A Method of Numerical Integration for Trajectories with Variational Equations

W. H. GOODYEAR*
IBM Federal Systems Division, Morris Plains, N.J.

A method of numerical integration is described which was developed for the computation of ballistic missile trajectories. But the technique is generally applicable to the initial-value problem for systems of ordinary differential equations. The associated variational equations are integrated simultaneously to provide the sensitivity coefficients of the solution. This facilitates the inclusion of an extra derivative in the numerical integration. The differences and advantages of the method as compared to standard integration techniques are discussed. A Newton-Raphson iterative solution of a single-step fourth-order Hermite corrector provides unconditional numerical stability. This very favorable property should make the method useful for stiff systems of ordinary differential equations. Polynomials are derived for interpolating the solution and its sensitivity coefficients, and also for extrapolating the solution to provide a predictor for the Hermite corrector. Integration errors and step-size control are also discussed. The end result is an accurate and efficient numerical integrator for solving ordinary differential equations and determining the sensitivity coefficients of the solution.

Introduction

THE equations of motion for a trajectory can be placed in the first-order form

$$ds/dt = f(s, t) \tag{1}$$

where s is the state vector of positions and velocities, and time t is the independent variable. The equations are said to be autonomous when

$$ds/dt = f(s) \tag{2}$$

so that f is not explicitly a function of t. The vector function f in Eq. (1) or its special case (2) is assumed to satisfy sufficient conditions so that the initial values

$$t = t_0 \Rightarrow s = s_0 \tag{3}$$

define a unique solution

$$s = \sigma(s_0, t_0, t) \tag{4}$$

in some interval of interest about t_0. In special cases tractable analytic formulas are available for the vector function σ, but in general this is not true.

For example, the equations of motion for a ballistic missile can be expressed in geocentric Earth-fixed Cartesian coordinates x, y, z with the z-axis through the North Pole and the x-axis through the Greenwich Meridian. In the standard first-order form, these equations are

$$\frac{d}{dt}\begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ -\mu x/r^3 + \omega^2 x + 2\omega\dot{y} + X \\ -\mu y/r^3 + \omega^2 y - 2\omega\dot{x} + Y \\ -\mu z/r^3 + Z \end{bmatrix} \tag{5}$$

$$r \equiv (x^2 + y^2 + z^2)^{1/2} > 0$$
$$v \equiv (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2} > 0$$

where the Earth has the gravitational constant μ (mass of the Earth times the universal gravitational constant) and the angular velocity ω (in radians per time unit about the z-axis). The position vector has components x, y, z and magnitude r while the velocity vector has components $\dot{x}, \dot{y}, \dot{z}$ and magnitude v. The acceleration vector has the gravitational components $-\mu(x, y, z)/r^3$, the centrifugal components $\omega^2 x, \omega^2 y, 0$, the coriolis components $2\omega\dot{y}, -2\omega\dot{x}, 0$, and perturbing components X, Y, Z due to effects such as gravitational anomalies, atmospheric drag, and rocket thrusting. When X, Y, Z are identically zero, a closed-form solution for x, y, z and $\dot{x}, \dot{y}, \dot{z}$ at any value of t in terms of the initial values x_0, y_0, z_0, and $\dot{x}_0, \dot{y}_0, \dot{z}_0$ at t_0 is readily available. If X, Y, Z depend only on $x, y, z, \dot{x}, \dot{y}, \dot{z}$ and do not involve explicit t, then Eqs. (5) are autonomous.

The square matrix

$$\partial s/\partial s_0 \equiv \partial\sigma(s_0, t_0, t)/\partial s_0 \tag{6}$$

of sensitivity coefficients (partial derivatives of elements of the vector s with respect to elements of the vector s_0) satisfies the associated variational equations

$$(d/dt)\,\partial s/\partial s_0 = [\partial f(s, t)/\partial s]\,\partial s/\partial s_0 \tag{7}$$

with initial conditions

$$t = t_0 \Rightarrow \frac{\partial s}{\partial s_0} = I \tag{8}$$

where I is the identity matrix. When a closed-form analytic function σ in (4) is available for the solution, direct differentiation using Eq. (6) gives the Jacobian matrix $\partial s/\partial s_0$ of the solution s with respect to the initial condition s_0. But if numerical integration is used to integrate Eqs. (1) or their special case (2), it may also be used to integrate the variational Eqs. (7).

