LINEAR MULTISTEP METHODS

Equations involving the derivative of a function of one variable occur in many branches of applied mathematics. Broadly speaking, any situation that concerns the rate of change of one variable with respect to another leads to a differential equation, and such situations are clearly very common.

Therefore, several theoretical methods had been introduced and suggested to solve such ordinary differential equations whenever the differential equation satisfies certain conditions for existence and uniqueness. But still there some problems which meets the existence and uniqueness theorem of ordinary differential equations that could not be solved analytically. Hence numerical methods are necessary to solve such equations.

Solution of differential equations occurs in two categories, the first one is the **L**inear **M**ultistep **M**ethods (LMM) and the second is the Runge-Kutta Methods.

1- BASIC CONCEPTS

I- Difference Equations:

Let $x_n = x_0 + nh$, n = 1, 2, …, k, where h is the step length. The difference equation of order k is an equation involving the unknown y_n together with its differences up to order k, i.e., difference equation has the form:

$$F(x_n, y_n, y_{n+1}, \ldots, y_{n+k}) = f(x_n) \quad \ldots\ldots(1)$$

Hence the order k of the difference equation is the difference between the highest and lowest index of y.

II- Classification of Finite Difference Equation:

Suppose that the finite difference equation is given by eq.(1), then:

1. The finite difference equation is linear in y_n if F appears to be linear in y.
2. The finite difference equation is said to be homogenous if $f(x_n) = 0$, otherwise it is non-homogenous.
3. The finite difference equation is said to be of constant coefficient if the coefficients of $y_n, y_{n+1}, \ldots, y_{n+k}$ are constants.

III- Solution of Linear Difference Equations with Constant Coefficients:

We shall occasionally need the general solution of the k-th order linear difference equation with constant coefficients:

$$a_k y_{n+k} + a_{k-1} y_{n+k-1} + \ldots + a_0 y_n = f_n \ldots \quad (2)$$

where $n = 1, 2, \ldots$, and a_j, $j = 0, 1, \ldots, k$, are constants independent of n, and $a_k \neq 0$, $a_0 \neq 0$.

A solution of such a difference equation will consist of a sequence $y_1, y_2, \ldots$, which will be indicated by $\{y_n\}$.

Let $\{\hat{y}_n\}$ be the general solution of the corresponding homogeneous difference equation:

$$a_k y_{n+k} + a_{k-1} y_{n+k-1} + \ldots + a_0 y_n = 0 \quad (3)$$

$n = 1, 2, \ldots$

If $\{\psi_n\}$ is any particular solution of eq.(2), then the general solution of eq.(2) is $\{y_n\}$; $y_n = \hat{y}_n + \psi_n$.

The solution of the difference equation can be evaluated easily by letting $y_n = \beta^n$, and considering the general solution of the difference equation depending on the roots of β whether it is a repeated or distinct real roots, or may be of complex roots, etc.

IV- Taylor's Theorem:

In general, numerical methods study the solution of the first order ordinary differential equations:

$$y'(x) = f(x, y(x)), \; y(x_0) = y_0$$

which satisfies the standard conditions for the existence and uniqueness theorem.

Now, expanding the solution y(x) using Taylor series expansion about x_n, one get:

$$y(x) = y(x_n) + (x - x_n)y'(x_n) + \ldots + \frac{(x - x_n)^m}{m!} y^{(m)}(x_n) + R_{m+1}$$

Now, put $x = x_{n+1}$

$$y(x_{n+1}) = y(x_n) + (x_{n+1} - x_n)y'(x_n) + \ldots + \frac{(x_{n+1} - x_n)^m}{m!} y^{(m)}(x_n) + R_{m+1}$$

$$= y(x_n) + hy'(x_n) + \ldots + \frac{h^m}{m!} y^{(m)}(x_n) + R_{m+1}$$

Truncate after m-terms

$$y_{n+1} = y_n + hy'_n + \ldots + \frac{h^m}{m!} y_n^{(m)}$$

which is called the Taylor method of order m, for all n = 0, 1, …; such that y_0 is given as initial condition.

Definition (1):

The ***local truncation error*** (LTE) of the m-th order Taylor method is:

$$LTE = \frac{h^{m+1}}{(m+1)!} y^{(m+1)}(\xi), \ \xi \in (x_n, x_{n+1}).$$

Definition (2):

The ***global truncation error*** (GTE) is defined to be:

$GTE = y(x_n) - y_n$

2- GENERAL LINEAR MULTISTEP METHODS

2-1 Formulation and Classification:

Consider the initial value problem for a single first-order differential equation:

$y'(x) = f(x, y(x)), y(x_0) = y_0, x \in [a,b] \subset \mathbb{R}$ .....(4)

We seek for a solution in the range $a \le x \le b$, and we assume that f satisfies the conditions of the existence and uniqueness theorem, which we shall indicate by y(x).

Consider the sequence of points $\{x_n\}$ defined by $x_n=a + nh$, n=0, 1, … The parameter h, which will always be regarded as constant, except where otherwise indicate as it will be seen in chapter two, is called the step length. An essential property of the majority of computational methods for the solution of eq.(4) is that of discritization; that is we seek for an approximate solution, not on the continuous interval $a \le x \le b$, but on the discrete point set $\{x_n \mid n = 0, 1, \ldots, (b - a)/h\}$.

Let y_n be an approximation to the theoretical solution at x_n that is, to $y(x_n)$ and let $f_n = f(x_n, y_n)$.

If a computational method for determining the sequence $\{y_n\}$ takes the form of a linear relationship between y_{n+j} and f_{n+j}, $\forall$ j = 0, 1, .., k, we call it a Linear Multistep Method of step number k, or a linear k-step method.

The general form of LMM may thus be written as:

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f_{n+j} \quad \text{.....(5)}$$

where α_j and β_j are constants to be determined. We assume that $\alpha_k \neq 0$ and that not both of α_0 and β_0 equals zero. Since eq.(5) can be multiplied on both sides by the same constant without altering the relationship, the coefficients α_j and β_j are arbitrary to the extent of a constant multiplier. We remove this arbitrariness by assuming throughout that $\alpha_k = 1$.

Thus the problem of determining the solution y(x) of the, in general, non-linear initial value problem we replace eq.(4) by that of finding the sequence $\{y_n\}$ which satisfies the difference eq.(5).

Note that, since f_n is in general non-linear function of y_n, then eq.(5) is a non-linear difference equation. Such equations are no easier to handle theoretically as in linear differential equations, but they have the particle advantage of permitting us to compute the sequence $\{y_n\}$ numerically.

In order to do this, we must first supply a set of starting values $y_0, y_1, \ldots, y_{k-1}$. (In the case of a one-step method, only one of such value y_0 is needed and we normally choose $y_0 = \eta$). As a classification to the L.M.M. we say that the L.M.M. is explicit if $\beta_k = 0$ and implicit if $\beta_k \neq 0$. For an explicit method eq.(5) yields the current value y_{n+k} directly in terms of previous y_{n+j}, f_{n+j}, j = 0, 1, …, k-1, which at this stage of the computation, have already been calculated. An implicit method, however, will call for the solution at each stage of the computation, of the equation:

$$y_{n+k} = h\beta_k f(x_{n+k}, y_{n+k}) + g \quad \text{.....(6)}$$

where g is a known function of the previously calculated values y_{n+j}, f_{n+j}, j = 0, 1, …, k−1. When the original differential eq.(4) is linear, then eq.(6) is also linear in y_{n+k}, and there exists a unique solution for y_{n+k} and there is no problem in solving such implicit problems, when f is non-linear, then there exists a unique solution for y_{n+k}, which can be approached arbitrarily closely by the iteration $y_{n+k}^{[s+1]} = h\beta_k f(x_{n+k}, y_{n+k}^{[s]}) + g$, $y_{n+k}^{[0]}$ arbitrary.

Thus implicit methods in general entail a substantially greater computational effort than do explicit methods; on the other hand, for a given step number, implicit methods can be made more accurate than explicit ones and, moreover, enjoy more favorable stability properties.

2-2 Order and Error Constant:

Next, we are in effect formalizing the Taylor series method, for the derivation of LMM's. With the LMM eq.(5), we associate the linear difference operator T defined by:

$$T[y(x_n), h] = \sum_{j=0}^{k} \alpha_j y(x_n + jh) - h\sum_{j=0}^{k} \beta_j f(x_n + jh, y(x_n + jh)) \quad \text{.....(7)}$$

where y(x) is arbitrary function, continuously differentiable on [a, b]. The reason for introducing this operator is that, by allowing it to operate on an arbitrary test function y(x), which we may assume to have as many higher derivatives as we require, we can formally define the order of accuracy of the operator and of the associated LMM, without invoking the solution of the initial value problem (4) which, as we have already observed, may posses only a first derivative.

Definition (3):

The local truncation error of the LMM. is defined by eq.(7).

After expanding the test function $y(x_n + jh)$ and it's derivative $y'(x_n + jh)$ as a Taylor series about x_n, and collecting termed in eq.(7), gives:

$$T[y(x_n); h] = C_0 y(x_n) + C_1 h y'(x_n) + C_2 h^2 y''(x_n) + \ldots + C_p h^p y^{(p)}(x_n) + \ldots$$

where $C_0, C_1, \ldots, C_p, \ldots$ are constants.

A simple calculation yields the following formula for the C_i's, i = 0, 1, …, p, …, in terms of the coefficients α_j and β_j:

$$\left.\begin{aligned} C_0 &= \alpha_0 + \alpha_1 + \ldots + \alpha_k \\ C_1 &= \alpha_1 + 2\alpha_2 + \ldots + k\alpha_k - (\beta_0 + \beta_1 + \ldots + \beta_k) \\ &\vdots \\ C_p &= \frac{1}{p!}\left(\alpha_1 + 2^p\alpha_2 + \ldots + k^p\alpha_k\right) - \frac{1}{(p-1)!}\left(\beta_1 + 2^{p-1}\beta_2 + \ldots + k^{p-1}\beta_k\right) \end{aligned}\right\} \quad \text{.....(8)}$$

where p = 2, 3, … This formula can be used to derive a certain LMM of a given structure and maximal order.

Definition (4):

The method of LMM is said to be of order p if $C_0 = C_1 = \ldots = C_p = 0$, but $C_{p+1} \neq 0$, and the local truncation error is given by:

$$T_n = C_{p+1} h^{p+1} y^{(p+1)}(\eta), \; x_n < \eta < x_{n+1}$$

Example (1):

Construct an implicit LMM (2-steps method) of maximal order, containing one free parameter, and find its order.

$$\alpha_2 y_{n+2} + \alpha_1 y_{n+1} + \alpha_0 y_n = h\{\beta_2 f_{n+2} + \beta_1 f_{n+1} + \beta_0 f_n\}$$

with $k = 2$, $\alpha_2 = 1$, hence:

$$y_{n+2} + \alpha_1 y_{n+1} + \alpha_0 y_n = h\{\beta_2 f_{n+2} + \beta_1 f_{n+1} + \beta_0 f_n\}$$

if we let ($\alpha_0 = a$) to be a free parameter, then we have to find α_1, β_2, β_1, β_0.

We can set $C_0 = C_1 = C_2 = C_3 = 0$, then:

$$1 + \alpha_1 + a = 0$$

$$\alpha_1 + 2 - (\beta_0 + \beta_1 + \beta_2) = 0$$

$$\frac{1}{2!}(\alpha_1 + 4) - (\beta_1 + 2\beta_2) = 0$$

$$\frac{1}{3!}(\alpha_1 + 8) - \frac{1}{2!}(\beta_1 + 2\beta_2) = 0$$

Solving these equations yields:

$$\alpha_1 = -1 - a,\ \beta_0 = -\frac{1}{12}(1 - 5a),\ \beta_1 = \frac{2}{3}(1 - a) \text{ and } \beta_2 = \frac{1}{12}(5 + a)$$

and the method is:

$$y_{n+2} - (1 + a)y_{n+1} + ay_n = \frac{h}{12}\{(5 + a)f_{n+2} + 8(1 - a)f_{n+1} - (1 - 5a)f_n\}$$

and:

$$C_4 = \frac{1}{4!}(\alpha_1 + 16) - \frac{1}{3!}(\beta_1 + 8\beta_2) = -\frac{1}{4!}(1 + a)$$

$$C_5 = \frac{1}{5!}(\alpha_1 + 32) - \frac{1}{4!}(\beta_1 + 16\beta_2) = -\frac{1}{3!5!}(17 + 13a)$$

Hence:

if $a \neq -1$, then $C_4 \neq 0$ and the method is of order 3

if $a = -1$, then $C_4 = 0$ and $C_5 \neq 0$ and the method is of order 4

i.e.

$$y_{n+2} - y_n = \frac{h}{12}(4f_{n+2} + 16f_{n+1} + 4f_n) \text{ (Simpson rule)}$$

if a = 0, we have:

$$y_{n+2} - y_{n+1} = \frac{h}{12}(5f_{n+2} + 8f_{n+1} - f_n) \text{ (Adam Moulton rule of 2-steps)}$$

if a = –5, then the method is explicit.

