RECENT DEVELOPMENT IN SOLAR CONVECTION THEORY

KWING L. CHAN
Applied Research Corporation

Å. NORDLUND
Copenhagen University Observatory

MATTHIAS STEFFEN
Universitat Kiel

and

R. F. STEIN
Michigan State University

In recent years, the theory of solar (and stellar) convection has made fundamental advances due to the increasing cost effectiveness of supercomputers and the constant improvement of numerical techniques. It is expected that the numerical approach will become a dominant trend for the future. Here, we report on these new advances. First, we provide a brief review of the subject. In Sec. I, references to theoretical studies on phenomena related to solar convection are compiled. The objective is to provide a view of the breadth, not to be exhaustive. The next three sections then discuss three numerical studies of solar convection in greater detail, so as to provide the readers with some general understanding of the numerical techniques being used and the results obtained. The discussion starts, in Sec. II, with a two-dimensional study of the spectroscopic properties of solar granules. While the two-dimensional limitation is severely detrimental to some important hydrodynamical processes, it is both economical and able to provide some initial understanding of the gross features of solar convection. Section III discusses the testing of the well-known mixing-

length theory with three-dimensional numerical experiments. It also gives an example of applying the numerically gained knowledge to analytical study, in this case the behavior of compressible convection as a heat engine. Section IV describes a realistic, three-dimensional simulation of solar granulation; many observational features of solar granules are faithfully reproduced. It is the most sophisticated numerical calculation of this sort today.

I. INTRODUCTION

As illustrated by discussions in various other chapters, it is difficult to find a solar phenomenon that is not somewhat affected by the solar convection zone. The need for accurate modeling of convective processes is great. In recent years, with the aide of supercomputers, the theoretical handling of many complicated, highly nonlinear phenomena associated with solar convection have made remarkable progress. In this chapter, we discuss some new developments in this very active area of solar research.

As solar convection touches upon many related areas and the length of discussion must be limited, we arrange this chapter in the following manner: In this section, a compilation of references to the theoretical studies of solar convection problems is given. It is intended to provide a bird's-eye-view of the subject and therefore the bibliography is not exhaustive. However, by putting the many related phenomena together, we hope to present an integrated picture of the solar convection system. In later sections, in depth discussions are presented for several numerical studies that illustrate the techniques and results of present-day efforts in the theory of solar convection.

Table I gives a list of problems closely associated with the solar convection zone. The other columns show different types of theoretical approaches. For each entry, examples of a few of the earliest works are given. The theoretical approaches are roughly divided into three groups: the analytical/one-dimensional group, the modal group and the multi-dimensional simulation group. The analytical and one-dimensional approaches are grouped together as the final steps of most analytical studies result in differential equations of one spatial dimension. The modal approach cannot be strictly separated from the multi-dimensional group as the modes are multi-dimensional. However, the number of modes used in actual calculations are usually very limited and the amplitudes of the modes are described by one-dimensional differential equations that are relatively simple to solve. Multi-dimensional simulations usually require much more computing resources than the other approaches and their results display much more complexity.

Approximations are sometimes used in multi-dimensional simulations to reduce the demand on computing resources. The most popular are the Boussinesq (1903; Spiegel and Veronis 1960) and anelastic approximations (Ogura and Phillips 1962; Gough 1969). The main reason for using these approximations is to eliminate the acoustic modes which generally limit the time steps of a calculation to very small values (the CFL condition; see Richtmyer

TABLE I
Theoretical Studies of Problems Related to Solar Convection

Problem	Analytical/1-D Approach	Modal Approach	Multi-Dimensional Simulation
Dynamics of stratified convection	Bierman (1932) Vitense (1953) Spiegel (1963)	Gough et al. (1975*b*); Latour et al. (1976)	Graham (1975) Deupree (1977) Hurlburt et al. (1984)
Overshooting	Unno (1957) Shaviv and Salpeter (1973)	Veronis (1963) Latour et al. (1981)	Sofia and Chan (1984); Hurlburt et al. (1986)
Granulation	Musman et al. (1976)	Nelson and Musman (1977) Van der Borght and Fox (1983)	Cloutman (1979) Nordlund (1980)
Photospheric effects	Voight (1956) Schroter (1957)	Beckers and Nelson (1978)	Dravins et al. (1981)
Interaction with pulsation	Unno (1967) Stein (1967)	Gabriel et al. (1975); Goldreich and Kumar (1977)	Steffen (1988) Chan and Sofia (1988); Stein et al. (1988)
Interaction with rotation	Wasiutynski (1946) Kippenhahn (1963)	Durney (1970) Busse (1970)	Gilman (1972) Young (1974) Glatzmaier (1984)
Interaction with small-scale magnetic field	Parker (1963) Weiss (1964)	Weiss (1966); Clark and Johnson (1967)	Schussler (1979) Galloway and Moore (1979) Nordlund (1985*b*)
Interaction with large-scale magnetic field	Parker (1955*b*) Babcock (1961) Steenbeck et al. (1966)	Yoshimura (1972) Stix (1973)	Gilman et al. (1981); Glatzmaier (1985*a*)

and Morton 1968) and therefore make the total number of steps prohibitively large. The Boussinesq approximation is a straightforward extension of the technique used in studying liquids; it eliminates all the effects of compressibility except the idealized buoyancy term. In recent years, it has generally been abandoned for computations involving large density stratifications. The anelastic approximation eliminates sound waves by neglecting the Eulerian

variation of density in the continuity equation, but the Lagrangian variation of density is kept; it is applicable to situations with some density stratification. However, it is not valid unless the Mach number is much less than 1 (not always true in the upper region of the solar convection zone) and when the acoustic waves are unimportant to the dynamics. More and more researchers now chose to solve the fully compressible fluid equations. For conciseness, most of the references listed below will not be differentiated according to their levels of approximation as discussed here.

A. Convection Dynamics

The most direct problem of solar convection concerns the hydrodynamics of the process (1st row of Table I). The following questions need to be addressed: what are the patterns of the flows; how can the convective turbulence be described; and how should the flows be related to the structure of the convection zone (i.e., depth distributions of the mean temperature, pressure, etc.)? The last question is particularly important for understanding the internal structure and evolution of the Sun. Up to now, almost all such calculations use the mixing-length theory (Biermann 1932; Vitense 1953). This theory has the advantage of being simple to manipulate analytically and easy to implement computationally. However, it is based on a hypothetical picture of heat carrying bubbles which is not rigorously derived from the fluid equations. Therefore, this theory is only applied with some reluctance, and the results are viewed with caution.

Numerous attempts to improve on the mixing-length theory have been made. The most significant improvement is the generalization to nonlocal theories (Spiegel 1963; Unno 1969; Ulrich 1970; Travis and Matsushima 1973; Nordlund 1976; Xiong 1981; Kuhfuss 1986). Furthermore, modification of the theory to account for the opacity difference in upward and downward flows has been considered (Deupree 1979). Variable mixing-length ratios have also been proposed (Deupree and Varner 1980; Chan et al. 1981; Cloutman 1987).

To study the problem on a more fundamental level, attempts have been made to reconstruct a uniform convective flux by a linear superposition of unstable modes generated by the superadiabatic structure (Hart 1973; Bogart et al. 1980; Narashima and Antia 1982; Antia et al. 1983). However, effects of the neglected nonlinear interactions cannot be properly assessed.

The most reliable way to study convection theory is to solve the fluid dynamical equations. To reduce the computational load, the modal approach uses a very limited number (usually 1 or 2) of planforms to represent the horizontal patterns of the convective motions; the vertical variation of the amplitudes of the modes can be computed in high resolution (Gough et al. 1975*b;* Van der Borght 1975; Latour et al. 1976,1983; Toomre et al. 1976; Fox and Van der Borght 1985; Fox 1985; Legait 1986).

Multi-dimensional solutions of the Navier Stokes equations for convective layers traversing several pressure scale heights were first obtained by Graham (1975,1977). Since then, the simulation approach flourishes in the study of convection dynamics (Deupree 1977; Marcus 1979,1980; Nordlund 1980,1982; Chan et al. 1982,1987; Hurlburt et al. 1984,1986; Yamagushi 1984,1985; Woodward and Porter 1987; Gigas 1989; Cataneo et al. 1990; Nordlund and Dravins 1990). However, the stratification of the solar convection zone is much too large for any single computation; so far, numerical computations can only study different regions of the solar convection zone.

B. Overshooting

The study of convective overshooting into neighboring stable regions is a natural extension of the study of dynamics. It has significant implication in the elemental distribution and evolution of the Sun (Bohm 1963; Weymann 1965; Straus et al. 1976; Chiosi 1986). Most of the theories are based on some extension of the mixing-length theory (Unno 1957; Shaviv and Salpeter 1973; Cogan 1975; Maeder 1975; Roxburgh 1978; Cloutman and Whitaker 1980; Van Ballegooijen 1982; Schmitt et al. 1984; Pidatella et al. 1986). Enlightening results have been obtained by modal calculations (Veronis 1963; Latour et al. 1981; Massaguer et al. 1984) and multi-dimensional simulations (Sofia and Chan 1984; Hurlburt et al. 1986).

C. Granulation

Granulation is the most prominent convective feature on the solar surface. A one-dimensional model based on energy balance has been considered by Musman and Nelson (1976), and two-dimensional single-mode models have been computed by Nelson and Musman (1977) and Van der Borght and Fox (1983). By solving the Navier Stokes equations in two dimensions, Cloutman's model (1979) included a treatment of the thermodynamics of the partially ionized gas. Nordlund's (1980,1982,1985*b*) anelastic model solved the Navier Stokes equations in three dimensions and included a sophisticated treatment of radiative transfer. Detailed behavior of axisymmetric two-dimensional models have been studied by Steffen and Muchmore (1988) and Steffen et al. (1989); also see Sec. II. Recently, Stein and Nordlund extended Nordlund's code to compute fully compressible three-dimensional models (see Sec. IV).

D. Effects on the Photosphere

The effects of convection on the photosphere is closely related to the granulation, the most important aspect being the influence on the radiation spectrum, because that is the fundamental diagnostic tool for observation. The significance of the convective turbulence on the radiative transfer in the photosphere was recognized very early (Voight 1956; Schroter 1957). Using

a sinusoidal model to represent the motions of the granules, Beckers and Nelson (1978) presented a theory for the solar limb effect based solely on the effects of convective motions. Using the kinematic information from Nordlund's granulation model, Dravin et al. (1981,1986,1990*a,b*) have computed details of line asymmetries and wavelength shifts of photospheric spectral lines. Steffen (1987) made a similar study on the spectroscopic properties of solar granulation with his axisymmetric stationary models.

E. Interaction with Pulsation

The study of pulsations or waves in the Sun has grown into a very large field. The 5-min oscillations discovered by Leighton et al. (1962) and Noyes and Leighton (1963) have become a very important diagnostic tool for the internal condition of the Sun (Ulrich 1970; Leibacher and Stein 1971; Wolff 1972; Unno et al. 1979; Cox 1980; Duval and Harvey 1983; Ulrich and Rhodes 1983; Deubner and Gough 1984; Gough 1985*b;* Christensen-Dalsgaard et al. 1985*a;* Christensen-Dalsgaard 1986). However, studies on the connection of solar convection and waves were originally developed for the heating of the chromosphere and corona (Whitaker 1963; Athay 1966; Schatzman and Souffrin 1967; Ulmschneider 1971). The convective turbulence was proposed to be the source of waves that supply mechanical energy to the upper atmosphere of the Sun (Unno 1967; Stein 1967,1968). This branch of development is merging with the study of convective effects on nonradial pulsations (Gabril et al. 1975; Goldreich and Keeley 1977*b;* Gonczi and Osaki 1980; Stellingwerf 1982; Goldreich and Kumer 1988). The generation of *p*-mode oscillations by convection was recently simulated in multidimensional calculations (Steffen 1988; Chan and Sofia 1988; Stein et al. 1988,1989). Excitation of gravity waves by penetrative convection was illustrated by the calculation of Hurlburt et al. (1986).

F. Interaction with Rotation

Like most astrophysical objects, the Sun is rotating. Rotation introduces much complication in the theory of convection. One important problem is the generation of the latitudinal differential rotation (Newton and Nunn 1951; Howard and Harvey 1970). Many theories on the origin of solar differential rotation have been proposed; most of them were based on the anisotropic action of the convective turbulence or on the latitudinal variation of convective heat transfer (Wasiutynski 1946; Kippenhahn 1963; Weiss 1965; Durney 1970; Busse 1970; Durney and Roxburgh 1971; Belvedere and Paterno 1977; Durney and Spruit 1979; Rudiger 1980,1989; Stix 1981; Pidatella et al. 1986; Tuominen and Rudiger 1989; Durney 1989). It also has been proposed that the differential rotation is the zonal velocity field of an axisymmetric convective mode (Chan et al. 1987). Numerical simulation of solar differential rotation was pioneered by Gilman (1972,1977*a*); elaborations were made in

later years (Young 1974; Gilman and Glatzmaier 1981; Glatzmaier and Gilman 1981; Glatzmaier 1984). However, recent results from helioseismology (Brown et al. 1989; Libbrecht 1989; Dziembowski et al. 1989) indicate that the distribution of angular velocity inside the Sun is substantially different from those predicted by the numerical models. Thus a reconsideration of the numerical problem is being undertaken (Gilman et al. 1989; Chan and Mayr 1989).

G. Interaction with the Global Magnetic Field: the Dynamo Model

The magnetic field adds another dimension of difficulty to the problem of solar convection. The origin of the 11-yr sunspot cycle is a problem that has occupied the minds of many generations of solar physicists (Hale 1908; Cowling 1934; Alfvén 1950). The prediction of the magnitude of a cycle has very practical application in space aeronautics (Schatten and Sofia 1987). In most solar dynamo models, the azimuthal field is generated by the nonuniform rotation stretching the poloidal field, and the poloidal field is produced by the cyclonic convective motions twisting the azimuthal field (Parker 1955*b;* Babcock 1961; Steenbeck et al. 1966; Steenbeck and Krause 1969*a,b;* Leighton 1969; Kohler 1970,1973; Roberts 1972). Global kinematic dynamo models compute the evolution of the global magnetic field with assumed patterns of the global scale convection (Yoshimura 1971,1972,1975*a;* Stix 1973,1976*a*). Dynamically self-consistent dynamo models further take account of the feedback of magnetic actions on the fluid motions (Gilman and Miller 1981,1986; Gilman 1983; Glatzmaier 1985*b;* Brandenburg et al. 1990). So far, such numerical models have not been successful in generating solutions compatible with the essential features of the solar magnetic cycle. Recently, attentions have been brought to locating the dynamo region in a thin layer near the bottom of the convection zone (DeLuca and Gilman 1986,1988; Gilman et al. 1989).

H. Interaction with the Magnetic Field in Smaller Scales

Studying the interaction of convection with the magnetic field in the small scales is fundamental to the understanding of sunspots, active regions and the fibrils that are the constituents of the global field (Parker 1955*b*,1975*a*,1979*a*,1982*a,b;* Weiss 1964; Leighton 1964; Gurm and Wentzel 1967; Mullan 1974; Piddington 1975; Schussler 1979,1980; Moffatt 1978; Spruit and Zweibel 1979*b;* Schatten et al. 1986). Many numerical studies have been made on the different aspects of the problem (Weiss 1966,1981*a;* Clark and Johnson 1967; Proctor and Weiss 1982; Galloway and Moore 1979; Schussler 1979; Nordlund 1983,1985*b*,1986; Deinzer et al. 1984*a,b;* Cattaneo and Hughes 1986; Hurlburt and Toomre 1988; Knolker et al. 1988; Nordlund and Stein 1989; Brandenburg et al. 1989).

II. TWO-DIMENSIONAL NUMERICAL SIMULATIONS AND SPECTROSCOPIC PROPERTIES OF SOLAR GRANULES

A. Introduction

In this section we describe two-dimensional (cylindrical) hydrodynamical models of solar granular convection cells, obtained from numerical simulations of turbulent compressible convection in a stratified medium, including realistic thermodynamics and a detailed treatment of radiative energy transfer. Based on these models, we discuss the dynamical and thermal structure of the solar photosphere. For a direct comparison with various spectroscopic observations, spatially resolved and horizontally averaged synthetic spectra have been computed from the models. We conclude that basic properties of the solar granulation can be reproduced and understood by our two-dimensional simulations.

