ASSIMILATION OF ASYNOPTIC DATA AND THE INITIALIZATION PROBLEM

K. P. Bube* and M. Ghil**
Courant Institute of Mathematical Sciences
New York University, New York, N. Y. 10012

* Also Department of Mathematics, Univesity of California, Los Angeles, CA 90024.
** Also Laboratory for Atmospheric Sciences, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

Prepared for:
The Seminar on Data Assimilation Methods, 15-19 September 1980,
The European Centre for Medium Range Weather Forecasts,
Reading RG2 9AX, England.

ABSTRACT

We discuss a mathematical framework for the use of asynoptic data in determining initial states for numerical weather prediction (NWP) models. A set of measured data, synoptic and asynoptic, is termed complete if it determines the solution of an NWP model uniquely. We derive theoretical criteria for the completeness of data sets. The practical construction of the solution from a complete data set by intermittent updating is analyzed, and the rate of convergence of some updating procedures is given.

It is shown that the time history of the mass field constitutes a complete data set for the shallow-water equations. Given that the time derivatives of the mass field are small at initial time, we prove that the velocity field obtained by the diagnostic equations we derive will also have small time derivatives. Hence our diagnostic equations also solve the initialization problem for this system, namely they provide an initial state which leads to a slowly evolving solution to the system.

Finally, we review the bounded derivative principle of Kreiss. It states that in systems with a fast and a slow time scale, initial data can be chosen so that the solution starts out slowly. For such initial data, the solution will actually stay slow for a length of time comparable to the slow time scale. The application of the principle to the initialization problem of NWP is discussed.

1. INTRODUCTION

Numerical weather prediction (NWP) is an initial-value problem for a system of nonlinear partial differential equations in which the initial values are known only incompletely and inaccurately. Data at initial time can be supplemented, however, by observations of the system distributed over a time interval preceding it.

A large number of observations is made by the conventional, ground-based meteorological network. They consist of point values of temperature, humidity, pressure and horizontal velocity. These observations are produced at the so-called synoptic times, 0000 GMT and 1200 GMT. It is customary therefore to choose a synoptic time as initial time for a numerical forecast. Conventional observations are insufficient in number in order to determine the initial state of the model atmosphere. Furthermore, they are very unevenly distributed in space, being concentrated over the continents of the Northern Hemisphere, and much sparser over the oceans and over the Southern Hemisphere.

A large number of additional observations are made at the so-called subsynoptic times, 0600 GMT and 1800 GMT, as well as in an essentially time-continuous manner, using geostationary satellites, polar-orbiting satellites and other nonconventional measuring platforms. All these observations together are called asynoptic.

The object of four-dimensional (4-D) data assimilation is to construct a set of complete, accurate initial data for a NWP model from the measured data, synoptic as well as asynoptic. The completeness question for the set of measured data can be formulated as follows: do these measured data, with their distribution over a time interval and over space, determine the solution of the model equations uniquely? Furthermore, do they determine this solution in a way which depends continuously on the data, so that small errors in the data only cause small errors in the solution?

To understand better the accuracy question, we have to recall that the primitive equations, which govern most operational NWP models today, admit two types of solutions: slow, quasi-geostrophic, meteorologically significant motions, and fast inertia-gravity waves. The latter are present in the real atmosphere only with very small amplitudes. In a model they interfere with short-range forecasts, up to 12h, of all fields of motion, and with longer-range forecasts of vertical velocity and precipitation. It is desirable therefore to eliminate entirely these waves from the model at initial time, or to reduce them as much as possible. To achieve this elimination or reduction is the purpose of initialization.

The goal of preparing initial data in NWP can be restated thus as that of going from a set of inaccurate data, both synoptic and asynoptic, to a complete set of synoptic data which will not generate fast waves when starting a forecast from it. Until rather recently, the meteorological literature has handled separately the two problems of 4-D assimilation, i.e., of passing from nonstandard, combined synoptic and asynoptic data to complete synoptic data, and that of initialization, i.e., of modifying synoptic data to prevent the growth of fast waves in the ensuing forecast.

The purpose of this review is to outline a number of approaches which have contributed to put both aspects of data handling in NWP on a more solid mathematical foundation, as well as pointing in the direction of their eventual unification. A companion article in this volume (Ghil et al., 1980) attempts to provide a systematic, unified theory of 4-D data assimilation and initialization.

In the next section, we shall deal mathematically with the problem of assimilating nonstandard data. The most common practical procedure to use data available at times preceding initial forecast time is intermittent updating. It was suggested by Charney et al. (1969) and independently by Smagorinsky et al. (1970). The model is provided the best available data at some preceding instant, e.g., 24 h or 48 h earlier, and is integrated forward in time. Additional data replace the model values as they become available: the model is updated. When the model integration reaches the initial instant for the next scheduled forecast, its guess of the initial state is blended with the data available at that instant to produce the desired estimate. Thus the model itself is used to assimilate the data available up to the initial instant. Variations of this procedure, as well as other procedures for the four-dimensional (4-D), space-time assimilation of meteorological observations are reviewed in Bengtsson (1975).

Our discussion in Section 2 will be based on the work of Bube (1978, 1980). Connections between the theoretical completeness question and the practice of intermittent updating will be pointed out.

In Section 3, another approach to completeness, based on the work of Ghil (1975; Ghil et al., 1977) will be discussed.

Connections with the initialization problem (Ghil, 1980) will also be outlined.

In Section 4, one particular aspect of the initialization question will be addressed. Various procedures for eliminating or reducing inertia-gravity waves exist, such as the nonlinear normal-mode approach of Baer and Tribbia (1977) or of Machenhauer (1977), and they are reviewed in this volume by Daley (1980). After applying any such procedure, we still have to ask: for how long after initial time has the rapid growth of the fast waves been prevented?

This question is answered by the bounded derivative principle of Kreiss (1979, 1980). Given that the time derivatives of the solution are small initially, these derivatives will stay small for a length of time comparable to the system's slow time scale, 6-24 h say, rather than merely to its fast time scale, O(10 min). The application of the principle to a system of equations of interest in NWP is illustrated. An initialization procedure based on the principle's formalism (Browning et al., 1979) is sketched and its connection with the completeness question and with other initialization procedures is commented upon.

Concluding remarks follow in Section 5.

2. COMPLETENESS AND UPDATING

In this section we review some mathematical results which address the completeness question for the shallow-water equations. These equations have solutions which share some of the essential properties of large-scale atmospheric flow. The initial-value problem for this system of partial differential equations is the problem of determining a solution of the system from complete initial data, i.e., from the global specification of all the state variables at a single time. For this system, the initial-value problem is well posed: given the initial data, there exists a unique solution of the system, and this solution depends continuously on the initial data.

