Solution of a Model Elliptic Problem

- Consider the following 1D second–order elliptic problem

$$\mathcal{L}u \equiv \nu u'' - au' + bu = f,$$

where ν, a and b are constant and $f = f(x)$ is a smooth 2π–periodic function.

- For $\nu = 10$, $a = 1$, $b = 5$ and the RHS function $f(x) = e^{\sin(x)}\left[\nu(\cos^2(x) - \sin(x)) - a\cos(x) + b\right]$ the solution is $u(x) = e^{\sin(x)}$

- We are interested in 2π–periodic solutions in the form

$$u_N(x) = \sum_{|k|\leq N} \hat{u}_k e^{ikx}$$

- To be analyzed:
  - Galerkin method
  - Collocation method (two variants)

SOLUTION OF AN ELLIPTIC PROBLEM — GALERKIN APPROACH (I)

- Residual

$$R_N(x) = \mathcal{L}u_N - f = \sum_{|k|\le N} \hat{u}_k \mathcal{L} e^{ikx} - f$$

- Cancellation of the residual in the mean (setting to zero projections on the basis functions $W_n(x) = e^{inx}$)

$$(R_N, W_n) = \sum_{k=-N}^{N} \hat{u}_k (\mathcal{L} e^{ikx}, e^{inx}) - (f, e^{inx}) = 0, \quad n = -N, \ldots, N$$

- Noting that $\mathcal{L} e^{ikx} = (-\nu k^2 - iak + b) e^{ikx} \triangleq \mathcal{G}_k e^{ikx}$ we obtain

$$\sum_{k=-N}^{N} \mathcal{G}_k \hat{u}_k \int_0^{2\pi} e^{i(k-n)}\, dx = \hat{f}_n, \quad n = -N, \ldots, N$$

- Assuming $\mathcal{G}_k \neq 0$, we obtain Galerkin equations for the coefficients $\hat{u}_k$

$$\mathcal{G}_k \hat{u}_k = \hat{f}_k, k = -N, \ldots, N$$

  - The Galerkin equations are decoupled
  - Since u is real, it is necessary to calculate $\hat{u}_k$ for $k \ge 0$ only

SOLUTION OF AN ELLIPTIC PROBLEM — COLLOCATION APPROACH (I)

- Residual (determining the expansion coefficients $\hat{u}_k$)

$$R_N(x) = \mathcal{L} u_N - f = \sum_{|k| \le N} \hat{u}_k \mathcal{L} e^{ikx} - f$$

- Canceling the residual pointwise at the collocation points $x_j,\ j = 1, \dots, M$

$$\sum_{k=-N}^{N} (\mathcal{G}_k \hat{u}_k - \tilde{f}_k) e^{ikx_j} = 0, \quad j = 1, \dots, M$$

where (note the aliasing error) $\tilde{f}_k = \hat{f}_k + \sum_{l \in \mathbb{Z} \setminus \{0\}} \hat{f}_{k+lM}$

- Thus, the collocation equations for the Fourier coefficients

$$\mathcal{G}_k \hat{u}_k = \tilde{f}_k = \hat{f}_k + \sum_{l \in \mathbb{Z} \setminus \{0\}} \hat{f}_{k+lM}, \quad k = -N, \dots, N$$

  - Formally, the Galerkin and collocation methods are distinct
  - In practice, the projection (f, e^{ikx}) is evaluated using FFT and therefore also involves aliasing errors. Therefore, for the present problem, the two approaches are numerically equivalent.

SOLUTION OF AN ELLIPTIC PROBLEM — COLLOCATION APPROACH (II)

- Residual (determining the nodal values $u_N(x_j)$, $j = 1, \dots, M$)

$$R_N(x) = \mathcal{L} u_N - f$$

- Canceling the residual pointwise at the collocation points x_j, $j = 1, \dots, M$

$$[R_N(x_1), \dots, R_N(x_M)]^T = \mathbb{L} U_N - F = (\nu \mathbb{D}_2 - a\mathbb{D}_1 + b\mathbb{I}) U_N - F = 0,$$

where $U_N = [u_N(x_1), \dots, u_N(x_M)]^T$ and $\mathbb{D}_1$ and $\mathbb{D}_2$ are the differentiation matrices.

