PDE Solution via Spectral Methods

Example: Linear eqn w/ variable coefficient

e.g. $$\frac{\partial u}{\partial t} + c(x)\frac{\partial u}{\partial x} = 0$$ Periodic BCs, known ICs

$$u = \sum_{k=0}^{N-1} a_k \phi_k(x)$$

Galerkin: $$\left\langle \left(\sum_{k=0}^{N-1} \frac{da_k}{dt}\phi_k + c(x)\sum_{k=0}^{N-1} a_k \frac{\partial \phi_k}{\partial x} \right) \phi_m \right\rangle = 0$$

$m=0,1,\ldots N-1$

$$\sum_{k=0}^{N-1} \frac{da_k}{dt}\langle \phi_k \phi_m \rangle + \left\langle c(x)\frac{\partial \phi_k}{\partial x}\phi_m \right\rangle \sum_{k=0}^{N-1} a_k = 0$$

$$\frac{da_m}{dt} = -\sum_{k=0}^{N-1} a_k \left\langle c(x)\frac{\partial \phi_k}{\partial x}\phi_m \right\rangle \qquad m=0,1,\ldots N-1$$

Center Difference in Time:

$$a_m^{\ell+1} = a_m^{\ell-1} - 2\Delta t \underbrace{\sum_{k=0}^{N-1} a_k^{\ell}\left\langle c(x)\frac{\partial \phi_k}{\partial x}\phi_m \right\rangle}_{} \qquad m=0,1,\ldots N-1$$

$O(N^2)$ operations/time-step
involves <u>all</u> a_k's for single advance of <u>one</u> a_m

Compare w/ FD: $O(N)$ operations/step

Note: if $c(x) = c$ = constant ... extra overhead goes away ... e.g. Fourier $\Rightarrow \phi_k = e^{ikx}$, $[0, 2\pi]$

$$2\Delta t \sum_{k=0}^{N-1} a_k^{\ell} \left\langle c(x) \frac{\partial \phi_k}{\partial x} \phi_m \right\rangle = 2\Delta t c \sum_{k=0}^{N-1} a_k^{\ell} \left\langle ik e^{ikx} e^{-imx} \right\rangle$$

$$= 2\Delta t c\, i m a_m^{\ell}$$

$$a_m^{\ell+1} = a_m^{\ell-1} - \underbrace{2\Delta t c\, i m a_m^{\ell}}_{O(N) \text{ effort just like FD}} \qquad m = 0, 1, \dots N-1$$

Conclude: $c(x)$ "messes-up" orthogonality integrals causing the method to be impractical

What if we use our interpolating polynomial approach with collocation ??

$$u(x) = \sum_{k=0}^{N-1} u_k \phi_k(x) \Rightarrow u_j = u(x_j)$$

$$\frac{du_j}{dt} + c(x_j) \sum_{k=0}^{N-1} u_k \frac{\partial \phi_k}{\partial x}(x_j) = 0 \quad j = 0, 1, \dots N-1$$

$$u_j^{\ell+1} = u_j^{\ell-1} - 2\Delta t c_j \underbrace{\sum_{k=0}^{N-1} u_k^{\ell} \frac{\partial \phi_k}{\partial x}(x_j)}_{} \quad j = 0, 1, \dots N-1$$

$O(N^2)$ / Time-step $\Rightarrow$ same problem

essentially an N-pt approximation to $\frac{\partial u}{\partial x}_j$

Conclude: Neither Galerkin nor Collocation is very efficient (gain in spatial accuracy offset by computational costs!)

In Galerkin ... orthogonality lost due to $c(x)$, but no difficulty with derivative per se (e.g. derivative of trig function is another trig function)

In Collocation ... derivative evaluation involves an N-pt formula (all gridvalues of u) but no difficulty with $c(x)$ ... i.e. that coefficient varies with position

Would like to combine the best features of both

i.e. $$u_j^{\ell+1} = u_j^{\ell-1} - 2\Delta t\, c_j \underbrace{\frac{\partial u}{\partial x}^{\ell}_j}_{} \qquad j = 0, 1, \ldots N-1$$

