On the Nonexistence of a Slow Manifold

E. N. LORENZ AND V. KRISHNAMURTHY

Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, MA 02139

(Manuscript received 29 September 1986, in final form 13 April 1987)

ABSTRACT

We define the slow manifold S in the state space of a primitive-equation model as a hypothetical invariant manifold on which there is no gravity-wave activity, and on which unique velocity-potential and streamfunction fields correspond to each isobaric-height field. We introduce a five-variable forced damped model, and show that for this model the point H representing the Hadley circulation and the two orbits forming the unstable manifold of H must lie in S if S exists. We then show that in traveling along one of these orbits one eventually encounters gravity waves, whereupon it follows that S does not exist.

A measure G of gravity-wave activity is found to decrease very rapidly as the external forcing F decreases. An approximate formula is derived for G as a function of F.

We show that a particular nine-variable forced damped model with orography also fails to possess a slow manifold, and we speculate as to the existence of slow manifolds in larger and more realistic models.

1. Introduction

Although a fairly realistic numerical simulation of the global atmospheric circulation demands a rather large primitive-equation (PE) model, many of the basic qualitative features of the circulation may be captured by small PE models. Such models often consist of systems of $3N$ ordinary differential equations, governing the fields of three physical variables, each of which has been represented in some manner by a set of N numbers. The variables may be the pressure and the horizontal wind components, or some related quantities such as the velocity potential χ, the streamfunction ψ, and the isobaric height z.

It is standard practice to treat the $3N$ dependent variables as coordinates in a $(3N)$-dimensional "state space" (or phase space). States of the model atmosphere then become points in state space, while time-dependent solutions of the equations become orbits along which the points travel.

If the model is confined to middle and higher latitudes, points representing states of instantaneous geostrophic balance form an N-dimensional manifold in state space. This manifold is not invariant, i.e., a point on the manifold generally leaves the manifold as it follows its orbit. Equivalently, a state which is initially in geostrophic balance does not remain in geostrophic balance, and short-period inertial–gravity wave activity soon becomes evident.

States satisfying the nonlinear balance equation form another N-dimensional manifold, which is also not invariant, but comes closer to being invariant than the geostrophic manifold in the sense that points do not move away from it so quickly or so far (see Charney, 1955). Other relationships, such as those arising in nonlinear normal-mode initialization (e.g., Machenhauer, 1977; Baer and Tribbia, 1977), define other N-dimensional manifolds, which may be even more nearly invariant. These considerations have led to the concept of a *slow manifold* (see Leith, 1980)—an N-dimensional invariant manifold whose states are completely devoid of gravity-wave activity.

On the slow manifold, the fields of χ and ψ are supposed to be unique functions of z, or, alternatively, χ and z are supposed to be unique functions of ψ. Evidence supporting the reality of the slow manifold was provided by a study (Lorenz, 1980; hereafter referred to as L80) in which extended numerical solutions of a simple forced dissipative PE model failed to reveal the development of any gravity wave activity when none was present initially.

More recently the existence of slow manifolds as exact N-dimensional manifolds has been questioned, apparently first by T. Warn (personal communication, 1983), who noted that the various algorithms for locating the slow manifold appeared to be asymptotic rather than convergent. He therefore postulated a fuzzily defined manifold, or, equivalently, a $(3N)$-dimensional subset of state space that at each point was very thin in $2N$ directions.

Very recently, Warn and Menard (1986, hereafter referred to as WM86), confirmed the presence of gravity waves in certain solutions of the L80 model that had been thought to be free of them, using high-precision arithmetic and high-resolution output. Results favoring the "fuzziness" of the slow manifold were also obtained

by Krishnamurthy (1985) and Vautard and Legras (1986) (hereafter referred to as K85 and VL86), using the model of L80 but with stronger forcing than was used in L80. In both studies, extended numerical solutions that seemed for a while to be free of gravity-wave activity eventually encountered obvious gravity waves. This behavior could have occurred if undetectable gravity waves initially present or introduced by the computational procedure subsequently amplified manyfold because of the instability of the slow manifold; it could also have been produced by the absence of nontrivial solutions permanently devoid of gravity waves, i.e., by the nonexistence of a slow manifold.

As an initial state for an operational weather forecast, a point on a fuzzily defined slow manifold is as satisfactory as one on a sharply defined manifold, but the fuzziness or sharpness is of much theoretical interest. In an attempt to clarify the situation, we (Lorenz, 1986; hereafter referred to as L86) constructed a five-variable unforced undamped model by simplifying the equations of L80. We found that for this model the usual slow-manifold algorithms indeed failed to converge, but we succeeded in identifying a "slowest invariant manifold", on which χ and z were unique functions of ψ, and on much of which gravity waves were not evident, but on parts of which they were unmistakably present. The absence of forcing and dissipation greatly facilitated the study; we were unable to say how greatly it influenced the results.

The purpose of the present study is to demonstrate that a particular model resembling the L86 model, but with forcing and damping, does not possess an invariant slow manifold, and to suggest by inference that a similar result holds for a wide variety of models. To do this we shall show that a particular solution that must originate on the slow manifold, if such a manifold exists, subsequently develops pronounced gravity waves. We shall then show that these waves are not an instability phenomenon. In essence, the seeds of gravity-wave activity are present from the start.

2. The model

Our model is a modification of the model of L86, which in turn is a simplification of the PE model of L80. The latter model was derived from the shallow-water equations on an f-plane by expressing the fields of χ, ψ and z as double Fourier series, and then discarding all of the Fourier modes but three, designated as modes 1, 2 and 3, which formed an interacting triad. Scaled coefficients x_i, y_i and z_i in the truncated expressions for χ, ψ, and z, respectively, for $i = 1$, 2 and 3, became the nine dependent variables. An accompanying quasi-geostrophic (QG) model was produced by discarding the time derivatives and certain other terms in the equations governing x_i, after which x_i and z_i could be eliminated to yield a system of three equations in y_i.

In L86 the terms in the L80 PE model representing forcing, damping, and orographic effects, and also all nonlinear terms except those representing advection of vorticity by the portion of the wind derivable from ψ, were discarded. The QG model of L80 was similarly simplified. A five-variable model was then obtained by combining the QG equations for modes 1 and 2 with the PE for mode 3. Finally, the resulting system was transformed to normal-mode form by introducing variables U, V, W, X and Z, proportional respectively to y_1, y_2, $y_3 + b^2 z_3$, x_3, and $z_3 - y_3$, where b, the single adjustable constant in the system, represented the ratio of the wave length of mode 3 to the Rossby circumference of deformation. For further details the reader should consult L80 and L86.

