and analytic predictions display the same level of agreement.

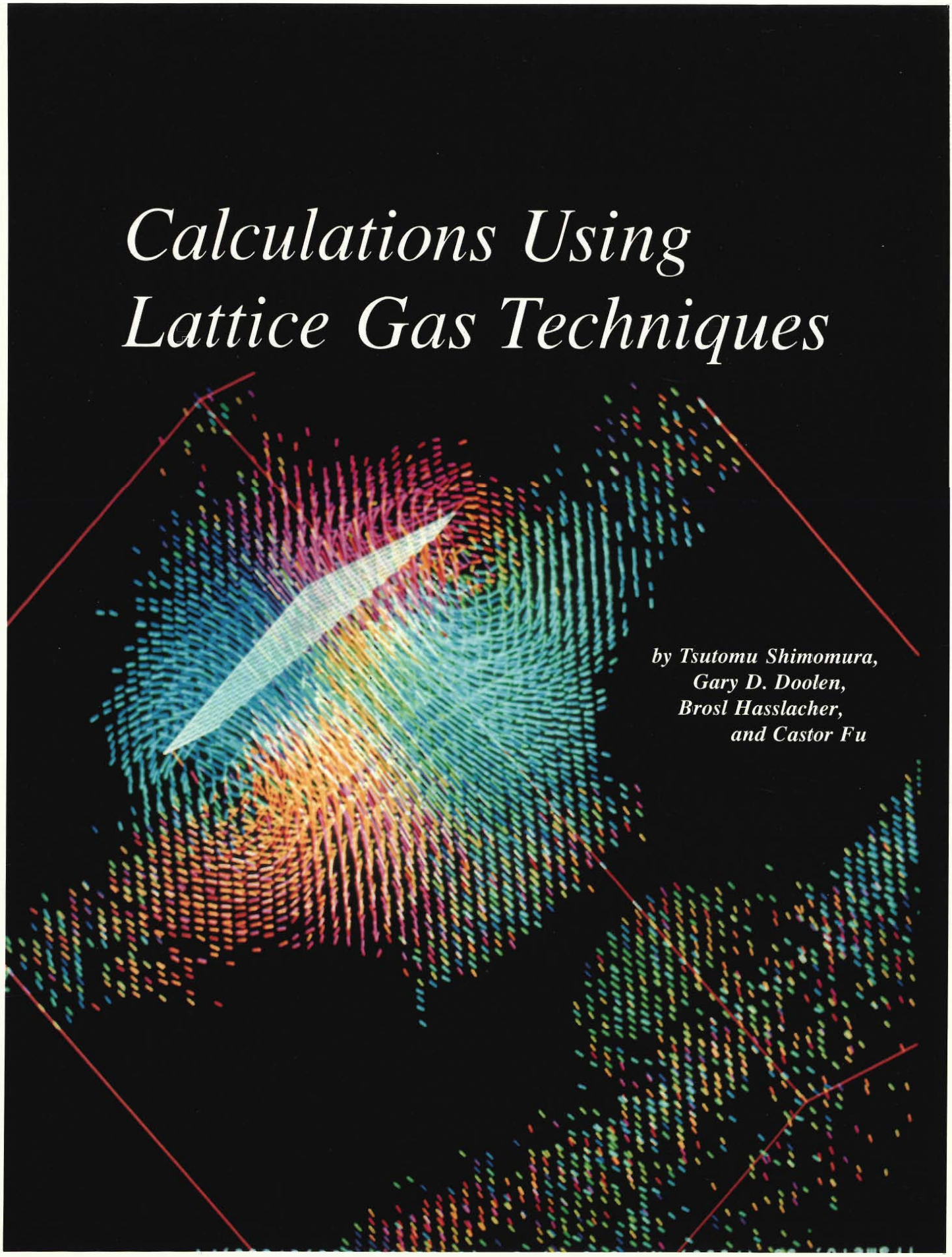
Detailed quantitative comparisons between conventional discretizations of the Navier-Stokes equations and lattice-gas simulations have yet to be done for several reasons. The simple lattice-gas automaton has a Fermi-Dirac distribution rather than the standard Maxwell-Boltzmann distribution. This difference alone causes deviations of $O(v^2)$ in the macrovelocity from standard results. For the same reason and unlike standard numerical spectral codes for fluid dynamics, the simple lattice-gas automaton has a velocity-dependent equation of state. A meaningful comparison between the two approaches requires adjusting the usual spectral codes to compute with a velocity-dependent equation of state. This rather considerable task has yet to be done. So far our simulations can be compared only to traditional two-dimensional computer simulations and analytic results derived from simple equations of state.

Some simple quantities such as the speed of sound and velocity profiles have been measured in the automaton model. The speed of sound agrees with predicted values and functional forms for channel velocity profiles and D'Arcy's law agree with calculations by standard methods. The automaton reaches local equilibrium in a few time steps and reaches global equilibrium at the maximum information-transmission speed, namely, at the speed of sound.

Simulations with the two-dimensional lattice-gas model hang together rather well as a simulator of Navier-Stokes dynamics. The method is accurate enough to test theoretical turbulent mechanisms at high Reynolds number and as a simulation tool for complex geometries, provided that velocity-dependent effects due to the Fermi nature of the automaton are correctly included. Automaton models can be designed to fit specific phenomena, and work along these lines is in progress.

Three-dimensional hydrodynamics is being simulated, both on serial and parallel machines, and early results show that we can easily simulate flows with Reynolds numbers of a few thousand. How accurately this model reproduces known instabilities and flows remains to be seen, but there is every reason to believe agreement will be good since the ingredients to evolve to Navier-Stokes dynamics are all present. We end Part II of this article with a graphical display of two- and three-dimensional simulations in "Calculations Using Lattice-Gas Techniques." My Los Alamos collaborators and I have accompanied this display with a summary of the known advantages and present limitations of lattice gas methods. (Part III begins on page 211.)

Calculations Using Lattice Gas Techniques



by Tsutomu Shimomura, Gary D. Doolen, Brosl Hasslacher, and Castor Fu

Over the last few years the tantalizing prospect of being able to perform hydrodynamic calculations orders-of-magnitude faster than present methods allow has prompted considerable interest in lattice gas techniques. A few dozen published papers have presented both advantages and disadvantages, and several groups have studied the possibilities of building computers specially designed for lattice gas calculations. Yet the hydrodynamics community remains generally skeptical toward this new approach. The question is often asked, "What calculations can be done with lattice gas techniques?" Enthusiasts respond that in principle the techniques are applicable to any calculation, adding cautiously that increased accuracy requires increased computational effort. Indeed, by adding more particle directions, more particles per site, more particle speeds, and more variety in the interparticle scattering rules, lattice gas methods can be tailored to achieve better and better accuracy. So the real problem is one of tradeoff: How much accuracy is gained by making lattice gas methods more complex, and what is the computational price of those complications? That problem has not yet been well studied. This paper and most of the research to date focus on the simplest lattice gas models in the hope that knowledge of them will give some insight into the essential issues.

We begin by examining a few of the features of the simple models. We then display results of some calculations. Finally, we conclude with a discussion of limitations of the simple models.

Features of Simple Lattice Gas Methods

We will discuss in some depth the memory efficiency and the parallelism of lattice gas methods, but first we will touch on their simplicity, stability, and ability to model complicated boundaries.

Computer codes for lattice gas methods are enormously simpler than those for other methods. Usually the essential parts of the code are contained in only a few dozen lines of FORTRAN. And those few lines of code are much less complicated than the several hundred lines of code normally required for two- and three-dimensional hydrodynamic calculations.

There are many hydrodynamic problems that cause most standard codes (such as finite-difference codes, spectral codes, and particle-in-cell codes) to crash. That is, the code simply stops running because the algorithm becomes unstable. Stability is not a problem with the codes for lattice gas methods. In addition, such methods conserve energy and momentum exactly, with no roundoff errors.

Boundary conditions are quite easy to implement for lattice gas methods, and they do not require much computer time. One simply chooses the cells to which boundary conditions apply and updates those cells in a slightly different way. One of three boundary conditions is commonly chosen: bounce-back, in which the directions of the reflected particles are simply reversed; specular, in which mirror-like reflection is simulated; or diffusive, in which the directions of the reflected particles are chosen randomly.

We consider next the memory efficiency of the lattice gas method. When the two-dimensional hydrodynamic lattice gas algorithm is programmed on a computer with a word length of, say, 64 bits (such as the Cray X-MP), two impressive efficiencies occur. The first arises because every single bit of memory is used equally effectively. Coined "bit democracy" by von Neumann, such efficient use of memory should be contrasted with that attainable in standard calculations, where each number requires a whole 64-bit word. The lattice gas is "bit democratic" because all that one needs to know is whether or not a particle with a given velocity direction exists in a given cell. Since the number of possible velocity directions is six and no two particles in the same cell can have the same direction, only six bits of information are needed to completely specify the state of a cell. Each of those six bits corresponds to one of the six directions and is set to 1 if the cell contains a particle with that direction and to 0 otherwise. Suppose we designate the six directions by A,B,C,D, E,F as shown on the next page. We associate each bit in the 64-bit word A with a different cell, say the first 64 cells in the first row. If the first cell contains (does not contain) a particle with direction A, we set the first bit in A to 1 (0). Similarly, we pack information about particles in the remaining 63 cells with direction A into the remaining 63 bits of A. The same scheme is used for the other five directions. Consequently, all the information for the first 64 cells in the first row is contained in the six words A, B, C, D, E, and F. Note that all bits are equally important and all are fully utilized.

