IMA Journal of Numerical Analysis (2001) **21**, 387–406

Global errors of numerical ODE solvers and Lyapunov's theory of stability

DIVAKAR VISWANATH[†]

Departments of Computer Science and Mathematics, University of Chicago, USA

[Received 4 January 1999 and in revised form 15 December 1999]

The error made by a numerical method in approximating the solution of the initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$, $t \geqslant 0$, $x(t) \in R^d$ varies with the time of integration. The increase of the global error $\|\tilde{x}(t; h) - x(t)\|$, where $\tilde{x}(t, h)$ is an approximation derived by a numerical method with time step h, with time t determines the feasibility of approximating the solution accurately for increasing t. However, the best available theoretical bounds involve the Lipshitz constant and are exponential in t for some problems where the actual increase of global error is only linear in time.

Using techniques from Lyapunov's theory of stability, we prove that the increase of global errors is linear in time for trajectories of dynamical systems which fall into a hyperbolic and attracting cycle or into a hyperbolic and attracting torus, with the flow on the torus being quasi-periodic. The increase is linear for non-linear problems when certain stability properties of the solution can be verified. The error analysis uses a conditioning function $E(t)$ associated with the exact solution, which captures the propagation and accumulation of global errors.

1. Introduction

The exact solution of the initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$, $t \geqslant 0$, $x(t) \in R^d$ is rarely obtainable exactly by analytic methods. However, the instances where it can be solved with sufficient accuracy by numerical methods, like Runge–Kutta, or backwards differentiation are numerous. The success of the numerical methods depends upon the rate of increase of the global error $\|\tilde{x}(t; h) - x(t)\|$ with t, where $\tilde{x}(t; h)$ denotes an approximation to the exact solution derived using a numerical method with constant step size of h. If the rate of increase is linear or polynomial of low degree in t, the solution can be approximated accurately for at least a moderately long period of time. However, if the rate is exponential in t, accurate approximation of the solution for even moderately long periods in t is a hopeless task. The global error depends on the time step h in proportion to h^r if the order of the numerical method is r. The exact solutions or trajectories are denoted by $x(t)$ or by $x(t; x_0)$, making the initial condition $x(0) = x_0$ explicit, and their approximations are denoted by $\tilde{x}(t; h)$ or by $\tilde{x}(t; x_0; h)$.

The function $E(t)$ below associated with the solution $x(t)$ can be thought of as a conditioning function which controls global errors. The function $E(t)$ is defined for $t \geqslant 0$

[†]Email: divakar@cs.uchicago.edu

by

$$E(t) = \sup_{v} \left\| \int_0^t \frac{\partial x(t)}{\partial x(s)} v(s)\, \mathrm{d}s \right\|, \tag{1.1}$$

where the supremum is taken over all continuous functions $v : [0, t] \to R^d$ with $\|v(s)\| \leqslant 1$ for $0 \leqslant s \leqslant t$. All vector norms in this paper are Euclidean norms and all matrix norms are the corresponding induced norms. The Jacobian $\frac{\partial x(t)}{\partial x(s)}$ in (1.1) is the sensitivity of $x(t)$ with respect to small changes to $x(s)$, and thus might be expected to control the propagation of the local discretization error, or a small perturbation to the exact solution, made at time s to time t. Since local discretization errors are made at every time step, $E(t)$ is defined by taking an integral which adds up to match the accumulation of errors. In (1.1), $v(s)$ gives the direction in R^d of the discretization error made at time s. Taking the supremum ensures that $E(t)$ picks up the worst possibility for the accumulation of errors.

Theorem 2.1 shows that, given $\epsilon > 0$ and $T > 0$,

$$\|x(t; x_0) - \tilde{x}(t; x_0; h)\| < (E(t) + \epsilon)Kh^r \tag{1.2}$$

for $0 \leqslant t \leqslant T$ and sufficiently small h, for a one-step method of order r. The one-step method could be, for example, a Runge–Kutta method. The constant K is determined by the local errors made by the numerical method. How small the time step has to be depends upon $f(t, x)$ and the numerical method, as well as ϵ and T. It is clear from (1.2) that the study of global errors can be reduced to a study of $E(t)$. The advantage is that $E(t)$ behaves like a conditioning function and is a mathematical property of the exact solution, which therefore does not contain any detail of the numerical method. The function $E(t)$ has been implicit in earlier work on global errors, especially in the work on asymptotic analysis (asymptotic with $h \to 0$) of global errors; see, for example, Henrici (1962), Henrici (1963), Gragg (1965), and Stetter (1973). We consider the clear and explicit definition of the conditioning function $E(t)$ to be one of the main contributions of this paper.

This work has two principal motivations. In their book (Stuart & Humphries, 1996, p. 240), Stuart and Humphries derive the available general purpose error bound

$$\|x(T; x_0) - \tilde{x}(T; x_0; h)\| \leqslant (\mathrm{e}^{LT} - 1)Kh^r,$$

where L, the Lipshitz constant of $f(t, x)$ in the region around the solution, is always positive. They point out the futility of this bound for large T and fixed time step h, however small h maybe, and go on to comment: *It is essentially for this reason that the interaction between the theories of dynamical systems and numerical analysis is an interesting and important area of investigation.* We replace $(\mathrm{e}^{LT} - 1)$ by $E(t)$ as in (1.2) and then bound $E(t)$ linearly in t for some classes of problems, thus deriving error bounds valid for much longer stretches of time. Indeed, it turns out that the talked about interaction between numerical analysis and dynamical systems theory is very much part of this.

The other motivation is from the work on global errors of symplectic integrators of Hamiltonian systems initiated by Calvo & Sanz-Serna (1993). They showed that the increase of global errors of some periodic solutions of Hamiltonian systems, for example the solution of the two-body Kepler problem, would be quadratic in time for the sort of general purpose integrators considered here, but only linear in time for symplectic integrators. This was convincingly explained later by Cano & Sanz-Serna (1997) following the papers (Calvo & Hairer, 1995) and (Estep & Stuart, 1995). These results have led to renewed interest in the analysis of global errors. There is numerical evidence due to Quispel & Dyt (1998) that the increase of errors is linear and not quadratic if symplectic or volume-preserving solvers are used to integrate quasi-periodic trajectories of Hamiltonian systems. Quasi-periodic trajectories wind around the Liouville tori of integrable Hamiltonian systems and appear in the remarkable KAM theory. Quasi- periodic trajectories play a far more important role in Hamiltonian dynamics than periodic trajectories. Thus, a proof of linear increase of global errors of symplectic solvers on quasi-periodic trajectories in Hamiltonian dynamics is desirable. We prove in Theorem 5.5 that the increase of global errors is linear in time for stable quasi-periodic trajectories of dissipative dynamical systems. This proof uses the normal hyperbolicity theory of Hirsh *et al.* (1977).

Sections 3 and 4 give examples to show that there is no direct connection between Lyapunov stability and linear bounds for $E(t)$. Propositions 4.2 and 4.3 obtain constant and linear upper bounds on $E(t)$ using the theory of inverse Lyapunov functions (Yoshizawa, 1966).

Section 5 gives yet another proof, different from the proofs in Stetter (1973) and Stuart & Humphries (1996), that $E(t)$ is bounded by a constant for trajectories that fall into a hyperbolic sink of a dynamical system $\dot{x} = f(x)$. The Hartman–Grobman theorem (Robinson, 1995) says that such trajectories approach the sink at an exponential rate once they are close to it. Therefore, small perturbations to the trajectory decay at an exponential rate.

Theorem 5.2 proves that $E(t)$ is bounded linearly in t for trajectories that fall into an attracting, hyperbolic cycle of a dynamical system. Perturbations transverse to the cycle approach the cycle at an exponential rate and perturbations along the cycle are propagated without any expansion or contraction along the cycle. However, only exponential contraction of perturbations towards the cycle is not enough to prove linear increase of $E(t)$. The result which is crucially needed is called convergence in phase and gives that if x_0 is close enough to the cycle γ, there is point p on γ such that $\|x(t; x_0) - x(t; p)\|$ decreases exponentially as $t \to \infty$; thus x_0 tracks p with exponentially increasing accuracy for increasing t. Convergence in phase is not involved when approximating a trajectory that falls into a sink because a sink is just one point. Theorem 5.5 bounds $E(t)$ linearly for trajectories falling into a torus if the flow is stable and quasi-periodic on the torus.

Trajectories of dissipative dynamical systems which are asymptotically stable as $t \to \infty$, and not chaotic, usually fall into either a sink, an attracting hyperbolic cycle or an attracting hyperbolic quasi-periodic torus. In the three main possibilities for stable asymptotic behaviour in dissipative dynamical systems, the bounds on $E(t)$ in Section 5 match the increase of global errors observed numerically.