The equations of the present model are

$$dV/dt = UW - bUZ - aV + aF, \tag{1a}$$

$$dU/dt = -VW + bVZ - aU, \tag{1b}$$

$$dW/dt = -VU - aW, \tag{1c}$$

$$dX/dt = -Z - aX, \tag{1d}$$

$$dZ/dt = bVU + X - aZ. \tag{1e}$$

They were obtained by copying Eqs. (5a)–(5e) from L86 (with the order of the first two equations interchanged) and appending damping terms $-aV$, ..., where a is a damping coefficient, and a forcing term aF. The model reduces to the L86 model when $a = 0$. The model describes a single gravity-wave complex, given by X and Z and oscillating with period 2π, nonlinearly coupled to a single quasi-geostrophically interacting triad or "Rossby wave" complex, given by V, U and W and oscillating with a period that is amplitude-dependent, but is much longer than 2π unless F is close to or exceeds unity. We shall call X and Z the fast variables, and V, U and W the slow variables, even though X and Z may sometimes oscillate slowly, or V, U and W may oscillate rapidly, because of the coupling. On the slow manifold, if it exists, X and Z should be unique functions of V, U and W.

Because of the damping and forcing, the equations possess no quadratic invariants, but the quantity $U^2 - W^2 - X^2 - Z^2$ decays exponentially to zero. In particular, if U^2 and $W^2 + X^2 + Z^2$ are equal initially, they remain equal.

A convenient value for b, which we have used in most of our numerical work, is 0.5. This corresponds to a wavelength of about 8000 km for mode 3. Our time unit, which makes the free gravity wave period equal to 2π, may be assumed to equal about 90 min, so that the convenient value $a = 0.02$, used in many of our computations, implies a damping time of about three days. We have used various values of F exceeding a critical value F_c, which we shall presently define.

Our model is like many forced dissipative models

in that, when the forcing is weak, there is a single steady solution—the directly forced solution—which is stable, but, when the forcing exceeds F_c, the directly forced solution is unstable. In that event there are two additional steady solutions, which are stable when the forcing is not too strong.

To confirm these statements, we observe that Eqs. (1) always possess the steady (directly forced) solution $V = F$, $U = W = X = Z = 0$. The absolute orientations of the Fourier modes do not affect the equations, so that we may, in this model, identify mode 2 with the zonally symmetric flow. The directly forced solution then represents the Hadley circulation, which we shall denote by H. We have chosen to identify mode 2 with the zonal flow, and to force mode 2, because if we forced mode 1 (as in L80) or mode 3, the directly forced solution would not become unstable.

Infinitesimal departures (v, u, w, x, z) from H are governed by the linear equations

$$\frac{d}{dt}\begin{pmatrix} v \\ u \\ w \\ x \\ z \end{pmatrix} = \begin{pmatrix} -a & 0 & 0 & 0 & 0 \\ 0 & -a & -F & 0 & bF \\ 0 & -F & -a & 0 & 0 \\ 0 & 0 & 0 & -a & -1 \\ 0 & bF & 0 & 1 & -a \end{pmatrix}\begin{pmatrix} v \\ u \\ w \\ x \\ z \end{pmatrix}. \quad (2)$$

The eigenvalues of the square matrix satisfy the characteristic equation

$$(\lambda + a)[(\lambda + a)^4 + (1 - F^2 - b^2F^2)(\lambda + a)^2 - F^2] = 0. \quad (3)$$

Equation (3) always possesses two complex conjugate roots λ_4, λ_5 with negative real parts (if $a > 0$), associated with gravity-wave activity, and at least two real negative roots λ_2, λ_3, associated with quasi-geostrophic behavior. The remaining root λ_1 is positive so that H is unstable when $|F| > F_c$, where

$$F_c^2 = a^2(1 + a^2)/(1 + a^2 + b^2a^2). \quad (4)$$

If a is small, F_c is slightly less than a, while if F is not too large, λ_1 is close to $F - F_c$.

When $F > F_c$, Eqs. (1) possess the additional steady solutions $V = F_c$, $U = \pm A$, $W = \mp F_cA/a$, $X = \mp bF_cA/(1 + a^2)$, $Z = \pm baF_cA/(1 + a^2)$, where $A^2 = F_c(F - F_c)$. A plot of U against F for the steady solutions would therefore exhibit a typical pitchfork bifurcation at F_c. The steady solutions with $U = \pm A$ represent flows with stationary Rossby waves, and we shall denote them by R and R'. They appear to be stable for all positive values of b, a and F so that almost all points are attracted to one or the other of them, but since the stability or instability of R and R' will not enter our basic arguments, we shall not seek a proof.

Again when $F > F_c$, the point H possesses a four-dimensional stable manifold S_H, which extends from H in the directions corresponding to $\lambda_2, \ldots, \lambda_5$, and is composed of the points whose orbits approach H instead of R or R' as $t \to \infty$. It separates points attracted to R' from points attracted to R. The point H also possesses a one-dimensional unstable manifold U_H, composed of the points on the two orbits that emanate from H, i.e., that approach H in the two opposite directions corresponding to λ_1 as $t \to -\infty$. On these orbits $U^2 = W^2 + X^2 + Z^2$. These orbits will play a major role in our subsequent arguments.

3. The basic result

We begin by observing that if there is a universal slow manifold S, i.e., one on which X and Z are defined for all combinations of V, U and W, the point H must lie on S. This conclusion seems self-evident, but we may verify it by noting that if the values $V = F$, $U = W = 0$—the values on H—are accompanied by any values of X and Z other than both zero—the values on H—short-period oscillations will ensue. We next maintain that if $F > F_c$, the unstable manifold U_H of H must be contained in S. Our argument is that a short enough segment of U_H originating at H approximates a straight-line segment, with no wiggles, as closely as desired. We shall presently consider this point in greater detail.

To show that there is no universal slow manifold it should therefore be sufficient to show that a point traveling along either of the orbits of U_H will eventually undergo gravity-wave activity. It would be difficult to demonstrate the absence of gravity waves numerically; the possibility that the waves are so weak as to be obscured by the computational uncertainty or even the round-off error always exists. It need not be difficult, however, to demonstrate the presence of gravity waves if they are reasonably strong; unless the Rossby waves are so strong that their period is comparable to the gravity-wave period, any observed oscillations with a period of about 2π can have no other interpretation.

To determine the portion of U_H near H, we note that on U_H, as $t \to -\infty$, the departure of each variable from its value at $t = -\infty$ varies as $\exp(\lambda_1 t)$. Accordingly, we may let $\rho = c\exp(\lambda_1 t)$, where c is arbitrary, and we may then express V as a power series

$$V = \sum_{n=0}^{\infty} V_n \rho^n, \quad (5)$$

with analogous series for U, W, X and Z. It follows that

$$dV/dt = \lambda_1 \sum_{n=0}^{\infty} nV_n\rho^n. \quad (6)$$

Since the constant terms $V_0, \ldots$ are simply the values of the variables when $\rho = 0$, or $t = -\infty$, they are the values on H; therefore, $V_0 = F$ and $U_0 = W_0 = X_0 = Z_0 = 0$.

