

Chapter 5

Stiff Systems

5.1 Understanding of Stiff Systems

Stiff ordinary differential systems arise frequently in the fields of chemical kinetics, nuclear reactors, control theory and electrical circuit theory. Generally speaking, whenever there involves a quickly changing dynamics, there is stiffness.

Definition 5.1.1 *A linear differential system*

$$y' = Ay + \phi(x) \quad (5.1)$$

where $A \in R^{n \times n}$, $y, \phi \in R^n$ is said to be stiff if and only if

1. For all i , $\Re\lambda_i < 0$,
2. (Stiffness ration) $\frac{\max |\Re\lambda_i|}{\min |\Re\lambda_i|} \gg 1$

where λ_i , $i = 1, \dots, n$ are eigenvalues of A .

Remark. For a general problem

$$y' = f(x, y), \quad (5.2)$$

we think it exhibits stiffness if the eigenvalues of the Jacobian matrix $\frac{\partial f}{\partial y}$ behaves in a similar fashion. This is because that near a particular solution $y = g(x)$ we may regard

$$f(x, y) = f(x, g(x)) + \frac{\partial f(x, g(x))}{\partial y} (y - g(x)) + O(\|y - g(x)\|^2). \quad (5.3)$$

Define $z(x) := y(x) - g(x)$. Then z solves the system

$$z' = J(x)z + O(\|z\|^2). \quad (5.4)$$

Example. Consider the heat equation

$$\begin{aligned}\frac{\partial u(x,t)}{\partial t} &= \frac{\partial^2 u(x,t)}{\partial x^2} \\ u(0,t) &= u(1,t) = 0 \\ u(x,0) &= g(x) \text{ (given).}\end{aligned}\tag{5.5}$$

Suppose we want to use the so called *method of line* to approximate the solution $u(x,t)$ at $x = x_k$ by $u_k(t)$ where

$$\begin{aligned}\frac{du_k(t)}{dt} &= \frac{u_{k+1}(t) - 2u_k(t) + u_{k-1}(t)}{(\delta x)^2} \\ u_0 &= u_N(t) = 0 \\ u_k(0) &= g(x_k).\end{aligned}\tag{5.6}$$

Then (5.6) is a stiff system because in matrix form

$$\begin{bmatrix} u'_1 \\ u'_2 \\ \vdots \\ u'_{N-1} \end{bmatrix} = \frac{1}{(\delta x)^2} \begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-1} \end{bmatrix},\tag{5.7}$$

it is well known that the coefficient matrix has eigenvalues

$$\lambda_{N-i} = -(2N \sin \frac{i\pi}{2N})^2.\tag{5.8}$$

To understand why stiff systems are difficult, we consider the following problem

$$\begin{aligned}y' &= A(y - p(x)) + p'(x) \\ y(0) &= v\end{aligned}\tag{5.9}$$

where $A, y, p \in R$ and $A < 0$. The exact solution is given by

$$y(x) = (v - p(0))e^{Ax} + p(x).\tag{5.10}$$

Assume that

$$|p''| \ll 1.\tag{5.11}$$

Suppose that the variable-step Euler method

$$y_{n+1} = y_n + h_n f(x_n, y_n)$$

is used to solve the problem. Since the local truncation error ERR is given by

$$ERR = \frac{h_n^2}{2} y''(x_n) + O(h_n^3),\tag{5.12}$$

the error control criterion requires

$$\epsilon \approx \left| \frac{h_n^2}{2} y''(x_n) \right|. \quad (5.13)$$

Thus the step size h_n should be chosen so that

$$h_n \approx \left(\frac{2\epsilon}{|y''(x_n)|} \right)^{1/2}. \quad (5.14)$$

Observe that

$$y'' = (v - p(0))A^2 e^{Ax} + p''(x).$$

We see that

$$h_n \approx \begin{cases} \left(\frac{2\epsilon}{|v-p(0)|A^2} \right)^{1/2} & \text{for } 0 < x \ll 1 \\ \left(\frac{2\epsilon}{p''(x)} \right)^{1/2} & \text{for } 1 \ll x. \end{cases} \quad (5.15)$$

