Tellus (1996), 48A, 74–95

TELLUS
ISSN 0280–6495

The tangent linear model for semi-Lagrangian schemes: linearizing the process of interpolation

By SAROJA POLAVARAPU*, MONIQUE TANGUAY[1], RICHARD MÉNARD[2]** and ANDREW STANIFORTH[1], *Data Assimilation and Satellite Meteorology Division, Atmospheric Environment Service, 4905 Dufferin Street, Downsview, Ontario, Canada M3H 5T4*; [1]*Division de Recherche en Prévision Numérique, Service de l'Environnement Atmosphérique, 2121, route Transcanadienne, Dorval, PQ, Canada H9P 1J3*; [2]*Department of Atmospheric and Oceanic Sciences, McGill University, Montréal, PQ, Canada H3A 2K6*

(Manuscript received 9 January 1995; in final form 16 May 1995)

ABSTRACT

The tangent linear model may be used in diverse applications such as Kalman filtering, variational assimilation using the adjoint method, sensitivity studies or predictability studies. A "correct" tangent linear variation contains all of the linear part of the nonlinear variation. This concept is used to show that simply differentiating a nonlinear model's code does not necessarily lead to a tangent linear model which is correct in all circumstances. The example of linearizing interpolation schemes is used. For infinitesimal variations, the linear variation is correct if and only if the first derivative of an interpolator is continuous. Even if the tangent linear variation is occasionally incorrect, the size of the error can be determined and may in fact be quite tolerable. Therefore, there should be no fundamental difficulty in linearizing semi-Lagrangian schemes if care is taken in choosing an appropriate interpolation scheme.

1. Introduction

4-dimensional variational assimilation refers to the process of determining the model initial state whose evolution over a given time interval variationally minimizes the distance to a set of observations. Since the introduction of four-dimensional variational assimilation using the adjoint method (4DVAR) to meteorology (Lewis and Derber, 1985; LeDimet and Talagrand, 1986; Courtier and Talagrand, 1987; Derber, 1987; Talagrand and Courtier, 1987), steady progress has been made in demonstrating the method's ability to propagate information and implicitly generate flow-dependent structure functions (Rabier and Courtier, 1992; Rabier et al., 1993; Thépaut et al., 1993; Andersson et al., 1994). Because of these encouraging results and the recent proposal of Courtier et al. (1994) which reduces the cost of 4DVAR by an order of magnitude, the operational implementation of this method is now viable and imminent. Nevertheless, the cost is still high and with the inclusion of physical processes it will be even higher so that the efficiency advantage offered by the large time steps of semi-Lagrangian time integration schemes is desirable (Chao and Chang, 1992). While this efficiency has been demonstrated in the context of 4DVAR (Li et al., 1993; 1994), such models have yet to be fully embraced by the data assimilation community. One reason for the reluctance to use such models for 4DVAR is due to the difficulty in linearizing interpolation schemes (Li et al., 1993; Ménard, 1994). Li et al. (1993) have also suggested that the range of validity of the tangent linear hypothesis is reduced where large gradients in the flow are combined with strong advection and this imposes an additional restriction on the time step.

* Corresponding author.
** Current Affiliation: Data Assimilation Office, Code 910.3, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

Given the investment of operational meteorological forecast models in this method, it is clear that there is a need to provide a mathematical description of the potential difficulty of linearizing semi-Lagrangian models as well as practical solutions and quantitative error estimates. This is the goal of the present work.

In this article, we examine the tangent linear model (hereafter TLM) for semi-Lagrangian schemes. The TLM can be used to describe the evolution of short range forecast errors (Lacarra and Talagrand, 1988; Vukicevic, 1991; Rabier and Courtier, 1992) for numerical weather prediction. As a result, this model finds application in the determination of unstable modes, the propagation of forecast error covariances for a Kalman filter, and the verification of adjoint models which may be used for four-dimensional variational assimilation or sensitivity analysis. References on all of these topics are found in Courtier et al. (1993). References on meteorological applications of the Kalman filter are found in Cohn (1993). In developing an adjoint model, one must verify its correctness (Thépaut and Courtier, 1991) to machine precision. Similarly, in developing a TLM, its correctness must be verified and the range of validity of the tangent linear hypothesis determined. Correctness is a property of the TLM which can be objectively ascertained (Thépaut and Courtier, 1991) and refers to the fact that all of the linear part of the nonlinear variation is explained by the TLM. On the other hand, the size of perturbations for which the TLM hypothesis is valid depends on the application within which the TLM is utilized. Thus, we focus our attention on the correctness of TLMs. As we shall see, simply differentiating a nonlinear code does not always lead to a "correct" TLM.

In Section 2 we isolate the nondifferentiable part of the semi-Lagrangian scheme and in Section 3 we clarify the concepts of "correctness" and "validity" with respect to a TLM. We then proceed to determine the linearization error for a general interpolation scheme as well as for specific, commonly used schemes in Section 4. Numerical examples illustrating properties of error are also presented. In Section 5, we re-integrate the results of the preceding sections back into the meteorological context of a simple one-dimensional semi-Lagrangian model. The results are summarized and discussed in Section 6.

2. The semi-Lagrangian time integration scheme

A semi-Lagrangian treatment of advection when combined with the semi-implicit time integration scheme has proven to be an efficient way of integrating the primitive equations (Robert et al., 1985). In particular, a factor of 6 increase in the time step over a corresponding Eulerian integration, may be obtained. For a review of the use of the semi-Lagrangian method in meteorology as well as a brief description of this method, see Staniforth and Côté (1991). Perhaps the simplest model with meteorological application for which we may implement semi-Lagrangian advection is Burgers' equation with no viscosity (Kuo and Williams, 1990):

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}=0. \tag{1}$$

A two-time-step semi-Lagrangian model of this equation is

$$u(x_i, t)=u(x_i-\alpha_i, t-\Delta t), \tag{2}$$

where the subscript i refers to the ith grid point and α_i is determined from an $O[(\Delta t)^2]$ accurate approximation to

$$\frac{dx}{dt}=u(x, t). \tag{3}$$

The right-hand side of (2) involves interpolation of velocity between grid points. As noted by Li et al. (1993), the interpolation may possess a non-differentiable operation. Specifically, for regular grid spacing and linear interpolation,

$$u(x_i-\alpha_i, t-\Delta t)=au(x_{j-1}, t-\Delta t)$$
$$+bu(x_j, t-\Delta t),$$

where $b=(x_i-\alpha_i-x_{j-1})/\Delta x$ and $a=1-b$. The difficulty arises due to the following operation:

$$j=\mathrm{INT}\left(\frac{x_i-\alpha_i}{\Delta x}\right)+1,$$

which locates the grid interval that contains the upstream interpolation point. INT is a Fortran intrinsic function which truncates a real number to obtain an integer. Note that j is a function of a control variable, α_i, but j is not differentiable. As a result, the tangent linear code is linearized about this grid interval although a perturbed interpolation point may very well lie outside it. The question then arises as to whether or not the TLM is correct in such a circumstance. Furthermore, we have thus far only mentioned linear interpolation and regular grid spacing. Does this problem arise for all interpolation schemes? In what follows we address these questions by examining the linearization error for various interpolation schemes. First, however, we must define what is meant by a "correct" TLM.

3. The concept of correctness in tangent linear models re-examined

In this section, we define the concept of "correctness" of a TLM, equating it with a linearization error which is asymptotically second order in perturbed variables. We also discuss how correctness differs from the validity of the TLM hypothesis. Correctness can be objectively established for a given discrete linear model (the TLM) and refers to the ability of this model to asymptote to the linear part of the nonlinear variation as the size of perturbations is diminished. On the other hand, the TLM hypothesis is valid when the difference between the linear part of the nonlinear variation and the total nonlinear variation is small when compared to its linear part. Thus the validity of the TLM hypothesis must be established for a given discrete TLM and in the context of a particular application and is not a property of a given model. In this article, we shall be mainly concerned with the issue of correctly deriving a TLM. As we shall see in Section 4, simply differentiating a nonlinear model does not always lead to TLM which is "correct" in all circumstances.

Consider a model, $\boldsymbol{M}$, which updates a state variable, $\boldsymbol{X}$, as follows:

$$\boldsymbol{X}^n = \boldsymbol{M}(\boldsymbol{X}^{n-1}), \tag{4}$$

where $\boldsymbol{X}^0$ is given and superscript n refers to time $n\,\Delta t$. $\boldsymbol{M}$ is a vector valued function of dimension N and N is the number of time-steps over which data will be assimilated. For a given space-time trajectory, $\bar{\boldsymbol{X}}^n$, $n = 0, 1, \ldots, N$, which satisfies (4), perturbations or variations about this reference trajectory evolve according to

$$\delta\boldsymbol{X}^n = \boldsymbol{M}(\bar{\boldsymbol{X}}^{n-1} + \delta\boldsymbol{X}^{n-1}) - \boldsymbol{M}(\bar{\boldsymbol{X}}^{n-1}). \tag{5}$$

Now the *tangent linear hypothesis* says that for small perturbations, i.e., for $|\delta\boldsymbol{X}| \ll |\bar{\boldsymbol{X}}|$, a good approximation to (5) is given by the linear part of this nonlinear variation:

$$\frac{\mathrm{d}\boldsymbol{M}}{\mathrm{d}\boldsymbol{X}}(\bar{\boldsymbol{X}}^{n-1})\,\delta\boldsymbol{X}^{n-1}. \tag{6}$$

By definition, the TLM describes the evolution of the linear (i.e., 1st order) variation which is tangent to the reference trajectory in phase space. Thus, the TLM should contain all of the linear part of the nonlinear variation and hence equal the above expression. Clearly, this is true for a continuous or theoretical model. However, for a discrete, numerical model, the TLM is sometimes obtained by differentiating the model code line by line (Navon et al., 1992; Li et al., 1993) and due to the existence of nondifferentiable or discrete programming structures, the TLM is not necessarily equal to the linear variation. Therefore, it is necessary to distinguish between the linear variation predicted by the TLM and denoted $(\delta\boldsymbol{X}^n)^L$ (see equation below), and the linear part of the nonlinear variation given by (6). The TLM may be written as

$$(\delta\boldsymbol{X}^n)^L = L(\bar{\boldsymbol{X}}^{n-1})\,\delta\boldsymbol{X}^{n-1},$$

where L is a linear operator which can be viewed as a matrix. Thus the TLM is correct when it is equal to the linear variation, i.e., $L = \mathrm{d}\boldsymbol{M}/\mathrm{d}\boldsymbol{X}$. The difference between the nonlinear variation and the linear variation described by the TLM is

$$\begin{aligned}\boldsymbol{E} \equiv \delta\boldsymbol{X}^n - (\delta\boldsymbol{X}^n)^L = \boldsymbol{M}(\bar{\boldsymbol{X}}^{n-1} + \delta\boldsymbol{X}^{n-1}) \\ - \boldsymbol{M}(\bar{\boldsymbol{X}}^{n-1}) - (\delta\boldsymbol{X}^n)^L.\end{aligned} \tag{7}$$