- Derivation of the differentiation matrices

$$\left.\begin{aligned} u_N^{(p)}(x_j) &= \sum_k (ik)^p \hat{u}_k e^{ikx_j} \\ \hat{u}_k &= \frac{1}{M} \sum_{j=1}^{M} u_N(x_j) e^{-ikx_j} \end{aligned}\right\} \Longrightarrow u_N^{(p)}(x_i) = \sum_{j=1}^{M} d_{ij}^{(p)} u_N(x_j)$$

Solution of an Elliptic Problem — Collocation Approach (III)

- Differentiation Matrices (for even collocation, i.e., $I_N = -N+1, \ldots, N$ and $M = 2N$)

$$d_{ij}^{(1)} = \begin{cases} \frac{1}{2}(-1)^{i+j}\cot(h_{ij}) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad d_{ij}^{(2)} = \begin{cases} \frac{1}{4}(-1)^{i+j}N + \frac{(-1)^{i+j+1}}{2\sin^2(h_{ij})} & \text{if } i \neq j \\ -\frac{(N-1)(N-2)}{12} & \text{if } i = j \end{cases}$$

- Remarks:
  - The differentiation matrices are full (and not so well–conditioned ...), so the system of equations for $u_N(x_j)$ is now coupled
  - For constant coefficient PDEs the present approach is therefore inferior to the first collocation approach where the Fourier coefficients are determined
  - Note the relationship to the banded matrices obtained when approximating differential operators using finite differences
- Question —Derive the above differentiation matrices, also for the case of odd collocation

NYQUIST–SHANNON SAMPLING THEOREM

- If a function $f(x)$ has a Fourier transform $\hat{f}_k = 0$ for $|k| > M$, then it is completely determined by giving the value of the function at a series of points spaced $\Delta x = \frac{1}{2M}$ apart. The values $f_n = f(\frac{n}{2M})$ are called the samples of $f(x)$.
- The minimum sample frequency that allows reconstruction of the original signal, that is $2M$ samples per unit distance, is known as the Nyquist frequency . The time in between samples is called the Nyquist interval .
- The Nyquist–Shannon sampling theorem is a fundamental tenet in the field of information theory (originally formulated by Nyquist in 1928, but formally proved by Shannon only in 1949)

PDEs with variable coefficients — Galerkin Approach (I)

- Consider again the problem $\mathcal{L}u = \nu u'' - au' + bu = f$, but assume now that the coefficient a is a function of space $a = a(x)$
- The following Galerkin equations are obtained for $\hat{u}_k$

$$-\nu k^2 \hat{u}_k - i \sum_{p=-N}^{N} p \hat{a}_{k-p} \hat{u}_p + b \hat{u}_k = \hat{f}_k, \quad k = -N, \dots, N,$$

where $a(x) \cong a_N(x) = \sum_{k=-N}^{N} \hat{a}_k e^{ikx}$ and $f(x) \cong f_N(x) = \sum_{k=-N}^{N} \hat{f}_k e^{ikx}$; Note that

$$\sum_{q=-N}^{N} \hat{a}_q e^{iqx} \sum_{p=-N}^{N} \hat{u}_p e^{ipx} = \sum_{q,p=-N}^{N} \hat{a}_q \hat{u}_p e^{i(q+p)x} = \sum_{k=-2N}^{2N} \sum_{\substack{q,p=-N \\ q+p=k}}^{N} \hat{a}_q \hat{u}_p e^{ikx}$$

$$= \sum_{k=-2N}^{2N} \sum_{p=-N}^{N} \hat{a}_{k-p} \hat{u}_p e^{ikx}, \text{ where } \hat{a}_q, \hat{u}_q \equiv 0, \text{ for } |q| > N$$

- Now the Galerkin equations are coupled (a system of equations has to be solved)

PDEs with variable coefficients — Collocation Approach (I)

- With Fourier coefficients $\hat{u}_k$ as unknowns, the collocation equations

$$-\sum_{k=-N}^{N}(\nu k^2+b)\,\hat{u}_k e^{ikx_j} - a(x_j)\sum_{k=-N}^{N} ik\hat{u}_k e^{ikx_j} = f(x_j), \quad j=1,\dots,M$$

