$H(\mathrm{div})$ and the mixed Laplacian

Mixed finite element methods — a crash course · Lecture 3 of 4

Daniele Boffi

23 July 2026

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

The specification left by Lecture 2

Lecture 2 established: a Galerkin pair (V_h, Q_h) is uniformly stable if and only if (ElKer$_h$) and (InfSup$_h$) hold with h-independent constants. For the mixed Laplacian, $a(\boldsymbol{\tau}, \boldsymbol{\tau}) = \|\kappa^{-1/2}\boldsymbol{\tau}\|_0^2$ controls only the L^2 part of $\|\cdot\|_{\mathrm{div}}$. The two hypotheses translate into a construction task:

- **(S1)** $V_h \subset H(\mathrm{div}; \Omega)$, conforming;
- **(S2)** $\mathrm{div}\, V_h \subset Q_h$ — then $\mathrm{div}\, \boldsymbol{v}_h = 0$ on K_h, so $K_h \subset K$ and (ElKer$_h$) holds with $\alpha_h \geq 1/\kappa_+$, exactly as on the continuous level;
- **(S3)** an interpolation operator Π_h with $\mathrm{div}\, \Pi_h \boldsymbol{\tau} = P_h \,\mathrm{div}\, \boldsymbol{\tau}$ and $\|\Pi_h \boldsymbol{\tau}\|_{\mathrm{div}} \leq C\|\boldsymbol{\tau}\|_1$ — then (F1$_W$), (F2$_W$) hold with $W = (H^1)^d$, and Corollary 12 yields (InfSup$_h$) uniformly.

This lecture constructs the spaces and the operator, proves the estimates, and turns the result into a working solver.

The space $H(\mathrm{div};\Omega)$

Definition. $H(\mathrm{div};\Omega) := \{\boldsymbol{\tau} \in L^2(\Omega)^d : \ \mathrm{div}\,\boldsymbol{\tau} \in L^2(\Omega)\}$, the divergence taken in the sense of distributions;

$$\|\boldsymbol{\tau}\|_{\mathrm{div}}^2 := \|\boldsymbol{\tau}\|_0^2 + \|\mathrm{div}\,\boldsymbol{\tau}\|_0^2.$$

$H(\mathrm{div};\Omega)$ is a Hilbert space, and

$$H^1(\Omega)^d \ \subsetneq \ H(\mathrm{div};\Omega) \ \subsetneq \ L^2(\Omega)^d.$$

- Membership requires only that *one* scalar combination of first derivatives is a function; the individual components need not be in H^1.
- In particular, fields with discontinuous *tangential* components are admitted. The next slide makes this precise on a mesh; the resulting characterization is the design principle for the degrees of freedom.

Membership is a condition on the normal jumps

Lemma 15 (patching). Let $\boldsymbol{\tau} \in L^2(\Omega)^d$ with $\boldsymbol{\tau}|_K \in H^1(K)^d$ for every $K \in \mathcal{T}_h$. Then

$$\boldsymbol{\tau} \in H(\mathrm{div};\Omega) \iff [\![\boldsymbol{\tau} \cdot \boldsymbol{n}]\!]_e = 0 \text{ on every interior facet } e,$$

and in that case $(\mathrm{div}\,\boldsymbol{\tau})|_K = \mathrm{div}(\boldsymbol{\tau}|_K)$.

Proof. For $\varphi \in C_c^\infty(\Omega)$, elementwise integration by parts gives

$$\int_\Omega \boldsymbol{\tau} \cdot \nabla\varphi = -\sum_K \int_K \mathrm{div}(\boldsymbol{\tau}|_K)\,\varphi + \sum_{e \text{ int.}} \int_e [\![\boldsymbol{\tau} \cdot \boldsymbol{n}]\!]_e\,\varphi.$$

The distributional divergence is therefore the piecewise divergence minus a distribution supported on the facets; the latter belongs to $L^2(\Omega)$ if and only if it vanishes, i.e. if and only if every normal jump is zero. $\square$

Conformity in $H(\mathrm{div})$ constrains only the normal component: the degrees of freedom must control normal traces — nothing more.

The normal trace, and Green's formula

Theorem 16 (stated; see Boffi–Brezzi–Fortin, Ch. 2). The normal trace $\boldsymbol{\tau} \mapsto \boldsymbol{\tau} \cdot \boldsymbol{n}|_{\partial\Omega}$, defined on smooth fields, extends to a bounded and surjective operator $H(\mathrm{div};\Omega) \to H^{-1/2}(\partial\Omega)$, and Green's formula holds:

$$(\mathrm{div}\,\boldsymbol{\tau},\, v) + (\boldsymbol{\tau},\, \nabla v) = \langle \boldsymbol{\tau} \cdot \boldsymbol{n},\, v \rangle_{\partial\Omega} \qquad \forall\, v \in H^1(\Omega).$$

- The normal trace is in general *not* a function in $L^2(\partial\Omega)$; the exponent $-1/2$ is sharp.
- Consequently it cannot be restricted to a part of the boundary: an element of $H^{-1/2}(\partial\Omega)$ acts on functions defined on all of $\partial\Omega$, and the extension of $\varphi \in H^{1/2}(e)$ by zero does not belong to $H^{1/2}(\partial\Omega)$ in general. In particular, the edge values $\boldsymbol{\tau} \cdot \boldsymbol{n}|_e$ are *not defined* for $\boldsymbol{\tau} \in H(\mathrm{div};\Omega)$ — the fact that will prevent the canonical interpolant of slide 14 from being defined on all of $H(\mathrm{div};\Omega)$.
- The right-hand side $F(\boldsymbol{\tau}) = \langle u_D, \boldsymbol{\tau} \cdot \boldsymbol{n} \rangle$ used since Lecture 1 is this duality pairing; it requires $u_D \in H^{1/2}(\partial\Omega)$.

The exchange of essential and natural conditions

The observation of Lecture 1 (frame 5), now systematic:

	primal ($u \in H^1$)	mixed ($\boldsymbol{\sigma} \in H(\mathrm{div})$)
$u = u_D$ on Γ_D	**essential** (in the space)	natural (enters F)
$\boldsymbol{\sigma} \cdot \boldsymbol{n} = g$ on Γ_N	natural (enters F)	**essential** (constrained in V_h)

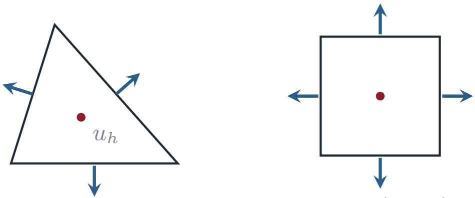
- This explains the absence of `DirichletBC` in every code shown so far: the Dirichlet datum enters the linear functional.
- Conversely, a prescribed flux must be imposed *in the space*: in code, a boundary condition object on the flux space, constraining the normal degrees of freedom introduced below.