The theoretical understanding of the phenomenon of solar granulation is one of the topics of present-day solar physics. Despite extensive observational and theoretical efforts, our knowledge about the velocity and temperature fluctuations associated with granulation in the solar photosphere is still incomplete. Commonly, granulation is interpreted as a pattern of surface convection cells at the boundary between the hydrogen convection zone and the photosphere, but from time to time even the convective character of granulation has been questioned, as in a recent study by Roudier and Muller (1986) who suggest that granules smaller than about 1000 km might actually be turbulent eddies, owing their existence to larger-scale flows of which they are just decay products, in contrast to convection cells, that are autonomous structures driven by buoyancy.

It has become evident over the years that it is impossible to obtain a consistent quantitative picture of the conditions prevailing in the solar granulation layers from a purely empirical analysis of existing observational material (see, e.g., Bray et al. 1984). To make progress, theoretical granulation models having the potential to provide data that can be compared directly to solar observations are indispensable.

Notoriously, the calculation of convection in stellar atmospheres is a complex problem, requiring the application of time-dependent, nonlinear hydrodynamics to a highly turbulent flow that, to complicate the situation, strongly interacts with the photospheric radiation field. Although in principle the problem is well defined by a few differential equations, an analytical solution is impossible when realistic background physics is to be included. Relevant results can only be obtained from numerical simulations on powerful computers.

The purpose of this section is to demonstrate what kind of results can be obtained from two-dimensional models of granular convection cells. In Sec. II.B we give some motivation for doing two-dimensional calculations. Section II.C briefly describes the basic foundations of the numerical simulations,

while in Secs. II.D and E some of the results are presented along with corresponding observations.

B. Why Two-Dimensional Hydrodynamics?

Turbulent convection is intrinsically a three-dimensional process. Strictly two-dimensional flows are nowhere found in nature; they are an idealization applied to flows in computer simulations. It is well known that two-dimensional turbulence has basically different cascade properties compared to three-dimensional turbulence (cf. review by Kraichnan and Montgomery 1980). While in three-dimensional systems the energy cascade is from larger to smaller spatial scales, it is from higher to lower wavenumber modes in two-dimensional flows, i.e., the largest possible spatial scales are preferred here. But it is not necessarily true that all the energy in two-dimensional flows resides at the largest wavelengths. If the flow is strongly driven at smaller wavelengths, there can be significant energy also at those smaller wavelengths. However, it is clear that some interesting phenomena are missed even qualitatively in two dimensions. For example, vertical vorticity cannot be modeled and there is no vortex stretching.

Nevertheless, there are good reasons to perform two-dimensional simulations of stellar convection. One important aspect is that in the case of the solar granulation the preferred horizontal scales are known observationally. Although granules down to very small sizes may exist, the overwhelming contribution to the surface area (and hence to the emergent spectrum) comes from granules measuring between 700 and 1500 km in diameter (Roudier and Muller 1986). Furthermore, the visual impression of the granulation pattern suggests that the structure of a single granule is roughly axially symmetrical, especially for undisturbed granules. On this basis, it seems reasonable to model isolated granular convection cells in cylindrical symmetry, with model diameters corresponding to scales that dominate the appearance of the solar granulation.

There are in fact examples of two-dimensional convection calculations that have been successfully applied to describe actual convective phenomena accessible to experimental verification. For example, the behavior of fireballs could be realistically predicted by two-dimensional models for a large range of events (Ruppel and Norton 1975). For a reasonable agreement between simulation and nature it seems important that the appropriate spatial scales are imposed on the models and that the actual mean flow is essentially two-dimensional.

Further justification for two-dimensional calculations comes from a comparison between otherwise identical simulations in two and three dimensions, as carried out, e.g., by Deupree (1984*a*,*b*) and by Chan and Sofia (1986). These authors find that the corresponding results are phenomenologically similar and many basic properties found in two-dimensional flows persist in three dimensions.

Finally, it is important to note that restricting the problem to two dimensions is an enormous advantage in terms of computer requirements. The calculation of three-dimensional models is much more severely restricted by available computer capacities than two-dimensional calculations, in particular if detailed radiative transfer is to be taken into account. This has important consequences: the spatial resolution achievable with two-dimensional models is at least an order of magnitude better than with three-dimensional models; or, alternatively, the simulated volume can be correspondingly larger in two dimensions. Moreover, it is often necessary to repeat a simulation with a variety of different parameters to understand the physical (or numerical) cause for a certain phenomenon found in the calculations. It is simply not feasible to do a large number of test runs with fully three-dimensional models. In practice, this is not merely a quantitative difference, but constitutes a qualitative advantage of two-dimensional calculations. For example, the three-dimensional compressible simulations described by Stein et al. (1989) took 600 CPU hr on a vector machine to cover 3 solar hr. In contrast, the two-dimensional models discussed here need approximately 5 to 10 hr of CPU time on a CRAY X-MP to simulate 3 hr of real time.

Apart from the problem of properly displaying three-dimensional information, the added complexity makes the numerical results more difficult to understand. In contrast, the two-dimensional (cylindrical) models described in the following are comparatively simple, especially if the steady-state solutions are considered. Under these circumstances it is more readily possible to study in some detail the physical mechanisms governing granular convection.

Summarizing, modeling granular convection cells in two dimensions is well motivated. It is clear that two-dimensional simulations cannot replace fully three-dimensional calculations. Rather, they must be understood as a complementary approach, which may prove advantageous for certain applications. Our aim is to investigate to what extent such an idealized description is useful to explain the observations. The least we can expect is that two-dimensional models including detailed thermodynamics and realistic radiative transfer will be able to give a much better representation of the dynamical and thermal structure of the solar photosphere than the commonly used one-dimensional solar atmospheric models based on mixing-length concepts.

C. The Numerical Simulations

1. Foundations of the Numerical Simulations. The framework of the model calculations is given by the time-dependent, nonlinear equations of hydrodynamics prescribing the conservation of mass, momentum and energy in a stratified compressible fluid. To account for the highly turbulent character expected for the solar granular flow (Reynolds number $\approx 10^9$) viscosity terms are included to model roughly the turbulent exchange of momentum and energy on subgrid scales. The corresponding subgrid scale eddy viscosity is

calculated according to the scheme given by Deardorff (1970,1971). In cylindrical coordinates (r,φ,z), assuming axial symmetry but not permitting a φ-component of the flow velocity, the conservation equations for mass and momentum read:

$$\frac{D\rho}{Dt} + \rho \operatorname{div} \mathbf{v} = 0 \tag{1}$$

$$\frac{Du}{Dt} + \frac{1}{\rho}\frac{\partial p}{\partial r} - g_{r,vis} = 0 \tag{2}$$

$$\frac{Dv}{Dt} + \frac{1}{\rho}\frac{\partial p}{\partial z} - g_{z,vis} + g = 0 \tag{3}$$

where $\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial r} + v\frac{\partial}{\partial z}$ is the Lagrangean (substantial) derivative along the particle path, t denotes time, r and z represent the horizontal and vertical spatial coordinates (z increasing upwards), ρ is the mass density, p the gas pressure; u and v are the horizontal and vertical velocity components, respectively, $\operatorname{div} \mathbf{v} \equiv \partial u/\partial r + u/r + \partial v/\partial z$; g is the acceleration of gravity (directed downwards) while $g_{r,vis}$ and $g_{z,vis}$ stand for the viscous acceleration in r and z direction, respectively, representing functions of the spatial derivatives of the velocity field.

The energy equation may be written as

$$\frac{Ds}{Dt} = \left.\frac{Ds}{Dt}\right|^{rad} + \left.\frac{Ds}{Dt}\right|^{dis} + \left.\frac{Ds}{Dt}\right|^{dif} \tag{4}$$

stating that the specific entropy s of a moving fluid element is in general not constant (as in the adiabatic case), but changes due to: (1) exchange of radiation, (2) viscous dissipation of kinetic energy, (3) turbulent diffusion of heat. The radiative damping term is given by

$$\left.\frac{Ds}{Dt}\right|^{rad} = \frac{4\pi}{\rho T}\int_0^\infty k_\nu (J_\nu - S_\nu)\mathrm{d}\nu \tag{5}$$

where T is temperature, k_ν is the monochromatic total absorption coefficient (cm^{-1}), J_ν is the angle-averaged monochromatic intensity and S_ν is the source

function which in LTE is equal to the Planck function B_ν. The dissipation term reads

$$\left.\frac{Ds}{Dt}\right|^{dis} = \frac{\Phi}{\rho T} \tag{6}$$

with

$$\Phi = \eta\left[2\left(\frac{\partial u}{\partial r}\right)^2 + 2\left(\frac{\partial v}{\partial z}\right)^2 + 2\left(\frac{u}{r}\right)^2 + \left(\frac{\partial u}{\partial z} + \frac{\partial v}{\partial r}\right)^2 - \frac{2}{3}(\text{div } \mathbf{v})^2\right]. \tag{7}$$

η is the dynamical viscosity which itself is a complicated function of the velocity field. The turbulent heat transfer is modeled as

$$\left.\frac{Ds}{Dt}\right|^{dif} = -\frac{\text{div } \mathbf{f}_{dif}}{\rho T} \tag{8}$$

where

$$\mathbf{f}_{dif} = -\frac{2}{Pr} T\eta \text{ grad } s. \tag{9}$$

The Prandtl number Pr is taken to be 1. Note that here the heat diffusion is proportional to the local entropy gradient, which is substantially different from the commonly used proportionality to the temperature gradient. This concept is more compatible with the idea that turbulence has the tendency to produce isentropic conditions, as opposed to the diffusive action based on temperature (see also Chan and Sofia 1986). It is worth pointing out that the convective energy flux according to mixing-length theory may be written, without restriction to the case of an ideal gas, as

$$F_c = -\frac{1}{2} T \rho v l \frac{ds}{dz}. \tag{10}$$

This is essentially the same expression as Eq. (9). The main quantitative difference is that while the mixing length l in Eq. (10) is of the order of one pressure scale height, it is of the order of the grid resolution in Eq. (9). The calculations show that in the solar photosphere, the effect of dissipation is about one order of magnitude smaller than that of turbulent heat transfer, which, in turn, is roughly one order of magnitude smaller than radiative damping, the dominant mode of energy exchange, at least in the layers around optical depth unity.

The thermodynamical relations entering the computations explicitly al-

low for temperature and pressure dependent ionization of hydrogen that has a critical influence on the specific heats, the adiabatic temperature gradient, and hence on the strength of convection in the upper layers of the solar convection zone. Helium ionization is negligible in the temperature/pressure domain represented by the current models.

Equally important, a realistic modeling of granular convection requires a reasonable description of the interaction of hydrodynamics and radiative transfer. Because the simulation of photospheric convection has to include the optically thick and optically thin (visible) layers at the same time, it is not appropriate to calculate the radiation field from the local temperature gradient by applying the often-used diffusion approximation; this is only acceptable in the deep, optically thick layers where, however, radiative transfer is of minor importance for the gas dynamics. In the transition region from optically thick to optically thin conditions, radiative cooling becomes a very important factor. Here and in the higher layers, the radiation field has a nonlocal character and the diffusion approximation is no longer valid. More accurate (but also more costly) methods are needed to derive the radiation field at the critical boundary between convective and radiative layers. Notably, the thermal and dynamical structure of this so-called overshoot region are of considerable interest, because it is here that most of the emergent spectrum originates.

Our approach is to solve the equation of radiative transfer.

$$dI_\nu = -k_\nu(I_\nu - S_\nu)dl \tag{11}$$

along a large number of rays crossing the model in various directions and with different inclinations. Angle-averaging of the intensities resulting from Eq. (11) yields J_ν needed in Eq. (5). In this way, nonlocal radiative exchange is taken into account both vertically and horizontally. Using a realistic Rosseland mean opacity as a function of pressure and temperature, the gray approximation in LTE has been adopted so far.

The equations of radiative transfer and hydrodynamics are solved simultaneously without introducing simplifications such as linearizations or the anelastic approximation. The numerical scheme uses an iterative procedure based on the method of bi-characteristics. The code was derived from that of Stefanik et al. (1984). Several extensions were applied to adapt it for the simulation of solar granulation, including the introduction of turbulent viscosity, hydrogen ionization and two-dimensional radiative transfer.

2. Model Parameters, Boundary and Initial Conditions. Typically, the models extend vertically from 250 km below to 600 km above the $\tau_{ross} = 1$ level, i.e., they span several pressure scale heights. The vertical grid distance ranges from about 20 km in the lower, convective part (more than 10 grid points per pressure scale height) to about 40 km in the upper, radiative region.

The diameter of the cylindrical model is a free parameter which has been

varied between 260 and 2600 km. Depending on the model diameter, the horizontal resolution of the grid lies between 10 and 40 km.

The time step is typically less than 1 s by the requirement that it must not exceed the sound-travel time between any two adjacent grid points (Courant condition).

Boundary conditions appropriate to the situation in the solar atmosphere must be imposed before the differential equations can be solved numerically. At the axis of symmetry ($r = 0$) the horizontal component of the velocity must be zero, while for all remaining variables the horizontal derivative must vanish. Similar conditions are imposed at the lateral boundary ($r = R$). For the upper boundary we have two options. Either we use a stress-free closed top ($v = \partial u/\partial z = 0$) or the upper boundary is made transmitting for simple acoustic waves as described by Stefanik et al. (1984), to permit (initial) pressure disturbances to leave the computational domain instead of being reflected.

For our purpose, the stratification of the solar convection zone is much too large to be included in a single simulation. We model only the very top of the convection zone. In this situation, the lower boundary is most critical because there is no way to place it at a position where conditions naturally allow a simple boundary condition to be used; considering that at the location of the lower boundary nearly all the energy is carried by convection, a closed bottom seems unreasonable. Much effort was needed to devise an open lower boundary condition, allowing a free flow of gas out of and into the model. The basic idea in the formulation of this boundary condition is to assume a spatially constant pressure p^* along those parts of the bottom where the flow is directed upward. Two principal versions have been tested. In the first one (a), p^* is constant also in time. Then the value of p^* fixes the depth of the lower boundary within the atmosphere. Alternatively (version b) p^* is adjusted from time step to time step such that the total mass within the model volume is conserved. In this case, the depth of the model is determined by its (initial) total mass. The entropy of the gas entering the model from below is automatically adjusted in such a way as to make the radiative flux through the *upper* boundary correspond to the specified effective temperature. The time constant for this entropy adjustment is chosen to be of the order of 1 turnover time. Both versions of the lower boundary condition are physically consistent and flexible enough to allow the flow itself to choose the horizontal positions of rising and sinking regions at any time during the simulation. While versions (a) and (b) are the same for steady-state situations, condition (b) seems in general to give more reasonable results.

For radiative transfer we assume no incident radiation at the top, while at the bottom we can safely apply the diffusion approximation as a boundary condition. The lateral boundary is chosen to be reflective in order to mimic closely conditions in the solar granulation where each convection cell is surrounded by several similar ones. This choice also guarantees that the net

energy flow through the side walls is zero, and no assumptions about the surroundings of the model need to be made.

In principle, any arbitrary configuration can be taken as an initial condition. In practice, however, the initial state must not be too far from the mean relaxed state that develops during the subsequent time evolution. Otherwise numerical problems can arise if the flow velocities go supersonic and shocks are generated in a too violent initial phase of relaxation.

3. Computation of the Emergent Spectrum. It is essential to have the possibility to calculate synthetic spectra from the numerical models in order to investigate their observational implications. Although in the numerical simulation itself radiative transfer is treated in the gray approximation, it is possible (somewhat inconsistently) to use the two-dimensional hydrodynamical model atmospheres for detailed line-formation calculations to derive basic spectroscopic properties of the models. Employing a modified version of the LTE package ATMOS/LINFOR, developed by the Kiel Group for the analysis of stellar spectra, the emergent spectrum at arbitrary wavelengths can be obtained as a function of the inclination of the line of sight against the vertical axis μ. Such spectrum synthesis calculations are performed only at selected instants of time for diagnostic purposes, e.g., to evaluate the continuum intensity contrast or the asymmetry of spectral line profiles.

D. Resulting Granular Flows

1. Steady-State Solutions. Starting with appropriate initial conditions, we can follow the time evolution of the flow. From a first series of models, we found that the flow developed towards a steady state if the chosen diameter of the cell was less than a critical upper limit of roughly 2000 km. Subsequently, the calculation of the radiation field was improved to give a better angular resolution, using more than 10 times as many rays as before. Furthermore, the new scheme was designed to assure a much better numerical conservation of energy. The resulting series of second-generation models shows no steady-state solutions down to cell diameters of about 1000 km. It is presently not known whether the new models become stationary if the cell size is further reduced; the corresponding runs have not yet been carried out.