- Approximations of the Fourier coefficients of $a(x)$ and $f(x)$, $\hat{a}^c_k$ and $\hat{f}^c_k$, respectively, are calculated using Discrete Fourier Transform;

$$a(x_j)\sum_{k=-N}^{N} ik\hat{u}_k e^{ikx_j} = \sum_{p=-N}^{N}\hat{a}^c_p e^{ipx_j}\sum_{q=-N}^{N} iq\hat{u}_q e^{iqx_j} =$$

$$i\sum_{k=-N}^{N}\left(\sum_{\substack{q,p=-N\\ q+p=k}}^{N} q\hat{a}^c_p\hat{u}_q + \sum_{\substack{q,p=-N\\ q+p=k+N}}^{N} q\hat{a}^c_p\hat{u}_q + \sum_{\substack{q,p=-N\\ q+p=k-N}}^{N} q\hat{a}^c_p\hat{u}_q\right)e^{ikx_j}$$

$$\triangleq i\sum_{k=-N}^{N}\hat{S}_k e^{ikx_j}$$

- The resulting algebraic system

$$-\nu k^2\hat{u}_k - i\hat{S}_k + b\hat{u}_k = \hat{f}_k, \quad k=-N,\dots,N,$$

PDEs WITH VARIABLE COEFFICIENTS — COLLOCATION APPROACH (II)

- Expressing (hypothetically) $a(x)$ and $f(x)$ with infinite Fourier series we obtain

$$
\begin{aligned}
au'\Big|_{x=x_j} &= i\sum_{k=-N}^{N}\left(\hat{S}_k^{(0)}+\hat{S}_k^{(1)}+\hat{S}_k^{(2)}+\hat{S}_k^{(3)}\right)e^{ikx_j}\\
&= i\sum_{k=-N}^{N}\left(\sum_{\substack{q,p=-N\\ q+p=k}}^{N} q\hat{a}_p^c\hat{u}_q+\sum_{\substack{m=-\infty\\ m\neq 0}}^{\infty}\sum_{\substack{q,p=-N\\ q+p=k}}^{N} q\hat{a}_{p+mM}^c\hat{u}_q\right.\\
&\left.+\sum_{m=-\infty}^{\infty}\sum_{\substack{q,p=-N\\ q+p=k+N}}^{N} q\hat{a}_{p+mM}^c\hat{u}_q+\sum_{m=-\infty}^{\infty}\sum_{\substack{q,p=-N\\ q+p=k-N}}^{N} q\hat{a}_{p+mM}^c\hat{u}_q\right)
\end{aligned}
$$

- The collocation equation

$$
-\nu k^2\hat{u}_k - i\hat{S}_k^{(0)} + i\left(\hat{S}_k^{(1)}+\hat{S}_k^{(2)}+\hat{S}_k^{(3)}\right) + b\hat{u}_k = \hat{f}_k^e + \sum_{\substack{m=-\infty\\ m\neq 0}}^{\infty}\hat{f}_{k+mM}^e, \quad k=-N,\dots,N,
$$

- Note that terms in red are absent in the corresponding Galerkin formulation

PDEs WITH VARIABLE COEFFICIENTS — COLLOCATION APPROACH (III)

- With the nodal values $u(x_j),\ j = 1,\ldots,M$ as unknowns, the collocation equations are (cf. 85)

$$(\nu\mathbb{D}_2 - \mathbb{D}' + b\mathbb{I})U_N = F,$$

  where the matrix $\mathbb{D}' = \left[a(x_j)d_{jk}^{(1)}\right],\ j,k = 1,\ldots,M$

- Again, solution of an algebraic system is required

Fourier transforms in higher dimensions

- Consider a function $u = u(x,y)$ 2π–periodic in both x and y;
  Direct Discrete Fourier Transform

$$\hat{u}_{k_x,k_y} = \frac{1}{2\pi}\int_0^{2\pi}\left[\frac{1}{2\pi}\int_0^{2\pi} u(x,y)e^{-ik_x x}\,dx\right]e^{-ik_y y}\,dy = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} u(x,y)e^{-i\mathbf{k}\cdot\mathbf{r}}\,dxdy,$$

  where $\mathbf{k} = [k_x, k_y]$ is the wavevector and $\mathbf{r} = [x, y]$ is the position vector.

