Chapter 6

Zero-Stability and Convergence for Initial Value Problems

6.1 Convergence

To discuss the convergence of a numerical method for the initial value problem, we focus on a fixed (but arbitrary) time $T > 0$ and consider the error in our approximation to $u(T)$ computed with the method using time step k. The method converges on this problem if this error goes to zero as $k \to 0$. Note that the number of time steps that we need to take to reach time T increases as $k \to 0$. If we use N to denote this value ($N = T/k$), then convergence means that

$$\lim_{\substack{k\to 0\\ Nk=T}} U^N = u(T). \tag{6.1}$$

In principle a method might converge on one problem but not on another, or converge with one set of starting values but not with another set. To speak of a *method* being *convergent* in general, we require that it converges on *all* problems in a reasonably large class with *all* reasonable starting values. For an r-step method we need r starting values. These values will typically depend on k, and to make this clear we will write them as $U^0(k),\ U^1(k),\ \ldots,\ U^{r-1}(k)$. While these will generally approximate $u(t)$ at the times $t_0 = 0,\ t_1 = k, \ldots,\ t_{r-1} = (r-1)k$, respectively, as $k \to 0$, each of these times approaches $t_0 = 0$. So the weakest condition we might put on our starting values is that they converge to the correct initial value η as $k \to 0$:

$$\lim_{k\to 0} U^\nu(k) = \eta \quad \text{for } \nu = 0,\ 1,\ \ldots,\ r-1. \tag{6.2}$$

We can now state the definition of convergence.

Definition 6.1. *An r-step method is said to be* convergent *if applying the method to any ODE* (5.1) *with $f(u,t)$ Lipschitz continuous in u, and with any set of starting values satisfying* (6.2)*, we obtain convergence in the sense of* (6.1) *for every fixed time $T > 0$ at which the ODE has a unique solution.*

To be convergent, a method must be *consistent*, meaning as before that the local truncation error (LTE) is $o(1)$ as $k \to 0$, and also *zero-stable*, as described later in this chapter. We will begin to investigate these issues by first proving the convergence of one-step methods, which turn out to be zero-stable automatically. We start with Euler's method on linear problems, then consider Euler's method on general nonlinear problems and finally extend this to a wide class of one-step methods.

6.2 The test problem

Much of the theory presented below is based on examining what happens when a method is applied to a simple scalar linear equation of the form

$$u'(t) = \lambda u(t) + g(t) \tag{6.3}$$

with initial data

$$u(t_0) = \eta.$$

The solution is then given by Duhamel's principle (5.8),

$$u(t) = e^{\lambda(t-t_0)}\eta + \int_{t_0}^{t} e^{\lambda(t-\tau)} g(\tau)\, d\tau. \tag{6.4}$$

6.3 One-step methods

6.3.1 Euler's method on linear problems

If we apply Euler's method to (6.3), we obtain

$$\begin{aligned} U^{n+1} &= U^n + k(\lambda U^n + g(t_n)) \\ &= (1 + k\lambda)U^n + kg(t_n). \end{aligned} \tag{6.5}$$

The LTE for Euler's method is given by

$$\begin{aligned} \tau^n &= \left(\frac{u(t_{n+1}) - u(t_n)}{k}\right) - (\lambda u(t_n) + g(t_n)) \\ &= \left(u'(t_n) + \frac{1}{2}ku''(t_n) + O(k^2)\right) - u'(t_n) \\ &= \frac{1}{2}ku''(t_n) + O(k^2). \end{aligned} \tag{6.6}$$

Rewriting this equation as

$$u(t_{n+1}) = (1 + k\lambda)u(t_n) + kg(t_n) + k\tau^n$$

and subtracting this from (6.5) gives a difference equation for the global error E^n:

$$E^{n+1} = (1 + k\lambda)E^n - k\tau^n. \tag{6.7}$$

Note that this has exactly the same form as (6.5) but with a different nonhomogeneous term: $-\tau^n$ in place of $g(t_n)$. This is analogous to equation (2.15) in the boundary value theory and again gives the relation we need between the local truncation error τ^n (which is easy to compute) and the global error E^n (which we wish to bound). Note again that *linearity* plays a critical role in making this connection. We will consider nonlinear problems below.