RT_0: shape functions and degrees of freedom

On a triangle K (statements hold on simplices, $d = 2, 3$; proofs for $d = 2$):

$$RT_0(K) := P_0(K)^2 + \boldsymbol{x}\, P_0(K) = \{\boldsymbol{a} + c\,\boldsymbol{x}\}, \qquad \dim = 3.$$

Degrees of freedom: the edge fluxes $N_e(\boldsymbol{\tau}) := \int_e \boldsymbol{\tau} \cdot \boldsymbol{n}_e \, ds$, one per edge, with a *fixed global* normal $\boldsymbol{n}_e$.



staggered (MAC) cell

Remark. Normal fluxes on the faces, one scalar per cell: the staggered-grid arrangement announced in Lecture 1 (frame 16), realized as a finite element.

RT_0: unisolvence and basic properties

Proposition 17. On every triangle K: (i) $\operatorname{div} RT_0(K) = P_0(K)$; (ii) $\boldsymbol{\tau} \cdot \boldsymbol{n}$ is constant on each edge; (iii) the three edge fluxes are unisolvent, with dual basis

$$\boldsymbol{\varphi}_e(\boldsymbol{x}) = \frac{s_{K,e}}{2|K|}\,(\boldsymbol{x} - \boldsymbol{p}_e),$$

where $\boldsymbol{p}_e$ is the vertex opposite e and $s_{K,e} = \boldsymbol{n}_e \cdot \boldsymbol{n}_{\partial K} \in \{\pm 1\}$.

Proof. (ii) $\boldsymbol{x} \cdot \boldsymbol{n}$ is constant along any line with normal $\boldsymbol{n}$; hence $(\boldsymbol{a} + c\boldsymbol{x}) \cdot \boldsymbol{n}_e$ is constant on e. (i) $\operatorname{div}(\boldsymbol{a} + c\boldsymbol{x}) = 2c$; the map is onto $P_0(K)$ with kernel the constants. (iii) Let all three fluxes vanish. By (ii), $\boldsymbol{\tau} \cdot \boldsymbol{n} \equiv 0$ on ∂K; then $2c\,|K| = \int_K \operatorname{div} \boldsymbol{\tau} = \int_{\partial K} \boldsymbol{\tau} \cdot \boldsymbol{n} = 0$, so $c = 0$; then $\boldsymbol{a} \cdot \boldsymbol{n}_e = 0$ for two linearly independent normals, so $\boldsymbol{a} = \boldsymbol{0}$. For the basis: $(\boldsymbol{x} - \boldsymbol{p}_e) \cdot \boldsymbol{n}_{\partial K}$ equals the distance $2|K|/|e|$ from $\boldsymbol{p}_e$ to the line of e, for every $\boldsymbol{x} \in e$; hence $\int_e \boldsymbol{\varphi}_e \cdot \boldsymbol{n}_e \, ds = 1$, while on the other two edges $(\boldsymbol{x} - \boldsymbol{p}_e) \parallel e'$ gives 0. $\square$

The family RT_k, $k \geq 0$

$$RT_k(K) := P_k(K)^2 \oplus \boldsymbol{x}\, \widetilde{P}_k(K), \qquad \dim = (k{+}1)(k{+}3)$$

($\widetilde{P}_k$: homogeneous of degree k); $\operatorname{div} RT_k(K) = P_k(K)$, $\boldsymbol{\tau} \cdot \boldsymbol{n}|_e \in P_k(e)$. Degrees of freedom: $\int_e (\boldsymbol{\tau} \cdot \boldsymbol{n})\, q$ for $q \in P_k(e)$ on each edge, and $\int_K \boldsymbol{\tau} \cdot \boldsymbol{q}$ for $\boldsymbol{q} \in P_{k-1}(K)^2$.

Theorem 18 (unisolvence; $d = 2$; for $d = 3$, faces replace edges; details on backup B2). Assume all degrees of freedom vanish.

1. On each edge, $\boldsymbol{\tau} \cdot \boldsymbol{n}|_e \in P_k(e)$; taking $q = \boldsymbol{\tau} \cdot \boldsymbol{n}|_e$ in the edge moments gives $\int_e (\boldsymbol{\tau} \cdot \boldsymbol{n})^2\, ds = 0$, so $\boldsymbol{\tau} \cdot \boldsymbol{n} \equiv 0$ on ∂K.
2. For $q \in P_k(K)$: $\int_K (\operatorname{div} \boldsymbol{\tau}) q = -\int_K \boldsymbol{\tau} \cdot \nabla q + \int_{\partial K} (\boldsymbol{\tau} \cdot \boldsymbol{n}) q = 0$ $(\nabla q \in P_{k-1}^2)$; since $\operatorname{div} \boldsymbol{\tau} \in P_k(K)$: $\operatorname{div} \boldsymbol{\tau} = 0$.
3. $\operatorname{div} \boldsymbol{\tau} = 0$ on the simplex gives $\boldsymbol{\tau} = \mathbf{curl}\, \psi := (\partial_2 \psi, -\partial_1 \psi)$, $\psi \in P_{k+1}(K)$; $\boldsymbol{\tau} \cdot \boldsymbol{n} = \partial_t \psi = 0$ on ∂K makes ψ constant there, so $\psi = b_K\, r$, $r \in P_{k-2}(K)$.
4. Choose $\boldsymbol{q} \in P_{k-1}^2$ with $\operatorname{rot} \boldsymbol{q} = r$; then $0 = \int_K \boldsymbol{\tau} \cdot \boldsymbol{q} = \int_K b_K\, r^2$, so $r = 0$, so $\boldsymbol{\tau} = \mathbf{0}$. $\square$

The Piola transform, and the orientation of edges

With $F_K(\hat{\boldsymbol{x}}) = \mathsf{B}_K \hat{\boldsymbol{x}} + \boldsymbol{b}_K$ the affine map onto K, the **Piola transform** of $\hat{\boldsymbol{\tau}}$ is $\boldsymbol{\tau}(\boldsymbol{x}) := (\det \mathsf{B}_K)^{-1} \mathsf{B}_K \, \hat{\boldsymbol{\tau}}(\hat{\boldsymbol{x}})$.

Proposition 19 (verified numerically).

