Saddle-point problems, and why naive finite elements fail

Mixed finite element methods — a crash course · Lecture 1 of 4

Daniele Boffi

20 July 2026

All numerical results in these slides were produced by the verified reference implementation that accompanies the course materials.

From $\mathsf{A}\mathbf{x} = \mathbf{f}$ to saddle-point systems

Assumed background: for a coercive bilinear form a, the problem $a(u,v) = F(v)\ \forall v \in V$ discretizes to a symmetric positive definite system $\mathsf{A}\mathbf{x} = \mathbf{f}$ (Lax–Milgram, Céa).

Subject of this course: *pairs* of variational equations. Find $(u,p) \in V \times Q$ such that

$$a(u,v) + b(v,p) = F(v) \quad \forall v \in V, \qquad b(u,q) = G(q) \quad \forall q \in Q,$$

with bilinear forms a, b and data F, G. Discretization yields

$$\begin{pmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} :$$

symmetric but indefinite; the zero block reflects a *constraint*; coercivity on the whole space is lost, and the standard theory does not apply directly.

This lecture: where such systems arise, and how naive discretizations fail.

Model problem 1: the Stokes equations

Incompressible viscous flow: $-\nu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f}$, $\operatorname{div}\boldsymbol{u} = 0$ in Ω, $\boldsymbol{u} = \boldsymbol{0}$ on $\partial\Omega$.

Weak form, with $\boldsymbol{u} \in V := H^1_0(\Omega)^d$, $p \in Q := L^2_0(\Omega)$:

$$\begin{aligned} \nu(\nabla\boldsymbol{u},\, \nabla\boldsymbol{v}) - (\operatorname{div}\boldsymbol{v},\ p) &= (\boldsymbol{f},\, \boldsymbol{v}) && \forall\, \boldsymbol{v} \in V, \\ -(\operatorname{div}\boldsymbol{u},\ q) &= 0 && \forall\, q \in Q, \end{aligned}$$

i.e. $a(\boldsymbol{u},\boldsymbol{v}) = \nu(\nabla\boldsymbol{u},\nabla\boldsymbol{v})$, $\quad b(\boldsymbol{v},q) = -(\operatorname{div}\boldsymbol{v},q)$,

$F(\boldsymbol{v}) = (\boldsymbol{f},\boldsymbol{v})$, $\quad G = 0$ — the structure announced on the previous slide.

The constraint $\operatorname{div}\boldsymbol{u} = 0$ is part of the model: for this problem the saddle-point structure is intrinsic.

The origin of the structure: constrained minimization

The Stokes system is the optimality system of

$$\min_{\substack{\boldsymbol{v}\in V\\ \operatorname{div}\boldsymbol{v}=0}} \quad J(\boldsymbol{v}) = \tfrac{\nu}{2}\|\nabla \boldsymbol{v}\|_0^2 - (\boldsymbol{f}, \boldsymbol{v}).$$

Introduce the Lagrangian $\mathcal{L}(\boldsymbol{v}, q) = \frac{1}{2}a(\boldsymbol{v}, \boldsymbol{v}) - F(\boldsymbol{v}) + b(\boldsymbol{v}, q)$, $b(\boldsymbol{v}, q) = -(\operatorname{div}\boldsymbol{v}, q)$:

- stationarity in $\boldsymbol{v}$: $a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) = F(\boldsymbol{v}) \;\; \forall \boldsymbol{v}$ (momentum equation)
- stationarity in q: $b(\boldsymbol{u}, q) = 0 \;\; \forall q$ (the constraint)

The pressure is the Lagrange multiplier of the incompressibility constraint.
Remark: p carries no boundary condition; p is determined up to a constant (hence $Q = L_0^2$); p belongs to a space of lower regularity than $\boldsymbol{u}$.

Model problem 2: mixed formulation of Laplace pb.