To appreciate the significance of such efficient use of memory, consider how many cells can be specified in the solid-state storage device presently used with the Cray X-MP/416 at Los Alamos. That device stores 512,000,000 64-bit words. Since the necessary information for $10\frac{2}{3}$ cells can be stored in each word, the device can store information for about 5,000,000,000 cells, which corresponds to a two-dimensional lattice with 100,000 cells along one axis and 50,000 cells along the other. That number of cells is a few orders of magnitude greater than the number normally treated when other methods are used. (Although such high resolution may appear to be a significant advantage of the lattice gas method, some averaging over cells is required to obtain smooth results for physical quantities such as velocity and density.)

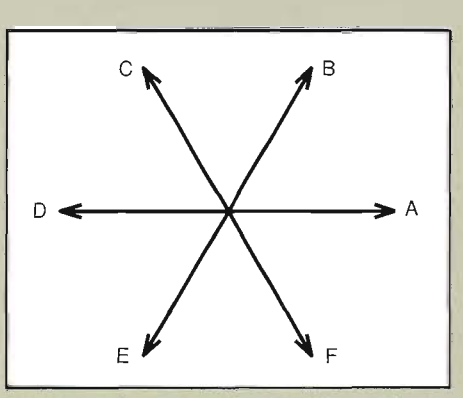
The second efficiency is related to the

fact that lattice gas operations are bit oriented rather than floating-point-number oriented and therefore execute more naturally on a computer. Most computers can carry out logic operations bit by bit. For example, the result of the logic operation AND on the 64-bit words A and B is a new 64-bit word in which the ith bit has a value of 1 only if the ith bits of both A and B have values of 1. Hence in one clock cycle a logic operation can be performed on information for 64 cells. Since a Cray X-MP/416 includes eight logical function units, information for 8 times 64, or 512, cells can be processed during each clock cycle, which lasts about 10 nanoseconds. Thus information for 51,200,000,000 cells can be processed each second. The two-dimensional lattice gas models used so far require from about thirty to one hundred logic operations to implement the scattering rules and about another dozen to move the particles to the next cells. So the number of cells that can be updated each second by logic operations is near 500,000,000. Cells can also be updated by table-lookup methods. The authors have a table-lookup code for three-dimensional hydrodynamics that processes about 30,000,000 cells per second.

A final feature of the lattice gas method is that the algorithm is inherently parallel. The rules for scattering particles within a cell depend only on the combination of particle directions in that cell. The scattering can be done by table lookup, in which one creates and uses a table of scattering results—one for each possible cell configuration. Or it can be done by logic operations.

Using Lattice Gas Methods To Approximate Hydrodynamics

In August 1985 Frisch, Hasslacher, and Pomeau demonstrated that one can approximate solutions to the Navier-Stokes equations by using lattice gas methods,



but their demonstration applied only to low-velocity incompressible flows near equilibrium. No one knew whether more interesting flows could be approximated. Consequently, computer codes were written to determine the region of validity of the lattice gas method. Results of some of the first simulations done at Los Alamos and of some later simulations are shown in Figs. 1 through 6. (Most of the early calculations were done on a Celerity computer, and the displays were done on a Sun workstation.) All the results indicate qualitatively correct fluid behavior.

Figure 1a demonstrates that a stable trailing vortex pattern develops in a two-dimensional lattice gas flowing past a plate. Figure 1b shows that without a three-particle scattering rule, which removes the spurious conservation of momentum along each line of particles, no vortex develops. (Scattering rules are described in Part II of the main text.)

Figure 2 shows that stable vortices develop in a lattice gas at the interface between fluids moving in opposite directions. The Kelvin-Helmholtz instability is known to initiate such vortices. The fact that lattice gas methods could simulate vortex evolution was reassuring and caused several scientists to begin to study the new method.

Figure 3 shows the complicated wake that develops behind a V-shaped wedge in a uniform-velocity flow.

Figure 4 shows the periodic oscillation of a low-velocity wake behind a cylinder. With a Reynolds number of 76, the flow has a stable period of oscillation that slowly grows to its asymptotic limit.

Figure 5 shows a flow with a higher Reynolds number past an ellipse. The wake here becomes chaotic and quite sensitive to details of the flow.

Figure 6 shows views of a three-dimensional flow around a square plate, which was one of the first results from Los Alamos in three-dimensional lattice gas hydrodynamic simulations.

Rivet and Frisch and other French scientists have developed a similar code that measures the kinematic shear viscosity numerically; the results compare well with theoretical predictions (see Fig. 8 in the main text).

The lattice gas calculations of a group at the University of Chicago (Kadanoff, McNamara, and Zanetti) for two-dimensional flow through a channel (Fig. 7 of the main text) agree with the known parabolic velocity profile for low-velocity channel flows.

The above calculations, and many others, have established some confidence that qualitative features of hydrodynamic flows are simulated by lattice-gas methods. Problems encountered in detailed comparisons with other types of calculations are discussed in the next section.

Limitations of Simple Lattice Gas Models

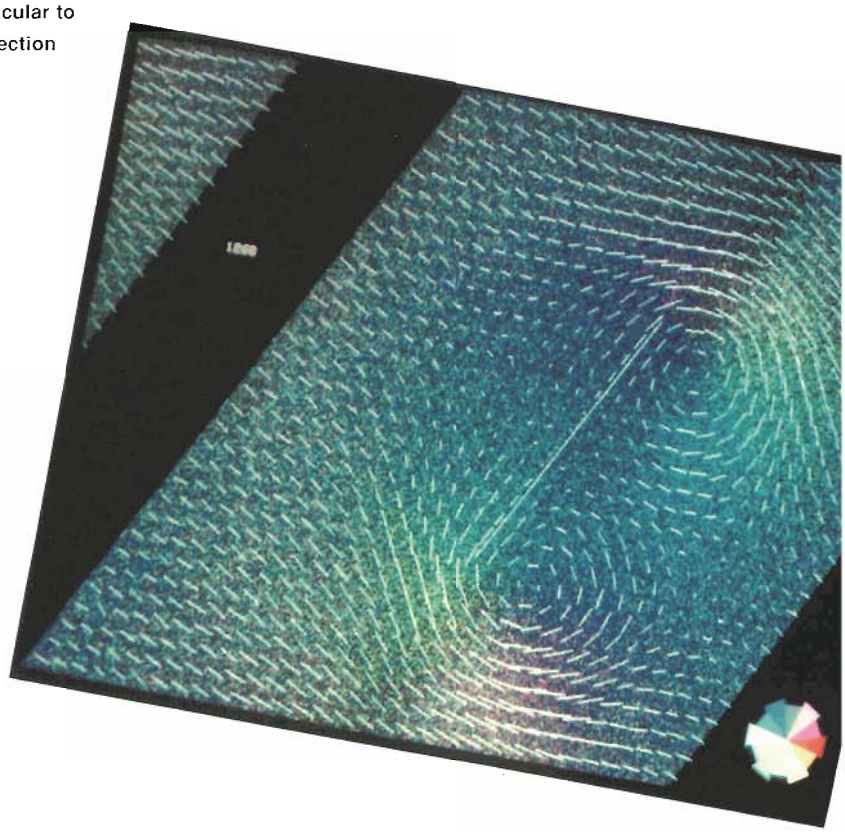
As we discussed earlier, lattice gas methods can be made more accurate by making them more complicated—by, for example, adding more velocity directions and magnitudes. But the added complications degrade the efficiency. We mention in this section some of the difficulties (associated with limited range of speed, velocity dependence of the equation of state, and noisy results) encountered in the simplest lattice-gas models.

The limited range of flow velocities is inherent in a model that assumes a

continued on page 210

Flow Past a Plate

Fig. 1a. Flow past a plate with periodic boundary conditions. This simulation, which was done in September 1985, shows vortices forming behind the plate. The average flow velocity has a magnitude of 0.2 lattice sites per time step and is perpendicular to the plate, pointing to the lower right. The direction of the flow velocity is color-coded.