To determine the remaining coefficients we substitute (5) and (6) into (1), and find that

$$\begin{pmatrix} -\mu_n & 0 & 0 & 0 & 0 \\ 0 & -\mu_n & -F & 0 & bF \\ 0 & -F & -\mu_n & 0 & 0 \\ 0 & 0 & 0 & -\mu_n & -1 \\ 0 & bF & 0 & 1 & -\mu_n \end{pmatrix} \begin{pmatrix} V_n \\ U_n \\ W_n \\ X_n \\ Z_n \end{pmatrix}$$

$$= \sum_{j=1}^{n-1} \begin{pmatrix} -U_{n-j}Y_j \\ V_{n-j}Y_j \\ U_{n-j}V_j \\ 0 \\ -bU_{n-j}V_j \end{pmatrix}, \quad (7)$$

where $\mu_n = n\lambda_1 + a$ and $Y_j = W_j - bZ_j$. For $n = 1$ the square matrix is singular, and is, in fact, the matrix whose determinant was equated to zero to obtain Eq. (3), while the summation is vacuous. Obviously $V_1 = 0$, while U_1 may be chosen at will, after which $W_1 = -FU_1/\mu_1$, $X_1 = -bFU_1/(1 + \mu_1^2)$, and $Z_1 = b\mu_1 FU_1/(1 + \mu_1^2)$. Changing the value of U_1 or changing c merely shifts the time at which any given point of U_H is encountered. For $n > 1$ the matrix in (7) is nonsingular and the summation is nonvacuous, and the values of $V_n, \ldots$ may be found by matrix inversion.

Since to evaluate the elements of the right side of (7) we sum $n - 1$ terms, while for large n to multiply by the inverse of the matrix we effectively divide by $n\lambda_1$, the values of $V_n, \ldots$ should increase no more rapidly than exponentially with n, and the series should have a finite radius of convergence. We find, in fact, from numerical computations, that, after suitably adjusting the value of U_1, the values of V_n for even n, and the values of U_n, W_n and Z_n for odd n are nearly constant, except that the signs alternate, while the magnitude of X_n for odd n decreases approximately as $1/n$. Values of V_n for even n and U_n, W_n, X_n and Z_n for odd n vanish. The series thus converge for $|\rho| < 1$, and, in fact, the series for V and U look much like those for $1/(1 + \rho^2)$ and $\rho/(1 + \rho^2)$.

With convergence established for small values of ρ we can strengthen our claim that U_H lies in the slow manifold S, if S exists. Our argument involves noting that small differences between points near H are governed approximately by Eq. (2).

We assume that S exists but does not contain U_H. Since the constant and linear terms in (5) would define a straight line in state space, any miniscule gravity-wave activity on U_H, when ρ is small, must be described by the terms of higher degree. The distance from a point P moving along U_H to the point on S having the same values of V, U and W must, according to (2), vary in proportion to the gravity-wave activity on U_H, and must therefore decrease with decreasing ρ at least as rapidly as ρ^2. In the limit the distance from P to the closest point on S must decrease equally rapidly. However, Eq. (2) also indicates that the distance from P to a point moving along an orbit in S can amplify with increasing ρ no more rapidly than $\exp(\lambda_1 t)$, or ρ. The distance from P to the closest point on S must then amplify equally slowly. This contradicts our previous finding; therefore, U_H lies in S or else S does not exist.

To construct the portion of U_H near H we have computed the first 64 terms in the series (5), and have then summed the series for various values of $\rho \leq 1/2$. The maximum error due to the omission of higher terms is therefore comparable to 2^{-64}, and so is no larger than the round-off errors in our "double precision" computations. We have let $c = 1/2$, so that $t = 0$ when $\rho = 1/2$. For $b = 0.5$, $a = 0.02$, and the fairly strong forcing $F = 0.2$, the variations of Z determined by series summation appear in the portion of Fig. 1 from $t = -30$ to $t = 0$. The expected quasi-exponential growth occurring near H is seen to persist beyond $t = -10$, while an abrupt decrease sets in before $t = 0$.

The remainder of U_H may easily be found by stepwise numerical integration, starting with the values already found at $t = 0$. We have used a fourth-order Taylor-series procedure, in most cases with the rather small time step $\Delta t = 0.05$, so that more than 100 steps occur within a single gravity-wave period. As a consistency check we have first integrated backward from $t = 0$ to $t = -30$, with $\Delta t = -0.05$, and have found that the values of Z so obtained differ by less than 2×10^{-8} units from the values already obtained by series summation.

The curve produced by forward integration from $t = 0$ to $t = 50$ forms the remainder of Fig. 1. A major minimum of Z occurs near $t = 4$, coincidentally at about the limit of convergence of the power series, and, for $t > 10$, there are unmistakable oscillations of approximate period 2π, which must be gravity waves, and which imply that there is no universal slow manifold. Figure 2 extends the solution to $t = 180$, and

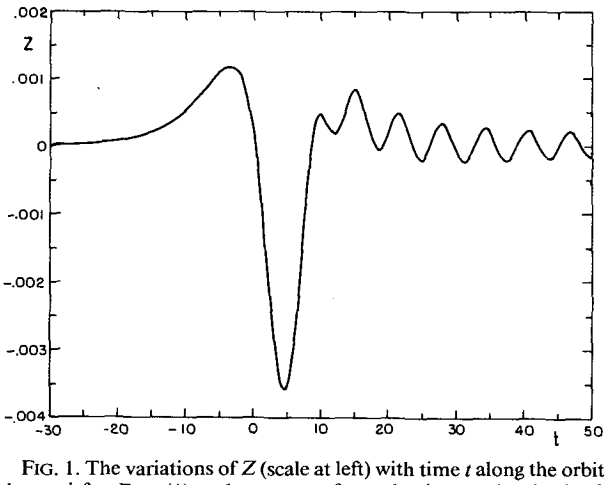

FIG. 1. The variations of Z (scale at left) with time t along the orbit that satisfies Eqs. (1) and emanates from the fixed point H, for $b = 0.5$, $a = 0.02$, and $F = 0.20$. The values of Z for $t < 0$ were obtained by power-series summation, while those for $t > 0$ were obtained by numerical integration.