Despite the accuracy consideration, the step size h_n given by (5.15) does not return reasonable numerical results. Recall that the global error $e_n = y(x_n) - y_n$ is given by

$$\begin{aligned} e_{n+1} &= (1 + h_n A)e_n + [y(x_{n+1}) - y(x_n) - h_n y'(x_n)] \\ &= \text{Propagated error} + \text{Local truncation error}. \end{aligned}$$

The propagated error will be amplified unless $|1 + h_n A| \leq 1$. That is, for the stability consideration, it is required that

$$0 < h_n \leq \frac{2}{|A|}.$$

This restriction dominates the selection (5.15) when x is large.

Remark. The geometric meaning of stiffness therefore is:

1. $\Re \lambda_i < 0$ for all i implies the existence of a steady-state solution.
2. Small $|\Re \lambda_i|$ implies longer range of integration to reach the steady state.
3. Large $|\Re \lambda_i|$ implies excessively small step size for stability.
4. Large stiffness ratio implies difficult computation.

In contrast, suppose the backward Euler method

$$y_{n+1} = y_n + h_n f(x_{n+1}, y_{n+1})$$

is used. Then

$$\begin{aligned} e_{n+1} &= y(x_{n+1}) - y_{n+1} \\ &= y(x_{n+1}) - \{y_n + h_n [A(y_{n+1} - p(x_{n+1})) + p'(x_{n+1})]\} \\ &= y(x_n) + h_n y'(x_{n+1}) + O(h_n^2) - \{y_n + h_n [A(y_{n+1} - p(x_{n+1})) + p'(x_{n+1})]\} \\ &= e_n + h_n A e_{n+1} + O(h_n^2). \end{aligned}$$

Therefore,

$$e_{n+1} = \frac{1}{1 - h_n A} e_n + \frac{O(h_n^2)}{1 - h_n A}. \quad (5.16)$$

Thus the propagated error is damped whenever $|1 - h_n A| > 1$. Since $A < 0$, we see that there is no restriction at all on the step size caused by the stability concern.

5.2 Stability Issues in Stiff Systems

We have seen that a basic difficulty (but not the only one) in the numerical solution of stiff ordinary differential systems is the satisfaction of the required absolute stability. Thus several definitions that call for methods to possess some substantially larger region of absolute stability have been proposed.

Definition 5.2.1 *A numerical method is said to be A-stable if its region of absolute stability contains the entire left-half complex plane for \bar{h} .*

Example. The R -stage implicit Runge-Kutta method of order $2R$ is always A-stable. (Review §2.4.)

Theorem 5.2.1 (*Dahlquist*)

1. An explicit linear multistep method cannot be A-stable.
2. The order of an A-stable implicit linear multistep method cannot exceed 2.
3. The second order A-stable implicit linear multistep method with smallest error constant is the Trapezoidal rule (with error constant $-\frac{1}{12}$).

Definition 5.2.2 *A numerical method is said to be $A(\alpha)$ -stable, $\alpha \in (0, \frac{\pi}{2})$, if its region of absolute stability contains the infinite wedge*

$$W_\alpha := \{\bar{h} \mid -\alpha < \pi - \arg \bar{h} < \alpha\}.$$

It is said to be $A(0)$ -stable if it is $A(\alpha)$ -stable for some sufficiently small $\alpha \in (0, \frac{\pi}{2})$.

Theorem 5.2.2 (*Widlund*)

1. An explicit linear multistep method cannot be $A(0)$ -stable.
2. There is only one $A(0)$ -stable linear k -step method whose order exceeds $k + 1$, namely, the Trapezoidal rule.
3. For all $\alpha \in (0, \frac{\pi}{2})$, there exists an $A(\alpha)$ -stable linear k -step method of order p for which $k = p = 3$ or $k = p = 4$.