Although the new, nonstationary models are certainly more realistic, we discuss the first-generation steady-state models here, because they reveal more clearly some of the basic physics governing granular convection. A typical steady-state solution is displayed in Fig. 1, where the model diameter is 1750 km. All stationary models show similar characteristic flow patterns in the lower, convectively unstable part of the model. A strong downdraft at the axis of symmetry with maximum velocities of the order of 6 km s^{-1} is surrounded by a broader ring-like upflow of hot gas with lower velocities ($\lesssim$2.5 km s^{-1}) that again turns into a narrow downflow near the side walls. The convective velocity field extends considerably into the stable layers, a

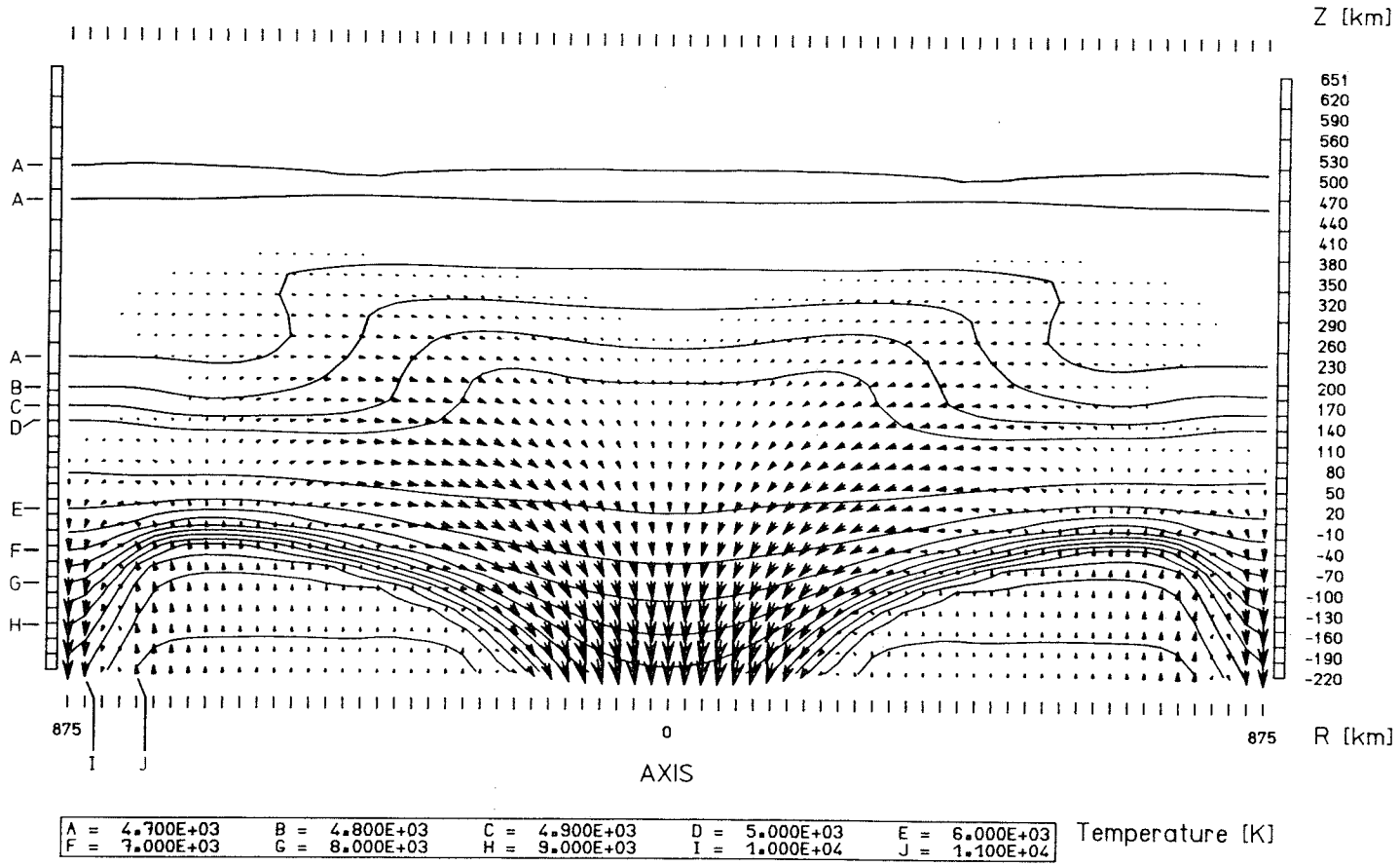

Fig. 1. Steady-state granular flow resulting from a two-dimensional numerical simulation in cylindrical symmetry. The convection cell measures 1750 km in diameter and extends from 220 km below to 650 km above $\tau = 1$ ($z = 0$). $\tau = 0.1$ corresponds to $z = 135$ km, $\tau = 0.01$ to $z = 280$ km. Arrows indicate direction and magnitude of the local velocity where $|v|$ exceeds 300 m s^{-1}; their lengths scale in proportion to $|v|$. The maximum velocity is 6.3 km s^{-1} in the central downdraft, while upward velocities are < 2.5 km s^{-1}. The thermal structure is represented by lines of constant temperature in steps of 500 K in the lower atmosphere, and in steps of 100 K where temperature is below

result that cannot be obtained within the framework of local mixing-length theories. The amount of overshooting is clearly a function of the horizontal size of the convection cells. While the absolute value of the *rms* vertical velocity v_{rms} is typically 1 km s^{-1} at a height of approximately 30 km above $\tau = 1$, quite independent of cell size, we find that the scale height of v_{rms} in the overshooting region H_v depends on the model diameter D roughly as $H_v \sim D^{0.7}$, i.e., vertical motions decline more rapidly with height above smaller granules as compared to the larger ones. Similarly, the magnitude of the maximum *rms* horizontal velocity, $u_{rms,max}$ depends on D approximately as $u_{rms,max} \sim D^{0.5}$. This means that the horizontal return flows occurring in the layers around $\tau = 1$, where the vertical motions are decelerated and turned into a horizontal direction, are more pronounced in large granules.

The corresponding temperature structure is also indicated in Fig. 1. The lines of constant temperature illustrate that granulation generates large horizontal temperature differences, typically more than 4000 K just 100 km below the visible surface ($\tau = 1$). At first sight, this value seems exceedingly large. However, it is easily verified that temperature differences of this magnitude are necessary to carry the solar energy flux by convection with flow velocities of a few km s^{-1}.

Equally remarkable, the calculations produce a very steep temperature gradient at the top of the ascending part of the flow ($\gtrsim$40 K km^{-1}) where the hot gas reaches the optically thin layers and loses its excess energy within a short time by efficient radiative cooling. The steep temperature gradient in concert with the recombination of hydrogen produces a local density inversion about 50 km below $\tau = 1$; i.e., a layer of relatively higher density lies on top of gas with lower density. In contrast, density increases monotonically with depth in the cool, intergranular regions.

In the overshooting layers, the temperature fluctuations change sign: these layers are relatively cooler above the ascending granular parts and relatively warmer above the descending intergranular regions. This behavior is a consequence of the penetration of the convective motions into stably stratified atmospheric layers.

The uppermost part of the model is essentially in radiative equilibrium. Here the temperature is nearly constant with height as expected for a gray radiative atmosphere. In these layers, horizontal temperature fluctuations become insignificant for the steady-state models. For a detailed study of the calculated steady-state velocity and temperature fields as a function of horizontal cell size see Steffen et al. (1989).

Test calculations have shown that details of the initial conditions are unimportant in the sense that the final steady state seems to be largely independent of the initial configuration. In particular, an initial model suggesting the flow to ascend at the axis of symmetry and descend in a ring surrounding the central part, resulted in a reversal of the motions after a short time. This behavior seems to be a consequence of the symmetry conditions required at

the axis, and not due to a problem with the code or the lower boundary condition. This notion is supported by the numerical simulation study of compressible convection by Chan and Sofia (1986). They found that downdrafts are attracted and enhanced by impenetrable lateral boundaries (centers of symmetry). Thus, the ring-shaped granules emerging from our calculations may well be related to the so-called "exploding granules," a common phenomenon in quiet granulation (Title et al. 1987*a*).

Finally, it is worthwhile mentioning that oscillations, superimposed on the convective flow, seem to be ubiquitous. Typically, we find periods of the order of 250 s (Steffen 1988). Current evidence suggests that the oscillation frequency is related to the acoustic cutoff frequency in the layers around $\tau = 1$. Test runs have confirmed the frequency to increase in proportion to g, the acceleration of gravity.

2. Time-Dependent Flows. Simulations with model diameters exceeding the critical upper limit mentioned above never reach a steady state, not even asymptotically; they are truly nonstationary. For diameters close to the critical limit, we find long quiet periods of time (about 20 to 30 min) where the topology of the flow is basically like that of the steady-state solutions with smaller horizontal size. After some time, however, oscillations with increasing amplitude develop and finally lead to a violent reorganization of the flow structure: the extended rising regions become separated by a downflow, implying that large convection cells temporarily split into smaller fragments (rings). This configuration typically lives for 5 to 10 min, after which adjacent regions of hot rising gas have the tendency to merge again. When the initial topology is restored, another quiet-time interval begins. The downdrafts at the axis of symmetry and at the side walls persist throughout the simulated time interval of approximately 2 hr of real time.

A similar behavior is found from the more recent simulations carried out with the improved version of the code (see above). For all cases studied so far, the resulting flows exhibit a distinctly nonstationary character.

The time evolution of the flow can be characterized as stochastic. Significant changes occur on time scales of the order 10 min (corresponding to approximately 1 turnover time), comparable to typical granular life times. Alterations of the flow topology may be interpreted as a continuous splitting and merging of granules under the constraints imposed by the cylindrical symmetry. Sometimes these changes are strong enough to cause supersonic flow velocities in the higher layers, leading to the formation of upward traveling shocks above the central downdraft.

A snapshot from a nonstationary simulation, using the more advanced version of the code, is shown in Fig. 2. Here the center of the rising part of the flow has collapsed, resulting in the formation of two "granules" separated by a downflow. In the subphotospheric layers, velocity and temperature fluctuations are comparable in magnitude to those found in the steady-state mod-

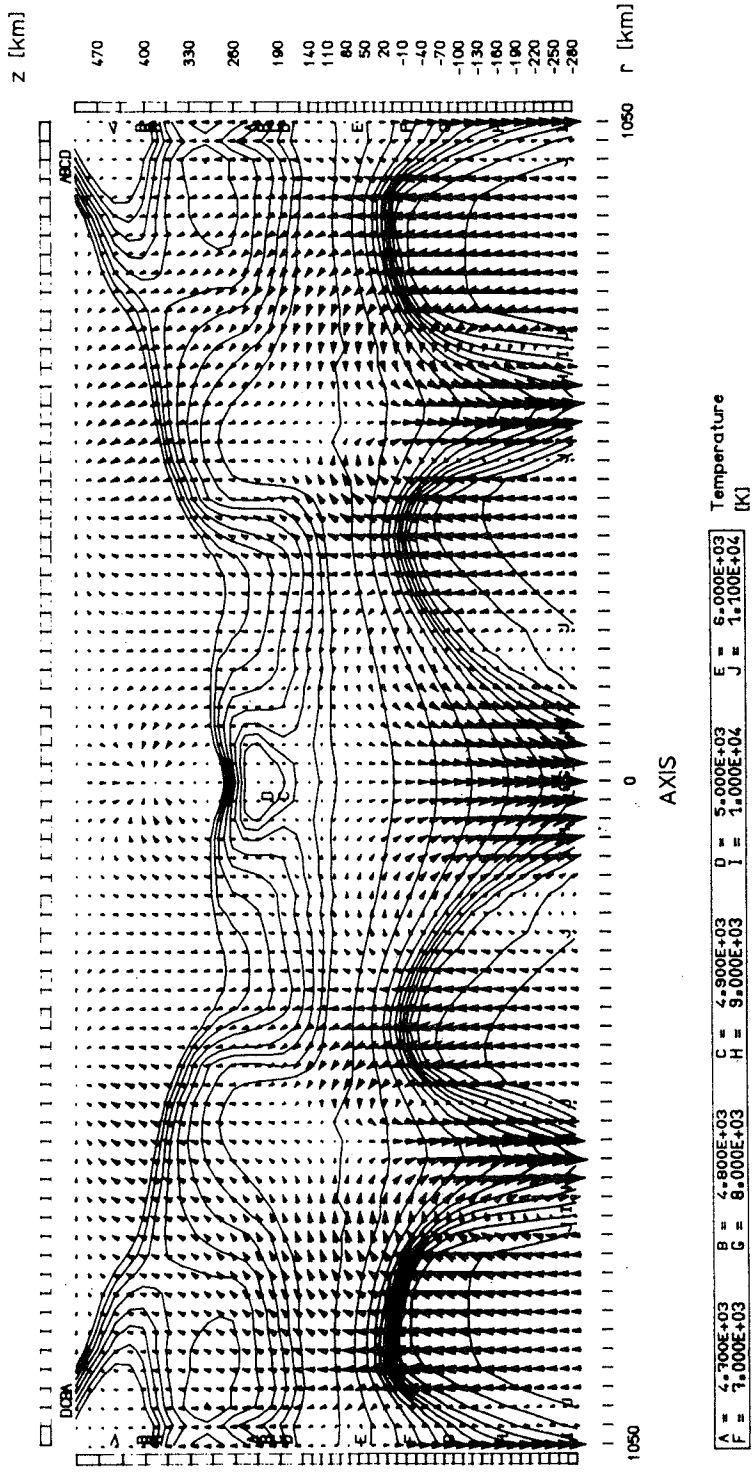

Fig. 2. Snapshot from a nonstationary numerical simulation of granular convection in a cell measuring 2100 km in diameter. The velocity and temperature field is indicated as in Fig. 1. The maximum velocity is $|v| = 7.6$ km s^{-1}.

els. However, the layers above $\tau = 1$ are much more strongly affected by granular convection. In contrast to the steady-state models, we have to conclude that the thermal structure of the higher layers (where the spectral lines are formed) is not exclusively determined by the condition of radiative equilibrium. Rather, the velocity field due to overshooting motions from the convection zone seems to have an important impact on the temperature stratification of the upper solar photosphere.

The development of more than one ring of rising gas may be an indication that the assumption of cylindrical symmetry is no longer a good choice. Perhaps a more reasonable alternative would be to use two-dimensional rectangular coordinates, where the granules are rolls of infinite length instead of rings. Of course, the real topology of granulation can only be modeled in three dimensions (see Sec. III and IV).

E. Spectroscopic Properties of the Two-Dimensional Model Granules

1. Continuum. From radiative transfer calculations along rays parallel to the axis of symmetry, we can simulate spectroscopic observations at disk center. Viewed in continuum light, the steady-state models show a dark center surrounded by a bright ring that is bounded by an outer dark lane (resembling an "exploding granule"). For steady-state models with diameters between 1000 and 2000 km the *rms* intensity contrast of the two-dimensional intensity pattern ranges between 14 and 16% in the continuum at 5000 Å, which seems to be in reasonable agreement with observational evidence (see, e.g., Wittmann 1979; Bray et al. 1984). For the larger, time-dependent models we obtain a somewhat higher value of the *rms* intensity contrast. Averaged over time, 20% at 5000 Å is a typical value.

Towards smaller granular scales the amplitude of the horizontal intensity fluctuations declines strongly. As discussed in detail by Steffen et al. (1989), the main reason is that horizontal radiative exchange becomes increasingly more efficient with decreasing cell size, reducing horizontal temperature fluctuations particularly in the continuum-forming layers. Observations and model calculations indicate that the continuum intensity contrast δI_{rms} depends on wavelength roughly as $\delta I_{rms} \sim 1/\lambda$, essentially reflecting the wavelength dependence of the Planck function.

Spectrum synthesis for different disk positions requires more extensive calculations with inclined rays ($\mu = \cos\theta < 1$). Table II gives $\delta I_{rms}(\mu)/\delta I_{rms}(1)$ at λ 5380 Å as a function of μ for a steady-state model measuring 1050 km in diameter (column A) and for the larger model shown in Fig. 2 (column B), representing a typical phase during the time evolution of a nonstationary simulation (no time average).

