PART ONE

FUNDAMENTALS OF LARGE EDDY SIMULATION

1

Some Historical Remarks on the Use of Nonlinear Viscosities

JOSEPH SMAGORINSKY

1.1 Introductory Remarks

This chapter reviews the early experiences of the application of computational methods to the hydrodynamics of large-scale atmospheric motions: the beginnings of numerical weather prediction (see, for example, Smagorinsky 1983).

In the early 1950s, the state of the art was a quasi-geostrophic, quasi-two-dimensional model of large-scale atmospheric dynamics. Of course, the models were time-dependent and nonlinear. We had just emerged from the era when we dealt exclusively with barotropic models that simulated atmospheric evolutions on time scales of one or, at the most, two days. In these models total kinetic energy was assumed to be conserved. Models were also being constructed for the next hierarchical step, that is, baroclinic models which allowed potential–kinetic energy conversions, thus including the processes related to storm development (baroclinic instability) and thereby extending the validity of the models by another day or two. The small, but nontrivial, large-scale vertical component of motion was implied, even quasi-geostrophically.

The crucial achievement in the mid-1950s was the successful construction by Norman Phillips (1956) of an energetically self-sustaining model which emulated the prime external radiative energy source and the viscous energy sink. This model was capable of being integrated for extended intervals and exhibited many of the atmosphere's nonlinear characteristics such as the fundamental energy or index cycle. Phillips had used a conventional linear viscosity and conductivity with the coefficient

$0.2L^{4/3}$ (in CGS units), determined in accordance with Richardson's (1926) empirical law, L being identified with the grid size. What was being observed in these longer-term integrations was the development of extended vortex lines, and at the Institute for Advanced Study the phenomenon was aptly dubbed "noodling."

John von Neumann, who oversaw Charney's group at the Institute, recalled some of his and Richtmyer's earlier experience (1950) with hydrodynamic shocks. In particular, they were dealing with one-dimensional flows, and were seeking a means "to introduce (artificial) dissipative terms into the equations so as to give the shocks a thickness comparable to (but preferably somewhat larger than) the spacing of the points of the network. Then the differential equations (more accurately, the corresponding difference equations) may be used for the entire calculation, just as though there were no shocks at all." Von Neumann and Richtmyer were guided by earlier physical insights on the roles of dissipative mechanisms, that is, viscosity and heat conduction, in the behavior of shocks. The form of the dissipation was derived heuristically and was "introduced for purely mathematical reasons." The viscosity was taken to be proportional to the magnitude of the divergence, which, in a one-dimensional flow, is indistinguishable from the magnitude of the deformation.

A second reference has to do with a study by Phillips (1959) on the occurrence of nonlinear computational instability in the numerical integration of a barotropic, nondivergent atmospheric model. It was motivated by his experience, a few years earlier (Phillips 1956), with the unprecedented long integrations of a quasi-geostrophic, baroclinic, two-level "general circulation model," the first of this genre. After several weeks of simulated time, the appearance of large truncation errors caused an almost explosive increase of the total energy of the system. This instability at the smallest resolvable scales could not be suppressed by reducing the time step, as is customary in the case of linear Courant–Friedrich–Lewy (CFL) instability. However, by periodically eliminating all components with wave lengths smaller than 4 times the grid size, that is, an artificial "smoothing," he was able to suppress the nonlinear instability.

About 1960, J. Charney and N. Phillips conveyed by personal communication that they had successfully used a two-dimensional version of the von Neumann–Richtmyer nonlinear viscosity, proportional to the deformation, to control grid-scale filamentation in trial numerical integrations. Presumably these were for quasi-geostrophic flows.

In a recently published interview with Charney, conducted in August 1980 (Lindzen, Lorenz and Platzman 1990, p. 61), Charney said that after using the nonlinear viscosity in 1955 or 1956 he tried a linear viscosity and "found the results were just about as good and, therefore, gave up the idea of using the non-linear viscosity, ... other techniques are probably preferable filtering techniques. Because the one problem with these viscosities is that in the long run they can give you unrealistic transports of momentum."

The present account is based on an attempt in the early 1960s to rationalize the derivation of a nonlinear viscosity based upon the principles of modern turbulence theory, particularly those related to Heisenberg–Kolmogorov similarity in the inertial subrange of three-dimensional, isotropic turbulence. What I tried to do at the time was to particularize the classical results from the theory of turbulence to three-dimensional horizontally isotropic turbulence, in which the vertical component is quasi-hydrostatically constrained. The formulation, however, in general is applicable to the primitive equations of motion, that is, where gravity waves are admissible modes.