$-\operatorname{div}(\kappa\nabla u) = f$ in Ω, $u = u_D$ on $\partial\Omega$.
Introduce the flux $\boldsymbol{\sigma} := \kappa\nabla u$:

$$\kappa^{-1}\boldsymbol{\sigma} - \nabla u = 0, \qquad \operatorname{div}\boldsymbol{\sigma} = -f.$$

Multiply the first equation by $\boldsymbol{\tau}$, integrate the ∇u term by parts, and insert $u = u_D$ on $\partial\Omega$.
With $\boldsymbol{\sigma} \in V := H(\operatorname{div};\Omega)$, $u \in Q := L^2(\Omega)$:

$$\begin{aligned} (\kappa^{-1}\boldsymbol{\sigma},\, \boldsymbol{\tau}) + (\operatorname{div}\boldsymbol{\tau},\, u) &= \langle u_D,\, \boldsymbol{\tau}\cdot\boldsymbol{n}\rangle \qquad && \forall\, \boldsymbol{\tau} \in V, \\ (\operatorname{div}\boldsymbol{\sigma},\, v) &= -(f,\, v) && \forall\, v \in Q, \end{aligned}$$

i.e. $a(\boldsymbol{\sigma},\boldsymbol{\tau}) = (\kappa^{-1}\boldsymbol{\sigma},\boldsymbol{\tau})$, $b(\boldsymbol{\tau},v) = (\operatorname{div}\boldsymbol{\tau},v)$,

$F(\boldsymbol{\tau}) = \langle u_D,\boldsymbol{\tau}\cdot\boldsymbol{n}\rangle$, $G(v) = -(f,v)$ — the same structure as for Stokes.

Observation: the Dirichlet datum has become a *natural* condition; neither space carries boundary conditions. (Discussed in Lecture 3.)

The two problems side by side

abstract	mixed Laplacian	Stokes
V	$H(\mathrm{div};\Omega) \ni \boldsymbol{\sigma}$	$H_0^1(\Omega)^d \ni \boldsymbol{u}$
Q	$L^2(\Omega) \ni u$	$L_0^2(\Omega) \ni p$
$a(\cdot,\cdot)$	$(\kappa^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau})$	$\nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v})$
$b(\cdot,\cdot)$	$(\mathrm{div}\, \boldsymbol{\tau},\, v)$	$-(\mathrm{div}\, \boldsymbol{v},\, q)$
$F,\ G$	$\langle u_D, \boldsymbol{\tau} \cdot \boldsymbol{n} \rangle,\ \ -(f, v)$	$(\boldsymbol{f}, \boldsymbol{v}),\ \ 0$
multiplier	scalar u	pressure p

Exchange of roles. In Stokes the minimizer is the physical velocity and the multiplier an auxiliary pressure. In the mixed Laplacian the multiplier is u — the original unknown of the problem — while the constrained minimizer is the flux (complementary energy; backup slide). Notation is overloaded across the two problems; each symbol is therefore always paired with its space.

Why a mixed formulation of the Laplace problem?

(Main application: Darcy flow in porous media.)

- **Local conservation.** With $\boldsymbol{q} := -\boldsymbol{\sigma}$ the physical flux, $\int_{\partial K} \boldsymbol{q} \cdot \boldsymbol{n}\, ds = \int_K f\, dx$ on every element. Mixed finite elements satisfy this identity exactly (proof: Lecture 3; experiment: Exercise 3); Lagrange elements do not.
- **Discontinuous coefficients.** Across a material interface, ∇u jumps while $\boldsymbol{q} \cdot \boldsymbol{n}$ is continuous; the mixed method approximates the continuous quantity.
- **Low regularity.** With $\boldsymbol{\sigma}$ as an unknown, $u \in L^2(\Omega)$ suffices.

Saddle-point problems also arise without being chosen, e.g. in nearly incompressible elasticity (volumetric locking; Lecture 4).

Plan. The simplest conforming discretization of the 1D mixed Laplacian converges. Two natural modifications of it fail, in two distinct ways. The analysis of these failures leads to the theory of Lecture 2.