2. A model for discretization errors

The model of discretization error in this section closely imitates discretization errors made by single-step methods with a constant step size. Stuart & Humphries (1996) model discretization errors of single-step methods in a similar manner. A similar model has been used by Stetter (1973). In this section, the exact solution of the initial value problem,

$\dot{x}(t) = f(t, x)$, $x(s) = x_0$, $t \geqslant s$, is denoted by $x(t; s, x_0)$, or abbreviated to $x(t; x_0)$ or $x(t)$ if $s = 0$.

Let $\alpha(h)$ be a continuous, strictly increasing function of h for $h \geqslant 0$. Assume also that $\alpha(0) = 0$. Then an approximation $\tilde{x}_\alpha(t; x_0; h)$ to $x(t; x_0)$ is defined as follows:

$$\begin{aligned} &\tilde{x}_\alpha(0; x_0; h) = x_0 \\ &\tilde{x}_\alpha(nh; x_0; h) = x(nh; (n-1)h, \tilde{x}_\alpha((n-1)h; x_0; h)) + h\alpha(h)v_n \quad n \geqslant 1, \end{aligned} \tag{2.1}$$

where $v_n \in R^d$ can be any vector with $\|v_n\| \leqslant 1$. In other words, an approximate solution at $t = nh$, $n \geqslant 1$ is obtained by exactly propagating the point $\tilde{x}_\alpha((n-1)h; x_0; h)$ at $t = (n-1)h$ under $\dot{x}(t) = f(t, x)$ until $t = nh$, and then adding the discretization error or discontinuity $h\alpha(h)v_n$, where $\|v_n\| \leqslant 1$. For $nh \leqslant t < (n+1)h$,

$$\tilde{x}_\alpha(t; x_0; h) = x(t; nh, \tilde{x}_\alpha(nh; x_0; h)). \tag{2.2}$$

Since v_n can be any vector with $\|v_n\| \leqslant 1$, this actually defines a whole family of approximate solutions which we denote by $\tilde{X}_\alpha(x_0; h)$. The exact solution $x(t)$ is the only member of this family which is continuous. Figure 1 gives an example of an approximate solution. Members of the family $\tilde{X}_\alpha(x_0; h)$ are written as $\tilde{x}(t; x_0; h)$, leaving α implicit, and sometimes as $\tilde{x}(t; h)$, leaving the initial condition $x(0) = x_0$ implicit. Approximate solutions where the initial condition is given at a point other than $t = 0$ are never considered, obviously, without any loss of generality.

Single-step numerical methods are related to approximations from the family $\tilde{X}_\alpha(x_0; h)$. Single-step methods can be thought of as propagating the solution exactly between integer multiples of the time step and committing a discretization error of $K_i h^{r+1} v_i$, $\|v_i\| = 1$, at the ith time step if the method is of order r. We now assume that the K_i are bounded by a constant K which does not depend upon h or i. This can be proven in some circumstances; see Stuart & Humphries (1996). Besides, if there is no such K, the numerical method will not in practice behave as if it were of order r. Thus, when the order of accuracy of the numerical method for approximating $x(t; x_0)$ is r, we can take $\alpha(h) = Kh^r$. Then there will always be an approximation in the family $\tilde{X}_\alpha(x_0; h)$ which is the same as that obtained by taking the numerically computed values at $t = nh$, $n = 1, 2, \ldots$, and interpolating in the intervals $t \in [nh, (n+1)h)$ by propagating the numerically computed values at $t = nh$. However, the family $\tilde{X}_\alpha(x_0; h)$ contains approximations other than the one obtained from the numerical method. As a result, the bound (2.3) is usually not sharp for the $\tilde{x}(t; h)$ corresponding to a numerical method.

The following theorem relates $E(t)$, defined in (1.1), to the errors $\|\tilde{x}(t; h) - x(t)\|$. The conditioning function $E(t)$ is defined when $f(t, x)$ is continuous in t and continuously differentiable in x. When additional assumptions about $f(t, x)$ are needed as in Theorem 2.1, they are stated explicitly.

THEOREM 2.1 Let $E(t)$ and the approximate solution $\tilde{x}_\alpha(t; x_0; h)$ be associated with the initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$, $x(t) \in R^d$. Assume that $f(t, x)$ is twice continuously differentiable with respect to t and x. Assume that the solution $x(t)$ exists for $t \geqslant 0$. Given $T > 0$ and $\epsilon > 0$ there exists $h_0 > 0$ such that $h < h_0$ implies

$$\sup_{\tilde{x} \in \tilde{X}_\alpha(x_0; h)} \|\tilde{x}(t; x_0; h) - x(t; x_0)\| \leqslant (E(t) + \epsilon)\alpha(h) \tag{2.3}$$

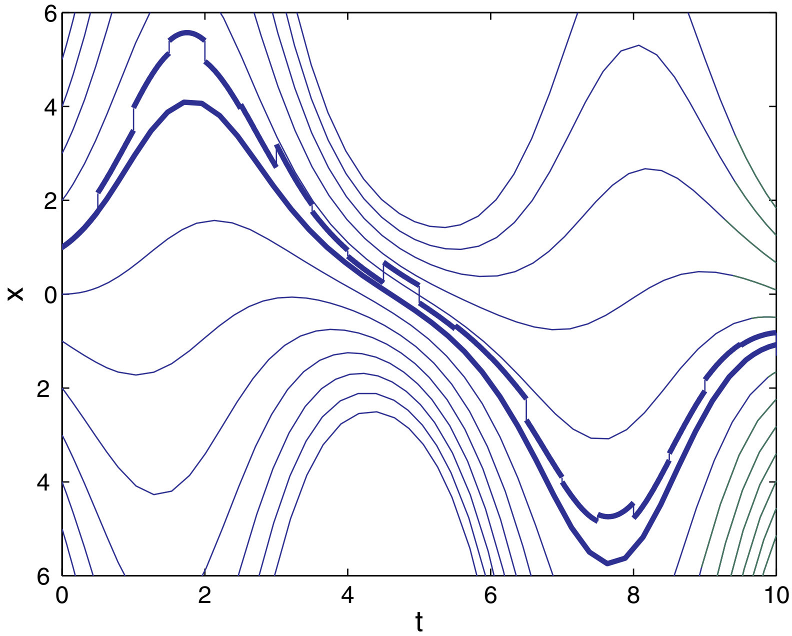

FIG. 1. The thick lines show an exact solution of the equation $\dot{x}(t) = \sin t + (\cos t)x$ and an approximation to it with $h = 0{\cdot}5$ and $\alpha(h) = 2h$.

for $0 \leqslant t \leqslant T$. The choice of h_0 depends on T, ϵ, the initial value problem and $\alpha(h)$.

Further, $E(t)$ cannot be replaced by a smaller number in the bound above because

$$\limsup_{h \to 0} \frac{\sup_{\tilde{x} \in \tilde{X}_\alpha(x_0;h)} \|\tilde{x}(t; x_0; h) - x(t; x_0)\|}{\alpha(h)} = E(t). \tag{2.4}$$

Proof. We denote $\tilde{x}(kh; h)$, an approximation to $x(kh)$, by $\tilde{x}_{kh}$. Let $n = \lfloor \frac{t}{h} \rfloor$. The proof of the first part is organized into three steps corresponding to (2.5)–(2.7). We take h small enough that all approximate solutions stay within δ of the exact solution for $0 \leqslant t \leqslant T$, i.e. $\|\tilde{x}(t; h) - x(t)\| < \delta$. The construction to realize this is given in Stuart & Humphries (1996), for example. Restrictions will be placed on δ later.

To begin with consider

$$\begin{aligned} x(t) - \tilde{x}(t; h) &= x(t) - x(t; nh, \tilde{x}_{nh}) \\ &= (x(t) - x(t; (n-1)h, \tilde{x}_{(n-1)h})) + (x(t; nh, \tilde{x}_{nh} - h\alpha(h)v_n) - x(t; nh, \tilde{x}_{nh})). \end{aligned}$$

The first equality follows from (2.2). By (2.1), $x(nh; (n-1)h, \tilde{x}_{(n-1)h}) = \tilde{x}_{nh} - h\alpha(h)v_n$, and the equality's two sides are added and subtracted from the right-hand side of the first equality to obtain the second equality. By similar manipulations, the sum below telescopes.

$$\begin{aligned} x(t) - \tilde{x}(t; h) = x(t) &- x(t; h, \tilde{x}_h) \\ &+ x(t; 2h, \tilde{x}_{2h} - h\alpha(h)v_2) - x(t; 2h, \tilde{x}_{2h}) \\ &+ \cdots \\ &+ x(t; nh, \tilde{x}_{nh} - h\alpha(h)v_n) - x(t; nh, \tilde{x}_{nh}). \end{aligned}$$

Let $M_k = \frac{\partial x(t;kh,\xi)}{\partial \xi}\big|_{\xi=\tilde{x}_{kh}}$. Take the kth term in the sum, which telescopes as $M_k h\alpha(h)v_k + e_k$, using a Taylor series approximation. The differentiability assumptions on f imply that $x(t; s, x_s)$ is twice continuously differentiable in x_s (Hale, 1969). Therefore, $\|e_k\| \leqslant C_1 h^2 \alpha(h)^2$, where the constant C_1 depends on the δ neighbourhood around the solution in $0 \leqslant t \leqslant T$. The terms of the sum which were expanded are rewritten to obtain

$$x(t) - \tilde{x}(t; h) = h\alpha(h)(M_1 v_1 + \cdots + M_n v_n) + E_1, \tag{2.5}$$

where $E_1 = e_1 + \cdots + e_n$ and consequently $\|E_1\| \leqslant C_1 n h^2 \alpha^2(h) \leqslant C_1 T h \alpha^2(h)$. Derivation of (2.5) is similar to the proof of the Alexseev–Grobner lemma in Hairer (1980) and to calculations by Iserles & Söderlind (1993).

