Convergence of Finite Difference Methods for BVPs

A. Donev, Courant

We showed last class how to use the FD method to convert the BVP

$$u''(x) = f(x) \qquad 0 < x < 1$$

$$u(0) = \alpha \qquad u(1) = \beta$$

to the system $AU = F$

$$\frac{1}{h^2}\begin{bmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix} \qquad F = \begin{bmatrix} f_1 - \alpha/h^2 \\ f_2 \\ \vdots \\ f_m - \beta/h^2 \end{bmatrix}$$

Let the true solution evaluated <u>pointwise</u> be

$$\hat{u} = [u(x_1), \ldots, u(x_m)]$$

Then the <u>global error</u> is

$$E = u - \hat{u}$$

<u>Question</u>: Does $\|E\| \to 0$ as $h \to 0$???
$m \to \infty$

Put the exact solution in the finite difference to get the <u>local truncation error</u> (LTE)

$$\tau_j = \frac{1}{h^2}\left[u(x_{j-1}) - 2u(x_j) + u(x_j)\right] - \underbrace{f(x_j)}_{u''(x_j) \text{ from PDE}}$$

$$\tau_j = \frac{h^2}{12} u''''(x_j) + O(h^4)$$ 3

$$\|\tau\| = O(h^2)$$

$$\Rightarrow \quad \tau = A\hat{U} - F = O(h^2) \Rightarrow$$

$$\left.\begin{cases} A\hat{u} = F + \tau \\ AU = F \end{cases}\right\} \text{take difference}$$

$$\Rightarrow \quad \boxed{AE = -\tau} + \begin{cases} E_0 = 0 \\ E_m = 0 \end{cases}$$

Back to this over & over again:

(homogeneous)

$\boxed{\text{Error satisfies same equation as solution but with LTE as a source term on the R.H.S.}}$

Now we can *guess* that the global error is also $O(h^2)$ = **second order accurate** by the following argument:

Since $A\bar{E} = -\bar{\tau}$ *approximates*

$$e''(x) = -\bar{\tau}(x), \quad e(0) = e(1) = 0$$

$$e''(x) = -\frac{h^2}{12} u^{(4)}(x) + O(h^4)$$

$$\Rightarrow \quad e(x) \approx -\frac{h^2}{12} u''(x) + \text{boundary terms } O(h^2)$$

$$\boxed{e(x) \approx -\frac{h^2}{12} f(x)} \quad \text{continuum estmate}$$

But this is not a discrete proof!

Discretely, $E = -A^{-1} \tau \Rightarrow$

$$\|E\| \leq \|A^{-1}\| \, \|\tau\|$$

so if we want both $\|E\|$ and $\|\tau\|$ to be $O(h^2)$ independent of h then

we want

$$\|A^{-1}\| \leq C \quad \text{for all small } h$$

A method to solve a linear BVP is ***stable*** if A^{-1} exists and

$$\|A^{-1}\| \leq C \quad \text{for all } h < h_0$$

A method is consistent if

$$\|\tau\| \to 0 \quad \text{as} \quad h \to 0$$

$$\boxed{\text{stability} + \text{consistency} \Rightarrow \text{convergence}}$$

$$\boxed{O(h^p)\ \text{LTE} + \text{stability} \Rightarrow O(h^p)\ \text{error}}$$

So now we just need to prove

$$\|A^{-1}\| \leq C$$

and we have proven second order convergence.

For this we need to pick our norm!

And now all norms are NOT equivalent as $h \to 0$

because infinite dimensional

Stability in L_2

$$\|A\|_2 = \rho(A) = \max_p |\lambda_p|$$

spectral radius

Here A is symmetric so all λ's are real, and same for inverse, so

$$\boxed{\|A^{-1}\|_2 = \left(\min_p |\lambda_p|\right)^{-1}}$$

We need to find the eigenvalues.

Now go back to the continuum PDE and recall that the eigenfunctions are just sin functions. So guess (anzatz)

$$u_j^p = \sin(p\pi j h)$$

p'th eigenvector

Plug into $Au^p = \lambda^p u^p$ and get

$$\lambda_p = \frac{2}{h^2}\left[\cos(p\pi h) - 1\right]$$

Observe that for small waveindex (wavenumber) p we have

$$\lambda_p \simeq -\pi^2 p^2 + \underbrace{\frac{1}{12}\pi^4 p^4 h^2}$$

also shows second-order convergence

Now smallest eigenvalue corresponds to the longest wavelength (9)

$$\lambda_1 = \frac{2}{h^2}\left(\cos(\pi h) - 1\right) \simeq -\pi^2 + O(h^2)$$
$$= \text{const}$$

$$\Rightarrow \quad \|E\| \le \|A^{-1}\| \, \|\tau\| \simeq \frac{\|\tau\|}{\pi^2}$$

and we have second order accuracy.