We call nonstandard any set of measured data other than a complete set of initial data. The completeness question for a nonstandard data set can be formulated as follows: is it possible to determine complete initial data at some time from the nonstandard data set, with the initial data depending continuously on the measured nonstandard data set? There are really two separate questions to be answered -- the theoretical uniqueness and stability question for the differential equations, and the practical question of computationally constructing initial data from measurements of a nonstandard data set.

For the sake of brevity, we consider only the case where the nonstandard data set consists mainly of measurements of the mass fields (i.e., temperature and surface pressure), with few measurements of the wind fields. This case is of some historical interest because the large number of mass field measurements available from polar-orbiting satellites was the first important set of nonconventional data used in NWP.

We shall consider in the sequel various forms of the shallow-water equations, in order to illustrate as directly as possible each one of a number of theoretical questions we wish to address. The full, nonlinear system can be written in cartesian coordinates on a plane tangent to the Earth at a latitude θ_0, say, as:

$$u_t + uu_x + vu_y = - \phi_x + fv, \tag{2.1a}$$

$$v_t + uv_x + vv_y = - \phi_y - fu , \tag{2.1b}$$

$$\phi_t + u\phi_x + v\phi_y = - \phi(u_x+v_y). \tag{2.1c}$$

Here x is the coordinate pointing in the zonal, West-East direction, y in the meridional, South-North direction, with u and v the corresponding velocity components. The height of the free surface is h, with $\phi = gh$ the geopotential, g being the acceleration of gravity. The Coriolis parameter f is taken to be constant, $f = 2\Omega \sin \theta_0$, Ω being the angular velocity of the Earth.

We shall be also interested in a linearized form of the equations; the linearization is made around a state with $u = U$, $v = 0$, $\phi = \Phi$, which satisfies the geostrophic constraint, $fU = - \Phi_y =$ const. (see also Ghil <u>et al</u>., 1980). The linear system is

$$u_t + Uu_x = - \phi_x + fv , \tag{2.2a}$$

$$v_t + Uv_x = - \phi_y - fu , \tag{2.2b}$$

$$\phi_t + U\phi_x = - \Phi(u_x + v_y) + fUv. \tag{2.2c}$$

2.1 The completeness question

Bube and Oliger (1977) and Bube (1978, 1980) have considered completeness questions generally for linear first-order hyperbolic systems where the nonstandard data set consists essentially of measurements of a fixed set of the state variables, such as the mass field. We present here the results for the model system

$$\phi_t = a\,\phi_x + b\,u_x\,, \tag{2.3a}$$

$$u_t = b\,\phi_x + a\,u_x\,, \tag{2.3b}$$

with periodic boundary conditions in x. The linearized shallow-water equations (2.2) for purely zonal, $v \equiv 0$, one-dimensional flow around the circle of latitude $\theta = \theta_0$ can be written in this form, where ϕ is the geopotential, u is the zonal velocity, and the constants $-a$ and $-b$ are U and $\sqrt{\Phi}$, the zonal velocity and the geopotential of the basic state, respectively.

We consider the theoretical question first. We want to find nonstandard data sets, consisting as much as possible of measurements of ϕ only, from which complete initial data, say at time $t = 0$, can be determined. Two important observations can be made immediately. First, in order to be able to infer anything about u from measurements of ϕ, the equations must have sufficient linkage between ϕ and u; for system (2.3), we must have $b \neq 0$. For the linearized shallow-water equations, this is satisfied sice $\Phi > 0$. Second, u can be determined at most up to a constant from measurements of ϕ alone, since $\phi = 0$, $u = u_0$ is a solution of system (2.3). It follows that at least one measurement of u is needed.

The periodic domain in x will be the unit interval $0 \leq x \leq 1$. We examine three nonstandard sets of data:

(i) measurements of $\phi(x,0)$ and $\phi_t(x,0)$ for $0 \leq x \leq 1$, and of $u(0,0)$;

(ii) measurements of $\phi(x,t)$ for $0 \leq x \leq 1$ and $t = -j\tau$ for some $\tau > 0$ and $j = 0,1,\ldots,m$, and of $\int_0^1 u(x,0)\,dx$;

(iii) measurements of $\phi(x,t)$ for $0 \leq x \leq 1$ and $-T \leq t \leq 0$ for some $T > 0$, and of $\int_0^1 u(x,0)\,dx$.

For data set (i), the "instantaneous time-history" of ϕ at time $t = 0$ is measured, i.e., its time derivative, in addition to $\phi(x,0)$.

Assuming $b \neq 0$, Eq. (2.2a) becomes an ordinary differential equation for $u(x,0)$, which we can solve using the measured value of $u(0,0)$. For the solution $u(x,0)$ to be periodic, the measured $\phi_t(x,0)$ must satisfy the compatibility condition

$$\int_0^1 \phi_t(x,0)\, dx = 0. \tag{2.4}$$

Thus data set (i) is a good nonstandard data set: we can determine from it $u(x,0)$, and $u(x,0)$ depends continuously on the measured data. This approach corresponds to the use of a generalized set of diagnostic equations to complete a set of initial data. It will be discussed further in the next section.

For data set (ii), we sample the time history of ϕ at discrete, periodic points in time, with our last measurement at the initial time for a forecast. This measurement pattern is closer than (i) to the way data are gathered operationally; in fact, data are measured on even more complicated space-time manifolds. To analyze the completeness question for data sets (ii) and (iii), we expand ϕ and u in Fourier series in x:

$$\phi(x,t) = \sum_{\xi=-\infty}^{\infty} \hat{\phi}(\xi,t)\, e^{2\pi i\xi x}, \tag{2.5a}$$

$$\hat{\phi}(\xi,t) = \int_0^1 e^{-2\pi i\xi x}\, \phi(x,t)\, dt, \tag{2.5b}$$

where ξ is the wave number. Note that

$$\hat{u}(0,0) = \int_0^1 u(x,0)\, dx. \tag{2.6}$$

It can be shown that if τb is a rational number, then data set (ii) does not determine $u(x,0)$ uniquely. If τb is irrational, then data set (ii) determines $u(x,0)$ uniquely, but $u(x,0)$ does not depend continuously in the L^2-norm on the measured data, i.e., root-mean-square (rms) errors in $u(x,0)$ will not be small if the rms errors in the measured data are small.

