

Maximum Simplification of the Dynamic Equations

By EDWARD N. LORENZ, Massachusetts Institute of Technology¹

(Manuscript received February 1, 1960)

Abstract

When the dynamic equations are to be used to further our understanding of atmospheric phenomena, it is permissible to simplify them beyond the point where they can yield acceptable weather predictions. Through the use of double Fourier series, and with the omission of all but the largest scales of motion, the barotropic vorticity equation may be reduced to a system of three ordinary nonlinear differential equations. The analytic solutions of these equations are elliptic functions of time. The equations may also be solved rapidly by numerical integration.

Particular solutions of the equations picture the motion of finite disturbances on a zonal flow, with exchanges of kinetic energy between the zonal flow and the disturbances accompanying the meridional transport of zonal momentum by the disturbances. Other solutions picture the initial growth and eventual cessation of growth of small disturbances on an unstable zonal current. Still further solutions picture the destruction of a stable zonal flow by large disturbances, and lead to a plausible hypothesis concerning index cycles in the atmosphere.

Less extreme simplifications of the dynamic equations may be used when more complicated atmospheric phenomena are to be studied.

1. Simplification of the dynamic equations and the initial conditions

The various phenomena which are observed in our atmosphere, and the changes in the state of the atmosphere from one time to another, are supposedly governed by a set of physical laws. The dynamic meteorologist does not usually regard the discovery of these laws as one of his tasks, being willing to concede that the laws have already been established, at least in approximate form, by workers in other fields. Instead, he includes among his problems the prediction of future states of the atmosphere by means of these laws, and the explanation of typical observable phenomena in terms of these laws. He ordinarily finds it convenient to express the laws as a set of mathematical equations.

¹ The research resulting in this work has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, under Contract No. AF 19 (604)-1000.

Tellus XII (1960), 3

In order to make the best attainable forecast of the future weather, it would be desirable to express the physical laws as exactly as possible, and determine the initial conditions as precisely as possible. Yet the ultimate achievement of producing perfect forecasts, by applying equations already known to be exact to initial conditions already known to be precise, if such a feat were possible, would not by itself increase our understanding of the atmosphere, no matter how important it might be from other considerations. For example, if we should observe a hurricane, we might ask ourselves, "Why did this hurricane form?" If we could determine the exact conditions at an earlier time, and if we should feed these conditions, together with a program for integrating the exact equations, into an electronic computer, we should in due time receive a forecast from the computer, which would show the presence of a hurricane. We then might still be justified in asking why

the hurricane formed. The answer that the physical laws required a hurricane to form from the given antecedent conditions might not satisfy us, since we were aware of that fact even before integrating the equations.

It is only when we use systematically imperfect equations or initial conditions that we can begin to gain further understanding of the phenomena which we observe. For if we omit the terms representing specified physical processes, such as friction, from the equations, or if we fail to include certain observable features, such as cloudiness, in the initial conditions, we may, by comparing the mathematical solutions with reality, gain some insight concerning the relative importance of the retained and omitted features. Of course, in so doing, we forgo the opportunity of simultaneously making the best attainable forecast.

Our present methods of weather observation, and also any foreseeable future methods, yield systematically incomplete initial conditions, and our present mathematical techniques do not allow us to solve the dynamic equations without previous systematic simplification, whether or not the equations may originally be expressed in exact form. Beyond these unavoidable inaccuracies, further simplifications have so far been necessary for the sake of economy. Thus it is that the recent studies in numerical weather prediction, besides yielding creditable although not optimum forecasts, have made vast contributions to our understanding of common weather phenomena. It should be added that these contributions are no accidental by-product; the investigators were concerned with explanation as well as prediction.

The customary simplifications of the dynamic equations are of two kinds. First, we omit or modify certain terms. By doing so, we neglect or alter certain physical processes believed to be of secondary importance. For example, the omission of terms representing the release of latent heat has become familiar, and the introduction of the filtering approximations, which prohibit the propagation of sound and gravity waves, has been one of the important recent advances in dynamic meteorology.

Further simplifications are then demanded by our present inability to solve the partial differential equations. Accordingly, by finite

difference methods or otherwise, we convert each partial differential equation into a specified finite number of ordinary differential equations. These equations generally become algebraic when finite differences in time are also introduced. In this process, phenomena of such a small scale that they are lost in the differencing process are automatically excluded. There is no theoretical limit to the number of algebraic equations which may be handled, but for reasons of economy and convenience the number must be restricted. Several hundred equations is typical in short-range dynamic forecasting.

The simplifications of the initial conditions are analogous to those of the dynamic equations. First, we may omit or modify the description of certain observable quantities. For example, it may be logical to omit the field of humidity altogether, when the condensation process is not included in the dynamic equations.

Further simplifications are then demanded by our present system of observations, which yields values of observable quantities only at specified locations. Thus the true state of the atmosphere with its nearly infinite number of degrees of freedom is replaced by a state with a specified finite number of degrees of freedom. Again, phenomena small enough to be lost between observing stations will not be described. For reasons of economy and convenience, the number of degrees of freedom retained is often far less than the maximum number allowed by the observations.

When the problem is finally ready for solutions, it is logical that the finite set of quantities describing the initial conditions should be the finite set of dependent variables in the modified equations.