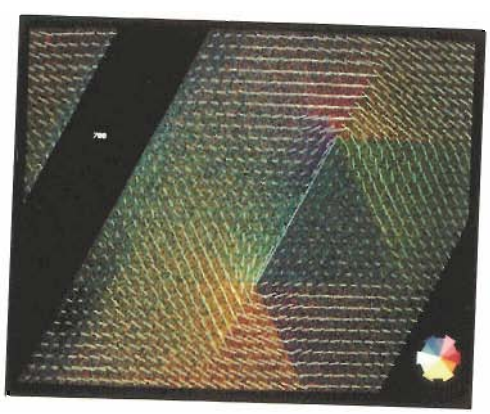
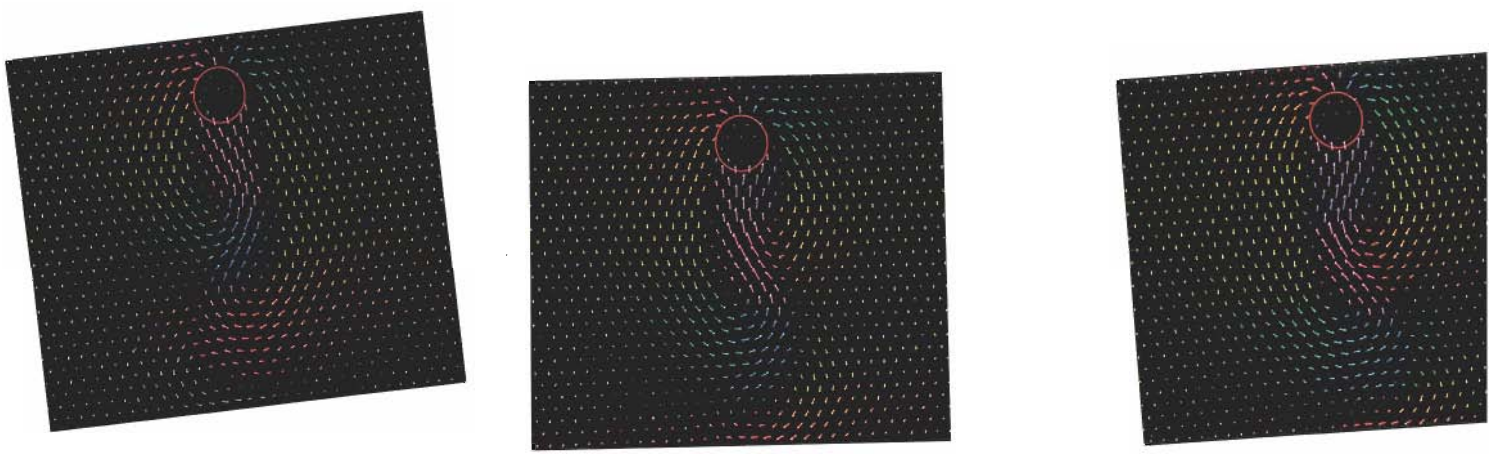


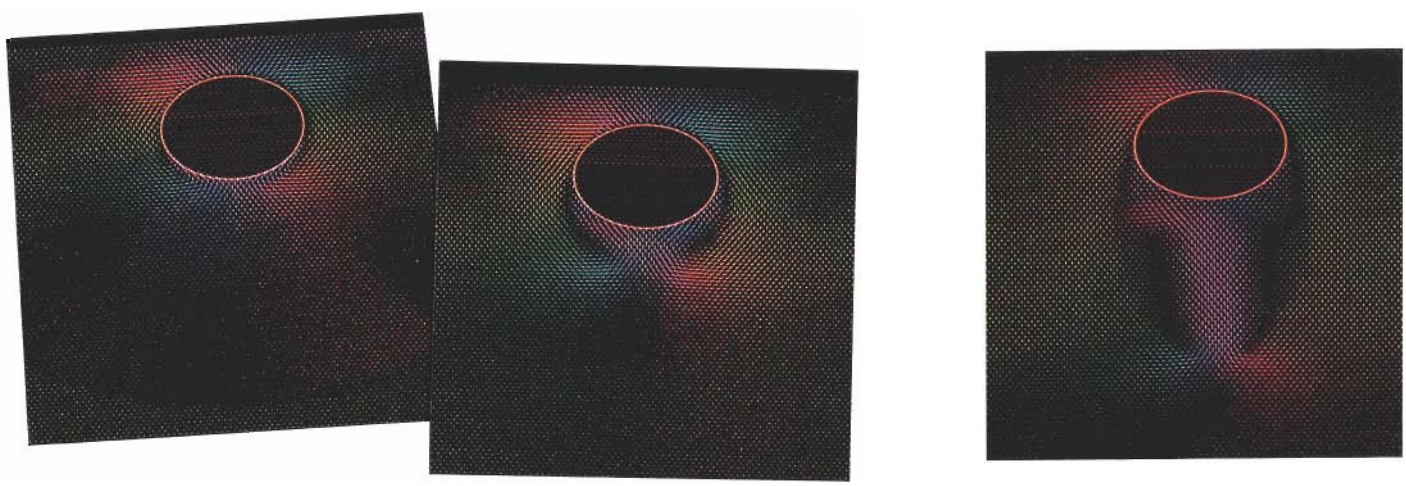
Fig. 1b. The same simulation as that described in Fig. 1a but with no three-body scattering rule. As a result, spurious laws of conservation of momentum along the lines of the grid prevent the development of hydrodynamics.

Kelvin-Helmholtz Instability

Fig. 2. A Kelvin-Helmholtz instability develops into vortices from initially opposing flows past a sinusoidal interface that is removed at $t = 0$. Periodic boundary conditions apply. For this simulation about 10,000,000 particles and 14,000,000 cells were used.

Flow Past a Wedge





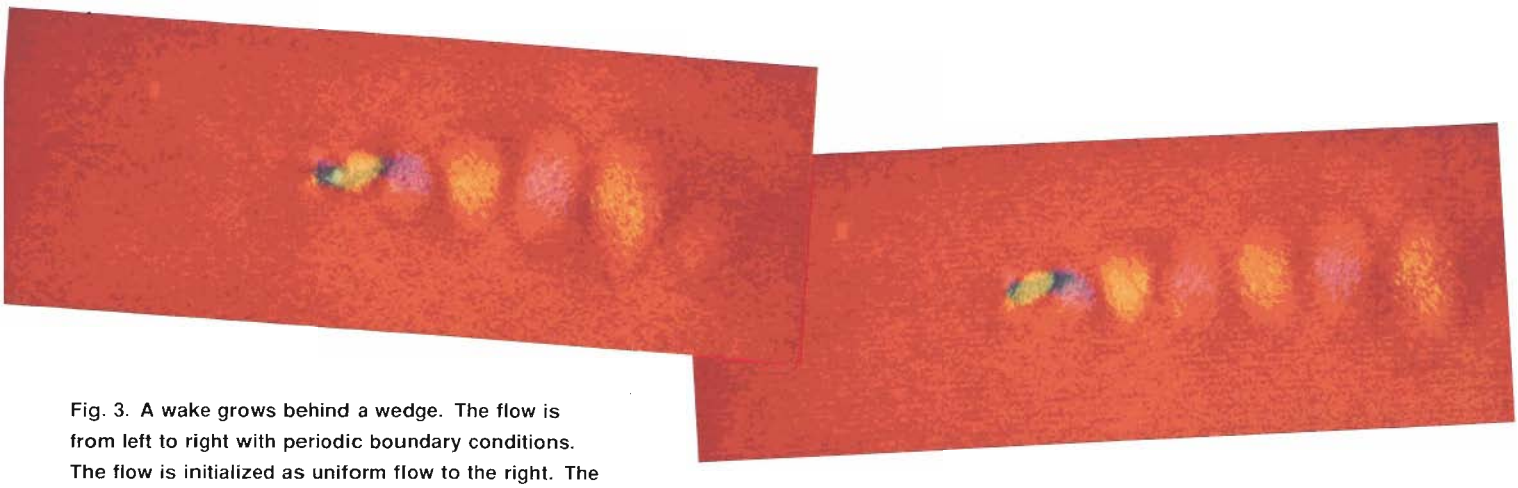
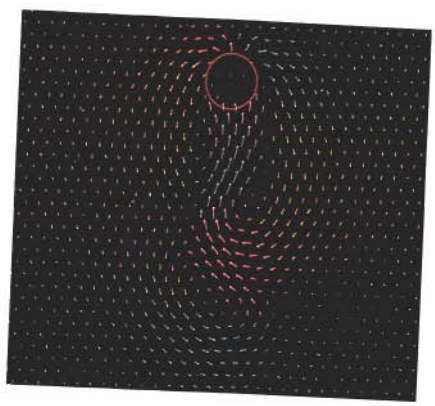
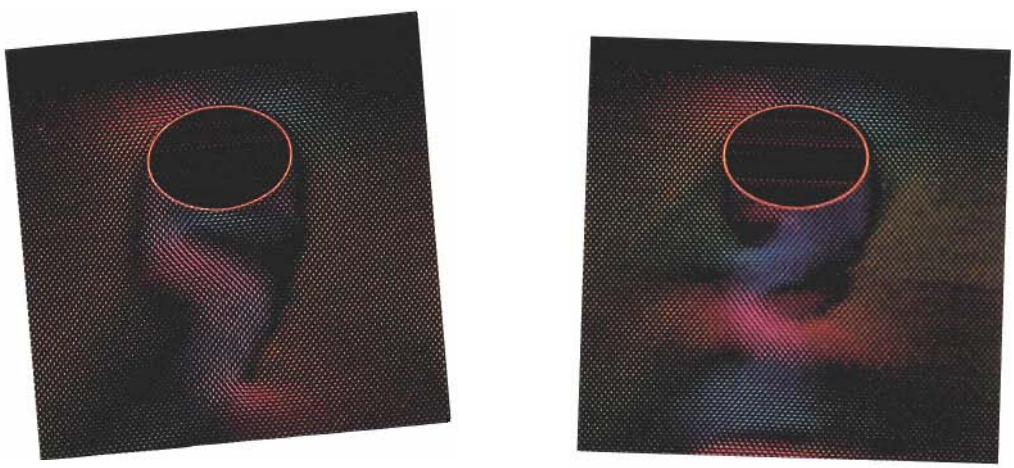


Fig. 3. A wake grows behind a wedge. The flow is from left to right with periodic boundary conditions. The flow is initialized as uniform flow to the right. The wedge is inserted at $t = 0$. Then vortices grow and are carried downstream. For this simulation 20 million particles and 16 million cells were used.