For data set (iii), however, we do have continuous dependence. In fact, for each $T > 0$, there is a constant C_T for which

$$\| u(x,0) \| \leq C_T(\hat{u}(0,0) + \max_{-T \leq t \leq 0} \| \phi(x,t) \|), \tag{2.7}$$

where $\| \ \|$ is the L^2 or rms norm,

$$\| y(x) \|^2 = \int_0^1 |y(x)|^2 \, dx. \tag{2.8}$$

Estimate (2.7) immediately implies both uniqueness and continuous dependence.

2.2 Computational construction of initial data.

We now address the second question, of actually constructing the initial data from nonstandard, measured data. In particular, we consider the method of intermittent updating, where the measured values of ϕ are inserted into a computation as the computation proceeds. At first it appears that our results on data set (ii) imply that intermittent updating does not use enough data to determine $u(x,0)$ even theoretically in a continuous manner. In practice, however, only a finite set of measurements can be used for any computation, and we cannot hope to determine a function of x completely for all x in the interval $0 \leq x \leq 1$ either. If our discrete measurements of $\phi(x,t)$ are sufficiently dense in x and t, then these measured data are a good approximation to $\phi(x,t)$ in the norm $\max_{-T \leq t \leq 0} \| \phi(x,t) \|$, and data set (iii) is actually the appropriate theoretical model for the measured data set. Thus, theoretically, enough data are being used. We now examine how intermittent updating uses this data set.

For this purpose, we shift the time origin back to the beginning of a numerical computation. Let τ be the updating interval; we have measurements of $\phi(x,t)$ for $t = j\tau$, $j = 0,1,2,\ldots$. We start with the initial measurement of ϕ and an initial approximation to u at $t = 0$. As we integrate forward in time, we replace the computed values of ϕ by the measured values at times $t = j\tau$.

Let $\varepsilon(x,t)$ denote the error in the computed u at time t. Notice that ε is not reduced at the updating times, since u is not updated:

only ϕ is. The error in u will decrease between updating times, as a result of the linkage in the system beween u and ϕ.

Expanding ε in its Fourier series and ignoring the distinction between the numerical solution and the corresponding solution of the differential equation, we can show that

$$\hat{\varepsilon}(\xi,(j+1)\tau) = \rho(\xi,\tau) \quad \hat{\varepsilon}(\xi,j\tau), \tag{2.9}$$

where

$$\rho(\xi,\tau) = |\cos(2\pi b\xi\tau)|. \tag{2.10}$$

So the effect of an update is to multiply each Fourier coefficient of the error in u by a decrease factor $\rho(\xi,\tau) \leq 1$. If $\rho(\xi,\tau) << 1$, then $\hat{\varepsilon}(\xi,j\tau)$ approaches 0 rapidly as j increases. If $\rho(\xi,\tau) \approx 1$, then $\hat{\varepsilon}(\xi,j\tau)$ approaches 0 slowly as j increases.

For $t = j\tau$,

$$|\hat{\varepsilon}(\xi,t)| = \sigma(\xi,\tau)^t|\hat{\varepsilon}(\xi,0)|, \tag{2.11}$$

where

$$\sigma(\xi,\tau) = \rho(\xi,\tau)^{1/\tau}. \tag{2.12}$$

It can be shown that, for $\xi \neq 0$, $\sigma(\xi,\tau)$ is a strictly decreasing function of τ as long as $0 < \tau < 1/(4|b\xi|)$, and

$$\lim_{\tau \to 0} \sigma(\xi,\tau) = 1. \tag{2.13}$$

In other words, as τ decreases, σ increases monotonically and tends to 1. Hence updating ϕ more frequently does not necessarily make $\varepsilon(x,t) \to 0$ faster as t increases. We must allow enough time for the new information to pass from ϕ to u; alternatively, the energy of the error has to have time to pass from u to ϕ, and then out of the system when ϕ is updated.

Note the important dependence of ρ and σ on both wave number ξ and updating interval τ. First, if $\xi = 0$, then $\rho = \sigma = 1$, so no improvement is made in the mean velocity $\hat{u}(0,t)$. If $\varepsilon(x,t)$ is to converge to 0, the value of $\hat{u}(0,0)$ in the initial approximation to u must be correct. For other values of ξ, the best updating interval is

$$\tau \approx \frac{1}{4|b\xi|} .$$

This result for τ has an intuitive explanation. Consider initial data for system (2.3) of the form

$$\phi = 0 , \qquad v = v_0 \sin (2\pi\xi x) .$$

In other words, the error in geopotential is zero, while the error in velocity is purely in wave number ξ. The corresponding solution of (2.3) has

$$\phi(x,t) = v_0 \sin (2\pi\xi bt) \cos [2\pi\xi(at+x)] .$$

The geopotential error at time t generated by the velocity error at time $t = 0$ is thus also a pure ξ-wave, travelling at phase speed a, but amplitude-modulated with frequency $\omega = |b\xi|$. Hence our result means that the best updating interval for this wave of frequency ω is $\tau \approx 1/4\omega$, i.e., 1/4 of the period of this amplitude modulation.

We are interested in what happens if the frequency of updates were increased. For fixed $\xi \neq 0$ and "assimilation cycle length" $t > 0$, letting the update interval τ decrease, $\tau \to 0$, Eqs. (2.11,2.13) imply that

$$|\hat{\varepsilon}(\xi,t)| \to |\hat{\varepsilon}(\xi,0)| ,$$

i.e., the error at time t is not decreased from that at $t = 0$, even though ϕ was updated more often and more data were used. In other words, $\tau \approx 1/4\omega$ is really an optimum, and further decrease of τ is counterproductive.

If the entire error is in only one wave number ξ, i.e., if $\hat{\varepsilon}(\xi,0) \neq 0$ only for one value of ξ, then τ can be chosen judiciously to decrease the error quickly. In practice, $\hat{\varepsilon}(\xi,0) \neq 0$ for many values of ξ. The trouble is that if $\tau >> 1/(4|b\xi|)$, then $\rho(\xi,\tau)$ is just as likely to be close to 1 as it is to be close to 0. The best τ for a small value of $|\xi|$ may give a $\rho \approx 1$ for some large values of $|\xi|$. The dilemma becomes: if τ is too large, waves with high wave number may converge slowly; if τ is too small, waves with low wave number will converge slowly.

One possible solution to this dilemma is to use several different updating intervals $\tau_1,\ldots,\tau_m$, and repeat them in a cycle. If there is a $\tau_j \approx 1/4\omega$ for each frequency ω of interest, then convergence should be reasonably rapid for all corresponding wave numbers. Numerical experiments with model system (2.3) confirm all the theoretical results above.