Definition (5):

The first characteristic polynomial of LMM (1.5) is given by:

$$\rho(r) = \sum_{j=0}^{k} \alpha_j r^j = \alpha_k r^k + \alpha_{k-1} r^{k-1} + \ldots + \alpha_0.$$

While the second characteristic polynomial is given by:

$$\sigma(r) = \sum_{j=0}^{k} \beta_j r^j = \beta_k r^k + \beta_{k-1} r^{k-1} + \ldots + \beta_0.$$

The necessary and sufficient conditions for LMM to have an order p can be studied by using the first and second characteristic polynomials.

Also, it is important to notice that if $\sigma(r)$ is given, then we can find a unique polynomial $\rho(r)$ of degree k such that the method ha an order $p \geq k$.

2-3 Classification of Linear Multistep Methods:

We can consider the LMM according to the roots of the first characteristic polynomial $\rho(r)$ and whether it is explicit or implicit.

(i) If the roots of $\rho(r)$ equal to 1 and 0, then the methods are called of Adam's types and if the LMM is explicit, then it is called of Adam Bashforth, while if it is implicit then it is called of Adam Moulton type, i.e., in Adam's methods, we have the following:

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

$$= r^{k-1}(r - 1) = 0$$

(ii) If the roots of $\rho(r)$ equals to –1, 0 and 1, then the methods are called of Nystrom type if it is explicit and if the method is implicit, then it is called of Milne-Simpson type, i.e., we have:

$$\rho(r) = r^k - r^{k-2}$$

$= r^{k-2}(r^2 - 1)$

$= r^{k-2}(r - 1)(r + 1)$

2-4 Derivation of Some Linear Multistep Methods:

Any specific LMM may be derived in a number of different ways, most important of them will be given next in details:

I-Derivation Through Taylor Expansion:

This method can be considered as the most simplest method for derivation of some LMM. Consider the Taylor series expansion for $y(x_n + h)$ about x_n:

$$y(x_n + h) = y(x_n) + hy'(x_n) + \frac{h^2}{2!}y''(x_n) + \ldots + \frac{h^p}{p!}y^{(p)}(x_n) + \ldots \quad (9)$$

If we truncate this expansion after two terms and substitute for $y'(x_n)$ from the differential eq.(1.4), we have:

$$y(x_n + h) = y(x_n) + hf(x_n, y(x_n)) \quad (10)$$

a relation which is in truncation error by:

$$\frac{h^2}{2!}y''(x_n) + \frac{h^3}{3!}y'''(x_n) + \ldots \quad (11)$$

equation (10) express an approximate relation between exact values of the solution of eq.(4). We can also interpret it as an exact relation between approximate values of the solution of eq.(4). If we replace $y(x_n)$, $y(x_n + h)$ by y_n and y_{n+1} respectively, yielding:

$$y_{n+1} = y_n + hf_n \quad (12)$$

which is an explicit linear one step method. It is in fact Euler's method, that can be considered as the simplest of all LMM's. The error associated with eq.(12) is expression in eq.(11) (multiplying by +1 or –1 according to the sense of definition of error) and is called the local truncation error or local discretization error, for this method the local truncation error is $O(h^2)$ (i.e., of order h^2). Consider now the Taylor series expansion for $y(x_n + h)$ and $y(x_n – h)$ about x_n:

$$y(x_n+h) = y(x_n) + hy'(x_n) + \frac{h^2}{2!}y''(x_n) + \frac{h^3}{3!}y'''(x_n) + \ldots$$

$$y(x_n–h) = y(x_n) – hy'(x_n) + \frac{h^2}{2!}y''(x_n) – \frac{h^3}{3!}y'''(x_n) + \ldots$$

subtracting, we get:

$$y(x_n + h) - y(x_n - h) = 2hy'(x_n) + \frac{h^3}{3!}y'''(x_n) + \ldots$$

Arguing as previously yields the associated linear two-steps method:

$$y_{n+1} - y_{n-1} = 2hf_n$$

This can be brought into the standard form eq.(5) by replacing n by n + 1 to give:

$$y_{n+2} - y_n = 2hf_{n+1} \quad \text{.....(13)}$$

Which is called also, the Mid-Point method and its local truncation error is given by:

$$\pm\frac{1}{3}h^3y'''(x_n) + \ldots$$

Similar techniques can be used to derive any LMM of given specification.

Thus if we wish to find the most accurate one-step implicit method:

$$y_{n+1} + \alpha_0 y_n = h(\beta_1 f_{n+1} + \beta_0 f_n)$$

We write down the associated approximate relation ship

$$y(x_n + h) + \alpha_0 y(x_n) = h(\beta_1 y'(x_n + h) + \beta_0 y'(x_n)) \quad \text{.....(14)}$$

and choose α_0, β_0 and β_1 so as to make the approximation as accurate as possible. The following expansions are used:

$$y(x_n + h) = y(x_n) + hy'(x_n) + \frac{h^2}{2!}y''(x_n) + \ldots$$

$$y'(x_n + h) = y'(x_n) + hy''(x_n) + \frac{h^2}{2!}y'''(x_n) + \ldots$$

Substituting in eq.(14) and collecting the terms on the left-hand side gives:

$$C_0y(x_n) + C_1hy'(x_n) + C_2h^2y''(x_n) + C_3h^3y'''(x_n) + \ldots = 0$$

where $C_0 = 1 + \alpha_0$, $C_1 = 1 - \beta_1 - \beta_0$, $C_2 = \frac{1}{2} - \beta_1$, $C_3 = \frac{1}{6} - \frac{1}{2}\beta_1$, …

Thus, in order to make the approximation in eq.(14) as accurate as possible we solve the given system $C_0 = C_1 = C_2 = 0$, to get $\alpha_0 = -1$, $\beta_1 = \beta_0 = \frac{1}{2}$. The implicit one-steps method is now:

$$y_{n+1} - y_n = \frac{h}{2}(f_{n+1} + f_n) \quad \text{.....(15)}$$

which is called also the trapeziodal method and it's local truncation error is given by:

$$\pm\frac{1}{12}h^3y'''(x_n) + \dots$$

the most difficulties of this method of derivation is that it leaves some unanswered questions. If we take Taylor series expansions about some new point say about $x_n + h$ in the last example do we get the same values for coefficients α_j and β_j ? Do we get same values for the coefficients in the infinite series representing the local truncation error? How can we remove the ambiguity of sign in the truncation error?

II- Derivation through numerical integration:

Let $x = x_n + th$ and consider the Newton-Gregory forward interpolation formula:

$$P(x) = P(x_n + th) = f_n + t\Delta f_n + \frac{t(t-1)}{2!}\Delta^2 f_n + \dots$$

If our aim is to derive, say a linear two-step method then the only available data for the approximate evaluation of the integral will be the values f_n, f_{n+1}, f_{n+2} where p(x) be the unique polynomial of degree two passing through the three points (x_n, f_n), (x_{n+1}, f_{n+1}), (x_{n+2}, f_{n+2}) for this objective, consider the initial value problem.

$$y' = f(x, y),\ y(x_0) = y_0$$

We now make the approximation:

$$\int_{x_n}^{x_{n+2}} y'(x)\,dx = \int_{x_n}^{x_{n+2}} f(x,y)\,dx$$

Hence:

$$y(x_{n+2}) - y(x_n) = \int_{x_n}^{x_{n+2}} y'(x)\,dx \quad \text{.....(16)}$$

$$= \int_{x_n}^{x_{n+2}} [f_n + t\Delta f_n + \frac{t(t-1)}{2!}\Delta^2 f_n]\,dx$$

if we let $x = x_n + th$, the $dx = hdt$. Hence:

$$y_{n+2} - y_n = h\int_0^2 [f_n + t\Delta f_n + \frac{t(t-1)}{2!}\Delta^2 f_n]\,dt$$

$$= h[2f_n + 2\Delta f_n + \frac{1}{3}\Delta^2 f_n]$$

Expanding Δf_n and $\Delta^2 f_n$ in terms of f_n, f_{n+1}, f_{n+2}, where Δ is the forward difference operator and substituting in eq.(16) gives:

$$y_{n+2} - y_n = \frac{h}{3}(f_{n+2} + 4f_{n+1} + f_n) \quad \text{.....(17)}$$

which is Simpson's method, the most accurate implicit linear two-steps method.

Now, if we replace eq.(16) by the identity:

$$y(x_{n+2}) - y(x_{n+1}) = \int_{x_{n+1}}^{x_{n+2}} y'(x)\,dx$$

and replaces $y'(x)$ by $P(x)$, defined as above, we derive the method:

$$y_{n+2} - y_{n+1} = \frac{h}{12}(5f_{n+2} + 8f_{n+1} - f_n) \quad \text{.....(18)}$$

which is the two-steps Adams-Moulton method.

The importance of the present technique is that it establishes a link between the concepts of polynomial interpolation and LMM's.

Other methods for derivations are also presented in literatures, such as the derivation through interpolation.

2-5 Consistency, Convergence and Zero stability of LinearMultistep Methods:

I- Convergence:

A basic property which we shall demand of an acceptable LMM is that the solution $\{y_n\}$ generated by the method converges, in some sense to the theoretical solution $y(x)$ as the step length h tends to zero. In converting this intuitive concept into a precise definition, the following points must be kept in mind:

i. It is inappropriate to consider n as remaining fixed while $h \longrightarrow 0$. For example, consider a fixed point $x = x^*$ and let the initial choice of step length h_0 be such that $x^* = a + 2h_0$. In the special case when the step length is successively halved, the situation illustrated by Fig.(1) is holds. If we use the notation $y_n(h)$ to denote the value y_n given by the LMM when the step length is h, then we are interested not in the convergence of the sequence $y_2(h_0)$, $y_2(h_0/2)$, $y_2(h_0/4)$, …, but in the convergence of the sequence $y_2(h_0)$, $y_4(h_0/2)$, $y_8(h_0/4)$, …, to $y(x^*)$ (in general, of course, we consider the case when h tends continuously to zero). We are thus led to the idea of fixed station convergence by which

we mean convergence in the limit as $h \longrightarrow 0$, $n \longrightarrow \infty$, $nh = x - a$, and a remaining fixed. Such limit will be written as $\lim_{\substack{h\to 0 \\ nh=x-a}}$.

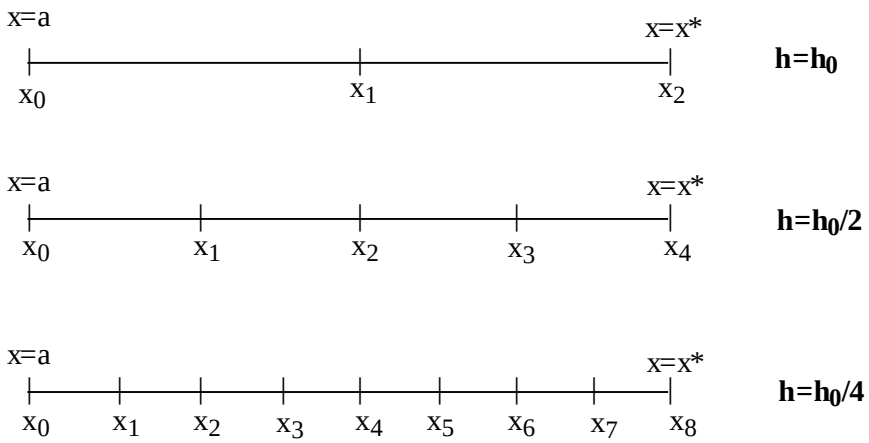


Fig.(1).

ii. The definition must take account of the additional starting values $y_1, y_2, \ldots, y_{k-1}$, which must be supplied when $k \geq 2$.

iii. If the term convergent is to be applied to the LMM, then the convergence property must hold for all initial value problems (4) subject to the hypothesis of the existence and uniqueness theorem.