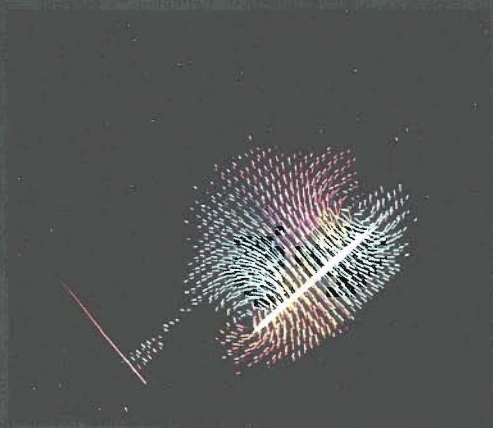
Flow Past a Cylinder

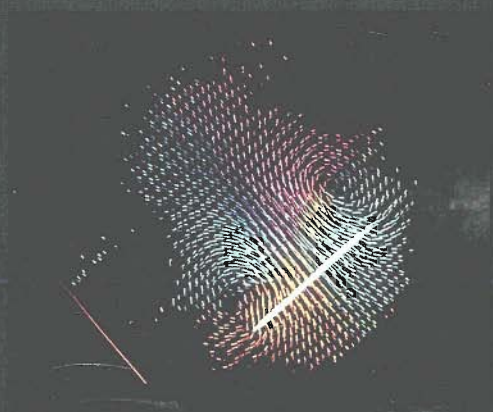
Fig. 4. Low-velocity flow (from top to bottom) past a cylinder creates a periodically oscillating wake. Four snapshots from one period of the oscillation are shown. In this simulation, which has periodic right and left boundaries, 1.4 million particles flowed through 1 million cells. The flow was initially uniform.

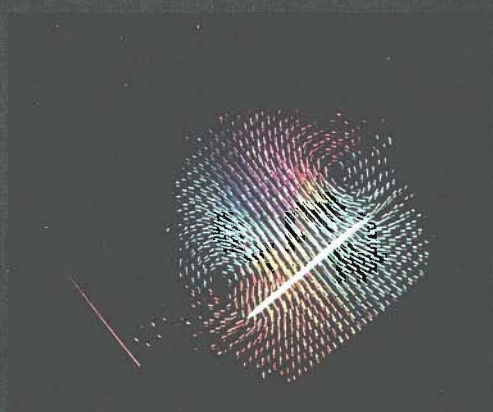


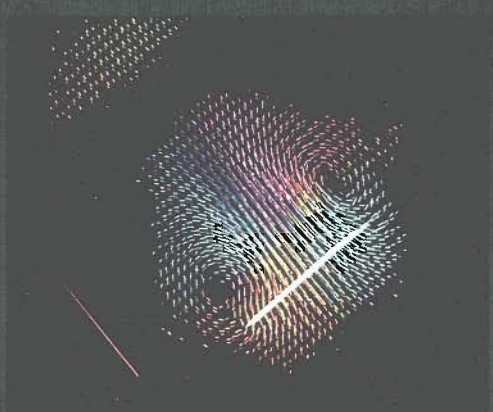
Turbulent Wake

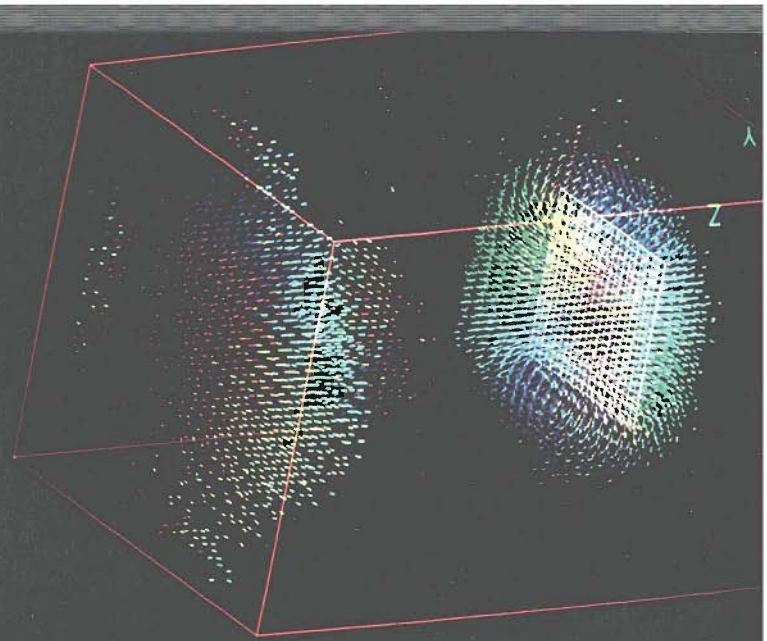
Fig. 5. A turbulent wake grows behind an ellipse being dragged through a fluid consisting of 11 million particles and 8 million cells. The ellipse is composed of about 2400 cells in which the velocity directions of the entering particles are reversed. The flow has periodic right and left boundaries. (An infinite sequence of equivalent ellipses exists to the left and right of the frame shown.) The Reynolds number in the flow is 1021.