We note a monotonic decline of the granular contrast towards the limb for both models. It is obvious that the nonstationary model (column B) produces a slower decrease of the granular contrast towards the limb, because it shows larger temperature inhomogeneities in the higher layers (see above).

TABLE II
Simulated and Observed Center-to-Limb Variation of *rms* Intensity Contrast at 5380 Å

	$\delta I_{rms}(\mu)/\delta I_{rms}(1)$		
μ	A	B	C
1.0	1.00	1.00	1.00
0.8	0.88	0.89	0.93
0.6	0.70	0.76	0.75
0.4	0.44	0.61	0.52
0.2	0.14	0.52	0.25

The original observational results found by Schmidt et al. (1979) are given in column C of Table II showing the same general trend as derived from the simulations. It must be mentioned, however, that according to Wiesmeier and Durrant (1981), the values given in column C should be corrected by a factor $1/\sqrt{\mu}$. Accordingly, the observed center-to-limb variation of the granular intensity contrast is even smaller than indicated by the simulated results listed in column B.

2. Line Spectrum. The profiles of spectral lines can be calculated with a spatial resolution that corresponds to the horizontal grid distance of the hydrodynamical models (typically between 10 and 40 km). Although currently no spectroscopic observation is capable of such an extremely high spatial resolution, it is nevertheless instructive to look at the predicted individual line profiles. The general picture emerging from the numerical simulations may be summarized as follows.

The cores of absorption lines originating from the bright granular regions are blue shifted (relative to the laboratory wavelength), and the blue wing of the line profile is depressed relative to the red wing, resulting in a considerable line asymmetry. Lines formed in the dark intergranular lanes exhibit an even stronger asymmetry, but in the opposite direction, their cores being red shifted (Fig. 3). The horizontally averaged line profile, obtained as a superposition of the spatially resolved profiles, turns out to be much less asymmetrical than most of the line profiles seen at high spatial resolution.

The residual intensity in the cores of weak spectral lines, which are formed near the continuum-forming layers, varies across the granulation pattern in accordance with the continuum intensity. In contrast, the intensity in the cores of the stronger lines tends to be anti-correlated with the continuum intensity: the cores of these lines are darker in the granules and brighter in the intergranular lanes (Fig. 3). This behavior is due to the change of sign of the temperature fluctuations in the overshooting layers mentioned in Sec. II.D.1.

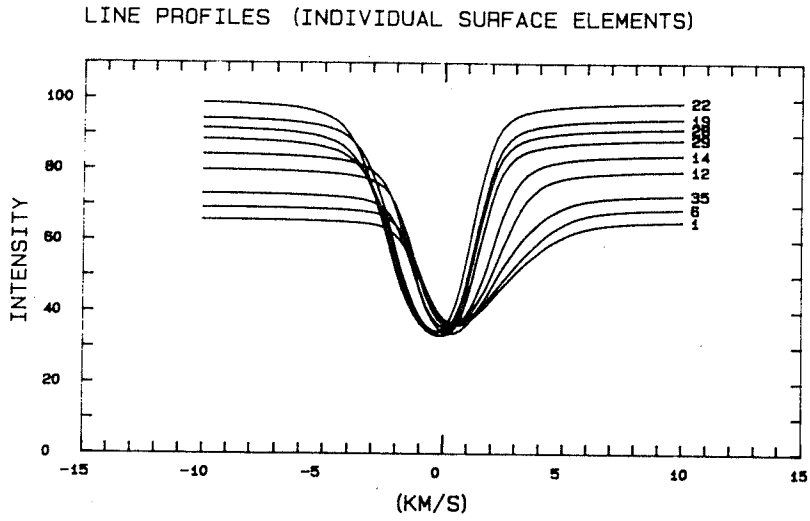

Fig. 3. Spatially resolved profiles of an artificial Fe I line at λ 6000 Å at disk center, calculated from a steady-state model measuring 1050 km in diameter. In granular regions with high continuum intensity, the line profiles are blue-shifted and exhibit a blue asymmetry, whereas the lines are red-shifted and show a strong red asymmetry in the intergranular regions with low continuum intensity. Numbers indicate the horizontal position within the cylindrical model seen from above (1 = near axis, 35 = near lateral boundary). The horizontally averaged line profile has an equivalent width of 53.5 mÅ. Referred to the local continuum, the line is weakest in the dark regions with an equivalent width of 42.5 mÅ and strongest in the bright granular parts, equivalent width 59.5 mÅ. From this model we obtain a continuum intensity contrast of δI_{rms} (6000 Å) = 11.5%.

It is encouraging to see that recent spectroscopic observations with high spatial resolution carried out at the Observatorio del Teide in the Canary Islands (Wiehr and Kneer 1988; Holweger and Kneer 1989) indeed confirm these spectroscopic characteristics predicted by our hydrodynamical models.

Figure 4 illustrates the situation in terms of the line bisectors. In the bright parts of the granulation, the corresponding line bisectors are inclined to the blue (the top portion near the continuum being blue shifted relative to the line core); in the dark intergranular regions, the line bisectors are inclined even stronger, but to the red. The different slopes of the spatially resolved bisectors reflect the different depth dependence of temperature and convective velocity at the various horizontal positions within a granular convection cell. Remarkably, the bisector of the horizontally averaged line profile exhibits the typical C-shape in close agreement with observation. Clearly, opposite asymmetries of the spatially resolved profiles cancel to a large degree when the spectrum is averaged over the granulation pattern. Spectroscopic evidence for such a behavior was found by Mattig et al. (1989).

A series of synthetic line spectra has been obtained from a nonstationary simulation for Fe II, λ 5197.6 Å. The bisector of the horizontally averaged

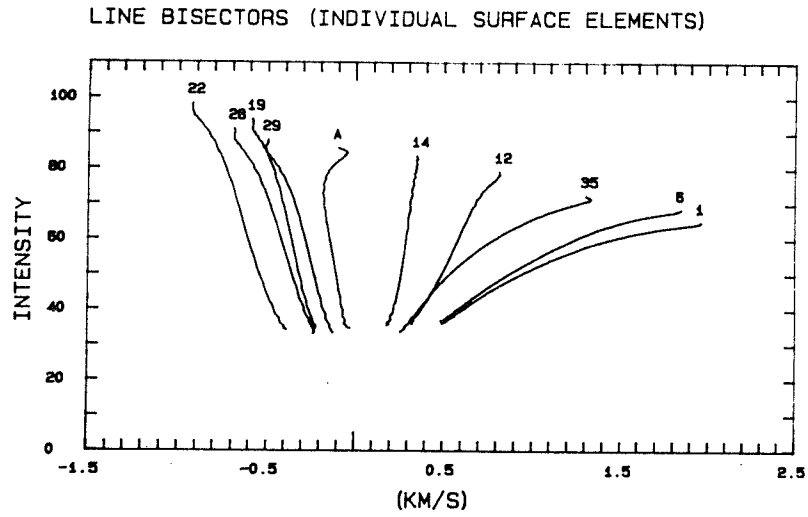

Fig. 4. Bisectors corresponding to the line profiles shown in Fig. 3. They are given on an absolute velocity scale (solar reference frame). Numbers indicate the horizontal position within the cylindrical model seen from above (1 = near axis, 35 = near lateral boundary). *A* denotes the bisector of the horizontally averaged line profile (C-shape).

line profile is shifted back and forth (typically ± 1 km s^{-1}) and heavily distorted as the flow evolves in time. However, the shape of a "C" is more or less retained; inverted "C's" were not found. Spatial and temporal averaging results in a line profile can be compared directly to standard solar spectra. In Fig. 5 we compare the observed line bisector of Fe II, λ 5197.6 Å with that of the mean synthetic line profile obtained by averaging 70 computed spectra (each one again an average of 35 spatially resolved spectra) separated in time by 100 s, covering a total interval of almost 2 hr. The agreement between observed and calculated bisector shape is excellent. However, the absolute convective blueshift of about 0.2 km s^{-1} (line core) seems too small if the value of $\Delta V = -0.8$ km s^{-1}, given by Dravins et al. (1986), is taken as a reference. On the other hand, the line broadening provided by the simulated photospheric velocity field is sufficient to account fully for the observed line width without invoking the usual *ad hoc* parameters micro- and macroturbulence.

Finally, an example of how the bisector of a synthetic spectral line varies across the solar disk is shown in Fig. 6, based on the snapshot model displayed in Fig. 2. Note that the large blueshift indicated by the bisectors is due to the fact that the whole upper atmosphere is moving upward at this instant. A time-averaged spectrum will result in a substantially smaller net blueshift (cf. Fig. 5). In qualitative agreement with observation, the bisector near the limb at $\mu = 0.2$ has the shape of an inverted "C".

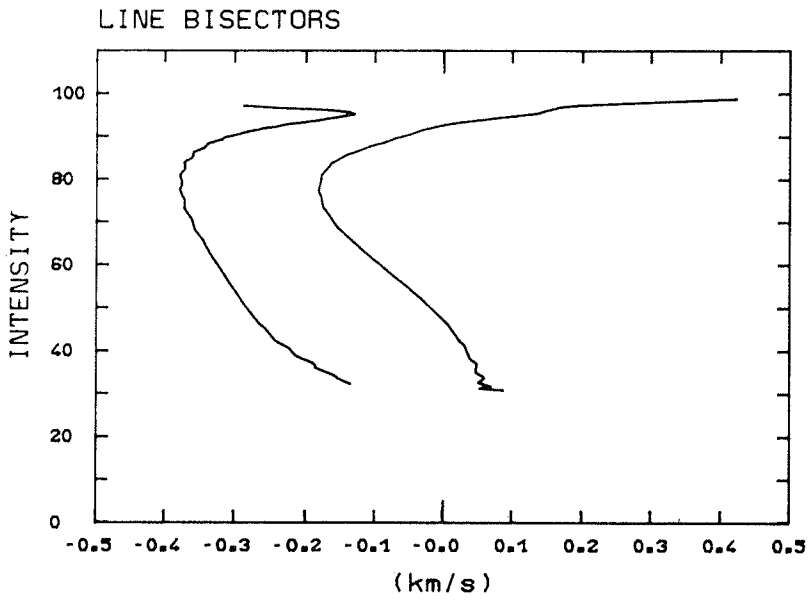

Fig. 5. Comparison of calculated and observed bisector of Fe II, λ 5197.6 Å at solar disk center, which has an equivalent width of roughly 90 mÅ. The observed bisector (right) is arbitrarily displaced relative to the calculated bisector (left, on an absolute velocity scale). Based on a nonstationary simulation in a cell measuring 2100 km in diameter, the synthetic bisector was obtained as a combined spatial and temporal average of 2450 individual line profiles. The differences near the continuum are due to a weak blend in the observed spectrum. (FTS observation courtesy of W. Livingston, National Solar Observatory, Tucson, U.S.A.)

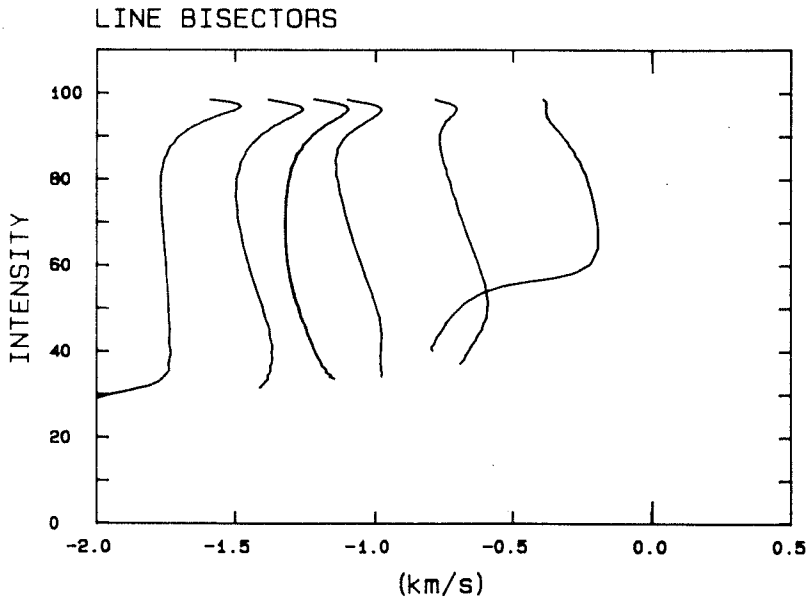

Fig. 6. Calculated center-to-limb variation of the bisector of Ni I, λ 6767.8 Å (GONG line). From left to right cos θ = μ is 1.0, 0.8, fd, 0.6, 0.4 and 0.2, respectively (fd = bisector of disk-integrated line profile). Spectrum synthesis is based on the model granule shown in Fig. 2 (no time average).

F. Conclusions and Future Prospects

We have calculated detailed hydrodynamical models of cylindrical photospheric convection cells. This two-dimensional problem requires considerably less computational effort than the full three-dimensional case. Based on the two-dimensional models, we have discussed the dynamical and thermal structure of the solar photosphere. The nonstationary simulations suggest that even the higher photosphere is not in radiative equilibrium but exhibits substantial temperature inhomogeneities induced by overshooting motions from the convection zone. Using spectrum synthesis techniques, observable quantities have been derived from our models. The results obtained so far indicate that basic properties of solar granular convection can be reproduced and understood by the two-dimensional model calculations. For example, the computed continuum intensity contrast and asymmetry of spectral lines compare well with observations.

We plan to perform further simulations of solar convection using two-dimensional rectangular geometry instead of axial symmetry, allowing the use of periodic boundary conditions in the horizontal direction. In this way, the development of exceedingly strong downdrafts at the lateral boundaries will be avoided. Another improvement under way is the implementation of non-gray radiative transfer. Simulations with the future code should be extended to somewhat deeper layers than the present models.

Apart from studying solar granulation, future applications of the two-dimensional simulations will include the exploration of the hydrodynamical conditions in the atmospheres of F- and A-type stars.

III. BEHAVIOR OF DEEP, EFFICIENT CONVECTION

A. The Search for Hydrodynamical Principles of Convection

In the central region of the solar convection zone (from 1 to 16 pressure scale heights below the photosphere), radiation is ineffective for energy transport. Over 98% of the outward energy flux is carried by convection. This is a region where convection is mainly controlled by purely hydrodynamical effects, little complicated by radiation; we call this *efficient convection*. It would be easier to extract the hydrodynamical principles (if any) of deeply stratified convection by studying this region. These principles are not only important for understanding the structure and evolution of the Sun, but similarly important for other stars.

The mixing-length theory of convection is a set of assumed hydrodynamical principles. It supposes that energy transport is performed by heat-carrying bubbles which travel and then dissolve in about 1 to 2 pressure or density scale heights (the mixing length). The extent of the bubbles is scaled by the mixing length. This picture is mainly based on considering the effects of stratification (Schwarzschild 1961). As neighboring rising bubbles travel

across some scale heights, their volumes and cross sections have to expand by a large factor. There would be no room for them to continue, and the pushing and squeezing would generate turbulent (twisting) motions which destroy the vertical coherence of the velocity (Chan et al. 1981). However, all numerical simulations of convection before 1980 did not reveal any scaling effects or dynamical significance associated with the scale heights. The only important vertical length scale found was the total depth of the convective region.

B. Controversy over Effects of Stratification

In 1982, Chan, Sofia, and Wolff reported that convection cells with sizes ranging from the total depth of the convection zone to the smallest scale height at the top of the zone co-exist in their two-dimensional simulation. Furthermore, the longitudinal correlation function of the vertical velocity was found to be scaled by the local scale heights. This report has generated controversies because for quite a few years after that, no other group obtained similar results. In particular, the flows of the other calculations would produce vertical auto-correlation coefficients close to unity throughout the total depths of the convection zones. Conclusive comparison could not be made between the calculations of the different groups because the model problems were very different. The study made by Chan et al. differed from the other studies in two important aspects. First, they studied efficient and turbulent convection. Effects of diffusion which tend to smother small-scale motions were minimized. Second, they paid special attention to resolving the scale heights at all depths. This of course is a necessary requirement for any study interested in the effects of scale heights.