We call E the *linearization error*. The component of $\boldsymbol{E}$ corresponding to the ith component of the state vector $\boldsymbol{X}$ is

$$\begin{aligned}E_i = \sum_{j=1}^{N}\left(\frac{\partial M_i}{\partial X_j}(\bar{\boldsymbol{X}}^{n-1}) - l_{ij}(\bar{\boldsymbol{X}}^{n-1})\right)\delta X_j^{n-1} \\ + \frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 M_i}{\partial X_j\,\partial X_k}(\bar{\boldsymbol{X}}^{n-1})\,\delta X_j^{n-1}\,\delta X_k^{n-1} \\ + \mathrm{O}((\delta\boldsymbol{X}^{n-1})^3),\end{aligned} \tag{8}$$

on expanding $M_i(\bar{X}^{n-1}+\delta X^{n-1})$ in Taylor series about $\bar{X}^{n-1}$. l_{ij} refers to the (i, j)th element of L. Thus the TLM is correct if and only if for all i and all δX^{n-1}, E_i is 2nd-order in perturbed variables as δX^{n-1} goes to zero. Also, when the TLM is correct, the linear variation asymptotically approaches the nonlinear variation as the size of the variation diminishes. To conduct a test that verifies the correctness of TLM, it turns out to be useful to write the perturbation in a manner which identifies its shape and amplitude:

$$\delta X = c\,\delta\hat{X}.$$

c is a positive real number defining the amplitude of the perturbation while $\delta\hat{X}$ defines its shape. The linearization error may then be written as a function of perturbation amplitude if the shape is held fixed:

$$\begin{aligned} E_i(c) &= c\sum_{j=1}^{N}\left(\frac{\partial M_i}{\partial X_j}(\bar{X}^{n-1}) - l_{ij}(\bar{X}^{n-1})\right)\delta\hat{X}_j^{n-1} \\ &\quad + c^2\frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 M_i}{\partial X_j\,\partial X_k}(\bar{X}^{n-1})\,\delta\hat{X}_j^{n-1}\,\delta\hat{X}_k^{n-1} \\ &\quad + \mathrm{O}((c\,\delta\hat{X}^{n-1})^3). \end{aligned} \tag{9}$$

Now if the TLM is correct then E_i/c^2 will be asymptotically independent of perturbation amplitude, c, for all i. Thus the test that E_i/c^2 be asymptotically independent of c establishes the continuity of the derivative of M_i in the direction of $\delta\hat{X}$. If this is true for all such directions, then the TLM is correct. Thépaut and Courtier (1991) use the test,

$$\lim_{\delta X\to 0} R_i^{\mathrm{TC}} = 1$$

where

$$R_i^{\mathrm{TC}} = 1 + E_i\left[\sum_{j=1}^{N} l_{ij}(\bar{X}^{n-1})\,\delta X_j^{n-1}\right]^{-1}. \tag{10}$$

If the TLM is correct and E_i is 2nd-order for all i and for all variations, δX, then R_i^{TC} is asymptotically linear in variations. Thus both tests use the same criterion for correctness and either test provides a necessary condition for correctness. A necessary and sufficient condition for correctness obtains only when either test is passed for all possible directions, $\delta\hat{X}$.

As we have seen, there exists a quantitative, objective means of determining whether a given TLM is indeed *correct*. However, even when the TLM is correct, the TLM hypothesis may not be valid. The TLM hypothesis is *valid* when the linear variation is a good approximation to the nonlinear variation. A test of validity is obtained by looking at the magnitude of the linearization error to the first order term:

$$R_i^{\mathrm{V}} = E_i\left[\sum_{j=1}^{N}\frac{\partial M_i}{\partial X_j}(\bar{X}^{n-1})\,\delta X_j^{n-1}\right]^{-1}. \tag{11}$$

When R_i^{V} is small for all i, the TLM hypothesis is valid. The concept of validity involves some subjectivity since one must decide just how small the 2nd-order term should be in relation to the 1st-order term in order for the TLM to be valid. The ratio test of Rabier and Courtier (1992) is similar (except that domain integrated errors are used) and the reader is referred to their paper for further discussion.

In this article we are mainly concerned with the correctness of a TLM, which as noted above, can be quantitatively verified for a given discrete linear model. While some conclusions can be drawn about the validity of the TLM's considered, in general, this must be determined by the application in which the TLM is used.

4. Linearizing the process of interpolation

We now proceed to determine the linearization error for a general piecewise-continuous interpolating function. We shall also consider some examples of interpolation schemes in order to determine the size of this error in practice. First, however, we isolate the nonlinearity of the interpolation process and reduce the problem to its essence.

Consider a continuously-differentiable function $y(x)$ with independent variable x. Interpolation of y at a given point ξ may be described as

$$\eta = y(\xi). \tag{12}$$

Because $y(x)$ is known everywhere, the process of interpolation has simply become function evaluation in this case. In reality, y would be known only at a finite number of points so $y(x)$ would be approximated by some interpolating function. Now assume $y(x)$ and ξ consist of two parts: $y(x)=\bar{y}(x)+\delta y(x)$ and $\xi=\bar{\xi}+\delta\xi$ where barred quantities are reference variables and the perturbations about these reference variables, $\delta y(x)$ and $\delta\xi$, are small. The tangent linear variation of η due to variations in the function itself, $\delta y(x)$ and the point of interpolation, $\delta\xi$ is:

$$\delta\eta^L=\delta y(\bar{\xi})+\left.\frac{\mathrm{d}\bar{y}}{\mathrm{d}x}\right|_{\bar{\xi}}\delta\xi. \tag{13}$$

We shall call (13) a *tangent linear equation* or TLE because (12) is not, strictly speaking, a model. From (12), it is easy to see that η is a linear function of $y(x)$ but, in general, a nonlinear function of ξ. Thus in the absence of variation in ξ, interpolation would be a purely linear process. Therefore, we shall now consider only variations of ξ, that is, we set $\delta y(x)$ to zero. *This simplification is made in order to isolate the nonlinearity of interpolation and to ultimately identify conditions of linearization.* (We will return to the case of nonzero $\delta y(x)$ in Section 5.) Now by analogy with (7) we can define a linearization error for interpolation as

$$E=\delta\eta-\delta\eta^L,$$

which for the special case of no variations in $y(x)$ is

$$E=\bar{y}(\bar{\xi}+\delta\xi)-\bar{y}(\bar{\xi})-\frac{\mathrm{d}\bar{y}}{\mathrm{d}x}(\bar{\xi})\,\delta\xi. \tag{14}$$

With this definition, we may now determine the correctness of the TLE for various common interpolation schemes starting from the general case of a piecewise differentiable interpolating function.

4.1. *The general principle*

Let us consider the problem of interpolation in one-dimension. We define a grid with $N+1$ grid points, $x_0, x_1, \ldots, x_N$. A function y is assumed to have known values at the grid points, namely, $y_0, y_1, \ldots, y_N$. Now consider a piecewise continuous interpolating function, $P(x)$, which takes values y_j at $P(x_j)$ (Fig. 1). Let us assume that the original interpolation point lies in the half-open interval

$$\bar{\xi}\in[x_{j-1}, x_j),$$

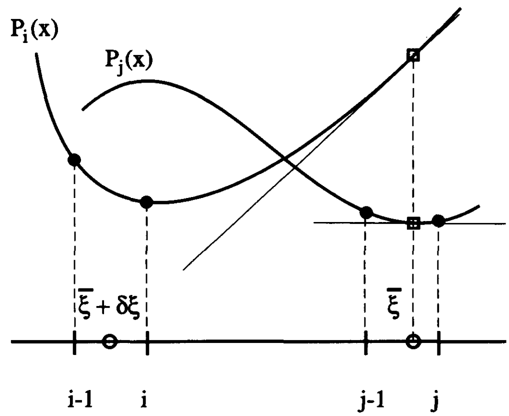

Fig. 1. Interpolation with a piecewise-continuous interpolator. The solid circles indicate locations where function values are provided. An interpolation point, $\bar{\xi}$, lies in the interval $[x_{j-1}, x_j)$ for which the interpolating function is $P_j(x)$. The perturbed interpolation point, $\bar{\xi}+\delta\xi$, lies in the interval $[x_{i-1}, x_i)$ for which the interpolating function is $P_i(x)$. The thin solid lines are tangent to the functions $P_j(x)$ and $P_i(x)$ at $\bar{\xi}$.

and that the perturbed interpolation point lies in the half-open interval:

$$\bar{\xi}+\delta\xi\in[x_{i-1}, x_i).$$

Then the nonlinear variation is determined by

$$P_i(\bar{\xi}+\delta\xi)-P_j(\bar{\xi}),$$

where the subscript refers to the interval for which the interpolating function is appropriate. The TLE is given by

$$\delta P_j(\bar{\xi})=\frac{\mathrm{d}P_j}{\mathrm{d}\xi}(\bar{\xi})\,\delta\xi,$$

so that according to (14), the linearization error is

$$E=P_i(\bar{\xi}+\delta\xi)-P_j(\bar{\xi})-\frac{\mathrm{d}P_j}{\mathrm{d}\xi}(\bar{\xi})\,\delta\xi,$$

which can be expanded using Taylor series to show that

$$E=[P_i(\bar{\xi})-P_j(\bar{\xi})]+\left[\frac{\mathrm{d}P_i}{\mathrm{d}\xi}(\bar{\xi})-\frac{\mathrm{d}P_j}{\mathrm{d}\xi}(\bar{\xi})\right](\delta\xi)$$
$$+\frac{1}{2}\frac{\mathrm{d}^2P_i}{\mathrm{d}\xi^2}(\bar{\xi})(\delta\xi)^2+\mathrm{O}[(\delta\xi)^3]. \tag{15}$$

Note that evaluation of P_i and its derivatives at $\bar{\xi}$ involves *extrapolation* of this function beyond the interval for which it is defined. Therefore, all further discussion pertains only to interpolating functions which are defined over the whole domain, such as polynomials. There are a number of points to note regarding the linearization error for piecewise continuous interpolating functions:

1. If the functions P_i and P_j are not equal at $\bar{\xi}$ then there is a zeroth order term in $\delta\xi$.
2. If the first derivatives of P_i and P_j are not equal at $\bar{\xi}$ then there is a 1st-order term in $\delta\xi$.
3. If the functions P_i and P_j as well as their first derivatives are equal at $\bar{\xi}$ then there is no linearization error even if higher-order derivatives are not equal at $\bar{\xi}$.
4. The 2nd- and higher-order terms involve extrapolating the function appropriate for the interval containing $\bar{\xi}+\delta\xi$ to the interval containing $\bar{\xi}$. Thus the better this extrapolation, the smaller the linearization error and the larger the range of validity of the tangent linear hypothesis.

Therefore, it is easy to see that the TLE is correct for all $\bar{\xi}$ if and only if $P_i = P_j$ for all i and j. This is true for spectral interpolation which we loosely define to be as the evaluation of a series expansion in terms of a finite number (corresponding to the number of meshpoints) of globally-defined and infinitely-differentiable basis functions whose sum exactly fits the data at meshpoints: these basis functions are usually the eigenfunctions associated with the spectrum of a linear operator, whence the nomenclature spectral. However, spectral interpolation is not well-suited for grids with variable spacing and it is expensive for a large number of grid points. Thus, piecewise-polynomial interpolation is often necessary. For this type of interpolation, $P_i = P_j$ for all i and j when the interpolation is exact, (i.e., if $y(x)$ is a polynomial of at most order n for a piecewise polynomial interpolant of order n), or when using a single interpolating polynomial over the whole domain. For a large number of grid points, a single interpolating polynomial would be of high order (which is expensive to compute and may oscillate wildly near boundaries), so low-order piecewise polynomials are preferable. Thus, linearization error is unavoidable. In reality, however, the perturbed interpolation point is determined by the variation of the velocity which is continuous, and for the purposes of determining forecast error growth, presumably small. Thus, in practice, the perturbed interpolation point is likely to remain within one or two grid intervals of the unperturbed interpolation point. Therefore, it is important to consider some special subcases of the general one. Moreover, for 4DVAR, because variations are *infinitesimal*, it is only the two cases described below that are relevant.

Case 1

If the interpolation point and the perturbed interpolation point lie in the same grid interval, then $i = j$ and

$$E = \frac{1}{2}\frac{\mathrm{d}^2 P_j}{\mathrm{d}\xi^2}(\bar{\xi})(\delta\xi)^2 + \mathrm{O}[(\delta\xi)^3]. \tag{16}$$

Thus the TLE is correct since the interpolating function and its first derivative are continuous within a grid interval. Of course, in practice, there is no guarantee that the perturbed and unperturbed interpolation points will lie in the same grid interval.