Although in meteorological practice the most common method of handling partial differential equations has been the use of finite differences, other methods are possible. An alternative method is the analysis of the field of a dependent variable into a series of orthogonal functions. Usually these orthogonal functions are eigenfunctions of a conveniently chosen equation (not the dynamic equation itself). A familiar example of such an analysis, in one dimension, is the Fourier series, in which the orthogonal functions are sines and cosines. The coefficients of these orthogonal functions

become the new dependent variables. The number of dependent variables and ordinary differential equations is then made finite by omitting reference to all but a finite number of variables. When the eigenfunctions can be arranged in the order of their eigenvalues, to represent features of successively smaller scale, this method also excludes the small-scale phenomena.

If our interest is confined to furthering our understanding of the atmosphere, we may simplify the equations and initial conditions to the point where good predictions can no longer be expected. In this study, we shall illustrate some of the advantages to be gained by simplifying the equations to the greatest extent possible, while still retaining the desired physical features. After simplifying the dynamic equations and initial conditions by omitting or modifying certain physical processes and features, we shall expand the equations and initial conditions into series of eigenfunctions. We shall then retain as dependent variables the coefficients of a minimum number of eigenfunctions, corresponding to features of the largest scale. As we shall see, in some cases the resulting set of ordinary differential equations is simple enough to be solved analytically. If this is not the case, we may convert the equations to algebraic equations by a finite difference approximation in time.

It should be evident that these remarks need not have been confined to the atmosphere. Similar remarks apply to a wide class of phenomena, which, in particular, includes controlled hydrodynamic experiments as well as such natural phenomena as oceanic flow.

2. The minimum hydrodynamic equations

As an illustration, we shall begin by letting the dynamic equations governing the atmosphere be simplified until they reduce to the familiar vorticity equation

$$\frac{\partial}{\partial t} \nabla^2 \psi = -\mathbf{k} \cdot \nabla \psi \times \nabla (\nabla^2 \psi), \quad (1)$$

where t is time, ψ is a stream function for two-dimensional horizontal flow, ∇ is a horizontal differential operator, $\nabla^2 = \nabla \cdot \nabla$ is the horizontal Laplacian operator, and \mathbf{k} is a unit vertical vector. Equation (1) is equivalent to the barotropic vorticity equation, which

approximately governs the vertically-averaged horizontal flow in the atmosphere. Equation (1) also governs the motion of a general two-dimensional homogeneous incompressible nonviscous fluid, and states that the vorticity of each material parcel of fluid is conserved.

We shall apply equation (1) to flow in a plane region, in which ψ is doubly periodic at all times, i.e.,

$$\psi(x + 2\pi/k, y + 2\pi/l, t) = \psi(x, y, t), \quad (2)$$

where the x and y axes point eastward and northward, respectively, and k and l are specified constants. In this way we have distorted the geometry of the spherical earth, but we have retained the important property that the effective area is finite but unbounded. We have also neglected the horizontal variations of the Coriolis parameter, although equation (1) is still consistent with a constant Coriolis parameter.

In such a plane region, the eigenfunctions of the equation

$$\nabla^2 \psi = c\psi \quad (3)$$

are trigonometric functions, or equivalently, complex exponential functions, of the variables $m k x + n l y$, where m and n are integers, and the corresponding eigenvalues c are the quantities $-(m^2 k^2 + n^2 l^2)$. Hence a series of eigenfunctions is in this case simply a double Fourier series. An expansion of $\nabla^2 \psi$ in such a series is

$$\nabla^2 \psi = \sum_{m=0}^{\infty} \sum_{n=n_0}^{\infty} [A_{mn} \cos(m k x + n l y) + B_{mn} \sin(m k x + n l y)] \quad (4)$$

where A_{mn} and B_{mn} are real coefficients (and $A_{00} = 0$), and the lower limit n_0 in the second sum is $-\infty$ if $m > 0$ and 0 if $m = 0$, or in a more concise notation,

$$\nabla^2 \psi = \sum_{\mathbf{M}} C_{\mathbf{M}} e^{i\mathbf{M} \cdot \mathbf{R}} \quad (5)$$

where $\mathbf{M} = i m k + j n l$ and $\mathbf{R} = i x + j y$ are the vector wave number, or wave vector, and the radius vector, i and j are unit vectors parallel to the x and y axes, m and n assume all integral value from $-\infty$ to ∞ , and

$$C_{\mathbf{M}} = \overline{C_{-\mathbf{M}}} = \frac{1}{2} (A_{mn} - i B_{mn}) \quad (6)$$

Here the bar ($\bar{}$) denotes a complex conjugate. The corresponding Fourier series for ψ is

$$\psi = - \sum_{\mathbf{M}} (\mathbf{M} \cdot \mathbf{M})^{-1} C_{\mathbf{M}} e^{i\mathbf{M} \cdot \mathbf{R}} \quad (7)$$

From (5), with the dummy index \mathbf{M} replaced by \mathbf{L} , and (7), with \mathbf{M} replaced by \mathbf{H} , it follows that

$$\begin{aligned} \mathbf{k} \cdot \nabla \psi \times \nabla (\nabla^2 \psi) &= \\ = \sum_{\mathbf{H}, \mathbf{L}} (\mathbf{k} \cdot \mathbf{H} \times \mathbf{L}) (\mathbf{H} \cdot \mathbf{H})^{-1} C_{\mathbf{H}} C_{\mathbf{L}} e^{i(\mathbf{H} + \mathbf{L}) \cdot \mathbf{R}} & \quad (8) \end{aligned}$$

Replacing $\mathbf{H} + \mathbf{L}$ in (8) by \mathbf{M} , and substituting (5) and (8) into the left and right hand sides, respectively, of (1), we find that

$$\frac{dC_{\mathbf{M}}}{dt} = - \sum_{\mathbf{H}} \frac{\mathbf{K} \cdot \mathbf{H} \times \mathbf{M}}{\mathbf{H} \cdot \mathbf{H}} C_{\mathbf{H}} C_{\mathbf{M}-\mathbf{H}} \quad (9)$$

Equation (9) is the Fourier transform of equation (1), and is actually the infinite set of ordinary differential equations which we sought to determine. The coefficients $C_{\mathbf{M}}$ are the dependent variables.