- Representation of a function $u = u(x,y)$ as a double Fourier series

$$u(x,y) = \sum_{k_x,k_y=-N}^{N} \hat{u}_{k_x,k_y} e^{i(k_x x + k_y y)} = \sum_{k_x,k_y=-N}^{N} \hat{u}_{k_x,k_y} e^{i\mathbf{k}\cdot\mathbf{r}}$$

- Fourier transforms in two (and more) dimensions can be efficiently performed using most standard FFT packages.

Nonlinear PDEs

- Replacing the term au' with the nonlinear the term uu' and applying Galerkin or collocation method leads to a system of nonlinear equations that need to be solved using iterative techniques
- From now on we will focus on time–dependent (evolution) PDEs and as a model problem will consider the Burgers equation

$$\begin{cases} \partial_t u + u\partial_x u - \nu\partial_{xx} u = 0 & \text{in } [0,2\pi]\times[0,T] \\ u(x) = u_0(x) & \text{at } t = 0 \end{cases}$$

  Note that steady problems can sometimes be solved as a steady limit of certain time–dependent problems.

- Looking for solution in the form

$$u_N(x,t) = \sum_{k=-N}^{N} \hat{u}_k(t) e^{ikx}$$

  Note that the expansion coefficients $\hat{u}_k(t)$ are now functions of time

- Denote by u_N^n the approximation of u_N at time $t_n = n\Delta t$, $n = 0, 1, \ldots$

Nonlinear PDEs — Galerkin approach (I)

- Time–discretization of the residual $R_N(x,t)$

$$R_N^n = \frac{u_N^{n+1} - u_N^n}{\Delta t} + u_N^n \partial_x u_N^n - \nu \partial_{xx} u_N^{n+1}$$

Points to note:

- explicit treatment of the nonlinear term avoids costly iterations
- implicit treatment of the linear viscous term allows one to mitigate the stability restrictions on the time step Δt
- here using for simplicity first–order accurate explicit/implicit Euler — can do much better than that

- system of equations obtained by applying the Galerkin formalism

$$\left(\frac{1}{\Delta t} + \nu k^2\right) \hat{u}_k^{n+1} = \frac{1}{\Delta t} \hat{u}_k^n - i \sum_{\substack{p,q=-N \\ p+q=k}}^{N} q \hat{u}_p^n \hat{u}_q^n, \quad k = -N, \dots, N$$

Note truncation of higher modes in the nonlinear term.

NONLINEAR PDES — GALERKIN APPROACH (II)

- Evaluation of the nonlinear $i\sum_{\substack{p,q=-N\\ p+q+k}}^{N} q\,\hat{u}_p^n\,\hat{u}_q^n$ term in Fourier space results in a convolution sum which requires $\mathcal{O}(N^2)$ operations – can do better that that?

- Pseudospectral approach —perform differentiation in Fourier space and evaluate products in real space; transition between the two representations is made using FFTs which cost ”only” $\mathcal{O}(N\log(N))$
  Outline of the algorithm:
  1. calculate (using FFT) $u_N^n(x_j)$, $j=1,\dots,M$ from $\hat{u}_k^n$, $k=-N\dots,N$,
  2. calculate (using FFT) $\partial_x u_N^n(x_j)$, $j=1,\dots,M$ from $ik\hat{u}_k^n$, $k=-N\dots,N$,
  3. calculate the product $w_N^n(x_j)=u_N^n(x_j)\partial_x u_N^n(x_j)$, $j=1,\dots,M$
  4. Calculate (using inverse FFT) $\tilde{w}_k^n$, $k=-N\dots,N$ from $w_N^n(x_j)$, $j=1,\dots,M$

- Note that, because of the aliasing phenomenon , the quantity $\tilde{w}_k^n$ is different from $\hat{w}_k^n = i\sum_{\substack{p,q=-N\\ p+q=k}}^{N} q\,\hat{u}_p^n\,\hat{u}_q^n$

NONLINEAR PDES — GALERKIN APPROACH (III)