(i) fluxes are invariant: $\int_e (\boldsymbol{\tau} \cdot \boldsymbol{n}) \, q \, ds = \int_{\hat{e}} (\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{n}}) \, \hat{q} \, d\hat{s}$;

(ii) $(\operatorname{div} \boldsymbol{\tau}) \circ F_K = (\det \mathsf{B}_K)^{-1} \widehat{\operatorname{div}} \, \hat{\boldsymbol{\tau}}$;

(iii) the Piola image of $RT_k(\hat{K})$ is $RT_k(K)$; bases and estimates are therefore constructed on the reference element.

Orientation. The dof $\int_e \boldsymbol{\tau} \cdot \boldsymbol{n}_e$ requires one global normal per edge; the two elements sharing e have opposite outward normals, so exactly one of the two local basis functions carries the factor -1. Libraries derive the sign from a global rule (for instance the low-to-high vertex direction); hand-written assembly must apply the same rule.

A wrong sign appears correct on one mesh and fails after a renumbering.

The global space; the constraint matrix, again

$$RT_k(\mathcal{T}_h) := \{\boldsymbol{\tau} : \ \boldsymbol{\tau}|_K \in RT_k(K), \ \text{shared edge dofs coincide}\},$$

paired with $Q_h := P_k^{\mathrm{disc}}(\mathcal{T}_h)$.

Corollary 20. $RT_k(\mathcal{T}_h) \subset H(\mathrm{div};\Omega)$: the normal trace on e lies in $P_k(e)$ and is determined by the shared moments, hence single-valued; Lemma 15 applies. □

Dimensions ($d = 2$): $\dim V_h = (k{+}1)\#E + k(k{+}1)\#T$, $\dim Q_h = \frac{1}{2}(k{+}1)(k{+}2)\#T$; for $k = 0$: $\#E$ and $\#T$.

The constraint matrix for $k = 0$. With the flux degrees of freedom, $\int_K \mathrm{div}\, \boldsymbol{\varphi}_e = s_{K,e}$, so $\mathsf{B}_{K,e} = \pm 1$ for the three edges of K and 0 otherwise: a signed incidence matrix, whose rows state that the signed edge fluxes of each element sum to $-\int_K f$, exactly.

This is the two-dimensional descendant of the matrix of Lecture 1, frame 9: the fundamental theorem of calculus has become the divergence theorem.

The pair of Lecture 1, identified

In one dimension, $\mathrm{div} = d/dx$, so $H(\mathrm{div};\Omega) = H^1(\Omega)$; the elements are intervals, the "edges" are their endpoints. Hence $RT_0(K) = P_0 + xP_0 = P_1(K)$, the degrees of freedom are the endpoint values, and the assembled conforming space is P_1^{cont}.

The pair that opened Lecture 1 is RT_0–P_0.

- The nodal increments of Lecture 1 (frame 9) are the ± 1 incidence rows; the Fortin operator of Lecture 2 is the canonical interpolant defined below.
- In one dimension the following families coincide: $RT_k(K) = P_k + x\widetilde{P}_k = P_{k+1}(K) = BDM_{k+1}(K)$, with the same boundary degrees of freedom. They separate for $d \geq 2$.

Warning (Firedrake / FEniCS indexing). `FunctionSpace(mesh, "RT", 1)` is RT_0 in the indexing of this course (Boffi–Brezzi–Fortin): the library index is shifted by one. The BDM indices agree: `("BDM", 1)` is BDM_1; cross-check on DefElement.

BDM_k, and the resolution of the 1D superconvergence

$BDM_k(K) := P_k(K)^2$ (complete polynomials), $k \geq 1$; $\dim = (k+1)(k+2)$; $\operatorname{div} BDM_k(K) = P_{k-1}(K)$. Dofs: the edge moments against $P_k(e)$ (all of them for $k = 1$; interior moments for $k \geq 2$: BBF §2.3); the natural pairing is BDM_k–P_{k-1}^{disc}.

L^2 errors of	$\boldsymbol{\sigma}$	$\operatorname{div} \boldsymbol{\sigma}$	u	$P_h u - u_h$
RT_0–P_0	h	h	h	h^2
BDM_1–P_0	h^2	h	h	h^2
Lecture 1, measured (1D)	2.00	1.00	1.00	2.00

The measured one-dimensional profile is the BDM_1–P_0 profile. Of the two phenomena of Lecture 1, frame 12: nodal exactness is specific to one dimension; the $O(h^2)$ flux rate survives, and its rate-preserving heir for $d \geq 2$ is BDM_1 (two dofs per edge), while RT_0 is the cost-preserving heir. The divergence columns are equal — explained by Theorem 25. Measured confirmation: backup B3.

The canonical interpolant

$\Pi_h^{RT}\boldsymbol{\tau} \in V_h$ is defined by matching all degrees of freedom of $\boldsymbol{\tau}$:

$$\int_e (\Pi_h^{RT}\boldsymbol{\tau}\cdot\boldsymbol{n})\, q\, ds = \int_e (\boldsymbol{\tau}\cdot\boldsymbol{n})\, q\, ds, \quad \int_K \Pi_h^{RT}\boldsymbol{\tau}\cdot\boldsymbol{q}\, dx = \int_K \boldsymbol{\tau}\cdot\boldsymbol{q}\, dx.$$

Domain of definition. For $\boldsymbol{\tau} \in H(\mathrm{div};\Omega)$ the edge moments are *not defined*: the normal trace lies in $H^{-1/2}(\partial\Omega)$ only, and by the non-restriction property of slide 5 it admits no restriction to a single edge. The operator Π_h^{RT} is defined — and, as shown below, bounded uniformly in h — on $(H^1(\Omega))^d$.

Π_h^{RT} is not bounded on $H(\mathrm{div};\Omega)$.
This is precisely the situation for which Lecture 2 proved the restricted-domain variant (InfSup$_W$) of Fortin's criterion.

The commuting diagram

Theorem 21. For every $\boldsymbol{\tau} \in (H^1(\Omega))^d$: $\quad \operatorname{div} \Pi_h^{RT} \boldsymbol{\tau} = P_h \operatorname{div} \boldsymbol{\tau}$.
Proof. Both sides belong to Q_h. For $q_h \in Q_h$ and $K \in \mathcal{T}_h$:

$$\begin{aligned}\int_K \operatorname{div}(\Pi_h \boldsymbol{\tau})\, q_h &= -\int_K \Pi_h \boldsymbol{\tau} \cdot \nabla q_h + \int_{\partial K} (\Pi_h \boldsymbol{\tau} \cdot \boldsymbol{n})\, q_h \\ &= -\int_K \boldsymbol{\tau} \cdot \nabla q_h + \int_{\partial K} (\boldsymbol{\tau} \cdot \boldsymbol{n})\, q_h = \int_K (P_h \operatorname{div} \boldsymbol{\tau})\, q_h,\end{aligned}$$

using the interior dofs ($\nabla q_h \in P_{k-1}(K)^2$), the edge dofs ($q_h|_e \in P_k(e)$), and integration by parts backwards. $\square$

For $k = 0$, $d = 1$, this is the argument of Lecture 1, frame 10.