The one-dimensional setting

On $\Omega = (0, 1)$, with $\kappa = 1$ and $u(0) = u(1) = 0$ (natural conditions: no `DirichletBC` in the code); in one dimension $H(\mathrm{div}) = H^1$:

$$(\sigma, \tau) + (\tau', u) = 0 \quad \forall \tau \in V_h,$$
$$(\sigma', v) = -(f, v) \quad \forall v \in Q_h.$$

- data: $f = 1$ ($u = \frac{1}{2}x(1-x)$, $\sigma = \frac{1}{2} - x$) for hand computation; $f = \pi^2 \sin(\pi x)$ ($u = \sin \pi x$, $\sigma = \pi \cos \pi x$) for convergence rates
- lowest-order conforming pair on a uniform mesh with n elements:

$$V_h = \text{continuous } P_1 \quad (\dim = n{+}1),$$
$$Q_h = \text{piecewise constants} \quad (\dim = n)$$

The constraint matrix is the fundamental theorem of calculus

Testing the second equation with the indicator of K_k gives $\int_{K_k} \sigma_h' = \sigma_h(x_k) - \sigma_h(x_{k-1})$ — an exact identity, free of quadrature error. Hence ($n = 4$):

$$\mathsf{B} = \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix} \qquad (\mathsf{B}\boldsymbol{\sigma})_k = -\int_{K_k} f.$$

$\ker \mathsf{B}^T$ by inspection: $(\mathsf{B}^T\mathbf{q})_0 = -q_1$, then $(\mathsf{B}^T\mathbf{q})_j = q_j - q_{j+1}$, hence $\mathbf{q} = \mathbf{0}$.

This matrix will be needed again shortly.

From exact increments to an $O(h^2)$ flux error

1. The constraint determines every increment exactly: $\sigma_h(x_k) - \sigma_h(x_{k-1}) = -\int_{K_k} f = \sigma(x_k) - \sigma(x_{k-1})$.
2. Hence $\sigma_h = I_h\sigma + c$ (nodal interpolant plus a constant).
3. Testing the first equation with $\tau = 1 \in V_h$ (note that $b(1, v_h) = 0$) gives $\int \sigma_h = 0 = \int \sigma$, hence $c = \int(\sigma - I_h\sigma) = O(h^2)$.

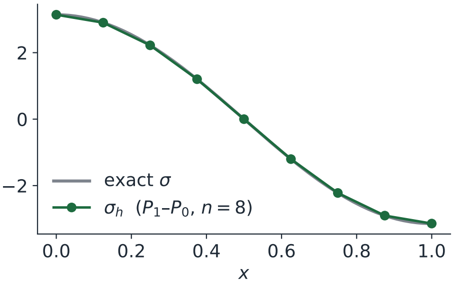
$$\boxed{\|\sigma - \sigma_h\|_0 = O(h^2)}$$

For both lecture data the constant vanishes (σ linear, resp. antisymmetric about $x = \frac{1}{2}$), so the demonstration will show exact nodal values; in general $c \neq 0$: for $f(x) = x$, $c = h^2/12 + O(h^4)$ (verified).

The lowest-order pair converges

Measured rates,
$f = \pi^2 \sin(\pi x)$:

quantity	rate
$\|\sigma - \sigma_h\|_0$	2.00
$\|(\sigma - \sigma_h)'\|_0$	1.00
$\|u - u_h\|_0$	1.00
u_h at midpoints	2.00



For $f = 1$: $\sigma_h = \sigma$ to machine precision; the midpoint error of u_h equals $h^2/24$.

▶ **live demonstration:** notebook `L1A`, cell 4

Two remarks

1. The $O(h^2)$ rate for σ is specific to one dimension: the constraint determines every nodal value through the fundamental theorem of calculus. In higher dimensions the corresponding rate is $O(h)$ (Lecture 3); this superconvergence should not be extrapolated.
2. **This pair should be kept in mind. It has a standard name and a natural generalization to dimensions $d \geq 2$; both are given in Lecture 3.**

The discrete saddle-point system

Galerkin discretization with any $V_h \subset V$, $Q_h \subset Q$ yields the structure announced at the beginning:

$$\begin{pmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma} \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix}, \qquad \mathsf{A}_{jk} = a(\varphi_k, \varphi_j), \quad \mathsf{B}_{ij} = b(\varphi_j, \psi_i).$$

- symmetric but indefinite; the zero block expresses that the multiplier contributes no energy (solution algorithms: Lecture 4)
- first question: under which conditions is this matrix invertible?