Let $N_k = \frac{\partial x(t;kh;\xi)}{\partial \xi}\big|_{\xi=x(kh)} = \frac{\partial x(t)}{\partial x(s)}\big|_{s=kh}$. Let $(M_k - N_k)v_k = e'_k$. Since $x(t; s, x_s)$ is twice continuously differentiable in x_s, $\|M_k - N_k\| < C_2\delta$ for some constant C_2 which depends on the δ neighbourhood around the solution in $0 \leqslant t \leqslant T$. Since $\|v_k\| \leqslant 1$, $\|e'_k\| \leqslant C_2\delta$. The M_ks in (2.5) are replaced by N_ks to get

$$x(t) - \tilde{x}(t; h) = h\alpha(h)(N_1 v_1 + \cdots + N_n v_n) + E_1 + E_2, \tag{2.6}$$

where $E_2 = h\alpha(h)(e'_1 + \cdots + e'_n)$ and consequently $\|E_2\| \leqslant C_2 n h \alpha(h) \leqslant C_2 T \alpha(h)\delta$.

Let $\tilde{v} : [0, t] \to R^d$ be the discontinuous function with $\tilde{v}(0) = 0$, $\tilde{v}(kh) = v_k$ and $\tilde{v}(s) = \tilde{v}(kh)$ for $kh \leqslant s < (k+1)h$. In

$$h(N_1 v_1 + \cdots + N_n v_n) = \int_0^t \frac{\partial x(t)}{\partial x(s)} \tilde{v}(s)\, \mathrm{d}s + \eta(h)$$

$\|\eta(h)\| \to 0$ (uniformly in t for $0 \leqslant t \leqslant T$) as $h \to 0$ because the vs are bounded in the norm by 1 and the Jacobian under the integral is continuous in s. We pick a continuous function $v(s)$ using Lusin's theorem (Rudin, 1987) which differs from $\tilde{v}(s)$ on a set of measure less than h. Then

$$\int_0^t \frac{\partial x(t)}{\partial x(s)} \tilde{v}(s)\, \mathrm{d}s = \int_0^t \frac{\partial x(t)}{\partial x(s)} v(s)\, \mathrm{d}s + e'',$$

where $\|e''\| \leqslant C_3 h$ for a constant C_3 taken to be the supremum of the norm of the Jacobian in the compact region $0 \leqslant s \leqslant t \leqslant T$. We replace $h\Sigma_{i=1}^n N_i v_i$ in (2.6) by an integral as above to get

$$x(t) - \tilde{x}(t; h) = \alpha(h) \int_0^t \frac{\partial x(t)}{\partial x(s)} v(s)\, \mathrm{d}s + E_1 + E_2 + E_3 + \alpha(h)\eta(h), \tag{2.7}$$

where $E_3 = \alpha(h)e''$ and consequently $\|E_3\| \leqslant C_3 h \alpha(h)$.

For a proof of (2.3), we take norms of all terms in (2.7) to get

$$\|x(t) - \tilde{x}(t; h)\| \leqslant \alpha(h) \left\| \int_0^t \frac{\partial x(t)}{\partial x(s)} v(s)\, \mathrm{d}s \right\| + \|E_1\| + \|E_2\| + \|E_3\| + \alpha(h)\|\eta(h)\|$$
$$\leqslant \left(E(t) + \frac{\|E_1\| + \|E_2\| + \|E_3\|}{\alpha(h)} + \|\eta(h)\| \right) \alpha(h).$$

By (2.5)–(2.7), the first factor is bounded above by $(E(t) + \epsilon_1)$, where $\epsilon_1 = C_2 T \delta$, in the limit $h \to 0$. We can take $\delta < \frac{\epsilon}{2C_2 T}$, for example, and pick an h_0 to complete the proof of (2.3). Although C_2 in (2.6) depends on δ, the C_2 which is fixed for a certain δ works for all smaller δs.

For a proof of (2.4), which is the second part, we observe that (2.3) implies that the lim sup is less than or equal to $E(t)$. It is sufficient to prove the inequality in the other direction. Let $v : [0, t] \to R^d$ be a continuous function with $\|v(s)\| \leqslant 1$. Given $h > 0$, we construct approximate solutions by taking $v_k = v(kh)$. The sum $h(N_1 v_1 + \cdots N_n v_n)$ approximates the Riemann integral $\int_0^t \frac{\partial x(t)}{\partial x(s)} v(s)\, \mathrm{d}s$ with an error $\eta'(h)$, $\|\eta'(h)\| \to 0$, as $h \to 0$. We can use the integral in place of the sum $h\Sigma_{i=1}^n N_i v_i$ in (2.6) and argue as in the previous paragraph to conclude that

$$\limsup_{h\to 0} \frac{\sup_{\tilde{x}\in\tilde{X}_\alpha(x_0;h)} \|\tilde{x}(t; x_0; h) - x(t; x_0)\|}{\alpha(h)} \geqslant \left\| \int_0^t \frac{\partial x(t)}{\partial x(s)} v(s)\, \mathrm{d}s \right\|,$$

for any continuous v with $\|v(s)\| \leqslant 1$. Taking a supremum over all v, it follows that the lim sup on the left in (2.4) is greater than or equal to $E(t)$.

3. $E(t)$ of linear systems

The relationship between $E(t)$ and stability properties of the exact solution for linear systems of the form $\dot{y}(t) = A(t)y$, $y(0) = y_0$ is easy to derive. However, the relationship is not as simple as one might wish. There are both asymptotically stable examples with exponentially increasing $E(t)$ and unstable examples with linearly bounded $E(t)$. However, Propositions 3.3 and 3.4 give conditions for $E(t)$ to be bounded by a constant or to be linearly bounded.

We assume $A(t) \in R^{d,d}$ to be continuous in t in the linear initial value problem $\dot{y}(t) = A(t)y$, $y(0) = y_0$. If $Y(t)$ is the principal fundamental matrix of this linear problem, then $y(t) = Y(t)y_0$ for any y_0 (Hale, 1969). The Jacobian $\frac{\partial x(t)}{\partial x(s)}$ in (1.1), the definition of $E(t)$, is equal to $Y(t)Y^{-1}(s)$. The $E(t)$ of a linear initial value problem does not depend upon the initial condition y_0.

For scalar linear systems $\dot{y}(t) = a(t)y$, $a(t) \in R$, we have the following lemma.

LEMMA 3.1 The $E(t)$ of the solution of $\dot{y}(t) = a(t)y$, $y(0) = y_0$,is given by

$$E(t) = \mathrm{e}^{g(t)} \int_0^t \mathrm{e}^{-g(s)}\, \mathrm{d}s,$$

where $g(t) = \int_0^t a(\tau)\, \mathrm{d}\tau$.

Proof. The fundamental matrix, which is scalar in this situation, is given by $Y(t) = \mathrm{e}^{g(t)}$. Since $Y(t)$ is always positive, the optimal choice of $v(s)$ in (1.1) is $v(s) \equiv 1$.

The following definitions of stability were put forward by Lyapunov (1949).

DEFINITION 1 The solution $x(t; x_0)$ of $\dot{x}(t) = f(t, x)$, $x(0) = x_0$ is *stable* if, given any $\epsilon > 0$, there exists a $\delta > 0$ such that $\|x_0' - x_0\| < \delta$ implies $\|x(t; x_0') - x(t; x_0)\| < \epsilon$ for $t \geqslant 0$. In fact, stability implies that given $\epsilon > 0$, there exists a $\delta(\tau) > 0$ for every $\tau \geqslant 0$ such that $\|x(\tau; x_0') - x(\tau; x_0)\| < \delta(\tau)$ implies that $\|x(t; x_0') - x(t; x_0)\| < \epsilon$ for $t \geqslant \tau$. Here we must emphasize that $\delta(\tau)$ can depend on τ.