The following definitions illustrated the sense of consistency and zero stability.

Definition (6):

The LMM is said to be consistent with the IVP $y' = f(x, y)$, $y(x_0) = y_0$ if it has an order $p \geq 1$, i.e., consistent methods implies $C_0=C_1=0$, but $C_2\neq 0$ ($\rho(1) = 0$ and $\rho'(1) = \sigma(1)$), or,

$$\sum_{j=0}^{k} \alpha_j = 0 \quad \text{and} \quad \sum_{j=0}^{k} j\alpha_j = \sum_{j=0}^{k} \beta_j$$

Definition (7):

The LMM is said to be zero-stable (0-stable) if all the roots (zeros) r_j's, $j = 1, \ldots, k$; of $\rho(r) = 0$ satisfy $|r_j| \leq 1$ and if r_j is a multiple zero of $\rho(r)$ then $|r_j| < 1$.

Theorem (1):

Assume the consistency conditions, then the LMM is convergent if and only if the zero-stability condition is satisfied.

Proof:

$\Rightarrow$ Suppose that the method is convergent. To prove it is zero-stable, i.e., $|r_j| \le 1$

For simplicity, consider the problem $y' \equiv 0$, $y(0) = 0$, which has the exact solution $y(x) = 0$.

The LMM takes the form:

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = 0, \text{ where } \alpha_k = 1$$

$$y_k = \sum_{j=0}^{k-1} \alpha_j y_{n+j} \quad \text{.....(19)}$$

Then the solution y_k will depends on $y_0, y_1, \ldots, y_{k-1}$ which is chosen to satisfy:

$$\eta(h) = \max_{0 \le n \le k-1} |y_n| \longrightarrow 0, \text{ as } h \longrightarrow 0 \quad \text{.....(20)}$$

For contrary, suppose that at least one of the roots of the characteristic equation $\rho(r)$ is greater than one, i.e., $|r_j| > 1$.

Then the solution of the finite difference eq.(19) is given by $y_n = h(r_j)^n$

Then condition (20) is satisfied for this solution, i.e.,

$$\eta(h) = \max_{0 \le n \le k-1} |y_n|$$

$$= \max_{0 \le n \le k-1} |h(r_j)^n|$$

$$= h \max_{0 \le n \le k-1} |r_j|^n$$

$$= h\,|r_j|^{k-1} \longrightarrow 0, \text{ as } h \longrightarrow 0$$

But the sequence of numerical solution $\{y_n\}$ does not converge, since:

$$\max_{0 \le n \le k-1} |y(x_n) - y_n| = h|r_j|^{N(h)}$$

Consider $h = (b - a)/N(h) = b / N(h)$. Then by using L'Hospital's rule can be used to show that:

$$\lim_{\substack{N \to \infty \\ h \to 0}} \frac{b}{N} |r_j|^N = \lim_{N \to \infty} bN |r_j|^{N-1} = \infty$$

which says that the numerical solution converge to ∞ and does not converges to the exact solution $Y(x) = 0$, which is a contradiction.

Hence the method is zero-stable.

$\Leftarrow$ Suppose that the LMM is zero-stable.

To prove it is converge

For simplicity, consider the equation $y' = \lambda y$, $y(0) = 1$.

To show that the term $C_0[r_0(\lambda h)]^n$ in its general solution:

$$y_n = \sum_{j=0}^{k} C_j[r_j(\lambda h)]^n$$

will be converge to the exact solution $y(x) = e^{\lambda x}$ on $[0, b]$. The remaining terms $C_j[r_j(\lambda h)]^n$, $j = 1, 2, \ldots, k-1$, they can be show to converge to zero as $h \longrightarrow 0$.

Expanding $r_0(\lambda h)$ using Taylor's theorem:

$$r_0(\lambda h) = r_0(0) + h\lambda\, r_0'(0) + O(h^2)$$

Since:

$$r_j'(u) = \frac{\delta(r_j(u))}{\rho'(r_j(u)) - u\sigma'(r_j(u))}$$

Hence:

$$r_0'(0) = \frac{\delta(r_0(0))}{\rho'(r_0(0)) - 0 \times \sigma'(r_0(0))}$$

$$= \frac{\delta(r_0(0))}{\rho'(r_0(0))}$$

Since $r_0(0) = 1$. Then:

$$r_0'(0) = \frac{\delta(1)}{\rho'(1)} = 1$$

Hence:

$$r_0(\lambda h) = 1 + \lambda h + O(h^2)$$

$$= e^{\lambda h} - O(h^2) + O(h^2)$$

$$\simeq e^{\lambda h}$$

Therefore:

$$[r_0(\lambda h)]^n \simeq [e^{\lambda h}]^n = e^{\lambda n h} = e^{\lambda x_n}$$

Hence:

$$\underset{0 \le x_n \le b}{\text{Max}} \left| [r_0(\lambda h)]^n - e^{\lambda x_n} \right| \longrightarrow 0 \ \text{ as } \ h \longrightarrow 0.$$

Now, we have to show that $C_0 \longrightarrow 1$ as $h \longrightarrow 0$. The coefficients of $C_0(h)$, $C_1(h)$, …, $C_{k-1}(h)$ satisfy the linear system:

$$C_0 + C_1 + \ldots + C_{k-1} = y_0$$

$$C_0[(r_0(\lambda h)] + C_1[r_1(\lambda h)] + \ldots + C_{k-1}[r_{k-1}(\lambda h)] = y_1$$

$$\vdots$$

$$C_0[r_0(\lambda h)]^{k-1} + C_1[r_1(\lambda h)]^{k-1} + \ldots + C_{k-1}[r_{k-1}(\lambda h)]^{k-1} = y_{k-1}$$

The initial values $y_0, y_1, \ldots, y_{k-1}$ depends on h and are assumed to satisfy the following:

$$\eta(h) = \max_{0 \le n \le k-1} \left| e^{\lambda x_n} - y_n \right| \longrightarrow 0 \text{ as } h \longrightarrow 0$$

But this implies that $\lim_{h \to 0} y_n = 1$

The coefficient C_0 can be obtained by using Cramer's rule to solve the above linear system.

$$C_0 = \frac{\begin{vmatrix} y_0 & 1 & \cdots & 1 \\ y_1 & r_1 & \cdots & r_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{k-1} & r_1^{k-1} & \cdots & r_{k-1}^{k-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ r_0 & r_1 & \cdots & r_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_0^{k-1} & r_1^{k-1} & \cdots & r_{k-1}^{k-1} \end{vmatrix}}$$

and as $h \longrightarrow 0$, we have:

$$C_0 = \frac{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ r_0 & r_1 & \cdots & r_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_0^{k-1} & r_1^{k-1} & \cdots & r_{k-1}^{k-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ r_0 & r_1 & \cdots & r_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_0^{k-1} & r_1^{k-1} & \cdots & r_{k-1}^{k-1} \end{vmatrix}} = 1$$

Hence $C_0 \longrightarrow 1$ as $h \longrightarrow 0$ and therefore the sequence $\{y_n\}$ converges to $y(x) = e^{\lambda x}$.

Example (2):

Suppose we have a two steps explicit method:

$$y_{n+2} - y_{n+1} = \frac{h}{2}(3f_{n+1} - f_n)$$

i.e., k =2, α_2 =1, $\alpha_1 = -1$, $\alpha_0 = 0$, $\beta_2 = 0$, $\beta_1 = 3/2$ and $\beta_0 = -1/2$

To test for convergence, first we must find the roots by using the first characteristic polynomial of LMM $\rho(r)$, which is:

$$\rho(r) = \sum_{j=0}^{2} \alpha_j r^j$$

$$= \alpha_0 + \alpha_1 r + \alpha_2 r^2$$

$$= r^2 - r$$

which have the roots r_1 =0 and $r_2 = 1$.

Then by using the definition of the zero-stability, we have:

$|r_1| = 0 \le 1$ and $|r_2| = 1 \le 1$

Hence the method is zero stable.

Second, we must apply the definition for consistency, hence:

$$C_0 = \sum_{j=0}^{2} \alpha_j = \alpha_0 + \alpha_1 + \alpha_2$$

$$= -1 + 1 = 0$$

$$C_1 = \sum_{j=0}^{2} j\alpha_j - \sum_{j=0}^{2} \beta_j = \alpha_1 + 2\alpha_2 - \beta_0 - \beta_1 - \beta_2$$

$$= -1 + 2 + \frac{1}{2} - \frac{3}{2} = 0$$

Therefore, the method is consistent, and as a result the method convergent

GENERAL RUNGE-KUTTA METHODS

The idea of extending the Euler method by allowing for a multiplicity of evolutions of the function f within each step was originally proposed by Runge (1895). Further contributions were made by Heun (1900) and by Kutta (1901). The latter completely characterized the set of Runge-Kutta method of order 4 and proposed the first methods of order 5. Special methods for second-order differential equations were proposed by Nyström (1925) who also contributed to the development of methods for first-order equations.

Since the advent of digital computers, fresh interest had been focused on Runge-Kutta methods, and a large number of research workers have contributed to recent extensions to the theory and the development of particular methods. Although, early studies were devoted entirely to explicit Runge-Kutta methods, interest has now extended to implicit methods, which are now recognized as appropriate for stiff differential equations.

1- FORMULATION OF RUNGE-KUTTA METHODS

The general form of an R-stages Runge-Kutta methods is given by:

$$y_{n+1} = y_n + h\sum_{r=1}^{R} c_r k_r$$

where:

$$k_r = f\left(x_n + ha_r, y_n + h\sum_{s=1}^{R} b_{rs}k_s \right)$$

and

$$a_r = \sum_{s=1}^{R} b_{rs}$$

where c_r, a_r and b_{rs}, for all r, s = 1, 2, …, K; are constants to be determined.

For convenience, we design the process by an array of constants, as follows:

b_{11}	b_{12}	…	b_{1k}	a_1
b_{21}	b_{22}	…	b_{2k}	a_2
⋮	⋮	⋱	⋮	⋮
b_{r1}	b_{r2}	…	b_{rk}	a_r
c_1	c_2	…	C_k	

and it is easy to classify Runge-Kutta methods, as follows:

- If $b_{ij} = 0$, $\forall$ $i < j$, then the method is called semi-explicit.
- If $b_{ij} = 0$, $\forall$ $i \leq j$, then the method is called explicit.
- Otherwise it is called implicit.

2- DERIVATION OF SOME RUNGE-KUTTA METHODS

We shall consider the derivation of some Runge-Kutta methods namely, 2-stages explicit, the 2-stages semi-explicit and the 2-stages implicit Runge-Kutta methods for the purpose of studying the accuracy of the methods.

2-1 Derivation of 2-Stages Explicit Runge-Kutta Method:

In order to derive two stages Runge-Kutta method; consider first the general form of two stages Runge-Kutta methods, which is given by:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

where:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f(x_n + a_2h, y_n + hb_{21}k_1)$$

and

$$a_2 = b_{21}$$

Hence, in this problem we have three unknown constants c_1, c_2, and a_2 which must be determined.