For the surface of a sphere, the eigenfunctions of equation (3) are spherical harmonics. An expansion of the vorticity equation into a series of spherical harmonics has been used by SILBERMAN (1954).

The final step in simplification is the omission of reference to all but a finite number of variables $C_{\mathbf{M}}$, corresponding to a specified set of values of m and n . If these values are small, only large-scale features will be retained. The summation in (9) then becomes a finite sum, but the equations are otherwise unaltered.

Under equation (1), two integral quantities, namely, the mean kinetic energy and the mean square of the vorticity, remain unchanged with time. Since equations (1) and (9) are equivalent, equation (9) must also conserve these values. According to (5) and (7), the mean kinetic energy is given by

$$E = \frac{1}{2} \sum_{\mathbf{M}} \frac{C_{\mathbf{M}} C_{-\mathbf{M}}}{\mathbf{M} \cdot \mathbf{M}}, \quad (10)$$

while the mean square of the vorticity is given by

$$V = \sum_{\mathbf{M}} C_{\mathbf{M}} C_{-\mathbf{M}}. \quad (11)$$

According to (9), the time derivatives of E and V each involve sums of terms containing products $C_{\mathbf{H}} C_{\mathbf{M}-\mathbf{H}} C_{-\mathbf{M}}$. In these sums each product occurs six times, corresponding to the six permutations of $C_{\mathbf{H}}$, $C_{\mathbf{M}-\mathbf{H}}$, and $C_{-\mathbf{M}}$, and the sum of the six coefficients may be readily verified to equal zero.

It follows that if all reference to specified variables is omitted, any specific product $C_{\mathbf{H}} C_{\mathbf{M}-\mathbf{H}} C_{-\mathbf{M}}$ will either be omitted altogether from the expressions for dE/dt and dV/dt or else it will occur six times, and the sum of the six coefficients will still vanish. Hence the simplified equations (9) still conserve the mean kinetic energy and the mean square of the vorticity, although the definition of these quantities has now been modified, in that the sums in (10) and (11) are now finite sums. Equations (9) are no longer able, however, to conserve the entire statistical distribution of vorticity. Equations (10) and (11) are two first integrals of (9).

Equation (9) may be regarded as describing a nonlinear interaction between the components whose coefficients are $C_{\mathbf{H}}$ and $C_{\mathbf{M}-\mathbf{H}}$, to alter a third coefficient $C_{-\mathbf{M}}$. Let us seek the maximum simplification of (9) which still describes this process. Clearly at least three terms with different eigenvalues must be retained.

A simple system may be formed by retaining only those eight significant terms for which m and n each assume a value of 1, 0, or -1 (the term $m = n = 0$ is superfluous). Then (4) reduces to

$$\begin{aligned} \nabla^2 \psi &= A_{10} \cos kx + A_{01} \cos ly + \\ &+ A_{11} \cos (kx + ly) + A_{1, -1} \cos (kx - ly) + \\ &+ B_{10} \sin kx + B_{01} \sin ly + B_{11} \sin (kx + ly) + \\ &+ B_{1, -1} \sin (kx - ly). \end{aligned} \quad (12)$$

Further simplifications appear when it is noted that if B_{10} , B_{01} , B_{11} , and $B_{1, -1}$ all vanish initially, they will remain zero, and if, in addition, $A_{1, -1} = -A_{11}$ initially, $A_{1, -1}$ will remain equal to $-A_{11}$. Letting $A_{01} = A$, $A_{10} = F$, $A_{1, -1} = G$, we obtain, as the ultimate simplification of the vorticity and the stream function,

$$\begin{aligned} \nabla^2 \psi &= A \cos ly + F \cos kx + \\ &+ 2G \sin ly \sin kx, \end{aligned} \quad (13)$$

$$\psi = -\frac{A}{l^2} \cos ly - \frac{F}{k^2} \cos kx - \frac{2G}{k^2 + l^2} \sin ly \sin kx. \quad (14)$$

The governing equations, obtained either from (9), or by substituting (13) and (14) directly into (1), are

$$\frac{dA}{dt} = -\left(\frac{1}{k^2} - \frac{1}{k^2 + l^2}\right) klFG, \quad (15)$$

$$\frac{dF}{dt} = \left(\frac{1}{l^2} - \frac{1}{k^2 + l^2}\right) klAG, \quad (16)$$

$$\frac{dG}{dt} = -\frac{1}{2} \left(\frac{1}{l^2} - \frac{1}{k^2}\right) klAF. \quad (17)$$

The coefficients of FG , AG and AF are actually determined by the ratio k/l . The mean kinetic energy and the mean square of the vorticity,

$$E = \frac{1}{4} \left(\frac{A^2}{l^2} + \frac{F^2}{k^2} + \frac{2G^2}{k^2 + l^2} \right), \quad (18)$$

and

$$V = \frac{1}{2} (A^2 + F^2 + 2G^2), \quad (19)$$

are readily seen to be conserved under equations (15—17).