Did some kind of numerical approximation or parameterization generate the reported behavior artificially? Specifically, the following questions have been raised: (1) Chan et al. used an alternating direction implicit method to increase the numerical speed of solving the Navier Stokes equations (Chan and Wolff 1982). The temporal truncation error of this method is relatively large due to the use of an operator-splitting technique (Chan 1983). Are the smaller cells simply numerical noise? (2) To parameterize the effects of subgrid-scale turbulence, Chan et al. used a nonlinear viscosity (Smagorinsky 1963) in their calculation. Can this added nonlinearity generate the smaller cells? These are valid concerns.

To examine these possibilities, Chan and Sofia (1986) made several three-dimensional calculations with different numerical techniques and viscosity models to intercompare. The scaling effects of the local scale heights persisted, even when a standard, explicit, time-marching scheme was used, as well as when a constant viscosity was used. Therefore, the behavior is robust.

The three-dimensional flows do not behave in a way identical to that conceived by the mixing-length picture (which is not surprising however).

Particularly, the upflows and downflows are highly asymmetric, and the stronger, more concentrated downflows can penetrate many scale heights, in agreement with those found by modal and two-dimensional calculations (see the references quoted in Sec. I). Recently, Stein and Nordlund reported that effects associated with stratification and scale heights appeared in their simulation of solar granules (see Sec. IV for a vivid description of such phenomena).

C. Dynamics of Efficient Convection

The demonstration of the effects of stratification, although important, is only one of the many aspects of efficient convection studied in the series of numerical work by Chan et al. To present an overall view on the behavior of this kind of convection, the following results are collected here to make a concise summary:

1. The energy-containing eddies of the convective turbulence decrease in size toward the top region where the lengths of the scale heights drop. The amplitudes and time scales of the fluctuation of the convective flux also decrease toward the top (Chan et al. 1982). The implication of this is that while the brightness fluctuation at the solar surface is only moderate, the absolute fluctuation can be much larger in the interior.
2. The convective velocity, temperature fluctuation and the enthalpy flux can be approximately computed from the mean superadiabatic gradient, with mixing-length type formulae (Chan and Sofia 1987). The result thus gives support to the mixing-length theory not only qualitatively, but also quantitatively.
3. The vertical correlation lengths of the vertical velocity and the temperature fluctuation are both scaled by the pressure scale height, not by the density scale height. This provides evidence that the mixing length in the mixing-length theory is indeed scaled by the pressure scale height, as most stellar evolution codes use.
4. In the upper part of the convection zone, vorticity in the vertical direction tends to associate with funnel-like downflows (Chan and Sofia 1986). This is similar to the behavior of the flows obtained by Nordlund (1985*b*). The correlation coefficient of the downward velocity and the absolute vertical vorticity is, however, found to be small, only ~ 0.1. This is consistent with the rare occurrence of fully developed vortices in the photosphere (Brandt et al. 1989).
5. Contrary to an implicit assumption of the mixing-length theory, the kinetic energy generated by the buoyancy can work form a large energy flux with amplitude approaching the size of the total flux. The direction of the mechanical flux is downward in most of the convection zone. This confirms the modal result of Massaguer and Zahn (1980) and the two-dimensional result of Hurlburt et al. (1984). The distribution of this flux is poorly

modeled by the diffusion approximation, but is found to be scaled by F/C_p where F is the total flux and C_p is the specific heat under constant pressure (Chan and Sofia 1989).

6. The production and dissipation of the kinetic energy do not parallel each other. Production is scaled by the total flux, and the local production rate is essentially a function of the local mean variables (see Sec.III.D). The dissipation is clearly nonlocal; a significant amount of kinetic energy is carried away from the production region, to be dissipated in lower regions.
7. The effective viscosity generated by the convective turbulence can be roughly estimated as $V_z'' H_p/3$ where V_z'' is the root-mean-square vertical velocity and H_p is the pressure scale height (Chan et al. 1987).
8. Long-lived oscillations co-exist with the convective turbulence (Chan and Sofia 1988). Their frequencies are shown to be almost identical to the eigenfrequencies of acoustic modes.

Chan and Sofia (1989) have compiled a list of numeral-empirical formulae which describe the quantitative relationships among the mean variables, root-mean-square fluctuations and correlations of the convective turbulence. However, the variables used in that paper are dimensionless. For the convenience of application, the list is recast in Table III in terms of dimensional variables. These formulae contain interesting information about the thermodynamical behavior of efficient convection. Below, two of them will be used to show that efficient convection is a very peculiar heat engine.

D. Efficient Convection as a Heat Engine

If no motions were allowed in the solar convection zone, the outward transfer of energy would have been very difficult. With convection, the process is much easier, but a price must be paid for moving the fluid against the turbulence (eventually the gas viscosity). This supply of energy comes from the heat entering at the bottom with a higher temperature T_b and leaving the top with a lower temperature T_t (see Fig. 7); considered in this way, the convecting gas is a heat engine. The rate of work W is the production rate of kinetic energy, which is eventually dissipated back into heat; therefore, the mean heat flux F going through the upper level must be the same as that going through the lower level.

At first glance, one may think that W is restricted by the Carnot limit: $(1-T_t/T_b)F$. However, we now show that this limit can apparently be exceeded. Applying the formulae 15 and 20 in Table III to relate the local production rate of kinetic energy w to the flux F, one obtains

$$w = -\langle V_z \rho' \rangle g = \langle V_z \rangle \langle \rho \rangle g \sim 0.8\,(F/C_p)\,(g/T) \qquad (12)$$

where V_z, ρ', g are the vertical velocity, density fluctuation and the gravitational acceleration, respectively; the brackets $\langle\ \rangle$ denote averaging. At the

TABLE III
A List of Approximate Relationships Obtained from fitting Three-Dimensional Numerical Results

Identifier	Approximate Formula
1	V_x'' (or V_y'') $\approx 0.61\ V_z''$
2	$\rho''/\langle\rho\rangle \approx 0.89\ T''/\langle T\rangle$
3	$p''/\langle p\rangle \approx 0.57\ T''/\langle T\rangle$
4	$S'' \approx 0.94\ C_p T''/\langle T\rangle$
5	$p'' \approx 0.26\ \langle\rho\rangle\ V''^2$
6	$p'' \approx 0.51\ \langle\rho\rangle\ V_z''^2$
7	$T''/\langle T\rangle \approx 0.90\ \langle\rho\rangle\ V_z''^2/\langle p\rangle$
8	$C[T',S'] \approx 0.99$
9	$C[\rho',S'] \approx -0.89$
10	$C[\rho',T'] \approx -0.82$
11	$C[p',T'] \approx 0.49$
12	$C[V_z,T'] \approx 0.81$
13	$C[V_z,S'] \approx 0.81$
14	$C[V_z,\rho'] \approx -0.74$
15	$\langle V_z\rho'\rangle \approx -\langle V_z\rangle\langle\rho\rangle$
16	$\langle V_z p\rangle \approx 1.24\ \langle V_z\rangle\langle p\rangle$
17	$\langle V_z T'\rangle \approx 1.26\ \langle V_z\rangle\langle T\rangle$
18	$\langle V_z S'\rangle \approx 1.20\ C_p\langle V_z\rangle$
19	$\langle V_z\rangle \approx 0.58\ \langle\rho\rangle\ V_z''^3/\langle p\rangle$
20	$F_{ep} \approx 1.25\ (C_p/R)\ \langle p\rangle\langle V_z\rangle$
21	$F_{ep} \approx 0.72\ (C_p/R)\ \langle\rho\rangle V_z''^3$
22	$T''/\langle T\rangle \approx 1.05\ \Delta\nabla + 0.0027$
23	$\langle\rho\rangle V_z''^2/\langle p\rangle \approx 1.17\ \Delta\nabla + 0.0032$
24	$\Delta\nabla \approx 1.04\ [(R/C_p)F]^{2/3}\ \langle\rho\rangle^{1/3}\langle p\rangle^{-1} - 0.002$
25	$\langle V_z\rho V^2\rangle \approx 1.03\ \langle\rho\rangle\langle V_z V^2\rangle$
26	$\langle V_z V^2\rangle \approx 1.13\ \langle V_z^3\rangle$
27	$\langle p'\nabla\cdot V\rangle \approx \langle V_z p'\rangle/H_\rho$

Notations

$'$	fluctuation with respect to the mean
$''$	root-mean-square (*rms*) deviation
$\langle\ \rangle$	the mean value at a certain depth
$C[,]$	correlation coefficient
C_p	specific heat under constant pressure
F	total flux
F_{ep}	enthalpy flux
H_ρ	density scale height
p	pressure
ρ	density
R	gas constant
T	temperature
S	entropy
$V_{x,y}$	horizontal velocities
V_z	vertical velocity
V''	*rms* velocity

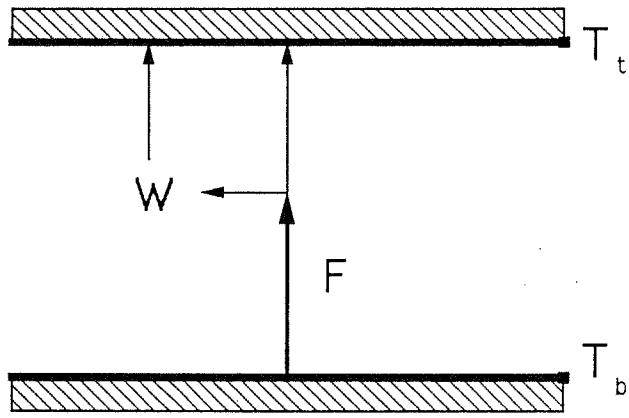

Fig. 7. The convection zone considered as a heat engine operating between two temperatures. The temperature at the bottom of the zone T_b is higher than the temperature at the top T_t. When a heat flux F flows through the system, a certain amount of mechanical power W is generated to drive the fluid motions.

moment, without affecting the validity of our argument, we ignore the difference between the total flux F and the enthalpy flux here. Integrating w over the depth z of the convection zone, one obtains

$$W = \int w \, dz = 0.8 \, (F/C_p) \int (g/T) \, dz \sim 0.8 \, F \, (R/C_p) \int d(\ln p) \sim 0.8 \, F \ln(T_b/T_t) \tag{13}$$

for which the hydrostatic approximation $d(\ln p)/dz \sim g/RT$ and the almost-adiabatic approximation $\ln(p_b/p_t) \sim (R/C_p) \ln (T_b/T_t)$ have been used. Equation (13) shows that W is proportional to the total number of temperature scale heights in the convection zone; it can make W greater than F. Does this mean that the laws of thermodynamics are being violated?

The Carnot limit given above is based on taking the whole convection zone as a single engine, but how about considering the zone as composed of a series of heat engines stacking one over another as shown in Fig. 8. For the convenience of analysis, let us suppose that the ratios of temperature drops that the engines operate on are the same; namely, $T_{i-1}/T_i = \lambda$, where $\lambda > 1$ is a constant. To satisfy the boundary conditions, $\lambda^N = T_b/T_t$ where N is the total number of engines. Now the total allowed power is

$$W = \Sigma \, W_i = F \, N \, (1 - 1/\lambda) = F \ln(T_b/T_t) \, (1 - 1/\lambda)/\ln(\lambda). \tag{14}$$

The optimal value of this sum is obtained as $\lambda \to 1$; the upper limit is now

$$W \leq F \ln(T_b/T_t). \tag{15}$$

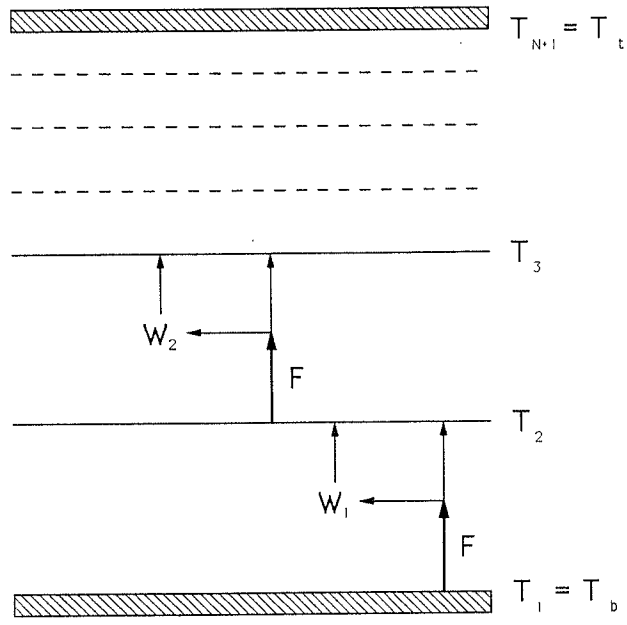

Fig. 8. The convection zone considered as a stack of heat engines. The i-th engine operates between temperatures T_i and T_{i+1}; it can generate mechanical work at the rate W_i.

Therefore, Eq. (13) is only at the 80% level of this limit and is in line with the laws of thermodynamics. One may consider such nonoptimal efficiency to be due to the existence of nonadiabatic heat diffusion. Alternatively, one could assume that the engines have 100% efficiency, then Eq(14) can be used to deduce a value of 1.6 for λ, implying that the engines operate between levels separated by about half a temperature scale height. The real situation should be somewhere between these two extremes.

The story does not stop here. When the buoyancy work integrals are evaluated directly with numerical data from the computations, some cases (with low C_p) show that W can even exceed slightly the limit given by Eq.(15) (on the order of 1%).

This turns out to be caused by the negative flux of kinetic energy that feeds energy to the lower portion of the convection zone. In the above derivation, we have ignored the difference between the enthalpy flux and the total flux. In fact, the enthalpy flux is not uniform and is larger than F in most of the convective region because it has to balance the substantial negative flux of kinetic energy. From another point of view, one can say that the back-feeding of mechanical energy to the lower portion of the convective region enhances the energy supply to the enthalpy flux. The sum over W_i in Eq.(14) can actually be larger.

The example shown in Fig. 9 illustrates the operation of the energy loop-back process. For simplicity, a single heat engine is considered here. The

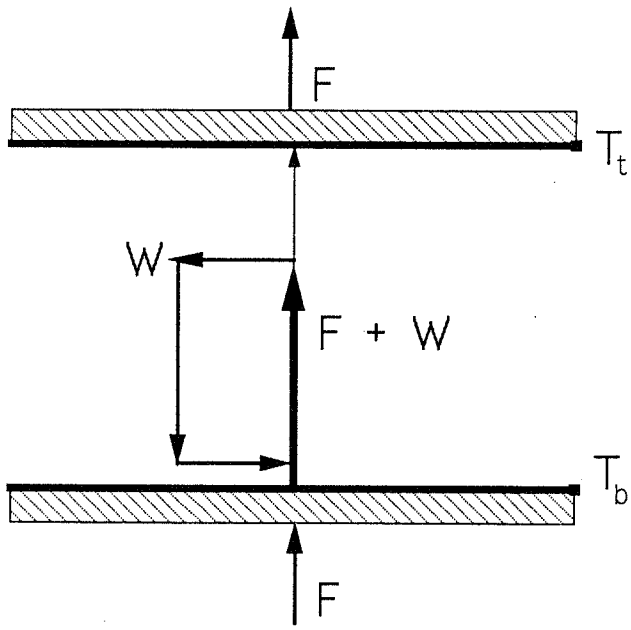

Fig. 9. An illustration of the loop-back process. Mechanical energy produced in the upper region is transported to the lower region, dissipated back to heat, and fed into the thermal flux inside the system.

mechanical energy, instead of being dissipated and fed to the low-temperature reservoir (top) as assumed previously, is now transferred to the bottom, dissipated there, and looped into the thermal flux. The Carnot limit for such a situation can be written as:

$$W < (1 - T_t/T_b)\,(F + W) \tag{16}$$

which can be rearranged as

$$W < (T_b/T_t)\,(1 - T_t/T_b)\,F. \tag{17}$$

This limit is much more liberal than (Eq. 15). The loop-back effect is more significant for cases with smaller C_p because the magnitudes of the mechanical fluxes are relatively larger (see point 5 in the previous section—III.C).

In all of the above arrangements, the real trick is that W does not deliver energy to outside the system. It does not generate real work for the external world. But taken at face value, the convection zone can indeed generate an amazing amount of mechanical power.