The notion of an inertial subrange of fluid motions implies a regime free from sources and sinks of kinetic energy and separated in wave number space from the processes responsible for molecular viscous decay. That is, an equilibrium should exist such that the total kinetic energy of the subrange is conserved, and that the flow of kinetic energy cascades from low to higher wave numbers purely by the nonlinear interactions of the inertial forces.

One purpose for seeking such a subrange in the atmosphere was that, in the numerical integration of the equations of motion, an artificial threshold was created by the grid size. For longer waves, the dynamics of the motions were dealt with explicitly, whereas for the shorter waves, the motions had to be dealt with statistically, that is, parametrically, or ignored entirely. Generally, unless the statistical dynamics of the turbulence are understood well enough, the explicit dynamics cannot adequately communicate with the implied viscous subrange. Fortunately, the results of similarity arguments (Taylor 1935; von Kármán and Howarth 1938; Kolmogorov 1941a,b; Onsager 1945, 1949; Batchelor 1946; von Weizsäcker 1948; Heisenberg 1948a,b; Bass 1949) provided some insight into the statistical properties of isotropic turbulent transfer in the inertial subrange which, as will be shown, appears to be formally applicable to atmospheric motions.

The present account includes a somewhat shortened version of an un-

published manuscript (Smagorinsky 1962). The latter was written at a time when a nonlinear viscosity was being used in an atmospheric primitive equation general circulation model (Smagorinsky 1963). However, the model was constrained to filter out external gravity wave modes. The essential differences between the derivations in this chapter and the early draft are in the justification of the quasi-hydrostatic impact on the viscosity coefficients; in the choice of, and the parametric relationships between, the numerical coefficients occurring in the turbulence formulation; in correcting the numerical coefficient from 6 to $\frac{2}{3}$ in (39) below; and in a rationale for the conditions governing the choice of the single nondimensional "constant" in the turbulence exchange coefficient.

Much has been learned since, both theoretically and experimentally. The present perspective reflects on some of the many developments of the past 28 years. Numerous applications of a nonlinear viscosity have enlarged the base of experience in meteorology, oceanography and a variety of other fluid dynamical problem areas.

One may infer the location of a relative inertial subrange for atmospheric motions, at least for those motions sufficiently removed from the lower boundary and the equator (Figure 1) (Smagorinsky 1974). It was already well-established in the early 1950s (Onsager 1949; Fjørtoft 1953) that because of the vorticity conservation by quasi-horizontal motions, the long barotropic waves (zonal or east–west wave numbers 1 to 3) have associated with them a net transfer of kinetic energy to lower wave numbers. This property of geostrophic turbulence is responsible for the -3 power falloff of spectral energy density for the largest atmospheric scales (Charney 1971). It is the range where enstrophy (half the squared vorticity) is cascaded to smaller scales, but where, in the net, kinetic energy decascades to longer scales, ultimately maintaining the jet stream against dissipation. Furthermore, the seasonal monsoon resulting from the planetary-scale continents provides a quasi-stationary thermal forcing with a zonal wave number of about 2. Also, the baroclinic instability process (Charney 1947; Eady 1949; Fjørtoft 1951) provides a major source of energy in zonal wave numbers 4 to 7 (see, for example, Holton 1979, pp. 5, 36). Large-scale condensation processes provide an additional source at somewhat higher wave numbers. The atmospheric mesoscale, nominally from about 5 to 500 km in horizontal dimensions, harbors many intense, but generally intermittent and sparse, phenomena in the extratropics, such as mesoscale convective complexes and fronts which may not appear in a spectrum taken at any one time. Proceed-