In the remainder of this work we shall be primarily concerned with the simplifications appearing in equations (13—17). These equations presumably contain the minimum number of degrees of freedom required to picture nonlinear barotropic phenomena. We shall call this set of equations the *minimum hydrodynamic equations*, or sometimes simply the *minimum equations*.

In equation (14), the first term on the right is independent of x , and therefore represents the zonal flow. Because there are so few degrees of freedom, the latitudes ($y = \pi/2l$, $3\pi/2l$, — — —) of the zonal wind maxima are fixed, but the intensity A/l may vary. Thus the variable A agrees with the meteorologist's concept of the zonal index.

The remaining terms represent disturbances superposed on the zonal flow. Together they describe a wave of a single wave number, but a variable shape, and, except at certain latitudes, a variable phase. Again because there are so few degrees of freedom, the shape of the wave depends upon the phase, so that the model

cannot picture the motion of disturbances without change of shape. It can, however, picture the nonlinear interaction between the zonal flow and the superposed disturbances.

In general, a set of nonlinear differential equations must be converted into a set of algebraic equations, by replacing time derivatives by finite differences, before a solution can be obtained. The minimum hydrodynamic equations (15—17), however, may be solved analytically. Either by eliminating two variables through the first integrals (18) and (19), and solving the remaining equation, or simply by observing that A , F , and G are three quantities, each of whose derivatives is proportional to the product of the remaining two quantities, we find that the solutions of (15—17) are elliptic functions of time. The particular elliptic functions depend upon the ratio $\alpha = k/l$, and the ratio V/E which is determined by the initial conditions.

If $\alpha > 1$, and if $V/E < 2k^2$, we find in view of (18) and (19) that A , F , and G have maximum values A^* , F^* , and G^* given by

$$\left. \begin{aligned} A^{*2} &= A^2 + \alpha^{-4} F^2 \\ F^{*2} &= F^2 + 2(1 - \alpha^{-4})^{-1} G^2 \\ G^{*2} &= G^2 + \frac{1}{2}(1 - \alpha^{-4}) F^2 \end{aligned} \right\} \quad (20)$$

The solutions of (16—18) are then

$$\left. \begin{aligned} A &= A^* \operatorname{dn} h(t - t^*) \\ F &= F^* \operatorname{sn} h(t - t^*) \\ G &= G^* \operatorname{cn} h(t - t^*) \end{aligned} \right\} \quad (21)$$

where the modulus k_0 of the elliptic functions is given by

$$k_0^2 = \frac{2}{\alpha^4(\alpha^4 - 1)} \frac{F^* G^*}{A^{*2}} \quad (22)$$

h is given by

$$h^2 = \frac{1}{2} \frac{\alpha^2(\alpha^2 - 1)}{\alpha^2 + 1} A^{*2}, \quad (23)$$

and t^* is the time when $A = A^*$, $F = 0$, and $G = G^*$, given by the elliptic integral

$$ht^* = \int_0^{\Theta} \frac{d\Theta}{\sqrt{1 - k_0^2 \sin^2 \Theta}} \quad (24)$$

$\tan^{-1} F_0/G_0$

The solution is periodic of period $4K/h$, where K is given by the complete elliptic integral

$$K = \int_0^{\pi/2} \frac{d\Theta}{\sqrt{1 - k_0^2 \sin^2 \Theta}}. \quad (25)$$

With different initial conditions, or a different ratio α , the elliptic functions dn , sn , and cn may correspond to different permutations of A , F , and G .

When a numerical solution is desired, it is often simpler to solve equations (15–17) numerically than to use a table of elliptic functions. To convert the differential equations into algebraic equations which may be solved numerically, we choose a fixed time interval Δt , and let A_n , F_n , and G_n be the values of A , F and G at time $t_0 + n\Delta t$. We may then use the centered difference formula

$$A_{n+1} = A_{n-1} + 2 \left(\frac{dA}{dt} \right)_n \Delta t, \quad (26)$$

with analogous equations for F_{n+1} and G_{n+1} .

At the first time step, we cannot use a centered difference formula. Accordingly, we may obtain a first approximation $A_{(1)}$ to A_1 by the uncentered difference formula

$$A_{(1)} = A_0 + \left(\frac{dA}{dt} \right)_0 \Delta t, \quad (27)$$

with analogous formulas for $F_{(1)}$ and $G_{(1)}$. We then obtain the final value A_1 by the formula

$$A_1 = A_0 + \frac{1}{2} \left[\left(\frac{dA}{dt} \right)_0 + \left(\frac{dA}{dt} \right)_{(1)} \right] \Delta t, \quad (28)$$

with analogous formulas for F_1 and G_1 .

When partial differential equations are integrated numerically through the use of finite differences in space and time, the phenomenon of computational instability may arise. In order to avoid such instability, the time interval Δt must be chosen to be not too large compared to the space intervals Δx and Δy . The corresponding condition for computational stability, when orthogonal functions are used, is that Δt be not too large a fraction of the period of oscillation of the most rapidly oscillating variable. Thus, whichever method is used, the smaller the scale of the phenomena

which are admitted, the smaller Δt must be in order to avoid computational instability.