This analysis extends directly to the case of forward and backward integration in an update cycle: one such cycle is equivalent to two updating intervals. Talagrand (1980) has also analyzed forward-backward updating schemes.

3. COMPLETENESS AND INITIALIZATION

In the previous section we have seen that, for the one-dimensional, linear system (2.3), instantaneous mass field data $\phi(x,0)$ and $\phi_t(x,0)$ determine the velocity field $u(x,0)$, provided the velocity at a single point, $u(0,0)$ say, is given and that (2.4) holds. In this section we shall show first that similar results hold for a two-dimensional linear version of the shallow-water equations, as well as for the full, nonlinear system (2.1). Then it will be shown that complete initial data determined by the procedure outlined here from nonstandard mass field data lead to a solution of (2.1) in which fast waves have moderate amplitude.

3.1 Diagnostic equations for compatible balancing

We consider first the shallow-water equations linearized around a state of rest, $U = 0$, Φ = const.:

$$u_t = - \phi_x + fv \, , \tag{3.1a}$$

$$v_t = - \phi_y - fu \, , \tag{3.1b}$$

$$\phi_t = -\Phi(u_x + v_y) \quad . \qquad (3.1c)$$

We deal here only with this special case of (2.2), in which $U = 0$, for the sake of simplicity.

Our purpose is to derive a set of equations for u and v at time $t = 0$, say, given ϕ and its "instantaneous time history" ϕ_t, as well as higher time derivatives, if necessary, at $t = 0$.

Clearly (3.1c) is such a diagnostic equation for u,v, given ϕ_t. In the present, two-dimensional case, this equation does not determine u and v completely, while (2.3a) did determine u completely for the previous, one-dimensional case.

A second diagnostic equation can be derived easily. Differentiate (3.1a) with respect to x, (3.1b) with respect to y and (3.1c) with respect to t. The quantities u_{xt} and v_{yt} appearing in the differentiation of (3.1c) can then be substituted from their values obtained when differentiating (3.1a) and (3.1b), respectively. This leads to the diagnostic system (Ghil, 1975, 1980):

$$u_x + v_y = -\phi_t/\Phi, \qquad (3.2a)$$

$$u_y - v_x = -(1/f)(\nabla^2\phi - \phi_{tt}/\Phi), \qquad (3.2b)$$

where $\nabla^2\phi = \phi_{xx} + \phi_{yy}$. Thus ϕ_{tt}, as well as ϕ_t, is necessary here to determine u and v at $t = 0$.

The linear system (3.2) is a set of inhomogeneous Cauchy-Riemann equations for the functions v, u; cross-differentiation of these equations would lead to a Poisson equation for either u or v. A boundary-value problem having a unique, stable solution for (3.2) is the Dirichlet problem: prescribing u, say, on a closed contour ∂D. These boundary conditions determine u completely and v up to an additive constant, in the domain D bounded by the contour ∂D; the value of this constant can be given by prescribing v at some point in D. Alternatively, one can prescribe v on ∂D and u at one point. Also u can be given on one part of ∂D and v on the rest. Either one of these possibilities corresponds in the present case to the prescription of $u(0,0)$ in the previous section. Appropriate compatibility conditions, similar to (2.4), are discussed in Ghil and Balgovind (1979).

Hence, within the framework of model (3.1), solving the Dirichlet problem for system (3.2) solves the completeness problem for the nonstandard data set comprised of ϕ, ϕ_t and ϕ_{tt} measured in D, and of u and v measured in a very small subset of D. Furthermore, this solution is stable, i.e., it depends continuously on the nonstandard data.

We are ready to turn our attention to the full, nonlinear system (2.1). A diagnostic system for u and v at t = 0, given ϕ, ϕ_t and ϕ_{tt} at t = 0 can also be derived (Ghil, 1975; Ghil et al., 1977). The procedure is similar to that leading from (3.1) to (3.2), with the material derivative $d/dt = \partial/\partial t + u\,\partial/\partial x + v\,\partial/\partial y$ replacing the partial time derivative $\partial/\partial t$ in the cross-differentiation. The resulting diagnostic system is:

$$u_x + v_y = -\Psi_x u - \Psi_y - \Psi_t, \tag{3.3a}$$

$$\begin{aligned} u_x^2 + 2u_y v_x + v_y^2 + f(u_y - v_x) \\ = f(\Psi_x v - \Psi_y u) + \frac{d\Psi_t}{dt} + u\frac{d\Psi_x}{dt} + v\frac{d\Psi_y}{dt} \\ - \frac{1}{\phi}(\phi_x^2 + \phi_y^2) - \nabla^2\phi, \end{aligned} \tag{3.3b}$$

where $\Psi = \log \phi$. More general diagnostic systems for the wind field, given instantaneous mass field history information, have been derived in Ghil (1975) for three-dimensional primitive-equation systems closely related to those used in operational NWP.

System (3.3) reduces, when $\delta = 0 = d\delta/dt$, with $\delta = u_x + v_y$ being the divergence of the wind field, to the classical Monge-Ampère equation. Its solution, which raises interesting mathematical, numerical and meteorological questions was studied in Ghil et al. (1977). Our system, which generalizes this equation, seems to have a unique, stable solution in most situations of meteorological importance. This settles for our purposes the completeness question with respect to system (2.1) and the nonstandard data set comprised of measurements of ϕ, ϕ_t and ϕ_{tt} at one time instant.

Given measurements of ϕ at discrete time intervals t_j, $\phi_j(x,y) = \phi(x,y,t_j)$, this approach would appear to suggest the possibility of constructing initial data in a way different from updating (cf. also Sec. 2.2, first paragraph). Namely, the history of the mass field, $\phi(x,y,t)$, could be interpolated from $\phi_j(x,y)$; furthermore, the partial time derivatives ϕ_t and ϕ_{tt} can be computed at the last measurement time, t_m ($m \geq 2$), which is the initial forecast time. Then u and v could be computed at $t = t_m$ from system (3.3).

In fact, the nonstandard data sets available today are much more complex, and the completeness question from mass field data is of interest mostly from the historical viewpoint and as a relatively simple illustration. From this illustrative perspective, we shall proceed to point out the connection with the initialization question. Specifically, we wish to show that a solution of the prognostic system (3.1), given initial data satisfying (3.2), will contain no fast waves of large amplitude, provided the mass field data had small ϕ_t and ϕ_{tt} at $t = 0$. A similar result will be outlined for (2.1) and (3.3).