Now, recall the Taylor series expansion for a function $g(x + h, y + k)$ of two variables about (x, y), we have:

$$g(x+h,y+k) = g(x,y) + hg_x(x,y) + kg_y(x,y) + \frac{h^2}{2!}g_{xx}(x,y) + hkg_{xy}(x,y) + \frac{k^2}{2!}g_{yy}(x,y) + \ldots$$

So expanding k_2 using Taylor series about (x_n, y_n), we have:

$$k_2 = f(x_n + a_2h, y_n + hb_{21}k_1)$$

$$= f(x_n, y_n) + a_2hf_x + hb_{21}k_1f_y + \frac{a_2^2h^2}{2!}f_{xx} + a_2b_{21}h^2k_2fxy + \frac{h^2b_{21}^2k_1^2}{2!}f_{yy} + \ldots$$

$$= f + a_2h(f_x + ff_y) + \frac{a_2^2h^2}{2!}(f_{xx} + 2ff_{xy} + f^2f_{yy}) + \ldots$$

So, the two stages Runge-Kutta method takes the form:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

$$= y_n + h[c_1f + c_2(f + a_2h(f_x + ff_y) + \frac{a_2^2h^2}{2!}(f_{xx} + 2ff_{xy} + f^2f_{yy}) + \ldots)]$$

$$= y_n + h[c_1f + c_2f + c_2a_2h(f_x + ff_y) + \frac{c_2a_2^2h^2}{2!}(f_{xx} + 2ff_{xy} + f^2f_{yy}) + \ldots]$$

$$= y_n + hf(c_1 + c_2) + h^2c_2a_2(f_x + ff_y) + \frac{c_2a_2^2h^3}{2!}(f_{xx} + 2ff_{xy} + f^2f_{yy}) + O(h^4) \ldots\ldots\ldots\ldots (1)$$

Since Taylor method takes the form:

$$y_{n+1} = y_n + h\,y_n' + \frac{h^2}{2!}y_n'' + \frac{h^3}{3!}y_n''' + O(h^4) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

and since $y' = f(x, y)$, then $y'' = f_x + ff_y$, and

$$y''' = f_{xx} + ff_{xy} + f_y(f_x + ff_y) + f(f_{xy} + ff_{yy})$$

$$= (f_{xx} + 2ff_{xy} + f^2f_{yy}) + (f_xf_y + ff_y^2)$$

Therefore, eq.(2) becomes:

$$y_{n+1} = y_n + hf + \frac{h^2}{2!}(f_x + ff_y) + \frac{h^3}{3!}[(f_{xx} + 2ff_{xy} + f^2f_{yy}) + (f_xf_y + ff_y^2)] + O(h^4) \ldots\ldots\ldots (3)$$

Now, in order to get an agreement between Runge-Kutta method and Taylor's method (i.e., between eq.(1) and eq.(3)), we must have:

$$c_1 + c_2 = 1 \text{ and } c_2a_2 = \frac{1}{2}$$

with local truncation error of order h^3. Also, the order for Runge-Kutta method with R = 2 equals to p = 2.

The above two equations have an infinite number of solutions, e.g., we can take also $c_1 = c_2 = \frac{1}{2}$ and $a_2 = 1$.

2-2 Derivation of 2-Stages Semi-Explicit Runge-Kutta Method:

Since R = 2, then Runge-Kutta formula takes the form:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

where:

$$k_1 = f(x_n + a_1h, y_n + ha_1k_1)$$

$$k_2 = f(x_n + a_2h, y_n + hb_{21}k_1 + hb_{22}k_2)$$

and $a_1 = b_{11}$, $a_2 = b_{21} + b_{22}$.

To find these constants, consider the following power series:

$$k_1 = A_1 + hB_1 + h^2C_1 + \ldots \quad \text{.....(4)}$$

$$k_2 = A_2 + hB_2 + h^2C_2 + \ldots \quad \text{.....(5)}$$

Since:

$$k_1 = f + h[a_1f_x + a_1(A_1 + hB_1 + h^2C_1 + \ldots)f_y + \frac{h^2}{2}[a_1^2f_{xx} + 2a_1(A_1 + hB_1 + h^2C_1)f_{xy} +$$

$$a_1^2(A_1 + hB_1 + h_2C_1 + \ldots)^2f_{yy}] + \frac{h^3}{6}[a_1^3f_{xxx} + 3a_1^3(A_1 + hB_1 + h^2C_1 + \ldots)f_{xxy} +$$

$$3a_1^3(A_1 + hB_1 + h^2C_1 + \ldots)^2f_{xyy} + a_1^3(A_1 + hB_1 + h^2C_1 + \ldots)^3f_{yyy} + O(h^4) \quad \text{.....(6)}$$

Comparing eqs.(4) and (6), we have:

$$A_1 = f$$

$$B_1 = a_1(f_x + f_yA_1) = a_1(f_x + ff_y) = a_1F$$

$$C_1 = \frac{a_1^2}{2}(f_{xx} + 2A_1f_{xy} + A_1^2f_{yy}) + a_1B_1fy$$

$$= \frac{a_1^2}{2}(f_{xx} + 2ff_{xy} + f^2f_{yy}) + a_1^2(f_x + ff_y)f_y$$

$$= \frac{a_1^2}{2}G + a_1^2Ff_y$$

where:

$G = f_{xx} + 2ff_{xy} + f^2 f_{yy}$, and

$$D_1 = a_1 C_1 f_y + a_1^2 B_1 f_{xy} + a_1^2 A_1 B_1 f_{yy} + \frac{a_1^3}{6} f_{xxx} + \frac{a_1^3}{2} A_1 f_{xxy} + \frac{a_1^3}{2} A_1 f_{xxy} + \frac{a_1^3}{6} A_1 f_{yyy}$$

$$= a_1^3 F f_y^2 + a_1^3 F(f_{xy} + f f_{yy}) + a_1^3 G f_y + \frac{a_1^3}{6} H$$

where:

$H = f_{xxx} + 3ff_{xxy} + 3f^2 f_{xyy} + f^3 f_{yyy}$

and similarly:

$$k_2 = f + h[a_2 f_x + (b_{21}(A_1 + hB_1 + h^2 C_1) + b_{22}(A_2 + hB_2 + h^2 C_2))f_y] + \frac{h^2}{2}[a_2^2 f_{xx} + 2a_2(b_{21}(A_1 + hB_1) + b_{22}(A_2 + hB_2))f_{xy} + (b_{21}(A_1 + hB_1) + b_{22}(A_2 + hB_2))^2 f_{yy}] + \frac{h^3}{6}[a_2^3 f_{xxx} + 3a_2^2(b_{21}A_1 + b_{22}A_2)f_{xxy} + 3a_2(b_{21}A_1 + b_{22}A_2)^2 f_{xyy} + (b_{21}A_1 + b_{22}A_2)^3 f_{yyy}] + O(h^4) \quad \text{.....(7)}$$

Comparing eqs.(5) and (7), we have:

$A_2 = f$

$B_2 = a_2 f_x + (b_{21}A_1 + b_{22}A_2)f_y$

$$C_2 = (b_{21}B_1 + b_{22}B_2)f_y + \frac{a_2^2}{2} f_{xx} + a_2(b_{21}A_1 + b_{22}A_2)f_{xy} + \frac{1}{2}(b_{21}A_1 + b_{22}A_2)^2 f_{yy}, \text{ and}$$

$$D_2 = (b_{21}C_1 + b_{22}C_2)f_y + (a_2(b_{21}B_1 + b_{22}B_2 f_{xy} + (b_{21}A_1 + b_{22}A_2)(b_{21}B_1 + b_{22}B_2)f_{yy} + \frac{a_1^3}{6} f_{xxx} + \frac{a_1^2}{6}(b_{21}A_1 + b_{22}A_2)f_{xxy} + \frac{a_2}{2}(b_{21}A_1 + b_{22}A_2)^2 f_{xyy} + \frac{1}{6}(b_{21}A_1 + b_{22}A_2)^2 f_{yyy}$$

Since $a_2 = b_{21} + b_{22}$, therefore, using a_2 to simplify and solve the last fore equations, we obtain the following solution:

$A_2 = f$

$B_2 = a_2 f_x + (b_{21}f + b_{22}f)f_y$

$= a_2 f_x + (b_{21} + b_{22})ff_y$

$= a_2 f_x + a_2 f f_y$

$= a_2(f_x + f f_y) = a_2 F$

$$C_2 = (b_{21}a_1 + b_{22}a_2)Ff_y + \frac{a_2}{2}G$$

where, $G = f_{xx} + 2ff_{xy} + f^2f_{yy}$, an similarly:

$$D_2 = [b_{21}(b_{11}a_1 + b_{12}a_2) + b_{22}(b_{21}a_1 + b_{22}a_2)]F f_y^2 + a_2(b_{21}a_1 + b_{22}a_2)F(f_{xy} + ff_{yy}) + (b_{21}a_1^2 + b_{22}a_2^2)Gf_y + \frac{a_1^3}{6}H$$

where $H = f_{xxx} + 3f_{xxy} + 3f^2f_{xyy} + f^3f_{yyy}$

Since $\phi(x, y, h) = \sum_{r=1}^{2} c_r k_r$. Hence:

$$\phi = c_1k_1 + c_2k_2$$

$$= c_1(A_1 + hB_1 + h^2C_1 + h^3D_1) + c_2(A_2 + hB_2 + h^2C_2 + h^3D_2) + O(h^4)$$

$$= c_1A_1 + c_2A_2 + h(c_1B_1 + c_2B_2) + h^2(c_1C_1 + c_2C_2) + h^3(c_1D_1 + c_2D_2) + O(h^4).....(8)$$

where the coefficients A_r, B_r, C_r and D_r, r = 1, 2, are given above.

Comparing with the total differential expansion of $\phi(x, y, h)$ of eq.(8), we have:

$$\phi(x,y,h)= f + \frac{1}{2}hF + \frac{1}{6}h^2(Ff_y+G) + \frac{1}{24}h^3[(3f_{xy}+3ff_{yy}+f_y^2)F + Gf_y + H]+O(h^4).....(9)$$

where:

$$F = f_x + ff_y$$

$$G = f_{xx} + 2ff_{xy} + f^2f_{yy}$$

$$H = f_{xxx} + 3ff_{xxy} + 3f^2f_{xyy} + f^3f_{yyy}$$

Comparing eqs.(8) and (9), we have the following case:

(i) Two stages implicit Runge-Kutta method of order one, if:

$$c_1 + c_2 = 1$$

(ii) Two stages implicit Runge-Kutta method of order two, if:

$$c_1 + c_2 = 1$$

$$c_1a_1 + c_2a_2 = 1/2$$

(iii) Two stages implicit Runge-Kutta method of order three, if:

$$c_1 + c_2 = 1$$

$$c_1a_1 + c_2a_2 = 1/2$$

$$c_1 a_1^2 + c_2(b_{21}a_1 + b_{22}a_2) = 1/6$$

$$c_1 a_1^2 + c_2 a_2^2 = 1/3$$

(iv) Two stages implicit Runge-Kutta method of order four, if:

$$c_1 + c_2 = 1$$

$$c_1a_1 + c_2a_2 = 1/2$$

$$c_1 a_1^2 + c_2(b_{21}a_1 + b_{22}a_2) = 1/6$$

$$c_1 a_1^2 + c_2 a_2^2 = 1/3$$

$$(c_1b_{11} + c_2b_{21})\ a_1^2 + c_2b_{22}(b_{21}a_1 + b_{22}a_2) = 1/24$$

$$c_1 a_1^3 c_2a_2(b_{21}a_1 + b_{22}a_2) = 1/8$$

$$c_1 a_1^3 + c_2(b_{21} a_1^2 + b_{22} a_2^2) = 1/12$$

$$c_1 a_1^3 + c_2 a_2^3 = 1/4$$

From the above results, one can see that more accurate methods (of higher order) could be used with small stages. For example, one of the solutions to the fourth order method is given by:

$$c_1 = c_2 = \frac{1}{2},\ a_{1,2} = \frac{1}{2} \mp \frac{\sqrt{3}}{6},\ b_{11} = \frac{1}{2} - \frac{\sqrt{3}}{6},\ b_{21} = a_2 - \frac{1}{4} = \frac{1}{4} + \frac{\sqrt{3}}{6}$$

and $b_{22} = \frac{1}{4}$

2-3 Derivation of 2-Stages Implicit Runge-Kutta Method:

Consider the 2-stages implicit Runge-Kutta method, which takes the form:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

where:

$$k_1 = f(x_n + ha_1,\ y_n + b_{11}hk_1 + b_{12}hk_2)$$

$$k_2 = f(x_n + ha_2,\ y_n + b_{21}hk_1 + b_{22}hk_2)$$

and $a_1 = b_{11} + b_{12}$, $a_2 = b_{21} + b_{22}$.