The phenomenon of computational instability is easily recognized when it arises, and may be eliminated by decreasing Δt . In the case of the minimum hydrodynamic equations, where only phenomena of very large scale are admitted, Δt may be moderately large.

3. Particular solutions of the minimum hydrodynamic equations

We have seen that the minimum equations (15–17) preserve the average kinetic energy. At the same time, since they admit variations of A , they allow the kinetic energy of the zonal motion to vary, and hence allow exchanges of kinetic energy between the zonal flow and the disturbances. In actual barotropic flow, such exchanges can only accompany a net transport of momentum by the disturbances to or from the zones of maximum flow. In our first numerical example, we shall show that this phenomenon is adequately described by the minimum equations.

The quantity $2\pi/l$ is the distance between successive zonal wind maxima, while $2\pi/k$ is the wave length of the disturbances. Let us choose $2\pi/l = 10,000$ kilometers and $2\pi/k = 5,000$ kilometers, so that $\alpha = 2$. In so doing we shall be choosing systems of a size comparable to the large-scale flow systems in the earth's atmosphere. Equations (15–17) then become

$$\left. \begin{aligned} \dot{A} &= -\frac{1}{10} FG \\ \dot{F} &= -\frac{8}{5} AG \\ \dot{G} &= -\frac{3}{4} AF \end{aligned} \right\} \quad (29)$$

where a dot denotes a time derivative.

The unit of time may be conveniently chosen equal to 3 hours, so that the unit of vorticity is $(3 \text{ hours})^{-1}$, or approximately the value of the Coriolis parameter in middle latitudes.

For initial conditions, let us choose $A = 0.12$ units, $F = 0.24$ units, and $G = 0$. The initial maximum zonal wind speed, occurring where $\gamma = 3\pi/2l$, is then 64 kilometers per

hour, while the initial maximum vorticity, occurring where $x = 0$, and $y = 0$, is 0.36 time the Coriolis parameter. Since $F = 2A$, the zonal kinetic energy and disturbance kinetic energy are initially equal.

Applying equations (20–25), we find that the analytic solution of equation (29) is

$$\left. \begin{aligned} A &= 0.1342 \operatorname{dn}(0.147t + 11.03) \\ F &= 0.2400 \operatorname{sn}(0.147t + 11.03) \\ G &= 0.1643 \operatorname{cn}(0.147t + 11.03) \end{aligned} \right\} \quad (30)$$

where the modulus of the elliptic functions is $k_0 = 0.2$, while the period is 44.1 time units, or 132.3 hours.

Instead of consulting tables of elliptic functions, we may solve equations (29) numerically. If time interval Δt is chosen to be 6 hours, or 2 time units, the algebraic approximations to (29) are

$$\left. \begin{aligned} A_{n+1} &= A_{n-1} - 0.4 F_n G_n \\ F_{n+1} &= F_{n-1} + 6.4 A_n G_n \\ G_{n+1} &= G_{n-1} - 3.0 A_n F_n \end{aligned} \right\} \quad (31)$$

with appropriate modifications for the first time step.

The numerical integrations of (31) may be carried out with a desk computer, or even a slide rule, at a computation rate comparable to one minute per time step. The computations are presented in table 1.

The products FG , AG and AF , which occur on the right hand side of (31), are included in the tabulation.

From the numerical point of view, we observe that F becomes zero, and A and G reach maxima, after about $5\frac{1}{2}$ time steps, or 33 hours, while A and G reach their original values after about 11 steps, after which the previous history must be repeated, except that F and G have opposite signs. The full period is therefore about 22 steps, or 132 hours, which agrees closely with the analytic solution, while the maxima of A and G also agree closely with A^* and G^* as determined by (20).

From the physical point of view, the initially symmetric disturbance is distorted by the zonal flow so as to transport momentum into the zones of maximum zonal winds, thereby increasing the zonal kinetic energy, at the expense of the disturbance kinetic energy. Eventually, the wave becomes distorted to a shape where it transfers momentum out of the zonal wind maxima, and the zonal kinetic energy decreases again to its original value.

From the synoptic point of view, the stream function field initially consists of alternate zones of high and low values of ψ , crossed by a series of troughs and ridges with axes oriented N–S, as shown in the first map in fig. 1. The troughs and ridges are carried eastward in the eastward flow, and westward in the westward flow, so that the axes acquire a NE–SW orientation across the latitudes of high ψ , and a NW–SE orientation across the latitudes of low ψ , as shown in the second map, which follows the first by 18 hours. As described by STARR (1948), eastward momentum is carried northward where the troughs and ridges tilt NE–SW, and southward where

Table 1. Numerical integration of the minimum equations, with the indicated initial conditions, for the case $\alpha = 2$, with time steps equal to two time units.

n	A_n	F_n	G_n	$A_n F_n$	$A_n G_n$	$F_n G_n$
0	0.120	0.240	0.000	0.0288	0.0000	0.0000
1	0.121	0.232	–0.043	0.0281	–0.0052	–0.0100
2	0.124	0.207	–0.084	0.0257	–0.0104	–0.0174
3	0.128	0.165	–0.120	0.0211	–0.0154	–0.0198
4	0.132	0.108	–0.147	0.0143	–0.0194	–0.0159
5	0.134	0.041	–0.163	0.0055	–0.0218	–0.0067
6	0.135	–0.032	–0.163	–0.0043	–0.0220	0.0052
7	0.132	–0.100	–0.150	–0.0132	–0.0198	0.0150
8	0.129	–0.159	–0.123	–0.0205	–0.0159	0.0196
9	0.124	–0.202	–0.088	–0.0250	–0.0109	0.0178
10	0.122	–0.229	–0.048	–0.0279	–0.0059	0.0110
11	0.120	–0.240	–0.004	–0.0288	–0.0005	–0.0010
12	0.122	–0.232	0.038	–0.0283	0.0046	–0.0088