Also recall $\tau \simeq \frac{h^2}{12} u^{(4)}(x) = \frac{h^2}{12} f^{(2)}(x)$

So the smoother $f(x)$ is, the less points we need, which makes sense physically.

Max norm stability in L_∞

If we just use finite-dimensional linear algebra, we would bound

$$\|E\|_\infty \le \frac{1}{\sqrt{h}} \|E\|_2 \Rightarrow$$

$$\|E\|_\infty = O(h^{3/2})$$

But this is overly pessmistic, in fact,

$\|E\|_\infty = O(h^2)$ as well.

So we need to show

$$\|A^{-1}\|_\infty \le C$$

For this we go back to Green's functions

What is the j'th column of A^{-1}? 12

Remember

$$\tilde{G}_j = A^{-1} e_j$$

j'th column of A^{-1}

zeros but 1 in entry j

$$\Rightarrow \quad \boxed{A \tilde{G}_j = e_j}$$

But A is a discretization of <u>Laplacian</u> with homogeneous BCs (Dirichlet), so this is a discretization of

$$\begin{cases} u''(x) = \delta(x - x_j) \cdot \underset{\substack{\uparrow \\ \text{grid} \\ \text{spacing}}}{h} \\ u(0) = u(1) = 0 \end{cases}$$ (13)

That is

$\boxed{\tilde{G}_j \text{ is a discrete Green's function}}$

and tells us how an error (local truncation error, roundoff, etc.) at node j translates to error at other nodes.

We can compute $\tilde{G}_j$ explicitly in this simple case:

$$(\tilde{G}_j)_i = h \begin{cases} (x_j - 1)\, x_i \,, & i = 1, \dots, j \\ (x_i - 1)\, x_j \,, & i = j, j+1, \dots, m \end{cases}$$ (14)

Which is exactly what we expect,

$$(\tilde{G}_j)_i = \boxed{h\, G(x_i, x_j) = A^{-1}_{ij}}$$

Actual Green's function of PDE

<u>Note</u>: Inhomogeneous Dirichlet BCs can be easily handled, see 2.11 in LeVeque.

Now go back to

$$AU = F \Rightarrow AE = \tau$$

We need to bound $\|A^{-1}\|$

$$\|A^{-1}\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{m} (\tilde{G}_j)_i$$

Note that $|(\tilde{G}_j)_i| \le h$ $\Rightarrow$

since $x(1-x) \le 1$

$$\|A^{-1}\|_\infty \le mh = 1$$

so indeed we have <u>stability</u>!

Observe that $\bar{\tau}_j = O(h^2)$ and so if we only made a localized error $\bar{\tau}_j = O(h^2)$ and a much smaller error at other points, then the global error

$$E_i = h \cdot O(h^2) \cdot G(x_i ; x_j)$$

$$E_i = O(h^3)$$

This shows an important fact:

We can make an error of order p at a few points (e.g., boundaries) and still get global order $q > p$, sometimes

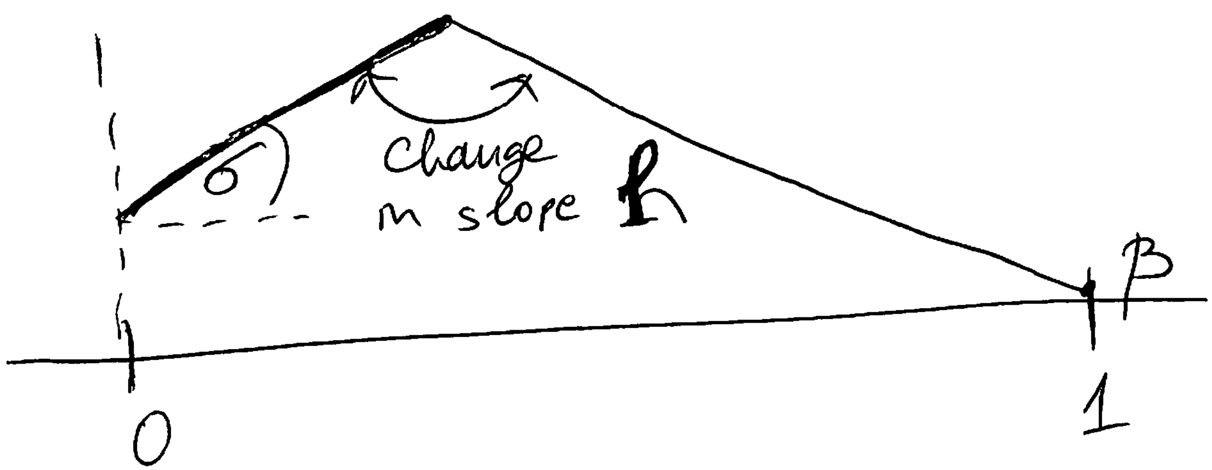
But <u>not always</u>.