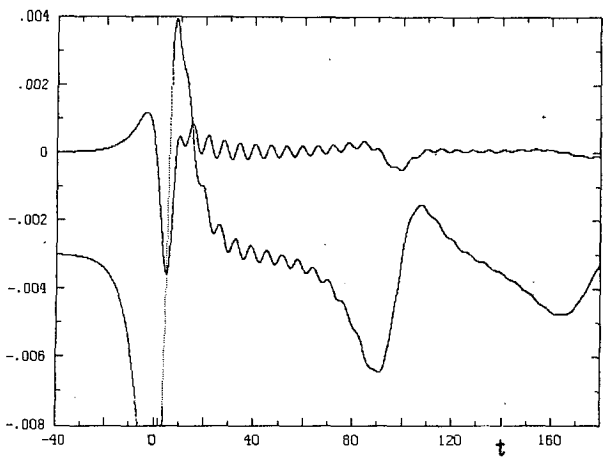

FIG. 2. The variations of Z (upper curve) and X (lower curve) with t for the conditions of Fig. 1, for an extended range of t. The scale at the left is for Z; the curve for X has been displaced downward by 0.003 units to reduce the overlap.

shows X as well as Z. The short-period oscillations of X are comparable in magnitude to those of Z, and lead them by a quarter period, while both the short-period and longer-period variations undergo damping as the orbit proceeds toward R or R'.

It remains for us to demonstrate that the gravity waves in Figs. 1 and 2 are not an artifact of the computational procedure. We have first repeated the integration from $t = 0$ to $t = 50$ with $\Delta t = 0.01$, and have found no detectable change in the graphical output, strongly suggesting that the gravity waves are genuine.

We may define the total error of a point on an orbit as the distance between its true and computed positions. With a reasonably small Δt, the errors in the individual variables added during a single time step by the fourth-order procedure should be comparable to the first omitted terms, i.e., to $(d^5V/dt^5)\Delta t^5/5!$, ... With $\Delta t = 0.05$ we find numerically that the total error introduced during a single time step reaches nearly equal peaks of about 1.9×10^{-12} units near $t = 2$ and $t = 6$, where the steepest slopes appear in Fig. 1, and that it exceeds 1.0×10^{-12} only when $0 < t < 10$. The accumulated total error during the 400 steps from $t = 0$ to $t = 20$ therefore cannot reach 10^{-9} units; with $\Delta t = 0.01$ it cannot even reach 10^{-12} units. On the other hand, according to Fig. 1, if the gravity-wave activity at $t = 20$ is spurious, the total error must exceed 10^{-4} units. It follows that if the gravity waves are spurious they must have been produced not by simple accumulation of errors but instead by amplification by several orders of magnitude of errors already present at $t = 0$ or introduced shortly afterward. We have already noted that such rapid amplification is not to be expected near H.

Since we are now farther from H, we have determined numerically the behavior of a sphere of radius ϵ, centered about the point on U_H where $t = 0$, as each point on the sphere travels along its orbit. The sphere is initially deformed into an ellipsoid, and, if ϵ is small enough, say 10^{-9} units, it is still an approximate ellipsoid when $t = 20$. We find that during this period the longest semi-axis of the ellipsoid never exceeds 3.5ϵ, i.e., there is not even a one-order-of-magnitude amplification of errors. We conclude that the gravity waves displayed in Fig. 1 are real, and that there is no slow manifold.

For good measure we have taken the solution of Figs. 1 and 2 at $t = 50$, and have "initialized" it by retaining the values of V, U and W but substituting the values of X and Z which make d^2X/dt^2 and d^2Z/dt^2 vanish simultaneously. This procedure does not eliminate gravity waves, but it should reduce their amplitude considerably. We have then integrated backward, with $\Delta t = -0.01$, from $t = 50$ to $t = -30$. The result is shown in Fig. 3. The gravity waves remain much weaker than those in Fig. 1 while $t > 0$, but when $t < 0$, there are strong gravity waves where there were none before. Integrating forward from the new values at $t = -30$, with $\Delta t = +0.01$, essentially reproduces the curve in Fig. 3. Evidently, on an orbit close to U_H, if we are to avoid gravity waves following the major minimum of Z, we must have unmistakable gravity waves before the minimum.

4. Other parameter values

Despite the positiveness of our main conclusion we have done nothing to establish its generality. We should like to know whether it is valid for other and perhaps almost all values of the constants in our model, and more importantly, whether it is valid for more general models.

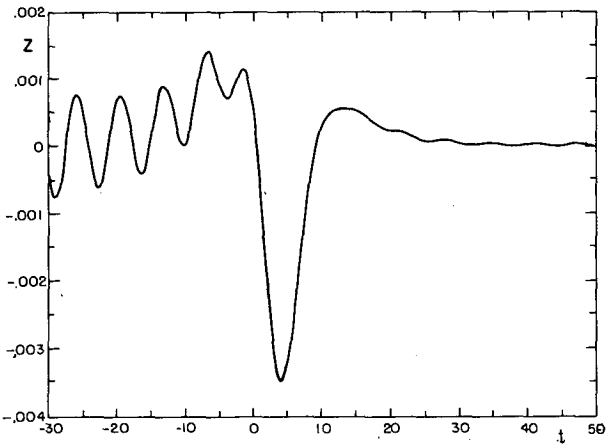

FIG. 3. The variations of Z with t along the orbit that passes through the point obtained by following the orbit of Fig. 1 to $t = 50$, and then replacing the values of X and Z by new values which make d^2X/dt^2 and d^2Z/dt^2 vanish.

We begin by varying F. Figure 4 shows three curves similar to the one in Fig. 1, again computed with $b = 0.5$ and $a = 0.02$, but with $F = 0.18$, 0.16 and 0.14. Evidently the gravity-wave activity falls off very rapidly as F decreases; for $F = 0.14$ it is no longer visible in the graph, although it is easily detected in the numerical output. For $F = 0.10$, not shown in Fig. 4, even the numerical values leave some doubt.

The question thus arises as to whether gravity waves simply become progressively weaker as $F \to F_c$, or whether they disappear altogether at some value of F exceeding F_c. Not the least of the difficulties in reaching an answer is that of precisely defining the presence or absence of gravity waves, when the waves are weak. The longer-period oscillations on which the gravity waves are superposed are generally not sinusoidal, and may therefore possess overtones in the gravity-wave frequency band. Since the longer-period variations have no unique analytic form, we cannot determine the precise magnitude of the overtones, and so we cannot subtract them from the total signal to obtain the gravity waves as a residual.

We can obtain a partial answer by considering the case when $a = 0$, i.e., when Eqs. (1) reduce to the model of L86. Here there is no longer a unique Hadley solution. For every value of F the point $V = F$, $U = W = X = Z = 0$ on the V-axis is a fixed point, and each of these points possesses a one-dimensional unstable manifold U_L, on which $V^2 + U^2 = F^2$ and $U^2 = W^2 + X^2 + Z^2$. Here F is meaningful only as the value of V when $t = -\infty$; it no longer appears in the equations.