DEFINITION 2 The solution $x(t; x_0)$ is *asymptotically stable* if given $\epsilon > 0$ there exists a $\delta(\tau) > 0$ for every $\tau \geqslant 0$ such that $\|x_\tau' - x(\tau; x_0)\| < \delta(\tau)$ implies not only that $\|x(t; \tau, x_\tau') - x(t; x_0)\| < \epsilon$ for $t \geqslant \tau$ but also that $\|x(t; \tau, x_\tau') - x(t; x_0)\| \to 0$ as $t \to \infty$.

Implicit in the definitions is an assumption about the existence of solutions which begin near the solution $x(t; x_0)$. Obviously, asymptotic stability implies stability. For the scalar, linear problem $\dot{y}(t) = a(t)y$, $y(0) = y_0$, a necessary and sufficient condition for asymptotic stability is $g(t) \to -\infty$ as $t \to \infty$, where $g(t) = \int_0^t a(s)\,\mathrm{d}s$. However, the following examples show that both these concepts of stability are insufficient for bounding $E(t)$.

EXAMPLE 3.1 Given a rate $r(t)$, we consider a continuously differentiable function $g(t)$, $t \geqslant 0$, $g(0) = 0$, such that

1. $g(t) \leqslant -t$ for all $t \geqslant 0$,
2. $\mathrm{e}^{g(k)} \int_0^k \mathrm{e}^{-g(s)}\,\mathrm{d}s > r(k)$ for $k = 1, 2, 3, \ldots$.

For the linear system, we take $a(t) = g'(t)$. The first condition ensures asymptotic stability of the solution, and the second condition implies $E(k) > r(k)$ for positive integers k. Such a $g(t)$ is easy to construct. We take $g(k) = -k$ for $k = 0, 1, 2, \ldots$. For $k - 1 < t < k$, $k \geqslant 1$, we define $g(t)$ so that $g(t) \leqslant -t$ and

$$\int_{k-1}^{k} \mathrm{e}^{-g(s)}\,\mathrm{d}s \geqslant r(k)\,\mathrm{e}^k.$$

This can be carried out for any continuous $r(t)$, for example $r(t) = \mathrm{e}^t$.

EXAMPLE 3.2 On the other hand, there are unstable solutions with linearly bounded $E(t)$. Consider the scalar, linear system $\dot{y}(t) = \frac{\alpha}{t+1}y$, $t \geqslant 0$. For this ODE, $y(t) = (1+t)^\alpha y(0)$ implying instability of the solution for $\alpha > 0$. Yet, for $0 < \alpha < 1$, $E(t)$, which is $(1-\alpha)^{-1}(1+t)^\alpha((1+t)^{1-\alpha} - 1)$, is linearly bounded. For $\alpha = 1$, $E(t)$ is $(1+t)\log(1+t)$.

EXAMPLE 3.3 In Example 3.1, $|a(t)|$ is unbounded. Even asymptotic stability of $\dot{y}(t) = a(t)y$, $t \geqslant 0$ and the boundedness of $|a(t)|$ do not imply a linear bound for $E(t)$. We sketch the construction of a $g(t)$ to show this. First we take $g_1(t) = -t$ and $g_2(t) = -2t$. We take $g(t) = g_2(t)$ for $0 \leqslant t \leqslant t_1$, and let $g(t)$ increase monotonically until $g(t_2) = g_1(t_2)$ for $t_2 \geqslant t_1$, and then let $g(t)$ decrease monotonically until $g(t_3) = g_2(t_3)$ for $t_3 \geqslant t_2$. We repeat the same construction from t_3 onwards with t_4, t_5, and t_6 in place of t_1, t_2, t_3, and so on. The construction may be arranged so that

1. if $g(\tau) = g_1(\tau)$ then $g(t) = g_2(t)$ for $f_1\tau \leqslant t \leqslant f_2\tau$ for any fixed $0 < f_1 < f_2 < 1$;
2. $|a(t)| = |g'(t)|$ is bounded.

It is easy to check that $E(\tau) \geqslant e^{-\tau/2}(e^{2f_2\tau} - e^{2f_1\tau})$, for τ such that $g(\tau) = g_1(\tau)$. Therefore, for $f_2 > 1/2$, $E(t)$ increases exponentially. The linear system in this example has a negative Lyapunov exponent of -1.

The definitions of uniform stability and uniform asymptotic stability that follow seem to have been introduced by Malkin (1956). Theorems which deduce the stability of a non-linear system from its linearization usually (always?) assume the linear first approximation to be uniformly stable or uniformly asymptotically stable. The uniformity assumptions are not always explicitly stated, for example in Bellman (1953). In these cases, the $A(t)$ in $\dot{y}(t) = A(t)y$ is either constant or periodic, which means that stability implies uniform stability and asymptotic stability implies uniform asymptotic stability.

DEFINITION 3 The solution $x(t; x_0)$ of $\dot{x}(t) = f(t, x)$ is *uniformly stable* if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\|x(\tau; x_0) - x'_\tau\| < \delta$ for $\tau \geqslant 0$ implies $\|x(t; x_0) - x(t; \tau, x'_\tau)\| < \epsilon$ for $t \geqslant \tau$.

DEFINITION 4 The solution $x(t; x_0)$ is *uniformly asymptotically stable* if it is uniformly stable and the choice of δ in the previous definition can be made in such a way that $\|x(t; x_0) - x(t; \tau, x'_\tau)\| \to 0$ as $\tau \to \infty$ in a uniform way, i.e. given $\epsilon' > 0$ there exists $T_{\epsilon'}$ such that $\|x(t; x_0) - x(t; \tau, x'_\tau)\| < \epsilon'$ for all $t > \tau + T_{\epsilon'}$ and x'_τ satisfying $\|x(\tau; x_0) - x'_\tau\| < \delta$.

LEMMA 3.2 For the initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$

$$E(t) \leqslant \int_0^t \left\| \frac{\partial x(t)}{\partial x(s)} \right\| \mathrm{d}s.$$

Proof. Follows from (1.1).

For linear systems $\dot{y}(t) = A(t)y$, the stability properties of the solution and its $E(t)$ do not depend upon the initial condition at $t = 0$. Assuming the initial condition to be given at $t = 0$, we may talk about the linear system itself as being uniformly stable or uniformly asymptotically stable, or as having a linear $E(t)$.

PROPOSITION 3.3 If the linear system $\dot{y}(t) = A(t)y$ is uniformly stable, its $E(t)$ is linearly bounded, i.e. $E(t) \leqslant Kt$ for some $K > 0$ and $0 \leqslant t < \infty$.

Proof. Uniform stability of the linear system is equivalent to boundedness of $\|Y(t)Y^{-1}(s)\|$ for $t \geqslant s \geqslant 0$ (Hale, 1969; Yoshizawa, 1966). If $\|Y(t)Y^{-1}(s)\| \leqslant K$ for $t \geqslant s \geqslant 0$, then Lemma 3.2 implies $E(t) \leqslant Kt$.

PROPOSITION 3.4 If the linear system $\dot{y}(t) = A(t)y$ is uniformly asymptotically stable, its $E(t)$ is bounded by a constant, i.e. $E(t) < K$ for some $K > 0$ and $0 \leqslant t < \infty$.

Proof. Uniform asymptotic stability of the linear system is equivalent to $\|Y(t)Y^{-1}(s)\| < M\,\mathrm{e}^{-\nu(t-s)}$ for $\nu > 0$, $M > 0$ and $t \geqslant s \geqslant 0$ (Hale, 1969; Yoshizawa, 1966). Again, we use Lemma 3.2 to complete the proof.

Proposition 3.3 implies that $E(t)$ for the solution of $\dot{y}(t) = Ay$, $y(0) = y_0$ is linearly bounded if all the eigenvalues of A have negative or zero real parts, and the ones with zero real part are simple. If all the eigenvalues of A have strictly negative real parts then $E(t)$ is bounded by a constant by Proposition 3.4. The necessary stability properties of $\dot{y}(t) = Ay$ are verified in numerous places.

4. $E(t)$ of non-linear systems

This section gives two approaches to the analysis of $E(t)$ of non-linear systems. The conditioning function $E(t)$ is defined using the Jacobian $\frac{\partial x(t)}{\partial x(s)}$, which is determined by the linear first approximation or the linearization or the equation of first variation of the non-linear initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$. One approach is to look at the linearization (Proposition 4.1). The other approach is to directly make stability assumptions on the solution of the non-linear problem (Theorems 4.2 and 4.3). The two approaches correspond to Lyapunov's method of first approximation and Lyapunov's direct method, respectively.

