7. Galerkin Methods

7.1 Introduction

Until now, finite-difference methods for solving partial differential equations were applied. These methods specify the dependent variables at arbitrary grid-points in space and time and the derivatives in the equations are evaluated using Taylor series expansions. The definitions of convergence. accumulated error, ... are based on comparing the solution

$U_f^n = U(f\Delta x, n\Delta t)$

to the *continuous* solution $\mu(x,t)$ at grid point locations. The Galerkin procedure represents the dependent variables with a sum of functions that have a prescribed spatial structure. The coefficient associated with each function is normally a function of time. This procedure transforms a partial differential equation for the coefficients which are usually solved with finite difference in time. The two most useful Galerkin methods are the spedmal* method and the finite element method.

We now look at grid point values U_f^n as being sepersenative of grid-box averages of $\mu(x,t)$. Thus, in the case of one spatial dimension, we now compose

$$U_j^n = \frac{1}{\Delta x}\int_{(f-\frac{1}{2})\Delta x}^{(f+\frac{1}{2})\Delta x} U_f^n dxt$$

rather than U_f^n to $\mu(n\Delta x, u\Delta t)$.

We can reformulate by defining

(1) $\phi_j(x) =$

$\rightarrow \dfrac{1}{\Delta x}$ for $(j-\dfrac{1}{2})\Delta x < x < (f+\dfrac{1}{2})\Delta x$

$\rightarrow 0$ elsewhere.

and the U_f^n values as a vehicle for defining $\surd$ a piecewise constant approximation $U(x, n\Delta t)$ such that

$$(2) U(x, n\Delta t) = \sum_{f=-\infty}^{\infty} U_f^n \phi_j(x)$$

A formal series expansion for the purpose of approximating $\mu(x,t$ can be carried out with an infinite variety of function ϕ_j. The ones from (1) are representative of the traditional grid point values associated with the *straeland* finite difference equations.

Example: Advection Equation. $\frac{\partial \mu}{\partial t} + c\frac{\partial \mu}{\partial x} = 0$

By using (2), the advection equation can be rewritten:

(3) $\sum_f \frac{\partial U_f^n}{\partial t}\phi_f = -c\sum_f U_f^n \frac{\partial \phi_j}{\partial x}$

The *basic* function (1) are orthogonal →

$\int_{-\infty}^{\infty} \phi_i(x)\phi_j(x)dx =$

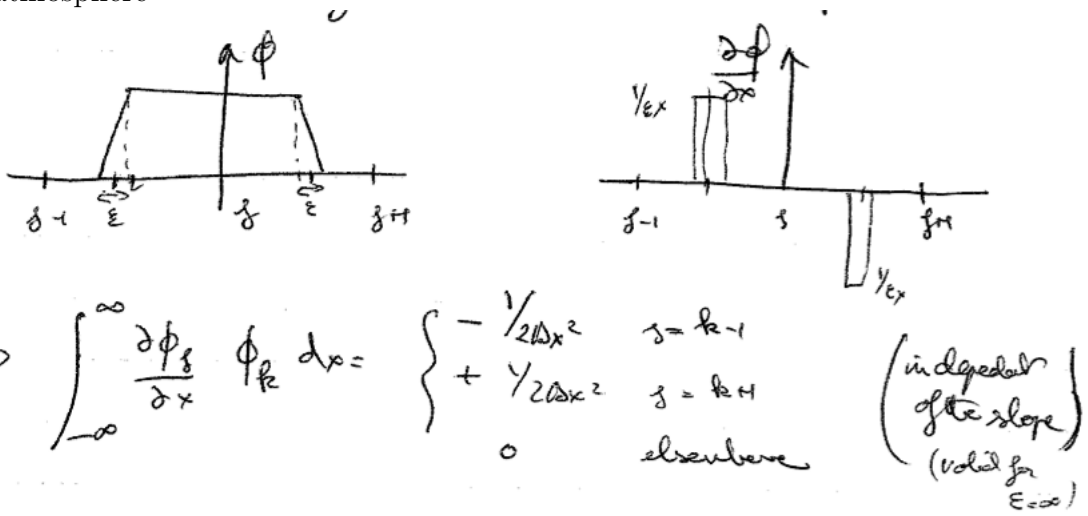
→ $(\Delta x)t*$ if $i = f$

→ 0 elsewhere

We can obtain an equation for $\frac{\partial U_f^n}{\partial t}$ by multiplying by ϕ_k and integrating over x.