The significance of a manifold U_L for our problem is that, except at $t = +\infty$, it is the limiting form of U_H as $a \to 0$ while b and F remain fixed. Fig. 5 is like Fig. 4, but with $a = 0$, and it shows the behavior of Z on U_L for $F = 0.18$, 0.16 and 0.14. For the included range of t the curves in Fig. 5 are much like those in Fig. 4, except that, for a given value of F, the gravity waves are stronger. The activity still decreases rapidly as F decreases. We shall attempt to determine whether or not it ceases altogether before F reaches zero.

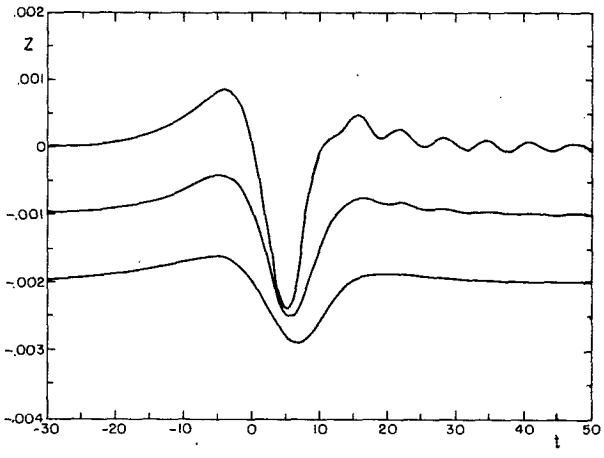

FIG. 4. The variations of Z with t for the conditions of Fig. 1, except that $F = 0.18$ (upper curve), 0.16 (middle curve), and 0.14 (lower curve). The scale at the left applies when $F = 0.18$; the curves for $F = 0.16$ and 0.14 have been displaced downward by 0.001 and 0.002 units, respectively, to reduce the overlap.

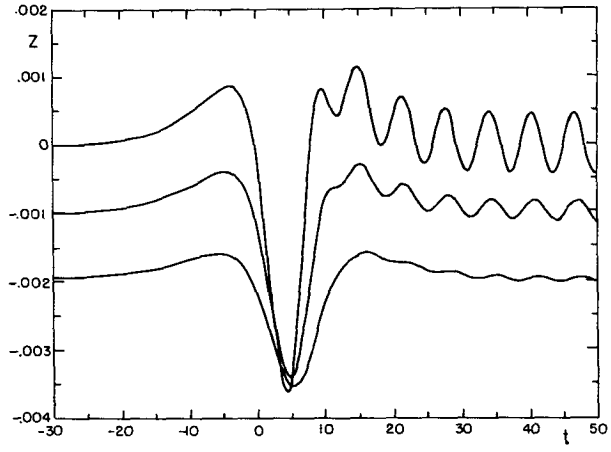

FIG. 5. The same as Fig. 4, except with $a = 0.0$, and with the orbits emanating from the points $V = F$, $U = W = X = Z = 0$.

Along one of the two orbits forming a typical unstable manifold U_L we find that U remains positive, while W, after first becoming negative, eventually becomes positive again. As a definable measure G of the gravity-wave activity that follows the major minimum of Z, we shall choose the value of $(X^2 + Z^2)^{1/2}$, which also equals the value of U, at the time t_W of the first zero-crossing of W. For $F = 0.18$ we find that $t_W = 45.0$; the accompanying gravity-wave activity, with Z oscillating between about -0.0005 and $+0.0005$, appears prominently in the upper right portion of Fig. 5. For $F = 0.16$ and 0.14 the crossings occur respectively at $t = 56.8$ and 73.9, beyond the range of t in Fig. 5.

The upper curve in Fig. 6 shows the variations of G, computed with $\Delta t = 0.02$, and plotted logarithmically, as F varies from 0.1 to 0.2. The activity varies by more than three orders of magnitude, and the decrease with decreasing F is most rapid when F is smallest. Superposed on the decrease there are small-amplitude wobbles, which are more closely spaced when F is smaller. These are evidently associated with the phase that the gravity waves reach at t_W. To see this, observe the lower curve in Fig. 6, which shows the ratio X/U at t_W. Recall that $X^2 + Z^2 = U^2$ at t_W, so that X/U and Z/U are simply the cosine and sine of the phase. It is clear that the wobbles in the upper and lower curves are phase-locked.

If F is not too large, we can remove most of the gravity-wave activity at time t_W by retaining the values of V, U and W and replacing X and Z by $-bVU$ and 0—the values which make d^2X/dt^2 and d^2Z/dt^2 vanish

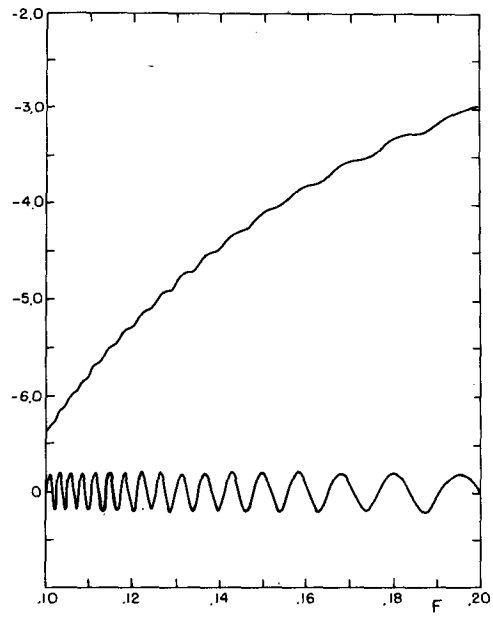

FIG. 6. The variations of the gravity-wave amplitude G (upper curve, with scale for $\log_{10} G$ at left) and the cosine X/U of the gravity wave phase (lower curve) with the external forcing F, for $b = 0.5$ and $a = 0.0$, at the time t_W of the first zero-crossing of W on the orbit that satisfies Eqs. (1) and emanates from the point $V = F$, $U = W = X = Z = 0$.

when $W = 0$. We shall then no longer be on the unstable manifold U_L, since now $W^2 + X^2 + Z^2 = b^2V^2U^2$, which, since V cannot exceed F, is decidedly smaller than U^2. The values of X and Z on U_L should therefore temporarily undergo oscillations of amplitude $G'(F)$, where $(G')^2 = (X + bVU)^2 + Z^2$ about the new solution. For a modified measure of gravity-wave activity, which should nearly eliminate the wobbles from Fig. 6, we shall choose G'.