Solvability is governed by $\ker \mathsf{B}^T$

Lemma 1. Let A be symmetric positive definite on all of V_h (true for both model problems: a weighted mass matrix, resp. a Dirichlet stiffness matrix). Then the block matrix is invertible $\iff \ker \mathsf{B}^T = \{0\} \iff$ B is surjective.

Proof. $\mathsf{A}\mathbf{x} + \mathsf{B}^T\mathbf{y} = 0$ and $\mathsf{B}\mathbf{x} = 0$ give $\mathbf{x}^T\mathsf{A}\mathbf{x} = -(\mathsf{B}\mathbf{x})^T\mathbf{y} = 0$, hence $\mathbf{x} = 0$, hence $\mathsf{B}^T\mathbf{y} = 0$. Conversely, $(\mathbf{0}, \mathbf{y})$ lies in the kernel for every $\mathbf{y} \in \ker \mathsf{B}^T$. $\square$

- **Definition.** A function $q_h \neq 0$ with $b(\tau_h, q_h) = 0\ \forall \tau_h \in V_h$ is a spurious mode.
- **Corollary (dimension count).** Surjectivity requires $\dim Q_h \leq \dim V_h$. This condition is necessary; the experiments below measure how far it is from sufficient.

Two natural modifications

P_1–P_0 satisfies the criterion of Lemma 1 ($\ker \mathsf{B}^T = \{0\}$, verified by inspection), and A is a mass matrix. Both spaces are of the lowest possible order; let's try to modify the method, one space at a time:

Modification 1

higher resolution for the *multiplier*

$$Q_h : \; P_0 \longrightarrow \text{continuous } P_1$$

$$P_1\text{–}P_1$$

Modification 2

higher resolution for the *flux*

$$V_h : \; P_1 \longrightarrow \text{continuous } P_2$$

$$P_2\text{–}P_0$$

Before computing: which of the two modified pairs do you expect to converge?

Modification 1: the effect of a richer multiplier space

Equal-order interpolation, the same space for both fields, is also the first choice most practitioners make in two dimensions.
Dimension count: $n+1$ vs. $n+1$ — satisfied with equality.

$$\underbrace{\begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix}}_{P_1 - P_0:\ \text{exact element increments}} \longrightarrow \underbrace{\frac{1}{2}\begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}}_{P_1 - P_1:\ \text{averaged differences}}$$

The richer test space replaces the element rows by averages.
Interior rows are central differences: node i is coupled only to $i-1$ and $i+1$, so odd and even indices decouple
(the same decoupling as for collocated finite differences; staggered grids are the classical remedy — Lecture 3).

The kernel: an alternating mode

- In general (two lines): $\{\tau_h' : \tau_h \in V_h\}$ consists of *all* piecewise constants; hence $q_h \in \ker B^T$ if and only if $\int_K q_h = 0$ on every element, i.e. $q_i = -q_{i-1}$:

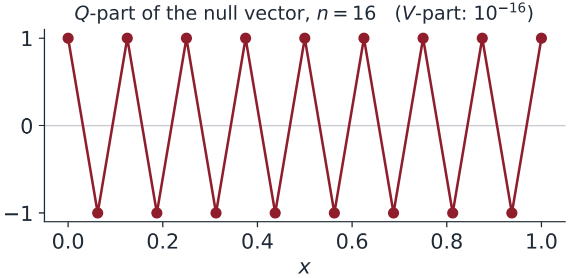
$$\ker B^T = \operatorname{span}\{\chi_h\}, \qquad \chi_i = (-1)^i, \qquad \text{for every } n.$$

- Verification on one row: $(\mathsf{B}^T \boldsymbol{\chi})_j = \frac{1}{2}(\chi_{j-1} - \chi_{j+1}) = 0$; boundary rows: $\mp\frac{1}{2}(\chi_0 + \chi_1) = 0$.
- By Lemma 1 the system is singular for every n.

