7 Optimal Control Problems

We discuss the use of the DWR method in the finite element discretization of optimal control problems. As examples, we consider boundary control in heat transfer and drag minimization in viscous flow. Optimal control problem with state space V and control space Q:

****

Abstract optimal control problem

$$J(u,q) = \min! \qquad A(u,q)(\psi) = 0 \quad \forall \psi \in V$$

Galerkin approximation in subspaces $V_h \times Q_h \subset V \times Q$:

$$J(u_h,q_h) = \min! \qquad A(u_h,q_h)(\psi_h) = 0 \quad \forall \psi_h \in V_h$$

Preliminary thoughts:

- *Notion of admissible states* $u = u(q)$ *?*
  Discretization introduces perturbation of solution operator.
  Accuracy in discretization of PDEs is expensive.
  To what extent is "admissibility" relevant for the optimization?

- *How to "measure" admissibility:*
  In ODEs, we may require the error to be uniformly "small".
  In PDEs, the choice of error measures is less clear: "energy" norm, L^2-norm, local max-norm, ... ?

Euler-Lagrange approach and Galerkin approximation:

Lagrangian functional $\quad \mathcal{L}(u,q,z) := J(u,q) - A(u,q)(z)$

(P) Determine stationary point $x := \{u,q,z\} \in X := V \times Q \times V$:

$$\left\{ \begin{array}{l} J'_u(u,q)(\varphi) - A'_u(u,q)(\varphi,z) \\ J'_q(u,q)(\chi) - A'_q(u,q)(\chi,z) \\ -A(u,q)(\psi) \end{array} \right\} = 0 \qquad \forall\{\varphi,\chi,\psi\}$$

(P_h) Galerkin approximation $x_h := \{u_h,q_h,z_h\} \in X_h := V_h \times Q_h \times V_h$:

$$\left\{ \begin{array}{l} J'_u(u_h,q_h)(\varphi_h) - A'_u(u_h,q_h)(\varphi_h,z_h) \\ J'_q(u_h,q_h)(\chi_h) - A'_q(u_h,q_h)(\chi_h,z_h) \\ -A(u_h)(\psi) \end{array} \right\} = 0 \qquad \forall\{\varphi_h,\chi_h,\psi_h\}$$

First idea for error control: Measure accuracy in terms of cost functional $J(u)-J(u_h)$ depending on residuals of $x_h := (u_h,q_h,z_h)$:

$$\begin{aligned} \rho^*(x_h)(\cdot) &:= J'_u(u_h,q_h)(\cdot) - A'_u(u_h,q_h)(\cdot,z_h) && \text{(dual)} \\ \rho^q(x_h)(\cdot) &:= J'_q(u_h,q_h)(\cdot) - A'_q(u_h,q_h)(\cdot,z_h) && \text{(control)} \\ \rho(x_h)(\cdot) &:= -A(u_h)(\cdot) && \text{(primal)} \end{aligned}$$

Proposition. *We have the a posteriori error representation*

$$\begin{aligned} J(u,q)-J(u_h,q_h) \;=\; & \frac{1}{2}\underbrace{\rho^*(z_h)(u-\varphi_h)}_{\textit{dual residual}} + \frac{1}{2}\underbrace{\rho^q(q_h)(q-\chi_h)}_{\textit{control residual}} \\ & + \frac{1}{2}\underbrace{\rho(u_h)(z-\psi_h)}_{\textit{primal residual}} + R_h^{(3)}(e^u,e^q,e^z) \end{aligned}$$

for arbitrary $\varphi_h, \psi_h \in V_h$ *and* $\chi_h \in Q_h$. *The remainder* $R_h^{(3)}(e,e^q,e^z)$ *is cubic in* $e^u := u-u_h$, $e^q := q-q_h$, $e^z := z-z_h$.

Remarks.

- The derivation of the error representation does not require the uniqueness of solutions (important for application to eigenvalue problems). The a priori assumption $x_h \to x$ $(h \to 0)$ makes the result meaningful for cases with non-unique solutions.
- The evaluation of the residual terms requires reliable guesses for the unknown solution $\{u, q, z\}$. These may be obtained by local high-order interpolation (e.g., biquadratic interpolation on 2×2 cell patches of bilinear solutions).
- The cubic remainder term may be neglected.
- This error estimation only uses available information (no extra dual problem has to solve).
- Practical solution process by nesting outer Newton iteration with mesh adaptation (successive "model enrichment").

Problem: Method results in "nonadmissible" solution q_h^{opt}, u_h^{opt}.

Solution: Recover an admissible state $\tilde{u}_h^{opt}$ from q_h^{opt} by solving the state equation on a finer mesh $\mathbb{T}_{h'}$:

$$A(\tilde{u}_h^{opt}, q_h^{opt})(\psi_{h'}) = 0 \quad \forall \psi_{h'} \in V_{h'}$$

Proof of the error representation

Abstract formulation

$$x := \{u, q, z\} \in X = V \times Q \times V, \qquad L(x) := \mathcal{L}(u, q, z)$$

Variational equation: $\quad L'(x)(y) = 0 \quad \forall y \in X$

$$x_h := \{u_h, q_h, z_h\} \in X_h = V_h \times Q_h \times V_h$$

Galerkin approximation: $\quad L'(x_h)(y_h) = 0 \quad \forall\, y_h \in X_h$

$$L(x) - L(x_h) = ???$$

Fundamental theorem of calculus

$$\begin{aligned} L(x) - L(x_h) = & \int_0^1 L'(x_h + se)(e)\,\mathrm{ds} + \frac{1}{2} L'(x_h)(e) \\ & - \frac{1}{2} L'(x_h)(e) - \frac{1}{2} \underbrace{L'(x)(e)}_{=0} \end{aligned}$$