A casual inspection of the numerical output that accompanied Fig. 6 had suggested that G, and so presumably G', fell off by a factor of about 5.0 whenever the reciprocal of F increased by one unit, so that, approximately,

$$G' = g \exp(-k/F) \tag{8}$$

for suitable constants g and k. In Table 1 we present values of $\log_{10}G'$, again computed with $\Delta t = 0.02$, corresponding to integer values of $1/F$. For the smaller values of F, the increments of $\log_{10}G'$ as $1/F$ increases by one unit are alike to three significant figures, so that (8) becomes a continually better approximation as F decreases to the smallest values for which our computational procedure will yield an answer. The increasing accuracy of (8) strongly suggests that some quantifiable gravity-wave activity persists all the way to $F = 0$. The indicated values of g and k are 2.76 and 1.574, respectively.

When one discovers within a set of numerical solutions a relationship as simple and as closely obeyed as the one apparently connecting G' and F, one often suspects that there is a simple way to deduce the relationship directly from the equations, without recourse to numerical procedures. In the present instance it is fairly simple to obtain an approximation G'' to G'. We first decouple the slow and fast variables by discarding the terms in Eqs. (1) containing b, and solve (1a–c) for V, U and W. We then recouple the variables, and solve (1d) and (1e) for X and Z, using the decoupled values of V and U to evaluate the term bVU in (1e). This proves to be equivalent to expressing each variable as a power series in b, and determining the first nonvanishing term in each series, which in turn is equivalent to expressing V, U, W, X/b, and Z/b as power series in b^2, and determining each leading term.

We therefore let

$$V = \sum_{n=0}^{\infty} v_n b^{2n}, \tag{9a}$$

with analogous expressions for U and W, and

$$X = \sum_{n=0}^{\infty} x_n b^{2n+1}, \tag{9b}$$

with an analogous expression for Z, and substitute the expressions into Eqs. (1), with $a = 0$; we have used lower-case symbols for the coefficients to distinguish them from the coefficients in (5). For $n = 0$ we obtain

$$dv_0/dt = u_0 w_0, \tag{10a}$$

$$du_0/dt = -v_0 w_0, \tag{10b}$$

$$dw_0/dt = -v_0 u_0, \tag{10c}$$

$$dx_0/dt = -z_0, \tag{10d}$$

$$dz_0/dt = v_0 u_0 + x_0. \tag{10e}$$

TABLE 1. Values of E, $L'(E)$, and $\Delta L'(E)$ corresponding to values of F for which E is an integer, where $E = 1/F$, $L'(E) = \log_{10}G'(F)$, and $\Delta L'(E) = L'(E) - L'(E-1)$.

F	E	$L'(E)$	$\Delta L'(E)$
0.500	2	−0.8863	—
0.333	3	−1.5875	−0.7012
0.250	4	−2.2841	−0.6966
0.200	5	−2.9706	−0.6865
0.167	6	−3.6594	−0.6888
0.143	7	−4.3452	−0.6858
0.125	8	−5.0284	−0.6832
0.111	9	−5.7133	−0.6849
0.100	10	−6.3970	−0.6837
0.091	11	−7.0804	−0.6834
0.083	12	−7.7641	−0.6837
0.077	13	−8.4473	−0.6832

From (10a)–(10c) we find, since $v_0 = F$ and $u_0 = w_0 = 0$ at $t = -\infty$, that

$$v_0 = -F \tanh(Ft), \tag{11a}$$

$$u_0 = F \operatorname{sech}(Ft), \tag{11b}$$

$$w_0 = -F \operatorname{sech}(Ft), \tag{11c}$$

after which, from (10d) and (10e),

$$x_0 = -P(t)\cos t - Q(t)\sin t, \tag{11d}$$

$$z_0 = Q(t)\cos t - P(t)\sin t + F \operatorname{sech}(Ft), \tag{11e}$$

where

$$P(t) = F \int_{-\infty}^{t} \operatorname{sech}(F\tau)\cos\tau\, d\tau, \tag{12a}$$

$$Q(t) = F \int_{-\infty}^{t} \operatorname{sech}(F\tau)\sin\tau\, d\tau. \tag{12b}$$

We note that w_0, unlike W, has no zero crossing; instead $w_0 \to 0$ as $t \to +\infty$.

In principle we could estimate the time t_W when W vanishes by finding w_1 and equating $w_0 + w_1 b^2$ to zero, but the equations seem rather intractable. Instead we merely assume that t_W is large, and note that, when t is large, the integrands in (12) are very small, so that $P(t_W)$ and $Q(t_W)$ may be approximated by $P(\infty)$ and $Q(\infty)$. Obviously $Q(\infty) = 0$, since the integrand is odd, while we find from tables of definite integrals or Fourier transforms (e.g., Campbell and Foster, 1948) that $P(\infty) = \pi \operatorname{sech}[\pi/(2F)]$. It follows from (11) that X and Z continue to oscillate as $t \to \infty$, with the approximate amplitude

$$G'' = b(x_0^2 + z_0^2)^{1/2} = \pi b \operatorname{sech}[\pi/(2F)]. \tag{13}$$

For the range of F considered, the hyperbolic secant is hardly distinguishable from twice the negative exponential. The numerically determined value $k = 1.574$ in (8) is closely approximated by $\pi/2$. The main discrepancy is in the coefficient $g = 2.76$, which falls short of $2\pi b$ by about 12 percent when $b = 1/2$. Additional numerical integrations with smaller values of b reveal that, for fixed values of F, the discrepancy between G'' and G' is closely proportional to b^3, suggesting that very close agreement could have been obtained by evaluating one additional term in each series. The existence of a simple approximation where $G'' > 0$ whenever $F > 0$ makes it seem even more likely that, on the actual unstable manifold U_L, gravity waves always develop when $F > 0$.

We note in passing that the hyperbolic secant formula (13) holds only for the first zero-crossing of W. For most values of F in the range that we have considered, the successive gravity-wave amplitudes at a succession of zero-crossings fluctuate aperiodically. Figure 7, which extends over five zero-crossings, illustrates the phenomenon. For values of F that make

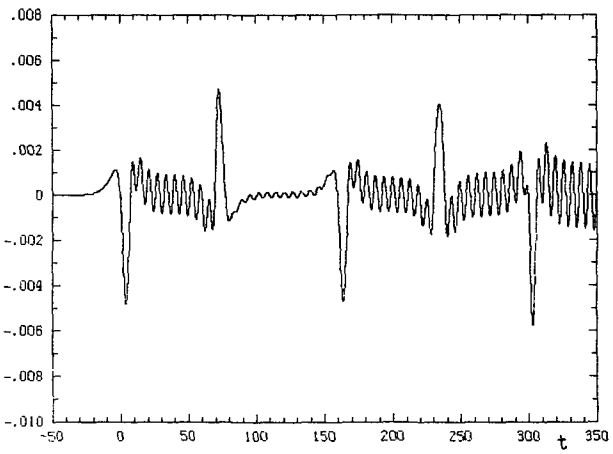

FIG. 7. The variations of Z (scale at left) with t along the orbit that satisfies Eqs. (1) and emanates from the point $V = F$, $U = W = X = Z = 0$, for an extended range of t, for $b = 0.5$, $a = 0.0$, and $F = 0.195$. Zero-crossings of W occur at $t = 38.5$, 118.3, 198.8, 268.4, and 334.1.

