Stability of a Two-Time-Level Semi-Implicit Integration Scheme
for Gravity Wave Motion

A. J. SIMMONS AND C. TEMPERTON

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, United Kingdom

(Manuscript received 26 March 1996, in final form 15 July 1996)

ABSTRACT

A study is made of the computational stability of semi-implicit treatments of gravity wave motion suitable for use with two-time-level advection schemes. The analysis is for horizontally uniform reference values of temperature and surface pressure, and for hybrid pressure-based vertical coordinates. Stability requires the use of reference temperatures that are warmer than those that can be used safely with the corresponding three-time-level scheme. The reference surface pressure should also be higher. When stable, the two-time-level scheme is damping, although the largest scales are damped less than by the three-time-level scheme if the latter uses a typical time filtering. The first-order decentered averaging of gravity wave tendencies used in a number of semi-Lagrangian models reduces the need for a relatively warm reference temperature profile but causes a quite substantial damping of otherwise well-represented low-wavenumber modes. The low-wavenumber damping can be avoided by using an alternative, second-order averaging involving a third (past) time level. For this alternative averaging, an economical spatial discretization is proposed that requires no additional departure point. Phase speeds show little sensitivity to these changes in formulation. All variants of the semi-implicit method substantially reduce the phase speeds of the fastest high-wavenumber modes when use is made of the large time steps possible with semi-Lagrangian advection.

1. Introduction

The introduction of the semi-implicit method of integrating the primitive equations forward in time was a landmark in the development of numerical models of atmospheric flow. As developed for sigma-coordinate baroclinic models by Robert et al. (1972), the method involves a separation of gravity wave terms linearized about a reference temperature profile and an averaging in time of these terms to enable use of a time step substantially longer than would be possible using a fully explicit scheme. The semi-implicit scheme became widely used in the years following Robert’s pioneering work, but as the vertical resolution of models was increased beyond a few layers it became clear that care had to be taken in the choice of reference temperature profile (Simmons et al. 1978). When the scheme was applied to models with hybrid vertical coordinates that changed from terrain-following at low levels to pressure at upper levels, a careful choice had also to be made of a reference surface pressure, and a smooth transition to the pressure coordinate was found to enhance stability (Simmons and Burridge 1981; Simmons and Strüning 1983). A more extensive analysis of the stability of the semi-implicit time scheme was reported by Côté et al. (1983), and further analysis was given by Simmons et al. (1989).

The early implementations of the semi-implicit method were generally in three-time-level integration schemes using an Eulerian treatment of advection. A further landmark in the development of modeling was the adoption of the semi-Lagrangian method for advection, also pioneered by Robert (1981). This method was first implemented in three-time-level schemes and was found to work well when used in conjunction with a semi-implicit treatment of gravity wave terms, despite the use of longer time steps that further tested the stability of the semi-implicit method. The European Centre for Medium-Range Weather Forecasts (ECMWF) currently uses a three-time-level, semi-implicit, semi-Lagrangian scheme for operational forecasting, using a spectral model with T213 horizontal resolution and a vertical resolution of 31 layers (Ritchie et al. 1995).

The semi-Lagrangian method is also well suited to use in two-time-level integration schemes (Temperton and Staniforth 1987). However, in some applications of the two-time-level approach it has been found necessary to implement a damping, first-order decentering in the semi-implicit averaging (McDonald and Haugen 1992, 1993; Bates et al. 1993). Moreover, early tests of a two-time-level version of the ECMWF model exhibited widespread noise in the subtropics. In these integrations,
2. The stability problem

We extend the analyses given by Simmons and Burridge (1981) and Simmons et al. (1989) for three-time-level schemes using hybrid vertical coordinates. The divergence, temperature, and surface pressure equations for small-amplitude gravity wave motion are written as follows:

\[
\frac{\partial \mathbf{D}}{\partial t} = \frac{n(n+1)}{a^2} \left[ (\gamma \mathbf{T}' + \delta(\ln p)') \right] \quad (1)
\]
\[
\frac{\partial \mathbf{T}'}{\partial t} = -\tau \mathbf{D} \quad (2)
\]
\[
\frac{\partial (\ln p)'}{\partial t} = -\nu \mathbf{D}. \quad (3)
\]

Here \( \mathbf{D} \) and \( \mathbf{T}' \) are column vectors representing the NLEV values of a particular spectral component of the divergence and of the deviation of temperature from the horizontally uniform temperature of the resting basic state; \((\ln p)'\) is the corresponding spectral component of the deviation of the logarithm of surface pressure from its uniform basic-state value. NLEV denotes the number of model levels; \( n \) is the total wavenumber of the spectral component in question, and \( a \) is the radius of the earth; \( \gamma \) and \( \tau \) are NLEV \( \times \) NLEV matrices, and \( \delta \) and \( \nu \) are column and row vectors, respectively, each depending on the temperature and surface pressure of the basic state. For \( \gamma, \tau, \delta, \) and \( \nu \) are given by Simmons and Burridge (1981).