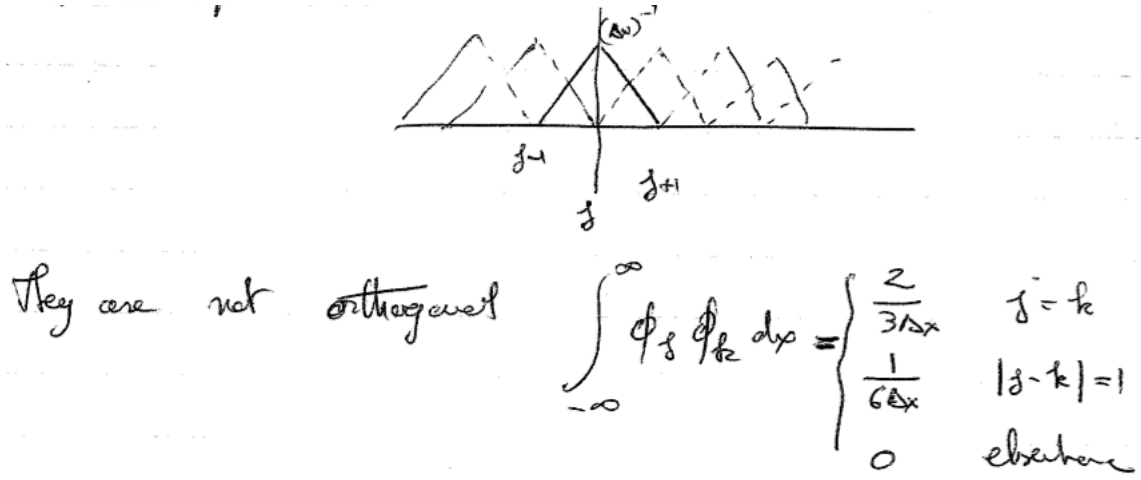
(4) $\sum_j \frac{\partial U_i^n}{\partial t}\int_{-\infty}^{\infty} \phi_f\phi_k dx = -c\sum_f U_f^n \int_{-\infty}^{\infty} \frac{\partial \phi_f}{\partial x}\phi_k dx$

In order to compute the RMS integral, we consider the step function to atmosphere*



(4) then becomes $\frac{\partial U_k^n}{\partial t} = -c\frac{U_{k+1}^n - U_{k-1}^n}{2\Delta x}$

* If we use a different ϕ such as a piecewise linear representation:



and seccall** in

(5) $\frac{1}{6}(\frac{\partial U_{k-1}^n}{\partial t} + 4\frac{\partial U_k^n}{\partial t} + \frac{\partial U_{k+1}^n}{\partial t}) = -c\frac{U_{k+1}^n - U_{k-1}^n}{2\Delta x}$

* We can generalize these ideas into a formed definition of the Galerkin Method.

Given a differential equation $L(\mu) > f(x)$ where L is a differential operator, μ the dependent variable and $f(x)$ a specified forcity* function in the domain R (x may be multidimensional). The Galerkin approximation is defined by:

(6)$U(x,t) = \sum_{j=1}^{N} a_f(t)\phi_j(x)$

where the coefficients $A_j(t)$ are determined by requiring that the error

(7)$e_N = L(U(x,t) - f(x)) = L(\sum_{f=1}^{N} A_j(t)\phi_f(x)) - f(x)$

be orthogonal to each basis* function.

(8) $\int_R e_N \phi_j(x)dx = a \qquad f = 1, ..., N$

The final form is:

(9)$\int_R \phi_k(\sum_{f=1}^{N} A_j(t)\phi_j(x))dx - \int_R \phi_k f(x)dx$

$$k = 1, ...N$$

This reduced to the problem of N algebraic equations that relate the unknown coefficients $A_j(t)$ to the "transforms" of the forcing* function. They are normally solved by finite difference in time.

there are various ways to interpret (8).