Case 2

Now, if the interpolation point lands on a grid point, say $\bar{\xi} \equiv x_{j-1}$, all negative perturbations less than the grid spacing will lie in the adjacent interval. Then $i = j-1$ and since the polynomials are continuous at x_{j-1}:

$$E = \left[\frac{\mathrm{d}P_{j-1}}{\mathrm{d}\xi}(x_{j-1}) - \frac{\mathrm{d}P_j}{\mathrm{d}\xi}(x_{j-1})\right](\delta\xi) + \frac{1}{2}\frac{\mathrm{d}^2 P_{j-1}}{\mathrm{d}\xi^2}(x_{j-1})(\delta\xi)^2 + \mathrm{O}[(\delta\xi)^3]. \tag{17}$$

For this situation, the TLE is correct if and only if the interpolating function has continuous first derivatives at grid points. If the interpolating point does not land exactly on a grid point, then not even the continuity of adjacent polynomials at $\bar{\xi}$ is ensured so the linearization error contains a zeroth-order term (which may be small) and is given by the general expression, (15), except that $i = j-1$.

Since piecewise-continuous interpolating functions may not always have correct TLEs, it is necessary to determine the size of these errors for the important case where the perturbed interpolation point lies in the adjacent interval to that containing the unperturbed one.

4.2. Some practical examples of linearization errors

In this section, we consider the linearization error of some actual, commonly-used interpolation schemes in order to show that while the TLE may be incorrect in some cases, the size of the error incurred may be, practically speaking, tolerable.

We shall consider interpolation schemes which can be written in the following form:

$$P(x) = a_j y_{j-1} + b_j y_j + \frac{h_j^2}{6} [(a_j^3 - a_j) y''_{j-1} + (b_j^3 - b_j) y''_j], \tag{18}$$

where

$$a_j = \frac{x_j - x}{h_j}, \quad b_j = \frac{x - x_{j-1}}{h_j}, \quad h_j = x_j - x_{j-1}. \tag{19}$$

For this type of interpolation, j is chosen such that

$$x \in [x_{j-1}, x_j).$$

By choosing how the $y''_j s$ are computed, we obtain various interpolation schemes. We consider four such schemes: linear, cubic Lagrange, cubic spline and deficient cubic spline. (See *Appendix* for details.) Linear interpolation is the simplest and cheapest interpolation scheme and is useful when accuracy is not of great concern. The cubic Lagrange and cubic spline schemes are formally of the same accuracy. The cubic Lagrange scheme is attractive because it is local in nature and hence relatively cheap and easy to implement. On the other hand, the cubic spline offers continuity of first and second derivatives everywhere and a better global fit. The price for this is the additional solution of a tridiagonal set of equations for each grid. The deficient cubic spline, provides a local fit to the data and has continuous 1st derivative everywhere. However, discontinuities at grid points occur in the second derivative and it is formally O(h) less accurate than the cubic Lagrange and cubic spline schemes.

We now determine the linearization error for the interpolation scheme described above. Let $\bar{\xi}$ be the point of interpolation which obtains in the absence of any variations and let $\delta\xi$ be the variation of this point. We assume no variation of the function, $y(x)$ and hence $\delta y_j = 0$ for all j. The linearization error may then be computed following the procedure used to obtain the error for the general case, or simply by substituting (18) into (15). In either case, the resulting error is

$$\begin{aligned} E = &\left[\bar{a}_i \bar{y}_{i-1} + \bar{b}_i \bar{y}_i + \frac{h_i^2}{6} ((\bar{a}_i^3 - \bar{a}_i) \bar{y}''_{i-1} + (\bar{b}_i^3 - \bar{b}_i) \bar{y}''_i) \right] - \left[\bar{a}_j \bar{y}_{j-1} + \bar{b}_j \bar{y}_j + \frac{h_j^2}{6} ((\bar{a}_j^3 - \bar{a}_j) \bar{y}''_{j-1} + (\bar{b}_j^3 - \bar{b}_j) \bar{y}''_j) \right] \\ &+ \left[\frac{\bar{y}_i - \bar{y}_{i-1}}{h_i} + \frac{h_i}{6} (-(3\bar{a}_i^2 - 1) \bar{y}''_{i-1} + (3\bar{b}_i^2 - 1) \bar{y}''_i) - \frac{\bar{y}_j - \bar{y}_{j-1}}{h_j} - \frac{h_j}{6} (-(3\bar{a}_j^2 - 1) \bar{y}''_{j-1} + (3\bar{b}_j^2 - 1) \bar{y}''_j) \right] (\delta\xi) \\ &+ \frac{1}{2} (\bar{a}_i \bar{y}''_{i-1} + \bar{b}_i \bar{y}''_i)(\delta\xi)^2 + \frac{1}{6} \frac{\bar{y}''_i - \bar{y}''_{i-1}}{h_i} (\delta\xi)^3, \end{aligned} \tag{20}$$

where

$$\bar{a}_k = \frac{x_k - \bar{\xi}}{h_k}, \quad \bar{b}_k = \frac{\bar{\xi} - x_{k-1}}{h_k}.$$

Case 1

First, consider the case where the original interpolating point and the perturbed interpolating point lie in the same grid interval. Then $i = j$ and the above error becomes:

$$E = \frac{1}{2} (\bar{a}_j \bar{y}''_{j-1} + \bar{b}_j \bar{y}''_j)(\delta\xi)^2 + \frac{1}{6} \frac{\bar{y}''_j - \bar{y}''_{j-1}}{h_j} (\delta\xi)^3. \tag{21}$$

The TLE is correct since there is no first-order term and this is so regardless of which particular interpolator is chosen from the family. The second-order term is simply the 2nd derivative of $P(x)$ evaluated at $\bar{\xi}$. The 3rd-order term is the third derivative of $P(x)$ at $\bar{\xi}$. Both terms are similar in form to those contained in the true variation, $\delta\eta$. However, because a cubic polynomial was used for the approximating function, terms higher than third order are identically zero since all derivatives higher than the third are zero everywhere.

By substituting the specified forms for the second derivatives, (31) to (35), into (21) we obtain the linearization errors found in the Appendix. To illustrate these errors, we have performed some numerical experiments. The experimental setup is as follows. A one-dimensional grid consisting of $N=91$ equally-spaced grid points is constructed where $-0.5 \leqslant x \leqslant 0.5$. A regular grid spacing was chosen for the sake of simplicity and since the correctness of the TLEs of these schemes does not depend on variability of resolution. For each grid interval, the interpolation point was chosen to be $0.75\ \Delta x$ from the left edge. The linearization error was computed at each interpolation point for the sequence of small perturbations:

$$-\delta\xi = 0.005, 0.0005, 0.00005, 0.000005.$$

Thus the perturbed and unperturbed interpolation points lie in the same grid interval. A known analytic function, $y(x)=x^4$, was specified and evaluated at grid points. These values were then supplied to the 4 interpolation schemes mentioned above and the linearization error computed. The graphs of the function $y(x)=x^4$ and its derivatives are shown in Fig. 2. In Fig. 3, we illustrate the spatial variation of the ratio $R2$, where

$$R2 = \frac{E}{(\delta\xi)^2}. \tag{22}$$

When a TLE is correct and $\delta\xi$ is sufficiently small, E is dominated by a 2nd-order term so $R2$ is a con-

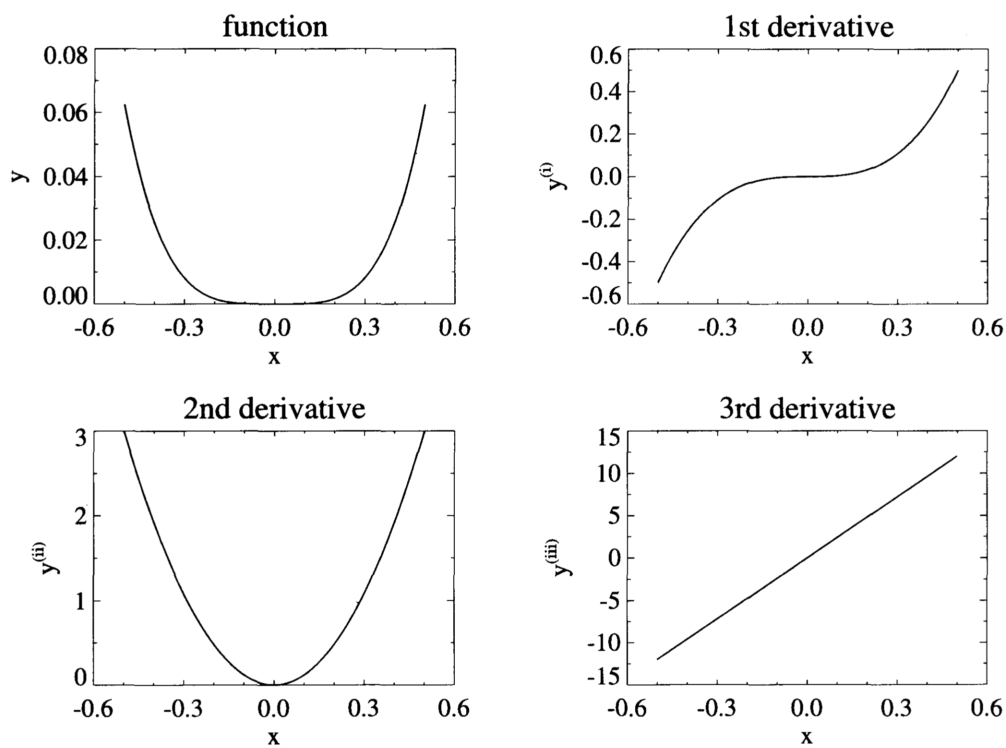

Fig. 2. The function $y=x^4$ and its first 3 derivatives.

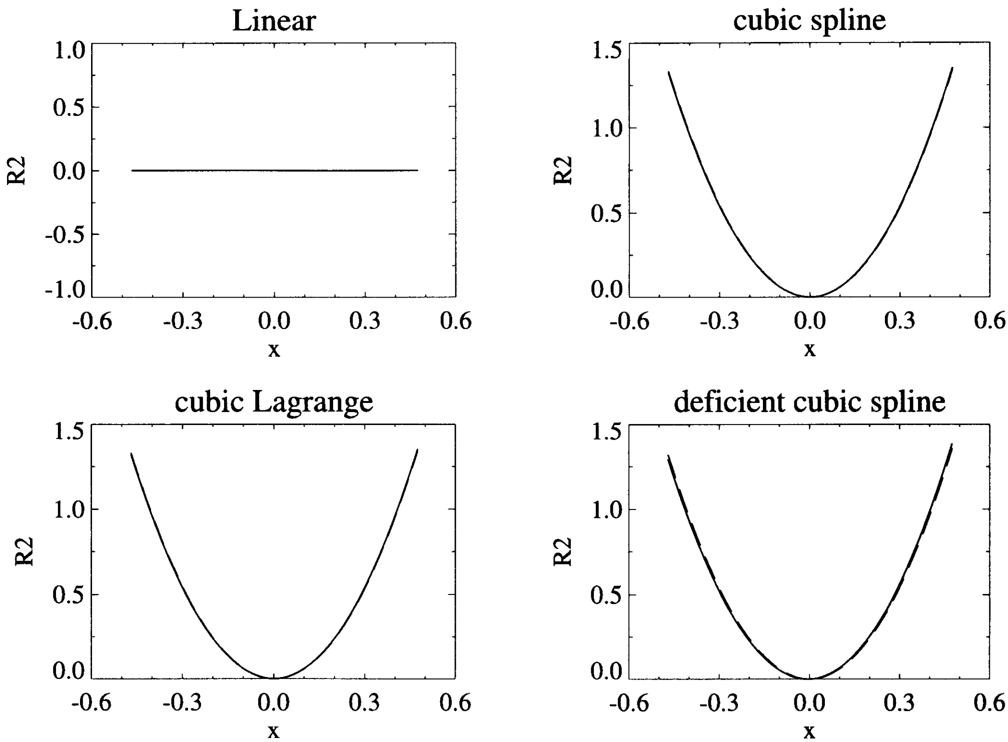

Fig. 3. The spatial distribution of the ratio, $R2 = E/(\delta\xi)^2$, for various interpolation schemes. The interpolation point is 0.75 Δx from the left edge of each grid interval. The solid, dash-dot, dash-dot-dot-dot and dashed curves correspond to $-\delta\xi = 0.000005$, 0.00005, 0.0005, 0.005 respectively. The function which is being approximated is $y = x^4$.

stant, independent of $\delta\xi$. Therefore plots of $R2$ for $\delta\xi$ of differing magnitude should almost exactly coincide. On the other hand, if the TLE is incorrect and there is a first- or zeroth-order term, then $R2$ will increase as $\delta\xi$ decreases and the various curves will spread apart. In Fig. 3, we see that all curves coincide and hence all interpolation schemes have correct TLEs when both perturbed and unperturbed interpolation points lie in the same grid interval. As expected from (21), all ratios describe one-half the second derivative of $y(x)$ (c.f., Fig. 2) except for the case of linear interpolation where the ratio is zero because its interpolator has zero second derivative. While all curves coincide, there is a hint of some error for the deficient spline. This is due to the modulation of the 2nd-order term by the 3rd derivative (see(40)). This term is identically zero when $a = b$ (i.e., at the midpoint of a grid interval) and is maximized when a or b is 1 (at the edge of an interval). Since the latter case is examined below, we shall see that even when this term is maximized, it represents only a tiny (barely perceptible) modulation of the dominant second-order term.