- Analysis of aliasing in the pseudospectral calculation of the nonlinear term

$$w_N^n(x_j) = \sum_{k=-N}^{N} \tilde{w}_k^n e^{ikx_j}, \quad \text{where} \quad w_N^n(x_j) = u_N^n(x_j)\partial_x u_N^n(x_j)$$

The Discrete Fourier Transform

$$\tilde{w}_k^n = \frac{1}{M}\sum_{j=1}^{M} w_N^n(x_j)e^{-ikx_j} = \frac{1}{M}\sum_{j=1}^{M}\left(\sum_{p=-N}^{N}\hat{u}_p^n e^{ipx_j}\right)\left(\sum_{q=-N}^{N} iq\,\hat{u}_q^n e^{iqx_j}\right)e^{-ikx_j}$$

$$= \frac{1}{M}\sum_{j=1}^{M}\sum_{p,q=-N}^{N} iq\,\hat{u}_p^n\,\hat{u}_q^n\,e^{i(p+q-k)x_j} = \frac{1}{N}\sum_{p,q=-N}^{N} iq\,\hat{u}_p^n\,\hat{u}_q^n\sum_{j=1}^{M} e^{i(p+q-k)x_j}$$

$$= \hat{w}_k^n + i\sum_{\substack{p,q=-N\\ p+q=k+M}}^{N} q\,\hat{u}_p^n\,\hat{u}_q^n + i\sum_{\substack{p,q=-N\\ p+q=k-M}}^{N} q\,\hat{u}_p^n\,\hat{u}_q^n \quad k=-N\ldots,N$$

The term $\hat{w}_k^n$ is the convolution sum obtained in the fully spectral Galerkin approach. The terms in red are the aliasing errors.

- Thus, the pseudospectral Galerkin equations are

$$\left(\frac{1}{\Delta t}+\nu k^2\right)\hat{u}_k^{n+1} = \frac{1}{\Delta t}\hat{u}_k^n - \tilde{w}_k^n, \quad k=-N,\ldots,N$$

NONLINEAR PDES — COLLOCATION APPROACH (I)

- Time–discretization of the residual $R_N(x,t)$

$$R_N^n = \frac{u_N^{n+1} - u_N^n}{\Delta t} + u_N^n \partial_x u_N^n - \nu \partial_{xx} u_N^{n+1}$$

- Canceling the residual at the collocation points

$$\frac{1}{\Delta t}\left[u_N^{n+1}(x_j) - u_N^n(x_j)\right] + u_N^n(x_j)\partial_X u_N^n(x_j) - \nu \partial_{xx} u_N^{n+1}(x_j) = 0 \quad j = 1, \ldots, M$$

- Straightforward calculation shows that the equation for the Fourier coefficients $\hat{\psi}_k$ is the same as in the pseudospectral Galerkin approach. Thus the two methods are numerically equivalent.

- Question —Show equivalence of pseudospectral Galerkin and collocation approaches to a nonlinear PDE

NONLINEAR PDES — ALIASING REMOVAL (I)

- "3/2 rule" —extend the spectrum, and therefore also the number of collocation points, of the quantities involved in the products, so that the aliasing errors arising in pseudospectral calculations are not present.
- Algorithm —consider two 2π–periodic functions

$$a_N(x) = \sum_{k=-N}^{N} \hat{a}_k e^{ikx}, \qquad b_N(x) = \sum_{k=-N}^{N} \hat{b}_k e^{ikx}$$

Calculated in a naive way, the coefficient of the product $w(x) = a(x)b(x)$ are

$$\tilde{w}_k = \hat{w}_k + \sum_{\substack{p,q=-N \\ p+q=k+M}}^{N} \hat{a}_p \hat{b}_q + \sum_{\substack{p,q=-N \\ p+q=k-M}}^{N} \hat{a}_p \hat{b}_q,$$

where $\hat{w}_k$ are the coefficients of the convolution sum that we want to obtain (only)

Nonlinear PDEs — Aliasing Removal (II)

1. Extend the spectra $\hat{a}_k$ and $\hat{b}_k$ to $\hat{a}'_k$ and $\hat{b}'_k$ according to

$$\hat{a}'_k = \begin{cases} \hat{a}_k & \text{if } |k| \le N \\ 0 & \text{if } N < |k| \le N' \end{cases}, \qquad \hat{b}'_k = \begin{cases} \hat{b}_k & \text{if } |k| \le N \\ 0 & \text{if } N < |k| \le N' \end{cases}$$

   The number N' will be determined later.