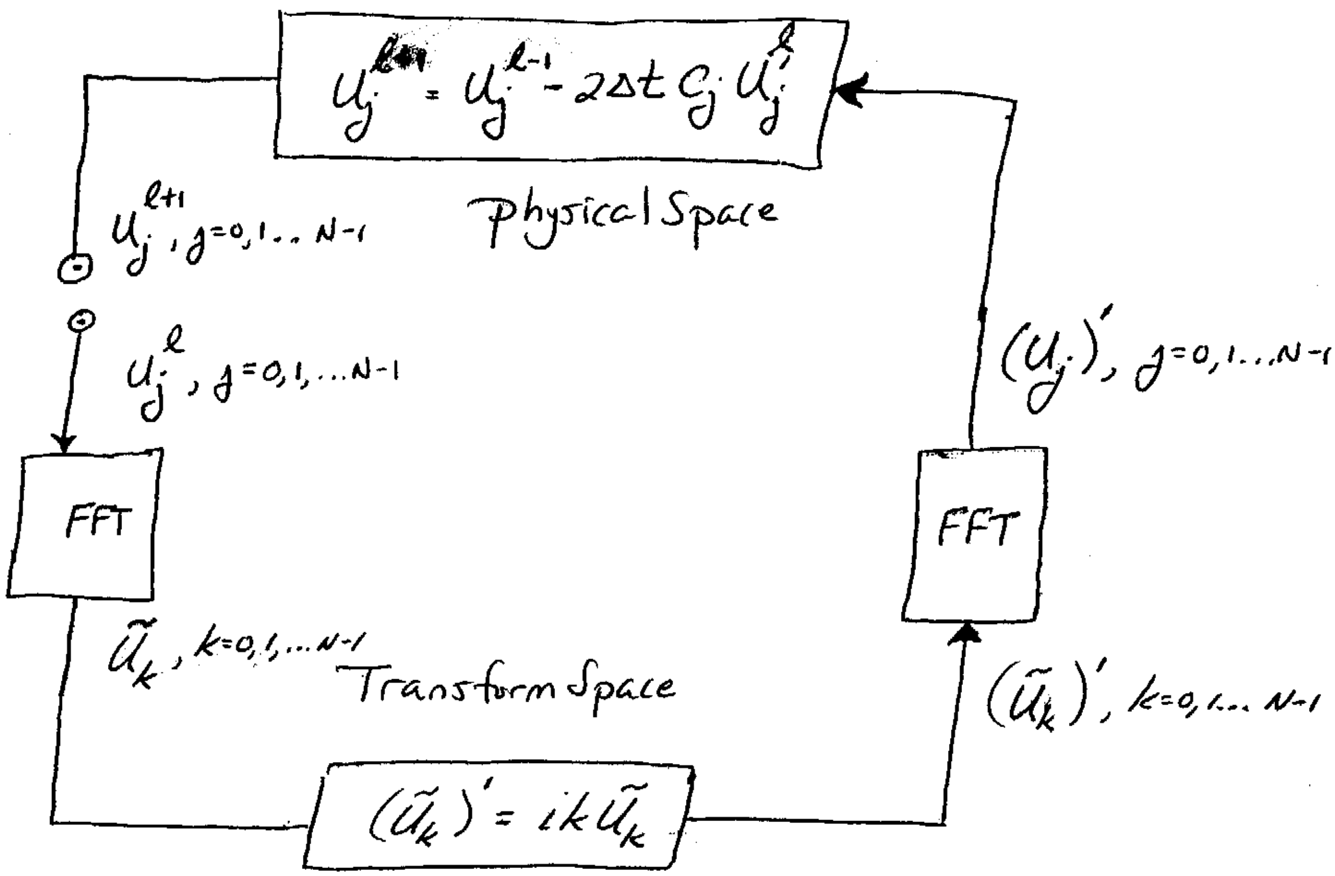
where this can be evaluated efficiently ... hopefully $O(N)$/step

Can accomplish this by appealing to our collocation derivative concepts but use a fast transform technique to compute $\Rightarrow O(N \log N)$

"pseudospectral method"

So we replace rate-limiting matrix multiply with FFT $\Rightarrow$ limited to Fourier and Chebyshev bases

Algorithm becomes something like:



We "jump back and forth" between the grid point representation of $u(x)$ (in order to multiply by $c(x)$) and the spectral coefficients (in order to differentiate $u(x)$)