Corollary 22. (i) $\operatorname{div} V_h = Q_h$ ($\subset$: Theorem 18; $\supset$: on convex polytopes the Poisson lifting of Lecture 2, slide 14, lies in $(H^1)^d$, and $\operatorname{div} \Pi_h \boldsymbol{\tau} = P_h q_h = q_h$); (ii) $K_h \subset K$: for $\boldsymbol{v}_h \in K_h$, $\operatorname{div} \boldsymbol{v}_h \in Q_h$ and $\operatorname{div} \boldsymbol{v}_h \perp Q_h$, hence $\operatorname{div} \boldsymbol{v}_h = 0$.

One level up: the de Rham complex

On simply connected $\Omega \subset \mathbb{R}^2$, rows are exact and squares commute:

$$\begin{array}{ccccccccc} \mathbb{R} & \to & H^1(\Omega) & \xrightarrow{\mathbf{curl}} & H(\mathrm{div};\Omega) & \xrightarrow{\mathrm{div}} & L^2(\Omega) & \to & 0 \\ & & \big\downarrow I_h & & \big\downarrow \Pi_h^{RT} & & \big\downarrow P_h & & \\ \mathbb{R} & \to & P_{k+1}^{\mathrm{cont}}(\mathcal{T}_h) & \xrightarrow{\mathbf{curl}} & RT_k(\mathcal{T}_h) & \xrightarrow{\mathrm{div}} & P_k^{\mathrm{disc}}(\mathcal{T}_h) & \to & 0 \end{array}$$

- Exactness at the last slot: top row = (InfSup), proved in Lecture 2 by the Poisson lifting; bottom row = Corollary 22. This lecture constructed a subcomplex that remains exact.
- Exactness at the middle slot was used in step 3 of Theorem 18. The right square is Theorem 21; the left square ($k = 0$) holds because $\int_e \mathbf{curl}\,\varphi \cdot \boldsymbol{n}\, ds = \varphi(b) - \varphi(a)$ depends only on vertex values.
- In $d = 3$: $H^1 \xrightarrow{\nabla} H(\mathbf{curl}) \xrightarrow{\mathbf{curl}} H(\mathrm{div}) \xrightarrow{\mathrm{div}} L^2$, discretized by Lagrange $\to$ **Nédélec** $\to$ Raviart–Thomas $\to$ discontinuous elements.

The systematic theory is *finite element exterior calculus*: Arnold–Falk–Winther, *Acta Numerica* 2006; Arnold, SIAM 2018. None of this is developed further here; the outlook of Lecture 4 returns to it with pointers.

(ElKer$_h$): inherited from the continuous problem

For $\boldsymbol{v}_h \in K_h$, Corollary 22 gives $\operatorname{div} \boldsymbol{v}_h = 0$, hence $\|\boldsymbol{v}_h\|_{\mathrm{div}} = \|\boldsymbol{v}_h\|_0$ and

$$a(\boldsymbol{v}_h, \boldsymbol{v}_h) = \|\kappa^{-1/2}\boldsymbol{v}_h\|_0^2 \;\geq\; \frac{1}{\kappa_+}\,\|\boldsymbol{v}_h\|_{\mathrm{div}}^2.$$

- (ElKer$_h$) holds with $\alpha_h \geq 1/\kappa_+ = \alpha$: because $K_h \subset K$, the discrete constant is not smaller than the continuous one, uniformly in h and independently of the mesh.
- Contrast with P_2–P_0 in Lecture 1: there $K_h \not\subset K$, and $\alpha_h = h^2/60$. The present construction removes the failure mode by removing the spurious kernel functions.
- Measured (Demo 1): for $\kappa = 1$ the Rayleigh quotient equals 1 on all of K_h, to machine precision.

(InfSup$_h$) via the restricted Fortin criterion

Corollary 12 (Lecture 2): if (InfSup$_W$) holds on $W \subset V$ with $\|\cdot\|_W$, and $\Pi_h : W \to V_h$ satisfies (F1$_W$) and (F2$_W$), then (InfSup$_h$) holds with $\beta_h \geq \beta_W / C_\Pi$. Here $W = (H^1(\Omega))^d$:

- **(F1$_W$)** $b(\boldsymbol{\tau} - \Pi_h \boldsymbol{\tau}, q_h) = (\operatorname{div} \boldsymbol{\tau} - P_h \operatorname{div} \boldsymbol{\tau},\, q_h) = 0$: *the commuting property is* (F1$_W$).
- **(F2$_W$)** $\|\Pi_h \boldsymbol{\tau}\|_{\mathrm{div}} \leq C \|\boldsymbol{\tau}\|_1$, $C = C$(shape regularity).
  Sketch: on the reference element the dofs are bounded on $(H^1)^2$ by the trace theorem; the Piola scaling gives $\|\Pi_K \boldsymbol{\tau}\|_{0,K} \leq C(\|\boldsymbol{\tau}\|_{0,K} + h_K |\boldsymbol{\tau}|_{1,K})$; and $\|\operatorname{div} \Pi_h \boldsymbol{\tau}\|_0 = \|P_h \operatorname{div} \boldsymbol{\tau}\|_0 \leq \|\operatorname{div} \boldsymbol{\tau}\|_0 \leq \sqrt{d}\, |\boldsymbol{\tau}|_1$.
- **(InfSup$_W$)** As delivered in Lecture 2, slide 29: on a convex polytope the Poisson lifting of $q \in L^2$ lies in $(H^1)^d$ with $\|\boldsymbol{\tau}\|_1 \leq C_\Omega \|q\|_0$ (elliptic regularity, stated without proof).

(InfSup$_h$) holds with $\beta_h \geq \beta_W / C_\Pi > 0$, independent of h

(uniform positivity, not the value; the measured value follows on Demo 1b).