The matrix $\partial s/\partial s_0$ is extremely important for many practical applications since the differential relation

$$\delta s = (\partial s/\partial s_0)[\delta s_0 - f(s_0, t_0)\cdot\delta t_0] + f(s, t)\cdot\delta t \tag{9}$$

may be used as a first-order approximation to give variations δs in the solution s produced by variations δs_0 in s_0, δt_0 in t_0, and δt in t. In the autonomous case (2) of Eqs. (1), Eq. (9) reduces to the simpler form

$$\delta s = (\partial s/\partial s_0)\delta s_0 + f(s)\cdot[\delta t - \delta t_0] \tag{10}$$

Since s and t may be treated as initial conditions and

$$s_0 = \sigma(s, t, t_0) \tag{11}$$

regarded as the solution at t_0,

$$\delta s_0 = (\partial s_0/\partial s)[\delta s - f(s, t)\cdot\delta t] + f(s_0, t_0)\cdot\delta t_0 \tag{12}$$

corresponds to Eq. (9) and reduces to

Presented as Paper 72-910 at the AIAA/AAS Astrodynamics Conference, Palo Alto, Calif., September 11–12, 1972; submitted September 18, 1972; revision received May 25, 1973. This work was sponsored under SAFEGUARD Contract DAHC-60-71-C-0005, Subcontract 601947, Bell Telephone Laboratories, Whippany, N.J.

Index categories: LV/M Trajectories; Earth-Orbital Trajectories; Lunar and Interplanetary Trajectories.

* Advisory Scientist/Engineer. Member AIAA.

$$\delta s_0 = (\partial s_0/\partial s)\,\delta s + f(s_0)\cdot[\delta t_0 - \delta t] \tag{13}$$

for the autonomous case (2). Here

$$\partial s_0/\partial s \equiv \partial\sigma(s,t,t_0)/\partial s = [\partial s/\partial s_0]^{-1} \tag{14}$$

follows from substitution of (9) into (12) since δs_0 and δt_0 are arbitrary. Thus $\partial s_0/\partial s$ is the matrix inverse of $\partial s/\partial s_0$.

Extra Derivative and Numerical Integration

The method of integration to be described utilizes the extra derivative

$$\dot{f}(s,t) \equiv \frac{df(s,t)}{dt} = \frac{\partial f(s,t)}{\partial s} f(s,t) + \frac{\partial f(s,t)}{\partial t} \tag{15}$$

of $f(s,t)$. The vector $\partial f(s,t)/\partial t$ of partials is identically zero in the case of autonomous differential equations (2), and is easily computed in many important examples of the nonautonomous differential equations (1). Thus it will be assumed that $\partial f(s,t)/\partial t$ is easily obtained. Then $\dot{f}(s,t)$ from Eq. (15) is easily obtained when the variational equations (7) are integrated since the square matrix $\partial f(s,t)/\partial s$ is readily available. Since the use of the extra derivative (15) is known to offer advantages for numerical integration, it is efficient and economical to utilize $\dot{f}(s,t)$ from (15) when the variational equations (7) are also integrated numerically.

The corrector

$$s_{n+1} = s_n + [(f_n + f_{n+1})/2]h + [(\dot{f}_n - \dot{f}_{n+1})/12]h^2 \tag{16}$$

will be used to determine the successive approximations

$$s_{n+1} \doteq \sigma(s_n, t_n, t_{n+1}) \tag{17}$$

to the true solution $\sigma(s_0, t_0, t_{n+1})$ at the successive times $t_n < t_{n+1}$ where $n = 0, 1, 2, \ldots$. Forward integration for which the step-size

$$h \equiv t_{n+1} - t_n \tag{18}$$

is positive will be assumed, but backward integration with $h < 0$ and $t_n > t_{n+1}$ is completely analogous. The compact notation

$$\begin{aligned} f_n &\equiv f(s_n, t_n) \qquad & f_{n+1} &\equiv f(s_{n+1}, t_{n+1}) \\ \dot{f}_n &\equiv \dot{f}(s_n, t_n) \qquad & \dot{f}_{n+1} &\equiv \dot{f}(s_{n+1}, t_{n+1}) \end{aligned} \tag{19}$$

has been used in the corrector (16) which is, in general, an implicit equation in s_{n+1} which must be solved at each step of the numerical integration. At any given step, s_n and t_n are known from the previous step (the initial condition when $n = 0$) and t_{n+1} is determined from a given value for h.