We assume a reference temperature profile and reference surface pressure, from which reference matrices and vectors \( \gamma, \tau, \delta, \) and \( \nu \) can be calculated. We test the stability of gravity wave perturbations to an idealized resting atmosphere characterized by matrices and vectors \( \gamma, \tau, \delta, \) and \( \nu \). We use a two-time-level semi-implicit scheme in which, when stepping from time level \( t \) to time level \( t + \Delta t \), terms on the right-hand sides of (1)–(3) that appear as gravity wave perturbations of the reference basic state are computed as time averages of terms at times \( t \), \( t + \Delta t \). The possibility of decentering, with parameter \( \epsilon \), is included in this time average (Tanguay et al. 1992; McDonald and Haugen 1992; Bates et al. 1993). A term linear in variable \( \mathbf{X} \) is evaluated using the value

\[
\frac{1}{2} \{(1 + \epsilon)\mathbf{X}(t + \Delta t) + (1 - \epsilon)\mathbf{X}(t)\}
\]

for \( \mathbf{X} \); \( \epsilon = 0 \) corresponds to the usual centered semi-implicit scheme. Remaining terms on the right-hand sides of (1)–(3) are computed by extrapolation of values at times \( t \) and \( t - \Delta t \). For these terms, the value

\[
\frac{3}{2} \mathbf{X}(t) - \frac{1}{2} \mathbf{X}(t - \Delta t)
\]

is used for the variable \( \mathbf{X} \). Equations (1)–(3) thus become

\[
\frac{\mathbf{D}(t + \Delta t) - \mathbf{D}(t)}{\Delta t} = \frac{n(n+1)}{2a^2} \left\{ \gamma\{[(1 + \epsilon)\mathbf{T}(t + \Delta t) + (1 - \epsilon)\mathbf{T}(t)] + \delta\{[(1 + \epsilon)\mathbf{l}(t + \Delta t) + (1 - \epsilon)\mathbf{l}(t)]
\right.
\]

\[
+ (\gamma - \gamma\{[3\mathbf{T}(t) - \mathbf{T}(t - \Delta t)] + (\delta - \delta\{[3\mathbf{l}(t) - \mathbf{l}(t - \Delta t)]\}
\}
\]

\[
+ (\gamma - \gamma\{[3\mathbf{T}(t) - \mathbf{T}(t - \Delta t)] + (\delta - \delta\{[3\mathbf{l}(t) - \mathbf{l}(t - \Delta t)]\}
\}
\]

\[
\frac{\mathbf{T}(t + \Delta t) - \mathbf{T}(t)}{\Delta t} = -\frac{1}{2} \{\tau\{[(1 + \epsilon)\mathbf{D}(t + \Delta t) + (1 - \epsilon)\mathbf{D}(t)] + (\tau - \tau\{[3\mathbf{D}(t) - \mathbf{D}(t - \Delta t)]\}
\}
\]

\[
+ (\nu - \nu\{[3\mathbf{D}(t) - \mathbf{D}(t - \Delta t)]\}
\}
\]

\[
\frac{\mathbf{l}(t + \Delta t) - \mathbf{l}(t)}{\Delta t} = -\frac{1}{2} \{\nu\{[(1 + \epsilon)\mathbf{D}(t + \Delta t) + (1 - \epsilon)\mathbf{D}(t)] + (\nu - \nu\{[3\mathbf{D}(t) - \mathbf{D}(t - \Delta t)]\}
\}
\]

\[
+ (\nu - \nu\{[3\mathbf{D}(t) - \mathbf{D}(t - \Delta t)]\}
\}
\]

Here, we have dropped primes and replaced \( \ln p \), by \( l \).

We look for modes for which

\[
\mathbf{X}(t + \Delta t) = \lambda\mathbf{X}(t), \quad (7)
\]

where \( \mathbf{X} \) denotes \( \mathbf{D}, \mathbf{T}, \) or \( l \). In general, \( \lambda \) will be complex. We shall mostly present results for parameters for which the numerical scheme is computationally stable, that is, for which \( |\lambda| \leq 1 \). We refer to \( |\lambda| \) as the damping factor of the scheme, the factor by which gravity wave amplitudes are reduced each time step. Instability occurs if \( |\lambda| > 1 \), in which case the “damping” factor actually represents the factor by which amplitudes increase over a time step. Writing \( \lambda \) in the form \( \lambda = \exp(i\theta) \), the time dependence of the mode can be expressed in terms of a phase speed \( c_p \), given by

\[
a \theta/[\ln(n+1)]^{1/2}\Delta t \]

and an exponential damping rate \( \chi \), given by

\[-\ln|\lambda|/\Delta t\].
Substituting (7) into (4), (5), and (6) gives

\[
(\lambda^2 - \lambda)\mathbf{D}(t - \Delta t) = \frac{n(n + 1)\Delta t}{2a^2} \left\{ (\gamma, (1 + e)\lambda^2 + (1 - e)\lambda) + (3\lambda - 1)(\gamma - \gamma_a) \right\} \mathbf{T}(t - \Delta t) \\
+ \left\{ \delta, (1 + e)\lambda^2 + (1 - e)\lambda \right\} l(t - \Delta t)
\]

Eliminating \( \mathbf{T}(t - \Delta t) \) and \( l(t - \Delta t) \), we obtain

\[
(\lambda^2 - \lambda)^2 \mathbf{D} = \frac{n(n + 1)\Delta t^2}{4a^2} \left\{ ([1 + e]\lambda^2 + (1 - e)\lambda]^{2} (\gamma, \tau, + \delta, \nu) \\
+ [(1 + e)\lambda^2 + (1 - e)\lambda](3\lambda - 1) \\
\times [\gamma, (\tau - \tau_a) + \delta, (\nu - \nu_a) + (\gamma - \gamma_a)\tau_a + (\delta - \delta_a)\nu_a] \\
+ (3\lambda - 1)^2[(\gamma - \gamma_a)(\tau - \tau_a) + (\delta - \delta_a)(\nu - \nu_a)] \right\} \mathbf{D}.
\]