One potential problem with this approach is "aliasing" errors

- If computational grid supports wavenumbers

$$k \in \left[-\frac{N}{2}, \frac{N}{2}\right]$$

then higher wavenumbers (outside the basis set) appear in the computations as if they were wavenumbers

$$k_A = k \pm mN \quad \text{where } k_A \in \left[-\frac{N}{2}, \frac{N}{2}\right]$$

i.e. wavenumbers outside the truncation are aliased to wavenumbers within

- This is a re-statement of the error occurring from interpolation rather than truncation

$$\text{Recall:} \quad \tilde{u}_k = \hat{u}_k + \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} \hat{u}_{k \pm mN}$$

e.g. $\cos((k+mN)x) = \cos(kx)\cos(mNx) - \sin(kx)\sin(mNx)$

$\sin((k+mN)x) = \sin(kx)\cos(mNx) + \cos(kx)\sin(mNx)$

Now if $x_i = \frac{2\pi i}{N}, \ i = 0, 1, \ldots N-1$

$$\left.\begin{aligned} \cos(mNx_i) &= \cos\left(mN\frac{2\pi i}{N}\right) = \cos(2\pi m i) = 1 \\ \sin(mNx_i) &= \sin(2\pi m i) = 0 \end{aligned}\right\} \text{all } m, i \text{ integer}$$

So $\cos(kx) = \cos((k+mN)x) = \cos((-k+mN)x)$

$\sin(kx) = \sin((k+mN)x) = -\sin((-k+mN)x)$

at the N grid points ... so we have linearly independent functions that are point-by-point equal on the grid $\Rightarrow$ aliasing error

- Can generate higher wavenumbers outside the basis set truncation through variable coefficients, nonlinear terms, etc.; these can lead to numerical instabilities (or inaccuracies)

e.g.
$$u\frac{\partial u}{\partial x} = \left(\sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1} \tilde{u}_l e^{ilx}\right)\left(\sum_{m=-\frac{N}{2}}^{N/2-1} im\,\tilde{u}_m e^{imx}\right)$$

$$= \sum_{l=-\frac{N}{2}}^{N/2-1}\sum_{m=-\frac{N}{2}}^{N/2-1} im\,\tilde{u}_l\tilde{u}_m e^{i(l+m)x}$$

$$= \sum_{k=-N}^{N-1} a_k e^{ikx}$$

Nonlinear interaction produces wavenumbers on the range $k \in [-N, N]$ which will get aliased onto $k \in [-\frac{N}{2}, \frac{N}{2}]$

- Recall... highest frequency on a discrete mesh "$2\Delta x$"-waves ()

Also recall Nyquist Sampling criterion... need to sample at twice the highest frequency for perfect reconstruction

- Obvious remedy... apply spatial filter so as to eliminate all waves with wavelengths less than $4\Delta x$ ... i.e. eliminate/remove all wavenumbers with $|k| > \frac{N}{4} \Rightarrow \frac{N}{4} < |k| \leq \frac{N}{2}$

Turns out can do better than this... obvious remedy is wasteful

"3/2 Rule": accurate, alias-free computation of a total of N wavenumbers for a quadratically nonlinear interaction, must use $\frac{3}{2} N$ samples

Also known as "padding" because we add $N/2$ zeros to the spectral coefficients before taking coefficient to grid transforms

- Alternately ... If use a total of N grid points, then only $2/3\,N$ wavenumbers should be retained in the truncation $\Rightarrow$ only necessary to filter waves w/ wavelengths between $2\Delta x$ and $3\Delta x$ (not $4\Delta x$) $\Rightarrow$ "$2/3$ Rule"