Case 2

Now consider the case where the original and perturbed interpolating points lie in adjacent grid intervals. Let us allow the original interpolation point to lie on a grid point. Then, in this case, $i = j - 1$, $\bar{\xi} = x_{j-1}$ and

$$E = \left[-\frac{\bar{y}_j - \bar{y}_{j-1}}{h_j} + \frac{\bar{y}_{j-1} - \bar{y}_{j-2}}{h_{j-1}} + \frac{1}{6}(h_{j-1}\bar{y}''_{(j-2)^+} + 2h_{j-1}\bar{y}''_{(j-1)^-} + 2h_j\bar{y}''_{(j-1)^+} + h_j\bar{y}''_{j^-}) \right] \delta\xi + \frac{1}{2}\bar{y}''_{(j-1)^-}(\delta\xi)^2 + \frac{1}{6}\frac{\bar{y}''_{(j-1)^-} - \bar{y}''_{(j-2)^+}}{h_{j-1}}(\delta\xi)^3. \tag{23}$$

Because the deficient spline has discontinuous second derivatives, left and right function values are indicated by subscripts + and −. Substituting the various forms for y'' into the above yields the specific errors:

$$E_{\text{lin}}=\left[-\frac{\bar{y}_j-\bar{y}_{j-1}}{h_j}+\frac{\bar{y}_{j-1}-\bar{y}_{j-2}}{h_{j-1}}\right](\delta\xi), \tag{24}$$

$$E_{\text{cl}}=\frac{h^3}{6}\left[\frac{\bar{y}_{j-3}-4\bar{y}_{j-2}+6\bar{y}_{j-1}-4\bar{y}_j+\bar{y}_{j+1}}{h^4}\right](\delta\xi)$$
$$+\frac{1}{2}\left[\frac{\bar{y}_{j-2}-2\bar{y}_{j-1}+\bar{y}_j}{h^2}\right](\delta\xi)^2$$
$$+\frac{1}{6}\left[\frac{-\bar{y}_{j-3}+3\bar{y}_{j-2}-3\bar{y}_{j-1}+\bar{y}_j}{h^3}\right](\delta\xi)^3, \tag{25}$$

$$E_{\text{cs}}=\frac{1}{2}\bar{y}''_{j-1}(\delta\xi)^2+\frac{1}{6}\frac{\bar{y}''_{j-1}-\bar{y}''_{j-2}}{h_{j-1}}(\delta\xi)^3, \tag{26}$$

$$E_{\text{dcs}}=\frac{1}{2}[(6\bar{y}_{j-2}-6\bar{y}_{j-1}+2h_{j-1}\bar{y}'_{j-2}$$
$$+4h_{j-1}\bar{y}'_{j-1})]\left(\frac{\delta\xi}{h_{j-1}}\right)^2$$
$$+\frac{1}{2}[(4\bar{y}_{j-2}-4\bar{y}_{j-1}+2h_{j-1}\bar{y}'_{j-2}$$
$$+2h_{j-1}\bar{y}'_{j-1})]\left(\frac{\delta\xi}{h_{j-1}}\right)^3. \tag{27}$$

The subscripts lin, cl, cs and dcs refer to the linear, cubic Lagrange, cubic spline and deficient cubic spline interpolation schemes, respectively. Note that the + and − have been dropped for function values and for derivatives which are continuous. When the 1st derivative is evaluated using (36), the error for the deficient spline is

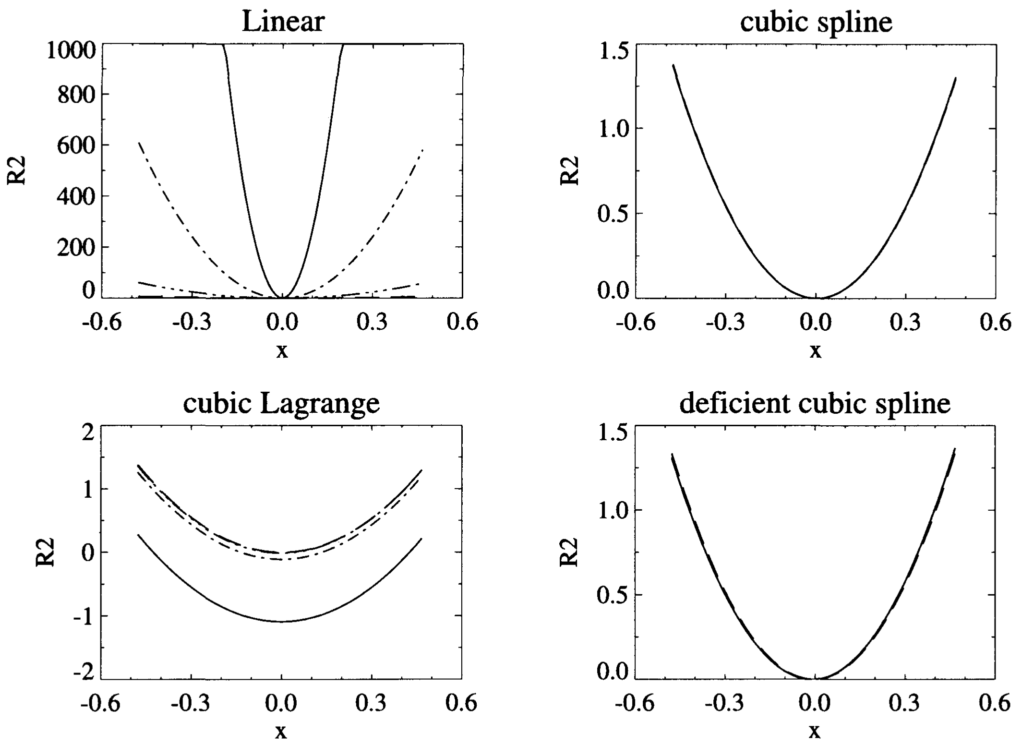

Fig. 4. The spatial distribution of the ratio, $R2=E/(\delta\xi)^2$, for various interpolation schemes. The interpolation point is $2.5\times10^{-6}\,\Delta x$ from the left edge of each grid interval. The solid, dash-dot, dash-dot-dot-dot and dashed curves correspond to $-\delta\xi=0.000005$, 0.00005, 0.0005, 0.005 respectively. The function which is being approximated is $y=x^4$.

$$E_{\mathrm{dcs}}=\frac{1}{2}\left[\frac{\bar{y}_{j-2}-2\bar{y}_{j-1}+\bar{y}_j}{h^2}+h\frac{-\bar{y}_{j-3}+3\bar{y}_{j-2}-3\bar{y}_{j-1}+\bar{y}_j}{h^3}\right](\delta\xi)^2+\frac{1}{2}\left[\frac{-\bar{y}_{j-3}+3\bar{y}_{j-2}-3\bar{y}_{j-1}+\bar{y}_j}{h^3}\right](\delta\xi)^3. \quad (28)$$

We may illustrate these errors using the same numerical procedure described above except that now the unperturbed interpolation point is a distance $2.5\times10^{-6}\,h$ from the left edge of each grid interval. Thus all perturbations lie in the left adjacent interval. Fig. 4 depicts the linearization errors for this case by illustrating $R2$ as a function of x. Because the curves are coincident for the spline interpolants, the error is second order and there is no linearization error. This is also verified by the absence of first order terms in (26) and (28). The linearization error for the linear and cubic Lagrange schemes is evidenced by the spread of the curves. The error is clearly worse for the linear case (note the change in scale). As expected from (24), the curves are simply the second derivative of $y(x)=x^4$ amplified by $\delta\xi$. For the cubic Lagrange case, substitution of the derivatives of $y(x)=x^4$ for the finite difference forms in (25) yields

$$R2\approx\frac{4h^3}{\delta\xi}+6\bar{\xi}^2.$$

Thus we expect to see a parabola vertically shifted downward by $4h^3/\delta\xi$, with the shift increasing in magnitude as $|\delta\xi|$ decreases and, indeed, this is what Fig. 4 depicts.

The effect of resolution on the linearization error for the cubic Lagrange scheme is depicted in Fig. 5. It is clear that as resolution improves, the curves are close together and the error is reduced. This

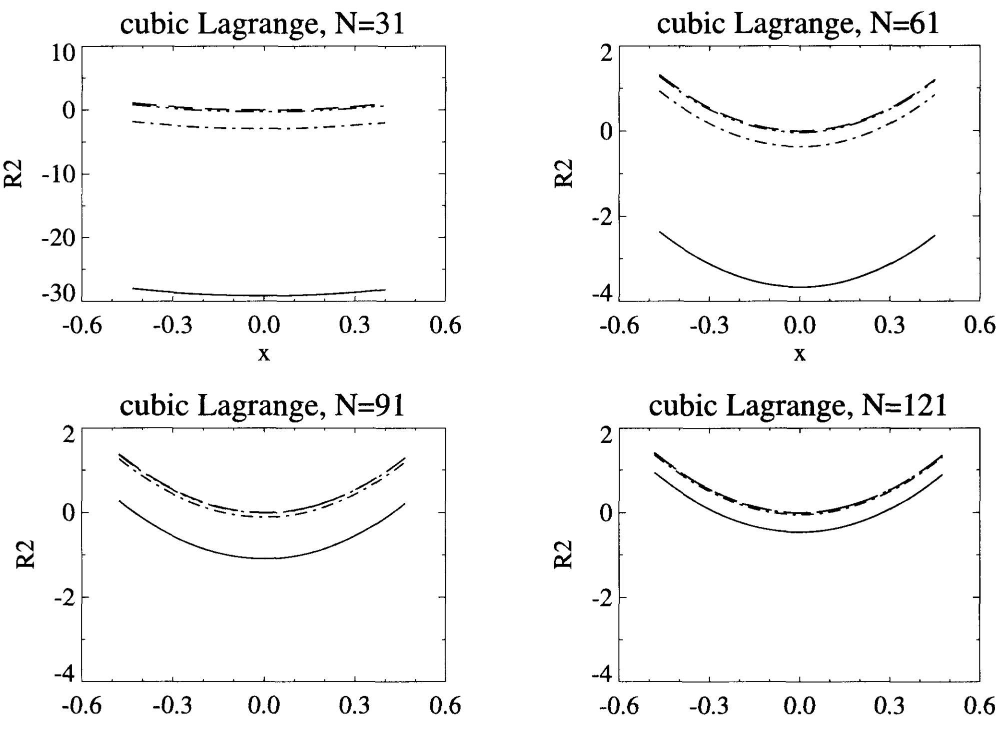

Fig. 5. The spatial distribution of the ratio, $R2=E/(\delta\xi)^2$, for the cubic Lagrange polynomial for various numbers of grid points, N, in the domain. The interpolation point is $2.5\times10^{-6}\,\Delta x$ from the left edge of each grid interval. The solid, dash-dot, dash-dot-dot-dot and dashed curves correspond to $-\delta\xi=0.000005$, 0.00005, 0.0005, 0.005 respectively. The function which is being approximated is $y=x^4$.