2. Calculate (via FFT) $a_{N'}$ and $b_{N'}$ in real space on the extended grid $x'_j = \frac{2\pi j}{N'}$, $j = 1, \ldots, N'$

$$a_{N'}(x'_j) = \sum_{k=-N'}^{N'} \hat{a}'_k e^{ikx'_j}, \qquad b_{N'}(x'_j) = \sum_{k=-N'}^{N'} \hat{b}'_k e^{ikx'_j}$$

3. Multiply $a_{N'}(x'_j)$ and $b_{N'}(x'_j)$: $w'(x'_j) = a_{N'}(x'_j)\, b_{N'}(x'_j)$, $j = 1, \ldots, N'$

4. Calculate (via FFT) the Fourier coefficients of $w'(x'_j)$

$$\tilde{w}'_k = \frac{1}{M'} \sum_{j=1}^{M'} w(x'_j) e^{-ikx'_j}, \quad k = -N', \ldots, N', \quad M' = 2N' + 1$$

   Taking the latter quantity for $k = -N, \ldots, N$ gives an expression for the convolution sum free of aliasing errors

NONLINEAR PDES — ALIASING REMOVAL (III)

- Making a suitable choice for N'

$$\tilde{w}'_k = \hat{w}_k + \sum_{\substack{p,q=-N' \\ p+q=k+M'}}^{N'} \hat{a}'_p \hat{b}'_q + \sum_{\substack{p,q=-N' \\ p+q=k-M'}}^{N'} \hat{a}'_p \hat{b}'_q$$

$$= \hat{w}_k + \sum_{\substack{p,q=-N \\ p+q=k+M'}}^{N} \hat{a}_p \hat{b}_q + \sum_{\substack{p,q=-N \\ p+q=k-M'}}^{N} \hat{a}_p \hat{b}_q$$

because $\hat{a}'_p, \hat{b}'_q = 0$ for $|p|, |q| > N$

- The alias terms will vanish, when one of the frequencies p or q appearing in each term of the sum is larger than N. Observe that in the first alias term $q = M' + k - p = 2N' + 1 + k - p$, therefore

$$\min_{|k|,|p| \leq N} (q) = \min_{|k|,|p| \leq N} (2N' + 1 + k - p) = 2N' + 1 - 2N > N$$

Hence $2N' > 3N - 1$. One may take $N' \geq 3N/2$ (the "3/2 rule")

- Analogous argument for the second aliasing error sum.

Hybrid Integration Schemes for ODEs with both linear and nonlinear terms)

- Consider a model ODE problem

$$\mathbf{y}' = \mathbf{r}(\mathbf{y}) + A\mathbf{y}$$

- One would like to use a higher–order ODE integrator with
  - explicit treatment of nonlinear terms
  - implicit treatment of linear terms (with high–order derivatives)
- Combining a three-step Runge–Kutta method with the Crank–Nicholson method results in the following approach:

$$\left(I - \frac{h_{rk}}{2}A\right)\mathbf{y}^{rk+1} = \mathbf{y}^{rk} + \frac{h_{rk}}{2}A\mathbf{y}^{rk} + h_{rk}\beta_{rk}\mathbf{r}(\mathbf{y}^{rk}) + h_{rk}\zeta_{rk}\mathbf{r}(\mathbf{y}^{rk-1}), \quad rk = 1,2,3$$

where

$$h_1 = \frac{8}{15}\Delta t \qquad h_2 = \frac{2}{15}\Delta t \qquad h_3 = \frac{1}{3}\Delta t$$

$$\beta_1 = 1 \qquad \beta_2 = \frac{25}{8} \qquad \beta_3 = \frac{9}{4}$$

$$\zeta_1 = 0 \qquad \zeta_2 = -\frac{17}{8} \qquad \zeta_3 = -\frac{5}{4}$$