Example: $u \cdot v$

Transform $\tilde{u}_k, \tilde{v}_k$ on N pts to u_j, v_j on M pts

Product $u_j v_j$ on M pts

Transform Back $(\widetilde{uv})_k$ on M pts

Truncate $(\widetilde{uv})_k$ to N (padding)

$$y_j = \frac{2\pi j}{M} \qquad j = 0, 1 \ldots M-1$$

$$u_j = \sum_{k=-M/2}^{M/2-1} \tilde{\tilde{u}}_k e^{iky_j}, \quad v_j = \sum_{k=-M/2}^{M/2-1} \tilde{\tilde{v}}_k e^{iky_j}$$

$$\text{where} \quad \tilde{\tilde{u}}_k = \begin{cases} \tilde{u}_k & k \le |\frac{N}{2}| \\ 0 & \text{otherwise} \end{cases},$$

$$\tilde{\tilde{V}}_k = \begin{cases} \tilde{V}_k & k \lesssim |\frac{N}{2}| \\ 0 & \text{otherwise} \end{cases}$$

then $\quad W_j = U_j \cdot V_j \qquad j = 0, 1, \ldots M-1$

and $\quad \tilde{W}_k = \frac{1}{M} \sum_{j=0}^{M-1} W_j e^{-iky_j} \qquad k = -\frac{M}{2}, \cdots \frac{M}{2} - 1$

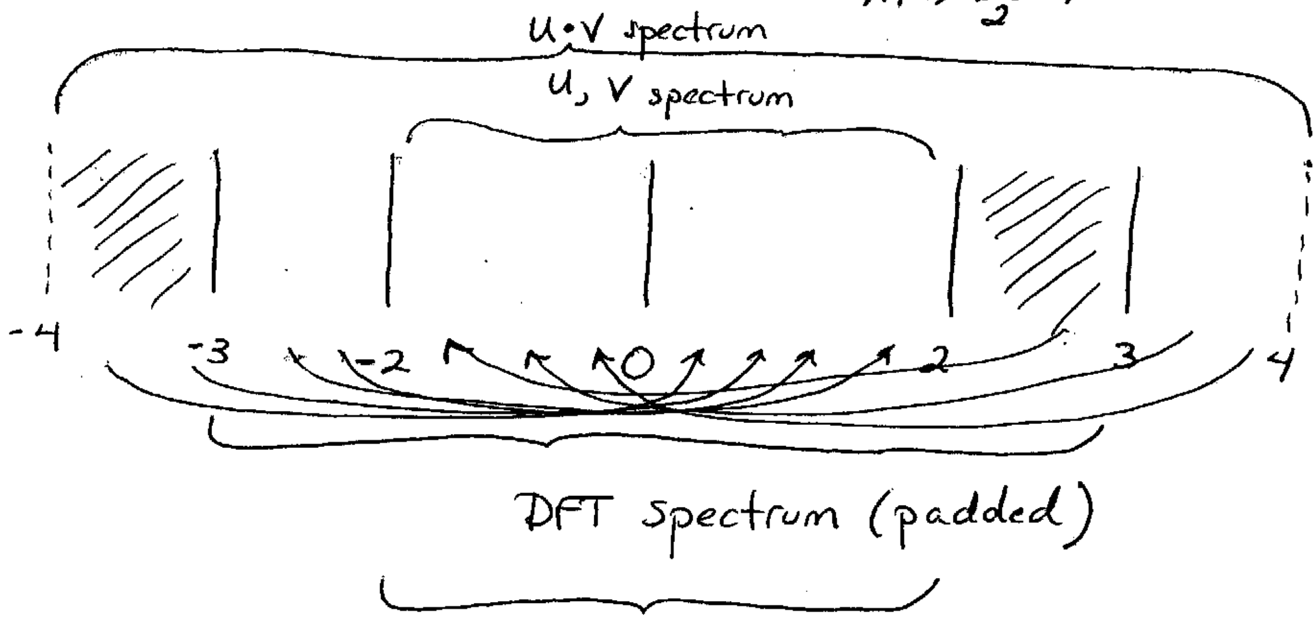
But $\quad \tilde{W}_k = \hat{W}_k + \underbrace{\sum_{m+n=k+M} \hat{U}_m \hat{V}_n}$

Choose M so that this term vanishes for $k \in [-\frac{N}{2}, \frac{N}{2}]$
i.e. when $|m|, |n| > N/2$

$\Rightarrow$ worst case $\; -\frac{N}{2} - \frac{N}{2} \leq \frac{N}{2} - 1 - M$

$$M \geq \frac{3N}{2} - 1$$

e.g.