**The dimension count is satisfied with equality, and the system is singular:
the count is necessary, but not sufficient.**

Demonstration: the singular system

- **live demonstration:** notebook `L1A`, cells 6–7
  - direct LU factorization ($f = 1$, $n = 64$) fails with a zero-pivot error (verified: smallest singular value 10^{-16}, second smallest $O(10^{-2})$)
  - dense SVD: one zero singular value; the Q-part of the corresponding null vector is shown; its V-part vanishes to machine precision



Questions deferred to Exercise 1

For the singular pair, three questions remain: when is the system solvable at all; which right-hand sides are admissible; and what a small regularization of the zero block returns. They are treated in Exercise 1 and in the self-study notebook cells `S1-S2`. Two results, to be derived there:

- solvability depends on the parity of n
- regularized solutions range from apparently correct to an alternating mode of amplitude $3.8 \cdot 10^5$, depending only on the datum f

(Both results are available on backup slides.)

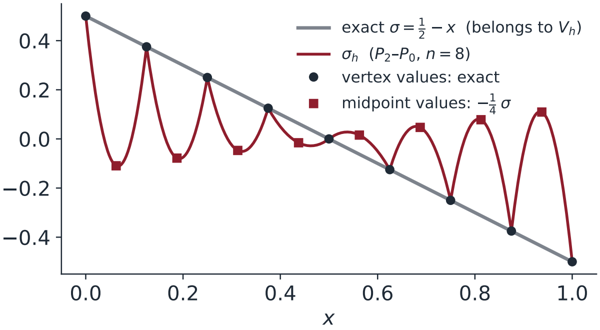
Modification 2: no algebraic defect

$V_h =$ continuous P_2 $(2n{+}1)$, $Q_h = P_0$ (n): the dimension count is satisfied with a margin of $n{+}1$.

- $\ker \mathsf{B}^T$: for $q_h \in P_0$, the function $\tau_h(x) = \int_0^x q_h$ belongs to $P_1 \subset V_h$ and gives $b(\tau_h, q_h) = \|q_h\|_0^2$; hence $\ker \mathsf{B}^T = \{0\}$.
- By Lemma 1 the matrix is invertible for every n.

According to every criterion introduced so far, this pair is preferable to the baseline.

The computed flux oscillates



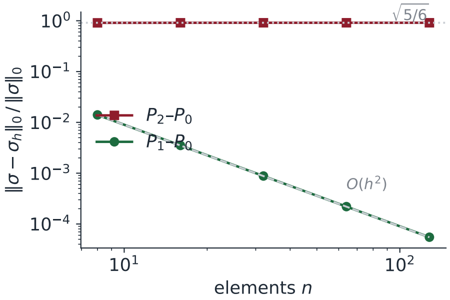
Measured in every element: vertex values exact; midpoint values equal to $-\frac{1}{4}$ of the exact flux (ratio -0.2500).

Derivation: Exercise 2(d).

▶ **live demonstration:** `L1A`, cell 10

The exact flux belongs to V_h; the computed flux does not reproduce it.

The errors converge — to nonzero limits



$f = \pi^2 \sin(\pi x)$:

n	$\frac{\|\sigma-\sigma_h\|_0}{\|\sigma\|_0}$	$\frac{\|u-u_h\|_0}{\|u\|_0}$
8	0.9094	0.8272
32	0.9127	0.8330
128	0.9129	0.8333

limits: $\sqrt{5/6}$ and $5/6$

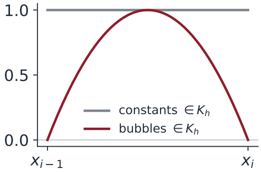
The scalar variable converges, with the correct shape, to $u/6$: without a reduced reference solution, its plot does not reveal the error.

(Exercise 2; backup slide.) ▶ **live demonstration:** `L1A`, cell 11

Diagnosis: the discrete constraint kernel

$$K_h = \{\tau_h : \int_K \tau_h' = 0 \ \ \forall K\} = \mathrm{span}\{1\} \oplus \mathrm{span}\{b_K\}_K$$

- the bubbles violate the *continuous* constraint $\tau' = 0$ — only their element means vanish: $K_h \not\subset K$
- on the bubbles, the ratio $a(\tau_h, \tau_h)/\|\tau_h\|^2_{H(\mathrm{div})}$ is of order $h^2/60$ — measured on every mesh (Exercise 2(b–c))



The modification added precisely those functions which satisfy all discrete constraints but on which a degenerates.

Made quantitative in Lecture 2.