Because the equation and method we are now considering are both so simple, we obtain an equation (6.7) that we can explicitly solve for the global error E^n. Applying the recursion (6.7) repeatedly we see what form the solution should take:

$$\begin{aligned} E^n &= (1+k\lambda)E^{n-1} - k\tau^{n-1} \\ &= (1+k\lambda)[(1+k\lambda)E^{n-2} - k\tau^{n-2}] - k\tau^{n-1} \\ &= \cdots . \end{aligned}$$

By induction we can easily confirm that in general

$$E^n = (1+k\lambda)^n E^0 - k\sum_{m=1}^{n}(1+k\lambda)^{n-m}\tau^{m-1}. \tag{6.8}$$

(Note that some of the superscripts are powers while others are indices!) This has a form that is very analogous to the solution (6.4) of the corresponding ordinary differential equation (ODE), where now $(1+k\lambda)^{n-m}$ plays the role of the solution operator of the homogeneous problem—it transforms data at time t_m to the solution at time t_n. The expression (6.8) is sometimes called the discrete form of Duhamel's principle.

We are now ready to prove that Euler's method converges on (6.3). We need only observe that

$$|1+k\lambda| \le e^{k|\lambda|} \tag{6.9}$$

and so

$$(1+k\lambda)^{n-m} \le e^{(n-m)k|\lambda|} \le e^{nk|\lambda|} \le e^{|\lambda|T}, \tag{6.10}$$

provided that we restrict our attention to the finite time interval $0 \le t \le T$, so that $t_n = nk \le T$. It then follows from (6.8) that

$$\begin{aligned} |E^n| &\le e^{|\lambda|T}\left(|E^0| + k\sum_{m=1}^{n}|\tau^{m-1}|\right) \\ &\le e^{|\lambda|T}\left(|E^0| + nk\max_{1\le m\le n}|\tau^{m-1}|\right). \end{aligned} \tag{6.11}$$

Let $N = T/k$ be the number of time steps needed to reach time T and set

$$\|\tau\|_\infty = \max_{0\le n\le N-1}|\tau^n|.$$

From (6.6) we expect

$$\|\tau\|_\infty \approx \frac{1}{2}k\|u''\|_\infty = O(k),$$

where $\|u''\|_\infty$ is the maximum value of the function u'' over the interval $[0, T]$. Then for $t = nk \le T$, we have from (6.11) that

$$|E^n| \le e^{|\lambda|T}(|E^0| + T\|\tau\|_\infty).$$

If (6.2) is satisfied then $E^0 \to 0$ as $k \to 0$. In fact for this one-step method we would generally take $U^0 = u(0) = \eta$, in which case E^0 drops out and we are left with

$$|E^n| \le e^{|\lambda|T} T \|\tau\|_\infty = O(k) \quad \text{as } k \to 0 \tag{6.12}$$

and hence the method converges and is in fact first order accurate.

Note where *stability* comes into the picture. The one-step error $\mathcal{L}^{m-1} = k\tau^{m-1}$ introduced in the mth step contributes the term $(1 + k\lambda)^{n-m}\mathcal{L}^{m-1}$ to the global error. The fact that $|(1 + k\lambda)^{n-m}| < e^{|\lambda|T}$ is uniformly bounded as $k \to 0$ allows us to conclude that each contribution to the final error can be bounded in terms of its original size as a one-step error. Hence the "naive analysis" of Section 5.5 is valid, and the global error has the same order of magnitude as the local truncation error.