DFT (aliased) get contamination from the u·v spectrum outside $[-2, 2]$ on $[-2, 2]$

Another Strategy: Dealiasing by phase shift

Aliased:

$$\begin{array}{l} \tilde{u}_k \Rightarrow u_j \longrightarrow \boxtimes\, w_j \Rightarrow \tilde{w}_k = \underbrace{\sum_{m+n=k} \tilde{u}_m \tilde{v}_n}_{\text{Correct}} + \underbrace{\sum_{m+n=k\pm N} \tilde{u}_m \tilde{v}_n}_{\text{Error}} \\ \tilde{v}_k \Rightarrow v_j \nearrow \end{array}$$

Shifted:

"Shifted Transforms" $\stackrel{\Delta}{\Longrightarrow}$

$$\left\{ \begin{aligned} u_j' &= \sum_{k=-\frac{N}{2}}^{N/2-1} \tilde{u}_k e^{ik(x_j+\Delta)} = \sum (\tilde{u}_k e^{ik\Delta}) e^{ikx_j} \\ \tilde{u}_k' &= \frac{1}{N} \sum_{j=0}^{N-1} u_j' e^{-ik(x_j+\Delta)} = e^{-ik\Delta} \frac{1}{N} \sum_{j=0}^{N-1} u_j' e^{-ikx_j} \\ e^{ik\Delta} \tilde{u}_k' &= \frac{1}{N} \sum_{j=0}^{N-1} u_j' e^{-ikx_j} \end{aligned} \right.$$

$$\begin{array}{l} \tilde{u}_k \stackrel{\Delta}{\Rightarrow} u_j' \longrightarrow \boxtimes\, w_j' \stackrel{\Delta}{\Rightarrow} \tilde{w}_k' = \sum_{m+n=k} \tilde{u}_m \tilde{v}_n + e^{\overbrace{\pm iN\Delta}^{\pi}} \sum_{m+n=k\pm N} \tilde{u} \tilde{v}_m \\ \tilde{v}_k \stackrel{\Delta}{\Rightarrow} v_j' \nearrow \end{array}$$

(i.e. $N\Delta = \pi$)

$e^{\pm i\pi} = -1$

dealiased $\tilde{w}_k = \dfrac{\tilde{w}_k + \tilde{w}_k'}{2}$

Cost turns out to be greater than padding technique so not used that often

Rule of thumb: dealiasing not really necessary for well-resolved simulations; may be helpful when simulation is marginally resolved (this can become important in large-scale problems)

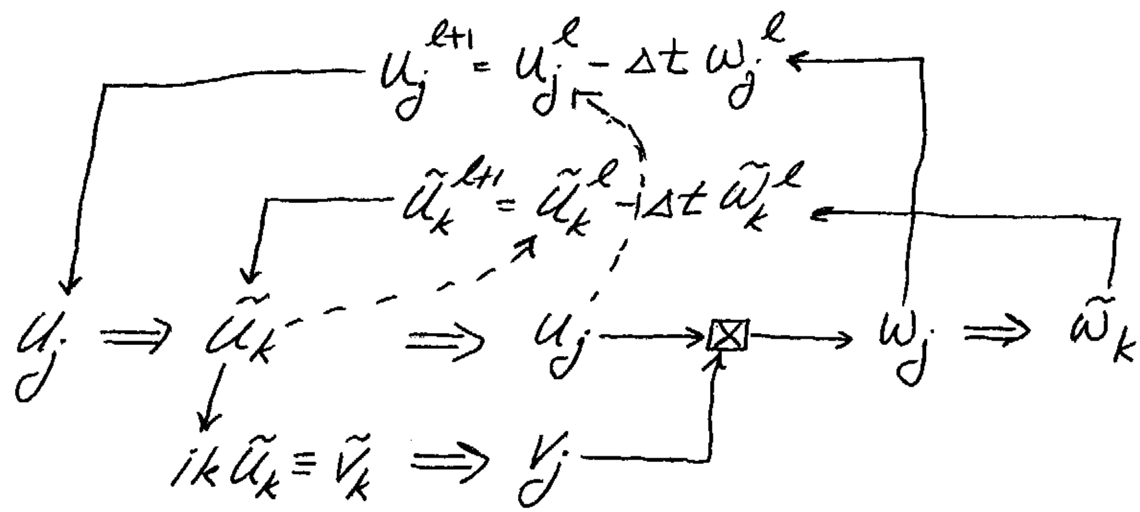
- Look at some possible strategies (and their relationship) for: $\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0$