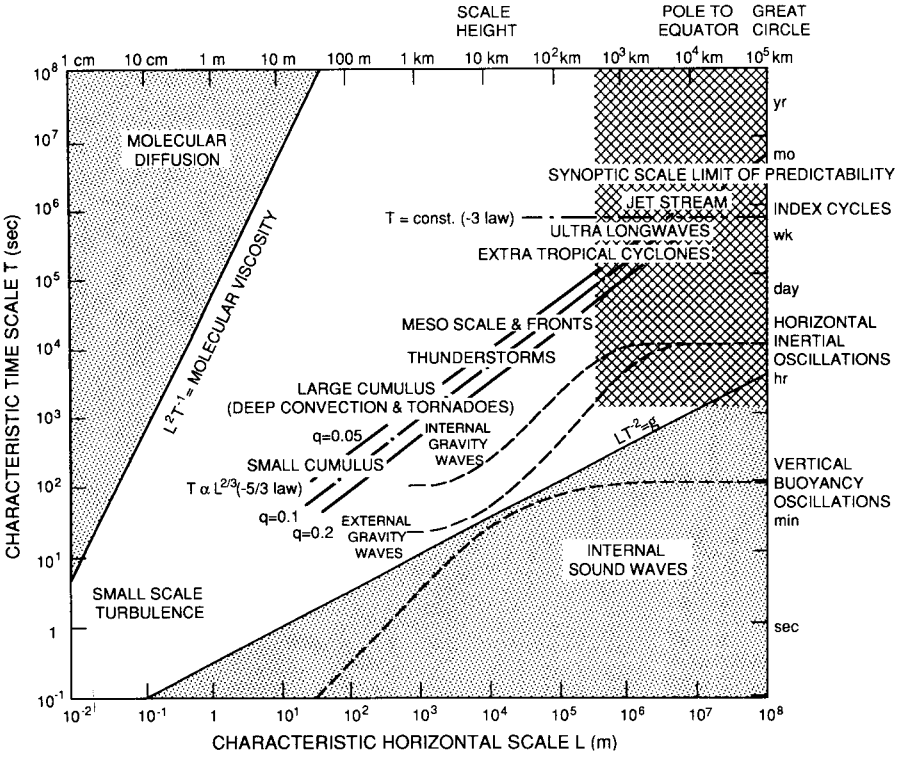

Fig. 1 A space–time domain for characteristic atmospheric phenomena. The unstippled region encompasses most of the kinetic-energy-containing phenomena, with the predominance of extratropical cyclones, ultralong waves and the jet stream. The cross-hatched area denotes scales and phenomena typically resolved by general circulation models, that is, the macroscale. The central heavy diagonal lines are $q = L^{2/3}T^{-1} =$ constant in units of $m^{2/3}\ s^{-1}$. (After Smagorinsky 1974, based on a figure by H. Fortag as modified by K. Ooyama.)

ing somewhat further downscale, we encounter convective phenomena of various types. However, it has been pointed out by Emanuel (1984) that "only slant-wise convection is capable of directly generating kinetic energy in the mesoscale and occurs only intermittently in strongly baroclinic situations." The mesoscale range has proven to be particularly

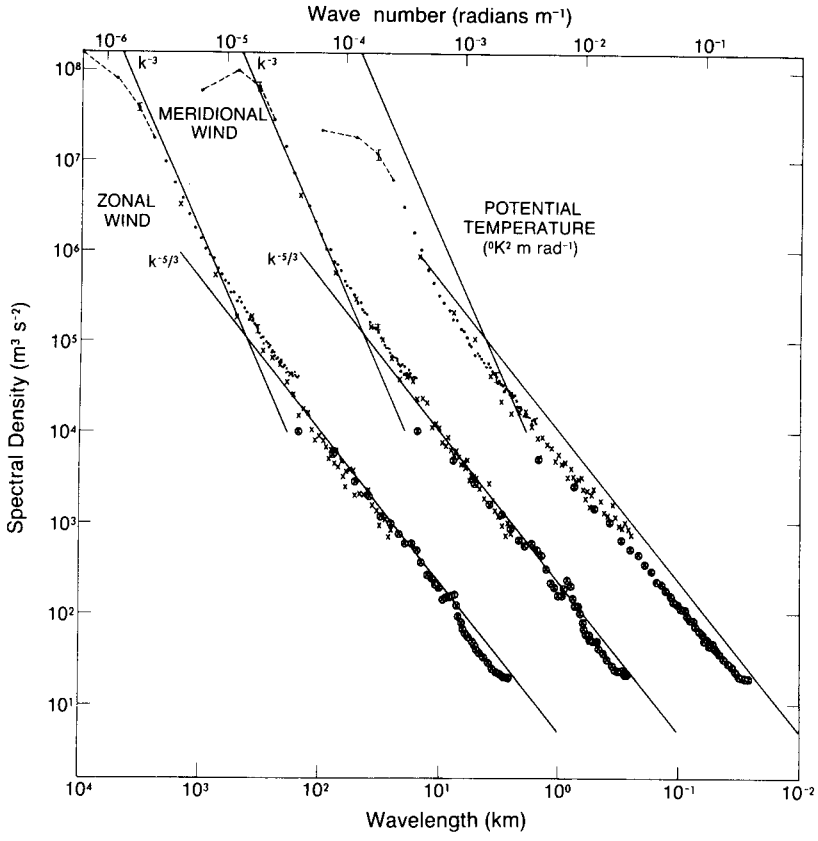

Fig. 2 Variance power spectra of wind and potential temperature near the tropopause from GASP aircraft data. The spectra for meridional wind and temperature are shifted one and two decades to the right, respectively; lines with slopes –3 and $-\frac{5}{3}$ are entered at the same relative coordinates for each variable for comparison. (After Nastrom and Gage, 1985.)

useful for setting the grid size in studying large-scale motions by numerical simulations.