Now, expanding k_r using Taylor series expansion about (x_n, y_n), we obtain that for r = 1, 2.

$$k_r = f(x_n, y_n) + h[a_r f_x + (b_{r1}k_1 + b_{r2}k_2)f_y] + \frac{h^2}{2}[a_r^2 f_{xx} + 2a_r(b_{r1}k_1 + b_{r2}k_2)f_{xy} + (b_{r1}k_1 + b_{r2}k_2)^2 f_{yy}] + \frac{h^3}{6}[a_r^3 f_{xxx} + 3a_r^2(b_{r1}k_1 + b_{r2}k_2)f_{xxy} + 3a_r(b_{r1}k_1 + b_{r2}k_2)^2 f_{xyy} + (b_{r1}k_1 + b_{r2}k_2)^3 f_{yyy}] + O(h^4) \quad \text{.....(10)}$$

which will gives two implicit equations for r = 1, 2. Hence, we can no longer proceed as in previous derivations and Runge-Kutta methods. Therefore, as in the semi-explicit methods, suppose that:

$$k_r = A_r + hB_r + h^2C_r + h^3D_r + O(h^4),\ r = 1, 2 \quad \text{.....(11)}$$

Equating eqs. (10) and (11), we get:

$$A_r + hB_r + h^2C_r + h^3D_r = f + h[a_r f_x + (b_{r1}(A_1 + hB_1 + h^2C_1) + b_{r2}(A_2 + hB_2 + h^2C_2))f_y] + \frac{h^2}{2}[a_r^2 f_{xx} + 2a_r(b_{r1}(A_1 + hB_1) + b_{r2}(A_2 + hB_2))f_{xy} + (b_{r1}(A_1 + hB_1) + b_{r2}(A_2 + hB_2))^2 f_{yy} + \frac{h^3}{6}[a_r^3 f_{xxx} + 3a_r^2(b_{r1}A_1 + b_{r2}A_2)f_{xxy} + 3a_r(b_{r1}A_1 + b_{r2}A_2)^2 f_{xyy} + (b_{r1}A_1 + b_{r2}A_2)^3 f_{yyy}] + O(h^4),\ r = 1, 2$$

Equating the coefficients of h^0, h^1, h^2 and h^3, we obtain that:

$$A_r = f$$

$$B_r = a_r f_x + (b_{r1}A_1 + b_{r2}A_2)f_y$$

$$C_r = (b_{r1}B_1 + b_{r2}B_2)f_y + \frac{1}{2}a_r^2 f_{xx} + a_r(b_{r1}A_1\ b_{r2}A_2)f_{xy} + \frac{1}{2}(b_{r1}A_1 + b_{r2}A_2)^2 f_{yy}$$

$$D_r = (b_{r1}C_1 + b_{r2}C_2)f_y + a_r(b_{r1}B_1 + b_{r2}B_2)f_{xy} + (b_{r1}A_1 + b_{r2}A_2)(b_{21}B_1 + b_{r2}B_2)f_{yy} + \frac{1}{6}a_r^3 f_{xxx} + \frac{1}{2}a_r^2(b_{r1}A_1 + b_{r2}A_2)f_{xxy} + \frac{1}{2}a_r(b_{r1}A_1 + b_{r2}A_2)^2 f_{xyy} + \frac{1}{6}(b_{r1}A_1 + b_{r2}A_2)^3 f_{yyy},\ r = 1, 2.$$

Since $a_1 = b_{11} + b_{12}$ and $a_2 = b_{21} + b_{22}$, or in general notation $a_r = b_{r1} + b_{r2}$. Therefore, using the notation for a_1 and a_2 to simplify and solve the last form equations, we obtain the following solution:

$$A_r = f$$

$$B_r = a_r f_x + (b_{r1}f + b_{r2}f)f_y$$

$$= a_r f_x + (b_{r1} + b_{r2})ff_y$$

$$= a_r f_x + a_r f f_y$$

$$= ar(fx + ff_y) = a_r F$$

$$C_r = (b_{r1}a_1F + b_{r2}a_2F)f_y + \frac{1}{2}a_r^2 f_{xx} + a_r(b_{r1}f + b_{r2}f)f_{xy} + \frac{1}{2}(b_{r1}f + b_{r2}f)^2 f_{yy}$$

$$= (b_{r1}a_1 + b_{r2}a_2)Ff_y + \frac{1}{2}a_r^2 f_{xx} + a_r^2 ff_{xy} + \frac{1}{2}a_r^2 f^2 f_{yy}$$

$$= (b_{r1}a_1 + b_{r2}a_2)Ff_y + \frac{1}{2}a_r^2 (f_{xx} + 2ff_{xy} + f^2 f_{yy})$$

$$= (b_{r1}a_1 + b_{r2}a_2)Ff_y + \frac{1}{2}a_r^2 G$$

where $G = f_{xx} + 2ff_{xy} + f^2 f_{yy}$.

Similarly:

$$D_r = [b_{r1}(b_{11}a_1 + b_{12}a_2) + b_{r2}(b_{21}a_1 + b_{22}a_2)]F f_y^2 + a_r(b_{r1}a_1 + b_{r2}a_2)F(f_{xy} + ff_{yy}) + (b_{r1}a_1^2 + b_{r2}a_2^2)Gf_y + \frac{1}{6}a_r^3 H$$

where $H = f_{xxx} + 3ff_{xxy} + 3f^2 f_{xyy} + f^3 f_{yyy}$.

Since:

$$\varphi(x, y, h) = \sum_{r=1}^{2} c_r k_r$$

Hence using eq.(11), one get:

$$\varphi = c_1k_1 + c_2k_2$$

$$= c_1(A_1 + hB_1 + h^2C_1 + h^3D_1) + c_2(A_2 + hB_2 + h^2C_2 + h^3D_2) + O(h^4)$$

$$= c_1A_1 + c_2A_2 + h(c_1B_1 + c_2B_2) + h^2(c_1C_1 + c_2C_2) + h^3(c_1D_1 + c_2D_2) + O(h^4) \quad .....(12)$$

where the coefficients A_r, B_r, C_r and D_r, $\forall$ r =1, 2, are given above.

Comparing with the total differential expansion of $\varphi(x, y, h)$ of eq.(12)

$$\varphi(x,y, h) = f + \frac{1}{2}hF + \frac{1}{6}h^2(Ff_y+G) + \frac{1}{24}h^3[(3f_{xy}+3ff_{yy}+f_y^2)F + Gf_y + H]+O(h^4) \quad ..(13)$$

where:

$$F = f_x + ff_y$$

$$G = f_{xx} + 2ff_{xy} + f^2 f_{yy}$$

$$H = f_{xxx} + 3ff_{xxy} + 3f^2 f_{xyy} + f^3 f_{yyy}$$

Comparing eqs.(12) and (13), we have the following cases:

(i) Two stages implicit Runge-Kutta method of order one, if:

$$c_1 + c_2 = 1$$

(ii) Two stages implicit Runge-Kutta method of order two, if:

$$c_1 + c_2 = 1$$

$$c_1a_1 + c_2a_2 = 1/2$$

(iii) Two stages implicit Runge-Kutta method of order p = 3, if:

$$c_1 + c_2 = 1$$

$$c_1a_1 + c_2a_2 = 1/2$$

$$c_1(b_{11}a_1 + b_{12}a_2) + c_2(b_{21}a_1 + b_{22}a_2) = 1/6$$

$$c_1 a_1^2 + c_2 a_2^2 = 1/3$$

(iv) Two stages implicit Runge-Kutta method of order p = 4, if:

$$c_1 + c_2 = 1$$

$$c_1a_1 + c_2a_2 = 1/2$$

$$c_1(b_{11}a_1 + b_{12}a_2) + c_2(b_{21}a_1 + b_{22}a_2) = 1/6$$

$$c_1 a_1^2 + c_2 a_2^2 = 1/3$$

$$(c_1b_{11} + c_2b_{21})(b_{11}a_1 + b_{12}a_2) + (c_1b_{12} + c_2b_{22})(b_{21}a_1 + b_{22}a_2) = 1/24$$

$$c_1a_1(b_{11}a_1 + b_{12}a_2) + c_2a_2(b_{21}a_1 + b_{22}a_2) = 1/8$$

$$c_1(b_{11} a_1^2 + b_{12} a_2^2) + c_2(b_{21} a_1^2 + b_{22} a_2^2) = 1/12$$

$$c_1 a_1^3 + c_2 a_2^3 = 1/4$$

From the above results, one can see that more accurate methods (of higher order) could be obtained with small stages. For example, one of the solutions to the fourth order method is given by:

$$c_1 = c_2 = \frac{1}{2},\ a_{1,2} = \frac{1}{2} \mp \frac{\sqrt{3}}{6},$$

$$b_{11} = b_{22} = \frac{1}{4},\ b_{12} = a_1 - \frac{1}{4} = \frac{1}{4} - \frac{\sqrt{3}}{6},\ b_{21} = a_2 - \frac{1}{4} = \frac{1}{4} + \frac{\sqrt{3}}{6}$$

3- STABILITY AND CONVERGENCE OF RUNGE-KUTTA METHODS

Since the purpose of numerical analysis is to represent the solution to actual problems. It is important that what could be called qualitative properties of the numerical solution should resemble those of the true solution.

By stability analysis, we shall mean study of such qualitative properties as boundedness and convergence to zero of numerical solutions, when these properties are passed by the exact solution. Given a slightly different emphasis, this type of analysis is appropriate for studying the growth of numerical errors in a computed solution to a differential equation.

Remark (1):

Recalling the general form of Runge-Kutta method, which is:

$$y_{n+1} = y_n + h\varphi(x_n, y_n, h) \quad \text{.....(14)}$$

which could be considered as a special case of the general Linear Multistep Methods (for short LMM), given by:

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f_{n+j} \quad \text{.....(15)}$$

(indeed, one step explicit LMM). Therefore, the stability of eq.(14) is equivalent to the convergence of eq.(14) (i.e., stability and convergence of Runge-Kutta are equivalent).

Theorem (1):

Assume the consistency condition and suppose that:

$$y_{n+1} - y_n = h\varphi(x_n, y_n, h) \quad \text{..... (1.19(16)}$$

which is a special case of eq.(15) as one step method with $\alpha_1 = 1$, $\alpha_0 = -1$, $\beta_0 = 1$, $\beta_1 = 0$ and $f = \varphi$. Then eq.(16) is converge if and only if it is zero stable.