E.g. Consider Neumann BCs

$$u'(0) = \sigma \qquad u(1) = \beta$$

(see 2.12 in Le Veque)

What is Green's function now?



If we introduce an error $O(h)$ at the boundary when we impose BC e.g. $\frac{u_1 - u_0}{h} = \sigma$

$$\frac{1}{h^2}\begin{bmatrix} -h & h & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & \ldots & \end{bmatrix}\begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \varepsilon_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} O(h) \\ O(h^2) \\ O(h^2) \\ \ldots \end{bmatrix}$$

Now we need the first column of A^{-1} – this corresponds to the BVP

$$\begin{cases} u''(x) = 0 \\ u'(0) = 1 \end{cases} \qquad u(1) = 0$$

no factor of h to help us

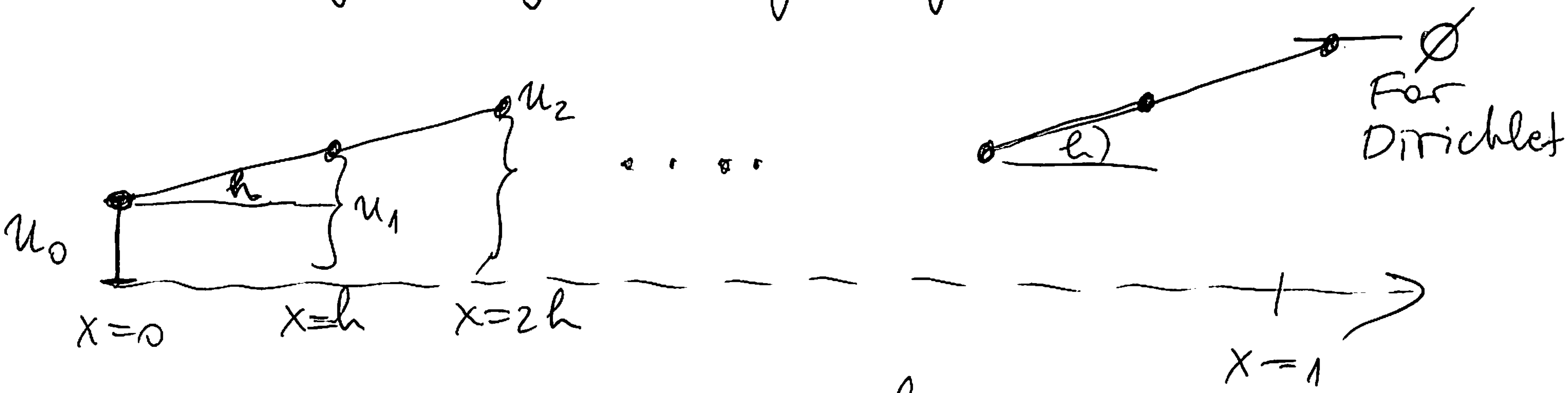
$\boxed{\text{Now } O(h) \text{ error on boundary causes } O(h) \text{ error everywhere}}$

Note that we can compute the Green's function here by hand algebraically or geometrically:

$$\frac{u_1 - u_0}{h} = 1 \quad \Rightarrow \quad u_1 = u_0 + h$$

After that, slopes are the same because discrete Laplacian is zero

$$u_{j+1} - u_j = u_j - u_{j-1}$$



$$\Rightarrow u_0 = -mh = 1$$

For interior points the difference is that we have a factor of

(20)

$$\frac{\frac{u_{j+1}-u_j}{h} - \frac{u_j-u_{j-1}}{h}}{h} = 1$$

So difference in <u>slopes</u> is $\underline{\underline{h}}$, which gave an extra factor of h.