Summary of the evidence

pair	count	$\ker \mathsf{B}^T$	matrix	σ error	u error	outcome
P_1–P_0	holds $(+1)$	$\{0\}$	invertible	$O(h^2)$	$O(h)$	converges
P_1–P_1	holds $(=)$	$\mathrm{span}\{\chi_h\}$	singular $\forall h$	—	—	spurious mode
P_2–P_0	holds $(\gg)$	$\{0\}$	invertible	$\to 0.91$	$\to 0.83$	wrong limit

The relevant balance

Each failure was produced by modifying a convergent pair in a direction that appeared advantageous — on opposite sides:

- **Enriching** Q_h adds rows to B: the constraint imposes more conditions than the flux space can match; surjectivity is lost; spurious modes appear. (P_1–P_1: equal dimensions.)
- **Enriching** V_h adds no conditions: the discrete kernel K_h grows — here by the bubbles, on which a degenerates. (P_2–P_0: a large dimension surplus.)

The relevant balance is not between the dim.s of V_h and Q_h.
It is a balance between the two spaces as coupled by b:
B **must remain (uniformly) surjective, and a must control what b does not see in V_h.**
Enriching either space may destroy it.

Two species of failure

(i) Singular constraint: $\ker B^T \neq \{0\}$; spurious modes. Detectable by finite linear algebra at a *single* value of h.

(ii) Degeneration with respect to h: every fixed-h certificate holds, while a constant decays with h (here $\sim h^2/60$). No computation at a fixed h can detect it.

Case (ii) is the more delicate one; it is the reason this subject requires analysis in addition to experiments.

Abstract saddle-point pb (recalled; dictionary: slide 6)

As announced at the opening: find $(u,p) \in V \times Q$ such that

$$a(u,v) + b(v,p) = F(v) \;\; \forall v \in V, \qquad b(u,q) = G(q) \;\; \forall q \in Q.$$

Constraint kernels:

$$K := \{v \in V : \; b(v,q) = 0 \; \forall q \in Q\},$$
$$K_h := \{v_h \in V_h : \; b(v_h,q_h) = 0 \; \forall q_h \in Q_h\}.$$

Measured above: $K_h \not\subset K$ in general — the P_2–P_0 bubbles satisfy every discrete constraint and violate the continuous one; the loss of control occurred precisely there.

Fewer test functions $\Rightarrow$ a weaker discrete constraint $\Rightarrow$ a larger discrete kernel: enriching V_h alone may enlarge K_h, as observed.

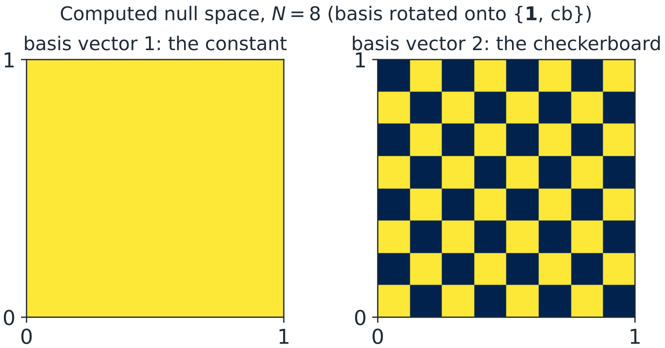
A two-dimensional example: Q_1–P_0 for Stokes

Stokes on the unit square, uniform $N \times N$ *quadrilateral* mesh; the lowest-order quadrilateral pair:

- continuous Q_1 velocities (bilinear per cell), P_0 pressures
- regularized lid: $\boldsymbol{u} = (16x^2(1-x)^2, 0)$ on the top side, $\boldsymbol{u} = \mathbf{0}$ on the remaining boundary (a smooth datum: corner singularities play no role)

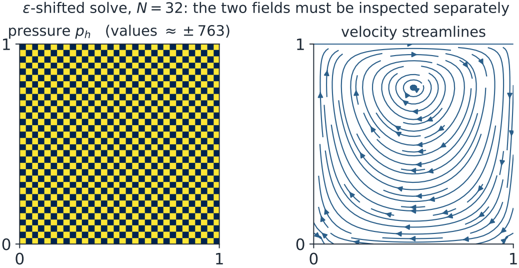
Dimension count: $2(N-1)^2$ interior velocity unknowns vs. N^2 pressures — satisfied for all $N \geq 4$ ($N = 32$: 1922 vs. 1024). Nevertheless, the system is singular.