Definition 5.2.3 A numerical method is said to be *stiffly stable* if in the region

$$R_1 := \{\bar{h}|\Re\bar{h} \leq a < 0\},$$

the method is absolutely stable, and in the region

$$R_2 := \{\bar{h}|a < \Re\bar{h} < 0 < \alpha, |\Im\bar{h}| < \theta\},$$

the method is accurate.

Theorem 5.2.3 (Gear)

1. The k -step method of order k with $\sigma(r) = \beta_0 r^k$ are stiffly stable for $k = 1, \dots, 6$. (These class of methods are called the backward differentiation formulas (BDF).)
2. The same class of BDF methods are not stiffly stable for $k = 7, \dots, 15$.
3. There are stiffly stable multistep methods of orders up to 11.

5.3 Backward Differentiation Formula

From theorems in the previous section, we conclude that the class of linear multistep methods in general does not provide good linear stability properties for stiff systems, except possibly the so called *backward differentiation formulas*.

Definition 5.3.1 A multistep method of the form

$$\sum_{i=0}^k \alpha_i y_{n-i} = h\beta_0 f_n \quad (5.17)$$

with $\beta_0 \neq 0$ is called a *backward differentiation formula (BDF)*.

The class of BDF's can be viewed to be the *dual* of the class of Adams-Moulton methods. Recall that the Adams methods are characterized by their first characteristic polynomial $\rho(\xi) = \xi^k - \xi^{k-1}$. In contrast, the BDF's have the simplest possible second characteristic polynomial $\sigma(\xi) = \beta_0 \xi^k$. Furthermore, recall that the Adams-Moulton methods are derived from the integral

$$y(x_n) - y(x_{n-1}) = \int_{x_{n-1}}^{x_n} y'(x) dx = \int_{x_{n-1}}^{x_n} f(x, y(x)) dx$$

where the integrand $f(x, y(x))$ is approximated by a polynomial with interpolating data (x_{n-i}, f_{n-i}) for $i = 0, 1, \dots, k$. In the BDF, we start from the differential system itself

$$y' = f(x, y) \quad (5.18)$$

and differentiate the k -th degree polynomial that interpolates the data (x_{n-i}, y_{n-i}) for $i = 0, 1, \dots, k$. More precisely, the Newton's divided difference formula affirms that the polynomial is given by

$$I(x) = y[x_n] + y[x_n, x_{n-1}](x - x_n) + \dots + y[x_n, x_{n-1}, \dots, x_{n-k}](x - x_n)(x - x_{n-1}) \cdots (x - x_{n-k}) \quad (5.19)$$

where the divided differences are defined by:

$$y[x_n] := y_n \quad (5.20)$$

$$y[x_n, x_{n-1}] := \frac{y[x_n] - y[x_{n-1}]}{x_n - x_{n-1}} \quad (5.21)$$

$$y[x_n, \dots, x_{n-k}] := \frac{y[x_{n-1}, \dots, x_{n-k}] - y[x_n, \dots, x_{n-k+1}]}{x_n - x_{n-k}}. \quad (5.22)$$

The BDF is obtained by replacing the left-hand side of (5.18) by $I'(x_n)$ and the right-hand side by f_n .

It can be proved, as for Adams-Moulton methods, that a k -step BDF is order k . Obviously, the theory of convergence, consistency and stability that we have discussed for general multistep methods applies to the class of BDF as well. In particular, we have the following theorem:

Theorem 5.3.1 (Cryer) *For $k = 1, 2, \dots, 6$, the methods of BDF are zero-stable, but for $k \geq 7$ the methods are all zero-unstable.*

The coefficients of the BDF together with the error constants are given in the table.