IV. REALISTIC, 3D SIMULATIONS OF SOLAR CONVECTION

A. Introduction

The nature of the upper solar convection zone can be explored using numerical simulations, and we have made such a model of the upper 2.5 Mm of the convection zone and the photosphere. By using a realistic equation of state and by including radiative energy transfer, we have been able to compare the results directly with observations. The simulation shows that the topology of solar convection is dominated by effects of stratification, and that convection consists of broad, gentle, structureless, warm, adiabatic, diverging upflows, with embedded filamentary, cool, fast, twisting, converging downdrafts. The flow topology is hierarchical, with downflows around many small cells close to the surface merging into fewer, filamentary downdrafts at greater depths. This merging of downdrafts into fewer, more widely separated plumes may persist through the entire depth of the convection zone. Radiative cooling at the surface provides the entropy deficient plasma that drives the circulation. A comparison of observable features from the simulation with recent granulation and mesogranulation observations shows that they are in accord.

B. Simulation

The upper solar convection zone is modeled by solving the equations of hydrodynamics, i.e., the equation of mass conservation

$$\frac{\partial \ln \rho}{\partial t} = -\mathbf{u} \cdot \nabla \ln \rho - \nabla \cdot \mathbf{u} \tag{18}$$

the equation of momentum conservation

$$\frac{\partial \mathbf{u}}{\partial t} = -\mathbf{u} \cdot \nabla \mathbf{u} + \mathbf{g} - \frac{P}{\rho} \nabla \ln P \tag{19}$$

and the equation of energy conservation

$$\frac{\partial e}{\partial t} = -\mathbf{u} \cdot \nabla e - \frac{P}{\rho} \nabla \cdot \mathbf{u} + Q_{\text{rad}} + Q_{\text{visc}}. \tag{20}$$

These are rewritten from Eqs. 1–4. Here ρ is the mass density, $\mathbf{u}$ is the velocity, e is the internal energy, P is the pressure, and Q_{rad} and Q_{visc} are the radiative heating and viscous dissipation, respectively. In order to model actual solar convection as realistically as possible, we use a tabular equation of state $P = P(\rho,T)$, which includes ionization and excitation of hydrogen and other

abundant atoms, and formation of H_2, CO and other molecules (Gustafsson 1973).

This form of the equations, in terms of per-unit-mass quantities rather than the conventional per-unit-volume quantities, is advantageous when dealing with a strongly stratified atmosphere. The per-unit-volume quantities (density, momentum and energy per unit volume), vary exponentially with depth, and we are presently performing simulations with density ratios bottom/top of the order of 10^5. The logarithm of the density and the energy per unit mass vary nearly linearly with depth, both in the optically thin atmosphere and in the subsurface layers. Also, the velocities vary much less with depth than the linear momenta. The per-unit-mass quantities therefore allow a more accurate finite difference representation of derivatives. Although the per-unit-volume variables offer the possibility of exact conservation of mass, momentum and energy, small errors in mass and energy conservation are harmless in a stratified model, because the hydrostatic and energy equilibria enforce a very well-defined mean state. These equations were tested for one-dimensional shock tubes and found to give a good representation of the solution for pressure jumps up to one million.

We evaluate the radiative energy exchange

$$Q_{\text{rad}} = \int_{\lambda}\int_{\Omega} \kappa_\lambda (I_{\lambda,\Omega} - S_\lambda) \mathrm{d}\Omega \mathrm{d}\lambda \tag{21}$$

by solving the transfer equation for the specific intensity $I_{\lambda,\Omega}$

$$\frac{\mathrm{d}I_{\lambda,\Omega}}{\mathrm{d}\tau_\lambda} = I_{\lambda,\Omega} - S_\lambda \tag{22}$$

along inclined rays. $\mathrm{d}\tau_\lambda$ is the optical depth increment along the ray, and S_λ is the source function. We assume that the source function is equal to the Planck function, and approximate the detailed wavelength integral with a sum over 4 bins. Absorption coefficients are calculated with a standard stellar atmosphere code (Gustafsson et al. 1975), and are sorted into bins representing continua, weak lines, intermediate lines and strong lines (cf. Nordlund 1982; Nordlund and Dravins 1990). This treatment is sufficiently elaborate to describe the sudden release of radiation by ascending hot gas in a thin layer at the solar surface, and the subsequent re-absorption of a small but energetically significant fraction of this energy in the upper layers of the photosphere.

Nonlinear time-stepping methods, such as van Leer's monotonic second-order upwind method (Van Leer 1977), and Colella and Woodward's piecewise parabolic method (Colella and Woodward 1984) are difficult to combine with radiative energy transfer and magnetic fields. As we are primarily inter-

ested in using our code as a test bed for understanding the interaction of convection, radiation and magnetic fields near stellar surfaces, conceptual simplicity and ease of implementation are important factors in our choice of numerical methods.

We advance the numerical solution in time with the Adams-Bashforth method, chosen for its high accuracy and modest memory requirements (cf. Gear 1971; Gazdag 1976). Because of a truncation error which corresponds to a weak but *negative* diffusion, Adams-Bashforth time stepping is weakly unstable and requires a small compensating positive diffusion to be added to all equations for stability. Stable methods were tested on model problems and found to offer no advantage, but to require more memory or input/output. Most stable methods (except those with large truncation errors) still need numerical diffusion near shocks. Independent of whether the diffusivity is provided explicitly (through spatial numerical diffusion) or implicitly (through truncation errors in the time stepping), the required net diffusivity is similar, and so are the actual results. As the magnitude of the numerical dissipation is decreased, the effective resolution of a code increases. If the dissipation is too small, structures smaller than the code can resolve and develop, and eventually the time evolution becomes unstable.

We use a diffusion coefficient with three types of contributions: (1) a term proportional to the sound speed, to stabilize the Adams-Bashforth time stepping; (2) a term proportional to the fluid velocity, to prevent ringing at sharp changes in advected quantities; and (3) a term proportional to the finite difference velocity convergence (where positive) to stabilize shock fronts.

One of the main difficulties with applying these equations to a specific simulation problem is the treatment of the "virtual boundaries"; i.e., boundaries of the computational domain that do not correspond to real boundaries, but just delimit the volume we choose (or can afford) to simulate. We deal with this problem by using periodic horizontal boundary conditions, and by constructing top and bottom boundary conditions which are as transmitting as we can easily make them while still preserving stability. At the top, we take an extra large boundary zone ($\geq$ scale height). In this zone, we impose the conditions that the amplitude of the velocity and the density fluctuations remain constant, while the energy density at the boundary is fixed at its initial average value. At the bottom, we impose constant pressure by adjusting the density, and we require $\partial \mathbf{u}/\partial z = 0$. The vertical heat flux is kept from drifting by specifying the internal energy of inflowing material at the bottom boundary. For additional discussion of our numerical methods see Nordlund and Stein (1990).

We simulated a region 6×6 Mm horizontally, in order to cover scales at least marginally larger than granulation, and extending vertically from the temperature minimum (-0.5 Mm) to a depth of 2.5 Mm (Stein and Nordlund 1989). We used a grid of $63 \times 63 \times 63$ points, which gave a resolution of

95 km horizontally and 50 km vertically. The calculations were performed on an Alliant FX/80 at the University of Colorado.

C. Results

1. Surface Topology: Granulation. The granulation pattern visible at the solar surface consists of hot upflowing plasma in disconnected cells, surrounded by cold downflowing plasma in narrow topologically connected intergranular lanes (Fig. 10). The primary process that occurs in the granulation is the radiation of energy from the hot upwelling fluid, which cools it and reduces its entropy. Higher pressure in the hot, ascending granule centers pushes the cooling fluid toward the intergranular lanes. The entropy deficient fluid then starts to sink under the pull of gravity.

The granulation pattern is asymmetrical with respect to the direction of time, as well as with respect to the connectivity of ascending and descending gas and, indeed, the two are closely related. The ascending fluid, which expands horizontally because of the density stratification, meets expanding fluid from neighboring cells along common borders. At these cell borders, the flow is deflected horizontally along the border and vertically downward. The flows along borders eventually converge at the common corners of three or more cells. The "arrow of time" is provided by the dissipative processes in conjunction with the asymmetry between the expansion of the ascending flow and the convergence of descending flow.

The horizontal flow is driven by horizontal pressure fluctuations, which in turn are caused by temperature fluctuations and the Bernoulli effect. In the anelastic approximation ($\nabla \cdot (\rho \mathbf{u}) = 0$), the pressure is determined by

$$\nabla^2 P = \nabla \cdot [\rho(P,T)\mathbf{g} - \rho \mathbf{u} \nabla \cdot \mathbf{u}] \tag{23}$$

which shows that the instantaneous pressure field is determined by the divergence of the forces. The first term on the right-hand side is the gravitational force, and the second term is the inertial "force".

Qualitatively, the resulting pattern of horizontal pressure fluctuations is easy to understand. For small velocities and sufficiently large horizontal scales the vertical derivatives dominate; the vertical pressure gradient force and gravity almost balance, i.e., large-scale fluctuations are close to hydrostatic equilibrium. In hydrostatic equilibrium, the change of pressure with height is determined by the local pressure scale height, which is proportional to temperature. As temperature fluctuations decrease rapidly with depth in a stratified convection zone, the pattern of horizontal pressure fluctuations at any given depth is dominated by the temperature fluctuations in the next few scale heights below that depth. Thus, the pressure excess in a granule (which provides the driving for the horizontal flow from the granule centers to the surrounding intergranular lanes), is a consequence of the temperature excess of the ascending granular flow just below the surface.

GRANULATION

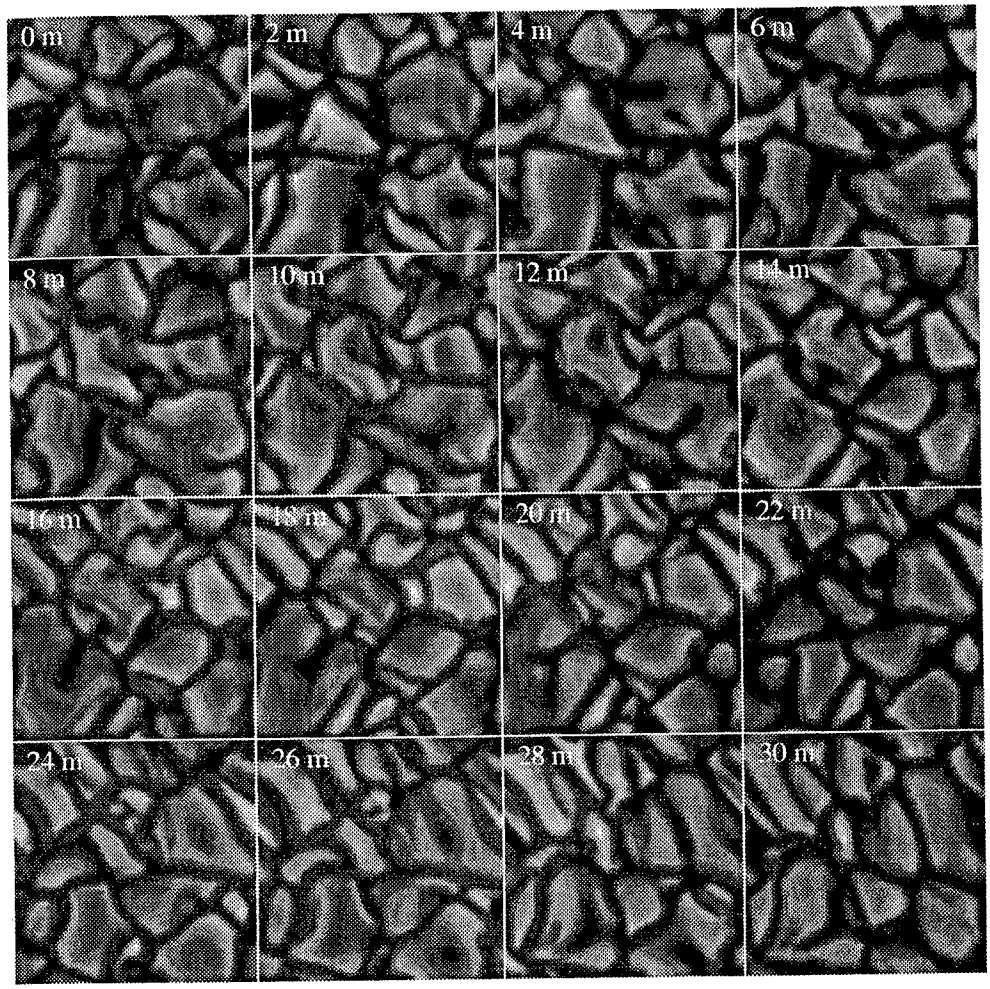

Fig. 10. Sixteen simulated snapshots of integrated radiation intensity at the solar surface. These approximate monochromatic continuum intensity snapshots. Each snapshot is 6 × 6 Mm (8″ × 8″). The sequence spans 30 min.

In the center of intergranular lanes, the inertial forces associated with the convergence of flows from neighboring cells also causes local pressure enhancements. This is obvious both from direct inspection of Eq. (23) above, and from physical grounds; the horizontal flow towards the intergranular lanes must be decelerated by a local pressure excess there. Hurlburt et al. (1984) have pointed out that this may be viewed as a Bernoulli effect along approximately horizontal streamlines near the solar surface—the pressure is a maximum where the horizontal velocity is a minimum, at the cell centers and in the intergranular lanes.

What determines the size of granules? When attempting to answer this question, one should keep in mind that there is no one size of granules, but

rather a continuous distribution of sizes (and shapes), with a relatively well-defined upper limit to the size of single bright patches. This upper limit may be understood in terms of the constraints from mass conservation, pressure and energy balance.

The larger the granule, the larger the horizontal velocities needed to carry the increasing amounts of overturning gas. The amount of ascending fluid is proportional to the horizontal surface area which scales quadratically with the linear size of the granule, whereas the granule boundary, through which it must flow, only scales linearly with the size. The pressure fluctuations required to accelerate this material horizontally scale as the square of the velocities, and hence as the square of the linear size of the granule. This scaling has two important, and in some sense, opposite consequences. On the one hand, the larger pressures achieved by a large granule will tend to force the common border with neighboring cells to expand, thus further increasing the size of the larger granule. On the other hand, the increasing pressure in the interior of the granule decreases the buoyancy and eventually leads to buoyancy braking of the ascending gas in the center of the cell. This happens when the pressure excess in the granule is sufficient to cancel the buoyancy due to the temperature excess. For the Sun, the temperature excess is of the order of 2 (about 11,000 K in granules relative to about 6000 K on the average), and hence the excess pressure can support close to sonic horizontal velocities. The temperature excess can be supported and the granule continues to grow, as long as the rate of energy advection to the surface

$$\rho(e + P/\rho)u_z + \frac{1}{2}\rho\mathbf{u}^2 u_z \tag{24}$$

exceeds the rate of radiation, σT^4. Once the vertical velocity has been reduced to

$$u_z \leq \frac{\sigma T^4}{\text{several} \times nkT} \tag{25}$$

the vertical advection of excess entropy to the surface is no longer able to supply the entropy lost through radiation, and the granule center begins to cool, thus strangling the granule through a lack of heat input (cf. Nordlund 1985*b*). The critical ascent velocity is of the order of 2 km s^{-1}. A large granule, which is permitted to grow with little influence from neighboring granules, may develop into a ring of hot, ascending material, surrounding a cool, dark and eventually descending granule center, produced by the buoyancy braking (cf. Fig. 10; Nordlund 1985*b*). This phenomenon has been called exploding granules in the literature.

For a rough estimate of the ratio of ascent vertical to horizontal velocity

as a function of granule size, assume that the ascending flow has cylindrical symmetry and a scale height of the vertical mass flux equal to $H_{\rho u_z}$. Then, if the medium is anelastic and the ascending and expanding fluid is overturning at a radius of r, continuity requires

$$\pi r^2 \rho u_z / H_{\rho u_z} = 2\pi r \rho u_r \tag{26}$$

or

$$d \equiv 2r = 4H_{\rho u_z} (u_r/u_z) \tag{27}$$

From the pressure equation, Eq. (23), it can be shown that large-scale pressure fluctuations (and hence velocity amplitudes) decrease relatively slowly with height as compared to the pressure itself (cf. Nordlund 1982), so the density factor dominates the mass flux scale height.

Collecting the constraints from the continuity equation (the ratio of ascent velocity to horizontal velocity), the momentum equation (the maximum horizontal velocity), and the energy equation (the minimum ascent velocity), we obtain an estimate of the maximum granule size by inserting the sound speed (8 km s^{-1}) for u_r, the density scale height (0.2 Mm) for $H_{\rho u_z}$, and 2 km s^{-1} for the minimum ascent velocity. The result is a diameter of about 3 Mm, in good agreement with observations of solar granulation (Bray et al. 1984).