Ralston[5] recommends the corrector (16) rather than Milne's method or Hamming's method when the extra derivative $\dot{f}(s,t)$ is easily computed. Lapidus and Seinfeld[2] point out that this corrector has a small truncation error and very favorable stability properties. Loscalzo[4] remarks that the formula (16) was derived by Hermite[1] in 1877, but that this early work is seldom referenced. He also discusses the history of Eq. (16) and establishes that this corrector has unconditional numerical stability (A-stability). Liniger and Willoughby[3] also consider the corrector (16) in deriving numerical integration formulas for stiff systems of ordinary differential equations. The main purpose here is to show how use of the corrector (16) is related to the simultaneous numerical integration of the variational equations (7).

If second-order partial derivatives of $f(s,t)$ with respect to s and t are neglected, the approximation

$$\partial\dot{f}(s,t)/\partial s \doteq [\partial f(s,t)/\partial s]\,\partial f(s,t)/\partial s \tag{20}$$

may be used to differentiate (16) with s_{n+1} considered as a function of s_n. This gives

$$\frac{\partial s_{n+1}}{\partial s_n} = \left[I - \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{h}{2} + \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{h^2}{12}\right]^{-1} \cdot \left[I + \frac{\partial f_n}{\partial s_n}\frac{h}{2} + \frac{\partial f_n}{\partial s_n}\frac{\partial f_n}{\partial s_n}\frac{h^2}{12}\right] \tag{21}$$

as an approximation for the partials $\partial\sigma(s_n, t_n, t_{n+1})/\partial s_n$ of the solution $\sigma(s_n, t_n, t_{n+1})$ at t_{n+1} with s_n at t_n considered as initial conditions. Thus (21) may be applied at each integration step after Eq. (16) has been solved to determine s_{n+1}. But Eq. (21) can also be interpreted as a corrector for the variational equations

$$\frac{d}{dt}\frac{\partial s}{\partial s_n} = \frac{\partial f(s,t)}{\partial s}\frac{\partial s}{\partial s_n} \tag{22}$$

with initial conditions

$$t = t_n \Rightarrow \partial s/\partial s_n = I \tag{23}$$

at each step of the integration. Neglect of the second-order partials of $f(s,t)$ with respect to s and t gives the approximation

$$\frac{d^2}{dt^2}\frac{\partial s}{\partial s_n} \doteq \frac{\partial f(s,t)}{\partial s}\frac{\partial f(s,t)}{\partial s}\frac{\partial s}{\partial s_n} \tag{24}$$

for the extra derivative of Eq. (22). Use of the same type of corrector as (16) for (22) then gives the same result (21).

With either interpretation

$$\partial s_0/\partial s_0 = I \tag{25}$$

so the chain rule in the form

$$\partial s_{n+1}/\partial s_0 = (\partial s_{n+1}/\partial s_n)\,\partial s_n/\partial s_0 \tag{26}$$

can be used at each integration step to approximate the partials $\partial\sigma(s_0, t_0, t_{n+1})/\partial s_0$. Also, the inverse of Eq. (21) gives

$$\frac{\partial s_n}{\partial s_{n+1}} = \left[\frac{\partial s_{n+1}}{\partial s_n}\right]^{-1} = \left[I + \frac{\partial f_n}{\partial s_n}\frac{h}{2} + \frac{\partial f_n}{\partial s_n}\frac{\partial f_n}{\partial s_n}\frac{h^2}{12}\right]^{-1} \cdot \left[I - \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{h}{2} + \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{h^2}{12}\right] \tag{27}$$

as an approximation for the partials $\partial\sigma(s_{n+1}, t_{n+1}, t_n)/\partial s_{n+1}$ of the solution $\sigma(s_{n+1}, t_{n+1}, t_n)$ at t_n with s_{n+1} and t_{n+1} considered as initial conditions. Thus the chain rule in the form

$$\partial s_0/\partial s_{n+1} = (\partial s_0/\partial s_n)\,\partial s_n/\partial s_{n+1} \tag{28}$$

can be used to approximate the inverse of $\partial\sigma(s_0, t_0, t_{n+1})/\partial s_0$. This seems preferable to a direct inversion of $\partial s_{n+1}/\partial s_0$ since h can always be made sufficiently small so that the inverses in (21) and (27) are well-defined.