(1) The residual error is orthogonal to orthogonal to ϕ_f, i.e. the error should leave no components in the space spanned* by ϕ_j.

(2) The coefficients A_j should be chosen to minimize the integral $\int_R e^2(x,t,)dx$ straightforward when L is a linear operator. In more complicated cases, not measurably valid.

(3) $L(\mu) = f(x)$ is approximated by $L(U) = f(x)$ as in the Introduction.

Schemes employing basis functions defined in terms of periodic functions are referred to as "spectrial*". The ones using more " *****" basis functions are "finite elements" schemes.

7.2 Energy Conservation

If we consider the simplified equation

(10) $$\frac{\partial \mu}{\partial t} + L(\mu) = 0$$

then the Galerkin form is:

(11) $$\sum_{f=1}^{N} \frac{\partial A_j}{\partial t} \int_R \phi_k \phi_f dx + \int_R \phi_k \alpha (\sum_{f=1}^{N} A_f \phi_j) dx = 0 \quad k = 1, .., N$$

this process gives N coupled ordinary differential equations in the coefficients $A_f(t)$. this can be solved by introducing finite differences in time.

We already discussed the importance of energy conserving schemes. The Galerkin method leads naturally to energy conversation in equations with quadratic energy invariants.

For an energy conserving system:

(12) $$\int_R \frac{\partial \frac{\mu^2}{2}}{\partial t} = -\int_R \mu L(\mu) dx,$$

the operator L must satisfy the condition $\int_R \psi L(\psi) dx = 0$ where

ψ is any reasonable function that satisfies the boundary conditions. Then (12) satisfies

(13) $$\frac{d}{dt}\int_R \frac{\mu^2}{2}dx = 0$$

which shows the energy conservation for the exact equation. We need to demonstrate that it holds for the finite sum.

We multiply the k^th equation (11) by A_k and sum from $k = 1$ to N

(14) $$\int_R (\sum_{k=1}^{N} A_k\phi_k)\frac{\partial}{\partial t}(\sum_{f=1}^{N} A_f\phi_f)dx = -\int_R (\sum_{k=1}^{N} A_k\phi_k)L(\sum_{f=1}^{N} A_j\phi_j)dx$$

The integral on the right side vanishes if we set $\psi = \sum_{f=1}^{N} A_f\phi_f = \sum_{k=1}^{N} A_k\phi_k$

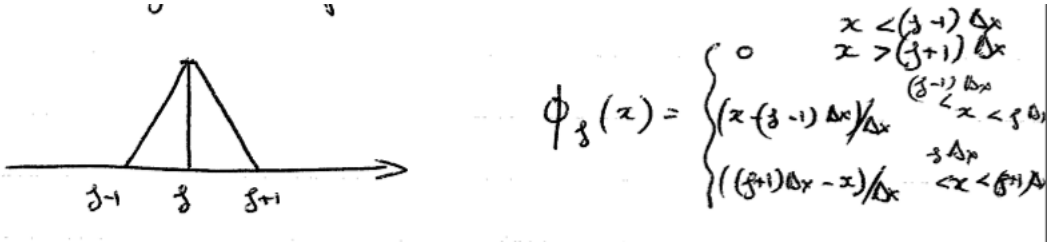
and (14) can be rewritten as:

$$\frac{\lambda}{2}\int_R \frac{\partial}{\partial t}(\sum_{k=1}^{N} A_k\phi_k)^2 dx = 0 \text{ (Energy conservation for the Galerkin Approximation)}$$

7.3 The Advection equation with Finite Elements

$$\frac{\partial \mu}{\partial t} + c\frac{\partial \mu}{\partial x} = 0$$

We use again piecewise linear elements such as



The Galerkin equations is obtained by setting

$$L = c\frac{\partial}{\partial \lambda *}, f(x) = 0$$

(15) $$\sum_{f=1}^{N} \frac{\partial A_f}{\partial t}\int_R \phi_k\phi_f dx + c\sum_{f=1}^{N} A_f \int_R \phi_k \frac{\partial \phi_f}{\partial x} = 0 \quad k = 1...N$$

The resulting equation is the one derived in the introduction:

(15) $$\frac{1}{6}(\frac{\partial A_{k+1}}{\partial t} + 4\frac{\partial A_k}{\partial t} + \frac{\partial A_{k-1}}{\partial t}) = -c\frac{A_{k+1} - A_{k-1}}{2\Delta x}$$

The advection term is the same as if obtained from *eiterated differencing, but the time derivative appears as a weighted everage over time periods*. This greatly increases the accuracy of the solution.