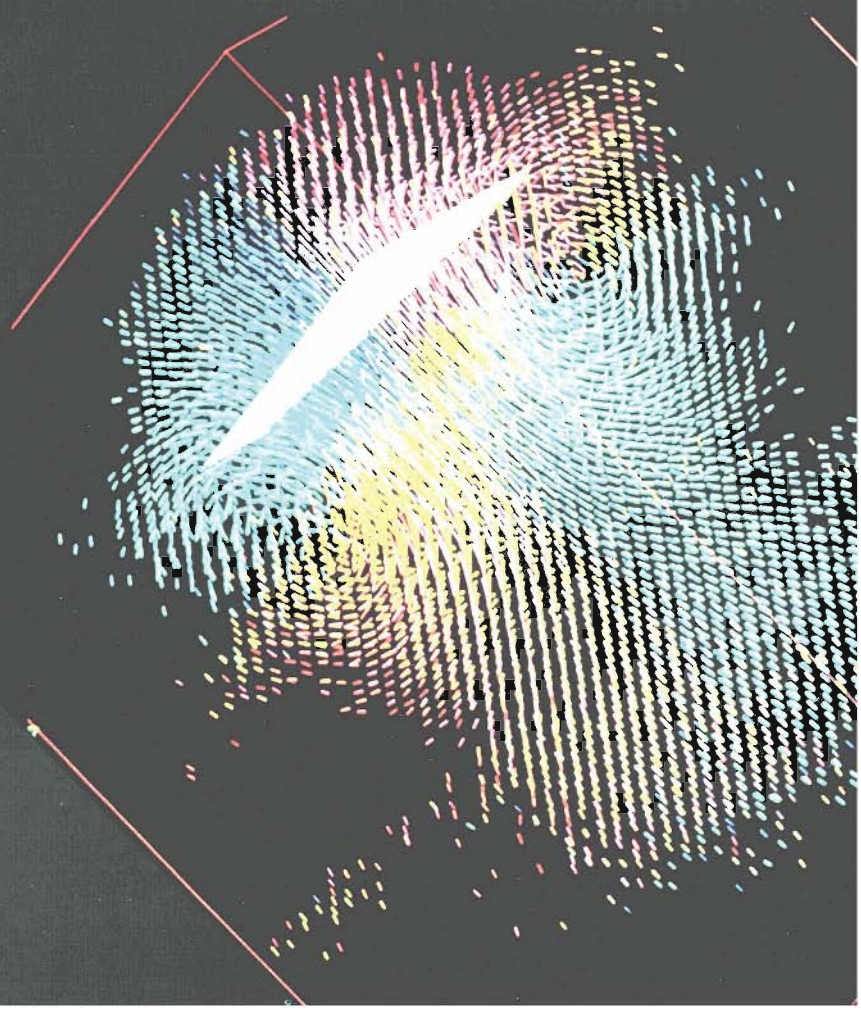


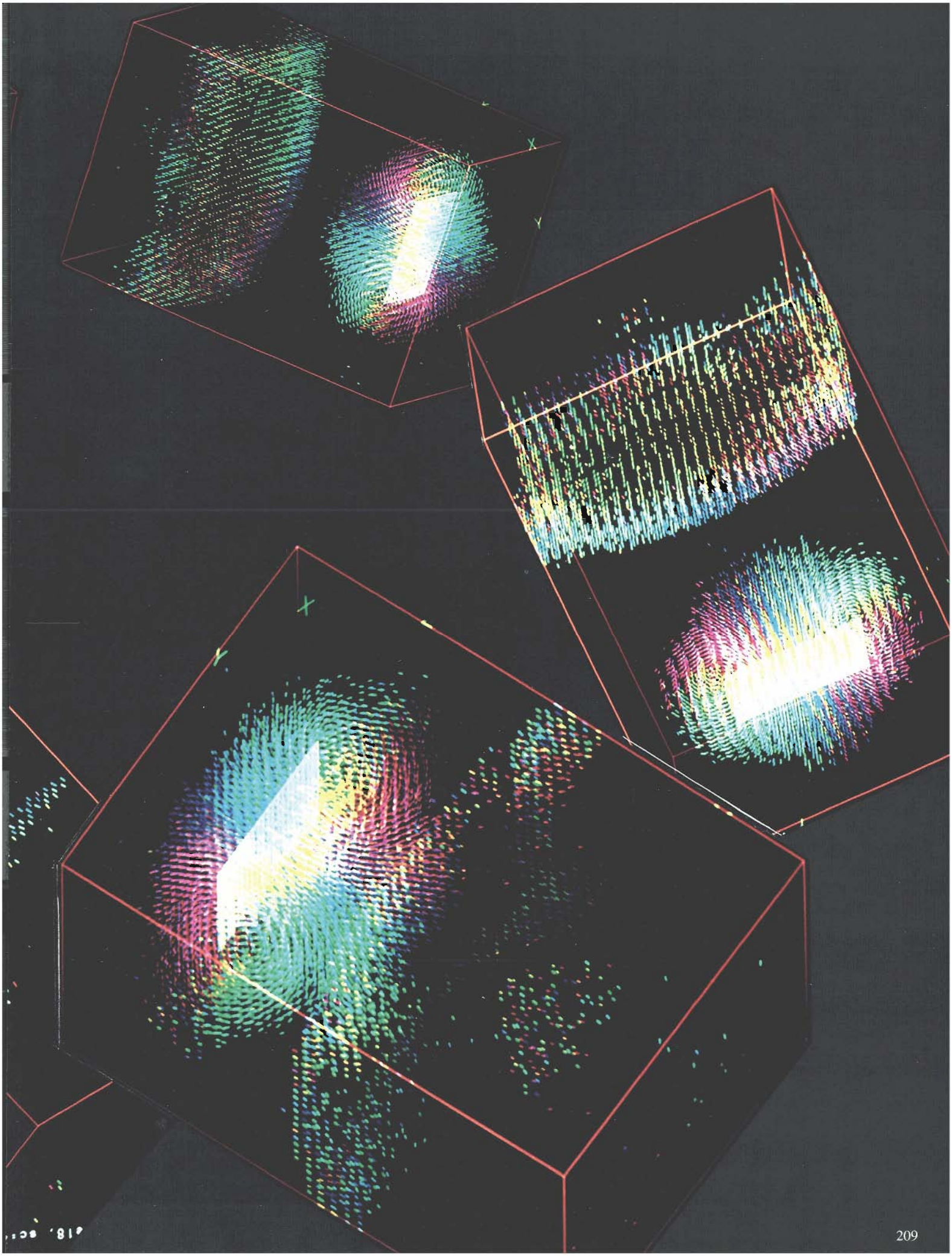




3-D Flow Past a Plate

Fig. 6. Cross sections (left) and three-dimensional views show the development of vortices behind a square plate and a spherical sound wave that propagates through the system. Mean flow was subtracted to highlight the dynamics of the flow. For this simulation 80 million particles and 10 million face- centered-cubic cells were used.




X
Y
X
Y

continued from page 203

single speed for all particles. The sound speed in such models can be shown to be about two-thirds of the particle speed. Hence flows in which the Mach number (flow speed divided by sound speed) is greater than 1.5 cannot be simulated. This difficulty is avoided by adding particles with a variety of speeds.

The limited range of velocities also restricts the allowed range of Reynolds numbers. For small Reynolds numbers (0 to 1000) the flow is smooth, for moderate Reynolds numbers (2000 to 6000) some turbulence is observed, and for high Reynolds numbers (10,000 to 10,000,000) extreme turbulence occurs. Since the effective viscosity, ν, is typically about 0.2 in two-dimensional problems, the Reynolds number scales with the characteristic length, l, allowed by computer memory. Currently the upper bound on l is of the order of 100,000.

The velocity dependence of the equation of state is unusual and is a consequence of the inherent Fermi-Dirac distribution of the lattice gas (see the section on Theoretical Analysis of the Discrete Lattice Gas in the main text). The low-velocity equation of state for a lattice gas can be written as $p = \frac{1}{2}\rho\left(1 - \frac{1}{2}v^2\right)$, where p is the pressure, ρ is the density, and v is the flow speed. Thus, for constant-pressure flows, regions of higher velocity flows have higher densities.

The velocity dependence of the equation of state is related to the fact that lattice gas models lack Galilean invariance. The standard Navier-Stokes equation for incompressible fluids is

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \nu \nabla^2 \mathbf{v}.$$

But in the incompressible, low-velocity limit the single-speed hexagonal lattice gas follows the equation

$$\frac{\partial \mathbf{v}}{\partial t} + g(\rho)\mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \nu \nabla^2 \mathbf{v},$$

where

$$g(\rho) = \frac{3 - \rho}{6 - \rho}$$

and ρ is the average number of particles per cell. The extra factor $g(\rho)$ requires special treatment. The conventional way to adjust for the fact that $g(\rho)$ does not equal unity (as it does in the Navier-Stokes equation) is to simply scale the time, t, and the viscosity, ν, by the factor $g(\rho)$ as follows: $t' = g(\rho)t$ and $\nu' = \nu/g(\rho)$. (The pressure must also be scaled.) Hence a density-dependent scaling of the time, the viscosity, and the pressure is required to bring the lattice gas model into a form that closely approximates the hydrodynamics of incompressible fluids in the low-velocity limit.

Finally, the discreteness of the lattice gas approximation introduces noise into the results. One method of smoothing the results for comparison with other methods is to average in space and time. In practice, spatial averages are taken over 64, 256, 512, or 1024 neighboring cells for time-dependent flows in two dimensions. For steady-state flows, time averaging is done. The details of noise reduction are complicated, but they must be addressed in each comparison calculation. The presence of noise is both a virtue and a defect. Noise ensures that only robust (that is, physical) singularities survive, whereas in standard codes, which are subject to less noise, mathematical artifacts can produce singularities. On the other hand, the noise in the model can trigger instabilities.

Conclusion

In the last few years lattice gas methods have been shown to simulate the qualitative features of hydrodynamic flows in two and three dimensions. Precise comparisons with other methods of calculation remain to be done, but it is believed that the accuracy of the lattice gas method can be increased by making the models more complicated. But how complicated they have to be to obtain the desired accuracy is an unanswered question.

Calculations based on the simple models are extremely fast and can be made several orders-of-magnitude faster by using special-purpose computers, but the models must be extended to get quantitative results with an accuracy greater than 1 percent. Significant research remains to be done to determine the accuracy of a given lattice gas method for a given flow problem. ■

Note added in proof: Recently Kadanoff, McNamara, and Zanetti reported precise comparisons between theoretical predictions and lattice gas simulations (University of Chicago preprint, October 1987). They used a seven-bit hexagonal model on a small automaton universe to simulate forced two-dimensional channel flow for long times. Three tests were used to probe the hydrodynamic and statistical-mechanical behavior of the model. The tests determined (1) the profile of momentum density in the channel, (2) the equation of state given by the statistical mechanics of the system, and (3) the logarithmic divergence in the viscosity (a famous effect in two- dimensional hydrodynamics and a deep test of the accuracy of the model in the strong nonlinear regime).

The results were impressive. First, to within the accuracy of the simulation, there is no discrepancy between the parabolic velocity profile predicted by macroscopic theory and the lattice gas simulation data. Second, the equation of state derived from theory fits the simulation data to better than 1 percent. Finally, the measured logarithmic divergence in the viscosity as a function of channel width agrees with prediction. These results are at least one order of magnitude more accurate than any previously reported calculations.

PART III THE PROMISE OF LATTICE GAS METHODS

In the sidebar "Calculations Using Lattice Gas Techniques" we displayed the results of generalizing the simple hexagonal model to three dimensions. Here, in the last part of the article, we will discuss numerous ways to extend and adapt the simple model. In particular, we emphasize its role as a paradigm for parallel computing.

Adjusting the Model To Fit the Phenomenon

There are several reasons for altering the geometry and rule set of the fundamental hexagonal model. To understand the mathematical physics of lattice gases, we need to know the class of functionally equivalent models, namely those models with different geometries and rules that produce the same dynamics in the same parameter range.

To explore turbulent mechanisms in fluids, the Reynolds number must be significantly higher than for smooth flow, so models must be developed that increase the Reynolds number in some way. The most straightforward method, other than increasing the size of the simulation universe, is to lower the effective mean free path in the gas. This lowers the viscosity and the Reynolds number rises in inverse proportion. Increasing the Reynolds number is also important for practical applications. In "Reynolds Number and Lattice Gas Calculations" we discuss the computational storage and work needed to simulate high-Reynolds-number flows with cellular automata.

To apply lattice gas methods to systems such as plasmas, we need to develop models that can support widely separated time scales appropriate to, for example, both photon and hydrodynamical modes. The original hexagonal model on a single lattice cannot do so in any natural way but must be modified to include several lattices or the equivalent (see below).

Within the class of fluids, problems involving gravity on the gas, multi-component fluids, gases of varying density, and gases that undergo generalized chemical reactions require variations of the hexagonal model. Once into the subject of applications rather than fundamental statistical mechanics, there is an endless industry in devising clever gases that can simulate the dynamics of a problem effectively.

We outline some of the possible extensions to the hexagonal gas, but do so only to give an overview of this developing field. Nothing fundamental changes by making the gas more complex. This model is very much like a language. We can build compound sentences and paragraphs out of simple sentences, but it does not change the fundamental rules by which the language works.

The obvious alterations to the hexagonal model are listed below. They comprise almost a complete list of what can be done in two dimensions, since a lattice gas model contains only a few adjustable structural elements.