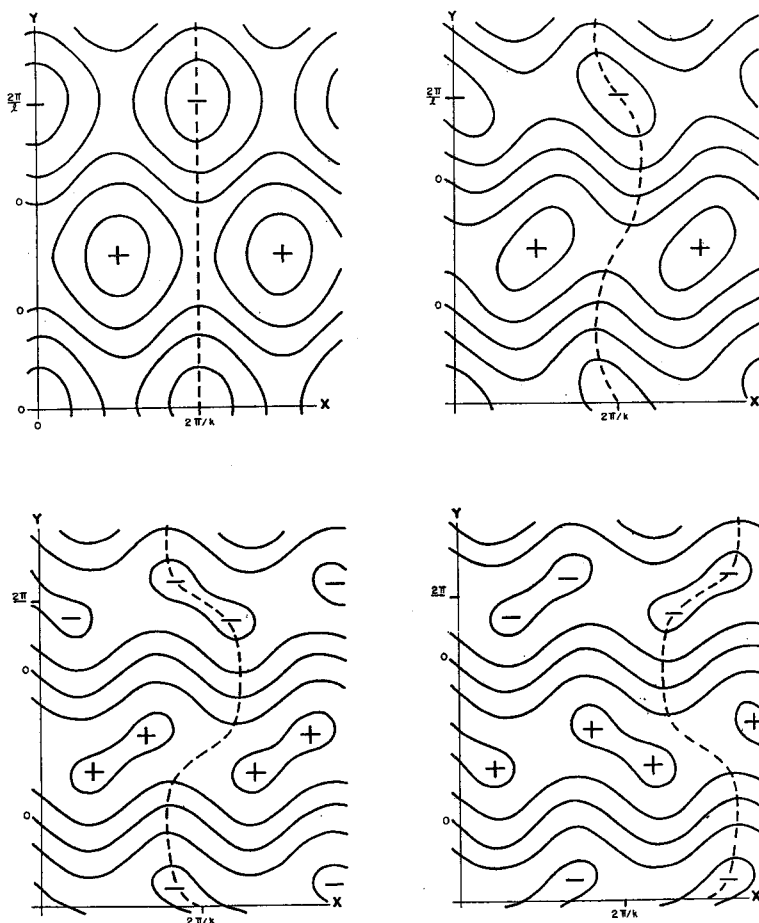


Fig. 1. Stream function patterns for the solutions of the minimum equations given in table 1. Upper left: 0 hours; upper right: 18 hours; lower left: 30 hours; lower right: 36 hours. Streamlines are drawn at intervals of 0.070 units. Dashed lines represent trough lines.

they tilt NW—SE, so that in this case eastward momentum is carried into the zones of eastward flow, in agreement with the increasing zonal index. After 30 hours, as seen in the third map, each center of high or low ψ breaks into two centers. Still later, the NE lobe of each high center departs from the SW lobe altogether, and merges with the SW lobe of the next high center to the east, while the low centers perform a similar process, so that, shortly after 36 hours, as seen in the fourth map, the streamline patterns become the mirror image, in the line $x = 3\pi/2k$, of the map at 30 hours. Patterns after 36 hours become the

mirror images of patterns before 30 hours, and, shortly after 66 hours, the pattern regains its initial appearance, except that it is displaced by one-half wave length. After 132 hours, the full cycle has taken place.

The solution presented in table 1 suggests that the zonal flow is stable, since the disturbances show no tendency for exponential growth. However, a stable flow strictly speaking is one which can exist by itself, and which does not tend to become further disturbed after being slightly disturbed. In a numerical study of stability we should therefore choose initial conditions where the

disturbance kinetic energy is small compared to the zonal kinetic energy.

According to the linear theory, in which the variable A is replaced by a constant, equations (16) and (17) will govern a stable zonal flow if the coefficients of AG and AF have opposite signs, so that the solutions for F and G are trigonometric, and they will govern an unstable flow if the coefficients have the same sign, so that the solutions are exponential. Examining these coefficients in (16) and (17), we see that the zonal flow is stable if $\alpha > 1$, and unstable if $\alpha < 1$. In particular, the zonal flow in the previous example is stable. A similar condition for stability holds even when an infinite number of orthogonal functions are present in the series for ψ and $\nabla^2\psi$, as shown by FJØRTOFT (1953).

Let us see what the nonlinear equations tell us about stability. Let us choose two cases, one where $\alpha = 1.05$, and one where $\alpha = 0.95$. The corresponding sets of differential equations are

$$\left. \begin{aligned} A' &= -0.453 FG \\ F' &= 0.551 AG \\ G' &= -0.049 AF \end{aligned} \right\} \quad (32)$$

and

$$\left. \begin{aligned} A' &= -0.553 FG \\ F' &= 0.451 AG \\ G' &= 0.051 AF \end{aligned} \right\} \quad (33)$$

We shall integrate these equations numerically by the same procedure as that used previously,

again choosing 3 hours as our unit of time, and 6 hours as the length of a time step. In each case the initial eddy kinetic energy is about one per cent of the zonal kinetic energy. The integrations are presented in table 2.