PROPOSITION 4.1 Let $f(t, x)$ have continuous first-order partial derivatives with respect to t and x. The $E(t)$ of the solution $x(t; x_0)$ of the initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$ and the $E(t)$ of its linearization $\dot{y}(t) = A(t)y$, $y(0) = y_0$, where $A(t) = \frac{\partial f(t,x)}{\partial x}\Big|_{x=x(t;x_0)}$, are the same.

Proof. Let $Y(t)$ be the fundamental matrix of the linear equation $\dot{y}(t) = A(t)y$. Then $\frac{\partial x(t)}{\partial x(s)}$ in (1.1) is equal to $Y(t)Y^{-1}(s)$; for a proof see Hale (1969). The Jacobian $\frac{\partial y(t)}{\partial y(s)}$ corresponding to the linearization is also equal to $Y(t)Y^{-1}(s)$.

By Proposition 4.1, the solution of a non-linear initial value problem and the linearization around it have the same $E(t)$. However, the solution of the non-linear problem and its linearization can have very different stability properties; for an example see (Bellman, 1953, p. 87). This difference arises because, in bounding global errors, the h_0 in Theorem 2.1 is allowed to depend upon T, the length of integration. In definitions of stability, in contrast, all perturbations smaller than a certain magnitude must stay "close" to the unperturbed solution until $t = \infty$.

We introduce a technique for error analysis of numerical approximations to non-linear problems which uses Lyapunov functions. But first we show an example to point out some difficulties in bounding the $E(t)$ of non-linear initial value problems by making stability assumptions on the solution.

EXAMPLE 4.1 Consider the zero solution of $\dot{x}(t) = x - \mathrm{e}^t x^3$, $x(0) = 0$, $t \geqslant 0$. Its $E(t)$, by Propositions 3.3 and 4.1, is $\mathrm{e}^t - 1$. But we show that the zero solution is actually uniformly asymptotically stable. It is even exponentially asymptotically stable.

Figure 2 is the portrait of trajectories of $\dot{x}(\tau) = x - \mathrm{e}^t x^3$. The portrait for $x \leqslant 0$ is a reflection about $x = 0$. Thus, we can restrict ourselves to trajectories which are always

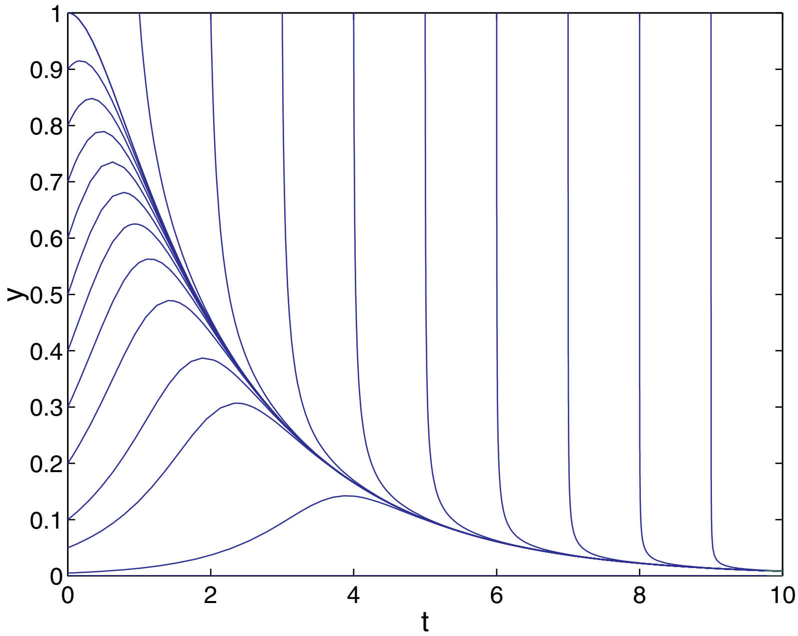

FIG. 2. The portrait of trajectories of $\dot{x}(t) = x - e^t x^3$. All solutions tend towards the curve $\dot{x}(t) = e^{-t/2}$.

in the upper half-plane. Every point (t, x) with $x \geqslant e^{-t/2}$ is on a trajectory that is pointed downwards.

We now verify uniform stability of the zero solution using the last observation. Given $\epsilon > 0$, we choose t_ϵ large enough that $e^{-t_\epsilon/2} < \epsilon$. By continuity properties, we can choose a $\delta > 0$ so that if $\tau \leqslant t_\epsilon$ and $x_\tau \leqslant \delta$, the trajectory through (τ, x_τ) stays below ϵ until t_ϵ. The maximum possible height (along x) of a trajectory beginning at (τ, x_τ), $\tau \geqslant t_\epsilon$ and $x_\tau < \delta$ is bounded by the larger of $e^{-\tau/2}$ and $|x_\tau|$. Since $e^{-\tau/2} < \epsilon$ and $\delta < \epsilon$, uniform stability is verified.

The verification of uniform asymptotic stability will be sketchy. It is based on the following facts:

1. the solution of $\dot{x}(t) = x - e^t x^3$, $x(0) = x_0$ tends to zero as $t \to \infty$ for $0 \leqslant x_0 \leqslant 1$;
2. further, $x(t; x_0) \leqslant x(t; 1)$ for $t \geqslant 0$, $0 \leqslant x_0 \leqslant 1$;
3. and, $0 \leqslant x(t + \tau; \tau; x_0) \leqslant x(t; 0; x_0)$ for any $x_0 \geqslant 0$, $\tau \geqslant 0$, $t \geqslant 0$.

The proof of item 1 involves a bit of elementary work which is done at the end. Item 2 is trivial. For item 3, we think of $x(t; 0, x_0)$ as the solution of $\dot{x}(t) = x - e^t x^3$, $x(0) = x_0$, and of $x(t + \tau; \tau, x_0) = z(t)$ as the solution of $\dot{z}(t) = z - e^{t+\tau} z^3$, $z(0) = x_0$, and use a differential inequality (Hartman, 1973, p. 27).

Now let T_ϵ be such that $\|x(t; x_0)\| < \epsilon$ for $t \geqslant T_\epsilon$ and $x_0 = 1$. Then $\|x(t+\tau; \tau, x_0)\| < \epsilon$ for any $\tau \geqslant 0$, $t \geqslant T_\epsilon$ and $|x_0| \leqslant 1$. Thus, the zero solution is uniformly asymptotically stable.

To prove item 1, it is enough to verify that the solution with the initial condition

$x(0) = 1$ satisfies $x(t) < 2\,\mathrm{e}^{-t/2}$ for $t \geqslant 0$. This is obvious from the portrait of trajectories: every trajectory that cuts the curve $x = 2\,\mathrm{e}^{-t/2}$ is in the downward direction; therefore any trajectory that starts below that curve has to stay below it for $t \geqslant 0$. In fact, $x(t) < 2\,\mathrm{e}^{-t/2}$ for the trajectory with $x(0) = 1$ implies that the zero solution is *exponentially asymptotically stable*, according to Yoshizawa's definition of exponential asymptotic stability (Yoshizawa, 1966).

The next two theorems are non-linear analogues of Propositions 3.3 and 3.4. The proofs rely on the theory of Lyapunov functions. Following the theorems is a comment about why Example 4.1 has an exponentially increasing $E(t)$ in spite of being exponentially asymptotically stable.

Within the proofs of Theorems 4.2 and 4.3, the non-linear initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$ is replaced by the zero solution of $\dot{z}(t) = F(t, z)$, $z(0) = 0$, where $F(t, z) = f(t, z + x(t; x_0)) - f(t, x(t; x_0))$. This reduction is standard in stability theory and goes back to Lyapunov. The reduced equation controls the propagation of perturbations and is called the perturbed equation. The original solution and the zero solution of the perturbed equation have exactly the same linearizations and hence the same $E(t)$. One can move back and forth between the approximate solutions of the perturbed equation and the approximate solutions of the original initial value problem for any $\alpha(h)$ by adding and subtracting $x(t; x_0)$. Replacement with the perturbed equation allows the employment of results from stability theory in the proofs without modification.

The Lyapunov functions $V(t, z)$ are always assumed to be continuous in t and locally Lipshitz in z. $V'_F(t, z)$ is the rate of increase of $V(t, z)$ along a solution of $\dot{z}(t) = F(t, z)$ which goes through (t, z). More precisely, if $z(t + \tau; t, z)$ is such a solution, then

$$V'_F(t, z) = \limsup_{\delta \to 0^+} \frac{V(t + \delta, z(t + \delta; t, z)) - V(t, z)}{\delta}.$$

If $V'_F(t, z) \leqslant aV(t, z)$ then $V(t + \delta, z(t + \delta; z, t)) \leqslant \mathrm{e}^{a\delta} V(t, z)$, and if $V'_F(t, z) \leqslant 0$ then $V(t + \delta, z(t + \delta; z, t)) \leqslant V(t, z)$; these two facts can be inferred from differential inequalities (Hartman, 1973).