Using an earlier, anelastic version of the present code, Nordlund and Dravins (1990) found that, for stars in the vicinity of the Sun in the Hertzsprung-Russell diagram, the size of granules scales roughly as the density scale height in the photosphere.

Apart from this type of "self-inflicted" death of relatively large granules, the simulations show granules breaking up because of influence from neighboring granules, whose expanding flows and pressure fluctuations constitute a highly time-dependent environment. The external perturbations experienced by an individual granule are not arranged in a nicely symmetric pattern around the granule. Rather, the neighbors surrounding any particular granule are likely to be of different strengths; some strong, some weak. Neighboring granules with large expansion velocities inhibit granule growth in that direction. Thus, the external influence from neighboring granules results in distortion of the shape, or even break up, of a granule.

From the perspective of an individual granule, the influence from surrounding granules is a pseudo-random function of space and time; although the surrounding granules may undergo similar evolutions, the evolutions are not in phase. As a consequence, a pattern consisting of many granules evolves chaotically. The evolution appears subjectively chaotic, and is indeed most likely mathematically chaotic in the sense that two neighboring points

in phase space depart exponentially with time. A pattern of granules evolves through many "points of decision," where the evolution may go one way or the other, and a slight perturbation may tip the balance in favor of one granule or the other. Two slightly different configurations (phase-space points) will take different paths through such "points of decision," and thus rapidly diverge. Such "points of decision" occur, for example, when granules of nearly equal size compete for space. They also occur when a granule is breaking up, and details of the external velocity field may have a decisive influence. For instance, tongues of protruding cool material from surrounding granules aid in the break up of a granule. An increase of the small-scale structure in granules increases the interaction between them; highly structured flows have a harder time arranging a "peaceful coexistence."

Initially, we used overly conservative estimates of the necessary numerical diffusion (away from shocks), which caused a lack of small-scale details in granules (Lites et al. 1989). Tests showed that the coefficients that control viscosity in nonshocking parts of the flow could be reduced by about a factor of 3, while still retaining stability. As expected, the reduction of the viscosity resulted in increased small-scale structure, and increased the number of small granules in the simulation. However, we found that the size of even the largest granules depended slightly on the amount of numerical diffusion. Diffusion increases the smoothness of the simulated granules, and hence delays the break up of granules.

2. *Subsurface Topology: Mesogranulation.* Beneath the surface, the connected intergranular lane downflow converges into topologically disconnected, finger-like structures (Fig. 11). The flow topology becomes large-scale, slow, diverging, structureless upflow of warm plasma, with embedded, twisting, narrow, isolated, fast, converging downdrafts of cool plasma (cf. Graham 1975; Nordlund 1985*b*; Chan and Sofia 1986). This change of topology takes place over a vertical distance ($\simeq$ 0.5 Mm) which is only a fraction of the typical horizontal cell size. To understand this remarkable change of topology, we proceed to discuss qualitatively the properties of convection below the solar surface, basing the discussion partly on the numerical results from our simulations, and partly on inspection of the governing equations.

The flow below the visible surface may be characterized as almost pure advection. Mass is advected with only small Eulerian changes of the density (i.e., almost anelastically), and entropy is advected with negligible influence from radiation and dissipation (i.e., almost adiabatically). Given the negligible energy exchange, the flow may be understood in terms of *fluid parcel trajectories;* the properties of a fluid parcel at a certain time is the result of the histories of its constituent parcels.

The flow topology is primarily controlled by the density stratification. The continuity equation, Eq. (18), may also be written

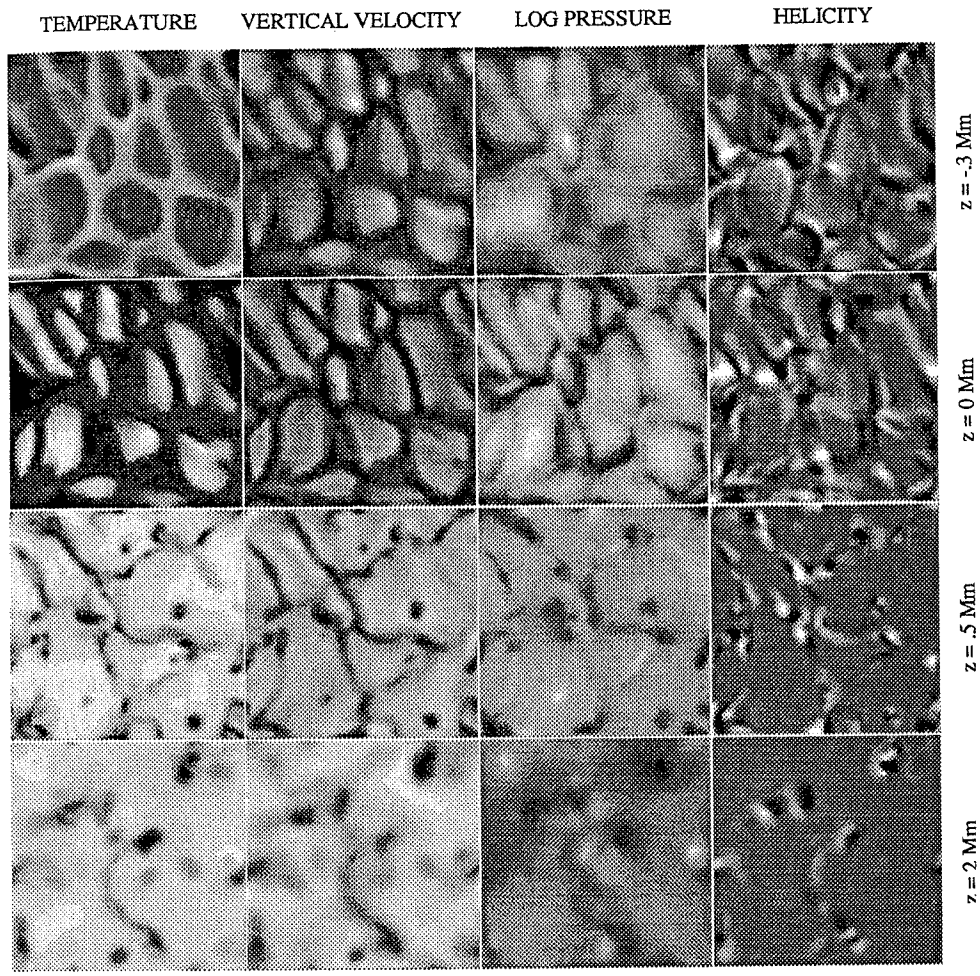

Fig. 11. Horizontal slices, showing temperature, vertical velocity, log pressure and helicity at depths -0.35, 0 (visible surface), 0.5 and 2.0 Mm. The flow topology changes radically below the surface, from isolated cells of warm upflowing plasma surrounded by connected lanes of cool downflowing plasma, to isolated downdrafts of cool plasma embedded in a broad gentle upflow of warm plasma. The pressure at the surface is high in the granule centers, low at the granule boundaries, and often has a secondary maximum in the intergranule lanes. In addition, the pressure shows a larger-scale variation corresponding to the meso-scale subsurface topology.

$$\mathrm{div}(\mathbf{u}) = \frac{\mathrm{D}}{\mathrm{D}t}(-\ln\rho) \tag{28}$$

which simply states that the expansion of the fluid flow is given by the relative rate of decrease of the density, following the motion. To be specific, a fluid parcel ascending from the bottom of our computational box (at $z = 2.5$ Mm)

to the solar surface (at $z = 0$ Mm) expands by a factor of about 90; e.g., its linear size would increase by about a factor of 4.5 if the expansion were isotropic. In principle, one could imagine the expansion to be entirely in the vertical, but this would not be consistent with the energy equation (constant total energy flux is obtained with a much smaller variation of vertical velocity with height). The ascent velocity actually varies with depth approximately as $\rho^{-0.5}$, and thus the linear horizontal size of a fluid parcel increases with about a factor of 3 over this 2.5 Mm depth interval. Because of the horizontal expansion, only a small fraction (about 10%) of the ascending fluid at the lower boundary ever makes it to the surface. Conversely, descending fluid at the surface contracts as it descends. Descending fluid parcels have a strong tendency to be sheared and stretched out into long, twisting filaments. Thus the horizontal size of a descending fluid parcel decreases even more rapidly than the horizontal size of an ascending fluid parcel increases. Consequently, the entropy deficient fluid from the surface rapidly becomes a smaller and smaller fraction of the descending fluid. The rest of the descending flow is made up of overturning fluid that never made it to the surface (and hence never lost any entropy).

Thus, the topology of the flow is primarily controlled by the continuity equation. The role of the energy equation is subsidiary; below a thin layer near the surface, the flow is very nearly adiabatic, and the energy equation simply traces the path of the entropy-deficient fluid from the surface as it descends, and the (mostly numerical) spatial diffusion of the entropy fluctuations. The tendency for the entropy fluctuation to spread horizontally is counteracted by the horizontal convergence of the descending flow. This keeps the descending cool filaments well defined.

Since the entropy-deficient fluid from the surface becomes a smaller and smaller fraction of the descending flow, the relative temperature fluctuations decrease rapidly with depth. These relative temperature fluctuations determine the pressure fluctuations (via the pressure scale height; cf. earlier discussion), which in turn determine the velocity amplitudes, which closes the causal loop. A selfconsistent (average) state is characterized by a vanishing average mass flux (pressure equilibrium), and a constant total energy flux (energy equilibrium; sum of radiative, convective and kinetic energy flux is constant).

Figure 12 shows the origin and destination of fluid parcels ascending through the visible surface (plane $z = 0$) at time $t = 0$. Most of these parcels were also ascending at 9 solar min earlier. They ascend slowly, with nearly the same speed, and all originate from a small source volume, both vertically (because of the nearly uniform vertical velocity) and horizontally (because of the diverging topology of the upflow). Going back further in time, we find the particles in progressively smaller source volumes. Only a small subset of all the ascending fluid at depth reaches the surface. Most of the fluid that does reach the surface, turns around and descends rapidly, concentrating into a few

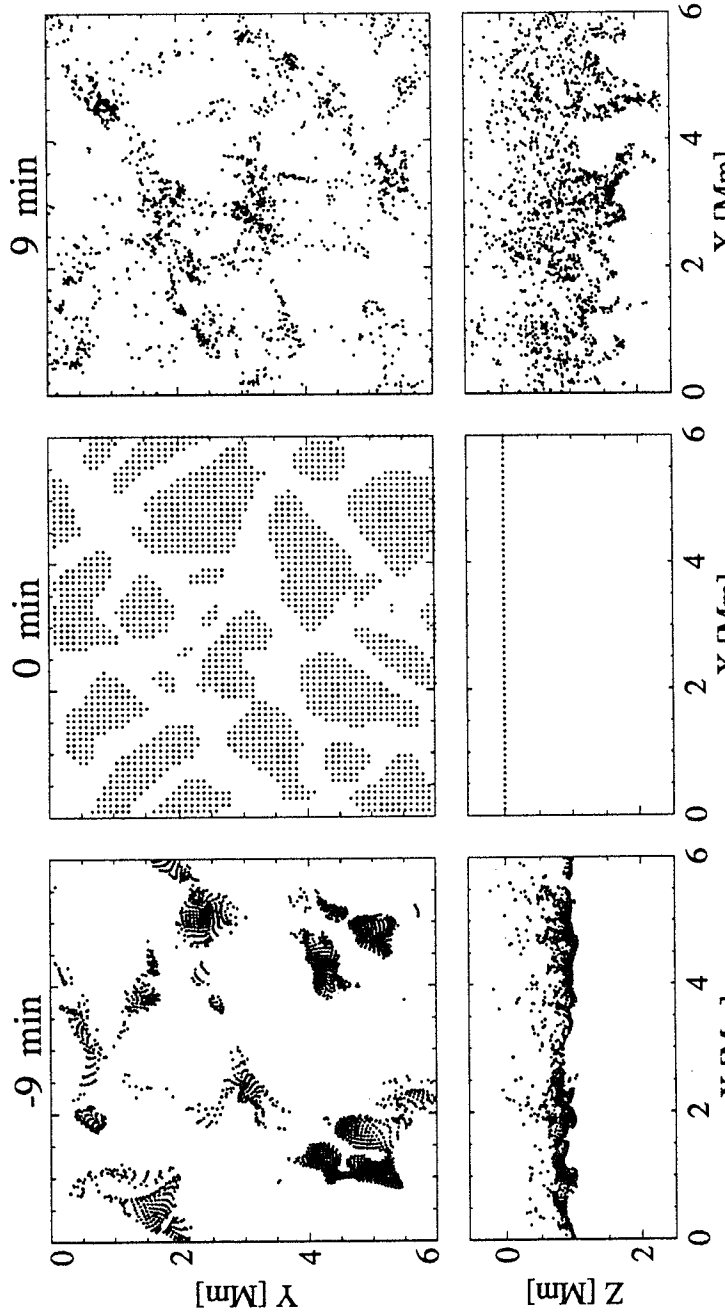

Fig. 12. Location of fluid parcels ascending through the solar surface at time $t = 0$ are shown 9 solar min earlier and later. Note the slow, diverging, nearly structureless nature of the ascending flow, and the fast descending flow, that converges into downdrafts that merge into a few widely separated filaments at large depth.

narrow downdrafts. Nine solar min later, most of the originally ascending fluid has descended a substantial distance and outlines these filamentary downdrafts.

The amplitude of entropy fluctuations decreases rapidly with depth in the convection zone. The mapping of very small source volumes of these nearly isentropic deeper layers onto all of the ascending flow close to the surface explains why the ascending flow is so nearly isentropic, and also justifies the use of isentropic inflow at the lower boundary of the computational domain.

We now have the necessary conceptual building blocks to understand the topology of convection qualitatively. Ascending flow is gentle and featureless because it is expanding; small-scale features are constantly being washed out by the expansion of the ascending flow. Overturning flows from neighboring updrafts collide along common borders, and are deflected towards corners between several updrafts. Fluid that reaches the surface, and *only fluid that reaches the surface,* looses entropy before overturning. Thus, at the surface, the characteristic cell topology is made clearly visible, with bright disconnected islands of ascending fluid, and connected dark lanes of descending fluid. Below the surface, the entropy-deficient fluid from the surface rapidly contracts, and is engulfed in overturning isentropic fluid that never reached the surface. Thus the entropy contrast between ascending and descending fluid decreases rapidly with depth, and the connectivity of the intergranular lanes is lost because entropy-deficient fluid in the lanes is replaced by overturning entropy neutral fluid.

Continuity allows fluid overturning below the surface to do so on increasingly larger scales at larger depths. At a depth of 2.5 Mm, the density scale height is significantly larger than at the surface (about 1 Mm), because of the ($\simeq$ 3 times) larger temperature, and also because of the ($\simeq$ 2 times) smaller mean molecular weight. Hence, Eq. (27) indicates that cell diameters of the order of 8 to 12 Mm are permitted, if the ratio of horizontal to vertical velocities remain of order 2 to 3. Our box was not large enough to allow this, but we do see the scale changing from many downdrafts at vertices near the surface to one or a few near the bottom. The scale change is caused by the advection of the small-scale downdrafts by the horizontal velocity field associated with the expansion of ascending fluid and is, of course, limited by the horizontal extent of our periodic box. In the Sun, no such restrictions occur and we expect that the scale of the horizontal flows continues to increase with depth.

Experiments where the fluid ascending through the lower boundary was not given a uniform entropy did not yield qualitatively different results. We have recently started a much larger and deeper ($12 \times 12 \times 9$ Mm) simulation. Similar topology occurs in this deeper simulation.

The vertical extent of the convective flows is greater than the depth of our computational box. The granulation pattern does *not* correspond to a

closed circulation pattern on the scale of the granulation. In fact, there is no evidence for multiple cells in the vertical direction. The flows do not close within our domain, and the merging downdrafts of cool material may possibly extend the entire depth of the convection zone (Fig. 13). However, the solar plasma is much less viscous than our simulated plasma, and so has a huge Reynolds number. Hence, although our upflows are smooth, and our downflows are only moderately turbulent, in the Sun they are likely to be strongly turbulent with significant generation of small-scale vorticity and eddies.