The kernel, computed



Computed for $N = 4, \ldots, 16$: exactly *two* zero singular values; the null space equals $\mathrm{span}\{\mathbf{1},\ \text{checkerboard}\}$ to 10^{-15}. The constant mode is physical (enclosed flow, $Q = L_0^2$); the checkerboard is spurious — the one-dimensional alternating mode, in two dimensions. Membership in $\ker B^T$: Exercise 4(a).

The two fields must be inspected separately



▶ **live demonstration:** `L1B`: the LU factorization fails as in 1D; the ε-regularized solve returns the fields above.

The velocity appears correct while the pressure is dominated by the spurious mode: each field must be inspected separately.

Checkerboard filters were used in production codes in the 1970s. The analogous triangular pair (P_1–P_1) is examined in Exercise 4(c) and classified in Lecture 2.

The evidence, completed

pair	count	$\ker \mathsf{B}^T$	matrix	outcome
P_1–P_0 (1D)	holds $(+1)$	$\{0\}$	invertible	converges $2/1/1/2$
P_1–P_1 (1D)	holds $(=)$	$\mathrm{span}\{\chi_h\}$	singular	spurious mode
P_2–P_0 (1D)	holds $(\gg)$	$\{0\}$	invertible	oscillates; $u_h \to u/6$
Q_1–P_0 (2D)	holds $(N \geq 4)$	$\mathbf{1} \oplus$ checkerboard	singular	spurious mode

Solvability at a fixed h guarantees little; invertibility at every h is compatible with convergence to a wrong limit. The missing conditions must be *quantitative* and *uniform in* h.

The two questions

Q1. When is the *continuous* saddle-point problem well-posed?

Q2. When does a Galerkin discretization *inherit* that well-posedness, uniformly in h?

These two questions are answered in Lecture 2.

The programme of Lecture 2

- **Two conditions** on (a, b, V_h, Q_h) answering Q2: a quantitative, h-uniform surjectivity condition on b, and control of a on K_h (the ratio measured on the bubbles will be given a name and a central role)
- An **error estimate** in which the reciprocal of the stability constant appears explicitly — the nonzero limits observed today become a theorem
- A practical, **projection-based criterion** for verifying the hard condition, pair by pair

The experiments of this lecture return as the examples of Lecture 2.

Exercise sheet 1, and outlook

1. the alternating mode: kernel, solvability, parity of n, effect of regularization
2. derivation of the $-\frac{1}{4}$ identity, of the ratio $h^2/60$, and of the limit $u_h \to u/6$
3. local conservation and discontinuous coefficients: mixed vs. primal formulation
4. the two-dimensional kernel; the triangular case (anticipates Lecture 2)

Notebook cells tagged `S1`–`S4` contain the material omitted from the lecture.

The pair that opened the lecture will be identified — and generalized — in Lecture 3.

Backup slides

Backup: the mixed Laplacian as constrained minimization

Complementary-energy principle:

$$\min_{\substack{\boldsymbol{\tau}\in H(\mathrm{div})\\ \mathrm{div}\,\boldsymbol{\tau}=-f}} \ \tfrac{1}{2}(\kappa^{-1}\boldsymbol{\tau},\boldsymbol{\tau}) - \langle u_D, \boldsymbol{\tau}\cdot\boldsymbol{n}\rangle,$$

$$\mathcal{L}(\boldsymbol{\tau}, v) = \tfrac{1}{2}(\kappa^{-1}\boldsymbol{\tau},\boldsymbol{\tau}) - \langle u_D, \boldsymbol{\tau}\cdot\boldsymbol{n}\rangle + (\mathrm{div}\,\boldsymbol{\tau} + f,\ v).$$

- stationarity in $\boldsymbol{\tau}$: $(\kappa^{-1}\boldsymbol{\sigma},\boldsymbol{\tau}) + (\mathrm{div}\,\boldsymbol{\tau},\ v) = \langle u_D, \boldsymbol{\tau}\cdot\boldsymbol{n}\rangle \quad \forall\boldsymbol{\tau}$ — the first equation, with the multiplier v in the position of u
- stationarity in v: $(\mathrm{div}\,\boldsymbol{\sigma}, v) = -(f, v) \quad \forall v$ — the constraint

The scalar variable u is the Lagrange multiplier of the constraint $\mathrm{div}\,\boldsymbol{\tau} = -f$.