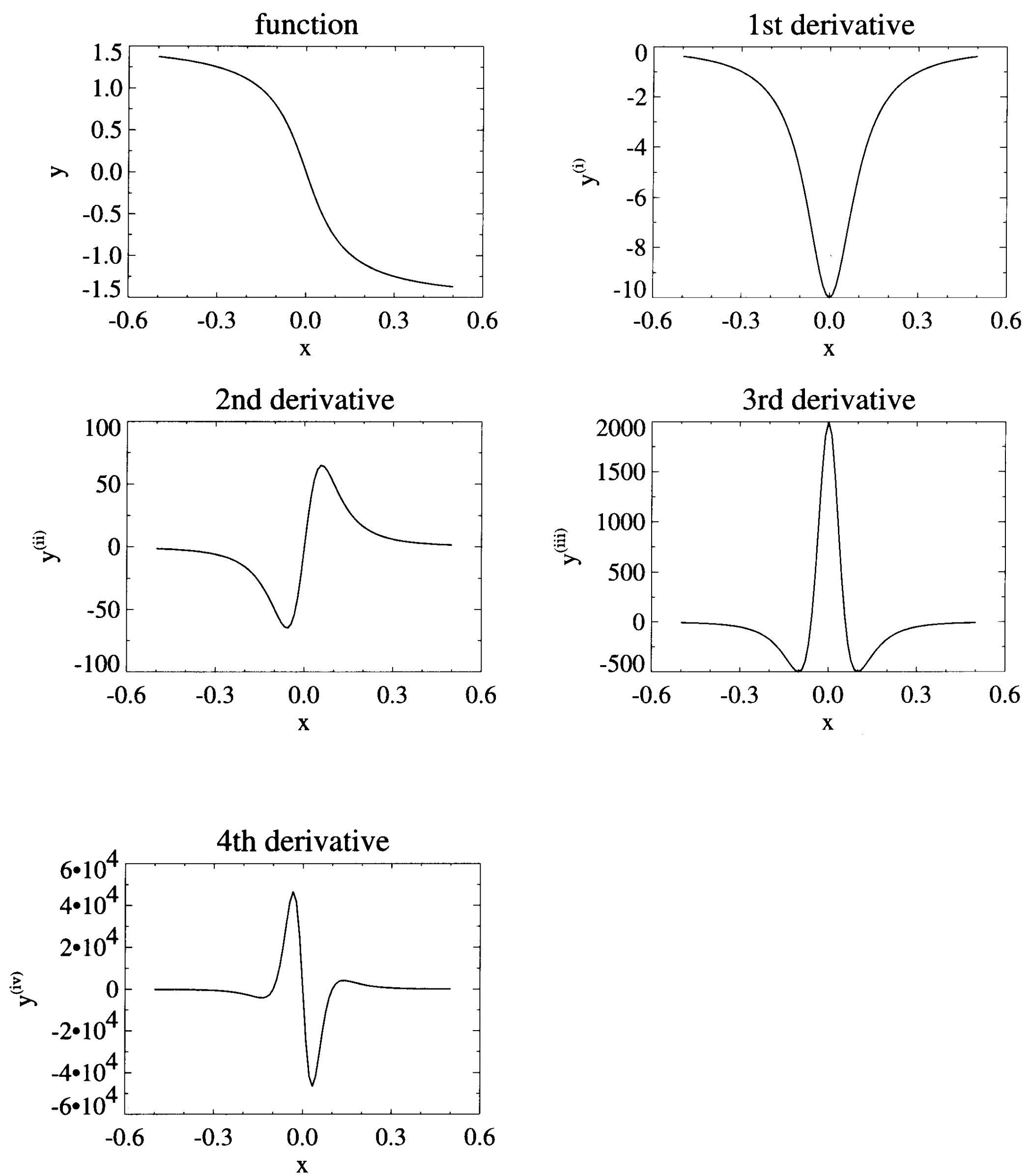

Fig. 6. The function $y = -\tan^{-1}(10x)$ and its 1st 4 derivatives.

resolution dependence is due to the presence of h^3 in the first order term of (25). Thus the higher the resolution, the smaller the magnitude of this term. The resolution dependence of the linearization error is understandable since a finite segment of a continuous function appears smoother with increasing resolution and the difference between left and right derivatives at grid points is smaller the smoother the function. Note that the error for linear interpolation is also resolution dependent because of the multiplication by h but an increase in resolution will have a smaller impact on this error than on that for the cubic Lagrange scheme.

Clearly, for this case, the cubic Lagrange has an incorrect TLE while the cubic spline does not. How large is this error in practice? The ratio, $R2$, while useful for determining the presence of linearization error, is not the best indicator of the size of the error. Better is the ratio, R, where

$$R = \frac{E}{|\delta\xi|}.$$

This ratio describes the relative error in the function to that of the interpolation point and is similar to that of Ménard (1994) except that we use the absolute value of $\delta\xi$. Now, if curves corresponding to various $\delta\xi$ coincide in a plot of R versus x, then there is a 1st-order error term and the TLE is incorrect. If there is no linearization error, each successive curve will diminish in amplitude as $\delta\xi$ diminishes. To illustrate the size of the linearization error in practice, consider the function $y = -\tan^{-1}(10x)$ which is illustrated in Fig. 6. This function might represent temperature or velocity in the cross-front direction typical of frontogenetic flows. Now with this ratio, $R2$, we can determine, at a glance, that the cubic splines have correct TLEs while the cubic Lagrange and linear schemes do not (Fig. 7). However, the same information presented differently (Fig. 8) shows similar behaviour of the curves for the cubic Lagrange and cubic splines: the curves are reduced in magnitude by a factor of 10 as $\delta\xi$ is reduced by a factor of 10. (Because of the presence of the first-order term in the TLE error for linear interpolation, all curves coincide.) For the cubic Lagrange scheme, based on (25) we expect

$$R \approx \left[\frac{h^3}{6}\,\bar{y}^{iv}(\bar{\xi}) + \frac{1}{2}\,\bar{y}''(\bar{\xi})(\delta\xi)\right]\frac{\delta\xi}{|\delta\xi|}.$$

Thus as $|\delta\xi|$ goes to zero, we see the relative error due to the first-order term. In Fig. 8, the ratio is zero for the smallest $|\delta\xi|$ (solid line), indicating the

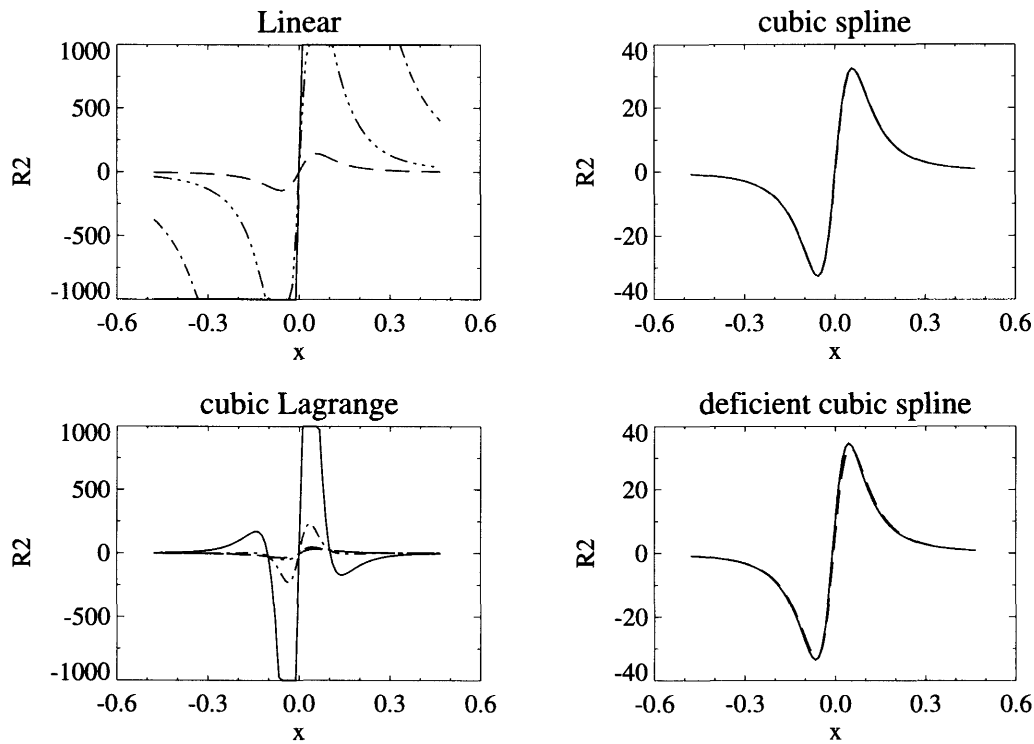

Fig. 7. The spatial distribution of the ratio, $R2 = E/(\delta\xi)^2$, for various interpolation schemes. The interpolation point is $2.5 \times 10^{-6}\,\Delta x$ from the left edge of each grid interval. The solid, dash-dot, dash-dot-dot-dot and dashed curves correspond to $-\delta\xi = 0.000005$, 0.00005, 0.0005, 0.005 respectively. The function which is being approximated is $y = -\tan^{-1}(10x)$.

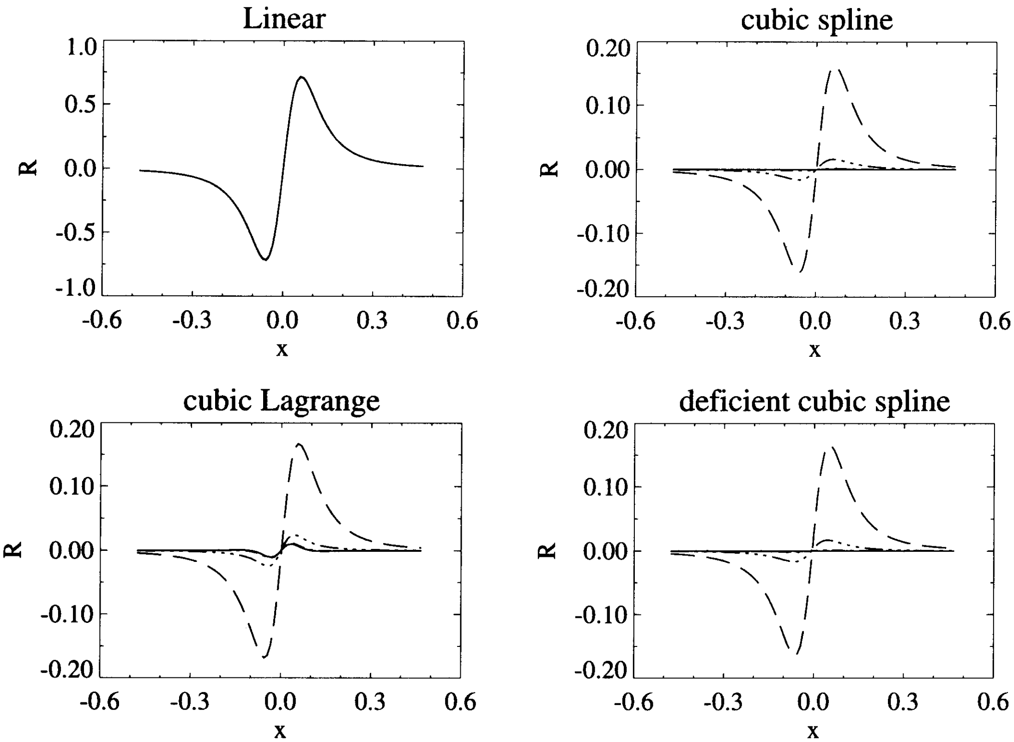

Fig. 8. The spatial distribution of the ratio, $R = E/|\delta\xi|$, for various interpolation schemes. The interpolation point is $2.5 \times 10^{-6}\,\Delta x$ from the left edge of each grid interval. The solid, dash-dot, dash-dot-dot-dot and dashed curves correspond to $-\delta\xi = 0.000005$, 0.00005, 0.0005, 0.005 respectively. The function which is being approximated is $y = -\tan^{-1}(10x)$.

splines have correct TLEs while the cubic Lagrange scheme is only slightly worse. Thus, the first order term is small compared to the second order term except when the perturbation size is very small. Therefore, in practice, and for moderate resolution and finite perturbations, the cubic Lagrange TLE error is not large and is perhaps tolerable.