Recent definitive measurements of atmospheric energy spectra in the span of horizontal scales 2 to 10,000 km (Figure 2) (Nastrom and Gage 1985) reaffirm that while a –3 slope fits the spectrum for scales larger than 300–400 km in the "geostrophic turbulence" range, $-\frac{5}{3}$ fits best for the mesoscales, that is, for wavelengths between 2.6 km and 300–400 km.

1.2 The Mean Equations of Motion

If we average the Navier–Stokes equations over a wave number interval $k \leq k' \leq \infty$, we find

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\frac{1}{\rho}\nabla p^* + \frac{1}{\rho}\nabla \cdot \boldsymbol{\Gamma}^* - \mathbf{g}. \tag{1}$$

Here, $D()/Dt$ is the individual time rate of change on a particle, ∇ is the three-dimensional gradient vector operator, $\mathbf{v}$ is the average three-dimensional relative velocity vector, p^* is the average molecular pressure, $\boldsymbol{\Gamma}^*$ is the Reynolds stress dyadic, $\mathbf{g}$ is the gravitational acceleration vector, ρ is the density (its Reynolds fluctuations having been neglected) and $\boldsymbol{\Omega}$ is the earth's angular velocity vector. We have assumed that the wave number, k, is sufficiently small that the Reynolds stresses are large compared with the molecular viscous stresses, which have therefore been omitted.

$\boldsymbol{\Gamma}^*$ may be written as the sum of its average value in all directions and a deviation, $\boldsymbol{\Gamma}$,

$$\boldsymbol{\Gamma}^* = \frac{|\boldsymbol{\Gamma}^*|}{3}\boldsymbol{\Upsilon} + \boldsymbol{\Gamma} \qquad \text{and} \qquad |\boldsymbol{\Gamma}| = 0, \tag{2}$$

where $\boldsymbol{\Upsilon}$ is the idemfactor and $|\boldsymbol{\Upsilon}| = 3$.

By analogy with the kinetic theory of gases, we interpret $|\boldsymbol{\Gamma}^*|/3$ as the pressure due to the eddy kinetic energy. Hence, the macropressure is (Hinze 1959, p. 15)

$$p = p^* - |\boldsymbol{\Gamma}^*|/3, \tag{3}$$

and (1) becomes

$$\frac{D\mathbf{v}}{Dt} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\frac{1}{\rho}\nabla p + \frac{1}{\rho}\nabla \cdot \boldsymbol{\Gamma} - \mathbf{g}, \tag{4}$$

in which $(1/\rho)\nabla \cdot \boldsymbol{\Gamma} \equiv \mathbf{F}$ is the frictional force vector per unit mass.

1.3 Application of Elasticity Theory

The general theory of viscous fluids conventionally suggests that the strains of individual fluid elements give rise to stresses analogous to those in elastic media. However, in fluids, the pressures are proportional to the rate of strain, since a fluid yields to a shearing stress. One can, therefore, adapt the formalisms developed in the theory of elasticity provided that the fluid velocity, $\mathbf{v}$, is used instead of the elastic displacements.

Therefore, we assume, for small deformations, that $\boldsymbol{\Gamma}/\rho$ is proportional to the rate of strain dyadic, $\mathbf{S} = \nabla\mathbf{v}$, through a tensor, $\mathbf{C}$, such that

$$\boldsymbol{\Gamma} = \rho|\mathbf{C} \cdot \mathbf{S}|, \tag{5}$$

in analogy to the elastic function in elasticity. In general, $\mathbf{C}$ is a tetradic, which has 3^4 $(= 81)$ components.

The kinetic energy equation may be formed by the scalar multiplication of $\mathbf{v}$ and (4) and then integrating over the entire volume. Assuming closed boundaries, one finds

$$\frac{\partial}{\partial t}\iiint \frac{\rho \mathbf{v}^2}{2}\, dV = \iiint p\nabla\cdot\mathbf{v}\, dV - \iiint \rho\mathbf{v}\cdot\mathbf{g}\, dV\,,$$

$$- \iint_{z=0} (\mathbf{v}\cdot\mathbf{\Gamma})\cdot d\mathbf{A} - \iiint \rho\epsilon\, dV, \tag{6}$$

where

$$\epsilon \equiv |\mathbf{S}\cdot\mathbf{\Gamma}|/\rho \tag{7}$$

is the local energy dissipation or decay, $\iiint \rho\epsilon\, dV$ is the energy transformation function between the explicit and eddy flow, dV is a volume element, $d\mathbf{A}$ is a horizontal area element, $\iiint (p\nabla\cdot\mathbf{v} - \rho\mathbf{v}\cdot\mathbf{g})\, dV$ is the potential–kinetic energy conversion, and $\iint_{z=0}(\mathbf{v}\cdot\mathbf{\Gamma})\cdot d\mathbf{A}$ is the energy dissipation due to the eddy stresses tangential to the lower boundary at $z = 0$, z being the vertical coordinate, opposite to the direction of $\mathbf{g}$.