3.2 Compatible balancing and initialization

Various systems of diagnostic equations have been considered in the past under the name of balance equations (Bengtsson, 1975, Ch. 6). The idea of static balancing is precisely that of obtaining a complete initial state which leads to a balanced, i.e., quasi-geostrophic, slowly evolving solution to the prognostic equations. We shall show that the solution to system (3.2), which is dynamically compatible with (3.1), does indeed produce such a slow evolution of the solution to (3.1), when it is used as its initial state.

Define the quantities

$$G \equiv \phi_x - fv, \tag{3.4a}$$

$$H \equiv \phi_y + fu. \tag{3.4b}$$

It follows from (3.4) and (3.2) that G and H satisfy (Ghil, 1980):

$$G_x + H_y = \phi_{tt}/\Phi \;, \tag{3.5a}$$

$$G_y - H_x = (f/\Phi)\phi_t. \tag{3.5b}$$

Eqs. (3.5) are a system with the same structure as (3.2). Under suitable boundary conditions, its solution (G,H) depends continuously on the right-hand side (ϕ_{tt}, ϕ_t). That is, smallness of ϕ_t and ϕ_{tt} will imply smallness of G and H.

But (3.1a,b) state that $u_t = G$, $v_t = H$. Hence, provided ϕ_t and ϕ_{tt} are small at $t = 0$, u_t and v_t for a solution of (3.1) obtained from an initial state given by (3.2) will also be small at $t = 0$. For a discussion of cases in which ϕ_t and ϕ_{tt} are not small at $t = 0$, see Ghil (1980). We conclude that the diagnostic system (3.2) solves the initialization problem for (3.1), given a set of nonstandard data ϕ, ϕ_t and ϕ_{tt} at $t = 0$. It can be shown further that small ϕ_{ttt} will lead to small u_{tt} and v_{tt}, and so on: higher temporal smoothness of the instantaneous mass field data will lead to higher smoothness of the initial velocity tendencies.

For the nonlinear shallow-water equations (2.1) and the corresponding, dynamically compatible diagnostic system (3.3), the proof is similar. It is technically more complex, and will only be sketched here (cf. Ghil, 1980).

Let the variables be nondimensionalized by

$$x = Lx', \quad y = Ly', \quad u = Uu', \quad v = Uv', \tag{3.6a}$$

and

$$\phi = \Phi + \phi_0\phi', \tag{3.6b}$$

with Φ the mean geopotential and ϕ_0,

$$\phi_0 = LfU,$$

a characteristic magnitude of the geopotential's deviation from Φ. Introduce also the Rossby radius of deformation

$$\lambda = \sqrt{\Phi}/f \tag{3.7}$$

and the nondimensional Rossby number

$$\varepsilon = U/Lf, \tag{3.8}$$

which is small for midlatitude large-scale flow. We shall assume that (A0):

$$L/\lambda = O(1), \tag{3.9}$$

which in particular means that inertia-gravity waves have velocities roughly equal to pure gravity waves. This permits us to introduce two nondimensional time scales, a <u>fast time</u> t^* and a <u>slow time</u> τ, by

$$t^* = ft, \tag{3.10a}$$

$$\tau = (U/L)t = \varepsilon t^*. \tag{3.10b}$$

For brevity, we shall use the symbol ∂_t for

$$\partial_t = \partial/\partial t^* + \varepsilon\, \partial/\partial\tau , \tag{3.11}$$

and drop primes in the sequel.

We make now the two assumptions that: (A1) the geopotential data satisfy

$$\phi_x, \phi_y = O(1) , \tag{3.12a}$$

$$\partial_t\phi, \partial_t^2\phi = O(\varepsilon) ; \tag{3.12b}$$

and that (A2) the solution (u,v) of (3.3) with such data has a characteristic length $L_{(u,v)}$ which equals that of the data, $L_{(\phi)}$,

$$L_{(u,v)} = L_{(\phi)} = L, \tag{3.13a}$$

and a characteristic magnitude U such that the corresponding Rossby number R_0 is indeed small

$$R_0 \equiv U/fL = \varepsilon << 1 \ . \qquad (3.13b)$$

Assumption (A2) actually follows from (A1) in certain cases, and seems to be plausible in general.

Defining G and H as before, in nondimensional form though,

$$G = \phi_x - v \ , \qquad (3.14a)$$

$$H = \phi_y + u \ , \qquad (3.14b)$$

yields immediately

$$G_x + H_y = \nabla^2\phi + u_y - v_x \ , \qquad (3.15a)$$

$$G_y - H_x = -\ (u_x + v_y) \ . \qquad (3.15b)$$

The right-hand sides of (3.15), however, cannot be equated directly with ϕ_{tt} and ϕ_t, respectively, as was the case for the linear system. Instead, system (3.3) has to be rewritten in nondimensional form, and Eqs. (3.7–3.13) used to show that in fact

$$G_x + H_y = \{1 + O(1)\}\ O(\varepsilon) \ , \qquad (3.16a)$$

$$G_y - H_x = O(\varepsilon) \ . \qquad (3.16b)$$

From (3.16) we can then conclude, as for (3.5), that $G,H = O(\varepsilon)$. Using (2.1) and (3.14), it follows that

$$du/dt = \partial_t u + \varepsilon(u\ \partial u/\partial x + v\ \partial u/\partial y) = O(\varepsilon) \ ,$$

$$dv/dt = \partial_t v + \varepsilon(u\ \partial v/\partial x + v\ \partial v/\partial y) = O(\varepsilon) \ .$$

This in turn implies $\partial_t u = O(\varepsilon)$, $\partial_t v = O(\varepsilon)$ at $t = 0$.

We saw in Sec. 3.1 that the nonstandard data set comprised of the "instantaneous mass field history" $(\phi, \phi_t, \phi_{tt})$ at $t = 0$

determines completely an initial data set (ϕ, u, v) for the nonlinear shallow-water equations (2.1). The wind field is determined from the diagnostic system (3.3). We have shown here that (3.3) is dynamically compatible with the prognostic system (2.1), which is equivalent to the statement that the smallness of ϕ_t and ϕ_{tt} measured at t = 0 will result in small u_t and v_t at t = 0.

This leaves open the question of how long will (ϕ_t, u_t, v_t), after initialization by compatible balancing, as above, or any other correct initialization procedure (Daley, 1980), stay small. This question is addressed in the next section.