$$
\begin{tikzcd}
 & & \tilde{u}_k^{\ell+1} = \tilde{u}_k^{\ell} - \Delta t\, \tilde{w}_k^{\ell} \arrow[dll] & & \\
\tilde{u}_k \arrow[r, Rightarrow] \arrow[d] & u_j \arrow[r] & \boxtimes \arrow[r] & w_j \arrow[r, Rightarrow] & \tilde{w}_k \arrow[ull] \\
ik\tilde{u}_k \equiv \tilde{v}_k \arrow[r, Rightarrow] & v_j \arrow[ur] & & &
\end{tikzcd}
$$

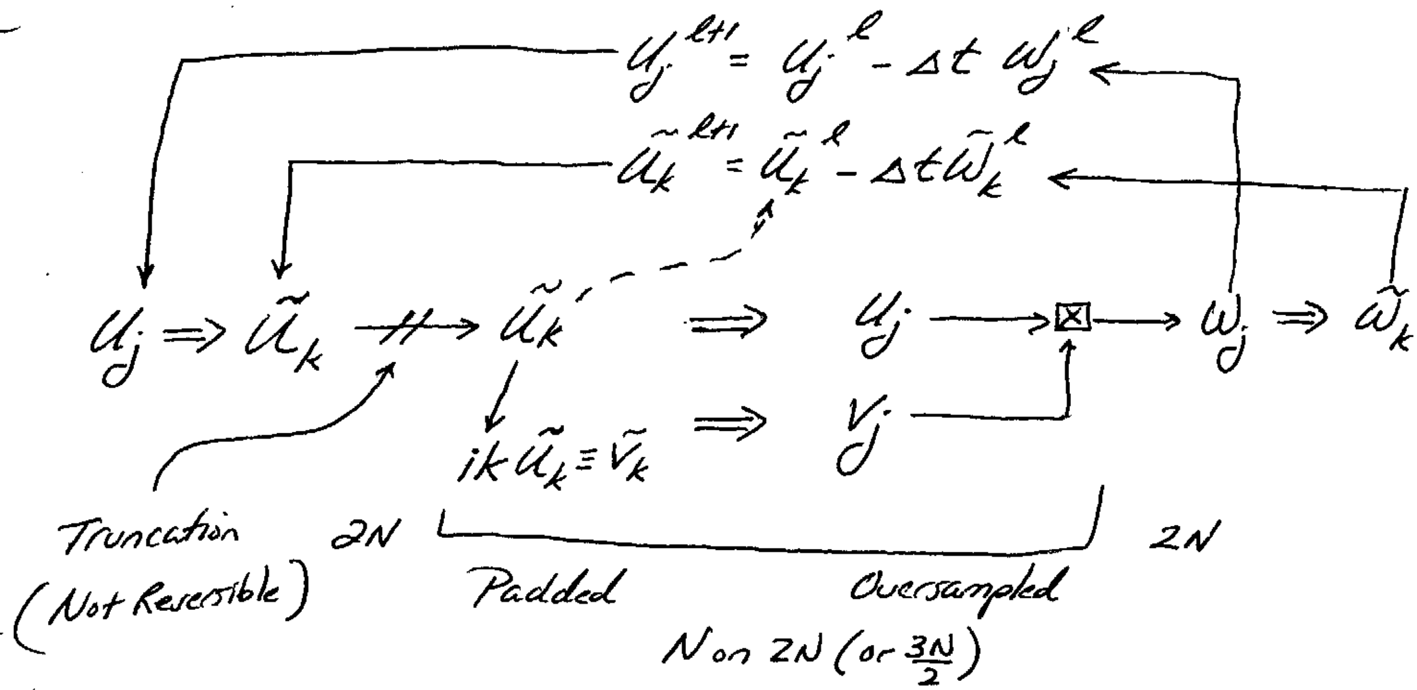
"Pseudospectral"

$$
\begin{tikzcd}
 & & & u_j^{\ell+1} = u_j^{\ell} - \Delta t\, w_j^{\ell} \arrow[dlll] & & \\
u_j \arrow[r, Rightarrow] & \tilde{u}_k \arrow[r, Rightarrow] \arrow[d] & u_j \arrow[r] & \boxtimes \arrow[r] & w_j \arrow[r, Rightarrow] \arrow[ul] & \tilde{w}_k \\
 & ik\tilde{u}_k \equiv \tilde{v}_k \arrow[r, Rightarrow] & v_j \arrow[ur] & & &
\end{tikzcd}
$$

"Collocation"



Collocation = Pseudospectral (aliased)
(reversible!!)