This equation can be put in the standard form

\[
(1 - \lambda^4 - M_1 \lambda^3 - M_2 \lambda^2 - M_3 \lambda - M_4) \mathbf{D} = 0,
\]

where

\[
M_1 = -[1 + (1 + e)^2d_s\mathbf{B}]^{-1}\{-2\mathbf{l} \\
+ d_s[2(1 - e)^2\mathbf{B} + 3(1 + e)\mathbf{E}_1]]
\]

\[
M_2 = -[1 + (1 + e)^2d_s\mathbf{B}]^{-1}[1 + d_s[(1 - e)^2\mathbf{B} \\
+ (2 - 4e)\mathbf{E}_1 + 9\mathbf{E}_2]]
\]

\[
M_3 = [1 + (1 + e)^2d_s\mathbf{B}]^{-1}[1 + d_s[(1 - e)\mathbf{E}_1 + 6\mathbf{E}_2]]
\]

\[
M_4 = -[1 + (1 + e)^2d_s\mathbf{B}]^{-1}d_s\mathbf{E}_2,
\]

with

\[
\mathbf{B} = \gamma, \tau, + \delta, \nu, \\
\mathbf{E}_1 = \gamma, (\tau - \tau_a) + \delta, (\nu - \nu_a) + (\gamma - \gamma_a)\tau_a \\
+ (\delta - \delta_a)\nu_a, \\
\mathbf{E}_2 = (\gamma - \gamma_a)(\tau - \tau_a) + (\delta - \delta_a)(\nu - \nu_a)
\]

and

\[
d_n = \frac{n(n + 1)\Delta t^2}{4a^2}.
\]

As for the corresponding problem for the three-time-level scheme (Simmons et al. 1989), the eigenvalue \( \lambda \) can be readily determined numerically as an eigenvalue of the \( 4 \text{NLEV} \times 4 \text{NLEV} \) matrix:

\[
\begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
M_4 & M_3 & M_2 & M_1
\end{bmatrix}.
\]

The eigenvectors of this matrix take the special form

\[
\begin{bmatrix}
\mathbf{D} \\
\lambda \mathbf{D} \\
\lambda^2 \mathbf{D} \\
\lambda^3 \mathbf{D}
\end{bmatrix},
\]

where the column vector \( \mathbf{D} \) of dimension NLEV is the divergence structure of the mode.

It is straightforward to extend the analysis to include a simple representation of horizontal diffusion. We assume diffusion to be applied through an implicit scheme in which provisional values of variables at time \( t + \Delta t \) computed using the semi-implicit treatment of gravity wave terms are then diffused by division by a factor \((1 + k_n)\). This is the approach adopted in the ECMWF model (Ritchie et al. 1995). For the two-time-level scheme and fourth-order diffusion of form \( k_n = \frac{K\Delta m^2(n + 1)^2}{a^4} \).

If the scheme is applied to each of the prognostic variables \( \mathbf{D}, \mathbf{T}, \) and \( l \), then (8)–(10) are simply modified by multiplying the terms involving \( \lambda^2 \) by \((1 + k_n)\). The stability problem again reduces to an eigenvalue problem of form (11), where now
and 1 a reference profile by a multiplicative factor (1 + e). Dependence on reference temperature

3. Special cases

a. Dependence on reference temperature

We consider the case of an exact reference pressure, and an actual temperature profile that differs from the reference profile by a multiplicative factor (1 + \(\alpha\)), where \(\alpha\) is a constant. Then \(\gamma = \gamma_0, \nu = \nu_0, \tau = (1 + \alpha)\tau_0\), and \(\delta = (1 + \alpha)\delta_0\). It follows that \(E_1 = \alpha B, E_2 = 0\), and \(M_1 = 0\). The modal divergence structures derived from (11) are in this case the same as the eigenfunctions of \(B\), and we consider one of the NLEV modes, denoting the corresponding eigenvalue of \(B\) by \(c^2\) (with \(c\) real corresponding to a stably stratified reference state). In the absence of decentering (\(e = 0\)) and horizontal diffusion (\(k_n = 0\)), the corresponding values of \(\lambda\) satisfy the cubic equation

\[
\lambda^3 - m_1 \lambda^2 - m_2 \lambda - m_3 = 0,
\]

where

\[
m_1 = \frac{2(1 - \beta^2)}{1 + \beta^2} - 3\mu
\]

\[
m_2 = -1 - 2\mu
\]

\[
m_3 = \mu,
\]

with

\[
\beta^2 = d_n c^2 = \frac{n(n + 1)c^2\Delta t^2}{4d^2}
\]

and

\[
\mu = \frac{\alpha\beta^2}{1 + \beta^2}.
\]

One solution of (21) is necessarily a real number and corresponds to a computational mode. The other two solutions form a complex conjugate pair, which represent eastward- and westward-moving gravity wave modes.

For \(\mu = 0\), the computational mode vanishes and the eigenvalues for the eastward- and westward-moving modes are given by the usual formula for a semi-implicit scheme with exact reference state:

\[
\lambda_0 = \frac{1 - \beta^2}{1 + \beta^2} \pm i \frac{2\beta}{1 + \beta^2} = \frac{1 \pm i\beta}{1 + \beta^2}.
\]

This is an approximation to the solution \(\exp(\pm i2\beta\Delta t)\) that holds in the absence of time truncation error. It is easily shown that \(|\lambda_0| = 1\), indicating the well-known stability of the basic scheme.