4. THE BOUNDED DERIVATIVE PRINCIPLE

In this section we address the question of how long a solution will continue to vary on the slow time scale after initialization. Kreiss (1978, 1979, 1980) has considered this question for both ordinary and partial differential equations and has formulated an easily stated principle, the bounded derivative principle. We will explain this principle for a system of ordinary differential equations (ODEs) with two time scales, give an example demonstrating the correctness of the principle, and illustrate its application to a system of partial differential equations (PDEs) of interest in NWP.

The occurrence of both slow quasi-geostrophic motions and fast inertia-gravity waves in NWP models indicate the presence of two different time scales in the same system of equations. After an appropriate scaling of the equations of motion, the slow waves have time derivatives which are O(1) times their space derivatives and the fast waves have time derivatives which are O(1/ε) times their space derivatives, where ε is the nondimensional Rossby number (Sec. 3), $0 < \varepsilon \ll 1$. Both the slow and fast time scales are associated with purely oscillatory behavior: no exponential growth or decay occurs. This fact is important in the application of the bounded derivative principle.

4.1 The bounded derivative principle for systems of ODEs

To present the principle in its simplest form, we consider a system of ODEs with two time scales,

$$d\underset{\sim}{y}/dt = A(t)\underset{\sim}{y} + \underset{\sim}{f}(t), \tag{4.1}$$

where $\underset{\sim}{y}$ is an n-vector and A is an n $\times$ n matrix. We assume that A(t) is diagonalizable with purely imaginary eigenvalues $\lambda_1,\ldots,\lambda_n$. We also assume that solutions $\underset{\sim}{z}(t)$ of the homogeneous equation

$$d\underset{\sim}{z}/dt = A(t)\underset{\sim}{z} \tag{4.2}$$

satisfy an estimate of the form

$$|\underset{\sim}{z}(t)| \leq K|\underset{\sim}{z}(0)| \ . \tag{4.3}$$

These assumptions imply that solutions of system (4.1) behave like solutions of a hyperbolic system of PDEs, with motions which are essentially oscillatory. We suppose that the eigenvalues split into two sets

$$M_1 = \{\lambda_1,\ldots,\lambda_m\}$$

$$M_2 = \{\lambda_{m+1},\ldots,\lambda_n\}$$

where $\lambda_j = O(1/\varepsilon)$ if $\lambda_j \in M_1$ and $\lambda_j = O(1)$ if $\lambda_j \in M_2$, with $0 < \varepsilon << 1$. Eigenvalues with two different orders of magnitude yield motions with two time scales: the motions associated with M_1 are fast, i.e., their time derivatives are $O(1/\varepsilon)$, and the motions associated with M_1 are slow, i.e. their time derivatives are $O(1)$. The initialization problem is to choose $y(0)$ so that the motions with the fast time scale stay small in amplitude.

The <u>bounded</u> <u>derivative</u> <u>principle</u> arises from a simple observation: If $\underset{\sim}{y}(t)$ varies slowly, then its first few time derivatives satisfy

$$\frac{d^\nu \underset{\sim}{y}}{dt^\nu} = O(1), \text{ for } \nu = 1,\ldots,p-1 \tag{4.4}$$

and some suitable $p > 0$. In particular, equation (4.4) must hold at time t = 0. This leads to the principle: choose the initial value $\underset{\sim}{y}(0) = \underset{\sim}{y}_0$ so that

$$\frac{d^\nu \underset{\sim}{y}}{dt^\nu} = O(1) \quad \text{at } t = 0 \quad \text{for } \nu = 0,1,\ldots,p-1. \tag{4.5}$$

The result which Kreiss has proved for ODEs and hyperbolic PDEs is that initial values which satisfy the bounded derivative principle generate solutions which vary only on the slow time scale for t in some finite interval $0 \leq t \leq T$, with $T = O(1)$. His results depend on assumptions on the structure of the systems which we will not discuss in detail. Some of these assumptions involve the existence of transformations of the systems into normal forms. Such transformations correspond roughly to the construction of normal modes in nonlinear normal-mode initialization procedures. One important feature of the principle is that these transformations do not have to be carried out in practice, when applying the principle to a given system.

To illustrate the validity of this principle, consider the <u>scalar</u> equation

$$\frac{dy}{dt} = \frac{i}{\varepsilon}\,(y + e^{it}), \tag{4.6}$$

$$y(0) = y_0.$$

The solution of this equation is

$$y(t) = y_1(t) + y_2(t)\ , \tag{4.7a}$$

where

$$y_1(t) = -\frac{1}{1-\varepsilon}\, e^{it} \tag{4.7b}$$

is the slow part of the solution, and

$$y_2(t) = \left(y_0 + \frac{1}{1-\varepsilon}\right) e^{it/\varepsilon} \tag{4.7c}$$

is the fast part of the solution. The derivatives of y are, correspondingly,

$$\frac{d^{\nu}y}{dt^{\nu}} = -i^{\nu}\left(\frac{1}{1-\varepsilon}\right) e^{it} + \left(\frac{i}{\varepsilon}\right)^{\nu}\left(y_0 + \frac{1}{1-\varepsilon}\right) e^{it/\varepsilon} .$$

For each time derivative $d^{\nu}y/dt^{\nu}$, $\nu = 0,1,2,\ldots$, to be $O(1)$, we need

$$y_0 = -\frac{1}{1-\varepsilon} + O(\varepsilon^{\nu}). \tag{4.8}$$

Hence a particular derivative will be dounded, $d^{\nu}y/dt^{\nu} = O(1)$ for all t, <u>if and only if</u> $d^{\nu}y/dt^{\nu} = O(1)$ at $t = 0$.

The homogeneous equation corresponding to equation (4.6) has solutions which vary only on the fast time scale $O(1/\varepsilon)$. For a solution of the homogeneous equation to have p bounded derivatives, its initial value must be $O(\varepsilon^{p})$. Equation (4.6) has really only a fast time scale, so we expect at most one solution of the inhomogeneous equation, to within $O(\varepsilon^{p})$, to have p bounded derivatives. Because of the forcing term, this solution $y_1(t)$ is not zero to within $O(\varepsilon^{p})$; its scale of motion is like a slow time scale. The choice of y_0 in equation (4.8) depends in fact on the forcing term in equation (4.6).

For this simple example, in which the explicit solution (4.7) is known, the initialization problem can be handled completely by setting

$$y_0 = -\frac{1}{1-\varepsilon} .$$

Then all derivatives of y are bounded. In general, the initialization problem is not as straightforward as this. We now illustrate how equation (4.6) can be used, together with the bounded derivative principle, to derive equation (4.8) without using the explicit solution (4.7).