The Corrector's Solution and Numerical Stability

A standard procedure for solving the corrector (16) for s_{n+1} is to use an approximation s^*_{n+1} in place of s_{n+1} to evaluate the right side which is then treated as a new approximation to iterate the procedure. But this approach has convergence difficulties relative to a Newton-Raphson type of iterative procedure which is very convenient to use here. Defining

$$s_{n+1} \doteq s^*_{n+1} + \Delta s^*_{n+1} \tag{29}$$

gives the first-order approximations

$$\begin{aligned} f(s_{n+1}, t_{n+1}) &\doteq f(s^*_{n+1}, t_{n+1}) + \frac{\partial f(s^*_{n+1}, t_{n+1})}{\partial s^*_{n+1}}\Delta s^*_{n+1} \\ &\equiv f^*_{n+1} + \frac{\partial f^*_{n+1}}{\partial s^*_{n+1}}\Delta s^*_{n+1} \\ \dot{f}(s_{n+1}, t_{n+1}) &\doteq \dot{f}(s^*_{n+1}, t_{n+1}) + \frac{\partial\dot{f}(s^*_{n+1}, t_{n+1})}{\partial s^*_{n+1}}\Delta s^*_{n+1} \\ &\equiv \dot{f}^*_{n+1} + \frac{\partial f^*_{n+1}}{\partial s^*_{n+1}}\frac{\partial f^*_{n+1}}{\partial s^*_{n+1}}\Delta s^*_{n+1} \end{aligned} \tag{30}$$

where the last equation utilizes (20). Substitution of Eq. (30) into the corrector (16) gives

$$\Delta s^*_{n+1} = \left[I - \frac{\partial f^*_{n+1}}{\partial s^*_{n+1}}\frac{h}{2} + \frac{\partial f^*_{n+1}}{\partial s^*_{n+1}}\frac{\partial f^*_{n+1}}{\partial s^*_{n+1}}\frac{h^2}{12}\right]^{-1} \cdot \left[s_n + \frac{f_n + f^*_{n+1}}{2}h + \frac{\dot{f}_n - \dot{f}^*_{n+1}}{12}h^2 - s^*_{n+1}\right] \tag{31}$$

for the correction Δs^*_{n+1} to be added to s^*_{n+1}. If the successive approximations obtained by treating $s^*_{n+1}+\Delta s^*_{n+1}$ as a new approximation converge, the last term in brackets and thus the successive corrections Δs^*_{n+1} approach zero. Two applications of Eq. (31) requiring two evaluations of f^*_{n+1}, $\partial f(s^*_{n+1}, t^*_{n+1})/\partial s^*_{n+1}$, and $\partial f(s^*_{n+1}, t_{n+1})/\partial t_{n+1}$ should be satisfactory when the step-size h is sufficiently small. The reason that (31) is so convenient here is that the same matrix inverse occurs in the corrector (21) for the variational equations.

Numerical stability is conventionally defined in terms of a particular integrator's performance for the particular case

$$ds/dt = As \tag{32}$$

of Eqs. (2) and thus of (1). The square matrix A is constant and real so that

$$\begin{aligned} f(s,t) &= As \\ \partial f(s,t)/\partial s &= A \\ \dot{f}(s,t) &= A^2 s \end{aligned} \tag{33}$$

occur in the preceding integration formulas. Substitution of Eq. (33) into the corrector (16) gives

$$s_{n+1} = [I - Ah/2 + A^2h^2/12]^{-1} \cdot [I + Ah/2 + A^2h^2/12]s_n \tag{34}$$

as the solution. But (31) gives this solution (34) directly for any approximation s^*_{n+1}. Thus there are no convergence difficulties in application of the iterative solution (31) to the particular case (32) of Eqs. (1), and hopefully this favorable property carries over to the case where $f(s,t)$ is nonlinear.

The correct solution for Eq. (32) at each integration step is

$$s_{n+1} = e^{Ah}s_n \equiv [I + Ah + A^2h^2/2! + A^3h^3/3! + \cdots]s_n \tag{35}$$

so the solution (34) for an integration step is equivalent to

$$e^{Ah} \doteq [I - Ah/2 + A^2h^2/12]^{-1}[I + Ah/2 + A^2h^2/12] \tag{36}$$

which is the second-order diagonal Padé approximation to the matric exponential. In the particular case where A is diagonalizable, there exists a square nonsingular (complex in general) matrix R such that

$$RAR^{-1} = \Lambda = \text{diagonal}\,\{\lambda_i\} \tag{37}$$

where the diagonal elements λ_i (complex in general) of the diagonal matrix Λ are the corresponding eigenvalues. Then

$$\begin{aligned} s_{n+1} &= R^{-1} \cdot \text{diag}\,\{e^{\lambda_i h}\} \cdot Rs_n \\ &\doteq R^{-1} \cdot \text{diag}\left\{\frac{1+\lambda_i h/2 + \lambda_i^2h^2/12}{1-\lambda_i h/2 + \lambda_i^2h^2/12}\right\} \cdot Rs_n \end{aligned} \tag{38}$$

follows from Eqs. (35–37).