Proof:

Suppose that the method is convergent and to prove that the method is zero stable

Consider for simplicity the problem, $y' = 0$, $y(0) = 0$, which has the exact solution $y(x) = 0$

Since $f = 0$, then for all r, we have $k_r = 0$ and hence $\varphi = 0$

Therefore, the method takes the form:

$$y_{n+1} - y_n = 0$$

Therefore, the first characteristic polynomial $\rho(r)$ is given by:

$$\rho(r) = r - 1 = 0$$

then $r = 1$.

and since $|r| \leq 1$, then the method is zero stable

Conversely, suppose that the Runge-Kutta method is zero stable and to prove that the method is converge

Similarly, for simplicity purpose consider $y' = \lambda y$, $y(0) = 1$

To show that the term $[r_0(\lambda h)]^n$ in the general solution

$$y_0 = [r_0(\lambda h)]^n$$

will be converge to the exact solution $y(x) = e^{\lambda x}$ on $[0, b]$, and then can be shown to be converge to the zero solution as $h \longrightarrow 0$

Expanding $r_0(\lambda h)$ using Taylor's theorem:

$$r_0(\lambda h) = r_0(0) + h\lambda r'_0(0) + O(h^2)$$

Hence:

$$r'_0(0) = \frac{\sigma(r_0(0))}{\rho'(r_0(0))}$$

Since $r_0(0) = 1$. Then:

$$r'_0(0) = \frac{\sigma(1)}{\rho'(1)} = 1$$

Hence:

$$\begin{aligned} r_0(\lambda h) &= 1 + \lambda h + O(h^2) \\ &= e^{\lambda h} - O(h^2) + O(h^2) \\ &\simeq e^{\lambda h} \end{aligned}$$

Therefore:

$$[r_0(\lambda h)]^n = [e^{\lambda h}]^n = e^{\lambda_n h} = e^{\lambda x_n}$$

Hence:

$$\underset{0 \leq x_n \leq b}{\text{Max}} |[r_0(\lambda h)]^n - e^{\lambda x_n}| \longrightarrow 0 \text{ as } h \longrightarrow 0.$$

3-1 Stability of Explicit Runge-Kutta Method:

To obtain intervals of stability of 2-stages Runge-Kutta method applied to the test problem $y' = \lambda y$, $\lambda < 0$, we have:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

where:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f(x_n + a_2h, y_n + a_2hk_1)$$

and for $f(x, y) = \lambda y$, we have:

$$k_1 = \lambda y_n$$

$$k_2 = \lambda(y_n + a_2\lambda h y_n)$$

$$= \lambda y_n + a_2\lambda^2 h y_n$$

since:

$$y_{n+1} = y_n + h(c_1\lambda y_n + c_2\lambda y_n + c_2a_2h\lambda^2 y_n)$$

$$= y_n + (c_1 + c_2)\hbar\, y_n + c_2a_2\hbar^2 y_n,\ \hbar = \lambda h$$

$$= y_n[1 + (c_1 + c_2)\ \hbar\ + c_2a_2\hbar^2]$$

Hence, to find the roots of the first characteristic polynomial $\rho(r) = 0$, we have to letting $y_n = r^n$, so:

$$r^{n+1} - r^n[1 + (c_1 + c_2)\hbar\ + c_2a_2\hbar^2] = 0$$

$$r^n\{r - [1 + (c_1 + c_2)\ \hbar\ + c_2a_2\hbar^2]\} = 0$$

and since $r^n \neq 0$, then:

$$r - [1 + (c_1 + c_2)\ \hbar\ + c_2a_2\hbar^2] = 0$$

i.e.,

$$r = 1 + (c_1 + c_2)\ \hbar\ + c_2a_2\hbar^2,\ |r| < 1$$

since for 2-stages Runge-Kutta method, p = 2, we have:

$$c_1 + c_2 = 1 \text{ and } c_2a_2 = \frac{1}{2}$$

then:

$$r = 1 + \hbar\ + \frac{1}{2}\hbar$$

In order to get a method which is stable, we must have $|r| < 1$, so:

$$|1 + \hbar + \frac{1}{2}\hbar^2| < 1$$

Then

$$-1 < 1 + \hbar + \frac{1}{2}\hbar^2 < 1$$

Which implies that $-2 < \hbar < 0$. So all R = p = 2 Runge-Kutta methods have an interval of absolute stability to be (–2, 0).

3-2 Stability of Semi-Explicit Runge-Kutta Method:

To obtain intervals of stability of 2-stages Runge-Kutta methods. Also, we consider the test problem $y' = \lambda y$, where $\mathrm{Re}(\lambda) < 0$:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

where:

$$k_1 = \lambda(y_n + ha_1k_1)$$

$$k_2 = \lambda(y_n + hb_{21}k_1 + hb_{22}k_2)$$

so:

$$k_1 = \frac{\lambda y_n}{1-\hbar a_1}$$

$$k_2 = \frac{\lambda y_n[(1-\hbar a_1)+\hbar b_{21}]}{(1-\hbar b_{22})(1-\hbar a_1)}$$

so:

$$y_{n+1} = y_n + \frac{h}{2}\left[\frac{\lambda y_n}{1-\hbar a_1} + \frac{\lambda y_n[1-\hbar a_1+\hbar b_{21}]}{(1-\hbar b_{22})(1-\hbar a_1)}\right]$$

$$= y_n + \frac{\hbar}{2}y_n\left[\frac{1}{1-\hbar a_1} + \frac{1-\hbar a_1+\hbar b_{21}}{(1-\hbar b_{22})(1-\hbar a_1)}\right]$$

$$= y_n + \frac{\hbar}{2}y_n\left[\frac{(1-\hbar b_{22})+(1-\hbar a_1+\hbar b_{21})}{(1-\hbar b_{22})(1-\hbar a_1)}\right]$$

$$= y_n + \frac{\hbar}{2}y_n\left[\frac{2(1-\hbar b_{22})}{(1-\hbar b_{22})(1-\hbar a_1)}\right]$$

$$= y_n\left[1+\frac{\hbar}{1-\hbar a_1}\right]$$

So, the corresponding root is given by:

$$r = 1 + \frac{\hbar}{1-\hbar a_1} = \frac{1-\hbar a_1+\hbar}{1-\hbar a_1} = \frac{s+\hbar}{s},\ \hbar < 0$$

So $|r| = \left|\frac{s+\hbar}{s}\right| < 1$

Hence, the interval of stability is given by $(-\infty, 0)$.

3-3 Stability of Implicit Runge-Kutta Method:

To obtain intervals of stability of 2-stages, fourth order Runge-Kutta method. Consider the test problem $y' = \lambda y$, where $Re(\lambda) < 0$, hence:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

where:

$$k_1 = f(x_n + ha_1, y_n + b_{11}hk_1 + b_{12}hk_2)$$

$$k_2 = f(x_n + ha_2, y_n + b_{21}hk_1 + b_{22}hk_2)$$

so for the test problem $y' = \lambda y$, we have:

$$k_1 = \lambda y_n + \frac{1}{4}\hbar k_1 + (\frac{1}{4} - \frac{\sqrt{3}}{6})\hbar k_2$$

$$k_2 = \lambda y_n + (\frac{1}{4} + \frac{\sqrt{3}}{6})\hbar k_1 + \frac{1}{4}\hbar k_2$$

Therefore, when $\hbar = \lambda h$, we have:

$$k_1 = \frac{\left[1-\left(\frac{1}{4}+\frac{\sqrt{3}}{6}\right)\hbar+\left(\frac{\sqrt{3}}{24}\right)\hbar^2\right]\lambda y_n}{1-\frac{3}{4}\hbar+\frac{15}{72}\hbar^2-\frac{3}{144}\hbar^3}$$

$$k_2 = \frac{\left[1+\frac{\sqrt{3}}{6}\hbar\right]\lambda y_n}{1-\frac{1}{2}\hbar+\frac{1}{12}\hbar^2}$$

and since:

$$y_{n+1} = y_n + h(c_1k_1 + c_2k_2)$$

Therefore:

$$y_{n+1} = y_n + \frac{\hbar y_n}{2}\left[\frac{1-\left(\frac{1}{4}+\frac{\sqrt{3}}{6}\right)\hbar+\left(\frac{\sqrt{3}}{24}\right)\hbar^2}{1-\frac{3}{4}\hbar+\frac{15}{72}\hbar^2-\frac{3}{144}\hbar^3}+\frac{\left(1+\frac{\sqrt{3}}{6}\hbar\right)}{1-\frac{1}{2}\hbar+\frac{1}{12}\hbar^2}\right]$$

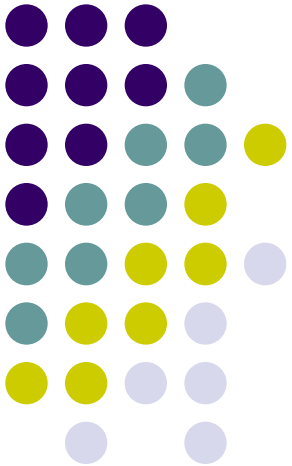
$$= y_n\left[1+\frac{\hbar-\left(\frac{1}{4}+\frac{\sqrt{3}}{6}\right)\hbar^2+\frac{\sqrt{3}}{24}\hbar^3}{2-\frac{3}{2}\hbar+\frac{15}{36}\hbar^2-\frac{3}{72}\hbar^3}+\frac{\hbar+\frac{\sqrt{3}}{6}\hbar^2}{2-\hbar-\frac{1}{6}\hbar^2}\right]$$

So the corresponding characteristic root is given by:

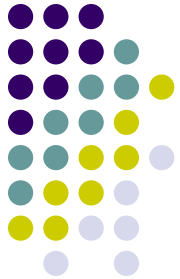
$$r = 1+\frac{\hbar-\left(\frac{1}{4}+\frac{\sqrt{3}}{6}\right)\hbar^2+\frac{\sqrt{3}}{24}\hbar^3}{2-\frac{3}{2}\hbar+\frac{15}{36}\hbar^2-\frac{3}{72}\hbar^3}+\frac{\hbar+\frac{\sqrt{3}}{6}\hbar^2}{2-\hbar-\frac{1}{6}\hbar^2}$$

Therefore, upon applying computer programming, the following interval of absolute stability is obtained, which is $\hbar \in (-1, -0.95) \cup (-0.64, 0)$.

Boundary Value Problems

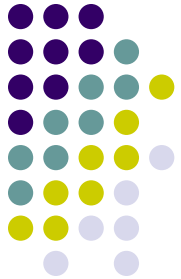


Boundary value problems



- Side conditions prescribing solution or derivative values at specified points are required to make solution of ODE unique
- For initial value problem, all side conditions are specified at single point, say t_0
- For *boundary value problem* (BVP), side conditions are specified at more than one point
- kth order ODE, or equivalent first-order system, requires k side conditions
- For ODEs, side conditions are typically specified at endpoints of interval $[a, b]$, so we have *two-point boundary value problem* with boundary conditions (BC) at a and b.

Boundary value problems



- General *first-order two-point BVP* has form

$$\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}), \qquad a < t < b$$

with BC

$$\boldsymbol{g}(\boldsymbol{y}(a), \boldsymbol{y}(b)) = \boldsymbol{0}$$

where $\boldsymbol{f} : \mathbb{R}^{n+1} \to \mathbb{R}^n$ and $\boldsymbol{g} : \mathbb{R}^{2n} \to \mathbb{R}^n$

- Boundary conditions are *separated* if any given component of $\boldsymbol{g}$ involves solution values only at a or at b, but not both
- Boundary conditions are *linear* if they are of form

$$\boldsymbol{B}_a \, \boldsymbol{y}(a) + \boldsymbol{B}_b \, \boldsymbol{y}(b) = \boldsymbol{c}$$

where $\boldsymbol{B}_a, \boldsymbol{B}_b \in \mathbb{R}^{n \times n}$ and $\boldsymbol{c} \in \mathbb{R}^n$

- BVP is *linear* if ODE and BC are both linear

Example

- Two-point BVP for second-order scalar ODE

$$u'' = f(t, u, u'), \qquad a < t < b$$

with BC

$$u(a) = \alpha, \qquad u(b) = \beta$$

is equivalent to first-order system of ODEs

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} y_2 \\ f(t, y_1, y_2) \end{bmatrix}, \qquad a < t < b$$

with separated linear BC

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1(a) \\ y_2(a) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_1(b) \\ y_2(b) \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Existence and uniqueness



- Unlike IVP, with BVP we cannot begin at initial point and continue solution step by step to nearby points
- Instead, solution is determined everywhere simultaneously, so existence and/or uniqueness may not hold
- For example,

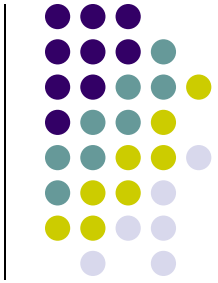
$$u'' = -u, \qquad 0 < t < b$$

with BC

$$u(0) = 0, \qquad u(b) = \beta$$

with b integer multiple of π, has infinitely many solutions if $\beta = 0$, but no solution if $\beta \neq 0$

Existence and uniqueness



- In general, solvability of BVP

$$\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}), \qquad a < t < b$$

with BC

$$\boldsymbol{g}(\boldsymbol{y}(a), \boldsymbol{y}(b)) = \boldsymbol{0}$$

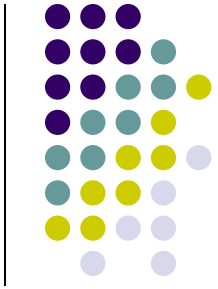
depends on solvability of algebraic equation

