13.1. **Numerical instability.** Question: How high an order can be achieved in a $(p+1)$ step method if it is consistent and zero-stable? In seeking high order methods, we automatically get consistency; zero-stability poses a more difficult constraint. Recall that a $(p+1)$ step method has $(2p+3)$ coefficients, $p+1$ a_is and $p+2$ b_i's. If the method is explicit, this number is reduced by one. Hence, we can expect at most order $2p+2$ for an implicit method and $2p+1$ for an explicit method (recall that if a method has order r, it is exact for all polynomials of degree $\leq r$). However, the following result is known.

Theorem 18. *(Dahlquist) No zero-stable $p+1$ step linear multistep method can have order exceeeding $p+2$ when p is even or exceeding $p+3$ when p is odd.*

A zero-stable $p+1$ step method which has order $p+3$ is called an optimal method. It can be shown that for an optimal method, all the roots of $\rho(z)$ lie on the unit circle.

Example: Simpson's rule: $y_{n+1} = y_{n-1} + (h/3)[f_{n+1} + 4f_n + f_{n-1}]$. Since $p = 1$, this is a two-step method. The local truncation error is $-(1/90)h^5 y^{(5)}(\xi)$. It is a fourth order method, so Simpson's rule is an optimal method. However, we shall see that Simpson's rule has computational disadvantages that make it unsuitable as a general purpose method. These disadvantages are shared by all optimal order methods. Hence, we will not choose the coefficients in a multistep method solely to achieve maximum order.

To understand this issue, consider the problem $y' = -y$, $y(0) = 1$, whose exact solution is $y(x) = e^{-x}$. We apply the midpoint rule method $y_{n+1} = y_{n-1} + 2hf_n$, which in this case becomes $y_{n+1} + 2hy_n - y_{n-1} = 0$. Since this is a linear difference equation with constant coefficients, we solve it by first computing the roots of the characteristic polynomial $\rho(z) = z^2 + 2hz - 1 = 0$. Then $z = -h \pm \sqrt{1 + h^2}$, so the general solution has the form

$$y_n = C_1(-h + \sqrt{1 + h^2})^n + C_2(-h - \sqrt{1 + h^2})^n.$$

Set $y_0 = 1$ and leave y_1 arbitrary for the moment and solve for C_1 and C_2.

$$y_0 = 1 = C_1 + C_2, \qquad y_1 = C_1(-h + \sqrt{1 + h^2}) + C_2(-h - \sqrt{1 + h^2}).$$

Then

$$C_1 = \frac{1}{2} + \frac{y_1 + h}{2\sqrt{1 + h^2}}, \qquad C_2 = \frac{1}{2} - \frac{y_1 + h}{2\sqrt{1 + h^2}}.$$

Inserting this result, we get

$$y_n = \left(\frac{1}{2} + \frac{y_1 + h}{2\sqrt{1 + h^2}}\right)(-h + \sqrt{1 + h^2})^n + \left(\frac{1}{2} - \frac{y_1 + h}{2\sqrt{1 + h^2}}\right)(-h - \sqrt{1 + h^2})^n.$$

Observe that $|-h - \sqrt{1 + h^2}| > 1$, so that unless we choose y_1 so that $\frac{1}{2} - \frac{y_1+h}{2\sqrt{1+h^2}} = 0$, $\lim_{n\to\infty} = \pm\infty$. The other term $|-h + \sqrt{1 + h^2}| = |1 - h + O(h^2)| < 1$ and hence $(-h + \sqrt{1 + h^2})^n \to 0$ as $n \to \infty$.

The above is an example of numerical instability. The true solution $e^{-x} \to 0$ as $x = nh \to \infty$, while for fixed h the approximate solution $\to \infty$ as $n \to \infty$.