Therefore

$$\begin{aligned} s_{n+1} &= R^{-1} \cdot \text{diag}\,\{e^{\lambda_i(t_{n+1}-t_0)}\} \cdot Rs_0 \\ &\doteq R^{-1} \cdot \left[\prod^{n} \frac{1+\lambda_i h/2 + \lambda_i^2h^2/12}{1-\lambda_i h/2 + \lambda_i^2h^2/12}\right] \cdot Rs_0 \end{aligned} \tag{39}$$

gives the true solution for s_{n+1} and its approximation produced by the integration. But when

$$\text{Real}\,\{\lambda_i h\} < 0 \;\forall\; i \Rightarrow s_{n+1} \to 0 \quad \text{as} \quad n \to \infty \tag{40}$$

it also follows that

$$\begin{aligned} \text{Real}\,\{\lambda_i h\} < 0 &\Rightarrow \left|\frac{1+\lambda_i h/2 + \lambda_i^2h^2/12}{1-\lambda_i h/2 + \lambda_i^2h^2/12}\right| < 1 \\ &\Rightarrow \left|\prod^{n} \frac{1+\lambda_i h/2 + \lambda_i^2h^2/12}{1-\lambda_i h/2 + \lambda_i^2h^2/12}\right| \to 0 \quad \text{as} \quad n \to \infty \end{aligned} \tag{41}$$

so the solution and its approximation both approach zero. But this is the definition of unconditional numerical stability (A-stability), so the method of integration described here has this very favorable property, which applies not only to the corrector (16) itself, but to its Newton-Raphson solution (29) and (31) as well.

Interval Approximations for the Solutions

The corrector (16) is based on a fourth-order polynomial approximation

$$s \doteq c_0 + c_1(t-t_n) + c_2(t-t_n)^2 + c_3(t-t_n)^3 + c_4(t-t_n)^4 \tag{42}$$

and its two derivatives

$$\begin{aligned} f(s,t) &\doteq c_1 + 2c_2(t-t_n) + 3c_3(t-t_n)^2 + 4c_4(t-t_n)^3 \\ \dot{f}(s,t) &\doteq 2c_2 + 6c_3(t-t_n) + 12c_4(t-t_n)^2 \end{aligned} \tag{43}$$

over the interval from t_n to t_{n+1}. Derivations of the corrector (16) often omit explicit determination of the coefficients c_0, c_1, c_2, c_3, c_4, but they will be obtained here so that (42) may be used to interpolate for values of s at arbitrary values of t inside the interval $[t_n, t_{n+1}]$. The condition that (42) and (43) agree with the solution $\sigma(s_n, t_n, t)$ through s_n, t_n and its two derivatives at $t = t_n$ gives

$$c_0 = s_n, \quad c_1 = f_n, \quad c_2 = \dot{f}_n/2 \tag{44}$$

for three of the coefficients.

Although s_{n+1} must be determined from s_n, t_n, and t_{n+1}, the similar condition that (42) and (43) agree with the solution $\sigma(s_{n+1}, t_{n+1}, t)$ through s_{n+1}, t_{n+1} at $t = t_{n+1}$ gives

$$s_{n+1} = s_n + f_n h + \dot{f}_n h^2/2 + c_3h^3 + c_4h^4 \tag{45}$$

and

$$\begin{aligned} f_{n+1} &= f_n + \dot{f}_n h + 3c_3h^2 + 4c_4h^3 \\ \dot{f}_{n+1} &= \dot{f}_n + 6c_3h + 12c_4h^2 \end{aligned} \tag{46}$$

where (44) has been used. The three equations (45) and (46) then determine s_{n+1} with c_3 and c_4 at each step of the integration. Solution of (46) for c_3 and c_4 then gives

$$\begin{aligned} c_3 &= (f_{n+1} - f_n)/h^2 - (\dot{f}_{n+1} + 2\dot{f}_n)/3h \\ c_4 &= (f_n - f_{n+1})/2h^3 + (\dot{f}_n + \dot{f}_{n+1})/4h^2 \end{aligned} \tag{47}$$

for c_3 and c_4 as functions of s_{n+1}.

Substitution from (47) into (45) then gives the Hermite corrector (16). This derivation is important here because (44) and (47) determine the coefficients c_0, c_1, c_2, c_3, c_4 in (42) after the corrector (16) has been solved to determine s_{n+1}. Thus (42) can be used to interpolate s for any value of t inside the interval $[t_n, t_{n+1}]$, which in turn gives any function of s for the corresponding value of t. Also, values of t at which given functions of s have specified values can be determined by regarding the polynomial (42) as the solution within the interval $[t_n, t_{n+1}]$.