Pseudospectral with dealiasing
(Collocation)

Rules of Thumb on Accuracy:

To achieve 5% accuracy for wave-like sol'n

20 samples/wavelength FD $O(\Delta x^2)$
(40 samples/wavelength $\Rightarrow$ 1% error)

10 samples/wavelength FD $O(\Delta x^4)$
(15 samples/wavelength $\Rightarrow$ 1% error)

3.5 polys/wavelength Chebyshev pseudospectral
(3.5 polys/wavelength $\Rightarrow$ 1% error)

If Δx is smallest scale to resolve on $x \in [-1,1]$
then an estimate of number of wavelengths

$$M \approx \frac{1}{\pi \Delta x}$$

Errors are based on smooth solutions (i.e. no discontinuities on the interior or boundary)

$\Rightarrow N \approx 3/\sqrt{\epsilon}$ polynomials for 1% accuracy for boundary layer width ϵ at $x = \pm 1$

Simple Time-stepping

e.g. $\dfrac{d\hat{u}_k}{dt} + \nu k^2 \hat{u}_k = 0 \qquad -\dfrac{N}{2} \le k \le \dfrac{N}{2} - 1$

$\Delta x = \dfrac{2\pi}{N}$

Euler: $\dfrac{\hat{u}_k^{\,\ell+1} - \hat{u}_k^{\,\ell}}{\Delta t} = -\nu k^2 \hat{u}_k^{\,\ell}$

let $\hat{u}_k^{\,\ell} = \lambda \hat{u}_k^{\,\ell-1}$

$$\frac{\lambda - 1}{\Delta t} = -\nu k^2 \Rightarrow \lambda = 1 - \nu k^2 \Delta t$$

want $-1 \le \lambda \le 1$ for stability

$$1 - \nu k^2 \Delta t \geq -1$$

$$2 \geq \nu k^2 \Delta t \Rightarrow \text{worst case } k^2 = \left(\frac{N}{2}\right)^2 = \left(\frac{\pi}{\Delta x}\right)^2$$

$$2 \geq \frac{\nu \pi^2 \Delta t}{\Delta x^2}$$

so $\boxed{\dfrac{\nu \Delta t}{\Delta x^2} \le \dfrac{2}{\pi^2}}$ (Compare w/ FD: $\dfrac{\nu \Delta t}{\Delta x^2} \le 1/2$)

2-level Implicit

$$\frac{\hat{u}_k^{\ell+1} - \hat{u}_k^{\ell}}{\Delta t} = -\nu k^2\left(\theta \hat{u}_k^{\ell+1} + (1-\theta)\hat{u}_k^{\ell}\right)$$

$$\lambda - 1 = -\nu k^2 \Delta t\left(\theta\lambda + (1-\theta)\right)$$

$$\lambda\left[1 + \theta \nu k^2 \Delta t\right] = 1 - (1-\theta)\nu k^2 \Delta t$$

$$\lambda = \frac{1 - (1-\theta)\nu k^2 \Delta t}{1 + \theta \nu k^2 \Delta t} \Rightarrow \text{always} < 1$$

can become < -1
(unstable)

$\theta > 1/2$ unconditionally stable!