$Z = 0$ (or $X/U = \pm 1$; see the lower curve in Fig. 6) at $t = t_W$, the orbits that form U_L are homoclinic, i.e., they converge at $t = \infty$ to the point from which they emanated at $t = -\infty$. One such value, $F = 0.19518$, is close to the value 0.195 used in Fig. 7; this accounts for the weakness of the gravity waves at the second zero-crossing.

We have still not answered the question as to whether gravity waves disappear before F reaches either 0 or F_c, when damping is present. The procedure that we used when $a = 0$ is not universally applicable when $a > 0$; for the unstable manifolds shown in Figs. 1 and 4, for example, W has no zero-crossings. The conclusion that gravity waves will always develop on U_H when $F > F_c$, if $a > 0$, is not incompatible with our result for $a = 0$; neither is the conclusion that they will fail to develop when F is less than some value $F_0 > F_c$, provided that $F_0 \to 0$ as $a \to 0$.

5. Other models

The procedure with which we have investigated the existence of a slow manifold can be applied to a wide variety of PE models. If all of the variables possess continuous time derivatives, and if the external forcing is steady, the equations should possess at least one steady solution or fixed point. If the matrix in the equation [like Eq. (2)] that governs infinitesimal departures from one of the fixed points, say H, possesses at least one real positive eigenvalue, the two orbits emanating from H in the directions corresponding to one of these eigenvalues, say λ_1, will form a one-dimensional manifold U_H. Just as in the model governed by Eqs. (1), U_H must lie in the slow manifold S, if S exists. It is not necessary for H to represent a Hadley circu-

lation. Neither is it necessary for U_H to constitute the entire unstable manifold of H, i.e., for λ_1 to be the only eigenvalue with a positive real part. It is not even necessary for H to be unstable; if λ_1 is real and negative, the orbit that converges to H, rather than emanating from H, in the direction corresponding to λ_1 should lie in S, if S exists.

We may again determine the portion of U_H close to H by summing power series in $\exp(\lambda_1 t)$, if the series converge, and we may then determine the remainder of U_H by numerical integration. If unmistakable gravity waves appear, S does not exist.

We have applied our procedure to one additional model: the 9-variable PE model of L80. As we noted earlier, the variables in the model—x_i, y_i, and z_i, for $i = 1$, 2 and 3—are coefficients occurring in double-Fourier-series expressions for χ, ψ, and z, truncated to three modes. In most of the work of L80 the constants assumed a fixed set of values, which we shall call the "standard" values. In particular, the external forcing, represented by the constants F_1, F_2, and F_3, was confined to mode 1, which in this model represented the zonally symmetric component of the circulation, so that F_2 and F_3 vanished, while the topographic height of the underlying surface, represented by h_1, h_2 and h_3, was zonally oriented, so that h_2 and h_3 vanished. There was therefore a steady Hadley-circulation solution H in which the variables with subscripts 2 and 3 vanished. The time unit was chosen to be 3 h, and the gravity-wave period varied from $2\pi/5$–$2\pi/3$ units.

Much of the work of K85, VL86, and WM86 used the standard values of the constants, except for F_1, which covered a considerable range. The exact form of the equations and the precise values of the constants other than F_1 appear to be of minor concern for the interpretation of our results, and we refer the reader who may wish to perform additional experiments to L80, VL86, or WM86.

With the standard values of the remaining constants, the solution H becomes unstable when F_1 exceeds 0.015, in which case there are two additional steady solutions R and R', which become unstable when F_1 exceeds 0.053. In L80 we dealt mainly with the case where $F_1 = 0.10$. The general solution was aperiodic. Gravity waves, when initially present, seemed to disappear permanently; they could not be detected in the five-decimal-place printout. We had not supposed at that time that a higher-resolution printout was called for; the results of WM86 indicate that such a printout would have revealed weak gravity waves.

With the higher values of F_1 in K85 and VL86, gravity waves often failed to disappear, or they would appear when they were not obviously present initially. As we have noted, such behavior does not by itself contradict the existence of a slow manifold, since it might arise from the instability of such a manifold.

In the present study we have examined the orbits emanating from H in the direction corresponding to the largest real eigenvalue λ_1, for various values of F_1 and for standard values of the other constants. Figure 8 shows the variations of the geostrophic departure z_3–y_3 [the analogue Z in Eqs. (1)] for $F_1 = 0.20$, 0.18 and 0.16. Although the details differ from those in Figs. 4 and 5, the general resemblance is apparent. Again there is a rapid decrease of gravity-wave activity as F_1 decreases. (Somewhat similar curves, obtained by numerical integration from initial points close to H, appear in Figs. 3.30 and 3.31 of K85.)

Auxiliary computations show that for $F_1 = 0.20$ a sphere of small radius ϵ at $t = 0$ becomes an ellipsoid whose longest semi-axis never exceeds 25.0ϵ while $t < 20$, indicating that, as with Eqs. (1), the gravity waves are not an instability phenomenon. We feel that the nonexistence of a universal invariant slow manifold for the model of L80 has been confirmed.

6. Concluding remarks

We have examined two simple forced dissipative models in which the forcing is constant with time, and in the second of which there are orographic effects. We have determined that in neither case is there a universal invariant slow manifold, at least when the forcing is moderately strong. Our method has been to identify an orbit that must lie in the slow manifold if such a manifold exists, and then to show that gravity waves ultimately appear as one follows the orbit.