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}(b; \boldsymbol{x})) = \boldsymbol{0}$$

where $\boldsymbol{y}(t; \boldsymbol{x})$ denotes solution to ODE with initial condition $\boldsymbol{y}(a) = \boldsymbol{x}$ for $\boldsymbol{x} \in \mathbb{R}^n$

- Solvability of latter system is difficult to establish if $\boldsymbol{g}$ is nonlinear

Existence and uniqueness



- For *linear* BVP, existence and uniqueness are more tractable
- Consider linear BVP

$$\boldsymbol{y}' = \boldsymbol{A}(t)\,\boldsymbol{y} + \boldsymbol{b}(t), \qquad a < t < b$$

where $\boldsymbol{A}(t)$ and $\boldsymbol{b}(t)$ are continuous, with BC

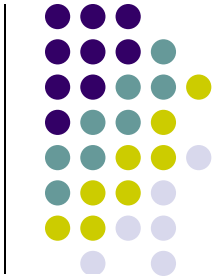
$$\boldsymbol{B}_a\,\boldsymbol{y}(a) + \boldsymbol{B}_b\,\boldsymbol{y}(b) = \boldsymbol{c}$$

- Let $\boldsymbol{Y}(t)$ denote matrix whose ith column, $\boldsymbol{y}_i(t)$, called ith *mode*, is solution to $\boldsymbol{y}' = \boldsymbol{A}(t)\boldsymbol{y}$ with initial condition $\boldsymbol{y}(a) = \boldsymbol{e}_i$, ith column of identity matrix
- Then BVP has unique solution if, and only if, matrix

$$\boldsymbol{Q} \equiv \boldsymbol{B}_a\boldsymbol{Y}(a) + \boldsymbol{B}_b\boldsymbol{Y}(b)$$

is nonsingular

Existence and uniqueness



- Assuming $\boldsymbol{Q}$ is nonsingular, define

$$\boldsymbol{\Phi}(t) = \boldsymbol{Y}(t)\,\boldsymbol{Q}^{-1}$$

and *Green's function*

$$\boldsymbol{G}(t,s) = \begin{cases} \boldsymbol{\Phi}(t)\boldsymbol{B}_a\boldsymbol{\Phi}(a)\boldsymbol{\Phi}^{-1}(s), & a \le s \le t \\ -\boldsymbol{\Phi}(t)\boldsymbol{B}_b\,\boldsymbol{\Phi}(b)\boldsymbol{\Phi}^{-1}(s), & t < s \le b \end{cases}$$

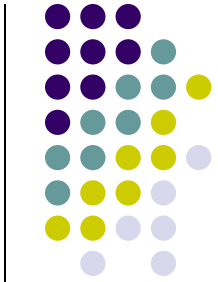
- Then solution to BVP given by

$$\boldsymbol{y}(t) = \boldsymbol{\Phi}(t)\,\boldsymbol{c} + \int_a^b \boldsymbol{G}(t,s)\,\boldsymbol{b}(s)\,ds$$

- This result also gives absolute condition number for BVP

$$\kappa = \max\{\|\boldsymbol{\Phi}\|_\infty,\ \|\boldsymbol{G}\|_\infty\}$$

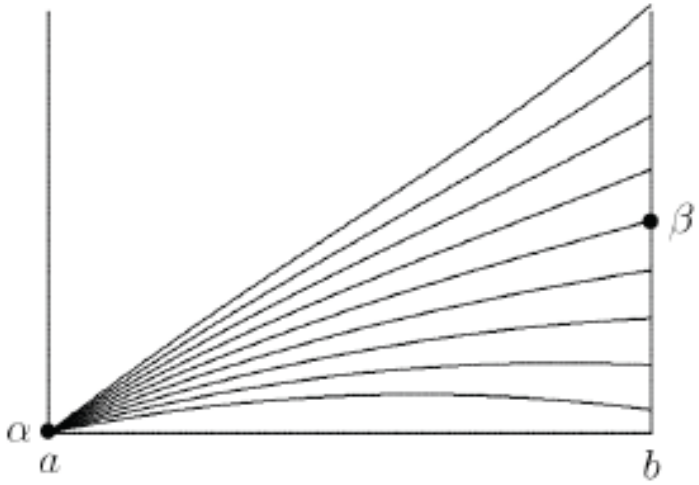
Numerical methods for BVPs



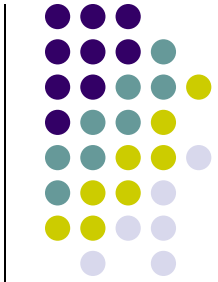
- For IVP, initial data supply all information necessary to begin numerical solution method at initial point and step forward from there
- For BVP, we have insufficient information to begin step-by-step numerical method, so numerical methods for solving BVPs are more complicated than those for solving IVPs
- We will consider four types of numerical methods for two-point BVPs
  - Shooting
  - Finite difference
  - Collocation

Shooting method

- In statement of two-point BVP, we are given value of $u(a)$
- If we also knew value of $u'(a)$, then we would have IVP that we could solve by methods discussed previously
- Lacking that information, we try sequence of increasingly accurate guesses until we find value for $u'(a)$ such that when we solve resulting IVP, approximate solution value at $t = b$ matches desired boundary value, $u(b) = \beta$



Shooting method



- For given γ, value at b of solution $u(b)$ to IVP

$$u'' = f(t, u, u')$$

with initial conditions

$$u(a) = \alpha, \qquad u'(a) = \gamma$$

can be considered as function of γ, say $g(\gamma)$

- Then BVP becomes problem of solving equation $g(\gamma) = \beta$
- One-dimensional zero finder can be used to solve this scalar equation

Example

- Consider two-point BVP for second-order ODE

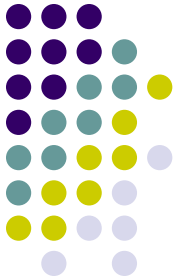
$$u'' = 6t, \qquad 0 < t < 1$$

with BC

$$u(0) = 0, \qquad u(1) = 1$$

- For each guess for $u'(0)$, we will integrate resulting IVP using classical fourth-order Runge-Kutta method to determine how close we come to hitting desired solution value at $t = 1$
- For simplicity of illustration, we will use step size $h = 0.5$ to integrate IVP from $t = 0$ to $t = 1$ in only two steps
- First, we transform second-order ODE into system of two first-order ODEs $\boldsymbol{y}'(t) = \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} y_2 \\ 6t \end{bmatrix}$

Example



- We first try guess for initial slope of $y_2(0) = 1$

$$
\begin{aligned}
\boldsymbol{y}^{(1)} &= \boldsymbol{y}^{(0)} + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\
&= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \frac{0.5}{6}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} + 2\begin{bmatrix} 1.375 \\ 1.500 \end{bmatrix} + \begin{bmatrix} 1.75 \\ 3.00 \end{bmatrix} \right) = \begin{bmatrix} 0.625 \\ 1.750 \end{bmatrix} \\
\boldsymbol{y}^{(2)} &= \begin{bmatrix} 0.625 \\ 1.750 \end{bmatrix} + \frac{0.5}{6}\left(\begin{bmatrix} 1.75 \\ 3.00 \end{bmatrix} + 2\begin{bmatrix} 2.5 \\ 4.5 \end{bmatrix} + 2\begin{bmatrix} 2.875 \\ 4.500 \end{bmatrix} + \begin{bmatrix} 4 \\ 6 \end{bmatrix} \right) = \begin{bmatrix} 2 \\ 4 \end{bmatrix}
\end{aligned}
$$

- So we have hit $y_1(1) = 2$ instead of desired value $y_1(1) = 1$

Example

- We try again, this time with initial slope $y_2(0) = -1$

$$
\begin{aligned}
y^{(1)} &= \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \frac{0.5}{6}\left(\begin{bmatrix} -1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} -1.0 \\ 1.5 \end{bmatrix} + 2\begin{bmatrix} -0.625 \\ 1.500 \end{bmatrix} + \begin{bmatrix} -0.25 \\ 3.00 \end{bmatrix}\right) \\
&= \begin{bmatrix} -0.375 \\ -0.250 \end{bmatrix} \\
y^{(2)} &= \begin{bmatrix} -0.375 \\ -0.250 \end{bmatrix} + \frac{0.5}{6}\left(\begin{bmatrix} -0.25 \\ 3.00 \end{bmatrix} + 2\begin{bmatrix} 0.5 \\ 4.5 \end{bmatrix} + 2\begin{bmatrix} 0.875 \\ 4.500 \end{bmatrix} + \begin{bmatrix} 2 \\ 6 \end{bmatrix}\right) \\
&= \begin{bmatrix} 0 \\ 2 \end{bmatrix}
\end{aligned}
$$

- So we have hit $y_1(1) = 0$ instead of desired value $y_1(1) = 1$, but we now have initial slope bracketed between -1 and 1

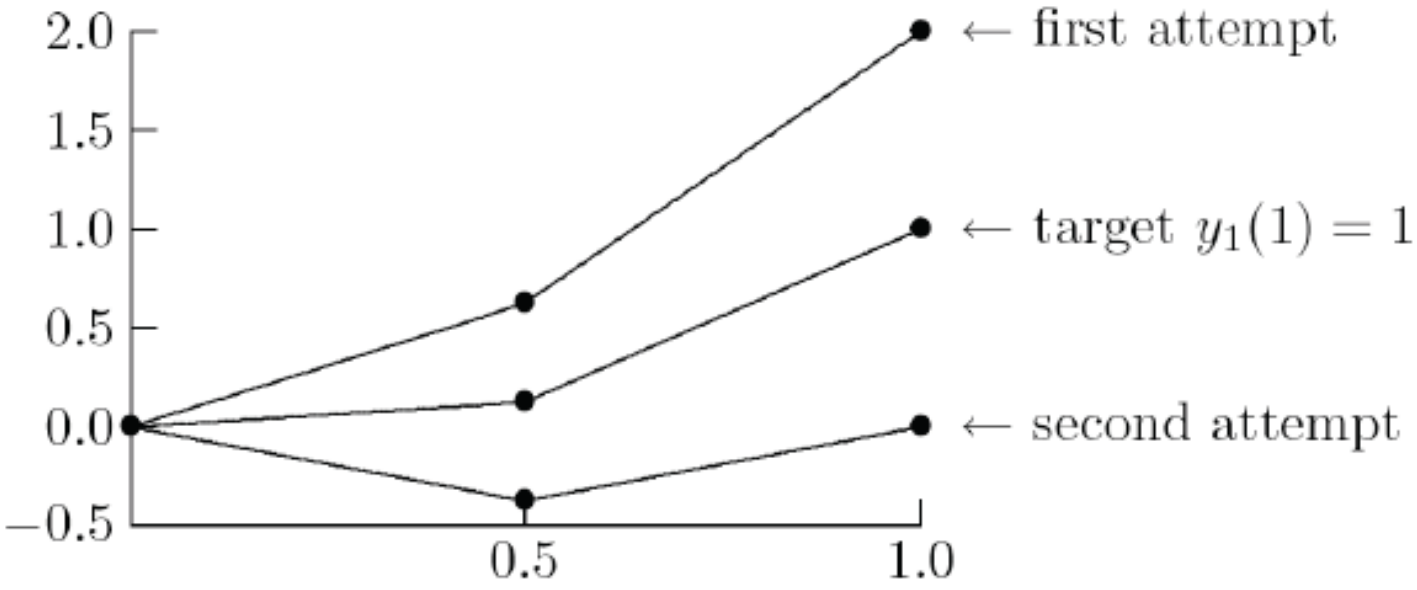
Example

- We omit further iterations necessary to identify correct initial slope, which turns out to be $y_2(0) = 0$

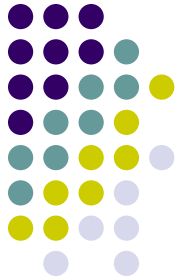
$$
\begin{aligned}
y^{(1)} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \frac{0.5}{6}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0.0 \\ 1.5 \end{bmatrix} + 2\begin{bmatrix} 0.375 \\ 1.500 \end{bmatrix} + \begin{bmatrix} 0.75 \\ 3.00 \end{bmatrix} \right) \\
&= \begin{bmatrix} 0.125 \\ 0.750 \end{bmatrix} \\
y^{(2)} &= \begin{bmatrix} 0.125 \\ 0.750 \end{bmatrix} + \frac{0.5}{6}\left(\begin{bmatrix} 0.75 \\ 3.00 \end{bmatrix} + 2\begin{bmatrix} 1.5 \\ 4.5 \end{bmatrix} + 2\begin{bmatrix} 1.875 \\ 4.500 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \end{bmatrix} \right) \\
&= \begin{bmatrix} 1 \\ 3 \end{bmatrix}
\end{aligned}
$$

- So we have indeed hit target solution value $y_1(1) = 1$

Example (cont.)