6.3.2 Relation to stability for boundary value problems

To see how this ties in with the definition of stability used in Chapter 2 for the BVP, it may be useful to view Euler's method as giving a linear system in matrix form, although this is not the way it is used computationally. If we view the equations (6.5) for $n = 0, 1, \ldots, N-1$ as a linear system $AU = F$ for $U = [U^1, U^2, \ldots, U^N]^T$, then

$$A = \frac{1}{k}\begin{bmatrix} 1 & & & & & \\ -(1+k\lambda) & 1 & & & & \\ & -(1+k\lambda) & 1 & & & \\ & & & \ddots & & \\ & & & -(1+k\lambda) & 1 & \\ & & & & -(1+k\lambda) & 1 \end{bmatrix}$$

and

$$U = \begin{bmatrix} U^1 \\ U^2 \\ U^3 \\ \vdots \\ U^{N-1} \\ U^N \end{bmatrix}, \quad F = \begin{bmatrix} (1/k + \lambda)U^0 + g(t_0) \\ g(t_1) \\ g(t_2) \\ \vdots \\ g(t_{N-2}) \\ g(t_{N-1}) \end{bmatrix}.$$

We have divided both sides of (6.5) by k to conform to the notation of Chapter 2. Since the matrix A is lower triangular, this system is easily solved by forward substitution, which results in the iterative equation (6.5).

If we now let $\hat{U}$ be the vector obtained from the true solution as in Chapter 2, then subtracting $A\hat{U} = F + \tau$ from $AU = F$, we obtain (2.15) (the matrix form of (6.7)) with solution (6.8). We are then in exactly the same framework as in Chapter 2. So we have convergence and a global error with the same magnitude as the local error provided that the method is stable in the sense of Definition 2.1, i.e., that the inverse of the matrix A is bounded independent of k for all k sufficiently small.

The inverse of this matrix is easy to compute. In fact we can see from the solution (6.8) that

$$A^{-1} = k\begin{bmatrix} 1 \\ (1+k\lambda) & 1 \\ (1+k\lambda)^2 & (1+k\lambda) & 1 \\ (1+k\lambda)^3 & (1+k\lambda)^2 & (1+k\lambda) & 1 \\ \vdots & & & & \ddots \\ (1+k\lambda)^{N-1} & (1+k\lambda)^{N-2} & (1+k\lambda)^{N-3} & \cdots & (1+k\lambda) & 1 \end{bmatrix}.$$

We easily compute using (A.10a) that

$$\|A^{-1}\|_\infty = k\sum_{m=1}^{N} |(1+k\lambda)^{N-m}|$$

and so

$$\|A^{-1}\|_\infty \le kNe^{|\lambda|T} = Te^{|\lambda|T}.$$

This is uniformly bounded as $k \to 0$ for fixed T. Hence the method is stable and $\|E\|_\infty \le \|A^{-1}\|_\infty \|\tau\|_\infty \le Te^{|\lambda|T}\|\tau\|_\infty$, which agrees with the bound (6.12).

6.3.3 Euler's method on nonlinear problems

So far we have focused entirely on linear equations. Practical problems are almost always nonlinear, but for the initial value problem it turns out that it is not significantly more difficult to handle this case if we assume that $f(u)$ is Lipschitz continuous, which is reasonable in light of the discussion in Section 5.2.

Euler's method on $u' = f(u)$ takes the form

$$U^{n+1} = U^n + kf(U^n) \tag{6.13}$$

and the truncation error is defined by

$$\begin{aligned}\tau^n &= \frac{1}{k}(u(t_{n+1}) - u(t_n)) - f(u(t_n)) \\ &= \frac{1}{2}ku''(t_n) + O(k^2),\end{aligned}$$

just as in the linear case. So the true solution satisfies

$$u(t_{n+1}) = u(t_n) + kf(u(t_n)) + k\tau^n$$

and subtracting this from (6.13) gives

$$E^{n+1} = E^n + k(f(U^n) - f(u(t_n))) - k\tau^n. \tag{6.14}$$

In the linear case $f(U^n) - f(u(t_n)) = \lambda E^n$ and we get the relation (6.7) for E^n. In the nonlinear case we cannot express $f(U^n) - f(u(t_n))$ directly in terms of the error E^n in general. However, using the Lipschitz continuity of f we can get a bound on this in terms of E^n:

$$|f(U^n) - f(u(t_n))| \le L|U^n - u(t_n)| = L|E^n|.$$

Using this in (6.14) gives

$$|E^{n+1}| \le |E^n| + kL|E^n| + k|\tau^n| = (1+kL)|E^n| + k|\tau^n|. \tag{6.15}$$

From this inequality we can show by induction that

$$|E^n| \le (1+kL)^n|E^0| + k\sum_{m=1}^{n}(1+kL)^{n-m}|\tau^{m-1}|$$

and so, using the same steps as in obtaining (6.12) (and again assuming $E^0 = 0$), we obtain

$$|E^n| \le e^{LT}T\|\tau\|_\infty = O(k) \quad \text{as } k \to 0 \tag{6.16}$$

for all n with $nk \le T$, proving that the method converges. In the linear case $L = |\lambda|$ and this reduces to exactly (6.12).