Therefore for elliptic PDEs it seems OK to make $O(h)$ error at a few interior points and still be second-order, but <u>NOT</u> at boundaries.

[NOTE: PARABOLIC is EASIER for B.C.s]

We can of course use Richardson extrapolation to get higher-order (21)

$$U_j \approx u(jh) \quad , \; j = 1, \dots, m$$

$$V_j \approx u(jh/2) \quad , \; j = 1, \dots, 2m+1$$

$$u_j \approx V_{2j}$$

Note:

$$e_c = U_j - u(jh) = c_2 h^2 + c_4 h^4 + O(h^6)$$

$$e_f = V_{2j} - u(jh) = c_2 \left(\frac{h}{2}\right)^2 + c_4 \left(\frac{h}{2}\right)^4 + O(h^6)$$

Therefore

$$\boxed{\bar{v}_j = \frac{1}{3}\left(4V_{2j} - v_j\right)}$$ is <u>fourth order</u>

$$= u(jh) + \frac{1}{3}\left(\frac{1}{4} - 1\right) C_4 h^4$$

But for the Poisson equation there is an even simpler trick:

$$AE = -\tau = -\frac{h^2}{12} u^{(4)} + O(h^4)$$

$$= -\frac{h^2}{12} f'' + O(h^4)$$

$$A\,U = F$$

$$A\,E \approx -\frac{h^2}{12}\,(D^2 f) + O(h^4)$$

numerical second derivative

$$\hat{u} = U - E$$

"true" solution

$$\boxed{A\,\hat{U} = F + \frac{h^2}{12} D^2 f}$$

is a <u>fourth-order</u> discretization that costs almost the same as the second-order discretization!

$$A\hat{U} = \left(I + \frac{h^2}{12} D^2\right) f$$ (24)

$$\underbrace{\left[\left(I + \frac{h^2}{12} D^2\right)^{-1} D^2\right]}_{\displaystyle \hat{A} = \left(I + \frac{h^2}{12} D^2\right)^{-1} D^2} \hat{U} = f$$

must be a fourth-order approximation of the Laplacian (?) – this is called a <u>"compact difference"</u>

How to show this?

For <u>Periodic BCs</u>, use <u>Fourier</u>

We know the Fourier basis (25)

diagonalizes D^2 (in fact, any finite difference in a periodic domain).

So let's work in the Fourier basis.

$$D^2 e^{ikx} = \underbrace{\left(\frac{e^{ikh} - 2 + e^{-ikh}}{h^2} \right)}_{\lambda_k = \text{symbol of } D^2 \text{ (eigenvalue)}} e^{ikx}$$

$$\boxed{\lambda_k = -\frac{\sin^2(kh/2)}{(h/2)^2} = -k^2 + \frac{h^2 k^4}{12} + O(h^4)}$$

second order

So in Fourier space (26)

$$\boxed{\hat{D}^2 = -\frac{\sin^2(kh/2)}{(h/2)}}$$

So

$$\left(I + \frac{h^2}{12}\hat{D}^2\right)^{-1}\hat{D}^2 = -k^2 + \frac{k^6 h^4}{240} + O(h^6)$$

Fourth order!

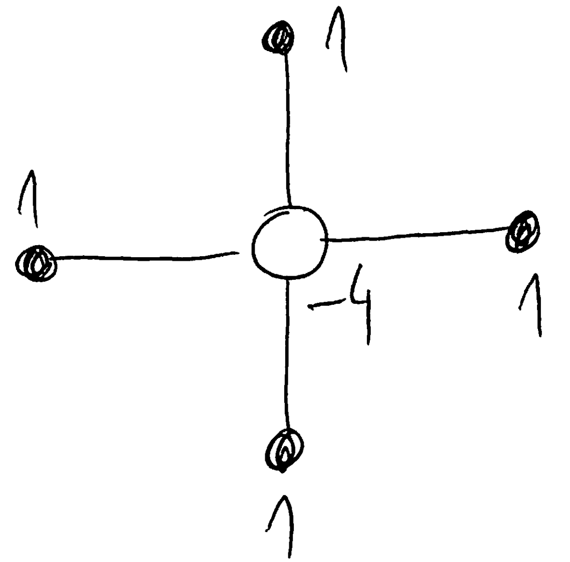
(I used Maple)