For small \(\mu\) there is a computational mode, but it has \(\lambda = \mu + O(\mu^2)\) and is rapidly damped. For the other modes, we seek a solution of the form

\[
\lambda = \lambda_0 + \lambda_1 \mu + O(\mu^2).
\]

Some algebra yields

\[
\lambda_1 = \frac{1 - 3\lambda_0}{\lambda_0 - 1}.
\]

The square of the damping factor is given by

\[
|\lambda|^2 = \lambda\lambda^* = \lambda_0\lambda_0^* + 2 \Re(\lambda_1\lambda_1^*)\mu + O(\mu^2),
\]

where the asterisk denotes the complex conjugate. Substituting (23) and (25) into (26) and simplifying, we obtain

\[
|\lambda|^2 = 1 + \frac{4\beta^2}{1 + \beta^2} \mu + O(\mu^2).
\]

Thus, \(|\lambda| > 1\) for \(\mu > 0\) and \(|\lambda| < 1\) for \(\mu < 0\). The semi-implicit scheme is unconditionally unstable if the reference temperature profile is slightly colder than the actual profile and unconditionally stable if the reference profile is slightly warmer than the actual profile. In the limit of large \(\Delta t, \beta \to \infty\) and \(\lambda_0 \to -1\). The damping factor \(|\lambda|\) is then approximately \(1 + 2\alpha\).

Equation (27) implies that, to ensure stability, the reference temperature profile should be chosen warmer than any likely actual temperature profile, so that \(\mu < \)
0 is automatically satisfied. It is instructive to compare this result with the analysis presented by Gravel et al. (1993). They too investigated the stability of a two-time-level scheme with extrapolated nonlinear terms, but in their numerical examples they considered only the case \( \mu > 0 \) and were therefore forced to invoke other damping mechanisms (such as decentering) in order to achieve stability.

It is of interest to compare the damping rate obtained with warm reference temperatures with that from introducing decentering with positive \( \epsilon \). In the latter case, if the reference temperature profile is exact and there is no horizontal diffusion,

\[
\lambda = \frac{1 - \beta^2(1 - \epsilon^2)}{1 + \beta^2(1 + \epsilon)^2} \pm i \frac{2\beta}{1 + \beta^2(1 + \epsilon)^2}, \tag{28}
\]

a result reported previously by Tanguay et al. (1992) and Bates et al. (1993). The square of the damping factor is given by

\[
|\lambda|^2 = \frac{1 + \beta^2(1 - \epsilon)^2}{1 + \beta^2(1 + \epsilon)^2},
\]

which, for comparison with (27), may be written as

\[
|\lambda|^2 = 1 - \frac{4\beta^2}{1 + \beta^2} \epsilon + O(\epsilon^2). \tag{29}
\]

In the limit of large \( \Delta t \), \( \beta \to \infty \), and the damping factor \( |\lambda| \) is approximately \( 1 - 2\epsilon \) for small \( \epsilon \). Comparison of (27) and (29) indicates that a decentering parameter of 0.1 gives damping of the same order as a 10% warm deviation of the reference temperature profile from the actual profile.

When horizontal diffusion of the form specified in the preceding section is included, there is an additional damping factor \((1 + k_n)^{-1}\) for an exact reference state; that is, the net damping factor is \( \lambda(1 + k_n)^{-1} \), where \( \lambda \) is given by (28).

### b. Dependence on reference surface pressure

Here, we consider the case of an exact reference temperature profile and assume that the actual surface pressure differs from the reference surface pressure by a multiplicative factor \((1 + \alpha)\). We shall discuss the case of small \( \alpha \) and assume no decentering or horizontal diffusion. Equations (12)–(15) become

\[
M_i = -(1 + d_\alpha B)^{-1}[-2I + d_\alpha(2B + 3\alpha F_i)] + O(\alpha^3), \tag{30}
\]

\[
M_2 = -(1 + d_\alpha B)^{-1}[(1 + d_\alpha B + 2\alpha F_1)] + O(\alpha^2), \tag{31}
\]

\[
M_3 = (1 + d_\alpha B)^{-1} d_\alpha F_1 + O(\alpha^2), \tag{32}
\]

\[
M_4 = -(1 + d_\alpha B)^{-1} d_\alpha F_2 + O(\alpha^3), \tag{33}
\]

where \( F_1 \) and \( F_2 \) are matrices that depend only on the reference state.

Inspection of (11) and (30)–(33) shows that in general there will be \( 2 \times NLEV \) rapidly damped computational modes with

\[
\lambda = \lambda_1 \alpha + O(\alpha^2),
\]

where \( \lambda_1 \) is an eigenvalue of

\[
[d_\alpha^2(1 + d_\beta B)\lambda^2 - F_1\lambda + F_2]D = 0.
\]

The remaining \( 2 \times NLEV \) modes represent the eastward- and westward-moving gravity wave modes. For these, \( \lambda = \lambda_0 + \lambda_1 \alpha + O(\alpha^2) \), where \( \lambda_0 \) is given by (23) for one of the NLEV pairs of modes. To obtain the solution to \( O(\alpha) \), we denote by \( D^{(1)} \) the eigenfunction of \( B \) corresponding to the solution \( \lambda_0 \) and expand \( F_1, D^{(1)} \) in terms of the full set of NLEV eigenfunctions, \( D^{(j)}, j = 1, 2, . . ., NLEV: \)

\[
F_1 D^{(j)} = c^2 \sum_{j=1}^{NLEV} f_j D^{(j)}.
\]

The component \( f_{11} \) determines \( \lambda_1 \), while the other NLEV–1 \( f_{ij} \) determine the \( O(\alpha) \) modification to the lowest-order mode structure \( D^{(1)} \), which involves terms in the other modes \( D^{(j)} \) for \( j = 2, 3, . . ., NLEV \). With \( \mu \) denoting the scaled value of \( \alpha \) as in (22), the evaluation of \( \lambda_1 \) proceeds as for the case considered in the preceding section, and we find

\[
|\lambda|^2 = 1 + \frac{4\beta^2}{1 + \beta^2} f_{11} \mu + O(\mu^2). \tag{34}
\]

The stability of the scheme thus depends on whether \( f_{11} \) is positive or negative. If \( f_{11} \) is positive, the semisimple scheme is unconditionally unstable for reference surface pressures slightly lower than the actual surface pressure and unconditionally stable for reference pressures slightly higher than the actual value. This is consistent with what is found in the numerical solutions presented below.