6.3.4 General one-step methods

A general explicit one-step method takes the form

$$U^{n+1} = U^n + k\Psi(U^n, t_n, k) \tag{6.17}$$

for some function Ψ, which depends on f of course. We will assume that $\Psi(u,t,k)$ is continuous in t and k and Lipschitz continuous in u, with Lipschitz constant L' that is generally related to the Lipschitz constant of f.

Example 6.1. For the two-stage Runge–Kutta method of Example 5.11, we have

$$\Psi(u,t,k) = f\left(u + \frac{1}{2}kf(u)\right). \tag{6.18}$$

If f is Lipschitz continuous with Lipschitz constant L, then Ψ has Lipschitz constant $L' = L + \frac{1}{2}kL^2$.

The one-step method (6.17) is *consistent* if

$$\Psi(u,t,0) = f(u,t)$$

for all u, t, and Ψ is continuous in k. The local truncation error is

$$\tau^n = \left(\frac{u(t_{n+1}) - u(t_n)}{k}\right) - \Psi(u(t_n), t_n, k).$$

We can show that any one-step method satisfying these conditions is convergent. We have

$$u(t_{n+1}) = u(t_n) + k\Psi(u(t_n), t_n, k) + k\tau^n$$

and subtracting this from (6.17) gives

$$E^{n+1} = E^n + k\left(\Psi(U^n, t_n, k) - \Psi(u(t_n), t_n, k)\right) - k\tau^n.$$

Using the Lipschitz condition we obtain

$$|E^{n+1}| \le |E^n| + kL'|E^n| + k|\tau^n|.$$

This has exactly the same form as (6.15) and the proof of convergence proceeds exactly as from there.

6.4 Zero-stability of linear multistep methods

The convergence proof of the previous section shows that for one-step methods, each one-step error $k\tau^{m-1}$ has an effect on the global error that is bounded by $e^{L'T}|k\tau^{m-1}|$. Although the error is possibly amplified by a factor $e^{L'T}$, this factor is bounded independent of k as $k \to 0$. Consequently the method is stable: the global error can be bounded in terms of the sum of all the one-step errors and hence has the same asymptotic behavior as the LTE as $k \to 0$. This form of stability is often called *zero-stability* in ODE theory, to distinguish it from other forms of stability that are of equal importance in practice. The fact that a method is zero-stable (and converges as $k \to 0$) is no guarantee that it will give reasonable results on the particular grid with $k > 0$ that we want to use in practice. Other "stability" issues of a different nature will be taken up in the next chapter.

But first we will investigate the issue of zero-stability for general LMMs, where the theory of the previous section does not apply directly. We begin with an example showing a consistent LMM that is *not* convergent. Examining what goes wrong will motivate our definition of zero-stability for LMMs.

Example 6.2. The LMM

$$U^{n+2} - 3U^{n+1} + 2U^n = -kf(U^n) \tag{6.19}$$

has an LTE given by

$$\tau^n = \frac{1}{k}[u(t_{n+2}) - 3u(t_{n+1}) + 2u(t_n) + ku'(t_n)] = \frac{5}{2}ku''(t_n) + O(k^2),$$

so the method is consistent and "first order accurate." But in fact the global error will not exhibit first order accuracy, or even convergence, in general. This can be seen even on the trivial initial-value problem

$$u'(t) = 0, \quad u(0) = 0 \tag{6.20}$$

with solution $u(t) \equiv 0$. In this problem, equation (6.19) takes the form

$$U^{n+2} - 3U^{n+1} + 2U^n = 0. \tag{6.21}$$

We need two starting values U^0 and U^1. If we take $U^0 = U^1 = 0$, then (6.21) generates $U^n = 0$ for all n and in this case we certainly converge to correct solution, and in fact we get the exact solution for any k.