So indeed the compact difference is ~~fourth~~ fourth order

Generalizing to <u>Higher Dimensions</u> is <u>in principle</u> straightforward

$$\nabla^2 u = u_{xx} + u_{yy} = f(x,y)$$

$$\frac{1}{h^2}\left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{ij}\right) = f_{ij}$$



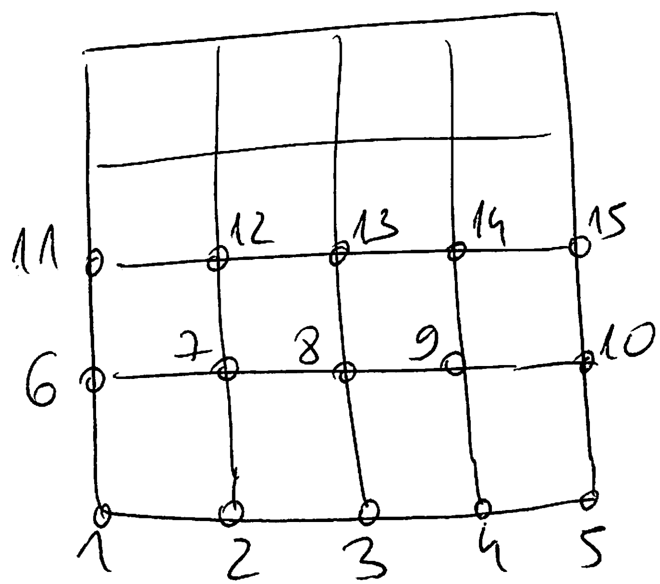
5^{pt} Laplacian stencil

Now take a box of $m \times m$ nodes.
The linear system

$$AU = F \text{ has } m^2 \text{ variables}$$

so A is $[m^2 \times m^2]$ & <u>very sparse</u>

Number of non-zeros is $5m^2$ only



$$A = \frac{1}{h^2}\begin{bmatrix} -4 & 1 & & & 1 & & \\ 1 & \ddots & \ddots & \underbrace{}_{m-1 \text{ zeros}} & & \ddots & 1 \\ & \ddots & & & & & \\ 1 & \underbrace{}_{m-1 \text{ zeros}} & \ddots & & \ddots & & 1 \\ & \ddots & & & & 1 & -4 \\ & & 1 & & & & \end{bmatrix}$$

What is the <u>local truncation error</u>?

Since x and y directions are 1D:

$$\tau_{ij} = \frac{1}{12} h^2 \left(u_{xxxx} + u_{yyyy} \right) + O(h^4)$$

To prove second order accuracy we need to bound the norm of $\|A^{-1}\|$

With <u>Dirichlet BCs</u> on all sides the <u>eigenvectors</u> of A are

$$u^{p,q}_{j,j'} = \sin(p\pi j h)\sin(q\pi j' h)$$

or $e^{i(2\pi(pjh + qj'h))}$ for <u>periodic BCs</u>

Since $D_{2D}^2 = D_x^2 + D_y^2 \Rightarrow$ 30

$$\lambda_{2D} = \lambda_x + \lambda_y$$

$$\boxed{\lambda_{p,q} = -\frac{\sin^2(k_x h/2)}{(h/2)^2} - \frac{\sin^2(k_y h/2)}{(h/2)^2}}$$

where $k_x = \frac{2\pi}{L} p$ and $k_y = \frac{2\pi}{L} q$

So all of the properties / analysis from 1D carries through directly

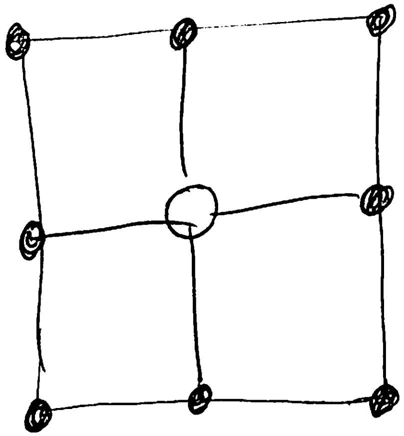
The conditioning number of A is

$$K_2(A) = \frac{|\lambda_{max}|}{|\lambda_{min}|} = \frac{\frac{8}{h^2}}{2\pi^2} = \frac{4}{\pi^2 h^2}$$

$$\boxed{K_2(A) \sim \frac{1}{h^2} \sim \frac{1}{m^2}}$$

grows very fast with the resolution m

So iterative methods for solving $AU=F$ will not converge fast unless we use a preconditioner