Error estimates

Theorem 23. For RT_k–P_k^{disc} on shape-regular meshes, with $\boldsymbol{\sigma} \in H^{k+1}(\Omega)^d$, $\operatorname{div} \boldsymbol{\sigma} \in H^{k+1}(\Omega)$, $u \in H^{k+1}(\Omega)$:

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathrm{div}} + \|u - u_h\|_0 \;\le\; C\, h^{k+1}\big(|\boldsymbol{\sigma}|_{k+1} + |\operatorname{div} \boldsymbol{\sigma}|_{k+1} + |u|_{k+1}\big).$$

Proof: Theorem 7 of Lecture 2 (constants α_h, β_h are h-uniform), combined with the interpolation estimates

$$\|\boldsymbol{\tau} - \Pi_h \boldsymbol{\tau}\|_0 \le C h^m |\boldsymbol{\tau}|_m, \;\; \|v - P_h v\|_0 \le C h^m |v|_m, \;\; 1 \le m \le k+1$$

(stated; proofs by Piola scaling: BBF §2.5). No separate estimate is needed for the divergence: $\operatorname{div}(\boldsymbol{\sigma} - \Pi_h \boldsymbol{\sigma}) = (I - P_h) \operatorname{div} \boldsymbol{\sigma}$ — the origin of the middle term. □

- For $k = 0$: rates $1/1/1$ for $\boldsymbol{\sigma}$, $\operatorname{div} \boldsymbol{\sigma}$, u — the rate announced in Lecture 1, frame 12, as the honest higher-dimensional statement.
- The next two slides sharpen this in two directions specific to the structure $\operatorname{div} V_h = Q_h$.

Local conservation, proved

Proposition 24. $\operatorname{div} \boldsymbol{\sigma}_h = -P_h f$, exactly, at every h.

Proof. $\operatorname{div} \boldsymbol{\sigma}_h \in Q_h$ by Corollary 22; the second discrete equation states $(\operatorname{div} \boldsymbol{\sigma}_h + f, v_h) = 0$ for all $v_h \in Q_h$; these two facts identify $\operatorname{div} \boldsymbol{\sigma}_h$ as $-P_h f$. □

Corollary. With the Darcy velocity $\boldsymbol{q}_h := -\boldsymbol{\sigma}_h$:

$$\int_{\partial K} \boldsymbol{q}_h \cdot \boldsymbol{n} \, ds = \int_K f \, dx \quad \text{on every element,}$$

the normal flux is single-valued across interior edges, and by summation the same balance holds on every union of elements (backup B4). This is the property promised in Lecture 1, frame 7.

The Lagrange comparison. For P_1 elements, $\boldsymbol{q}_h = -\kappa \nabla u_h$ is elementwise constant, so $\int_{\partial K} \boldsymbol{q}_h \cdot \boldsymbol{n} \, ds = 0$ on *every* element: the elementwise defect equals $|\int_K f|$ identically — the two-dimensional instance of Exercise 3, Lecture 1. A weaker balance on vertex patches survives (test with the hat functions); no single-valued flux exists.

The flux error does not see the scalar space

Theorem 25. If $\boldsymbol{\sigma} \in (H^1(\Omega))^d$, then

$$\|\kappa^{-1/2}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 \;\le\; \|\kappa^{-1/2}(\boldsymbol{\sigma} - \Pi_h^{RT}\boldsymbol{\sigma})\|_0 :$$

the constant equals 1, and no term in u appears.

Proof. Set $\boldsymbol{e}_h := \boldsymbol{\sigma}_h - \Pi_h\boldsymbol{\sigma} \in V_h$. Then $\operatorname{div}\boldsymbol{e}_h = -P_h f - P_h \operatorname{div}\boldsymbol{\sigma} = 0$, since $\operatorname{div}\boldsymbol{\sigma} = -f$ (Proposition 24 and Theorem 21); hence $\boldsymbol{e}_h \in K_h$, and in fact $\boldsymbol{e}_h \in K$. The first error equation gives $a(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, u - u_h) = 0$ for all $\boldsymbol{\tau}_h \in V_h$; the choice $\boldsymbol{\tau}_h = \boldsymbol{e}_h$ removes the second term, because $b(\boldsymbol{e}_h, u - u_h) = (\operatorname{div}\boldsymbol{e}_h,\, u - u_h) = 0$. Therefore $a(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{e}_h) = 0$, so that $\|\kappa^{-1/2}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0^2 = a(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma} - \Pi_h\boldsymbol{\sigma})$, and the Cauchy–Schwarz inequality concludes. □

- The proof uses only $\operatorname{div} V_h \subset Q_h$ and the commuting interpolant: it applies verbatim to BDM_k–P_{k-1} — the justification of the profile table.
- Likewise $\|\operatorname{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_0 = \|\operatorname{div}\boldsymbol{\sigma} - P_h \operatorname{div}\boldsymbol{\sigma}\|_0$: the divergence error *equals* its best approximation whenever $\operatorname{div} V_h = Q_h$.

Superconvergence of the scalar means

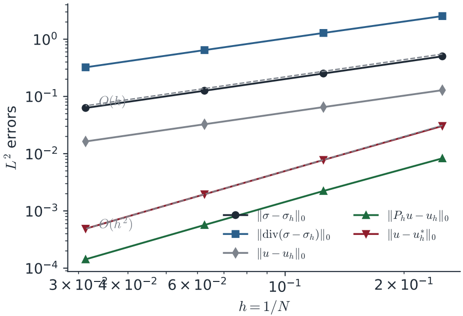
Theorem 26 (stated; duality argument: Douglas–Roberts 1985; BBF §7.4). On a convex polytope, for smooth data,

$$\|P_h u - u_h\|_0 \;\le\; C\,h^{k+2} \qquad (k \ge 0;\ \text{for } k = 0:\ O(h^2)).$$

- Reading: u_h is a first-order approximation of u, but a *second*-order approximation of the element means $P_h u$ — the error $u - u_h$ is dominated by the projection error $u - P_h u$, which carries no information about the discretization.
- Compare Lecture 1: for P_2–P_0 the scalar converged to the wrong limit $u/6$; here it converges to the right limit, and to the element means at a higher rate.
- Consequence: the accuracy stored in $\boldsymbol{\sigma}_h$ and in the means can be extracted by a local postprocessing — slide 29.

Measured (Demo 1a): rate 1.99 at $N = 32$.