The function $F(t, z)$ is *uniformly Lipshitz* with respect to z in a neighbourhood of zero if $\|F(t, z_1) - F(t, z_2)\| < L\|z_1 - z_2\|$, for a constant $L > 0$ and any z_i with $\|z_1\| < r$, $\|z_2\| < r$ where $r > 0$; L is the same constant for any $t \geqslant 0$. Similarly, $f(t, x)$ is uniformly Lipshitz with respect to x in a neighbourhood of the solution $x(t; x_0)$ if and only if $F(t, z) = f(t, z + x(t; x_0)) - f(t, x(t; x_0))$ is uniformly Lipshitz with respect to z around zero.

THEOREM 4.2 Let $f(t, x)$ be uniformly Lipshitz in x in a neighbourhood of the solution of the initial value problem $\dot{x}(t) = f(t, x)$, $x(0) = x_0$. Assume that the solution is exponentially stable in the sense that

$$\|x(t; s, x(s) + \delta) - x(t; x_0)\| \leqslant K\,\mathrm{e}^{-c(t-t_0)}\|\delta\|$$

for $\|\delta\| < r, s \geqslant 0, t \geqslant s, c > 0$ and $K > 0$. Then $E(t)$ of the solution $x(t; x_0)$ is bounded above by a constant.

Proof. As explained in the paragraphs preceding the theorem, the $E(t)$ of the zero solution of the perturbed problem $\dot{z}(t) = F(t, z)$ is equal to the $E(t)$ of $x(t; x_0)$. The proof applies to the perturbed equation and its zero solution.

Stability assumptions in the theorem imply existence of the Lyapunov function with the following properties (Yoshizawa (1966, p. 97), Hale (1969)):

1. $\|z\| \leqslant V(t, z) \leqslant C\|z\|$, where $C > 0$ is a constant;
2. $|V(t, z_1) - V(t, z_2)| \leqslant L\|z_1 - z_2\|$;
3. $V'_F(t, z) \leqslant -qcV(t, z)$ for some $0 < q < 1$.

The domain of $V(t, z)$ is $t \geqslant 0$ and $\|z\| \leqslant r$ for $r > 0$.

We take $h < h_0$ small enough that all approximate solutions $\tilde{z}(t; h)$ to the perturbed problem stay within a radius r of zero. Because of item 3, $V(t, \tilde{z}(t; h))$ decreases at least by a factor e^{-qch} along the approximate solution when t increases from kh to $(k + 1)h$; on that interval of t the approximate solution follows the exact solution until the discontinuity at $t = (k + 1)h$. The discontinuity can cause an increase in $V(t, z)$ of at most L times its magnitude by item 2. Therefore,

$$V(kh; \tilde{z}(kh; h)) \leqslant \mathrm{e}^{-qch} V((k - 1)h, \tilde{z}((k - 1)h; h)) + Lh\alpha(h),$$

for $k = 1, 2, \ldots, n$ and

$$V(t, \tilde{z}(t; h)) \leqslant \mathrm{e}^{-qch_r} V(nh, \tilde{z}(nh; h)),$$

where $h_r = t - nh$. Combining these inequalities, we get

$$\begin{aligned} V(t, \tilde{z}(t; h)) &\leqslant \mathrm{e}^{-qch_r} L\alpha(h)\left(\frac{1 - \mathrm{e}^{-nqch}}{1 - \mathrm{e}^{-qch}}\right) \\ &\leqslant K\alpha(h). \end{aligned}$$

That the K above can be a constant independent of h and n can be deduced from basic calculus using $h < h_0$. By item 1, $\|\tilde{z}(t; h)\| \leqslant C\alpha(h)$. Equation (2.4) of Theorem 2.1 implies a constant upper bound for $E(t)$.

The difficulty in proving Theorem 4.2 directly using the norm $\|\cdot\|$ is that when $K > 1$ the discretization error might actually be amplified by a factor greater than 1 over any given time step. Since we have to make the worst possible assumption at every time step, the final bound on $E(t)$ obtained this way will actually be exponential in t when $K > 1$. The proof of Theorem 4.2 uses a carefully constructed Lyapunov function to get around this difficulty.

THEOREM 4.3 Assume as in the previous theorem that $f(t, x)$ is uniformly Lipshitz with respect to x in a neighbourhood of $x(t; x_0)$. If the solution $x(t; x_0)$ of $\dot{x}(t) = f(t, x)$ is uniformly asymptotically stable, then $E(t) \leqslant Kt$ for some constant K.

Proof. Stability assumptions in this theorem imply the existence of a Lyapunov function with the following properties for the pertubed equation: (Hale (1969, Theorem 4.2, Chapter X) and Yoshizawa (1966))

1. $\|z\| \leqslant V(t, z)$;
2. $V(t, 0) \equiv 0$;
3. $V'_F(t, z) \leqslant 0$, where $V'_F(t, z)$ is defined as in the previous proof;
4. $|V(t, z_1) - V(t, z_2)| \leqslant K\|z_1 - z_2\|$ for some constant $K > 0$.

The domain of definition of $V(t, z)$ is the same as in the previous proof.

The proof is similar to that of Theorem 4.2, but this time

$$V(kh, \tilde{z}(kh; h)) \leqslant V((k-1)h, \tilde{z}((k-1)h; h)) + Kh\alpha(h)$$

for $k = 0, 1, \dots, n-1$ and

$$V(t, \tilde{z}(t; h)) \leqslant V(nh, \tilde{z}(nh; h)).$$

Combining these inequalities, we have $V(t, \tilde{z}(t; h)) \leqslant Kt\alpha(h)$. As before, $\|\tilde{z}(t; h)\| \leqslant Kt\alpha(h)$, which this time implies that $E(t) \leqslant Kt$.

The uniform Lipshitz assumption on $f(t, x)$ is crucial in Theorems 4.2 and 4.3. Example 4.1, which is uniformly asymptotically stable, does not satisfy the uniform Lipshitz assumption and has an $E(t)$ which increases exponentially. We do not know if Theorem 4.3 is still true if the assumption of uniform asymptotic stability is weakened to just uniform stability. If such a theorem were true, its wider applicability might be of use. Table 1 summarizes Sections 3 and 4.

5. Three applications to dynamical systems

We give three examples to illustrate the applicability of our methods for bounding the accumulation of global error. All three examples exploit stability theory of perturbed linear systems in Chapter III, Section 2, of Hale (1969).

5.1 *Hyperbolic sinks of $C^{1+\epsilon}$ dynamical systems*

Let p be a fixed point of a C^1 dynamical system $\dot{x}(t) = f(x)$; i.e. let $f(p) = 0$. Then p is a hyperbolic sink, if all the eigenvalues of $\frac{\partial f}{\partial x}\big|_{x=p}$ have strictly negative real parts. The following theorem can be derived using Chapter 6 of Stuart & Humphries (1996); it is originally due to Stetter (1973). We give the theorem here because our method of proof is different.

THEOREM 5.1 Let $x(t; x_0)$ be a trajectory of the dynamical system $\dot{x}(t) = f(x)$, $f \in C^{1+\epsilon}(R^d)$, which falls into a hyperbolic sink p as $t \to \infty$. Then its $E(t)$ is bounded above by a constant.

Proof. Let $A = \frac{\partial f(x)}{\partial x}|_{x=p}$. All the eigenvalues of A are in the left half-plane and consequently the linear system $\dot{y}(t) = Ay$ is uniformly asymptotically stable. By Robinson (1995, p. 150), there is a neighbourhood U_0 of p such that $x_0 \in U_0$ implies

$$\|x(t; x_0) - p\| < c\,\mathrm{e}^{-at}$$

TABLE 1

Summary of part of Sections 5 and 6. The second column is the stability assumption about the solution; the last column says what is known about the conditioning function $E(t)$ corresponding to such a solution

Linear Problems	Asymptotic stability	$E(t)$ can increase exponentially even if $\|A(t)\|$ is bounded; Example 3.3
	Uniform stability	$E(t)$ must be linearly bounded; Proposition 3.3
	Uniform asymptotic stability	$E(t)$ must be bounded by a constant; Proposition 3.4
Non-linear problems	Uniform stability	$E(t)$ can increase exponentially
	Uniform stability with uniform Lipshitz assumption	Not known if $E(t)$ must be linearly bounded
	Uniform asymptotic stability	$E(t)$ can increase exponentially; Example 4.1
	Uniform asymptotic stability with uniform Lipshitz assumption	$E(t)$ must be linearly bounded; Theorem 4.3
	Exponential stability as in Theorem 4.2	Not known if $E(t)$ must be linearly bounded
	Exponential stability as in Theorem 4.2 with uniform Lipshitz assumption	$E(t)$ must be bounded by a constant; Theorem 4.2

for constants $a > 0, c > 0$, and for $t \geqslant 0$. If x_0 is not in U_0, $x(t; x_0)$ enters U_0 and stays in U_0 after finite time. Therefore, $\|x(t; x_0) - p\| < c\,\mathrm{e}^{-at}$ holds as long as x_0 is in the basin of attraction of p, although c and a may have to be adjusted depending upon x_0.