In all cases considered so far, both the cubic spline and the deficient cubic spline have performed equally well. Both have continuous first derivatives at grid points. What then is the advantage, if any, of the cubic spline over the deficient spline as far as linearization error is concerned? To differentiate between these two schemes, we examine the quantity, $R3$, where

$$R3 = \frac{E}{(\delta\xi)^2} - \frac{1}{2}\frac{d^2\bar{y}}{dx^2}(\bar{\xi}). \qquad (29)$$

The 2nd term is evaluated using the matrix solution of (33) for the cubic spline case while the first term of (28) is used for the deficient spline case (in order to isolate the last two terms of (28)). $R3$ can be used to examine terms smaller than the dominant second-order term and is illustrated in Fig. 9 for both schemes. For the regular cubic spline we expect to see

$$R3 = \frac{1}{6}\frac{d^3\bar{y}}{dx^3}(\bar{\xi})\,\delta\xi.$$

Fig. 9 shows that we do indeed see the third derivative appearing in the third-order term. (This term is negative because $\delta\xi$ is negative.) It is also apparent that each curve is multiplied by $\delta\xi$. For the deficient spline, we expect the ratio $R3$ to be dominated by a constant term modulated by the third derivative due to the presence of an additional second-order term in (28). In Fig. 9, we see that this is indeed the case. All curves coincide except for the one corresponding to the largest $\delta\xi$ (dashed line). (This is because the 2nd-order term containing a third derivative is positive but since

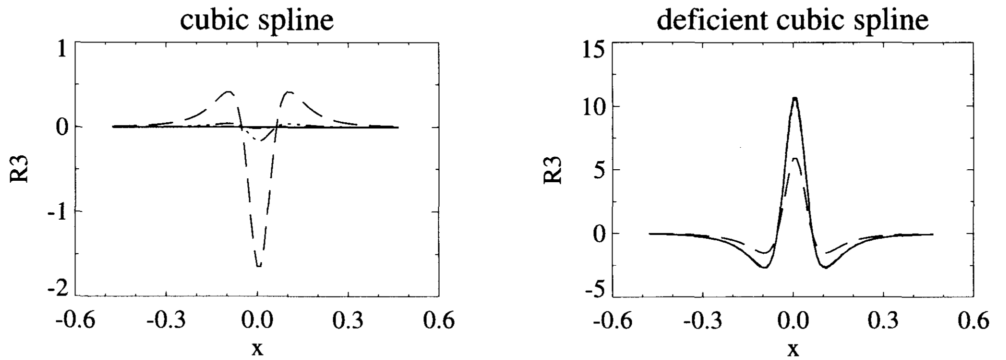

Fig. 9. The spatial distribution of the ratio, $R3$ for the cubic spline and the deficient cubic spline schemes. The interpolation point is $2.5 \times 10^{-6}\,\Delta x$ from the left edge of each grid interval. The solid, dash-dot, dash-dot-dot-dot and dashed curves correspond to $-\delta\xi = 0.000005$, 0.00005, 0.0005, 0.005 respectively. The function which is being approximated is $y = -\tan^{-1}(10x)$.

$\delta\xi$ is negative, the 3rd-order term is negative. Thus as $\delta\xi$ increases, the third-order term, being of opposite sign, cancels the second-order term to some extent, with the result that the curve corresponding to the largest $\delta\xi$ in Fig. 9 has a smaller value than the other curves.) Thus the lack of continuity of the second derivative for the deficient cubic spline has resulted in a larger second order term and therefore a smaller range of validity of the tangent linear hypothesis when compared to that for the cubic spline.

In summary, *for infinitesimal variations*, only the subcases considered are relevant and *an interpolation scheme has a correct TLE if and only if the interpolator has a continuous 1st derivative. For finite perturbations, the TLE can be incorrect whenever the perturbed and unperturbed interpolation points lie in different grid intervals.* Because higher-order schemes can better extrapolate function second derivatives, the range of validity of the TLM hypothesis, is larger for these schemes, as Li et al. (1993) have speculated. However, among interpolators of a given order, those with continuous first derivatives can further extend this range of validity.

5. The linearization error for semi-Lagrangian schemes

Thus far we have only examined the correctness of the TLE's of interpolation schemes. While this is of pedagogical importance, our ultimate interest lies in assessing the correctness of the TLM for semi-Lagrangian models. In the context of such models, the variation of the point of interpolation is due to the variation in the velocity so that $\delta\xi$ depends on $\delta y(x)$. If the latter variable is set to zero, as in the previous section, then there is no variation in the point of interpolation! Thus, to place these results in the meteorological context of nonlinear advection, it is necessary to add a further complication to our analysis of the linearization error for interpolation schemes: a nonzero variation of the unknown function which is related to the variation of the point of interpolation.

Let us consider the following experiment. We shall specify a known analytic function for $\bar{y}(x)$ as before and also a second known analytic function for $\delta y(x)$. These functions will be evaluated at grid points and supplied to the various interpolation schemes. The interpolating functions for $\bar{y}(x)$ and $\delta y(x)$ shall be called $P(x)$ and $Q(x)$, respectively. The general linearization error is

$$E = [P^i(\bar{\xi}) - P^j(\bar{\xi})] + \left[\frac{dP^i}{dx}(\bar{\xi}) - \frac{dP^j}{dx}(\bar{\xi})\right]\delta\xi$$

$$+ \frac{1}{2}\frac{d^2P^i}{dx^2}(\bar{\xi})(\delta\xi)^2 + \frac{1}{6}\frac{d^3P^i}{dx^3}(\bar{\xi})(\delta\xi)^3 + \cdots$$

$$+ [Q^i(\bar{\xi}) - Q^j(\bar{\xi})] + \frac{dQ^i}{dx}(\bar{\xi})\,\delta\xi$$

$$+ \frac{1}{2}\frac{d^2Q^i}{dx^2}(\bar{\xi})(\delta\xi)^2 + \frac{1}{6}\frac{d^3Q^i}{dx^3}(\bar{\xi})(\delta\xi)^3 + \cdots,$$

where i (j) is the index of the grid interval containing $\bar{\xi}+\delta\xi(\bar{\xi})$. Because $P(x)$ and $Q(x)$ are determined from the same interpolation scheme they are either both continuous or both discontinuous. Thus if $P^i(\bar{\xi})=P^j(\bar{\xi})$, then $Q^i(\bar{\xi})=Q^j(\bar{\xi})$ and the results of the previous section immediately follow. Therefore, a nonzero $\delta y(x)$ adds no new criteria for correctness of a TLE.

In addition to having nonzero $\delta y(x)$, a second complication introduced by semi-Lagrangian models over the interpolation process of the last section is the relationship between the variation of the point of interpolation and the variation of the function. This relationship is a consequence of advection and implies that $\delta\xi$ vary in space, instead of being a constant, as was assumed in Section 4. The ratio, R would then have its second-order term multiplied by $\delta\xi$, a spatially varying function. In contrast, $R2$, would have its 1st-order term divided by a spatially varying function. To provide a specific example, $\delta y(x)$ and $\delta\xi$ shall be defined as follows:

$$\delta y(x)=cg(x), \qquad \delta\xi=0.005cg(x)/g_{\max},$$

where $g_{\max}$ is the maximum of $|g(x)|$, and c is the perturbation amplitude which takes the values 1, 0.1, 0.01, 0.001. Note that because $\delta\xi$ changes sign, the original interpolation points were placed at the left (right) edge of the interval for $x\leqslant 0$ ($x>0$). Thus all perturbed interpolation points lie in adjacent intervals to the unperturbed ones except for the one near $x=0$. In this case, the linearization error simplifies to

$$E\approx\left[\frac{\mathrm{d}P^i}{\mathrm{d}x}(\bar{\xi})-\frac{\mathrm{d}P^j}{\mathrm{d}x}(\bar{\xi})\right]\delta\xi+\frac{\mathrm{d}Q^i}{\mathrm{d}x}(\bar{\xi})\,\delta\xi + \frac{1}{2}\frac{\mathrm{d}^2P^i}{\mathrm{d}x^2}(\bar{\xi})(\delta\xi)^2, \tag{30}$$

on retaining only terms of less than 3rd order in perturbation quantities. In the previous section, the function $\bar{y}(x)=-\tan^{-1}(10x)$ was used to represent the process of frontogenesis. If the rate of frontogenesis were in error, a perturbation might resemble the second derivative of $\bar{y}(x)$ (Fig. 6) in that the error would be zero away from the front but large and of opposite sign on either side of the front. Thus, for this case, we choose

$$g(x)=\frac{2x}{(1+100x^2)^2}.$$

The perturbation function resembles $\bar{y}''(x)$ but is reduced in amplitude by a factor of 1000.

Now when $\delta\xi$ varies in space but $\delta y(x)=0$, plots of $R2$ (not shown) are visually identical to those of Fig. 7 since the change in magnitude of c as $\delta\xi$ decreases dominates the effect of spatial variation of $\delta\xi$. On the other hand, the spatial variation of $\delta\xi$, is seen in plots of R for the cubic interpolation schemes since for those schemes, the 2nd-order term dominates.

When $\delta y(x)$ is defined as above and when the TLE is correct, we expect

$$E\approx\tfrac{1}{2}\,\bar{y}''(\bar{\xi})(\delta\xi)^2+cg'(\bar{\xi})\,\delta\xi.$$

Now the ratio, $R2=E/(\delta\xi)^2$, is not well behaved at $x=0$, so a more appropriate quantity is

$$R/c\approx\frac{1}{2}\,\bar{y}''(\bar{\xi})\,\frac{0.005\,|g(\bar{\xi})|}{g_{\max}}+g'(\bar{\xi})\,\frac{\delta\xi}{|\delta\xi|}.$$

Then if the TLM is correct, R/c is independent of perturbation amplitude, c, and all curves for various $\delta\xi$ will coincide. R/c is shown in Fig. 10. The curves increase in magnitude for the linear case, reflecting the first-order term which arises due to the difference in first derivatives of $P(x)$ at $\bar{\xi}$. Both splines have correct TLMs in this case since all curves coincide but instead of reflecting $\bar{y}''(x)$ as in Fig. 8, they reflect the dominant second term of (30). Thus they modulate as $\delta y'(x)$ or a 3rd derivative of $\bar{y}$ multiplied by the sign of $\delta\xi$. The cubic Lagrange scheme has an incorrect TLE, with the 2nd term of (30) dominating for $\delta\xi$ small. All of these results are in agreement with those of the previous section. Thus, the addition of a nonzero $\delta y(x)$ which is related to $\delta\xi$ only adds extra terms to E, making ratios more difficult to identify, but does not change the situations in which interpolation schemes have correct or incorrect TLEs.