### 4. Numerical solutions

Solutions of (11) have been determined numerically. Calculation of the matrices and vectors \( \gamma, \tau, \delta, \nu, \gamma', \tau', \delta', \) and \( \nu \), was made using subroutines extracted from the ECMWF model. Three idealized actual temperature profiles were used to represent a range of atmospheric conditions. These profiles are plotted as functions of pressure in Fig. 1. Profile 1 is chosen as representative of a midlatitude profile. Profile 2 is a “tropical” profile with an extremely warm surface temperature, and profile 3 is a cold “polar” profile. All results presented here are for the hybrid coordinate and 31-layer vertical resolution currently used operationally at ECMWF (Simmons 1991). The locations of full model levels when the surface pressure is 1013.2 hPa are indicated on the right-hand axis of Fig. 1. Except where stated otherwise, results shown in subsequent figures are for the smallest resolved wavenumber, \( n = 213 \), of the operational
ECMWF model, with no decentering or horizontal diffusion. We present results for the fastest moving non-computational mode. For each case for which this mode is stable, we have checked that no other mode is unstable.

Results are presented for time steps of 15, 20, and 30 min. The operational ECMWF (three-time-level, semi-Lagrangian, T213) model currently uses a 15-min time step; a 30-min time step has been used successfully in tests of the two-time-level scheme in the full T213 model. In the absence of horizontal diffusion, time step and wavenumber appear in the stability problem only in the combination \( n(n + 1) \Delta t^2 \). The results shown for \( n = 213 \) and 15- and 30-min time steps are thus very similar to those obtained for \( n = 106 \) with 30- and 60-min steps, apart from a reduction in damping rates in the latter case due to the \((\Delta t)^{-1}\) term in the expression relating the exponential damping rate \( \chi \) to the damping factor \( \beta \).

**a. Dependence on surface pressure**

Figure 2 shows the dependence of phase speed and damping rate on actual surface pressure for a reference surface pressure of 800 hPa. The reference temperature profile is set to be the same as the actual profile, taken to be profile 1. The reference pressure of 800 hPa is the default value used in the three-time-level semi-implicit scheme, based on the conclusions of stability analyses reported by Simmons and Burridge (1981) and Simmons et al. (1989). It was integrations of the full two-time-level model with this value of the reference pressure that became noisy and prompted the stability analyses presented here.

These numerical calculations confirm the preceding analysis. The damping factor has the value one when the actual surface pressure equals the reference pressure of 800 hPa. Instability (negative damping rate) occurs for surface pressures higher than the reference value, and there is stability, with damped modes, for lower surface pressures. Longer time steps give lower growth rates in the unstable case and lower damping rates in the stable case.

The dependence of damping (or growth) rate on time step can be seen from (22) and (34). The damping factor depends on time step through the term \( \beta^2(1 + \beta^2)^{-1} \), which appears in (34) both explicitly and implicitly through the definition (22) of \( \mu; \beta^2 = n(n + 1)e^2\Delta t^2(4 \Delta x)^{-1} \). The fastest gravity wave speed \( c \) varies from 313 to 302 m s\(^{-1}\) for the range of surface pressures illustrated, giving values of \( \beta \) in the range from 20.7 to 22.2 for a time step of 15 min. For the time steps illustrated, the term \( \beta^2(1 + \beta^2)^{-1} \) is close to its limiting (large time step) value of one. The damping (or growth) per time step thus depends only weakly on time step, and the damping (or growth) rate is thus approximately proportional to \((\Delta t)^{-1}\), as illustrated in Fig. 2. For small time steps the rate is \( O(\Delta t^2) \), but this time step depen-
Figure 3 shows results of a set of calculations similar to those shown in Fig. 2, except that the reference surface pressure is 1013.2 hPa rather than 800 hPa. The numerical integration scheme is now stable for all surface pressures in the range illustrated, from 1013.2 to 500 hPa. For pressures above about 650 hPa results are similar to those in the range 800–550 hPa shown in Fig. 2. For lower surface pressures, damping rates increase sharply and phase speeds decrease slightly, particularly for longer time steps.

Figure 4 shows results corresponding to those in Fig. 3 except that in this case the reference temperature profile is an isothermal 300 K, rather than equal to the actual temperature profile. Comparison of Figs. 3 and 4 shows that the introduction of the isothermal reference profile causes a small additional slowing of phase speeds and a substantial further damping of wave amplitude. The slowing of phase speeds is a characteristic of corresponding calculations for the three-time-level scheme, but the latter only damps modes if (as is usual in practice) time filtering and horizontal diffusion are included.
A comparison of two- and three-time-level results will be presented in section 4c.

b. Dependence on temperature profile

Figure 5 shows phase speeds and damping rates as functions of the reference temperature for an isothermal reference state. The actual temperature is profile 1 as before, and the actual and reference surface pressures are both taken to be 1013.2 hPa. The slowing of phase speeds increases with increasing reference temperature, more so at warmer temperatures. Damping rates increase as the reference temperature increases, at least for temperatures up to around 330–350 K. The two-time-level scheme becomes unstable for sufficiently cold temperatures, with a somewhat more stringent stability limit for longer time steps.

Corresponding results for the warm actual profile, profile 2, are shown in Fig. 6. Phase speeds are generally rather higher than for profile 1 and show less sensitivity to the reference temperature. Damping rates show the same qualitative behavior as for profile 1, but curves are shifted to the right. The isothermal reference temperature has to be increased by about 30 K to produce a damping for profile 2 that is similar to that for profile 1. This shift in temperature is of the same order as the difference in surface temperature between profile 2 and profile 1. For the warm profile, the two-time-level scheme is unstable for reference temperatures of about 300 K or lower.