Indistinguishable particles can be colored to create distinguishable species in the gas, and the collision rules can be appropriately modified. Rules can be weighted to different outcomes; for example, one can create a chiral gas (left- or right-handed) by biasing collisions to make them asymmetric. In three dimensions there is an instability at any Reynolds number caused by lack of microscopic parity, so the chiral gas is an important model for simulating this instability.

At the next order of complexity, multi-speed particles can be introduced, either alone or with changes in geometry. The simplest example is a square neighborhood in two dimensions in which the collision domain is enlarged to include next-to-nearest neighbors, and a diagonal particle with speed $\sqrt{2}$ is introduced to force an isotropic lattice gas. In general, any lattice model with only two-body collisions and a single speed will contain spurious conservation laws. But if multiple speeds are allowed, models with binary collisions can maintain isotropy. In other words, models with multiple speeds are equivalent to single-speed models with a higher order rotation group and extended collision sets. Many variations are possible and each can be designed to a problem where it has a special advantage.

Finally, colored multiple-speed mod-

els are in general equivalent to single-species models operating on separate lattices. Colored collision rules couple the lattices so that information can be transferred between them at different time scales. Certain statistical-mechanical phenomena such as phase transitions can be done this way.

By altering the rule domain and adding gas species with distinct speeds, it is possible to add independent energy conservation. This allows one to tune gas models to different equations of state. Again, we gain no fundamental insight into the development of large collective models by doing so, but it is useful for applications.

In using these lattice gas variations to construct models of complex phenomena, we can proceed in two directions. The first direction is to study whether or not complex systems with several types of coupled dynamics are described by skeletal gases. Can complex chemical reactions in fluids and gases, for example, be simulated by adding collision rules operating on colored multi-speed lattice gases? Complex chemistry is set up in the gas in outline form, as a gross scheme of closed sets of interaction rules. The same idea might be used for plasmas. From a theoretical viewpoint one wants to study how much of the known dynamics of such systems is reproduced by a skeletal gas; consequently both qualitative and quantitative results are important.

Exploring Fundamental Questions. Models of complex gas or fluid systems, like other lattice gas descriptions, may either be a minimalist description of microphysics or simply have no relation to microphysics other than a mechanism for carrying known conservation laws and reactions. We can always consider such gas models to be pure computers, where we fit the wiring, or architecture, to the problem, in the same fashion that ordinary discretization schemes have no relation to the microphysics of the problem. However for lattice gas models, or cellular-

continued on page 214

REYNOLDS NUMBER *and* Lattice Gas Calculations

The only model-dependent coupling constant in the Navier-Stokes equation is the viscosity. Its main role in lattice gas computations is its influence on the Reynolds number, an important scaling concept for flows. Given a system with a fixed intrinsic global length scale, such as the size of a pipe or box, and given a flow, then the Reynolds number can be thought of as the ratio of a typical macrodynamic time scale to a time scale set by elementary molecular processes in the kinetic model.

Reynolds numbers characterize the behavior of flows in general, irrespective of whether the system is a fluid or a gas. At high enough Reynolds numbers turbulence begins, and turbulence quickly loses all memory of molecular structure, becoming universal across liquids and gases. For this reason and because many interesting physical and mathematical phenomena happen in turbulent regimes, it is important to be able to reach these Reynolds numbers in realistic simulations without incurring a large amount of computational work or storage.

Some simple arguments based on dimensional analysis and phenomenological theories of turbulence indicate, at first glance, that any cellular automaton model has a high cost in computer resources when simulating high-Reynolds-number flows. These arguments appeared in the first paper on the subject (Frisch, Hasslacher, and Pomeau 1987) and were later elaborated on by other authors. We will go through the derivation of some of the more severe constraints on simulating high-Reynolds-number flows with

cellular automata, then discuss some possible ways out, and finally estimate the seriousness of the situation for a realistic large-scale simulation.

The turbulent regime has many length scales, bounded above by the length of the simulation box and below by the scale at which turbulent dynamics degenerates into pure dissipation, the so-called dissipation scale. We focus on these extreme scales and, with a few definitions, derive a bound on the computational storage and work needed for simulating high-Reynolds-number flows with cellular automata.

The Reynolds number R is usually defined not in terms of time but simply as $R = vL/\nu$ where L is a characteristic length, v is characteristic speed, and ν is the kinematic shear viscosity. One sees immediately why calculating viscosity functions for particular models is important. It is the only variable one can adjust in a flow problem, given a fixed flow in a fixed geometry. First, we calculate a rough upper bound on Reynolds numbers attainable with lattice models. If the speed of sound in the lattice gas is c_s and the spacing between lattice nodes is ℓ, then by definition the kinematic viscosity $\nu \geq c_s\ell$. Now viscosity estimated this way must agree with that fixed by the scale of hydrodynamic modes. Given a global length L and a global velocity V associated with these modes, $R = VL/\nu$ at best. In terms of the Mach number ($M = V/c_s$), the Reynolds number is equal to ML/ℓ. But M also characterizes fluid flow, and L and ℓ are model-dependent. In a lattice gas we can relate the ratio L/ℓ to the number of nodes in the gas simulator, namely $n = (L/\ell)^d$, where d is the space dimension of the model. Therefore, the number of nodes in a lattice model must grow at least as $n \sim (R/M)^d$. Computational work is the number of lattice nodes per time step multiplied by the number of time steps required to resolve hydrodynamical features. This is $L/\ell M$ steps (to cross the hydrodynamical feature at the given Mach number), and so we find the computational work is of order R^{d+1}/M^{d+2}. For a so-called normal simulation based on the usual ways of discretizing the Navier-Stokes equation, the growth in storage is roughly proportional to one power lower in the Reynolds number than the growth in storage for the lattice gas. So at first it seems that simulating high-Reynolds-number flows by lattice gas techniques is costly compared to ordinary methods.

This argument is not only approximate; it is also tricky and must be applied with great care. The normal way of simulating flows escapes power-law penalties by cutting off degrees of freedom at the turbulence-dissipation scale, which the lattice gas does not do. The gas computes within these scales and so wastes computational resources for some problems. Actually computation of these very small scales is the source of the noisy character of the gas and is responsible for its power to avoid spurious mathematical singularities. One way around this is to find an effective gas with new collision rules for which the dissipation length scales are averaged out. A possible technique uses the renormalization group, but it is useful only if the effective gas is not too complex and has the attributes that made the original gas attractive, including locality. Work is going on at present to explore this possibility, and it seems likely that some such method will be developed.

The more serious consideration is what happens in a realistic large-scale simulation, and here we will find the lattice gas does very well indeed.

First, we note that a dissipation length l_d with the behavior $l_d \to \infty$ as $R \to \infty$ is actually required to guarantee the scale separation between the lattice spacing and the hydrodynamic modes that is necessary to develop hydrodynamic behavior.

The actual Reynolds number in lattice gas models is much more complex than in normal fluid models. An accurate form is $R = Lvg(\rho_0)/\nu(\rho_0)$, where v is an averaged velocity and the fundamental unit of distance (the lattice spacing ℓ) and the fundamental unit of time (the speed required to traverse the lattice spacing ℓ) have been set to 1. To remain nearly incompressible, the velocities in the model should remain small compared to the speed of sound c_s, but c_s in lattice gases is model-dependent. So we factor the Reynolds number into model-dependent and invariant factors this way: we define $\hat{R}(\rho_0) = c_s\left(g(\rho_0)/\nu(\rho_0)\right)$ so that $R = ML\hat{R}(\rho_0)$. The value of $\hat{R}$ depends critically on the model used. In two dimensions it ranges from 0.39 to about 6 times that, depending on the amount of the state table we want to include. For the three-dimensional projection of the four-dimensional model, it is known that $\hat{R}$ is about 9.

By repeating essentially the same dimensional arguments, only more carefully, we find that the dissipation length $l_d = (M\hat{R})^{-1}R^{-1/2}$ for two dimensions and $l_d = (M\hat{R})^{-1}R^{-1/4}$ for three dimensions.

For a typical simulation in three dimensions, we take $M = 0.3$ for incompressibility, $\hat{R} \sim 9$, and $L = 10^3$, which is a large simulation, possible only on the largest Cray-class machines. Then l_d is about three lattice spacings, and the simulation wastes very little computational power. The subtle point is that the highly model-dependent factor $\hat{R}$ is not of order 1, as is usually estimated. It depends critically on the complexity of the collision set, going up a factor of 20 from the elementary hexagonal model in two dimensions to the projected four-dimensional case with an optimal collision table.

There is a great deal of work to be done on the high-Reynolds-number problem, but it is clear that the situation is complicated and rich in possibilities for evading simple dimensional arguments. ■

continued from page 212

automaton models in general, there always seems to be a deep relation between the abstract computer embodying the gas algorithm for a physical problem and the mathematical physics of the system itself.

This duality property is an important one, and it is not well understood. One of the main aims of lattice gas theory is to make the underlying mathematics of dynamical evolution clearer by providing a new perspective on it. One would, for example, like to know the class of all lattice gas systems that evolve to a dynamics that is, in an appropriate sense, nearby the dynamics actually evolved by nature. Doing this will allow us to isolate what is common to such systems and identify universal mathematical mechanisms.