In the first case, there is no tendency for the disturbance to grow. It simply progresses, with a period of about 18 time steps, or 108 hours, while the fluctuations of the zonal flow are very small. The solution differs only slightly from the trigonometric solution of the linearized equations.

In the second case, the disturbance begins to grow exponentially (F resembles a hyperbolic cosine, while G resembles a hyperbolic sine). Likewise, the rate of decrease of A increases exponentially. The exponential increase of F and G does not proceed without limit, however, as in the linear case, but is inhibited as A becomes smaller, and ceases altogether when A reaches zero.

Thus, as might have been expected, the linearized equations describe the changes of the system very well while the disturbances remain small, but not after they become comparable in size to the zonal flow itself.

The unstable solution throws considerable light upon the whole phenomenon of instability, which is not apparent in considering the linear equations. Instability itself is pictured here as a nonlinear phenomenon. The growth of F is due to the nonlinear interaction of A and G , while the growth of G is due to the nonlinear interaction of A and F , or, taken as a whole, the disturbance grows because of

Table 2. Numerical integration of the minimum equations, in the case of a disturbed stable zonal flow ($\alpha = 1.05$), and a disturbed unstable zonal flow ($\alpha = 0.95$).

n	stable case, $\alpha = 1.05$			unstable case, $\alpha = 0.95$		
	A_n	F_n	G_n	A_n	F_n	G_n
0	1.000	0.100	0.000	1.000	0.100	0.000
1	1.000	0.094	— 0.010	0.999	0.105	0.010
2	1.002	0.078	— 0.018	0.998	0.118	0.021
3	1.003	0.054	— 0.025	0.994	0.143	0.034
4	1.004	0.023	— 0.029	0.987	0.179	0.050
5	1.004	— 0.010	— 0.030	0.974	0.232	0.070
6	1.004	— 0.043	— 0.027	0.951	0.302	0.096
7	1.002	— 0.070	— 0.021	0.910	0.397	0.129
8	1.001	— 0.089	— 0.013	0.838	0.514	0.170
9	1.000	— 0.099	— 0.003	0.717	0.654	0.217
10	1.000	— 0.096	0.006	0.524	0.795	0.266
11	1.001	— 0.086	0.016	0.249	0.908	0.303
12	1.003	— 0.061	0.023	— 0.085	0.931	0.312

its nonlinear interaction with the zonal flow. Eventually the disturbance ceases to grow, and begins to decay, also because of its nonlinear interaction with the zonal flow. The source or sink of disturbance kinetic energy during the entire process is the zonal kinetic energy.

We have seen that the minimum equations are capable of picturing the motion of disturbances in a zonal current, and the stability or instability of a zonal current. Another related atmospheric phenomenon is the index cycle, i.e., the tendency for the mean strength of the middle-latitude zonal westerly winds to undergo rather marked fluctuations in intensity, usually with periods of a few weeks. It is interesting to see what light may be thrown upon the index cycle by the minimum equations.

In the example appearing in table I, the zonal index A undergoes fluctuations, but not the violent fluctuations of an index cycle. Moreover, the constraints of the model require some fluctuations of A if the disturbances are to move at all. With different initial conditions, even though $\alpha > 1$, the fluctuations of A need not be so weak. From (18) and (19), we observe that if $V/E < 2k^2$, and $\alpha > 1$, A can never become zero, but if $V/E > 2k^2$, there is no apparent reason why A cannot become zero. For $\alpha = 2$, the condition $V/E = 2k^2$ implies that $G^2 = \frac{15}{2} A^2$. Accordingly, we shall integrate equations (15–17) numerically for

two cases of large disturbances, one case where $G^2 < \frac{15}{2} A^2$, and one where $G^2 > \frac{15}{2} A^2$. We shall choose all quantities except the initial conditions as in the first example. The solutions appear in table 3.

In the first case, where initially $G = 5/2 A$, the solution is qualitatively like the example in table 1, although the fluctuations of the zonal index are more pronounced. In the second case, where $G = 10/3 A$ initially, the situation is different. Here the zonal index changes sign, so that easterlies occupy the region where westerlies were originally present. Moreover, G does not change sign, so that the disturbances do not continue to progress in one direction, but move back and forth, according to the sign of the zonal index. The index shows a period of about 26 time steps, or $6\frac{1}{2}$ days.

Separating these cases there is a discontinuity in the amplitude of the fluctuations of A . Either the total range of A is not greater than A^* , if $V/E < 2k^2$, or it is twice A^* , if $V/E > 2k^2$. We shall say that A undergoes minor fluctuations in the former case, and major fluctuations in the latter. It seems reasonable to identify major fluctuations with an index cycle. It should be noted that, in contrast to the solution in table 2, the major fluctuations do not result from the instability of the zonal flow with respect to small disturbances. It is only because the disturbance is large that major fluctuations can occur.