Figure 7 shows the corresponding results for profile 3, the cold profile. Phase speeds again decrease with increasing reference temperature, but the rate of decrease is largest for a range of temperatures colder than for profile 1. For profile 3 there is a considerable damping even for the lowest reference temperature plotted. However, the increase in damping with increasing reference temperature stops at around 290–310 K for this cold profile, and the damping decreases slightly as reference temperatures are increased beyond these values. A decrease in reference temperature of the order of 40 K is needed to produce a damping factor for profile 3 that is close to what is found for profile 1. Again, this change in reference temperature is of the order of the difference in surface temperature between the two profiles.

c. Comparison of two- and three-time-level schemes

A comparison of results from two- and three-time-level schemes is shown in Fig. 8. Here phase speeds...
and damping rates are plotted as functions of wavenumber. The actual temperature is given by profile 1, and the actual surface pressure is 1013.2 hPa. An isothermal reference temperature of 300 K is used for both schemes. The reference surface pressure is 1013.2 hPa for the two-time-level scheme and 800 hPa for the three-time-level scheme. The calculations for the three-time-level scheme include time filtering (Asselin 1972), using the coefficient 0.1 used operationally at ECMWF (Simmons et al. 1989). The time steps are 15 min for the three-time-level scheme and 30 min for the two-time-level scheme.

Figure 8 illustrates how there is little time truncation error for the largest scales, for which the gravity wave modes have periods much longer than the time steps examined. For wavenumbers below about 20, the damping in the two-time-level scheme due to an inexact reference state is less than the damping in the three-time-level scheme due to the use of time filtering. Conversely, the two-time-level scheme damps smaller scales more strongly than the three-time-level scheme. The dependence of phase speed on wavenumber is largely similar, but there is a little more slowing with the two-time-level scheme.

Additional calculations (not shown) indicate that the differences in damping of high-wavenumber components become larger for colder actual temperature profiles and lower actual surface pressures. For warm actual temperatures, the high-wavenumber damping becomes smaller in the two- than in the three-time-level scheme, and as seen before instability sets in with the two-time-level scheme if the actual temperature is sufficiently high. In all of these cases there is less damping of the largest scales of motion with the two-time-level scheme than with the three-time-level scheme with time filtering. The instability that occurs with the two-time-level scheme for profile 2 for reference temperatures below about 300 K is not found for the three-time-level scheme, even when time filtering is suppressed. Conversely, the three-time-level scheme is unstable for profile 1 for reference temperatures below 285 K when the actual surface pressure is 500 hPa and the time step is 15 min. The two-time-level scheme is stable with a 30-min time step in this case, at least for all reference temperatures tested, which spanned the range from 270 to 350 K.
d. Effect of decentering and horizontal diffusion

The effect of decentering is illustrated in Fig. 9. Phase speeds and damping rates are shown for $\epsilon = 0$ (solid lines) and $\epsilon = 0.1$ (dashed lines) as functions of isothermal reference temperature for the two-time-level scheme with reference surface pressure of 1013.2 hPa. The curves represented by the dotted lines are for an alternative to the decentering that will be discussed in the following section. The actual temperature distribution is profile 2, as this provides the most severe test of stability, and the actual surface pressure is 1013.2 hPa. The time step is 30 min. Decentering causes increased damping and a slight additional slowing of phase speeds. The extra damping from using $\epsilon = 0.1$ is of the same order as that from an increase of about 30 K in reference temperature. This is in accord with the analysis presented earlier, which showed that a 10% increase in reference temperature gives about the same damping as decentering with $\epsilon = 0.1$ in the special case of similar reference and actual temperature profiles.

The above calculations have been repeated for sigma coordinates with levels specified to be coincident with those of the hybrid coordinate for the surface pressure of 1013.2 hPa. The results are very similar to those shown for hybrid coordinates. This similarity holds, however, only for the case of an exact reference surface pressure. When actual surface pressures are lower than the reference value, there is more damping in the case of the hybrid coordinate. For the sigma coordinate, the matrices and vectors $\gamma$, $\pi$, $\delta$, and $\nu$ depend only indirectly on surface pressure, through dependence of the temperature profile on the pressures of the model coordinate surfaces. The only contribution to damping comes from differences between the reference and actual temperature profiles. For hybrid coordinates, there is an additional damping when the actual surface pressure is lower than the reference value.

A set of calculations showing dependence of phase speed and damping rate on wavenumber is presented in Fig. 10. We have again taken the actual temperature distribution to be profile 2. The actual and reference temperature distributions are taken to be coincident with those of the hybrid coordinate for the surface pressure of 1013.2 hPa. The results are very similar to those shown for hybrid coordinates. This similarity holds, however, only for the case of an exact reference surface pressure. When actual surface pressures are lower than the reference value, there is more damping in the case of the hybrid coordinate. For the sigma coordinate, the matrices and vectors $\gamma$, $\pi$, $\delta$, and $\nu$ depend only indirectly on surface pressure, through dependence of the temperature profile on the pressures of the model coordinate surfaces. The only contribution to damping comes from differences between the reference and actual temperature profiles. For hybrid coordinates, there is an additional damping when the actual surface pressure is lower than the reference value.
surface pressures are both taken to be 1013.2 hPa and the time step is 30 min. The dashed curves are for an isothermal reference temperature of 300 K and decentering with $\epsilon = 0.1$ and are to be compared with the solid curves, which are for an isothermal reference temperature of 330 K without decentering. The warmer reference temperature was chosen for the latter case to ensure a broadly similar level of damping for high wavenumbers. For wavenumbers higher than about 80, this calculation without decentering in fact gives a little more damping than that with decentering. In contrast, the decentering causes distinctly more damping of low and medium wavenumbers than does the use of a reference temperature that is 30 K warmer.