But in general we will not have the exact value U^1 available and will have to approximate this, introducing some error into the computation. Table 6.1 shows results obtained by applying this method with starting data $U^0 = 0, U^1 = k$. Since $U^1(k) \to 0$ as $k \to 0$, this is valid starting data in the context of Definition 6.1 of convergence. If the method is convergent, we should see that U^N, the computed solution at time $T = 1$, converges to zero as $k \to 0$. Instead it blows up quite dramatically. Similar results would be seen if we applied this method to an arbitrary equation $u' = f(u)$ and used any one-step method to compute U^1 from U^0.

The homogeneous linear difference equation (6.21) can be solved explicitly for U^n in terms of the starting values U^0 and U^1. We obtain

$$U^n = 2U^0 - U^1 + 2^n(U^1 - U^0). \tag{6.22}$$

Table 6.1. *Solution* U^N *to* (6.21) *with* $U^0 = 0$, $U^1 = k$ *and various values of* $k = 1/N$.

N	U^N
5	6.2
10	1023
20	5.4×10^4

It is easy to verify that this satisfies (6.21) and also the starting values. (We'll see how to solve general linear difference equations in the next section.)

Since $u(t) = 0$, the error is $E^n = U^n$ and we see that any initial errors in U^1 or U^0 are magnified by a factor 2^n in the global error (except in the special case $U^1 = U^0$). This exponential growth of the error is the instability that leads to nonconvergence. To rule out this sort of growth of errors, we need to be able to solve a general linear difference equation.

6.4.1 Solving linear difference equations

We briefly review one solution technique for linear difference equations; see Section D.2.1 for a different approach. Consider the general homogeneous linear difference equation

$$\sum_{j=0}^{r} \alpha_j U^{n+j} = 0. \tag{6.23}$$

Eventually we will look for a particular solution satisfying given initial conditions

$$U^0, U^1, \dots, U^{r-1},$$

but to begin with we will find the general solution of the difference equation in terms of r free parameters. We will hypothesize that this equation has a solution of the form

$$U^n = \zeta^n \tag{6.24}$$

for some value of ζ (here ζ^n is the nth power!). Plugging this into (6.23) gives

$$\sum_{j=0}^{r} \alpha_j \zeta^{n+j} = 0$$

and dividing by ζ^n yields

$$\sum_{j=0}^{r} \alpha_j \zeta^j = 0. \tag{6.25}$$

We see that (6.24) is a solution of the difference equation if ζ satisfies (6.25), i.e., if ζ is a root of the polynomial

$$\rho(\zeta) = \sum_{j=0}^{r} \alpha_j \zeta^j.$$

Note that this is just the first characteristic polynomial of the LMM introduced in (5.49). In general $\rho(\zeta)$ has r roots $\zeta_1, \zeta_2, \ldots, \zeta_r$ and can be factored as

$$\rho(\zeta) = \alpha_r(\zeta - \zeta_1)(\zeta - \zeta_2)\cdots(\zeta - \zeta_r).$$

Since the difference equation is linear, any linear combination of solutions is again a solution. If $\zeta_1, \zeta_2, \ldots, \zeta_r$ are distinct ($\zeta_i \neq \zeta_j$ for $i \neq j$), then the r distinct solutions ζ_i^n are linearly independent and the general solution of (6.23) has the form

$$U^n = c_1\zeta_1^n + c_2\zeta_2^n + \cdots + c_r\zeta_r^n, \tag{6.26}$$

where $c_1, \ldots, c_r$ are arbitrary constants. In this case, every solution of the difference equation (6.23) has this form. If initial conditions $U^0, U^1, \ldots, U^{r-1}$ are specified, then the constants $c_1, \ldots, c_r$ can be uniquely determined by solving the $r \times r$ linear system

$$\begin{aligned} c_1 + c_2 + \cdots + c_r &= U^0, \\ c_1\zeta_1 + c_2\zeta_2 + \cdots + c_r\zeta_r &= U^1, \\ \vdots\quad &\quad \vdots \\ c_1\zeta_1^{r-1} + c_2\zeta_2^{r-1} + \cdots + c_r\zeta_r^{r-1} &= U^{r-1}. \end{aligned} \tag{6.27}$$