In reality, not all interpolation points can land near grid points unless the velocity field is constant and the time step results in a Courant number which is an integer. Thus this experiment was specifically designed to expose the spatial variation of the linearization error. Experiments performed with a semi-Lagrangian model of Burgers' equation in a periodic domain reveal that, following intuition, at most only a few trajectory calculations will involve incorrect TLEs, for a given time step. Thus the TLM is correct for most trajectories most

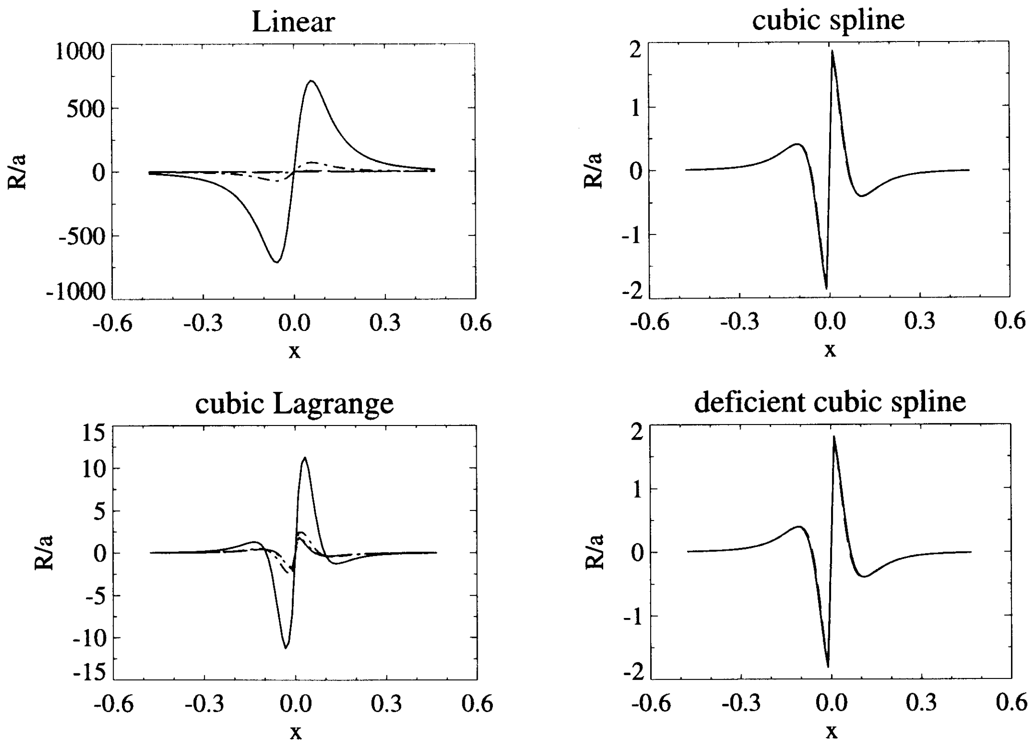

Fig. 10. The spatial distribution of the ratio R/c for various interpolation schemes. The interpolation point is $2.5 \times 10^{-6}\, \Delta x$ from the left (right) edge of each grid interval for $x \leqslant 0$ ($x > 0$). The solid, dash-dot, dash-dot-dot-dot and dashed curves correspond to $c = 0.001$, 0.01, 0.1, 1 respectively. The functions being interpolated are $y = -\tan^{-1}(10x)$ and $\delta y = c2x/(1 + 100x^2)^2$.

of the time even when using interpolators that have discontinuous first derivatives at meshpoints. This was also noted by Li et al. (1993) and explains the success of their 4DVAR experiments (Li et al., 1993; 1994).

6. Discussion

While the analyses presented here have been concerned with one-dimensional advection, it is clear that the results obtained apply equally well to the case of two or more dimensions. The general principle remains the same: for infinitesimal variations, a continuous interpolator and the equality of its derivatives normal to an interelemental face is required for a correct TLE. For interpolators which can be written as tensor products, the results of our one-dimensional analyses are applicable to each direction. For interpolators which cannot be so written, the linearization error needs to be calculated as above and numerical examples are also necessary to determine the size of the first order term in the event of an incorrect TLE.

Because the TLM obtained by differentiating the nonlinear code is not always correct, it would be useful if programs which generate adjoint code such as ODYSSEE (Rostaing et al., 1993) or AMC (Giering, 1993) would alert the user of a potential difficulty. Using our example of interpolation, the difficulty is recognized by the presence of a control variable in the argument of a truncation function for the case of uniform resolution. For variable resolution grids, this statement would be replaced by a search in which control variables appear in the condition of an IF sequence. However, not all IF statements having control variables in their conditions would pose a problem, so it is not clear whether such programs could automatically detect potential linearization problems.

Li et al. (1993) suggest that linearizing semi-Lagrangian schemes imposes an additional constraint on the time step. As we have shown, there is no linearization condition which constrains a perturbed interpolation point to lie in the same grid interval as the unperturbed one. If it does lie in a different grid interval, the TLE of the interpolation scheme may be incorrect and the range of validity of the tangent linear hypothesis reduced. However, this error is not catastrophic and can be controlled by carefully choosing an interpolation scheme. In fact, for infinitesimal perturbations (such as those used by 4DVAR) an interpolator need only have a continuous first derivative to be always correct. For finite perturbations, the variation of the point of interpolation does depend on the time step (Li et al., 1993, their eq. (3.6)) so that a larger time steps results in a reduced range of validity for a given TLM. However, the size of the time step is dictated by the magnitude of gradients in the flow (Appendix, Pudykiewicz et al., 1985) so that only when the gradients of the perturbed flow are significantly larger than those of the reference flow will a further *constraint* on the time step result. Because applications employing TLMs presume perturbations are small, such a restriction will not likely occur. Indeed, Li et al. (1994) used a time step of 30 minutes, a value commensurate with time steps currently in use for integrations of the global primitive equations, for their 4DVAR experiments with a three-dimensional semi-Lagrangian semi-implicit primitive equations model.

The desire for a TLE which is correct in most situations leads to the conclusion that schemes possessing continuous first derivatives everywhere are preferable. For example, the cubic spline, deficient or otherwise, is preferable to the cubic Lagrange scheme. By contrast, the issue of accuracy of the nonlinear model points to the equivalence of the cubic Lagrange and cubic spline schemes, both being superior to the deficient cubic spline. Thirdly, another consideration is the parallelizability of the code (Côté and Thomas, 1994). In this respect, local schemes are preferred to global ones so that the Lagrange polynomial and deficient cubic spline schemes are preferred to the cubic spline. Clearly these three issues of accuracy of the nonlinear model, the correctness and accuracy of the tangent linear model and parallelizability are at odds with one another so that the best solution may compromise at least one of these aspects. A possible compromise is the deficient cubic spline using a more accurate estimate of the first derivative than the one presented here. Such a scheme could be equivalent in accuracy to the cubic spline for the nonlinear model, produce a correct TLM in most circumstances and also be easily parallelizable. However, its drawback would be the discontinuity of the second derivative which might result in a reduced range of validity of the TLM hypothesis. Of course this speculation needs to be tested for a realistic model and for realistic meteorological cases. Nevertheless, it is now clear that when numerical weather prediction models are being designed, it is necessary to choose discretization operators which are not only accurate, but which lead to TLMs which are usually correct (and if not correct are only slightly inaccurate) and code which is parallelizable.

Finally, it is possible that the linearization error which was determined here for a single time step could accumulate in time for some meteorologically realistic conditions (such as that of a persistent zonal wind). The size of this type of accumulated error for varying interpolation schemes is currently being investigated. Such an accumulation of errors will be of particular concern to applications using the TLM with finite perturbations such as Kalman filtering, the generation of perturbations, or 4DVAR in its incremental incarnation (Courtier et al., 1994). 4DVAR in its "ideal" setting involves only infinitesimal variations for which the continuity of the first derivative of the interpolator determines the correctness of its TLE. Because the TLM is used to propagate increments in time in the incremental approach, finite but small variations are presumed. For finite variations, the TLE can be incorrect even for those interpolation schemes with continuous 1st derivatives. If such errors can accumulate then the appropriateness of semi-Lagrangian models to the incremental approach is a valid concern.

7. Conclusions

In this work, we have examined the correctness of tangent linear equations (TLEs) for general and specific piecewise interpolation schemes which arise in the context of models that employ a semi-Lagrangian treatment of advection. The correctness of a TLM refers to the fact that the linearization error (the difference between the nonlinear and linear variations) is second order in variations and this property can be objectively ascertained for a given model. The validity of the TLM hypothesis is ascertained when the ratio of the linearization error to its linear part is small, where "small" must necessarily be subjectively determined. It was demonstrated that straight differentiation of a nonlinear code does not always lead to a TLM which is correct in all circumstances.

The nondifferentiable aspect of the semi-Lagrangian advection scheme was traced to the interpolation process and the expressions for linearization error were derived for various interpolation schemes as well as for the general case. For infinitesimal variations, the general principle is that the TLE is correct if and only if an interpolating function and its 1st derivative are continuous everywhere. The continuity of second and higher order derivatives is irrelevant with respect to the correctness of the TLE. However, the validity of the tangent linear hypothesis depends on the value of the 2nd derivative of the polynomial appropriate for the perturbed point of interpolation *extrapolated* to the original point of interpolation. The better this extrapolation, the larger the range of validity of the tangent linear hypothesis. Thus higher-order interpolation schemes will have a larger range of validity for the TLE than lower order schemes and those with correct TLE's (continuous first derivatives) can have an even greater range of validity. For finite variations and piecewise interpolating functions, the TLE is incorrect whenever the reference and perturbed interpolation points lie in different grid intervals. Again, higher order schemes provide a larger range of perturbation magnitudes for which the tangent linear hypothesis is valid.

8. Acknowledgements

The 1st author is grateful to Dave Steenbergen for facilitating a five month visit to Montréal during which a collaboration was begun and this work was completed. Roger Daley and Pierre Gauthier provided valuable comments on a draft of this manuscript.

9. Appendix

Some examples of interpolation schemes

Various interpolation schemes may be obtained from the scheme, (18), simply by choosing how the $y''s$ are calculated. Four examples are given below.

1. Linear interpolation.

$$y_j'' = 0, \qquad j = 0, 1, ..., N. \tag{31}$$

2. Cubic Lagrange polynomial interpolation for a regular grid.

$$y_j'' = \frac{y_{j-1} - 2y_j + y_{j+1}}{h^2}, \qquad j = 1, ..., N-1. \tag{32}$$

The linearization error for the case of variable resolution cubic Lagrange interpolation is found below.

3. Cubic spline interpolation. By enforcing continuity of first derivatives at x_j we have

$$\frac{h_j}{6} y_{j-1}'' + \frac{h_j + h_{j+1}}{3} y_j'' + \frac{h_{j+1}}{6} y_{j+1}'' = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j}. \tag{33}$$

at interior points of the grid. Together with boundary conditions, this system of equations can be solved to simultaneously produce all $y_j''s$. For an exposition on splines, see Ahlberg et al. (1967).