Engineering Design Applications. The second direction of study is highly applied. In most engineering-design situations with complicated systems, one would like to know first the general qualitative dynamical behavior taking place in some rather involved geometry and then some rough numerics. Given both, one can plot out the zoo of dynamical development within a design problem. Usually, one does not know what kinds of phenomena can occur as a parameter in the system varies. Analytic methods are either unavailable, hard to compute by traditional methods, or simply break down. Estimating phenomena by scaling or arguments depending on order-of-magnitude dimensional analysis is often inaccurate or yields insufficient information. As a result, a large amount of expensive and scarce supercomputer time is used just to scan the parameter space of a system.

Lattice gas models can perform such tasks efficiently, since they simulate at the same speed whether the geometry and system are simple or complex. Complicated geometries and boundary conditions for massively parallel lattice gas simulators involve only altering collision rules in a region. This is easily coded and can be done interactively with a little investment in expert systems. There is no question that for complex design problems, lattice gas methods can be made into powerful design tools.

Beyond Two Dimensions

In two dimensions there exists a single-speed skeletal model for fluid dynamics with a regular lattice geometry. It relies on the existence of a complete tiling of the plane by a domain of sufficiently high symmetry to guarantee the isotropy of macroscopic modes in the model. In three dimensions this is not the case, for the minimum appropriate domain symmetry is icosahedral and such polyhedra do not tile three-space. If we are willing to introduce multiple-speed models, there may exist a model with high enough rotational symmetry, as in the square model with nearest and next-to-nearest neighbor interaction in two dimensions, but it is not easy to find and may not be efficient for simulations.

A tactic for developing an enlarged-neighborhood, three-dimensional model, which still admits a regular lattice, is to notice that the number of regular polyhedra as a function of dimension has a maximum in four dimensions. Examination of the face-centered four-dimensional hypercube shows that a single-speed model connected to each of twenty-four nearest neighbors has exactly the right invariance group to guarantee isotropy in four dimensions. So four-dimensional single-speed models exist on a regular tiling. Three-dimensional, or regular, hydrodynamics can be recovered by taking a thin one-site slice of the four-dimensional case, where the edges of the slice are identified. Projecting such a scheme into three-dimensional space generates a two-speed model with nearest and next-to-nearest neighbor interactions of the sort guaranteed to produce three-dimensional Navier-Stokes dynamics.

Such models are straightforward extensions of all the ideas present in the two-dimensional case and are being simulated presently on large Cray-class machines and the Connection Machine 2. Preliminary results show good agreement with standard computations at least for low Reynolds numbers. In particular, simulation of Taylor-Green vortices at a Reynolds number of about 100 on a $(128)^3$ universe (a three-dimensional cube with 128 cells in each direction) agrees with spectral methods to within 1 percent, the error being limited by Monte Carlo noise. The ultimate comparison is against laboratory fluid-flow experiments. As displayed at the end of Part II, three-dimensional flows around flat plates have also been done.

A more intriguing strategy is to give up the idea of a regular lattice. Physical systems are much more like a lattice with nodes laid down at random. At present, we don't know how to analyze such lattices, but an approximation can be given that is intermediate between regular and random grids. Quasi-tilings are sets of objects that completely tile space but the grids they generate are not periodic. Locally, various types of rotation symmetry can be designed into such lattices, and in three dimensions there exists such a quasi-tiling that has icosahedral symmetry everywhere. The beauty of quasi-tilings is that they can all be obtained by simple slices through hypercubes in the appropriate dimension. For three dimensions the parent hypercube is six-dimensional.

The idea is to run an automaton model containing the conservation laws with as simple a rule set as possible on the six-dimensional cube and then take an appropriately oriented three-dimensional slice out of the cube so arranged as to generate the icosahedral quasi-tiling. Since we only examine averaged quantities, it is enough to do all the averaging in six dimensions along the quasi-slice and image the results. By such a method we guarantee exact isotropy everywhere in three

dimensions and avoid computing directly on the extremely complex lattices that the quasi-tiling generates. Ultimately, one would like to compute on truly random lattices, but for now there is no simple way of doing that efficiently.

The simple four-dimensional model is a good example of the limits of present super-computer power. It is just barely tolerable to run a $(1000)^3$ universe at a Reynolds number of order a few thousand on the largest existing Cray's. It is far more efficient to compute in large parallel arrays with rather inexpensive custom machines, either embedded in an existing parallel architecture or on one designed especially for this class of problems.

Lattice Gases as Parallel Computers

Let us review the essential features of a lattice gas. The first property is the totally discrete nature of the description: The gas is updated in discrete time steps, lattice gas elements can only lie on discrete space points arranged in a space-filling network or grid, velocities can also have only discrete values and these are usually aligned with the grid directions, and the state of each lattice gas site is described by a small number of discrete bits instead of continuous values.

The second crucial property is the local deterministic rules for updating the array in space and time. The value of a site depends only on the values of a few local neighbors so there is no need for information to propagate across the lattice universe in a single step. Therefore, there is no requirement for a hardwired interconnection of distant sites on the grid.

The third element is the Boolean nature of the updating rules. The evolution of a lattice gas can be done by a series of purely Boolean operations, without ever computing with radix arithmetic.

To a computer architect, we have just described the properties of an ideal concurrent, or parallel, computer. The identical nature of particles and the locality of the rules for updating make it natural to update all sites in one operation—this is what one means by concurrent or parallel computation. Digital circuitry can perform Boolean operations naturally and quickly. Advancing the array in time is a sequence of purely local Boolean operations on a deterministic algorithm.

Most current parallel computer designs were built with non-local operations in mind. For this reason the basic architecture of present parallel machines is overlaid with a switching network that enables all sites to communicate in various degrees with all other sites. (The usual model of a switching network is a telephone exchange.) The complexity of machine architecture grows rapidly with the number of sites, usually as $n \log n$ at best with some time tradeoff and as $O(n^2)$ at worst. In a large machine, the complexity of the switching network quickly becomes greater than the complexity of the computing configuration.

In a purely local architecture switching networks are unnecessary, so two-dimensional systems can be built in a two-dimensional, or planar configuration, which is the configuration of existing large-scale integrated circuits. Such an architecture can be made physically compact by arranging the circuit boards in an accordion configuration similar to a piece of folded paper. Since the type of geometry chosen is vital to the collective behavior of the lattice gas model and no unique geometry fits all of parameter space, it would be a design mistake to hardwire a particular model into a machine architecture. Machines with programmable geometries could be designed in which the switching overhead to change geometries and rules would be minimal and the gain in flexibility large (Fig. 9).

In more than two dimensions a purely two-dimensional geometry is still efficient, using a combination of massive parallel updating in a two-dimensional plane and pipelining for the extra dimensions. As technology improves, it is easy to imagine fully three-dimensional machines, perhaps with optical pathways between planes, that have a long mean time to failure.

The basic hardware unit in conventional computers is the memory chip, since it has a large storage capacity (256 K bytes or 1 M bytes presently) and is inexpensive, reliable, and available commercially in large quantities. In fact, most modern computers have a memory-bound architecture, with a small number of controlling processors either doing local arithmetic and logical operations or using fast hashing algorithms on large look-up tables. An alternative is the local architecture described above for lattice gas simulators. In computer architecture terms it becomes very attractive to build compact, cheap, very fast simulators which are general over a large class of problems such as fluids. Such machines have a potential processing capacity much larger than the general-purpose architectures of present or foreseen vectorial and pipelined supercomputers. A number of such machines are in the process of being designed and built, and it will be quite interesting to see how these experiments in non-von Neumann architectures (more appropriately called super-von Neumann) turn out.

At present, the most interesting machine existing for lattice gas work is the Connection Machine with around 65,000 elementary processors and several gigabytes of main memory. This machine has a far more complex architecture than needed for pure lattice-gas work, but it was designed for quite a different purpose. Despite this, some excellent simulations have been done on it. The simulations at Los Alamos were done mainly on Crays with SUN workstations serving as code generators, controllers, and graphical units. The next generation of machines will see specialized lattice gas machines whether parallel, pipelined, or some combination, running either against