Two other configurations, both without decentering, are included in Fig. 10. It has been seen in Fig. 6 that the two-time-level scheme is unstable for wavenumber 213 with a reference temperature of 300 K and time step of 30 min. The dash–dotted curve in Fig. 10 shows the dependence on wavenumber in this case. Instability occurs only for the highest wavenumbers, and there is weak damping for much of the wavenumber range. Comparison of the solid and dash–dotted curves shows the impact of a 30-K increase in reference temperature, and comparison of the dashed and dash–dotted curves shows the impact of decentering. Finally, the dotted curve in Fig. 10 shows how the high-wavenumber instability in the case of a 300-K reference temperature and no decentering may be controlled by use of horizontal diffusion. Fourth-order diffusion was used in this calculation, as set out in (17)–(20): $K = 2.937 \times 10^{14}$ m$^4$ s$^{-1}$, the value used operationally at ECMWF.

5. An alternative averaging of gravity wave terms

The preceding analysis has been for a scheme in which a gravity wave tendency that is linearized about the reference state and involves variable $X$ is evaluated using the value

$$\frac{1}{2}[(1 + \epsilon)X(t + \Delta t) + (1 - \epsilon)X(t)],$$

(36)

for $X$. It has been shown how decentering with $\epsilon > 0$ can provide an effective alternative to use of a warmer reference state in stabilizing the semi-implicit scheme for high wavenumbers. However, decentering reduces the formal accuracy of the scheme from second to first order in time, and it causes a significant damping of low-wavenumber modes, which otherwise are treated relatively accurately by the semi-implicit scheme. We have thus investigated the stability of an alternative, second-order-accurate averaging of gravity wave terms. Specifically, the averaging given by (36) is replaced by

$$\frac{1}{2}[(1 + \xi)X(t + \Delta t) + (1 - 2\xi)X(t) + \xi X(t - \Delta t)],$$

(37)

where the parameter $\xi$ can be adjusted to determine the level of damping. Note that (37) can be rewritten as

$$\frac{1}{2}(1 + \xi)[X(t + \Delta t) + X(t)] - \xi X^\ast\left(t + \frac{1}{2}\Delta t\right),$$

(38)

where

$$X^\ast\left(t + \frac{1}{2}\Delta t\right) = \frac{3}{2}X(t) - \frac{1}{2}X(t - \Delta t),$$

showing how (37) may be viewed as the equivalent for the two-time-level scheme of the more general semi-implicit averaging,

$$\frac{1}{2}(1 + \xi)[X(t + \Delta t) + X(t) - \xi X(t)],$$

available as an option (and used) in the three-time-level ECMWF model (Simmons et al. 1989; Ritchie et al. 1995).

Use of (37) rather than (36) does not change the form of the stability problem. The matrices defined by (12)–(15) become
Fig. 12. The 48-h forecast of the 500-hPa height (contour interval 6 dam) from 1200 UTC 15 January 1994, using the three-time-level scheme with 15-min time step, 300-K reference temperature, and 800-hPa reference surface pressure. The model resolution is T213 and 31 layers.

\[ M_1 = -[1 + (1 + \xi)^2d_{1}\beta]^{-1}\{-2i + d_{1}[2(1 - \xi - 2\xi^3)\beta + 3(1 + \xi)E_{1}] \} \]

\[ M_2 = -[1 + (1 + \xi)^2d_{2}\beta]^{-1}\{1 + d_{2}[(-1 - 2\xi + 6\xi^3)\beta + (2 - 7\xi)E_{1} + 9E_{2}] \} \]

\[ M_3 = -[1 + (1 + \xi)^2d_{3}\beta]^{-1}\{d_{3}[2\xi(1 - 2\xi)\beta - (1 - 5\xi)E_{1} - 6E_{2}] \} \]

\[ M_4 = -[1 + (1 + \xi)^2d_{4}\beta]^{-1}\{d_{4}(\xi^2\beta - \xi E_{1} + E_{2}) \} \]

and numerical solutions can be obtained as before.

For an exact reference state, the stability problem reduces to

\[ \lambda[1 + i\beta(1 + \xi)] - \lambda[1 + i\beta(1 - 2\xi)] = i\beta\xi = 0, \]

where \( \beta \) is as defined previously. For small \( \xi \) there are two rapidly damped computational modes with

\[ \lambda = \mp \frac{i\beta\xi}{1 \mp i\beta} + O(\xi^2). \]

The eastward- and westward-moving gravity wave modes are approximated by

\[ \lambda = \frac{1 \mp i\beta}{1 \mp i\beta} + \frac{4\beta(\beta \mp i)}{(1 + \beta^2)^2} \xi + O(\xi^2). \]

For these modes, \( \lambda \) is approximately \(-1 + 4\xi\) for large \( \beta \). Thus for high wavenumbers and large time steps, \( \xi = 0.05 \) gives about the same damping as decentering with \( \epsilon = 0.1 \).

The dotted curves in Fig. 9 show the dependence of phase speed and damping rate on isothermal reference temperature for the case \( \xi = 0.05 \). The damping rate is indeed close to that for decentering with \( \epsilon = 0.1 \). These numerical solutions show that the analytical result for the case of an exact reference state holds to good ac-
accuracy for a range of reference states. Using $\xi = 0.05$ slows the phase speed by a little more than using $\epsilon = 0.1$, but this is noticeable only for warmer reference temperatures.