The polynomial (42) can also of course be used for extrapolation outside the interval, but will be defined relative to t_{n+1} for this purpose. This gives

$$s = c_0^* + c_1^*(t-t_{n+1}) + c_2^*(t-t_{n+1})^2 + c_3^*(t-t_{n+1})^3 + c_4^*(t-t_{n+1})^4 \tag{48}$$

where c_0^*, c_1^*, c_2^*, c_3^*, c_4^* are the corresponding coefficients. An approach similar to the preceding then gives

$$c_0^* = s_{n+1}, \quad c_1^* = f_{n+1}, \quad c_2^* = \dot{f}_{n+1}/2 \tag{49}$$

for the first three coefficients and

$$\begin{aligned} c_3^* &= (f_n - f_{n+1})/h^2 + (\dot{f}_n + 2\dot{f}_{n+1})/3h \\ c_4^* &= (f_n - f_{n+1})/2h^3 + (\dot{f}_n + \dot{f}_{n+1})/4h^2 \end{aligned} \tag{50}$$

so $c_4^* = c_4$ but $c_3^* \neq c_3$.

At the conclusion of the current step, the new step is started with $n+1 \to n$ and perhaps a different value of h. Then Eq. (49) transforms back to (44) and (48) transforms to

$$s^*_{n+1} = c_0 + c_1h + c_2h^2 + c_3^*h^3 + c_4^*h^4 \tag{51}$$

for the approximation s^*_{n+1} in (31). In the extrapolation formula (51), it is understood that c_3^* and c_4^* have been determined from (50) at the end of the previous step, where $c_3^* = c_4^* = 0$ may be used at the start of the integration if no better approximations are available. The formula (51) then functions as a predictor for the iterative solution (31) of the corrector since it gives a starting approximation s^*_{n+1} for s_{n+1}. It can be shown that (51) is equivalent to standard predictor formulas for the corrector (16), but this is not important here.

After the corrector has been solved to determine s_{n+1} at t_{n+1}, Eqs. (21) and (27) determine $\partial s_{n+1}/\partial s_n$ and $\partial s_n/\partial s_{n+1}$, and then (26) and (28) may be used to determine $\partial s_{n+1}/\partial s_0$ and $\partial s_0/\partial s_{n+1}$. But some method of interpolation is required to determine $\partial s/\partial s_n$ and thus

$$\begin{aligned}\partial s/\partial s_0 &= (\partial s/\partial s_n)\,\partial s_n/\partial s_0\\ \partial s_0/\partial s &= (\partial s_0/\partial s_n)[\partial s/\partial s_n]^{-1}\end{aligned} \tag{52}$$

at arbitrary values of t within the interval. However c_0, c_1, c_2, c_3, c_4 are functions of s_n and s_{n+1} while s_{n+1} is in turn a function of s_n through the corrector (16). Therefore

$$\frac{\partial s}{\partial s_n} = \frac{\partial c_0}{\partial s_n} + \frac{\partial c_1}{\partial s_n}(t-t_n) + \frac{\partial c_2}{\partial s_n}(t-t_n)^2 + \frac{\partial c_3}{\partial s_n}(t-t_n)^3 + \frac{\partial c_4}{\partial s_n}(t-t_n)^4 \tag{53}$$

follows from differentiation of (42) with respect to s_n. Here differentiation of (44) gives

$$\begin{aligned}\partial c_0/\partial s_n &= I\\ \partial c_1/\partial s_n &= \partial f_n/\partial s_n\\ \partial c_2/\partial s_n &= \tfrac{1}{2}(\partial f_n/\partial s_n)\,\partial f_n/\partial s_n\end{aligned} \tag{54}$$

while differentiation of (47) gives

$$\begin{aligned}\frac{\partial c_3}{\partial s_n} &= \frac{1}{h^2}\left[\frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial s_{n+1}}{\partial s_n} - \frac{\partial f_n}{\partial s_n}\right] - \frac{1}{3h}\left[\frac{\partial f_{n+1}}{d s_{n+1}}\frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial s_{n+1}}{\partial s_n} + 2\frac{\partial f_n}{\partial s_n}\frac{\partial f_n}{\partial s_n}\right]\\ \frac{\partial c_4}{\partial s_n} &= \frac{1}{2h^3}\left[\frac{\partial f_n}{\partial s_n} - \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial s_{n+1}}{\partial s_n}\right] + \frac{1}{4h^2}\left[\frac{\partial f_n}{\partial s_n}\frac{\partial f_n}{\partial s_n} + \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial s_{n+1}}{\partial s_n}\right]\end{aligned} \tag{55}$$

where the approximation (20) has been used.