Demonstration 1a: convergence of RT_0–P_0



$u = \sin(\pi x)\sin(\pi y)$, measured rates at $N = 32$:

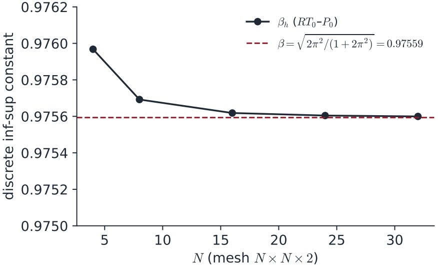
quantity	rate
$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0$	1.00
$\|\mathrm{div}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_0$	1.00
$\|u-u_h\|_0$	1.00
$\|P_h u-u_h\|_0$	1.99
$\|u-u_h^*\|_0$	2.00

u_h^*: the postprocessed scalar of slide 29.

The divergence error equals $\|f - P_h f\|_0$ to all digits, for RT_0 and for BDM_1 alike (Theorem 25).

▶ **live demonstration:** notebook `L3A`, cell 4

Demonstration 1b: the constants, measured



- β_h flat: 0.97597 at $N{=}4$, decreasing to 0.97560 at $N{=}32$, *from above*, toward

$$\beta = \sqrt{\tfrac{2\pi^2}{1+2\pi^2}} = 0.9755932.$$

- $\alpha_h = 1$ to machine precision on all of K_h ($\kappa = 1$); $\dim K_h = \#E - \#T = N^2{+}2N$.

In Lecture 2 the same computation in one dimension gave $\beta = \pi/\sqrt{1+\pi^2}$. Both values are instances of

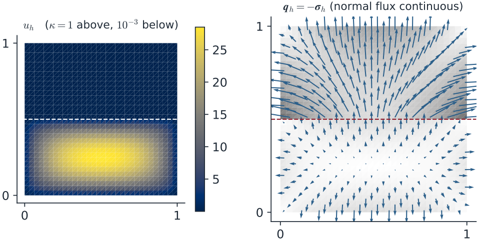
$$\beta = \sqrt{\mu_1/(1+\mu_1)}, \qquad \mu_1 = \text{first Dirichlet eigenvalue of } \Omega$$

(Exercise 4, starred — the two-dimensional continuation of Lecture 2, Exercise 2(e*)).

▶ **live demonstration:** `L3A`, cell 5

Demonstration 1c: conservation; a discontinuous coefficient

solve (N=32)	defect ($\max_K$)	$\max_K \lvert \int_K f \rvert$	
RT_0–P_0	$1.3 \cdot 10^{-17}$	$9.6 \cdot 10^{-3}$	Prop. 24
P_1 Lagrange	$9.6 \cdot 10^{-3}$	$9.6 \cdot 10^{-3}$	$\equiv \lvert \int_K f \rvert$
RT_0–P_0, layered κ	$3.8 \cdot 10^{-16}$	$4.9 \cdot 10^{-4}$	κ-independent



Layered κ (1 above, 10^{-3} below): the normal flux is single-valued across the interface by construction; the constraint rows are κ-independent.

▶ **live demonstration:** `L3A`, cells 7, 9

Solving the system: why hybridization

The assembled system is symmetric indefinite,

$$\begin{pmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{pmatrix}, \qquad \text{size } \dim V_h + \dim Q_h \quad (N{=}64:\ 20608),$$

and the standard machinery for symmetric positive definite systems does not apply to it directly (iterative methods for the indefinite form: Lecture 4).

Observation. The only coupling between neighbouring elements enters through the continuity of the normal flux built into V_h.

Idea (Fraeijs de Veubeke 1965; Arnold–Brezzi 1985): remove the continuity from the space and restore it as an explicit constraint, with a Lagrange multiplier on the interior edges. All flux and scalar unknowns then become element-local, and can be eliminated element by element.

The hybridized formulation

Spaces: $V_h^{\mathrm{b}} := \{\boldsymbol{\tau} : \boldsymbol{\tau}|_K \in RT_k(K)\}$ (no continuity), Q_h as before, $\Lambda_h := \{\mu : \mu|_e \in P_k(e),\ e \text{ interior}\}$; find $(\boldsymbol{\sigma}_h, u_h, \lambda_h)$:

$$
\begin{aligned}
a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) + \textstyle\sum_K \big[(\operatorname{div} \boldsymbol{\tau}, u_h)_K - \langle \boldsymbol{\tau}\cdot \boldsymbol{n}_K, \lambda_h\rangle_{\partial K \setminus \partial\Omega}\big] &= 0 && \forall\, \boldsymbol{\tau} \in V_h^{\mathrm{b}},\\
\textstyle\sum_K (\operatorname{div} \boldsymbol{\sigma}_h,\, v)_K &= -(f, v) && \forall\, v \in Q_h,\\
\textstyle\sum_K \langle \boldsymbol{\sigma}_h \cdot \boldsymbol{n}_K,\, \mu\rangle_{\partial K} &= 0 && \forall\, \mu \in \Lambda_h.
\end{aligned}
$$

Theorem 27. The problem is uniquely solvable; $(\boldsymbol{\sigma}_h, u_h)$ coincides with the conforming solution, and λ_h enters each element problem in the position of the Dirichlet datum.

Proof (sketch). (a) The third equation states that the normal fluxes are single-valued, so $\boldsymbol{\sigma}_h \in V_h$. (b) For conforming $\boldsymbol{\tau}$ the multiplier terms cancel edge by edge, leaving the conforming problem, which determines $(\boldsymbol{\sigma}_h, u_h)$. (c) The first equation on one element determines all moments $\langle \boldsymbol{\tau}\cdot\boldsymbol{n}, \lambda_h\rangle_e$, and the edge dofs of $RT_k(K)$ realize every moment against $P_k(e)$: λ_h exists and is unique, edge by edge. ☐ Verified: broken and conforming solutions agree to $9.7 \cdot 10^{-15}$ at $N = 8$.

Static condensation; the condensed system is SPD

Ordered by element, the $(\boldsymbol{\sigma}, u)$-block is block diagonal: one 4×4 block per triangle for $k = 0$, each invertible (Lemma 1 of Lecture 1, applied on one element). Eliminating elementwise leaves $\mathsf{S}\,\boldsymbol{\lambda} = \mathbf{r}$ with S the Schur complement of the element blocks.

Proposition 28. S is symmetric positive definite. *Sketch:* $\boldsymbol{\mu}^T \mathsf{S} \boldsymbol{\mu} = a(\boldsymbol{\sigma}(\mu), \boldsymbol{\sigma}(\mu))$, where $\boldsymbol{\sigma}(\mu)$ solves the local problems with datum μ and $f = 0$; it vanishes only if every local response vanishes, and this forces μ to equal the local scalar on each element and zero at the boundary, hence $\mu = 0$. $\square$

$N = 64$	unknowns	symmetry	solver
mixed system	20608	indefinite	factorization / Lecture 4
condensed system	12160	positive definite	CG, multigrid

Verified at $N{=}8$: $\|\mathsf{S} - \mathsf{S}^T\|_{\max} = 4 \cdot 10^{-16}$ and $\lambda_{\min}(\mathsf{S}) = 0.102 > 0$; the condensed and the full hybrid solve agree to $4.3 \cdot 10^{-15}$.