k	1	2	3	4	5	6
α_0	1	1	1	1	1	1
α_1	-1	$-\frac{4}{3}$	$-\frac{18}{11}$	$-\frac{48}{25}$	$-\frac{300}{137}$	$-\frac{360}{147}$
α_2		$\frac{1}{3}$	$\frac{9}{11}$	$\frac{36}{25}$	$\frac{300}{137}$	$\frac{450}{147}$
α_3			$-\frac{2}{11}$	$-\frac{16}{25}$	$-\frac{200}{137}$	$-\frac{400}{137}$
α_4				$\frac{3}{25}$	$\frac{75}{137}$	$\frac{225}{147}$
α_5					$-\frac{12}{147}$	$-\frac{72}{147}$
α_6						$\frac{10}{147}$
β_0	1	$\frac{2}{3}$	$\frac{6}{11}$	$\frac{12}{25}$	$\frac{60}{137}$	$\frac{60}{147}$
c_{k+1}	$-\frac{1}{2}$	$-\frac{2}{9}$	$-\frac{3}{22}$	$-\frac{12}{125}$	$-\frac{10}{137}$	$-\frac{20}{343}$

The most important feature of the BDF is the size of their regions of absolute stability. For $1 \leq k \leq 6$ these regions contain the whole of the negative real axis. In particular, these methods are stiffly-stable.

5.4 Handling Implicitness

Since no explicit multistep or explicit Runge-Kutta method can possess good stability properties. Solving an implicit algebraic is a necessary task when dealing stiff systems.

One might wonder if a *PC* method will facilitate the computation. Consider a general implicit multistep method

$$\sum_{i=0}^k \alpha_i y_{n-i} = h \sum_{i=0}^k \beta_i f_{n-i} \quad (5.23)$$

applied to a general problem $y' = f(x, y)$. Rewrite (5.23) in the form

$$y_n = h\beta_0 f(x_n, y_n) + \Psi_n \quad (5.24)$$

where

$$\Psi_n := \sum_{i=1}^k (-\alpha_i y_{n-i} + \beta_i f_{n-i}) \quad (5.25)$$

is a known quantity from past values. Recall that a predictor-corrector method is a fixed point iteration

$$y_n^{(s+1)} = h\beta_0 f(x_n, y_n^{(s)}) + \Psi_n. \quad (5.26)$$

Such an iteration will converge if

$$h < \frac{1}{|\beta_0|L} \quad (5.27)$$

where L is the Lipschitz constant of f with respect to y . For differentiable f , L can be estimated from the largest modulus of eigenvalues of $\frac{\partial f}{\partial y}$. In particular, if the system is stiff, then we expect that L is large, implying that h must be small. This restriction inherited in the iteration (5.26) persists regardless whether the corrector used in the iteration has a better region of absolute stability. Consequently, the only way out of this difficulty is to abandon fixed point iteration in favor of Newton iteration. The cost of computing the Jacobian of the system (5.24) and solving the resulting linear equation is, of course, quite expensive. In order to cut down the cost, a modified Newton iteration such as

$$\left(I - h\beta_0 \frac{\partial f}{\partial y}(x_n, y_n^{(0)}) \right) \delta y_n^{(s)} = -y_n^{(s)} + h\beta_0 f(x_n, y_n^{(s)}) + \Psi_n \quad (5.28)$$

is sometimes used. It is quite common to keep the Jacobian matrix constant not only throughout the iteration, but to use it for the next one or two integration

step as well. The matrix is updated and a new LU decomposition computed only when the iteration fails to converge.

Since calculating the Jacobian of f is inevitable for solving stiff systems, one might consider using the Jacobian information in the method itself. One such approach is the so called Obrechhoff method,

$$\sum_{i=0}^k \alpha_i y_{n-i} = h \sum_{i=0}^k \beta_i f_{n-i} + h^2 \sum_{i=0}^k \gamma_i \frac{\partial f_{n-i}}{\partial y}. \quad (5.29)$$

That is, we consider a linear multistep method that involves derivatives of y . In particular, the subclass of methods

$$y_n - y_{n-1} = h \sum_{i=0}^k \beta_i f_{n-i} + h^2 \gamma_0 \frac{\partial f_n}{\partial y} \quad (5.30)$$

can be shown to be stiffly stable for $k \leq 7$ with error constants much smaller than the corresponding BDF.