Figure 11 shows the dependence of damping rate on wavenumber for warm and cold distributions of actual temperature, profiles 2 and 3. Reference and actual surface pressures are 1013.2 hPa. The solid and dashed curves are as in Fig. 10, and show results using a standard semi-implicit scheme ($\epsilon = \xi = 0$) with 330-K reference temperature and those using decentering with $\epsilon = 0.1$ and a 300-K reference temperature. The dash-dotted curve is for the averaging of form (37) with $\xi = 0.05$ and 300-K reference temperature. The latter option evidently gives much less damping of low and intermediate wavenumbers than decentering with $\epsilon = 0.1$. For low wavenumbers it gives about the same amount of damping as the standard semi-implicit scheme with a reference temperature that is 30 K warmer. This similarity in damping holds for all wavenumbers for profile 3, while for profile 2 the scheme with $\xi = 0.05$ gives about the same high-wavenumber damping as decentering with $\epsilon = 0.1$, as discussed previously. Generally, similar results have been obtained for profile 1 and for lower actual surface pressures, with significant variations only in the extent of the damping of high-wavenumber components.

As far as the time discretization is concerned, the scheme described in this section is the same as the second-order decentering proposed by Rivest et al. (1994) and generalized by Côté et al. (1995). However, (38) suggests a considerably more economical spatial discretization than that of Rivest et al. Denoting the “arrival” grid point by $g$ and the “departure” point by $d$, (38) may be spatially discretized as

$$\frac{1}{2}(1 + \xi)[X(t + \Delta t, g) + X(t, d)] - 0.5\xi \left[ X^*(t + \frac{1}{2} \Delta t, g) + X^*(t + \frac{1}{2} \Delta t, d) \right].$$

Thus, the $X^*$ term in (38) is first extrapolated in time to $(t + \frac{1}{2}\Delta t)$ and then averaged along the trajectory. The term in $X(t - \Delta t)$ is thereby included without the inconvenience of having to compute an extra departure point corresponding to $(t - \Delta t)$ and without the expense of additional interpolations.
6. Experiments

A two-time-level version of the “fully interpolating” semi-Lagrangian model described by Ritchie et al. (1995) is currently under development. This work will be reported in detail elsewhere. Here, we show some preliminary results to illustrate the application of the stability analysis.

Figures 12–15 show 48-h forecasts of the 500-hPa height field, starting from the ECMWF analysis for 1200 UTC 15 January 1994. The horizontal resolution is T213 and the vertical resolution is 31 layers. Figure 12 shows the forecast using the standard three-time-level scheme with a reference temperature of 300 K, a reference surface pressure of 800 hPa, and a time step of 15 min. Figure 13 shows the corresponding forecast using the two-time-level scheme with the same reference temperature and surface pressure, and a time step of 30 min. Developing instabilities are clearly seen in the subtropics, where the actual temperature profile is warm and the surface pressure high.

Figure 14 demonstrates the stabilizing effect of changing the reference temperature and surface pressure to 350 K and 1000 hPa, respectively, while Fig. 15 shows that the same effect can be achieved while retaining the original reference temperature and surface pressure by introducing the alternative second-order averaging, spatially discretized as in (39), with a coefficient of $\xi = 0.05$.

7. Conclusions

The stability properties of the semi-implicit method of treating gravity wave terms in two-time-level semi-Lagrangian models have been shown to be different from those of the method as used in three-time-level models. Warmer reference temperatures should be used in two-time-level schemes if standard second-order accurate averaging of gravity wave tendencies is used. For hybrid pressure-based vertical coordinates, the reference surface pressure in two-time-level schemes should be set to a value near the high end of the range of pressures encountered in the model, rather than to the intermediate value that is required for three-time-level schemes. First-order decentering in the averaging of gravity wave tendencies reduces the need for a relatively warm reference temperature (or high reference surface pressure), but...
causes a quite substantial damping of otherwise well-represented low-wavenumber modes. The second-order averaging involving a third (past) time level given by (37) offers a promising alternative that gives substantially less low-wavenumber damping. An economical spatial discretization has been proposed and demonstrated for this alternative second-order averaging.

The two-time-level version of the semi-implicit scheme, when stable, is generally damping. In most of the calculations we have made there is more damping of higher wavenumber components by the two-time-level scheme than by the three-time-level scheme with time filtering. It is not clear, however, that the larger high-wavenumber damping is any cause for concern. A gross misrepresentation of the phase speeds of fast-moving high-wavenumber modes does not appear to be detrimental to the application of semi-Lagrangian models in numerical weather prediction, so additional unphysical damping might also be tolerated.

Conversely, there can be less damping of large-scale motion by the two-time-level scheme than by the three-time-level scheme when the latter includes time filtering with the coefficient that is used operationally at ECMWF. In the representative set of numerical solutions presented here the fastest gravity wave with total wavenumber 10 is damped with an e-folding time of around 1 day by the three-time-level scheme. The corresponding e-folding time is in the range from 1.1 to 1.5 days for the two-time-level scheme with either a warm reference temperature or the alternative second-order averaging. It is between 0.4 and 0.5 days for the two-time-level scheme with first-order decentering. These are for a gravity wave mode with phase speed of the order of 300 m s\(^{-1}\). Damping of the slower, meteorological modes of similar spatial scale should be much weaker. The absence of a damping time filtering of the vorticity equation in the two-time-level scheme is worth noting in this context.

Acknowledgments. We are grateful to Aidan McDonald and Andrew Staniforth for their valuable comments on a preliminary draft of this paper. Comments from Tony Hollingsworth on a later draft are also acknowledged.

REFERENCES