With sufficient effort our procedure could be applied to much larger models, including some global circu-

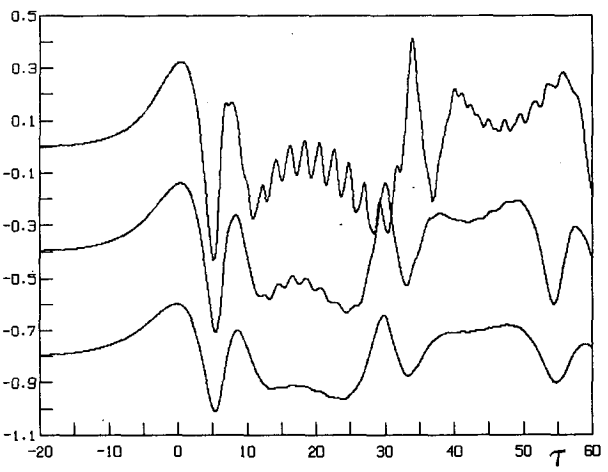

FIG. 8. The variations of z_3–y_3 with time τ along the orbit that satisfies the equations of the PE model of L80 and emanates from the Hadley solution H in the most unstable direction, for standard values of the constants other than F_1, and for $F_1 = 0.20$ (upper curve), 0.18 (middle curve), and 0.16 (lower curve). The scale at the left applies when $F_1 = 0.20$; the curves for $F_1 = 0.18$ and 0.16 have been displaced downward by 0.4 and 0.8 units, respectively, to reduce the overlap.

lation models that provide moderately realistic simulations of atmospheric behavior. We suspect that these models would yield similar results. For the real atmosphere, or for any model which is sufficiently realistic to include the local diurnal variations of solar heating, we would not expect to encounter solutions without gravity waves in any event.

The absence of a slow manifold should not be equated with the presence of gravity waves in the attractor (as illustrated, for example, by Fig. 11 of VL86), although the two phenomena are closely related. In our five-variable model, which has no slow manifold, the attractor set consists of two points representing steady flows, which certainly do not possess gravity waves. Conversely, it must still be considered possible, unless and until it is shown otherwise, that some models possess true slow manifolds which are unstable with respect to gravity-wave perturbations, so that gravity waves become established as the attractors are approached. Some of the numerical solutions in K85, VL86, and WM86 appear to verify the existence of gravity waves in the attractor, but not necessarily the nonexistence of a slow manifold, since a slow manifold need not be an attractor.

A simple qualitative explanation for the eventual presence of gravity waves that are not present initially, whether they subsequently die out or persist, is suggested by our procedure for approximating G' by G''. It appears to be valid for a wide variety of models where the fast variables may be regarded as constituting a forced damped linear oscillator, with the "forcing" F' supplied entirely or mainly by the slow variables. For Eqs. (1) the fast variables are X and Z, and F' is simply the term bVU in (1e). If, as in (1), the orbit emanating from an unstable fixed point is attracted to a fixed point, F' should be expressible as a Fourier integral. If instead the orbit is attracted to a periodic limit cycle, the transient and permanent parts of F' should be given by a Fourier integral and a Fourier series. If the orbit is aperiodic, F' should possess a continuous spectrum, i.e., its serial covariance should be given by a Fourier integral.

Because of the nonlinear advective terms in the equations governing the slow variables, a Fourier series may be expected to possess overtones in the gravity-wave frequency band, while a Fourier integral should overlap the band. In either event the response of the fast variables to the slowest frequencies in F' may be weak, but, close to the resonant frequencies, the response should be greatly enhanced. A possible outcome is that the fast oscillations of the fast variables will have all the properties of separate oscillations superposed on the slow variations, rather than overtones, i.e., they will be gravity waves. There is probably no unequivocal bounding condition separating the cases where gravity waves appear from those where the forcing is so weak that the fast oscillations still qualify as overtones.

The production of one type of motion from another type is by no means a newly recognized phenomenon in fluid dynamics, or even in atmospheric dynamics. In studies somewhat like ours, Errico (1982, 1984) has obtained solutions of multilayer models which appear to be quasi-geostrophic for extended periods, but where gravity waves finally develop and amplify until equipartitioning is reached. On a smaller spatial scale, Lilly (1983) has examined the apparent production of both gravity waves and quasi-two-dimensional turbulence from decaying three-dimensional cumulonimbus convection. Outside of meteorology, an example is the production of audible sound waves from a jet (see Lighthill, 1978).

We note in closing that our principal result is not incompatible with the existence and structure of the "slowest invariant manifold" in L86. For Eqs. (1) with $a = 0$, which are also the equations of L86, there appears to be an extensive region of state space where all solutions are periodic or almost periodic, possessing line spectra. The strictly periodic solutions in L86 which form the slowest invariant manifold occupy this region. With fixed initial values of all of the variables but one, spectral lines coincide with the gravity-wave frequency for discrete initial values of the remaining variable, and near-resonance, marked by prominent gravity waves, occurs in narrow intervals surrounding these values.

There is also an extensive region of state space where almost all solutions are aperiodic, possessing continuous spectra. The portion of the V-axis that we have considered, and the orbits emanating from it, as typified by Fig. 7, lie in this region. Here, with fixed initial values of all variables but one, the spectral continuum overlaps the gravity-wave frequency for all initial values of the remaining variable, and the gravity-wave activity varies continuously with this variable.

Acknowledgments. This work has been principally supported by the Climate Dynamics Program of the Atmospheric Sciences Section of the National Science Foundation, under Grant ATM-8515010. A portion of the work was performed while one of us (E.N.L.) was a summer visitor at the National Center for Atmospheric Research; NCAR is supported by the National Science Foundation. We wish to thank Ho-Chun Huang for critically reading and commenting upon the manuscript.

REFERENCES

Baer, F., and J. J. Tribbia, 1977: On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.*, **105**, 1536–1539.

Campbell, G. A., and R. M. Foster, 1948: *Fourier Integrals for Practical Applications,* D. Van Nostrand, 177 pp. (see p. 73).

Charney, J., 1955: The use of the primitive equations of motion in numerical forecasting. *Tellus,* **7,** 22–26.

Errico, R. M., 1982: Normal mode initialization and the generation of gravity waves by quasi-geostrophic forcing. *J. Atmos. Sci.,* **39,** 573–586.

——, 1984: The statistical equilibrium of a primitive-equation model. *Tellus,* **36A,** 42–51.

Krishnamurthy, V., 1985: The slow manifold and the persisting gravity waves. Ph.D. thesis, Massachusetts Institute of Technology, 146 pp.

Leith, C. E., 1980: Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.,* **37,** 958–968.

Lighthill, J., 1978: *Waves in Fluids,* Cambridge University Press, 504 pp.

Lilly, D. K., 1983: Stratified turbulence and mesoscale variability in the atmosphere. *J. Atmos. Sci.,* **40,** 749–761.

Lorenz, E. N., 1980: Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.,* **37,** 1685–1699.

——, 1986: On the existence of a slow manifold. *J. Atmos. Sci.,* **43,** 1547–1557.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialization. *Beitr. Phys. Atmos.,* **50,** 253–271.

Vautard, R., and B. Legras, 1986: Invariant manifolds, quasi-geostrophy, and initialization. *J. Atmos. Sci.,* **43,** 565–584.

Warn, T., and R. Menard, 1986: Nonlinear balance and gravity-inertial wave saturation in a simple atmospheric model. *Tellus,* **38A,** 285–294.