4. Deficient cubic spline interpolation.

$$y_{(j-1)^+}'' = S''(x_{(j-1)^+}) = 6\frac{y_j - y_{j-1}}{h_j^2} - 4\frac{y_{j-1}'}{h_j} - 2\frac{y_j'}{h_j}, \tag{34}$$

$$y_{(j)^-}'' = S''(x_{(j)^-}) = -6\frac{y_j - y_{j-1}}{h_j^2} + 2\frac{y_{j-1}'}{h_j} + 4\frac{y_j'}{h_j}, \tag{35}$$

for $j=1, ..., N-1$ where

$$S(x)=a^2(2b+1)\,y_{j-1}+b^2(2a+1)\,y_j + h_j[a^2by'_{j-1}-b^2ay'_j].$$

The derivatives y'_j may be computed by any method that ensures continuity at grid points but here we shall use

$$y'_j=\frac{1}{h_j+h_{j+1}}\left[\frac{y_{j+1}-y_j}{h_{j+1}}h_j+\frac{y_j-y_{j-1}}{h_j}h_{j+1}\right],$$

which for the case of a regular grid simplifies to

$$y'_j=\frac{\bar{y}_{j+1}-\bar{y}_{j-1}}{2h}. \tag{36}$$

This case is also known as cubic Bessel interpolation (DeBoor, 1978). Note that because the deficient spline possesses discontinuities in the second derivative at grid points, we have explicitly referred to left $(j)^-$ and right $(j-1)^+$ derivatives in (34) and (35).

By substituting the above forms for the second derivatives (31) to (35), into (21) and assuming no variations in $y(x)$, we obtain the following errors pertaining to *Case 1:*

$$E_{\text{lin}}=0, \tag{37}$$

$$E_{\text{cl}}=\frac{1}{2}\left(\bar{a}\left[\frac{\bar{y}_{j-2}-2\bar{y}_{j-1}+\bar{y}_j}{h^2}\right] +\bar{b}\left[\frac{\bar{y}_{j-1}-2\bar{y}_j+\bar{y}_{j+1}}{h^2}\right]\right)(\delta\xi)^2 +\frac{1}{6}\left[\frac{-\bar{y}_{j-2}+3\bar{y}_{j-1}-3\bar{y}_j+\bar{y}_{j+1}}{h^3}\right](\delta\xi)^3, \tag{38}$$

$$E_{\text{dcs}}=[\bar{a}(-3\bar{y}_{j-1}+3\bar{y}_j-2h_j\bar{y}'_{j-1}-h_j\bar{y}'_j) +\bar{b}(3\bar{y}_{j-1}-3\bar{y}_j+h_j\bar{y}'_{j-1}+2h_j\bar{y}'_j)]\left(\frac{\delta\xi}{h_j}\right)^2 +[2\bar{y}_{j-1}-2\bar{y}_j+h_j\bar{y}'_j+h_j\bar{y}'_{j-1}]\left(\frac{\delta\xi}{h_j}\right)^3, \tag{39}$$

where the subscripts lin, cl and dcs refer to the linear, cubic Lagrange and deficient cubic spline interpolation schemes, respectively. Of course the linearization error for the cubic spline is given by (21). For the choice of y'_j given by (36), the error of the deficient spline is

$$E_{\text{dcs}}=\frac{1}{2}\left[\bar{a}\,\frac{\bar{y}_{j-2}-2\bar{y}_{j-1}+\bar{y}_j}{h^2} +\bar{b}\,\frac{\bar{y}_{j-1}-2\bar{y}_j+\bar{y}_{j+1}}{h^2}\right](\delta\xi)^2 +\frac{1}{2}\left[(-\bar{a}+\bar{b})\,h \times\frac{-\bar{y}_{j-2}+3\bar{y}_{j-1}-3\bar{y}_j+\bar{y}_{j+1}}{h^3}\right](\delta\xi)^2 +\frac{1}{2}\left[\frac{-\bar{y}_{j-2}+3\bar{y}_{j-1}-3\bar{y}_j+\bar{y}_{j+1}}{h^3}\right](\delta\xi)^3. \tag{40}$$

9.1. *Variable resolution cubic Lagrange interpolation*

The cubic Lagrange interpolation scheme may be expressed as (18) where

$$y''_{(j-1)^+}=\frac{A_jF''_{(j-1)^+}-B_{j-1}F''_{(j)^-}}{A_j-B_{j-1}}, \tag{41}$$

$$y''_{(j)^-}=\frac{(1-B_{j-1})\,F''_{(j)^-}-(1-A_j)\,F''_{(j-1)^+}}{A_j-B_{j-1}}, \tag{42}$$

$$A_j=\frac{1}{3}\left(2+\frac{h_{j+1}}{h_j}\right), \tag{43}$$

$$B_{j-1}=\frac{1}{3}\left(1-\frac{h_{j-1}}{h_j}\right), \tag{44}$$

$$F''_{(j)^-}=\frac{2}{h_j+h_{j+1}} \times\left[\frac{y_{j+1}-y_j}{h_{j+1}}-\frac{y_j-y_{j-1}}{h_j}\right], \tag{45}$$

$$F''_{(j-1)^+}=\frac{2}{h_{j-1}+h_j}\times\left[\frac{y_j-y_{j-1}}{h_j}-\frac{y_{j-1}-y_{j-2}}{h_{j-1}}\right]. \tag{46}$$

Assuming no variations in $y(x)$, the linearization error for *Case 1* where the perturbed and unperturbed interpolation points lie in the same grid interval, (21) becomes:

$$E_{\text{cl}}=\frac{1}{2}\left(\bar{a}_j\left[\frac{A_jF''_{(j-1)^+}-B_{j-1}F''_{(j)^-}}{A_j-B_{j-1}}\right]+\bar{b}_j\left[\frac{(1-B_{j-1})\,F''_{(j)^-}-(1-A_j)\,F''_{(j-1)^+}}{A_j-B_{j-1}}\right]\right)(\delta\xi)^2+\frac{1}{6}\left[\frac{F''_{(j)^-}-F''_{(j-1)^+}}{h_j(A_j-B_{j-1})}\right](\delta\xi)^3. \tag{47}$$

For *Case 2*, where the unperturbed interpolation point lies on grid point x_{j-1} and the perturbed interpolation point lies in the interval $[x_{j-2}, x_{j-1})$, the linearization error is

$$E_{\text{cl}}=\left[-\frac{\bar{y}_j-\bar{y}_{j-1}}{h_j}+\frac{\bar{y}_{j-1}-\bar{y}_{j-2}}{h_{j-1}}+\frac{1}{6}\left(h_{j-1}\times\frac{(3A_{j-1}-2)\,F''_{(j-2)^+}+(-3B_{j-2}+2)\,F''_{(j-1)^-}}{A_{j-1}-B_{j-2}}+h_j\frac{(3A_j-1)\,F''_{(j-1)^+}+(-3B_{j-1}+1)\,F''_{(j)^-}}{A_j-B_{j-1}}\right)\right]\times(\delta\xi)+\frac{1}{2}\times\left[\frac{(1-B_{j-2})\,F''_{(j-1)^-}-(1-A_{j-1})\,F''_{(j-2)^+}}{A_{j-1}-B_{j-2}}\right]\times(\delta\xi)^2+\frac{1}{6}\left[\frac{F''_{(j-1)^-}-F''_{(j-2)^+}}{h_{j-1}(A_{j-1}-B_{j-2})}\right](\delta\xi)^3. \tag{48}$$

It is easily verified that for the case of uniform resolution, (47) and (48) reduce to (38) and (25), respectively. Note that the TLM is correct for *Case 1* but incorrect for *Case 2*, as in the case of uniform resolution.

REFERENCES

Ahlberg, J. H., Nilson, E. N. and Walsh, J. L. 1967. *The theory of splines and their applications*. Academic Press.

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J. R., McNally, A. P., Kelly, G. A. and Courtier, P. 1994. Use of cloud-cleared radiances in three/four dimensional variational assimilation. *Q. J. R. Meteorol. Soc.* **120**, 627–653.

Chao, W. C. and Chang, L.-P. 1992. Development of a four-dimensional variational analysis system using the adjoint method at GLA. Part 1: Dynamics. *Mon. Wea. Rev.* **120**, 1661–1673.

Cohn, S. E. 1993. Dynamics of short-term univariate forecast error covariances. *Mon. Wea. Rev.* **121**, 3123–3149.

Côté, J. and Thomas, S. 1994. Parallel semi-Lagrangian advection on the sphere using PVM. *Scalable high-performance computing conference*, Knoxville, Tennessee. May 23–25, 1994. IEEE Computer Society, 470–477.

Courtier, P. and Talagrand, O. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. Part II. Numerical results. *Q. J. R. Meteorol. Soc.* **113**, 1329–1368.

Courtier, P., Derber, J., Errico, R., Louis, J.-F. and Vukicevic, T. 1993. Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. *Tellus* **45A**, 342–357.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367–1387.

De Boor, Carl, 1978. *A practical guide to splines*. Springer-Verlag, New York. 392 pp.

Derber, J. C. 1987. Variational four-dimensional analysis using quasi-geostrophic constraints. *Mon. Wea. Rev.* **115**, 1479–1502.

Giering, R. 1993. *The AMC compiler manual users*. Max-Planck Institut fur Meteorologie. Hamburg. Available from author at giering@dkz.d400.de.

Kuo, H.-C. and Williams, R. T. 1990. Semi-Lagrangian solutions to the inviscid Burgers equation. *Mon. Wea. Rev.* **118**, 1278–1288.

Lacarra, J.-F. and Talagrand, O. 1988. Short-range evolution of small perturbations in a barotropic model. *Tellus* **40A**, 81–95.

LeDimet, F. X. and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* **38A**, 97–110.

Lewis, J. M. and Derber, J. C. 1985. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus* **37A**, 309–322.

Li, Y., Navon, I. M., Courtier, P. and Gauthier, P. 1993. Variational data assimilation with a semi-Lagrangian semi-implicit global shallow-water equation model and its adjoint. *Mon. Wea. Rev.* **121**, 1759–1769.

Li, Y., Navon, I. M., Yang, W., Zou, X., Bates, J. R., Moorthi, S. and Higgins, R. W. 1994. Four-dimensional variational data experiments with a multilevel semi-Lagrangian semi-implicit general circulation model. *Mon. Wea. Rev.* **122**, 966–983.

Ménard, R. 1994. *Kalman filtering of Burgers' equation and its application to atmospheric data assimilation.* PhD thesis, McGill University. Stormy Weather Group scientific report MW-100. 211 pp.

Navon, I. M., Zou, X., Derber, J. and Sela, J. 1992. Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.* **120**, 1433–1446.

Pudykiewicz, R. Benoit and Staniforth, A. 1985. Preliminary results from a partial LRTAP model based on an existing meteorological forecast model. *Atmosphere-Ocean* **23**, 267–303.

Rabier, F. and Courtier, P. 1992. Four-dimensional assimilation in the presence of baroclinic instability. *Q. J. R. Meteorol. Soc.* **118**, 649–672.

Rabier, F., Courtier, P., Pailleux, J., Talagrand, O. and Vasiljevic, D. 1993. A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional analysis. *Q. J. R. Meteorol. Soc.* **119**, 845–880.

Robert, A., Yee, T. L. and Ritchie, H. 1985. A semi-Lagrangian and semi-implicit numerical integration scheme for multilevel atmospheric models. *Mon. Wea. Rev.* **113**, 388–394.

Rostaing, N., Dalmas, S. and Galligo, A. 1993. Automatic differentiation in Odyssee. *Tellus* **45A**, 558–568.

Staniforth, A. and Côté, J. 1991. Semi-Lagrangian integration schemes for atmospheric models, a review. *Mon. Wea. Rev.* **119**, 2206–2223.

Talagrand, O. and Courtier, P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. Part 1: Theory. *Q. J. R. Meteorol. Soc.* **113**, 1311–1328.

Thépaut, J.-N. and Courtier, P. 1991. Four-dimensional data assimilation using the adjoint of a multilevel primitive equation model. *Q. J. R. Meteorol. Soc.* **117**, 1225–1254.

Thépaut, J.-N., Vasiljevic, D., Courtier, P. and Pailleux, J. 1993. Variational assimilation of conventional meteorological observations with a multilevel primitive equation model. *Q. J. R. Meteorol. Soc.* **119**, 153–186.

Vukicevic, T. 1991. Nonlinear and linear evolution of initial forecast errors. *Mon. Wea. Rev.* **119**, 1602–1611.