Consider also the

9

9^{pt} Laplacian with stencil

that includes the next-nearest neighbors:

$$
\left(\nabla_9^2 u\right)_{i,j} = \frac{1}{6h^2}\Big[4\left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}\right)
$$
$$
+ \left(u_{i-1,j-1} + u_{i-1,j+1} + u_{i+1,j+1} + u_{i+1,j-1}\right)
$$
$$
- 20\, u_{ij}\Big]
$$
$$
= \nabla^2 u(x_i, y_j) + \frac{h^2}{12}\left(u_{xxxx} + u_{yyyy} + 2u_{xxyy}\right) + O(h^4)
$$

$$(\nabla_g^2 u)_{i,j} = \nabla^2 u(x_i, y_j) + \underbrace{\frac{h^2}{12}(\nabla^2)^2 u}_{\text{isotropic}} + O(h^4)$$

This Laplacian is <u>isotropic</u> to $O(h^4)$, meaning, it is <u>rotationally-invariant</u>.
This can be important to preserve physics and symmetries.
Furthermore,

$$(\nabla^2)^2 u = \nabla^2 f$$

from the PDE

$$\nabla_g^2 u = \nabla^2 u + \frac{h^2}{12}\nabla^2 f + O(h^4)$$

So we can get a <u>fourth-order</u> discretization of ∇^2 in 2D, i.e., from a <u>compact finite difference</u>,

$$(\nabla_g^2 u)_{i,j} = f_{i,j} + \frac{h^2}{12}\nabla_{i,j}^2 f \quad , \text{e.g.}$$

$$\nabla_g^2 U = F + \frac{h^2}{12}\nabla_5^2 F = \left(I + \frac{h^2}{12}\nabla_5^2\right)F$$

<u>or</u>

$$\boxed{\nabla_g^2 U = \left(I + \frac{h^2}{12}\nabla_g^2\right)F}$$

<u>fourth order</u>

Which is related to

$$\nabla^2_{compact} = \left(I + \frac{h^2}{12}\nabla^2_9\right)^{-1} \nabla^2_9$$

Observe that if $F=0$, i.e., we are solving the Laplace equation, then

$\nabla^2_9 U = 0$ is a 4th order scheme!

So in this sense ∇^2_9 is a <u>better</u> discretization than ∇^2_5. However, it leads to a <u>denser more coupled</u> linear system that is harder to solve efficiently. So often we use ∇^2_5

How expensive is it to solve

$$AU = F$$

in 1D, 2D & 3D ?

→ Discuss on board

(P)CG note:

$$\|e_k\|_A \le 2\left(\frac{\sqrt{K_2}-1}{\sqrt{K_2}+1}\right)^k \|e_0\|_A$$

$$\approx 2\left(1-\frac{2}{\sqrt{K_2}}\right)^k \|e_0\|_A$$

Theorem (George)

Any direct method for $AU=F$ in 2D requires at least $O(m^3)$ operations

This bound is achieved by a nested-dissection algorithm

We will show in future that multigrid can solve in time $m^2\log(m)$ in 2D and $m^3\log(m)$ in 3D — same as FFT & optimal

$\boxed{\text{Factorizing } (a+b) \text{ diagonals matrix takes } N\cdot a\cdot b \text{ operations}}$

Conjugate gradients:

$$\frac{\|e_k\|_A}{\|e_0\|_A} \leq 2\left(\frac{\sqrt{K}-1}{\sqrt{K}+1}\right)^k \simeq 2\left(1-\frac{2}{\sqrt{K}}\right)^k$$

$$\simeq 2e^{-2k/\sqrt{K}}$$

so $\log \frac{\|e_k\|_A}{\|e_0\|_A} \approx -\frac{2k}{\sqrt{K}}$

So to get a fixed number of <u>accurate digits</u> we need a number of iterations $\sim \sqrt{K}$

iterations $\sim \sqrt{\frac{1}{h^2}} \sim m$

Each iteration costs $O(m)$ in 1D, $O(m^2)$ in 2D or $O(m^3)$ in 3D