The postprocessed scalar u_h^*

On each element ($k = 0$): define $u_h^*|_K \in P_1(K)$ by

$$\nabla u_h^*|_K = \text{elementwise mean of } \kappa^{-1}\boldsymbol{\sigma}_h, \qquad \int_K u_h^* = |K|\, u_h|_K.$$

Both conditions are local; the construction is elementwise, independent, and performed after the global solve. (General k: an elementwise P_{k+1} problem driven by $\kappa^{-1}\boldsymbol{\sigma}_h$ and the mean; Arnold–Brezzi 1985, Stenberg 1991.)

Proposition 29 (stated). On a convex polytope, for smooth data: $\|u - u_h^*\|_0 = O(h^{k+2})$; for $k = 0$: $O(h^2)$.

Mechanism: the flux carries the gradient information at first order; the mean carries the location information at second order (Theorem 26); the local problem combines the two.

The scalar accuracy of a richer pair, at the cost of the lowest-order one.

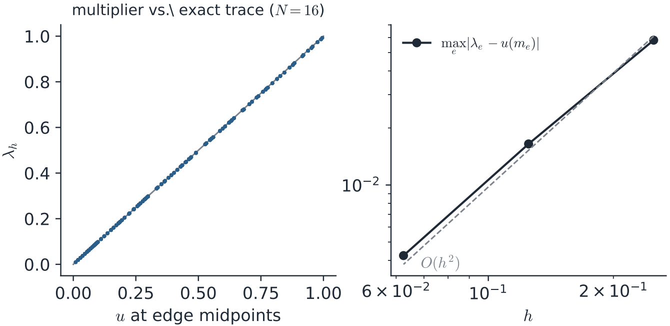
Measured: rate 2.00; $\|u - u_h^*\|_0 = 4.9 \cdot 10^{-4}$ at $N = 32$ (Demo 2b).

Demonstration 2a: hybridization in Firedrake

- **live demonstration:** notebook `L3B`, cells 2 and 4
  - Monolithic solve of the mixed system at $N = 64$ (20608 unknowns, direct factorization), then the same problem through `HybridizationPC`: Firedrake assembles the broken problem, condenses, and solves the trace system (12160 unknowns, symmetric positive definite) by the conjugate gradient method; iteration counts and timings are printed live.
  - Consistency check on screen: $\|\boldsymbol{\sigma}_h^{\text{mono}} - \boldsymbol{\sigma}_h^{\text{hyb}}\|_0$ at the solver tolerance — the two routes compute the same discrete solution (Theorem 27).
  - The back-substitution recovering $(\boldsymbol{\sigma}_h, u_h)$ from λ_h is elementwise and independent.

No timings are printed on this slide: they depend on the machine and are produced live by the notebook.

Demonstration 2b: the multiplier, and u_h^*



Measured: $\max_e |\lambda_e - u(m_e)| = 5.8\cdot10^{-2},\ 1.6\cdot10^{-2},\ 4.2\cdot10^{-3}$ at $N = 4, 8, 16$ — rates 1.82, 1.96: the multiplier is a *second-order* approximation of u at the edge midpoints, not a bookkeeping device. (For general k it feeds the postprocessing.)

The postprocessed u_h^* converges at rate 2.00 (table on slide 23). ▶

live demonstration: L3B, cells 6, 8

The pair of Lecture 1, fully explained

observed in Lecture 1 (1D)	explained in this lecture
constraint rows exact (frame 9)	B $=$ signed incidence matrix (slide 11)
nodal exactness of σ_h	specific to $d = 1$ (slide 13)
flux rate $O(h^2)$	BDM_1 profile; families coincide in 1D (slide 13)
rates $1/1$ for σ', u	Theorem 23 with $k = 0$
midpoint rate $O(h^2)$ for u_h	Theorem 26 ($P_h u$ superconvergence)
convergence, no further conditions	(ElKer$_h$), (InfSup$_h$) by construction

The identification promised in Lecture 1 (frames 12 and 34) is complete: the pair is RT_0–P_0, its generalization to $d \geq 2$ has been constructed, analyzed, and turned into a locally conservative solver.

What transfers to Stokes, and what does not

The pipeline of this lecture: conforming subspace $\to$ local degrees of freedom $\to$ commuting interpolant $\to$ (F1$_W$), (F2$_W$) $\to$ uniform stability.

- **Transfers:** the Fortin machinery of Lecture 2, the saddle solver question (avoided here by hybridization, unavoidable for Stokes), and the habit of measuring β_h.
- **Does not transfer:** for Stokes, $V = (H_0^1(\Omega))^d$: conformity requires *full* vector continuity. No standard pair achieves $\operatorname{div} V_h = Q_h$ together with an H^1-bounded commuting interpolant (constructions that do exist are referenced in the outlook).
- **Consequently:** Lecture 4 verifies (InfSup$_h$) pair by pair, by Fortin operators assembled from local corrections — the composition lemma (Lemma 13) of Lecture 2 was proved for this purpose.

For the mixed Laplacian, stability was obtained by construction.
For Stokes, the inf-sup condition itself becomes the construction problem.

Exercise sheet 3, and outlook

1. element calculus: the RT_0 dual basis, the incidence structure of B, one Piola identity by hand, the orientation sign rule
2. the commuting diagram in one dimension (the operator of Lecture 2, Problem 2, identified); counting exactness of the discrete complex by Euler's formula (starred)
3. conservation in practice (notebook): patch balances, layered coefficients, the Lagrange comparison
4. hybridization on two triangles, by hand: the 1×1 condensed system; (starred) the constant $\beta = \sqrt{\mu_1/(1+\mu_1)}$

Reading: Boffi–Brezzi–Fortin, Ch. 2 and 7; Gatica, Ch. 3; Arnold–Falk–Winther for the complex.

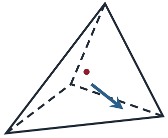
Lecture 4: Stokes — stable pairs, the solver question, and a map of what lies beyond this course.

Backup slides

Backup: RT_0 on a tetrahedron

$RT_0(K) = P_0(K)^3 + \boldsymbol{x}\,P_0(K)$, $\dim = 4$; degrees of freedom: the face fluxes $N_F(\boldsymbol{\tau}) = \int_F \boldsymbol{\tau}\cdot\boldsymbol{n}_F\,dA$.