Let $B(t) = \frac{\partial f(x)}{\partial x}|_{x=x(t;x_0)} - \frac{\partial f(x)}{\partial x}|_{x=p} = \frac{\partial f(x)}{\partial x}|_{x=x(t;x_0)} - A$. The linearization of $x(t; x_0)$ is given by $\dot{y}(t) = (A + B(t))y$. Since $f \in C^{1+\epsilon}$,

$$\begin{aligned}\|B(t)\| &= \left\| \frac{\partial f(x)}{\partial x}|_{x=x(t;x_0)} - \frac{\partial f(x)}{\partial x}|_{x=p} \right\| \\ &\leqslant c'\|x(t; x_0) - p\|^{\epsilon} \\ &\leqslant c''\mathrm{e}^{-\epsilon a t},\end{aligned}$$

where c' and c'' are constants independent of t. The first inequality above follows from $C^{1+\epsilon}$ Holder continuity of f since the trajectory $x(t; x_0)$ stays within a compact region of R^d. The second inequality follows from the inequality in the previous paragraph.

By Hale (1969, Theorem III.2.3), the linear system $\dot{y}(t) = (A + B(t))y$ is uniformly asymptotically stable. Propositions 3.3 and 4.1 imply $E(t)$ is bounded by a constant.

5.2 *Hyperbolic, attracting cycles of $C^{1+\epsilon}$ dynamical systems*

Let $x(t)$, $t \geqslant 0$ be a periodic orbit of the C^1 dynamical system $\dot{x}(t) = f(x)$ in R^d. Let $T > 0$ be its period so that $x(t+T) = x(t)$. We denote the set of points on this orbit by γ.

The characteristic multipliers of the cycle γ can be defined in two ways. One is to pick a point $p \in \gamma$, take a cross-section Σ at p, define a Poincaré map for Σ, and then define the characteristic multipliers as the $(d-1)$ eigenvalues of the linearization of the Poincaré map at p. The other way is to consider the linear first approximation $\dot{y}(t) = A(t)y$ on the cycle γ. Obviously, $A(t+T) = A(t)$ for $t \geqslant 0$. The Floquet numbers of this linear system can also be used to define characteristic multipliers. For a lucid account of these matters, see Robinson (1995).

The cycle γ is hyperbolic and attracting if all its characteristic multipliers are strictly less than 1 in magnitude.

THEOREM 5.2 Let $x(t; x_0)$, $t \geqslant 0$ be an orbit of a $C^{1+\epsilon}$ dynamical system $\dot{x}(t) = f(x)$ in R^d which falls into a hyperbolic, attracting cycle γ as $t \to \infty$. Then its $E(t)$ is linearly bounded from above.

LEMMA 5.3 Assume $x_0 \in \gamma$ so that $x(t; x_0)$ is a periodic orbit. Let its linearization be $\dot{y}(t) = A(t)y$, $t \geqslant 0$. If γ is hyperbolic and attracting, $\dot{y}(t) = A(t)y$ is uniformly stable, and the $E(t)$ associated with $x(t; x_0)$ is linearly bounded.

Proof. Uniform stability of $\dot{y}(t) = A(t)y$ is an easy consequence of the characteristic multipliers of γ being strictly less than 1; see Chapter VI of Hale (1969). The linear bound on $E(t)$ follows from Propositions 3.4 and 4.1.

Lemma 5.3 is contained in a different form in the work of Cano & Sanz-Serna (1997). However, Theorem 5.2 goes beyond Lemma 5.3 in a significant way. In practice, it is highly unlikely that x_0 itself is on the cycle γ, but it is often easy to find x_0 so that $x(t; x_0)$ falls into a cycle γ.

The following lemma is Theorem III.2.2 of Hale (1969). Its proof, which we omit, is short and simple, and illustrative of an important technique in stability theory.

LEMMA 5.4 Assume the linear system $\dot{y}(t) = A(t)y$, $t \geqslant 0$ is uniformly stable. Also assume that $B(t)$, $t \geqslant 0$ is continuous with $\int_0^\infty \|B(t)\|\,\mathrm{d}t < \infty$. Then the linear system $\dot{y}(t) = (A(t) + B(t))y$ is also uniformly stable.

Proof (Proof of Theorem 5.2). By Hartman (1973, p. 254), there exists a point $x_0' \in \gamma$ such that

$$\|x(t; x_0) - x(t; x_0')\| < c\,\mathrm{e}^{-at}, \tag{5.1}$$

for constants $a > 0$, $c > 0$ and for $t \geqslant 0$. This is called convergence in phase (Robinson, 1995).

Let $\dot{y}(t) = A(t)y$, where $A(t) = \frac{\partial f}{\partial x}\big|_{x=x(t;x_0')}$, be the first approximation along $x(t; x_0)$. By Lemma 5.3, this linear system is uniformly stable.

Let $\dot{y}(t) = (A(t)+B(t))y$, where $A(t)+B(t) = \frac{\partial f}{\partial x}\big|_{x=x(t;x_0)}$, be the first approximation along $x(t; x_0)$. The estimate (5.1) for convergence in phase implies

$$\|B(t)\| < c''\mathrm{e}^{-\epsilon a t},$$

for $\epsilon > 0$ and constants $a > 0$ and $c'' > 0$, as in the proof of Theorem 5.2. This is because both $x(t; x_0)$ and $x(t; x_0')$ approach each other exponentially in a compact region of R^d, and $f(x)$ is $C^{1+\epsilon}$.

By Lemma 5.4, $\dot{y}(t) = (A(t) + B(t))y$ is also uniformly stable.

Since $A(t) + B(t)$ gives the linearization of $x(t; x_0)$, the proof is completed using Propositions 3.3 and 4.1.

5.3 *Normally contracting manifolds with quasiperiodic flows*

The notation ϕ_t for the flow induced on R^d by $\dot{x}(t) = f(x)$ is standard in dynamical systems literature. With that notation $x(t; x_0) = \phi_t x_0$. Let V be a compact C^1 manifold which is invariant under this flow. We consider the situation when the flow on V is *differentiably conjugate* to the quasiperiodic flow on a torus, and V is normally hyperbolic and contracting, or briefly, *normally contracting*. We now explain the two italicized concepts in this paragraph.

A torus T^n is the product of n copies of the circle S^1. If the angle on the ith circle is parameterized by θ_i, a quasiperiodic flow on T^n is of the form $\theta_i(t) = (\theta_i(0) + \alpha_i t)$ mod 2π. The flow is periodic if the α_i are all mutually commensurable. We denote this flow by ψ_t.

When we say that the flow ϕ_t is differentiably conjugate to the quasiperiodic flow on a torus, we mean that there exists a C^1 homeomorphism $h : V \to T^n$ such that $h(\phi_t x) = \psi_t(hx)$ for $x \in V$ and $t \geqslant 0$.

To define normal contractivity (Hirsh *et al.*, 1977; Robinson, 1995), we associate a direct sum decomposition $T_x \oplus N_x$ of R^d with every x in V. In this splitting T_x is the tangent space of V at x, and N_x, the normal space, varies continuously with x. If the N_x can be chosen so that

$$\left\| \Pi_{N_y} \frac{\partial y}{\partial x} \right|_{N_x} \right\| < c\, \mathrm{e}^{-\mu t},$$

where $y = \phi_t x$, the matrix inside the norm is the restriction of the derivative $\frac{\partial y}{\partial x}$ to act from N_x to N_y, $c > 0$, and $\mu > 0$, then V is normally contracting. Usually, the definition of normal contractivity comes with another assumption which says contraction in the normal direction dominates any contraction on the manifold V. But since we have assumed that the flow on V is differentiably conjugate to quasiperiodic flow on a torus, this other assumption can be dropped.

Obviously, the tangent spaces T_x are invariant under the derivative map $\frac{\partial \phi_t x}{\partial x}$. It is actually possible to choose N_x so that they too are invariant under the derivative map (Hirsh *et al.*, 1977; Robinson, 1995). We take this to be the case. So the derivative map $\frac{\partial \phi_t x}{\partial x}$ maps T_x to $T_{\phi_t x}$ and N_x to $N_{\phi_t x}$.

THEOREM 5.5 Let $x(t; x_0)$ be a trajectory of the $C^{1+\epsilon}$ flow $\dot{x}(t) = f(x)$ which falls into a normally contracting and invariant manifold V. Assume that the flow on V is differentiably conjugate to the quasiperiodic flow on a torus. Then the $E(t)$ of $x(t; x_0)$ is linearly bounded.