However, if we consider the convergence of the sequence $\{y_n^h\}$ as $h \to 0$, $n \to \infty$ and $x = nh$ and make the assumption that $\lim_{h \to 0} y_1^h = y_0 = 1$, then

$$\lim_{h \to 0} \left(\frac{1}{2} + \frac{y_1 + h}{2\sqrt{1 + h^2}} \right) = 1, \qquad \lim_{h \to 0} \left(\frac{1}{2} - \frac{y_1 + h}{2\sqrt{1 + h^2}} \right) = 0.$$

Furthermore, for $x = nh$,

$$\lim_{h \to 0, n \to \infty} (-h + \sqrt{1 + h^2})^n = \lim_{h \to 0} [(-h + \sqrt{1 + h^2})^{1/h}]^x.$$

Let $y = \lim_{h \to 0}[(-h + \sqrt{1 + h^2})^{1/h}]$. Then

$$\ln y = \lim_{h \to 0}[\ln(-h + \sqrt{1 + h^2})/h] = \lim_{h \to 0} \frac{\frac{2h}{2\sqrt{1+h^2}} - 1}{\sqrt{1 + h^2} - h} = -1.$$

Hence, $\ln y = -1$ so $y = e^{-1}$. Then

$$\lim_{h \to 0, n \to \infty} C_1(-h + \sqrt{1 + h^2})^n = e^{-x}.$$

Thus, the first part of the solution of the difference equation gives an approximation to the true solution of the differential equation. One can easily show that $|(-h - \sqrt{1 + h^2})^n| \le e^x$. Hence, the second term is converging to zero, so the approximate solution is converging to the true solution.

To summarize, one root of the characteristic polynomial gives a solution that approximates the true solution. A second root gives a parasitic solution which for fixed h eventually blows up to give a bad overall approximation. Since the method converges, for any x and ϵ, one can find a value of h such that $|y_n^h - e^{-x}| < \epsilon$. However, since the parasitic solution grows like e^x, this h would have to be impractically small for any reasonable size x.

13.2. **Strong and weak stability.** To formalize the stability problem discussed above, we now define several concepts of stability that seek to differentiate between methods which exhibit numerical instability and those that do not. These definitions usually refer to the difference equations obtained by applying the multistep method to the model problem:

$$y' = \lambda y, \qquad y(x_0) = y_0,$$

whose exact solution is $y(x) = y_0 e^{\lambda(x - x_0)}$. In this case, the resulting difference equation is:

$$y_{n+1} = \sum_{i=0}^{p} a_i y_{n-i} + h \sum_{i=-1}^{p} b_i \lambda y_{n-i},$$

which may be rewritten as:

$$y_{n+1}[1 - h\lambda b_{-1}] = \sum_{i=0}^{p} [a_i + h\lambda b_i] y_{n-i}.$$

This is a linear constant coefficient difference equation. The associated characteristic polynomial is:

$$z^{p+1}[1 - h\lambda b_{-1}] = \sum_{i=0}^{p} [a_i + h\lambda b_i] z^{p-i}.$$

When $h = 0$, this becomes just $\rho(z) = 0$. In general, it is $\rho(z) - h\lambda\sigma(z) = 0$.

We previously defined zero-stability as requiring that all roots of $\rho(z)$ have modulus ≤ 1 and all roots of modulus one to be simple. Since we want our method to be consistent (necessary for convergence), $z = 1$ is always a root of $\rho(z) = 0$.

Definition: The roots of $\rho(z)$ of modulus one are called essential roots. The root $z = 1$ is called the principal root. The roots of $\rho(z)$ of modulus < 1 are called nonessential roots.

Definition: A linear multistep method is strongly stable if all roots of $\rho(z)$ are ≤ 1 in magnitude and only one root has magnitude one. If more than one root has magnitude one, the method is called weakly or conditionally stable. Note, we still require only simple roots of magnitude one. Also, note these definitions refer to the case $h = 0$.

Returning to the example $y_{n+1} = y_{n-1} + 2hf_n$, we have $\rho(z) = z^2 - 1$, so the roots are $z = \pm 1$. Hence, this is a weakly stable method. For the specific problem $y' = -y$, $y(0) = 1$, the roots of the difference equation were

$$z_1 = -h + \sqrt{1 + h^2}, \qquad z_2 = -h - \sqrt{1 + h^2}.$$

The problem was that since $|z_2| > 1$, the corresponding parasitic solution blew up. The basic idea of strong stability is that since the roots of a polynomial are continuous functions of the coefficients, for $h\lambda$ near zero, the roots of $\rho(z) - h\lambda\sigma(z) = 0$ are near the roots of $\rho(z) = 0$. If the method is strongly stable, all extraneous roots have magnitude < 1, so for $|h\lambda|$ small enough, all roots of $\rho(z) - h\lambda\sigma(z) = 0$ will also have magnitude < 1. Hence the parasitic solution corresponding to this root will decay as $n \to \infty$, instead of blowing up to ruin the approximate solution. Other definitions of stability try to more precise in defining the values of $h\lambda$ for which the parasitic solutions remain bounded.