By differentiating equation (4.6), we can derive equations for $d^\nu y/dt^\nu$ in terms of y, e.g.,

$$\frac{d^2y}{dt^2} = \frac{i}{\varepsilon}\left(\frac{dy}{dt} + ie^{it}\right)$$

$$= \frac{i}{\varepsilon}\left(\frac{i}{\varepsilon}(y + e^{it}) + ie^{it}\right). \qquad (4.9)$$

Evaluating (4.9) at t = 0,

$$\left.\frac{d^2y}{dt^2}\right|_{t=0} = -\frac{1}{\varepsilon^2}(y_0 + 1 + \varepsilon) ,$$

we see that $y''(0) = O(1)$ if and only if $y_0 + 1 + \varepsilon = O(\varepsilon^2)$. Differentiating equation (4.9) successively and using equation (4.6) after each differentiation, we obtain

$$\frac{d^\nu y}{dt^\nu} = \left(\frac{i}{\varepsilon}\right)^\nu y + \left(\sum_{j=1}^{\nu} \left(\frac{i}{\varepsilon}\right)^j i^{\nu-j}\right)e^{it}. \qquad (4.10)$$

Evaluating (4.10) at t = 0, we see that $(d^\nu y/dt^\nu)(0) = O(1)$ if and only if

$$y_0 = -\sum_{k=0}^{\nu-1} \varepsilon^j + O(\varepsilon^\nu), \qquad (4.11)$$

which agrees with equation (4.8), being just the power series of $-1/(1-\varepsilon)$.

In systems where there is only a fast time scale, in general only one solution to within $O(\varepsilon^p)$ has p bounded derivatives. An asymptotic expansion in powers of ε for the initial data of this single slow solution can be obtained from the system using the bounded derivative principle as we have just illustrated.

In systems where two time scales are present in the homogeneous equations, the "slow parts" of the initial data can be specified arbitrarily; then the bounded derivative principle can be used to determine an asymptotic expansion for the "fast parts" of the initial data. For example, consider the system

$$\frac{d\underset{\sim}{y}}{dt} = \frac{d}{dt}\begin{bmatrix} \underset{\sim}{y}^{I} \\ \\ \underset{\sim}{y}^{II} \end{bmatrix} = \begin{bmatrix} \frac{1}{\varepsilon} A_{11} & \frac{1}{\varepsilon} A_{12} \\ \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \underset{\sim}{y}^{I} \\ \\ \underset{\sim}{y}^{II} \end{bmatrix} + \begin{bmatrix} \frac{1}{\varepsilon} \underset{\sim}{f}^{I} \\ \\ \underset{\sim}{f}^{II} \end{bmatrix}, \qquad (4.12)$$

where $\underset{\sim}{y}^{I}$ is the fast part, and $\underset{\sim}{y}^{II}$ the slow part of the solution. Under certain assumptions, Kreiss (1979) has shown that for every $y^{II}(0)$, there is a $y^{I}(0)$ so that the solution of system (4.12) with these initial values has p bounded derivatives, and that given $y^{II}(0)$, this $y^{I}(0)$ is unique to within $O(\varepsilon^{p})$.

4.2 The bounded derivative principle in NWP

In the systems of PDEs governing NWP models, it is not clear *a priori* what the "slow and fast parts" of the initial data are. The bounded derivative principle can be used to determine a sequence of diagnostic equations which hold up to a certain order in ε. Initial data which satisfy these diagnostic equations will generate solutions of the model equations which vary on the slow time scale up to a time which is O(1). The question of how initial data which satisfy these diagnostic equations are constructed from asynoptic measurements does not have a simple answer; each pair of model and nonstandard data sets must be analyzed separately.

We conclude this section by presenting a part of the application of this principle to the shallow-water equations with orography H = H(x,y); this work is due to Browning, Kasahara, and Kreiss (1979). The equations are

$$\frac{du}{dt} + \phi_x - fv = 0 \, , \qquad (4.13a)$$

$$\frac{dv}{dt} + \phi_y + fu = 0 \, , \qquad (4.13b)$$

$$\frac{\partial\phi}{\partial t} + (u\phi)_x + (v\phi)_y + (\phi_0 - \Phi)\delta - (u\Phi_x + v\Phi_y) = 0 \ , \qquad (4.13c)$$

where $d/dt = \partial/\partial t + u\,\partial/\partial x + v\,\partial/\partial y$, ϕ_0 = mean geopotential, ϕ = deviation in geopotential from the mean, $\delta = u_x + v_y$ = divergence, $\zeta = -u_y + v_x$ = vorticity, and $\Phi = gH$. The nondimensionalized and scaled equations for the case of a midlatitude beta-plane become

$$\frac{du}{dt} + a = 0 \ , \qquad (4.14a)$$

$$\frac{dv}{dt} + b = 0 \ , \qquad (4.14b)$$

$$\frac{\partial\phi}{\partial t} + (u\phi)_x + (v\phi)_y + c = 0 \ , \qquad (4.14c)$$

where

$$a = \varepsilon^{-1}(\phi_x - fv) \ , \qquad (4.15a)$$

$$b = \varepsilon^{-1}(\phi_y + fu) \ , \qquad (4.15b)$$

$$c = \varepsilon^{-2}(\phi_0 - \varepsilon\Phi)\delta - \varepsilon^{-1}(u\Phi_x + v\Phi_y) , \qquad (4.15c)$$

and $f = f_0 + \varepsilon\beta y$, ε = Rossby number = $O(1/10)$. When we say in the sequel that a function is $O(1)$, we mean that the function and its spatial derivatives are $O(1)$.

The first-order time derivatives u_t, v_t, and ϕ_t are $O(1)$ if and only if a, b, and c are $O(1)$. Differentiating (4.15a) with respect to x and (4.15b) with respect to y, this requires

$$a_x + b_y = \varepsilon^{-1}\left(\nabla^2\phi - f\zeta + \varepsilon\beta u\right) = O(1) \ , \qquad (4.16a)$$

$$c = \varepsilon^{-2}\left(\phi_0\delta - \varepsilon\Phi\delta - \varepsilon(u\Phi_x + v\Phi_y)\right) = O(1) \ , \qquad (4.16b)$$

i.e.,

$$\nabla^2\phi - f\zeta + \varepsilon\beta u = O(\varepsilon) , \tag{4.17a}$$

$$\delta = \frac{\varepsilon(u\Phi_x + v\Phi_y)}{\phi_0 - \varepsilon\Phi} + O(\varepsilon^2) . \tag{4.17b}$$

Dropping terms $O(\varepsilon)$ in (4.17a) and $O(\varepsilon^2)$ in (4.17b), we obtain the diagnostic equations

$$\nabla^2\phi - f\zeta = 0 , \tag{4.18a}$$

$$\delta = \varepsilon\left(u\Phi_x + v\Phi_y\right)/\phi_0 . \tag{4.18b}$$

Equation (4.18b) implies that $\delta = O(\varepsilon)$, so the divergent part of the wind is $O(\varepsilon)$. Hence it is permissible to replace u and v by their rotational parts u^0 and v^0 in equation (4.18b), making an error of $O(\varepsilon^2)$. If the rotational part of the wind is measured, which is equivalent to the vorticity ζ being measured, then δ can be determined by (4.18b) and ϕ by (4.18a). Then u and v can be determined from δ and ζ by the Helmholtz theorem.