Example 6.3. The characteristic polynomial for the difference equation (6.21) is

$$\rho(\zeta) = 2 - 3\zeta + \zeta^2 = (\zeta - 1)(\zeta - 2) \tag{6.28}$$

with roots $\zeta_1 = 1,\ \zeta_2 = 2$. The general solution has the form

$$U^n = c_1 + c_2 \cdot 2^n$$

and solving for c_1 and c_2 from U^0 and U^1 gives the solution (6.22).

This example indicates that if $\rho(\zeta)$ has any roots that are greater than one in modulus, the method will not be convergent. It turns out that the converse is nearly true: if all the roots have modulus no greater than one, then the method is convergent, with one proviso. There must be no *repeated* roots with modulus equal to one. The next two examples illustrate this.

If the roots are not distinct, say, $\zeta_1 = \zeta_2$ for simplicity, then ζ_1^n and ζ_2^n are not linearly independent and the U^n given by (6.26), while still a solution, is not the most general solution. The system (6.27) would be singular in this case. In addition to ζ_1^n there is also a solution of the form $n\zeta_1^n$ and the general solution has the form

$$U^n = c_1\zeta_1^n + c_2 n\zeta_1^n + c_3\zeta_3^n + \cdots + c_r\zeta_r^n.$$

If in addition $\zeta_3 = \zeta_1$, then the third term would be replaced by $c_3 n^2 \zeta_1^n$. Similar modifications are made for any other repeated roots. Note how similar this theory is to the standard solution technique for an rth order linear ODE.

Example 6.4. Applying the consistent LMM

$$U^{n+2} - 2U^{n+1} + U^n = \frac{1}{2}k(f(U^{n+2}) - f(U^n)) \tag{6.29}$$

to the differential equation $u'(t) = 0$ gives the difference equation

$$U^{n+2} - 2U^{n+1} + U^n = 0.$$

The characteristic polynomial is

$$\rho(\zeta) = \zeta^2 - 2\zeta + 1 = (\zeta - 1)^2 \tag{6.30}$$

so $\zeta_1 = \zeta_2 = 1$. The general solution is

$$U^n = c_1 + c_2 n.$$

For particular starting values U^0 and U^1 the solution is

$$U^n = U^0 + (U^1 - U^0)n.$$

Again we see that the solution grows with n, although not as dramatically as in Example 6.2 (the growth is linear rather than exponential). But this growth is still enough to destroy convergence. If we take the same starting values as before, $U^0 = 0$ and $U^1 = k$, then $U^n = kn$ and so

$$\lim_{\substack{k\to 0\\ Nk=T}} U^N = kN = T.$$

The method converges to the function $v(t) = t$ rather than to $u(t) = 0$, and hence the LMM (6.29) is not convergent.

This example shows that if $\rho(\zeta)$ has a *repeated* root of modulus 1, then the method cannot be convergent.

Example 6.5. Now consider the consistent LMM

$$U^{n+3} - 2U^{n+2} + \frac{5}{4}U^{n+1} - \frac{1}{4}U^n = \frac{1}{4}hf(U^n). \tag{6.31}$$

Applying this to (6.20) gives

$$U^{n+3} - 2U^{n+2} + \frac{5}{4}U^{n+1} - \frac{1}{4}U^n = 0$$

and the characteristic polynomial is

$$\rho(\zeta) = \zeta^3 - 2\zeta^2 + \frac{5}{4}\zeta - \frac{1}{4} = (\zeta - 1)(\zeta - 0.5)^2. \tag{6.32}$$

So $\zeta_1 = 1,\ \zeta_2 = \zeta_3 = 1/2$ and the general solution is

$$U^n = c_1 + c_2\left(\frac{1}{2}\right)^n + c_3 n\left(\frac{1}{2}\right)^n.$$

Here there is a repeated root but with modulus less than 1. The linear growth of n is then overwhelmed by the decay of $(1/2)^n$.