We may now apply the leap frog time differencing scheme.

$$\frac{1}{12\Delta t}(A_{k+1}^{n+1} - A_{k+1}^{n-1} + 4(A_k^{n+1} - A_k^{n-1}) + A_{k-1}^{n+1} - A_{k-1}^{n-1}) = -c\frac{A_{k+1}^n - A_{k-1}^n}{2\Delta x}$$

The stability and phase error of this scheme can be investigated by substituting $A_k^n = Ae^{i(\mu\Delta x n k + \alpha n \Delta t)}*$ (Variation of the Fourier Transform)

Substitution into (17) leads to

(18) $$\sin(\alpha\Delta t) = -\frac{c\Delta t}{\Delta x}\frac{(3\sin(\mu\Delta x)}{2 + \cos(\mu\Delta x))}$$

(for the leap frog $\sin(\alpha\Delta t) = -\frac{c\Delta t}{\Delta x}$)

the solution is stable (neutral solutions with no damping or amplification) if α) is real or $\mid \sin(\alpha\Delta t \mid \leq 1$. To insure stability for all wavelengths, it is necessary to find the maximum magnitude of the RMS of (18).

Maximum when $\mu\Delta x = 120\,\mathrm{deg}$

$$\mid c\frac{\Delta t}{\Delta x} \mid \leq \frac{1}{\sqrt{3}}$$

Which is more restrictive than the lealfrog FD scheme. However it guves even better phase speed than the fourth-order leapfrog scheme.

Finite elements are an interesting alternative to classic FD methods. They offer a high level of slexibility offered for the use of guds* of variable give* shape and flexability and are attractive despite a higher cast* in coupler* tune*. They are popular in the engineering field and ceu*uld* daneri*. For a review, Le Provost* (1985) in O'Brian.

7.4 The Spectral and Pseudo- Spectral method applied to the non-linear advection equatio (Burger)

Before the advert of the fast Fourier Transform (FFT) spectral methods played only a minor role in fluid dynamics because they were far less economical* than grid parts methods.

We want to solve $L(\mu) = \dfrac{\partial \mu}{\partial t} + \mu \dfrac{\partial \mu}{\partial x}$

Let aronee* a cydyic* boundary condition at* the domain $-1 \leq x \leq 1$

We use for the basis function, the trigonometric*

(19) $\phi_j(x) = e^{i\pi f x} \quad (f = -J, J)$

The Galerkin approximation is then

(20) $\phi_m, \Sigma_f A'_f \phi_f > + < \phi_m, (\Sigma A_f \phi_f)(\Sigma A_k \phi'_k) >= 0 \quad m =_J, ...J$

and $\phi_i, \phi_f >= \int_R \phi_i \vec{\phi_j}$

The general form of this set of equations is:

(21) $\sum_f a_{mf} A'_f + \sum_f \sum_k b_{mjk} A_f A_k = 0 \quad m = -J \rightarrow J$

To advance the solutions by one step, we need $(2\sigma+1)^3$ multipliers for the "interaction" term (RMS of (21)) and an inversion* of the matrix components a_{mg}. The basic functions are orthogonal $\Rightarrow$

$< \phi_m, \phi_f >= 2$ when $m = f, = 0$ otherwise.

$< \phi_m, \phi_f \phi'_k >= 2i\pi k$ when $m = j + k$, 0 otherwise.

(21) then reduces to

(22) $A'M + \sum_{f+k=m} i\pi k A_f A_k = 0$

Thus, in practice, the inversion problem does not arise and the operation cout for the interaction terms in only σ^2. The operation count for grid points methods on the other is proportioned to the number of grid points $2\sigma + 1$. The difference is significant and acted as a major deterrent* in the past.