Table 3. Integration of the minimum equations for $\alpha = 2$, in the case of a subcritical disturbance ($G_0 = 5/2 A_0$) and a supercritical disturbance ($G_0 = 10/3 A_0$).

n	subcritical case			supercritical case		
	A_n	F_n	G_n	A_n	F_n	G_n
0	0.120	0.000	0.300	0.120	0.000	0.400
1	0.117	0.115	0.290	0.114	0.154	0.386
2	0.107	0.217	0.260	0.096	0.282	0.347
3	0.094	0.293	0.221	0.075	0.367	0.305
4	0.081	0.350	0.179	0.051	0.429	0.265
5	0.069	0.386	0.136	0.030	0.451	0.239
6	0.060	0.410	0.099	0.008	0.475	0.225
7	0.053	0.424	0.062	—0.013	0.463	0.228
8	0.049	0.431	0.031	—0.034	0.454	0.243
9	0.048	0.434	—0.001	—0.057	0.410	0.274
10	0.049	0.431	—0.031	—0.079	0.354	0.313
11	0.053	0.424	—0.064	—0.101	0.252	0.358
12	0.060	0.409	—0.099	—0.115	0.122	0.389
13	0.069	0.386	—0.138	—0.120	—0.034	0.400
14	0.081	0.348	—0.179	—0.110	—0.185	0.377

This solution suggests a plausible hypothesis for the existence of index cycles in the atmosphere, which we shall now describe. According to this hypothesis, the ratio V/E of the variance of vorticity to the kinetic energy is usually subcritical, so that only minor fluctuations of the zonal index take place. These fluctuations may occur within a high-index or low-index regime. Occasionally, as the result of a baroclinic process, the ratio V/E will become supercritical, and the index will change from high to low, or low to high. The index change itself will be a barotropic effect, even though the cause of the supercritical value of V/E will be baroclinic. Because of dissipative effects, the ratio V/E will soon become subcritical again, after which the index will undergo minor fluctuations about its new value, until another baroclinic process makes the value V/E supercritical again. If the time required for dissipative effects to make V/E subcritical, after it becomes supercritical, is about one-half of the natural period of a major fluctuation of the index, the result will be a change from a high to low, or low to high, index regime.

Although this hypothesis may appear plausible, it cannot be claimed that the solution presented in table 2 is any more than a piece of evidence in its favor. To place the hypothesis on a firmer basis, it is necessary to integrate less drastically simplified systems of equations, to see whether major fluctuations still appear in the solutions. In addition, it is necessary to perform a careful observational study, to see whether changes in the zonal index actually do follow large increases in the ratio V/E , or in some equally significant parameter.

Nevertheless, this example demonstrates unmistakably how a plausible hypothesis, capable of being tested by further study, can be formulated on the basis of a dynamic equation which has been simplified far beyond the point where it will yield an acceptable short-range forecast.

4. Further applications of simplified equations

We have shown that it is possible to simplify the dynamic equations governing the atmosphere to the point where they may either be solved analytically, or else be integrated nu-

merically with very little effort, while still retaining the nonlinear character of the equations. At the same time, these simplified equations are realistic enough to describe qualitatively some of the important physical phenomena in the atmosphere, and even to lead to plausible hypotheses concerning phenomena as yet not fully explained.

The degree of simplification which we are permitted to use depends upon the particular phenomena which we wish to investigate. For example, the belts of westerly winds pictured in fig. 1 do not resemble the atmospheric jet stream, since winds of nearly maximum strength occupy a large fraction of the belt. In order to study the behavior of a jet under barotropic flow, and in particular to observe such phenomena as the splitting of a jet into two streams, we must retain considerably more terms in the Fourier series, capable of picturing narrow bands of very strong flow.

Again, if we wish to study simple baroclinic flow, we may use one of the two-layer numerical prediction models in place of the barotropic vorticity equation. The maximum allowable simplification would then retain three degrees of freedom for the flow in each layer, or a total of six dependent variables. With such a system, the instability of zonal baroclinic flow, among other phenomena, could be studied.

As a final example, if we are interested in studying forced baroclinic flow, such as that which characterizes the general circulation of the atmosphere, we may again use the same six dependent variables, and modify the equations by appending terms representing non-adiabatic heating and friction. If, however, we wish to describe the increase in static stability which must accompany the conversion of potential and internal energy into kinetic energy through the sinking of cold air and rising of warm air, we must add at least one dependent variable, the space-averaged static stability. Such a system with seven degrees of freedom has been integrated by BRYAN (1959). Among other phenomena, the system is capable of duplicating the appearance or lack of appearance of waves superposed upon the symmetric flow occurring in the dishpan experiments (FULTZ, 1953), according to whether the external heating is weak or strong. Further refinements introduced by Bryan include the

spherical geometry of the earth, whereupon thirteen degrees of freedom are sufficient.

There is virtually no limit to the number of phenomena which one might study by means of equations simplified according to the manner we have described. In each case, the simplified

equations may seem to be rather crude approximations, but they should clarify our understanding of the phenomena, and lead to plausible hypotheses, which may then be tested by means of careful observational studies and more refined systems of dynamic equations.

REFERENCES

- BRYAN, K., 1959: A numerical investigation of certain features of the general circulation. *Tellus* **11**, pp. 163—174.
- FJØRTOFT, R., 1953: On the changes in the spectral distribution of kinetic energy for two dimensional nondivergent flow. *Tellus* **5**, pp. 225—230.
- FULTZ, D., 1953: A survey of certain thermally and mechanically driven systems of meteorological interest. *Proceedings of the first symposium on the use of models in geophysical fluid dynamics*. Johns Hopkins Univ., Baltimore.
- SILBERMAN, I., 1954: Planetary waves in the atmosphere. *J. Meteor.* **11**, pp. 27—34.
- STARR, V. P., 1948: An essay on the general circulation of the earth's atmosphere. *J. Meteor.* **5**, pp. 39—43.