For this three-step method we need three starting values $U^0,\ U^1,\ U^2$ and we can find $c_1,\ c_2,\ c_3$ in terms of them by solving a linear system similar to (6.27). Each c_i will be a linear combination of U^0, U^1, U^2 and so if $U^\nu(k) \to 0$ as $k \to 0$, then $c_i(k) \to 0$ as $k \to 0$ also. The value U^N computed at time T with step size k (where $kN = T$) has the form

$$U^N = c_1(k) + c_2(k)\left(\frac{1}{2}\right)^N + c_3(k)N\left(\frac{1}{2}\right)^N. \tag{6.33}$$

Now we see that

$$\lim_{\substack{k\to 0\\ Nk=T}} U^N = 0$$

and so the method (6.31) converges on $u' = 0$ with arbitrary starting values $U^\nu(k)$ satisfying $U^\nu(k) \to 0$ as $k \to 0$. (In fact, this LMM is convergent in general.)

More generally, if $\rho(\zeta)$ has a root ζ_j that is repeated m times, then U^N will involve terms of the form $N^s\zeta_j^N$ for $s = 0, 1, \dots, m-1$. This converges to zero as $N \to \infty$ provided $|\zeta_j| < 1$. The algebraic growth of N^s is overwhelmed by the exponential decay of ζ_j^N. This shows that repeated roots are not a problem as long as they have magnitude strictly less than 1.

With the above examples as motivation, we are ready to state the definition of zero-stability.

Definition 6.2. *An r-step LMM is said to be zero-stable if the roots of the characteristic polynomial $\rho(\zeta)$ defined by* (5.49) *satisfy the following conditions:*

$$\begin{aligned} &|\zeta_j| \le 1 \quad \textit{for } j = 1,\ 2,\ ,\dots,\ r. \\ &\textit{If } \zeta_j \textit{ is a repeated root, then } |\zeta_j| < 1. \end{aligned} \tag{6.34}$$

If the conditions (6.34) are satisfied for all roots of ρ, then the polynomial is said to satisfy the *root condition.*

Example 6.6. The Adams methods have the form

$$U^{n+r} = U^{n+r-1} + k\sum_{j=1}^{r}\beta_j f(U^{n+j})$$

and hence

$$\rho(\zeta) = \zeta^r - \zeta^{r-1} = (\zeta - 1)\zeta^{r-1}.$$

The roots are $\zeta_1 = 1$ and $\zeta_2 = \cdots = \zeta_r = 0$. The root condition is clearly satisfied and all the Adams–Bashforth and Adams–Moulton methods are zero-stable.

The given examples certainly do not prove that zero-stability as defined above is a sufficient condition for convergence. We looked at only the simplest possible ODE $u'(t) = 0$ and saw that things could go wrong if the root condition is *not* satisfied. It turns out, however, that the root condition is all that is needed to prove convergence on the general initial value problem (in the sense of Definition 6.1).

Theorem 6.3 (Dahlquist [22]). *For LMMs applied to the initial value problem for $u'(t) = f(u(t), t)$,*

$$\textit{consistency} + \textit{zero-stability} \iff \textit{convergence}. \tag{6.35}$$

This is the analogue of the statement (2.21) for the BVP. A proof of this result can be found in [43].

Note: A consistent LMM always has one root equal to 1, say, $\zeta_1 = 1$, called the *principal root*. This follows from (5.50). Hence a consistent one-step LMM (such as Euler, backward Euler, trapezoidal) is certainly zero-stable. More generally we have proved in Section 6.3.4 that any consistent one-step method (that is a Lipschitz continuous) is convergent. Such methods are automatically "zero-stable" and behave well as $k \to 0$. We can think of zero-stability as meaning "stable in the limit as $k \to 0$."

Although a consistent zero-stable method is convergent, it may have other stability problems that show up if the time step k is chosen too large in an actual computation. Additional stability considerations are the subject of the next chapter.