Theorem 5.5 generalizes Theorem 5.2. Its proof is exactly analogous. We begin with a lemma about trajectories that begin on V.

LEMMA 5.6 Let $x_0 \in V$ so that $x(t; x_0)$ stays on V for $t \geqslant 0$. We make the same assumptions about V as in Theorem 5.5. Then the linearization $\dot{y}(t) = A(t)y$ along $x(t; x_0)$ is uniformly stable, and the $E(t)$ of $x(t; x_0)$ is linearly bounded.

Proof. The principal fundamental matrix of the linearization in the lemma is given by $Y(t) = \frac{\partial \phi_t x_0}{\partial x_0}$, which is the derivative map. Therefore, it is enough if we show that $\| \frac{\partial \phi_t x}{\partial x} \|$ is bounded by a constant for any $x \in V$ and $t \geqslant 0$.

We already know that the maps induced by the derivative map between tangent spaces and between normal spaces are bounded in the norm because of differentiable conjugacy to flow on a torus and normal contractivity, respectively. Since the tangent spaces and normal spaces are both invariant under the derivative map, it is enough if we show that the angle between T_x and N_x (in the sense of the CS decomposition) is bounded away from 0. That this angle is bounded away from 0 is implied by the compactness of V.

The proof of Theorem 5.5 can be completed exactly as the proof of Theorem 5.2 using the following result about convergence in phase.

THEOREM 5.7 As in Theorem 5.5, let $x(t; x_0)$ be a trajectory of the C^1 flow $\dot{x}(t) = f(x)$ which falls into a normally contracting and invariant manifold V, and let the flow on V be differentiably conjugate to the quasi-periodic flow on a torus. Then there exists $x_0' \in V$ such that

$$\|x(t; x_0) - x(t; x_0')\| < c\, \mathrm{e}^{-at},$$

for positive constants c and a, and $t \geqslant 0$.

Proof. This theorem can be deduced from Theorem 4.1 and the remark following its proof in Hirsh *et al.* (1977). See in particular part (a) of that theorem about stable manifolds and part (g) about conjugacy to linearized flows.

6. Conclusion

Below are some remarks about multi-step methods and variable time stepping, and a brief discussion of one-sided Lipshitz conditions.

(i) *Multi-step methods and variable time stepping*. The model for discretization errors in Section 2 is adapted to single-step methods with constant step sizes. For linear multi-step methods with constant step sizes, we believe the accumulation of global error can be worse but not better (after excluding some trivial cases) than indicated by $E(t)$.

In most problems where variable time stepping is used, it works out as follows. There are some regions of the solution where the residual error of numerical approximations is large and the numerical methods take short time steps. In other regions of the solution, the time steps are longer. But as the tolerance for error control is increased, the time steps are refined roughly proportionally. For example, decreasing the tolerance for the absolute local error by a factor of 2^{r+1} will cause both the longer and shorter time steps to be roughly halved, if the numerical method is of order r. When variable time stepping behaves in this way, $E(t)$ will provide a good indication of the increase of global errors. The roughly proportional refinement of time steps would also imply that the ratio of the biggest time step to the smallest time step is bounded. This is true usually but not always (Stoffer & Nipp, 1991; Stuart, 1997).

(ii) *One-sided Lipshitz conditions*. There is an approach to global error analysis of the linear system $\dot{y}(t) = A(t)y$ using one-sided Lipshitz conditions. Since $\|y(t)\|^2 = y^T(t)y(t)$, we have

$$\frac{\mathrm{d}\|y(t)\|^2}{\mathrm{d}t} = y^T(t)(A^T(t) + A(t))y(t).$$

If $\lambda(t)$ is the maximum eigenvalue of $A^T(t) + A(t)$, then

$$\frac{\mathrm{d}\|y(t)\|^2}{\mathrm{d}t} \leqslant \lambda(t)\|y(t)\|^2.$$

Thus, upper bounds for $\|y(t)\|$, and hence for $\|Y(t)\|$ where $Y(t)$ is the principal fundamental matrix of the linear system, can be written down in terms of $\lambda(t)$. These can be plugged into Lemma 3.2 to get bounds on $E(t)$ and hence the accumulation of global error. For a detailed account, see Hairer (1980).

In our view, one-sided Lipshitz conditions are basically a way to get a handle on stability by looking at the evolution of the norm of $y(t)$. This is a far less general approach to stability than the two methods of Lyapunov we have used. It is also of far lesser applicability; we do not see a way to derive any of the results in Section 5 using one-sided Lipshitz conditions.

In conclusion, the main contributions of this paper are the analysis of global errors in terms of $E(t)$ defined in (1.1), the linear and constant uppper bounds on the accumulation of global errors using the theory of inverse Lyapunov functions in Section 4, and the linear upper bounds on the accumulation of global error for stable trajectories of dynamical systems which are asymptotically periodic or quasi-periodic.

Acknowledgements

The author is very grateful to Arieh Iserles and Mike Shub for encouragement and enlightening discussions. He has benefited from talking to several visitors to M.S.R.I., Berkeley, in the autumn of 1998 about this work. The author thanks Andrew Stuart for motivating this work. This paper was revised substantially following the suggestions made by the referees.

This work was carried out when the author was a postdoctoral fellow at M.S.R.I., Berkeley.

REFERENCES

BELLMAN, R. 1953 *Stability Theory of Differential Equations*. New York: McGraw–Hill.

CALVO, M. P. & HAIRER, E. 1995 Accurate long-term integration of dynamical systems. *Appl. Numer. Math.* **18**, 95–105.

CALVO, M. P. & SANZ-SERNA, J. M. The development of variable-step symplectic integrators, with application to the two-body problem. *SIAM J. Sci. Comput.* **14**.

CANO, B. & SANZ-SERNA, J. M. 1997 Error growth in the numerical integration of periodic orbits, with application to Hamiltonian and reversible systems. *SIAM J. Numer. Anal.* **34**, 1391–1417.

ESTEP, D. J. & STUART, A. M. 1995 The rate of error growth in Hamiltonian-conserving integrators. *Z. Agnew. Math. Phys.* **46**, 407–418.

GRAGG, W. B. 1964 Repeated extrapolation to the limit in the numerical solution of ordinary differential equations, *Thesis*, University of California; see also 1965 *SIAM J. Numer. Anal.* **2**, 384–403.

HAIRER, E., NØRSETT, S. P., & WANNER, G. 1980 *Solving Ordinary Differential Equations I*. New York: Springer.

HALE, J. 1969 *Ordinary Differential Equations*. New York: John Wiley.

HARTMAN, P. 1973 *Ordinary Differential Equations*. New York: John Wiley.

HENRICI, P. 1962 *Discrete Variable Methods in Ordinary Differential Equations*. New York: John Wiley.

HENRICI, P. 1963 *Error Propogation for Difference Methods*. New York: John Wiley.

HIRSH, M. W., PUGH, C. C., & SHUB, M. 1977 Invariant Manifolds. *Lecture notes in Mathematics 583*. New York: Springer.

ISERLES, A. & SÖDERLIND, G. 1993 Global bounds on numerical error for ordinary differential equations. *J. Complexity* **9**, 97–112.

LYAPUNOV, A. M. 1892 Problème générale de la stabilité du mouvement. *Comm. Soc. Math. Kharkov* **2**; 1893 *Ibid* **3**; 1907 *Ann. Fac. Sci. Toulouse* **9**, 204–474; 1949 *Ann. Math. Studies*, Vol. 17, Princeton: Princeton University Press.

MALKIN, J. G. 1956 Theory of stability of motion. *Atomic Energy Commision Technical Report, 3352*.

QUISPEL, G. R. W. & DYT, C. P. 1998 Volume-preserving integrators have linear error growth. *Phys. Lett.* A **242**, 25–30.

ROBINSON, C. 1995 *Dynamical Systems. Stability, Symbolic Dynamics, and Chaos*. Boca Roton: CRC press.

RUDIN, W. 1987 *Real and Complex Analysis*, 3rd edn, New York: McGraw-Hill.

STETTER, H. J. 1973 *Analysis of Discretization Methods for Ordinary Differential Equations*. New York: Springer.

STOFFER, D. & NIPP, K. 1991 Invariant curves for variable step size integrators. *BIT* **31**, 169–180.

STUART, A. M. 1997 Probabilistic and deterministic convergence proofs for software for initial value problems. *Numerical Algorithms* **14**, 227–260.

STUART, A. M. & HUMPHRIES, A. R. 1996 *Dynamical Systems and Numerical Analysis*. Cambridge: Cambridge University Press.

YOSHIZAWA, T. 1966 *Stability Theory by Liapunov's Second Method*. Tokyo: Mathematical Society of Japan.