ARCHITECTURE OF THE LATTICE GAS SIMULATOR

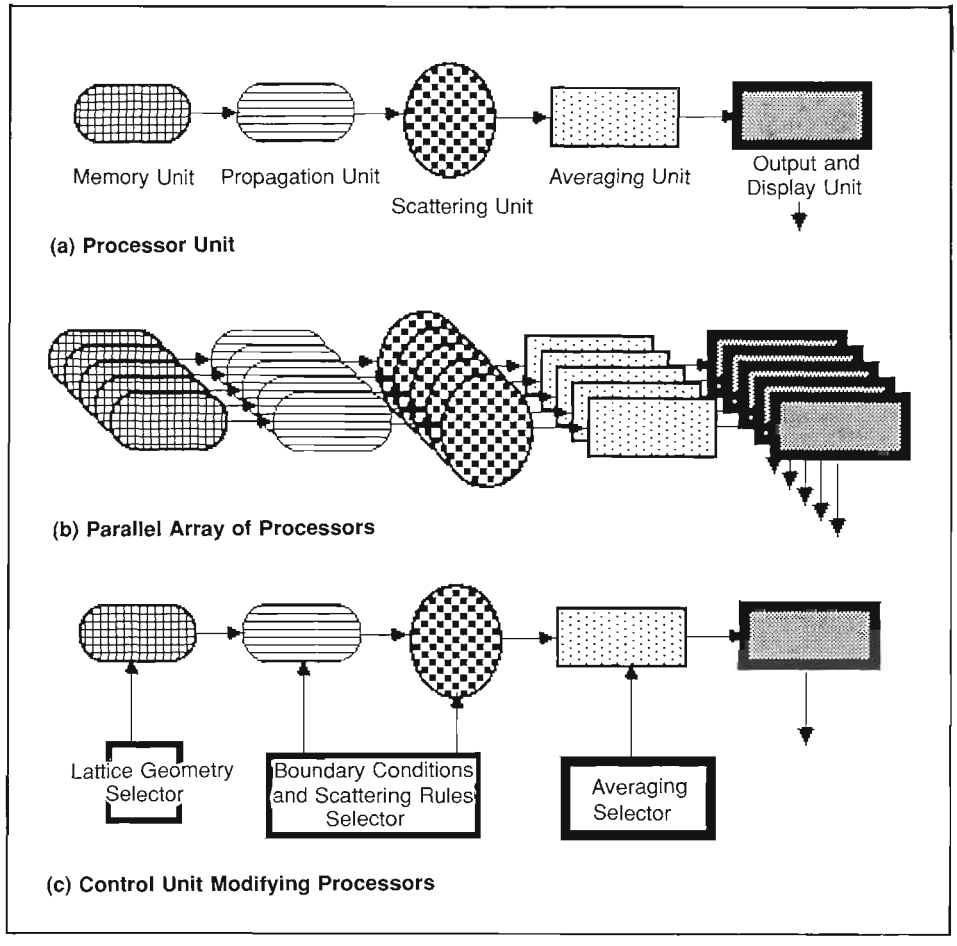


Fig. 9. The lattice gas code is a virtual machine in the sense that the way the code works is exactly the way to build a machine.

(a) The basic processor unit in a lattice gas simular has five units: (1) a memory unit that stores the state at each node of the lattice grid; (2) a propagation unit that advances particles from one node to the next; (3) a scattering unit that checks the state at each node and implements the scattering rules where appropriate; (4) an averaging unit that averages velocities over a preassigned region of the lattice universe; and (5) an output and display unit.

(b) Processors are arranged in a parallel array. Each processor operates independently except at nodes on shared boundaries of the lattice gas universe.

(c) Processor units are overlaid by units that can alter the geometry of the lattice, the collision rules and boundary conditions, and the type of averaging.

Connection Machine style architectures or using them as analyzing engines for processing data generated in lattice gas "black boxes." This will be a learning experience for everyone involved in massive simulation and provide hardware engines that will have many interesting physics and engineering applications.

Unfortunately, fast hardware alone is not enough to provide a truly useful exploration and design tool. A large amount of data is produced in a typical many degree of freedom system simulation. In three dimensions the problems of accessing, processing, storing, and visualizing such quantities of data are unsolved and are really universal problems even for standard supercomputer technology. As the systems we study become more complex, all these problems will also. It will take innovative engineering and physics approaches to overcome them.

Conclusion

To any system naturally thought of as classes of simple elements interacting through local rules there should correspond a lattice-gas universe that can simulate it. From such skeletal gas models, one can gain a new perspective on the underlying mathematical physics of phenomena. So far we have used only the example of fluids and related systems that naturally support flows. The analysis of these systems used the principle of maximum ignorance: Even though we know the system is deterministic, we disregard that information and introduce artificial probabilistic methods. The reason is that the analytic tools for treating problems in this way are well developed, and although tedious to apply, they require no new mathematical or physical insight.

A deep problem in mathematical physics now comes up. The traditional methods of analyzing large probabilistic systems are asymptotic perturbation expansions in various disguises. These contain no information on how fast large-scale collective behavior should occur. We know from computer simulations that local equilibrium in lattice gases takes only a few time steps, global equilibrium occurs as fast as sound propagation will allow, and fully developed hydrodynamic phenomena, including complex instabil-

ities, happen again as fast as a traverse of the geometry by a sound wave. One might say that the gas is precociously asymptotic and that this is basically due to the deterministic property that conveys information at the greatest possible speed.

Methods of analyzing the transient and invariant states of such complex multidimensional cellular spaces, using determinism as a central ingredient, are just beginning to be explored. They are non-perturbative. The problem seems as though some of the methods of dynamical systems theory should apply to it, and there is always the tempting shadow of renormalization-group ideas waiting to be applied with the right formalism. So far we have been just nibbling around the edges of the problem. It is an extraordinarily difficult one, but breaking it would provide new insight into the origin of irreversible processes in nature.

The second feature of lattice gas models, for phenomena reducible to natural skeletal worlds, is their efficiency compared to standard computational methods. Both styles of computing reduce to inventing effective microworlds, but the conventional one is dictated and constrained by a limited vocabulary of difference techniques, whereas the lattice gas method designs a virtual machine inside a real one, whose architectural structure is directly related to physics. It is not a priori clear that elegance equals efficiency. In many cases, lattice gas methods will be better at some kinds of problems, especially ones involving highly complex systems, and in others not. Its usefulness will depend on cleverness and the problem at hand. At worst the two ways of looking at the microphysics are complementary and can be used in various mixtures to create a beautiful and powerful computational tool.

We close this article with a series of conjectures. The image of the physical world as a skeletal lattice gas is essentially an abstract mathematical framework for creating algorithms whose dynamics spans the same solution spaces as many physically important nonlinear partial differential equations that have a microdynamical underpinning. There is no intrinsic reason why this point of view should not extend to those rich nonlinear systems which have no natural many-body picture. The classical Einstein-Hilbert action, phrased in the appropriate space, is no more complex than the Navier-Stokes equations. It should be possible to invent appropriate skeletal virtual computers for various gauge field theories, beginning with the Maxwell equations and proceeding to non-Abelian gauge models. Quantum mechanics can perhaps be implemented by using a variation on the stochastic quantization formulation of Nelson in an appropriate gas. When such models are invented, the physical meaning of the skeletal worlds is open to interpretation. It may be they are only a powerful mathematical device, a kind of virtual Turing machine for solving such problems. But it may also be that they will provide a new point of view on the physical origin and behavior of quantum mechanics and fundamental field-theoretic descriptions of the world. ■

Further Reading

G. E. Uhlenbeck and G. W. Ford. 1963. *Lectures in Statistical Mechanics*. Providence, Rhode Island: American Mathematical Society.

Paul C. Martin. 1968. *Measurements and Correlation Functions*. New York: Gordon and Breach Science Publishers.

Stewart Harris. 1971. *An Introduction to the Theory of the Boltzmann Equation*. New York: Holt, Rinehart, and Winston.

E. M. Lifshitz and L. P. Pitaevskii. 1981. *Physical Kinetics*. Oxford: Pergamon Press.

Stephen Wolfram, editor. 1986. *Theory and Applications of Cellular Automata*. Singapore: World Scientific Publishing Company.

L. D. Landau and E. M. Lifshitz. 1987. *Fluid Mechanics*, second edition. Oxford: Pergamon Press.

Complex Systems, volume 1, number 4. The entirety of this journal issue consists of papers presented at the Workshop on Modern Approaches to Large Nonlinear Systems, Santa Fe, New Mexico, October 27–29, 1986.

U. Frisch, B. Hasslacher, and Y. Pomeau. 1986. Lattice-gas automata for the Navier-Stokes equation. *Physical Review Letters* 56: 1505–1508.

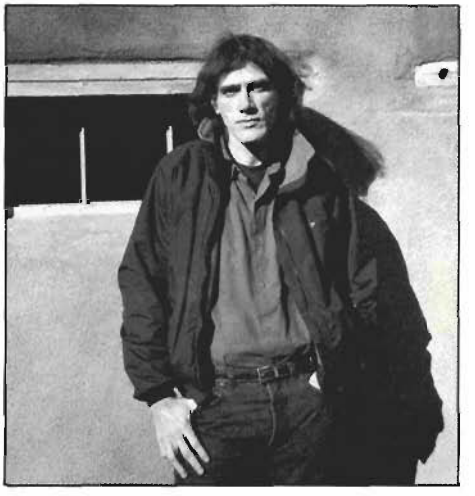
Stephen Wolfram. 1986. Cellular automaton fluids 1: Basic theory. *Journal of Statistical Physics* 45: 471–526.

Victor Yakhot and Stephen A. Orszag. 1986. *Physical Review Letters* 57: 1722.

U. Frisch, D. d'Humières, B. Hasslacher, P. Lallemand, Y. Pomeau, and J. P. Rivet. 1987. *Complex Systems* 1: 649–707.

Leo P. Kadanoff, Guy R. McNamara, and Gianluigi Zanetti. 1987. A Poiseuille viscometer for lattice gas automata. *Complex Systems* 1: 791–803.

Dominique d'Humières and Pierre Lallemand. 1987. Numerical simulations of hydrodynamics with lattice gas automata in two dimensions. *Complex Systems* 1: 599–632.



Brosl Hasslacher was born and grew up in Manhattan. He earned his Ph.D. in theoretical physics (quantum field theory) under Daniel Friedman, the inventor of supergravity, at State University of New York, Stonybrook. Most of his physics career has been spent at the Institute for Advanced Study in Princeton and at Caltech, with stops at the University of Illinois, École Normale Supérieure in Paris, and CERN. He is currently a staff physicist at the Laboratory. Says Brosl, "I spend essentially all my time thinking about physics. Presently I'm working on nonlinear field theories, chaotic systems, lattice gases, the architectures of very fast parallel computers, and the theory of computation and complexity."