Conditionally stable if: $\dfrac{1-(1-\theta)\nu k^2\Delta t}{1+\theta\nu k^2 \Delta t} > -1$

$$1 - (1-\theta)\nu k^2 \Delta t > -1 - \theta \nu k^2 \Delta t$$

$$2 > (1-2\theta)\nu k^2 \Delta t$$

$$\frac{2}{1-2\theta} > \nu k^2 \Delta t \Longrightarrow \underline{\underline{\frac{\nu \Delta t}{\Delta x^2} < \frac{2}{\pi^2(1-2\theta)}}}$$

Prototype Problem: Burger's Equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \nu\frac{\partial^2 u}{\partial x^2} = 0$$

$0 \le x \le 2\pi$
IC's
BC's given
ν = positive constant

Fourier Galerkin:

$$u^N(x,t) = \sum_{k=-\frac{N}{2}}^{N/2-1} \hat{u}_k(t)\, e^{ikx} \qquad \text{Fourier}$$

$$\int_0^{2\pi} \left(\frac{\partial u^N}{dt} + u^N \frac{\partial u^N}{\partial x} - \nu \frac{\partial^2 u^N}{\partial x^2} \right) e^{-ikx}\, dx = 0 \qquad \text{Galerkin}$$

$$\frac{\partial \hat{u}_k}{\partial t} + \underbrace{\widehat{\left(u\frac{\partial u}{\partial x}\right)}_k} + k^2 \nu \hat{u}_k = 0, \quad k = -\frac{N}{2} \ldots \frac{N}{2} - 1$$

Product Term: $$\widehat{(uv)}_k = \frac{1}{2\pi}\int_0^{2\pi} uv\, e^{-ikx}\, dx$$

$$= \sum_{p+q=k} \hat{u}_p \hat{v}_q \qquad \text{Convolution Sum}$$

$$\left(\frac{1}{2\pi} \int_0^{2\pi} \underbrace{\left(\sum_p u_p e^{ipx} \sum_q v_q e^{iqx}\right)}_{\sum\sum u_p v_q e^{i(p+q)x}} e^{-ikx}\, dx \right)$$

Implied creation of info outside the basis

Fourier Collocation:

- carried out in x-domain
- retain values of u at $x_j = \frac{2\pi j}{N}$, $j=0,1\dots N-1$

$$\frac{\partial u^N}{\partial t} + \underbrace{u^N \frac{\partial u^N}{\partial x}}_{u^N D_N u^N} - \underbrace{\nu \frac{\partial^2 u^N}{\partial x^2}}_{\text{Fourier Collocation Derivative } D_N^2 u^N \text{ (Full matrix)}}\Big|_{x_j} = 0 \quad j=0,1,\dots N-1$$

Pseudospectral Transform

$$\left.\begin{aligned} w(x) &= u(x)\,v(x) \\ \hat{w}_k &= \sum_{m+n=k} \hat{u}_m \hat{v}_n \end{aligned}\right\} \text{ convolution}$$

or

$$u_j = \sum_{k=-N/2}^{N/2-1} \hat{u}_k e^{ikx_j}$$

$$v_j = \sum_{k=-N/2}^{N/2+1} \hat{v}_k e^{ikx_j}$$

$$\Rightarrow \quad w_j = u_j v_j$$

$$\hat{w}_k = \frac{1}{N} \sum_{j=0}^{N-1} w_j e^{-ikx_j} \qquad k = -\frac{N}{2},\dots \frac{N}{2}-1$$

produces aliasing error:

$$\hat{w}_k = \hat{w}_k + \underbrace{\sum_{m+n=k\pm N} \hat{u}_m \hat{v}_n}_{\text{Error}}$$

So we have (pseudospectral)

$$\frac{d\hat{u}_k}{dt} + \sum_{m+n=k} \hat{u}_m \hat{v}_n + \sum_{m+n=k\pm N} \hat{u}_m \hat{v}_n + \nu k^2 \hat{u}_k = 0 \quad k = -\tfrac{N}{2}, \ldots \tfrac{N}{2}-1$$

Same as Collocation under DFT

$$\left(\frac{\partial u^N}{\partial t} + u^N v^N - \nu \frac{\partial^2 u^N}{dx^2} \right)\Bigg|_{x=x_j} = 0 \quad \text{all } x_j$$

$$\frac{d\hat{u}_k}{dt} + \sum_{m+n=k} \hat{u}_m \hat{v}_n + \underbrace{\sum_{m+n=k\pm N} \hat{u}_m \hat{v}_n}_{\text{eliminate aliasing error}} + \nu k^2 \hat{u}_k = 0 \qquad \text{same!!}$$

IF can dealias Collocation, then have Galerkin