Backup: parity and compatibility (P_1–P_1)

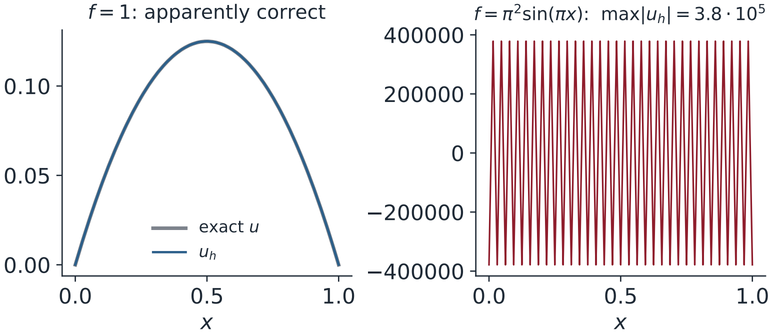
The singular system is solvable $\iff$ $(f, \chi_h) = 0$. One has $\int_0^1 \chi_h = 0$ for every n; moreover χ_h is symmetric about $x = \frac{1}{2}$ for even n and antisymmetric for odd n. Verified values of $(\mathbf{g}, \boldsymbol{\chi})$:

n	$f = 1$	$f = \pi^2 \sin(\pi x)$	$f = x$
4	0	0.344	0
5	0	10^{-16}	$6.7 \cdot 10^{-3}$
8	0	0.082	0

Solvability therefore depends on the parity of the number of elements. Derivation: Exercise 1(c).

Backup: regularized solves, two right-hand sides

Zero block replaced by $-10^{-8}\mathrm{M}_Q$; $n = 64$; identical code and mesh:

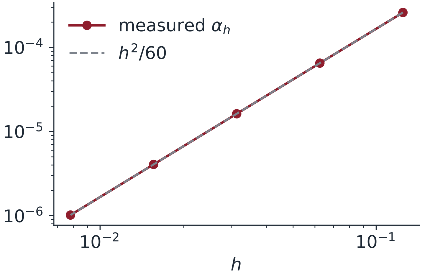


Left: compatible datum; the result is the representative of the family $u_h + t\chi_h$ selected by the solver, and its accuracy is not controlled. Right: incompatible datum. At $n = 65$ the two cases exchange roles (parity). Exercise 1(d).

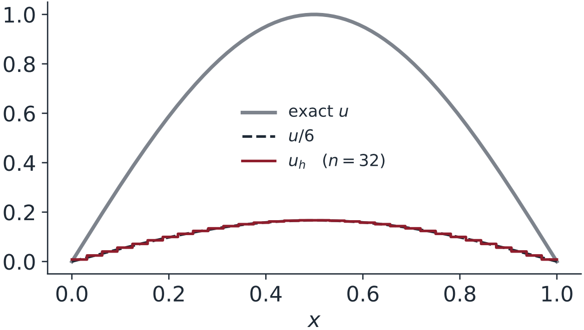
Backup: the kernel-coercivity ratio

$$\alpha_h := \min_{\tau_h \in K_h} \frac{\|\tau_h\|_0^2}{\|\tau_h\|_0^2 + \|\tau_h'\|_0^2}$$

A single bubble gives $\alpha_h \leq h^2/10$; the bubble minus its mean gives $h^2/60$, and the measured values show this bound to be sharp (Exercise 2(b–c)).



Backup: the limit $u_h \to u/6$ (P_2–P_0)



Measured: $\|u_h - u/6\|_0 / \|u\|_0 = 9.7 \cdot 10^{-3}$ at $n = 16$, $3.0 \cdot 10^{-4}$ at $n = 512$. Leading-order derivation: Exercise 2(e*).