The FFT can improve the speed of the spectral method. There is not approach need to perform Fourier Transform except at $t = 0$ when the initial conditions have to be transformed from physical* to phase space and vise-vera at the end the speed of the transforms does not effect the overall efficiency of the method.

The so-called "pseudo-spectral" method produces similar results to the spectral method by transforming the variables back and forth

between grid point and phase space every few steps* the transform (Pseudo-Spectral) method seaws* the revies* at certain spatrial* grid points and these fields are multiplied together at each point to form the non-linear terms. Then these terms are transformed back to spectral space. The usefulness of this method is enhanced by the existence of efficient transform methods such as the FTTs. this method is essentially a grid point method which uses spectral decomposition techniques to eliminiate the problem of FD in space, namely the phase retordation* of short waves by computing the spatial derivatives by differentiating the individual Fourier components.

We define the grid points by $x_f = \frac{f}{J}$ for $-\sigma < f < \sigma$. With $\phi_k(xf) = e^{i\pi kxf}$ as basis functions we have

(23) $U_f(t) = U)x_j, t) = \sum_{k=-J}^{J-1} A_k(t) e^{i\pi kxj}$

Note that the number of Fourier components matches the number of grid points (Cyclic). Also, if $k = J$ then we have two individual basis functions.

The orthogavelty* relaxation is then

(24)$< \phi_p, \phi_q >= \frac{1}{J} \sum_{f=-\sigma}^{\sigma-1} e^{i\pi(p-q)x_f} = 2$ when $p = q(2J), 0$ otherwise.

The inverse transform of (23) is then:

(25) $A_k(t) = \frac{1}{2S} \sum_{f=-\sigma}^{\sigma-1} U_f(t) e^{-i\pi kxj}$

Starting with grid points values U_f, the coefficients A_k can be computed from (25). The derivative $\frac{\partial U_f}{dx}$ is obtained by a second transform.

(26)$\frac{\partial U_j(t)}{\partial x} = \sum_{k=-J}^{\sigma-1} i\pi k A_k(t) e^{i\pi kx_f}$

Time integration is normally due by finite-defferenciy. An alternative method is to perform the time integration in phase space, but then transform the dependent variables back and forth to grid point space for the evaluation of the non-linear terms.

The final equation which approximates the advection equation is then

(27) $$\frac{\partial U_f(t)}{\partial t} + U_f(t) \sum_{k=J}^{\sigma-1} i\pi k A_k(t) e^{i\pi kxj} = 0$$

Let's now compare the two methods (Spectral and Pseudo).

We first transform (27) (PS) with* an equation for the A_k

(28) $(25)(A'_k(t) + \sum_p + \sum_q + \sum_f i\pi q A_p(t) A_q(t) e^{i\pi(p+q-k)******} =$

0

Using the orthgeniability* of the basis functions, we can rewrite (28) as:

(29) $A'_k(t) + \Sigma i\pi_q A_p A_q + \Sigma_i \pi_q A_p A_q$

$p+q=k \qquad p+q=k+2J$

$+\Sigma i\pi_q A_p A_q = 0$ (Since the orthogonably* is module* 2σ)

$p+q=k-2J$

The last two terms* show a significant difference between (P) and (PS). They are referred to as "alised" terms brought by the finite sampling interval of the discrete Fourier decomposition. This method is clearly much faster sure* the number of operations is $\Sigma + 1 \log_2(2\Sigma+1)$ versus $2\Sigma+1)^2$ for the uiberction* method (P).

There are two basic techniques for removing the alisasity* error introduced in (29).

(1) Aliasing removal by Padding or Tseucation.

The key is to use a discrete transform with M rather than N points where $M \geq \frac{3N}{2}$

(2) Alisaily* removal by phase shifts

Both methods can be escheated to two and three dimensions. For a complete description of spectral methods:

"Spectral Methods in Fluid Dynamics"
by Mauto, Hussaini, Quanteroni, Zaug
Spanger-Verlay 1988