13.3. **Absolute and relative stability.** The following definitions of stability attempt to give a more precise characterization of the values of $h\lambda$ for which parasitic solutions die out.

Definition: A linearly multistep method is said to be *absolutely stable* for those values of $h\lambda$ for which all roots r_s of $\pi(r, h\lambda) = \rho(r) - h\lambda\sigma(r) = 0$ satisfy $|r_s| \leq 1$ (and if $|r_s| = 1$, then r_s is simple).

In other words, all solutions of the test problem

$$y_{n+1}[1 - b_{-1}h\lambda] = \sum_{i=0}^{p}(a_i + h\lambda b_i)y_{n-i}$$

remain bounded as $n \to \infty$. If the method is absolutely stable for all $h\lambda \in (\alpha, \beta)$, the interval (α, β) is called the interval of absolute stability.

Example: midpoint rule $y_{n+1} = y_{n-1} + 2h\lambda y_n$. The characteristic polynomial is $r^2 - 2h\lambda r - 1 = 0$, so $r = h\lambda \pm \sqrt{h^2\lambda^2 + 1}$. Clearly if $h\lambda < 0$, then $|h\lambda - \sqrt{h^2\lambda^2 + 1}| > 1$ and if $h\lambda > 0$, then $|h\lambda + \sqrt{h^2\lambda^2 + 1}| > 1$. Hence, this method is only absolutely stable for $h\lambda = 0$, so there is no interval of absolute stability.

The definition of absolutely stable determines an interval in which parasitic solutions do not grow. However, if the true solution is increasing, i.e., $\lambda > 0$, then it is not a problem if parasitic solutions grow, provided they do not grow faster than the true solution.

Defintion: A linear multistep method is said to be relatively stable for those values of $h\lambda$ for which all roots r_s of $\pi(r, h\lambda)$ satisfy $|r_s| \leq |r_0|$, and if $|r_s| = |r_0|$, then r_s is simple. Here r_0 is the principle root, i.e., the root with the property that $\lim_{h\to 0} r_0(h) = 1)$. If the method is relatively stable for all $h\lambda \in (\alpha, \beta)$, the interval (α, β) is called the interval of relative stability.

Example: midpoint rule $r_0 = h\lambda + \sqrt{h^2\lambda^2 + 1}$, $r_1 = h\lambda - \sqrt{h^2\lambda^2 + 1}$. For relative stability, we require $|h\lambda - \sqrt{h^2\lambda^2 + 1}| \leq |h\lambda + \sqrt{h^2\lambda^2 + 1}|$, i.e., $h\lambda \geq 0$. So the interval of relative stability is $[0, \infty)$.

Remark: There are various similar defintitions that make slight changes (e.g., using $<$ instead of $\leq$ and not requiring simple roots). Note that this definition does not apply to one-step methods since there is only one root of $\rho(r)$, but absolute stability definition does apply.

Example: Euler's method: $y_{n+1} = y_n + hf_n$. When $f(x, y) = \lambda y$, we get $y_{n+1} = y_n + h\lambda y_n$, so $r_0 = 1 + h\lambda$. For absolute stability, we need $-1 \leq 1 + h\lambda \leq 1$, i.e., $-2 \leq h\lambda \leq 0$. Hence, interval of absolute stability is $[-2, 0]$.

Example: Trapezoidal rule $y_{n+1} = y_n + (h/2)(f_{n+1} + f_n)$. When $f(x, y) = \lambda y$, we get $y_{n+1} = y_n + (h\lambda/2)(y_{n+1} + y_n)$. Hence, $(1 - h\lambda/2)y_{n+1} = (1 + h\lambda/2)y_n$. So the only root of the characteristic polynomial is $r_0 = (1 + h\lambda/2)/(1 - h\lambda/2)$. For absolute stability, we need $|r_0| \leq 1$, i.e., $h\lambda \leq 0$. This is the best one can obtain, since one can show that $h\lambda > 0$ cannot belong to the interval of absolute stability.