Therefore (54) and (55) may be used in order to interpolate $\partial s/\partial s_n$ from (53) and thus obtain $\partial s/\partial s_0$ and $\partial s_0/\partial s$ from (52). Another interpretation of these formulas is that the matrices of partials in (53) are merely matric coefficients in a polynomial approximation for $\partial s/\partial s_n$. The condition that this polynomial satisfy the variational equations (22) with initial conditions (23) and the extra derivative (24) at t_n and t_{n+1} then gives (54) and (55). If (54) and (55) are substituted into (53) with $t = t_{n+1}$, the result is the corrector (21) for the variational equations (22). With either interpretation, (42) and (53) with (52) give s, $\partial s/\partial s_0$, and $\partial s_0/\partial s$ for arbitrary values of t in the interval $[t_n, t_{n+1}]$.

Integration Errors and Step-size Control

At each step of the integration, let

$$\begin{aligned}\sigma_n &\equiv \sigma(\sigma_0, t_0, t_n)\\ \sigma_{n+1} &\equiv \sigma(\sigma_0, t_0, t_{n+1})\end{aligned} \tag{56}$$

be the true values of the solution at t_n and t_{n+1} as distinguished from the approximations s_n and s_{n+1} produced by the integration. For generality it will be assumed that the true value σ_0 of the initial condition differs from the given value s_0. The integration error is defined as

$$\begin{aligned}\delta s_n &\equiv \sigma_n - s_n\\ \delta s_{n+1} &\equiv \sigma_{n+1} - s_{n+1}\end{aligned} \tag{57}$$

at t_n and t_{n+1}. Thus the approximations

$$\begin{aligned}f(\sigma_n, t_n) &= f(s_n + \delta s_n, t_n) \doteq f_n + \frac{\partial f_n}{\partial s_n}\delta s_n\\ f(\sigma_{n+1}, t_{n+1}) &= f(s_{n+1} + \delta s_{n+1}, t_{n+1}) \doteq f_{n+1} + \frac{\partial f_{n+1}}{\partial s_{n+1}}\delta s_{n+1}\\ \dot f(\sigma_n, t_n) &= \dot f(s_n + \delta s_n, t_n) \doteq \dot f_n + \frac{\partial f_n}{\partial s_n}\frac{\partial f_n}{\partial s_n}\delta s_n\\ \dot f(\sigma_{n+1}, t_{n+1}) &= \dot f(s_{n+1} + \delta s_{n+1}, t_{n+1}) \doteq \dot f_{n+1} + \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial f_{n+1}}{\partial s_{n+1}}\delta s_{n+1}\end{aligned} \tag{58}$$

can be used to substitute for f_n, $\dot f_n$, f_{n+1}, and $\dot f_{n+1}$ in the corrector (16).

Use of the corrector (21) for the variational equations gives this result in the form

$$\delta s_{n+1} = (\partial s_{n+1}/\partial s_n)\,\delta s_n + e_{n+1} \tag{59}$$

where

$$e_{n+1} \equiv \left[I - \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{h}{2} + \frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{\partial f_{n+1}}{\partial s_{n+1}}\frac{h^2}{12}\right]^{-1}\varepsilon_{n+1} \tag{60}$$

and where

$$\begin{aligned}\varepsilon_{n+1} &\equiv \sigma_{n+1} - \left[\sigma_n + \frac{f(\sigma_n, t_n) + f(\sigma_{n+1}, t_{n+1})}{2}h + \frac{\dot f(\sigma_n, t_n) - \dot f(\sigma_{n+1}, t_{n+1})}{12}h^2\right]\\ &= \overset{\cdots\cdot}{f}(\sigma(\sigma_n, t_n, \tau), \tau)\,h^5/720\end{aligned} \tag{61}$$

is the truncation error as usually defined. The fourth derivative $\overset{\cdots\cdot}{f}(\sigma, t)$ of $f(\sigma, t)$ with respect to t occurs at an intermediate value τ which differs in general for each component of ε_{n+1}.

The first term on the right in (59) is merely the propagation of the error δs_n from the beginning t_n of the integration step to the end t_{n+1} of the step. The truncation error is multiplied by the same matrix inverse that occurs in the corrector (21) for the variational equations. This gives the second term e_{n+1} in (59) as the error which is introduced by the integration step itself. But while Eqs. (59–61) show how the error accumulates with successive integration steps, it is difficult to obtain an accurate estimate for the truncation error ε_{n+1} in (61). Therefore Eq. (59) is not applied here to compute or bound δs_{n+1}, although this topic is well worth further investigation.