Hence, if we write the matrices of the tensor components symbolically as

$$\mathbf{S} \to S_{rs}, \quad \mathbf{\Gamma} \to \tau_{mn}, \quad \mathbf{C} \to C_{mnrs}, \tag{8}$$

then the equations for the stress components are

$$\tau_{mn}/\rho = \sum_{rs} C_{mnrs} S_{rs} \tag{9}$$

and

$$\epsilon = \sum_{mnrs} C_{mnrs} S_{mn} S_{rs}. \tag{10}$$

It follows from energy considerations (Sokolnikoff 1946, p. 63) that there exists a reflective symmetry along any arbitrary axis, so that

$$S_{rs} = S_{sr}, \quad \tau_{mn} = \tau_{nm}, \quad C_{mnrs} = C_{rsmn} = C_{nmrs} = C_{mnsr}, \tag{11}$$

in which case there are only 6 independent S_{rs} and τ_{mn}, and 21 independent C_{mnrs}.

1.4 Axial Symmetry

We will now confine our attention to a scale of motion sufficiently small that these motions can be regarded as horizontally isotropic, that is, symmetric with respect to the vertical axis. This definitely excludes the

barotropic long-wave regime and probably the energy-producing baroclinic instability regime.

The axial symmetry of τ_{mn} and S_{rs} contracts their matrices to 6×6, resulting in 9 nonzero C_{mnrs}, of which only 3 are independent when (2) is taken into account:

$$\sum_m \tau_{mm} = 0, \tag{12}$$

$$\begin{aligned} \alpha &\equiv -\tfrac{1}{2}C_{1133} = -\tfrac{1}{2}C_{2233} = \tfrac{1}{4}C_{3333} \ , \\ \beta &\equiv \ \tfrac{1}{2}(C_{1111} - C_{1122}) = \tfrac{1}{2}(C_{2222} - C_{1122}) = C_{1212} \ , \\ \gamma &\equiv \ C_{1313} = C_{2323} \ . \end{aligned} \tag{13}$$

The stress components may now be written

$$\begin{aligned} \tau_{11}/\rho &= \ \beta(S_{11} - S_{22}) + \alpha(\nabla \cdot \mathbf{v} - 3S_{33}) \ , \\ \tau_{22}/\rho &= -\beta(S_{11} - S_{22}) + \alpha(\nabla \cdot \mathbf{v} - 3S_{33}) \ , \\ \tau_{33}/\rho &= -2\alpha(\nabla \cdot \mathbf{v} - 3S_{33}) \ , \\ \tau_{23}/\rho &= \ \tau_{32}/\rho = 2\gamma S_{23} \ , \\ \tau_{13}/\rho &= \ \tau_{31}/\rho = 2\gamma S_{13} \ , \\ \tau_{12}/\rho &= \ \tau_{21}/\rho = 2\beta S_{12} \ , \end{aligned} \tag{14}$$

where

$$\nabla \cdot \mathbf{v} = S_{11} + S_{22} + S_{33} = -\frac{d}{dt} \ln \ \rho \tag{15}$$

by the equation of continuity.

Also, in the case of axial symmetry, one may calculate ϵ from (10) and (14):

$$\epsilon = \beta D^2 + \alpha M^2 + \gamma P^2, \tag{16}$$

where

$$\begin{aligned} D^2 &\equiv (S_{11} - S_{22})^2 + (2S_{12})^2 \ , \\ M^2 &\equiv (\nabla \cdot \mathbf{v} - 3S_{33})^2 \ , \\ P^2 &\equiv (2S_{13})^2 + (2S_{23})^2 \ . \end{aligned} \tag{17}$$

It is easy to see that for $\alpha, \beta, \gamma \geq 0$, ϵ is positive-definite, thus ensuring positive energy dissipation.

1.5 Spherical Curvilinear Coordinates

It will be useful to record the forms of the viscous terms in spherical curvilinear coordinates. We denote the longitude, λ (in the direction of the unit vector $\mathbf{i_1}$), the latitude, ϕ ($\mathbf{i_2}$), and the radial distance from the