3. Convective Driving. The flow topology has important consequences also for the *driving* of flows on different scales. Driving is provided by the buoyancy fluctuations associated with entropy fluctuations. But the ever-expanding, gentle upflow is almost isentropic, so the main source of entropy fluctuations is the entropy loss at the surface. The surface acts as a source of cool, relatively dense material, which descends into a nearly isentropic interior. From the discussion of the topology of granulation at the surface, we know the initial topology of the entropy-deficient material: a connected network of intergranular lanes, with accumulation of particularly cool and dense fluid at the corners between several granules. The situation is similar to the Rayleigh-Taylor instability of a dense fluid on top of a less dense fluid, and a similar evolution ensues: filaments of dense material are formed, with denser filaments descending faster; additional dense fluid is pulled down in the wakes of the descending filaments, and the denser filaments gain an additional advantage.

The accumulation of cool, dense fluid at the vertices is self-amplifying; the flow is driven by the pressure deficiency caused by the smaller pressure scale height of cooler material at the vertices, and causes further draining of cool material from the intergranular lanes into the vertices. The overturning material that replaces the intergranular material below the surface is not entropy deficient, and hence the entropy deficiency at the intergranular lanes vanishes rapidly below the surface. The topology thus changes from one with cool material in connected lanes to one with cool material in narrow, nearly vertical and descending filaments.

We showed earlier that larger-scale flows are *allowed* at larger depths by the continuity equation. The descending cool filamentary material also provides *driving* of these larger-scale flows. The overturning fluid of larger-scale flows advect cool descending filaments of smaller-scale flows towards the boundaries and vertices of these larger-scale flows. Thus, the small-scale downdrafts close to the surface merge into large-scale, more widely separated downdrafts further from the surface, and provide the supply of entropy-deficient material that is necessary to drive the large-scale flows. By this process, the horizontal distribution of entropy deficient material changes gradually with depth, from one with many closely spaced filaments near the surface to one with fewer, less closely spaced filaments at depth.

Fig. 13. Fluid velocity in the x-z plane. The ascending flow is broad and diverging; the descending flow is filamentary and converging. The cells span the entire vertical extent of the computational domain. There are no multiple cells in the vertical direction.

This can be clearly seen in the vertical kinetic energy flux, $\rho \mathbf{u}^2 u_z$ (Fig. 14). Only the downflow is visible, because the upflow is slow and structureless. The merging of the granular downflows into the isolated, filamentary, mesogranulation downflow is clearly revealed. Notice also, that in our simulation, with its transmitting lower boundary, there is no sign of the downflow turning around into upflow within our simulation domain.

At first glance, the situation may seem somewhat absurd, with the low-density surface layers apparently providing the driving for large-scale flows at great depth, where the density is orders of magnitude larger ("tail wagging the dog"). However, consider the distribution and evolution of kinetic energy density in the flow. Most of the kinetic energy is concentrated in the strong downflows (cf. Fig. 14), and its evolution is controlled by a balance of buoy-

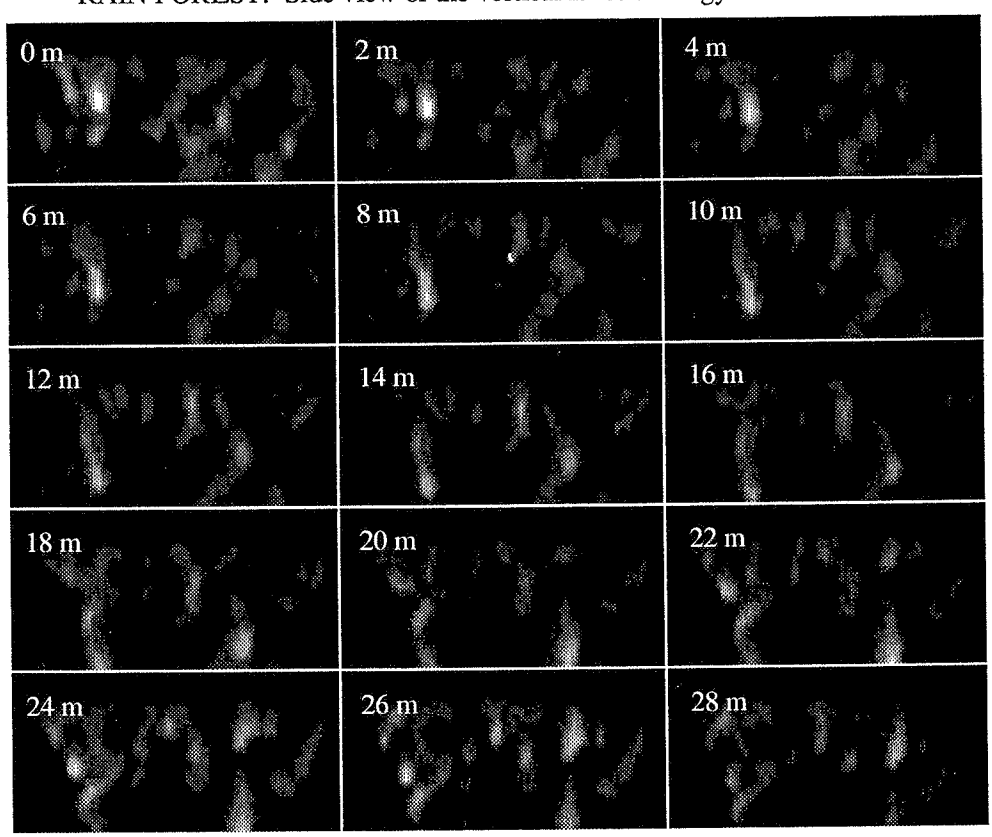

Fig. 14. Vertical kinetic energy flux, $\rho \mathbf{u} u_z$. Transparent side views through the computational domain, along the x-axis, are shown at 2-min intervals. Only the downward flux is visible, because the upflow is much slower, broader and featureless. The merging of downflows into isolated filaments, and the increasing horizontal separation of fewer filaments at greater depths is clearly revealed. The downdrafts penetrate through the transmitting lower boundary with little, if any, return flow.

ancy work, kinetic energy flux divergence and dissipation. Kinetic energy flux is even more concentrated into the downflows than the kinetic energy itself. The net kinetic energy flux is directed downwards, and almost entirely represents advection of kinetic energy density by the downflows. Buoyancy work (which is proportional to the product of density fluctuation and velocity) is positive in both the up and down flows, but is mainly performed on the downflows, where the density fluctuations and velocity are much larger. Hence, the energy input to the convective flow occurs primarily in the cool material from the surface layers.

This situation, where the "flow of information" is directed downwards (even though the total energy flux is directed outwards) is fortunate for the numerical simulation of solar surface layers, because it diminishes the consequences of having to introduce a computational lower boundary in the midst of the convection zone. Conversely, it necessitates a special treatment at a computational *upper* boundary in simulations of global solar convection (where these same surface layers have much too short characteristic time scales to be included).

4. Photospheric Phenomena. Our simulation results can be analyzed for other phenomena associated with the upper solar convection zone. We briefly indicate two of these.

Shocks developing from vertically propagating acoustic waves are thought to contribute to heating the chromosphere. Our simulation did not extend above the temperature minimum, so we can not investigate chromospheric heating. However, we do see shocks developing in the photosphere. These shocks are not due to vertically propagating acoustic waves, but rather to quasi-steady horizontal flows associated with the granulation pattern. Some of the ascending gas penetrates into the stable photosphere before being deflected sideways. The horizontal flow produces nearly stationary vertical shocks around bright granules which are being squeezed out of existence by their neighbors.

Helicity in the convection zone is thought to be essential in the dynamo process that produces the solar magnetic field (Parker 1955*b;* Gilman 1983). Although our simulation does not include rotation, we find that significant helicity occurs in the intergranule lanes and especially in the downdrafts (cf Fig. 11). When a magnetic field is present, this helicity will produce copious Alfvén waves, that will propagate along the field, and may contribute to chromospheric and coronal heating (Osterbrock 1961; Wentzel 1976; Ionson 1978; Holweg 1984).

D. Comparison with Observations

1. Granulation. Lites et al. (1989) used a sequence of simultaneous spectra in the Fe I 6302.5 Å line and narrowband slit-jaw images from the Swedish Vacuum Solar Telescope on La Palma, Canary Islands, Spain to

determine granular intensity and velocity fluctuations, and the granule size distribution. They compared these with the results of our simulation. The original simulations had a granule size distribution with significantly fewer small granules than observed. Because the amount of artificial viscosity needed for stabilization had initially been overestimated, a reduction of the magnitude of the terms that are proportional to the flow and sound speeds by a factor of 2.5 was possible, while still maintaining stability. This increased the effective spatial resolution of the code sufficiently to bring the sizes of the large simulated and observed granules into agreement.

The radiative energy transfer calculation, which is performed at each time step, yields, as a side effect, images in the total radiation intensity at the surface of the model. These images may be used as a convenient approximation of monochromatic continuum images. This has been verified by direct comparisons of images in total radiation intensity with monochromatic continuum images calculated with a separate program. Figure 10 shows a sequence of 16 such images. The time separation between images is 2 min. The size of each image is 6 × 6 Mm, corresponding to approximately 8″ × 8″. Because most features in an image change only moderately in a few minutes, there is some continuity from frame to frame in the sequence, and a single strip is in some sense analogous to an 8″ × 32″ image. More direct comparisons may be made, by smearing the synthetic images with point spread functions representing the telescope and atmospheric transmission functions (cf. Lites et al. 1989).

Figure 15 shows an 8″ × 24″ strip from an observed continuum image on the same scale as Fig. 10 (Lites et al. 1989). It is evident that the simulations do indeed produce surface patterns that qualitatively resemble the observed solar granulation. However, even with a reduced viscosity (the minimum allowed that maintains stability of the simulations), there are more

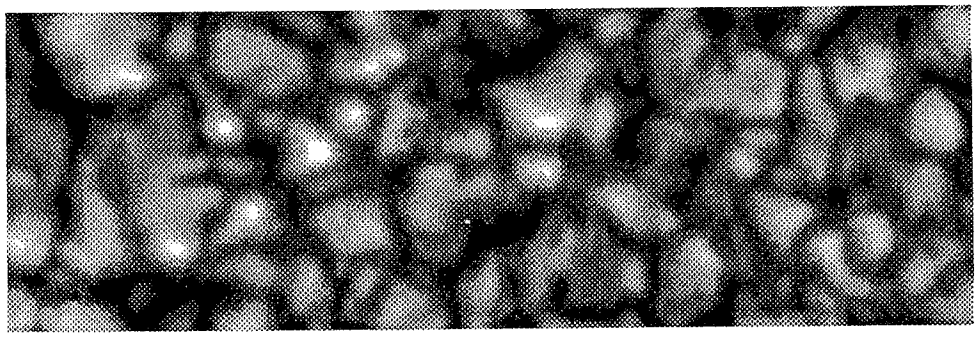

Fig. 15. Monochromatic continuum image of solar granulation, obtained at the Swedish Vacuum Solar Telescope on La Palma (Lites et al. 1989). The image is of a region 6 × 18 Mm (8″ × 24″). Compare this image with Fig. 10, which has the same scale. Note the similarity in large-scale features between the observation and simulation, but the presence of much more small-scale structure in the observation.

small-scale structures in the observations than in the simulations. Thus, the observations currently have a higher resolution than the simulations. A further increase of the spatial resolution of the simulations can only be obtained by decreasing the horizontal mesh spacing. Such simulations are under way.

Lites et al. also compared the amplitudes of the intensity and velocity fluctuations in the observed slit-jaw spectra to synthetic slit-jaw spectra. The magnitude of the observed intensity and velocity fluctuations depend on the (unknown) modulation transfer function (MTF) of the atmosphere (seeing) and telescope, so a direct comparison could not be made. However, by applying trial MTF's to the synthetic images, it was found that the *ratio* of intensity to velocity fluctuations is insensitive to the amount of smearing by the MTF, and that the ratio obtained from the simulations agrees with the observed ratio. On the other hand, the width of the (spatially unresolved) Fe I line provides a seeing independent check on the magnitude of the velocity fluctuations. From the average line width of the slit-jaw spectrum, Lites et al. conclude that the *rms* amplitude of the vertical velocities in the Fe I formation layers are consistent with the observations, and that consequently the intensity fluctuations are likely to be consistent too.

At the visible surface, the granule centers are warm and the intergranule lanes are cool. Our simulation shows that in the photosphere, due to adiabatic expansion, the diverging upflow above the bright granules becomes cooler than the converging downflow above the intergranule lanes (cf. Fig. 11). This reversal of bright and dark regions is observed (Evans and Catalano 1972; Canfield and Mehltretter 1973; Altrock and Musman 1976; Keil and Canfield 1978; Balthasar et al. 1990).

2. *Mesogranulation.* Granulation is the only scale for which the cell structure is directly visible at the surface, with cell boundaries delineated by cool, dark material. However, the larger scales are *indirectly* visible on the surface. The pressure fields induced by the larger-scale flows at depth extend to the surface. From the solutions of the anelastic pressure equation (Eq. 23) for large horizontal scales, Nordlund (1982, Eq. 46) has shown that the *relative* pressure fluctuations for horizontal wavenumber k in a medium with pressure scale height H have an $\exp(-k^2Hh)$ dependence on height h; i.e., cells with a horizontal size large relative to the scale height induce their relative pressure fluctuations over a height range comparable to or larger than their horizontal size. These relative pressure fluctuations drive the horizontal flows associated with the cell. Hence, the horizontal velocity fields associated with the larger-scale subsurface cells extend to the surface. The presence of large-scale subsurface cells thus becomes visible at the surface through the horizontal advection of smaller-scale cells.

The large (mesogranule) scale flow manifests itself at the visible solar surface by its effect on the growth as well as the motion of granules. Exami-

nation of a sequence of surface intensity images (Fig. 10) reveals areas where granule growth is enhanced and granules are large, and other areas where granule growth is suppressed and granules are small.

This pattern of granule motion from regions of enhanced growth towards regions of suppressed growth is just what is observed as mesogranular flows by November and Simon (1988) and Title et al. (1989). We find that the areas of horizontal convergence and small granules correspond to the persistent cool downdrafts and the areas of horizontally diverging motion and large granules correspond to the regions of warm upflows.

E. Conclusions

Only the small-scale cells at the surface (granulation) are made clearly visible by a large temperature contrast, with bright disconnected islands of ascending fluid, and connected dark lanes of descending fluid. The larger-scale flows are visible at the solar surface only by the advection of smaller-scale flows by their horizontal velocity fields. Our simulations reproduce the observed large end of granule scale structures, but are missing the smaller-scale structures. The simulated vertical velocities near the surface agree with those inferred from observations, and the ratio of intensity to velocity fluctuations in synthetic slit-jaw spectra agree with observed ones. The advection of granules and variation of granule sizes over larger regions, referred to as mesogranulation, are also reproduced.

The usual picture of turbulent convection has been of a hierarchy of eddies, or in the mixing length picture of bubbles that move some distance and then mix with their surroundings. Our simulations suggest a very different picture. The dominant topology of the outer solar convection zone appears to be one of broad, gentle, structureless, warm, adiabatic, diverging, upflows, with intermixed narrow, filamentary, cool, fast, twisting, converging downdrafts. The horizontal velocity field has a hierarchical structure, with small-scale cells at the surface, and successively larger-scale flows at larger depths, driven by the merging of the filamentary downdrafts of the smaller-scale cells closer to the surface. Our computational box, which is only 2.5 Mm deep, supports meso-scale flows with a horizontal extent comparable to the horizontal size of our box (6 Mm). Stronger turbulence in the downdrafts will increase the rate of mixing between ascending and descending fluid, but will not change the overall contraction of descending, entropy-deficient fluid. Thus, the filamentary downdrafts may possibly persist through the entire depth of the convection zone, merging into fewer, more widely separated plumes as they descend. Presumably, successively larger (supergranular) horizontal flows at the surface, with sizes ranging up to at least 50 Mm (Simon and Leighton 1964) reflect the successively increasing separation of the merging, descending plumes.

Although the descending vertical flows extend over many scale heights, estimates based on the mixing length concept may still be relevant, as the

overturning of most of the ascending flow within a scale height, and the associated dilution of descending fluid, in effect resembles a mixing.

Acknowledgments. KLC thanks the National Science Foundation for support. RFS and AN wish to express their appreciation for the support of the National Science Foundation, the Carlsberg Foundation and the Danish Natural Science Research Council. MS thanks the Deutsche Forschungsgemeinschaft for support. The authors wish to express their appreciation for this support.