To control the step-size h, a norm

$$\|s\| \equiv \left[\sum_i (w_i s^i)^2\right]^{1/2} \geqq 0 \tag{62}$$

for the vector s in the differential equations (1) will be assumed. Each weight $w_i \geqq 0$ can be interpreted as the reciprocal of a bound for the magnitude $|s^i|$ of the corresponding element s^i of the vector s since

$$\|s\| < 1 \Rightarrow \left[\sum_{j\in\{j\}\subset\{i\}} (w_j s^j)^2\right]^{1/2} < 1 \Rightarrow |s^i| < \frac{1}{w_i}\ \forall\, i \tag{63}$$

For example, in Eqs. (5) let

$$w_1 = w_2 = w_3 = 1/e_p, \quad w_4 = w_5 = w_6 = 1/e_v \tag{64}$$

where $e_p > 0$ and $e_v > 0$ are, respectively, position and velocity tolerances. Then

$$\|s\| < 1 \Rightarrow \begin{cases}(x^2+y^2+z^2)^{1/2} < e_p \Rightarrow |x|, |y|, |z| < e_p\\ (\dot x^2+\dot y^2+\dot z^2)^{1/2} < e_v \Rightarrow |\dot x|, |\dot y|, |\dot z| < e_v\end{cases} \tag{65}$$

so each position component does not exceed e_p in magnitude and each velocity component does not exceed e_v in magnitude.

The step-size h can be adjusted so that the norm of the fourth-order term in the polynomial (45) for s_{n+1} equals a given constant k. This gives

$$h = (k/\|c_4\|)^{1/4} \tag{66}$$

for the determination of the step-size after c_4 has been determined. Since the global error varies approximately as the fourth power of the step-size for small steps, changing k to qk can be expected to multiply the global error by q. The magnitude of the global error is an unknown function of k, but appropriate values of k for a given class of problems can be determined empirically by first using much smaller values of k than necessary for a sample case. The difference between such a result and the same case with a much larger value of k then indicates how the global error behaves as a function of k.

It is better to start the first integration step with $c_3{}^* = c_4{}^* = 0$ and h much smaller than necessary if better estimates are not available. But some restriction on the use of (66) is required to handle any integration step when sudden large changes in $f(s, t)$ in (1) occur. A good procedure is to recompute c_3 and

c_4 from (47) after each Newton-Raphson iteration (31) and (29). The current step-size should only be halved if (66) gives a lesser value, in which case the entire step should be reinitialized with the predictor (51). But the step-size should never be increased during the iterations of (31) within a step, and the current step-size should only be doubled for the next step if the use of $c_4{}^*$ from (50) for c_4 in (66) gives a greater value. When (66) is accepted so that h is changed for the next iteration of (31), (45) must be used to recompute s_{n+1} as a new approximation s^*_{n+1} for (31).

Conclusion

The method of numerical integration described here has been successfully applied to the equations of motion for ballistic missile trajectories where the perturbing accelerations included Earth oblateness, atmospheric drag, and rocket thrusting. This required a good deal of attention to practical numerical problems which have not been discussed here in any detail. Rather, the way in which the method of integration differs from standard techniques has been emphasized, so that the approach can be readily applied to other problems. This is particularly important because the numerical solution of the initial-value problem occurs in so many different fields and because the sensitivity coefficients of the solution have so many important applications. For such applications, the advantages described here and demonstrated for ballistic missile trajectories should be obtained from this method of numerical integration.

References

[1] Hermite, C., "Sur la formule d'interpolation de Lagrange," *Zeitschrift für Reine und Angewandte Mathematik*, Vol. 84, 1878, pp. 70–79.

[2] Lapidus, L. and Seinfeld, J. H., *Numerical Solution of Ordinary Differential Equations*, Academic Press, New York, 1971, p. 199.

[3] Liniger, W. and Willoughby, R. A., "Efficient Integration Methods for Stiff Systems of Ordinary Differential Equations," *SIAM Journal of Numerical Analysis*, Vol. 7, 1970, pp. 47–66.

[4] Loscalzo, F. R., "On the Use of Spline Functions for the Numerical Solution of Ordinary Differential Equations," Ph.D. thesis, 1968, Univ. of Wisconsin, Madison, Wis.

[5] Ralston, A., *A First Course in Numerical Analysis*, McGraw-Hill, New York, 1965, pp. 212–213.