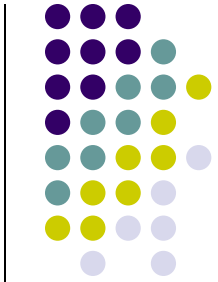


Multiple shooting



- Simple shooting method inherits stability (or instability) of associated IVP, which may be unstable even when BVP is stable
- Such ill-conditioning may make it difficult to achieve convergence of iterative method for solving nonlinear equation
- Potential remedy is *multiple shooting*, in which interval $[a, b]$ is divided into subintervals, and shooting is carried out on each
- Requiring continuity at internal mesh points provides BC for individual subproblems
- Multiple shooting results in larger *system* of nonlinear equations to solve

Finite difference method



- *Finite difference method* converts BVP into system of algebraic equations by replacing all derivatives with finite difference approximations
- For example, to solve two-point BVP

$$u'' = f(t, u, u'), \qquad a < t < b$$

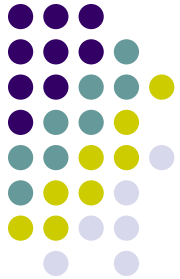
with BC

$$u(a) = \alpha, \qquad u(b) = \beta$$

we introduce mesh points $t_i = a + ih$, $i = 0, 1, \ldots, n+1$, where $h = (b-a)/(n+1)$

- We already have $y_0 = u(a) = \alpha$ and $y_{n+1} = u(b) = \beta$ from BC, and we seek approximate solution value $y_i \approx u(t_i)$ at each interior mesh point t_i, $i = 1, \ldots, n$

Finite difference method



- We replace derivatives by finite difference approximations such as

$$u'(t_i) \approx \frac{y_{i+1} - y_{i-1}}{2h}$$

$$u''(t_i) \approx \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}$$

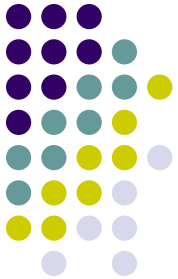
- This yields system of equations

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} = f\left(t_i, y_i, \frac{y_{i+1} - y_{i-1}}{2h}\right)$$

  to be solved for unknowns y_i, $i = 1, \ldots, n$

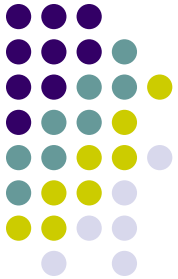
- System of equations may be linear or nonlinear, depending on whether f is linear or nonlinear

Finite difference method



- For these particular finite difference formulas, system to be solved is tridiagonal, which saves on both work and storage compared to general system of equations
- This is generally true of finite difference methods: they yield sparse systems because each equation involves few variables

Example



- Consider again two-point BVP

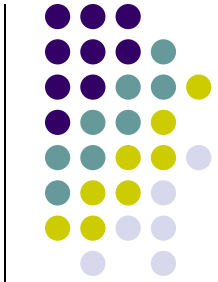
$$u'' = 6t, \qquad 0 < t < 1$$

with BC

$$u(0) = 0, \qquad u(1) = 1$$

- To keep computation to minimum, we compute approximate solution at one interior mesh point, $t = 0.5$, in interval $[0, 1]$
- Including boundary points, we have three mesh points, $t_0 = 0$, $t_1 = 0.5$, and $t_2 = 1$
- From BC, we know that $y_0 = u(t_0) = 0$ and $y_2 = u(t_2) = 1$, and we seek approximate solution $y_1 \approx u(t_1)$

Example



- Replacing derivatives by standard finite difference approximations at t_1 gives equation

$$\frac{y_2 - 2y_1 + y_0}{h^2} = f\left(t_1, y_1, \frac{y_2 - y_0}{2h}\right)$$

- Substituting boundary data, mesh size, and right hand side for this example we obtain

$$\frac{1 - 2y_1 + 0}{(0.5)^2} = 6t_1$$

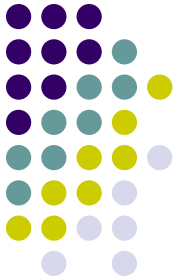
or

$$4 - 8y_1 = 6(0.5) = 3$$

so that

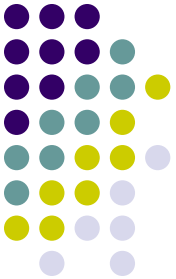
$$y(0.5) \approx y_1 = 1/8 = 0.125$$

Example (cont.)



- In a practical problem, much smaller step size and many more mesh points would be required to achieve acceptable accuracy
- We would therefore obtain *system* of equations to solve for approximate solution values at mesh points, rather than single equation as in this example

Collocation method



- *Collocation method* approximates solution to BVP by finite linear combination of basis functions
- For two-point BVP

$$u'' = f(t, u, u'), \qquad a < t < b$$

with BC

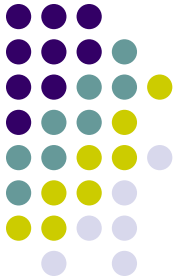
$$u(a) = \alpha, \qquad u(b) = \beta$$

we seek approximate solution of form

$$u(t) \approx v(t, x) = \sum_{i=1}^{n} x_i \phi_i(t)$$

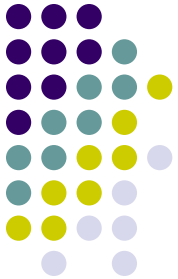
where ϕ_i are basis functions defined on $[a, b]$ and x is n-vector of parameters to be determined

Collocation method



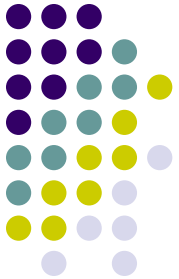
- Popular choices of basis functions include polynomials, B-splines, and trigonometric functions
- Basis functions with global support, such as polynomials or trigonometric functions, yield *spectral method*
- Basis functions with highly localized support, such as B-splines, yield *finite element method*

Collocation method



- To determine vector of parameters x, define set of n *collocation points*, $a = t_1 < \cdots < t_n = b$, at which approximate solution $v(t, x)$ is forced to satisfy ODE and boundary conditions
- Common choices of collocation points include equally-spaced points or Chebyshev points
- Suitably smooth basis functions can be differentiated analytically, so that approximate solution and its derivatives can be substituted into ODE and BC to obtain system of algebraic equations for unknown parameters x

Example



- Consider again two-point BVP

$$u'' = 6t, \qquad 0 < t < 1,$$

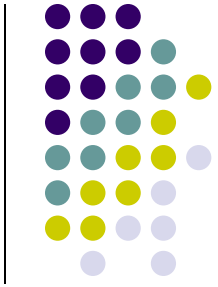
with BC

$$u(0) = 0, \qquad u(1) = 1$$

- To keep computation to minimum, we use one interior collocation point, $t = 0.5$
- Including boundary points, we have three collocation points, $t_0 = 0$, $t_1 = 0.5$, and $t_2 = 1$, so we will be able to determine three parameters
- As basis functions we use first three monomials, so approximate solution has form

$$v(t, \boldsymbol{x}) = x_1 + x_2 t + x_3 t^2$$

Example



- Derivatives of approximate solution function with respect to t are given by

$$v'(t, \boldsymbol{x}) = x_2 + 2x_3 t, \qquad v''(t, \boldsymbol{x}) = 2x_3$$

- Requiring ODE to be satisfied at interior collocation point $t_2 = 0.5$ gives equation

$$v''(t_2, \boldsymbol{x}) = f(t_2, v(t_2, \boldsymbol{x}), v'(t_2, \boldsymbol{x}))$$

  or

$$2x_3 = 6t_2 = 6(0.5) = 3$$

- Boundary condition at $t_1 = 0$ gives equation

$$x_1 + x_2 t_1 + x_3 t_1^2 = x_1 = 0$$

- Boundary condition at $t_3 = 1$ gives equation

$$x_1 + x_2 t_3 + x_3 t_3^2 = x_1 + x_2 + x_3 = 1$$

Example (cont.)

- Solving this system of three equations in three unknowns gives

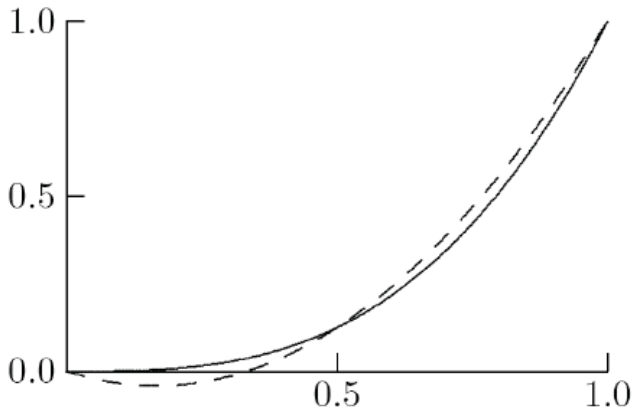
$$x_1 = 0, \qquad x_2 = -0.5, \qquad x_3 = 1.5$$

so approximate solution function is quadratic polynomial

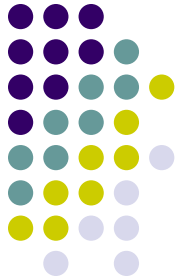
$$u(t) \approx v(t, \boldsymbol{x}) = -0.5t + 1.5t^2$$

- At interior collocation point, $t_2 = 0.5$, we have approximate solution value

$$u(0.5) \approx v(0.5, \boldsymbol{x}) = 0.125$$



Eigenvalue problems



- Standard eigenvalue problem for second-order ODE has form

$$u'' = \lambda f(t, u, u'), \qquad a < t < b$$

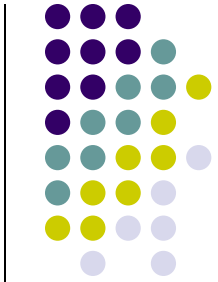
with BC

$$u(a) = \alpha, \qquad u(b) = \beta$$

where we seek not only solution u but also parameter λ

- Scalar λ (possibly complex) is *eigenvalue* and solution u is corresponding *eigenfunction* for this two-point BVP
- Discretization of eigenvalue problem for ODE results in algebraic eigenvalue problem whose solution approximates that of original problem

Example



- Consider linear two-point BVP

$$u'' = \lambda g(t)u, \qquad a < t < b$$

with BC

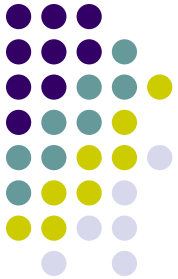
$$u(a) = 0, \qquad u(b) = 0$$

- Introduce discrete mesh points t_i in interval $[a, b]$, with mesh spacing h and use standard finite difference approximation for second derivative to obtain algebraic system

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} = \lambda g_i y_i, \quad i = 1, \ldots, n$$

where $y_i = u(t_i)$ and $g_i = g(t_i)$, and from BC $y_0 = u(a) = 0$ and $y_{n+1} = u(b) = 0$

Example (cont.)



- Assuming $g_i \neq 0$, divide equation i by g_i for $i = 1, \ldots, n$, to obtain linear system

$$\boldsymbol{A}\boldsymbol{y} = \lambda \boldsymbol{y}$$

where $n \times n$ matrix $\boldsymbol{A}$ has tridiagonal form

$$\boldsymbol{A} = \frac{1}{h^2}\begin{bmatrix} -2/g_1 & 1/g_1 & 0 & \cdots & 0 \\ 1/g_2 & -2/g_2 & 1/g_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1/g_{n-1} & -2/g_{n-1} & 1/g_{n-1} \\ 0 & \cdots & 0 & 1/g_n & -2/g_n \end{bmatrix}$$

- This standard algebraic eigenvalue problem can be solved by methods discussed previously