Notice that equation (4.18a) is the linear balance equation. The numerical experiments of Browning et al. (1979) indicate that requiring just the first time derivatives of u, v, ϕ to be O(1) at time t = 0 does not effectively prevent the growth of inertia-gravity waves. The improved diagnostic relations obtained by requiring that the second time derivatives of u, v, ϕ also be O(1) at time t = 0 do effectively prevent the growth of the first waves. These results corroborate those obtained in the practical implementation of nonlinear normal-mode initialization schemes (Daley, 1980).

We will not present the second-order relations here. In particular, the nonlinear balance equation can be derived from these second-order relations (compare also Sec. 3, and Leith (1980)). The reader is referred to Browning et al. (1979) for details.

5. CONCLUDING REMARKS

We have pointed out that there are two problems in passing from a set of measured data, distributed in space and time, to a solution of the governing equations for an NWP model: one is completeness, the other is, for lack of a better term, accuracy. We call a set of measured synoptic and asynoptic data complete if it determines uniquely the solution to the model equations, and if small errors in the data lead only to small errors in this solution. Criteria were outlined for the theoretical completeness of data sets considering some very simple model equations.

Given a complete data set, we studied the practical construction of the unique model solution by intermittent updating. The convergence of updating procedures was investigated, and its dependence on the updating interval was stressed.

The second problem, that of accuracy, is linked to the presence of two time scales, a slow and a fast one, in the model equations of NWP. For a model system, the shallow-water equations, we showed that the geopotential field and its first two time derivatives, or tendencies, form a complete set of data. The velocity field can be constructed by solving a set of diagnostic equations which use the mass field data. It was shown that the initial state thus constructed generates a slow solution to the prognostic model equations, provided the tendencies of the mass field at initial time are small. Hence this solution of the completeness problem also solves the initialization problem, that of ascertaining that the evolution of the model is on the slow time scale only.

It remained to show that, given small initial tendencies of the solution, the solution will continue to evolve slowly for a time interval comparable to the model's slow time scale. This is in fact the case, as shown by the bounded derivative principle. This principle was first illustrated for simple systems of ODEs with two time scales. It was then applied to the shallow-water equations.

In conclusion, we have addressed certain theoretical questions of 4-D data assimilation and initialization: completeness of data sets, convergence of intermittent updating, the time span of slow evolution of NWP model solutions. The questions and their solution were illustrated for the shallow-water equations.

We hope to have shown the close relationship between data assimilation and initialization. A unified treatment of both aspects of the data problem in NWP appears in the companion paper, Ghil *et al.* (1980).

Acknowledgements

It is a pleasure to acknowledge useful discussions and correspondence with F. Baer, L. Bengtsson, M. Halem, E. Isaacson,

E. Källén, H.-O. Kreiss, C. E. Leith, J. Oliger and O. Talagrand. Work on this review article was supported by the National Aeronautics and Space Administration under grants NSG-5034 (K. B.) and NSG-5130 (M. G.), monitored by the Goddard Laboratory for Atmospheric Sciences.

References

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

Bengtsson, L., 1975. 4-Dimensional Assimilation of Meteorological Observations. GARP Publications Series, No. 15, World Meteorological Organization, Geneva, 76 pp.

Browning, G., A. Kasahara, and H.-O. Kreiss, 1979. Initialization of the primitive equations by the bounded derivative method. NCAR Ms. 0501-79-4, National Center for Atmospheric Research, Boulder, CO 80307.

Bube, K. P., 1978. The construction of initial data for hyperbolic systems from nonstandard data. Ph.D. Dissertation, Department of Mathematics, Stanford University, Stanford, CA 94305, 114 pp.

__________, 1980. Determining solutions of hyperbolic systems from incomplete data. Comm. Pure Appl. Math., submitted.

__________ and J. Oliger, 1977. Hyperbolic partial differential equations with nonstandard data, in Advances in Computer Methods for Partial Differential Equations, II, R. Vichnevetsky (ed.), IMACS, Rutgers Univesity, New Brunswick, NJ 08903, pp. 256-263.

Charney, J., M. Halem, and R. Jastrow, 1969. Use of incomplete historical data to infer the present state of the atmosphere. J. Atmos. Sci., 26, 1160-1163.

Daley, R., 1980. Normal mode initialization, this volume.

Ghil, M., 1975. Initialization by compatible balancing, Report No. 75-16, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23665, 38 pp.

________, 1980. The compatible balancing approach to initialization, and four-dimensional data assimilation. Tellus, 32, 198-206.

________, and R. Balgovind, 1979. A fast Cauchy-Riemann solver. Math. Comput., 33, 585-635.

________, B. Shkoller and V. Yangarber, 1977. A balanced diagnostic system compatible with a barotropic prognostic model, Mon. Wea. Rev., 105, 1223-1238.

________, S. Cohn, J. Tavantzis, K. Bube and E. Isaacson, 1980. Applications of estimation theory to numerical weather prediction, this volume.

Kreiss, H.-O., 1978. Problems with different time scales. In Recent Advances in Numerical Analysis, C. de Boor and G. H. Golub (eds.), Academic Press, New York, pp. 95-106.

________, 1979. Problems with different time scales for ordinary differential equations. SIAM J. Numer. Anal. 16, 980-998.

________, 1980. Problems with different time scales for partial differential equations. Comm. Pure Appl. Math. 33, 399-439.

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

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.

Smagorinsky, J., K. Miyakoda and R. F. Strickler, 1970. The relative importance of variables in initial conditions for dynamical weather prediction. Tellus 22, 141-157.

Talagrand, O., 1980. A study of the dynamics of four-dimensional data assimilation. Tellus, to appear.