- $\mathrm{div}(\boldsymbol{a} + c\boldsymbol{x}) = 3c$: onto $P_0(K)$; $\boldsymbol{x}\cdot\boldsymbol{n}$ is constant on each face plane, so the normal trace is facewise constant.
- Unisolvence as in Proposition 17, with $2|K|$ replaced by $3|K|$ and edges by faces; dual basis $\boldsymbol{\varphi}_F = \frac{s_{K,F}}{3|K|}(\boldsymbol{x} - \boldsymbol{p}_F)$.
- Global orientation: one fixed normal per *face*; the sign rule is identical.



In three dimensions the pair RT_0–P_0 has one flux unknown per face and one scalar per cell; the staggered-grid reading of slide 7 persists.

Backup: Theorem 18, details

Dimension bookkeeping ($d = 2$): $\dim P_k^2 = (k{+}1)(k{+}2)$;
$\dim \boldsymbol{x}\widetilde{P}_k = k{+}1$ (direct sum: $\boldsymbol{x}\,\widetilde{p} \in P_k^2$ forces $\widetilde{p} = 0$); total $(k{+}1)(k{+}3)$.
Dofs: $3(k{+}1)$ edge + $k(k{+}1)$ interior = $(k{+}1)(k{+}3)$.

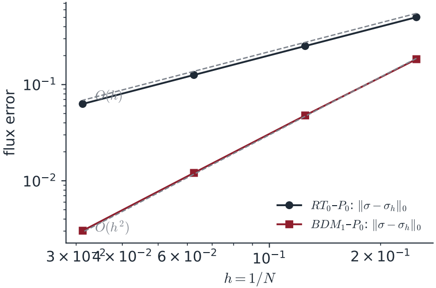
$\operatorname{div} RT_k = P_k$**:** $\operatorname{div} P_k^2 = P_{k-1}$, and for homogeneous $\widetilde{p}$ of degree k, Euler's identity gives $\operatorname{div}(\boldsymbol{x}\widetilde{p}) = (d+k)\widetilde{p}$; together they span P_k.

Step 3 in full: $\operatorname{div} \boldsymbol{\tau} = 0$ on the (simply connected) simplex implies $\boldsymbol{\tau} = \mathbf{curl}\, \psi$ with $\psi \in P_{k+1}$ (integrate the rotated field); $\boldsymbol{\tau} \cdot \boldsymbol{n} = \partial_t \psi$; a function whose tangential derivative vanishes on the connected boundary is constant there; normalizing, ψ vanishes on ∂K, hence $\psi = b_K r$, $r \in P_{k-2}$, with b_K the cubic bubble.

Step 4 in full: for $\boldsymbol{q} \in P_{k-1}^2$, $\int_K \mathbf{curl}\, \psi \cdot \boldsymbol{q} = \int_K \psi \operatorname{rot} \boldsymbol{q}$ (the boundary term carries $\psi|_{\partial K} = 0$), and $\operatorname{rot} : P_{k-1}^2 \to P_{k-2}$ is onto (choose $\boldsymbol{q} = (0, Q)$ with $\partial_1 Q = r$). Hence $\int_K b_K r^2 = 0$ and $r = 0$.

$d = 3$**:** faces carry moments against $P_k(F)$, the interior against $P_{k-1}(K)^3$; the representation in step 3 uses the Nédélec curl (BBF §2.3–2.5).

Backup: BDM_1 versus RT_0, measured



Same data, same meshes; measured rates at $N = 32$:

rate of	RT_0–P_0	BDM_1–P_0
$\boldsymbol{\sigma}$	1.00	2.00
$\operatorname{div}\boldsymbol{\sigma}$	1.00	1.00
u	1.00	1.00
$P_h u - u_h$	1.99	1.99

The divergence errors of the two pairs agree to all computed digits — $2.5348, 1.2857, 0.6452, 0.3229$ for $N = 4, \dots, 32$ — since both equal $\|f - P_h f\|_0$ (Theorem 25). The one-dimensional profile $2.00/1.00/1.00/2.00$ of Lecture 1 is the BDM_1 row.

▶ **live demonstration:** `L3A`, cell 11

Backup: balances on unions of elements

For any union ω of elements, summing Proposition 24 over $K \subset \omega$ and using the single-valuedness of $\boldsymbol{q}_h \cdot \boldsymbol{n}$ across interior edges (the contributions of shared edges cancel in pairs):

$$\int_{\partial\omega} \boldsymbol{q}_h \cdot \boldsymbol{n}\, ds = \int_{\omega} f\, dx \qquad \text{exactly, for every } \omega.$$

- The identity is algebraic: it holds at the level of the incidence matrix, independently of κ, the mesh, and the data.
- Notebook `L3A`, cell 13, verifies it on randomly chosen unions of elements (residuals at machine precision) — the discrete counterpart of the mass balance a reservoir engineer checks on a control volume.
- The Lagrange solution admits no such statement on arbitrary unions: no single-valued flux exists to integrate.

Backup: static condensation in code (Slate)

The self-study cell L3B, cell 8, assembles the three-field problem and condenses it explicitly; the essential lines:

```
el = FiniteElement("RT", triangle, 1)      # RT0 in course indexing
Vb = FunctionSpace(mesh, BrokenElement(el))
Q  = FunctionSpace(mesh, "DG", 0)
L  = FunctionSpace(mesh, "HDiv Trace", 0)
W  = Vb * Q * L
sigma, u, lam = TrialFunctions(W)
tau, v, mu    = TestFunctions(W)
n  = FacetNormal(mesh)
a  = (dot(sigma, tau) + div(tau)*u + div(sigma)*v)*dx
a -= (jump(tau, n)*lam('+') + jump(sigma, n)*mu('+'))*dS
A  = Tensor(a);  F = Tensor(-f*v*dx)
A00, A01 = A.blocks[:2, :2], A.blocks[:2, 2]
S  = A.blocks[2, 2] - A.blocks[2, :2] * A00.inv * A01
```

- `jump(tau, n)` realizes $[\![\boldsymbol{\tau} \cdot \boldsymbol{n}]\!]$; the `dS` terms are the multiplier equations of slide 27. `A00.inv` is the elementwise elimination: the block is cell-local, so the inverse is computed element by element at assembly.
- The condensed `S` is the matrix of Proposition 28; the notebook verifies symmetry, positive definiteness, and agreement with `HybridizationPC`.