Galerkin orthogonality

$$L'(x_h)(e) = L'(x_h)(x - y_h) + \underbrace{L'(x_h)(y_h - x_h)}_{=0}, \quad y_h \in X_h$$

Error representation of trapezoidal rule

$$\int_0^1 f(s)\,\mathrm{ds} - \frac{1}{2}\{f(0) + f(1)\} + \frac{1}{2} \int_0^1 f''(s) s(s-1)\,\mathrm{ds}$$

Hence, with an arbitrary $y_h \in X_h$, there holds

$$L(x) - L(x_h) = \frac{1}{2} \underbrace{L'(x_h)(x - y_h)}_{\text{Residual}} + \frac{1}{2} \int_0^1 L'''(x_h + se)(e, e, e)\, s(s-1)\,\mathrm{ds}$$

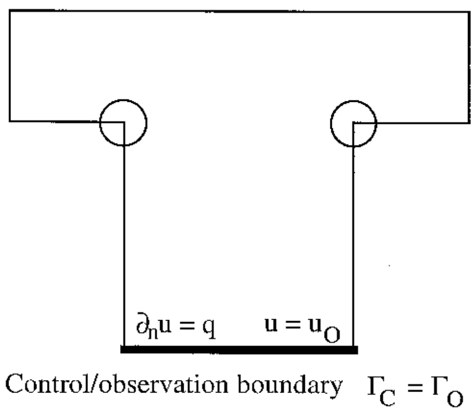
Application for $L(x) = L(u, p, z)$ yields the assertion.

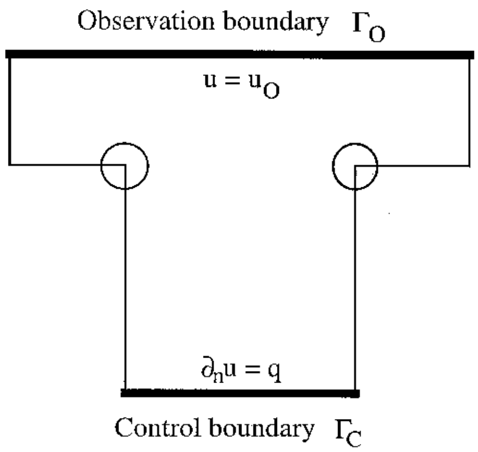
7.1 An example of boundary control (H. Kapp 1999)

$$-\Delta u + s(u) = f \quad \text{in } \Omega \subset \mathbb{R}^2, \quad s(u) = u^3 - u$$

$$\partial_n u = 0 \ \text{ on } \Gamma_N, \quad \partial_n u = q \ \text{ on } \Gamma_C \ \text{ (control)}$$

$$J(u,q) = \tfrac{1}{2}\|u - c_0\|_{\Gamma_O}^2 + \tfrac{\alpha}{2}\|q\|_{\Gamma_C}^2 \qquad (c_0 \equiv 1, \ \alpha = 1)$$





$$V = H^1(\Omega), \quad Q = L^2(\Gamma_C)$$

Necessary optimality conditions:

$$\begin{aligned} (u - c_0, \psi)_{\Gamma_O} + (\nabla \psi, \nabla z)_\Omega + (\psi, z)_\Omega &= 0 \quad \forall \psi \in V \\ \alpha(q,\chi)_{\Gamma_C} - (z,\chi)_{\Gamma_C} &= 0 \quad \forall \chi \in Q \\ (\nabla u, \nabla \varphi)_\Omega + (u,\varphi)_\Omega - (f,\varphi)_\Omega - (q,\varphi)_{\Gamma_C} &= 0 \quad \forall \varphi \in V \end{aligned}$$

Galerkin approximation in

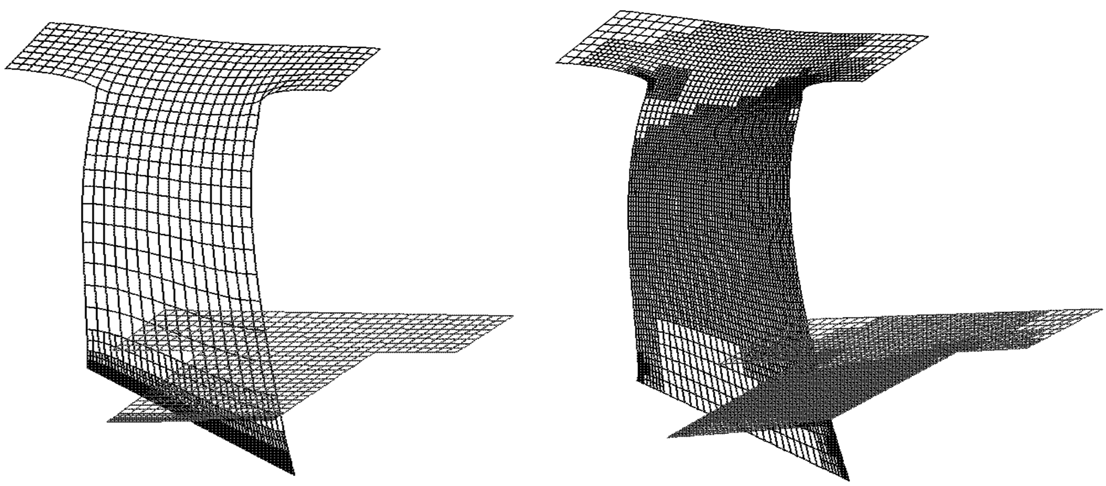
$$V_h = \text{"linear elements"}, \qquad Q_h := \text{"}\partial_n V_{h|\Gamma_C}\text{"}$$

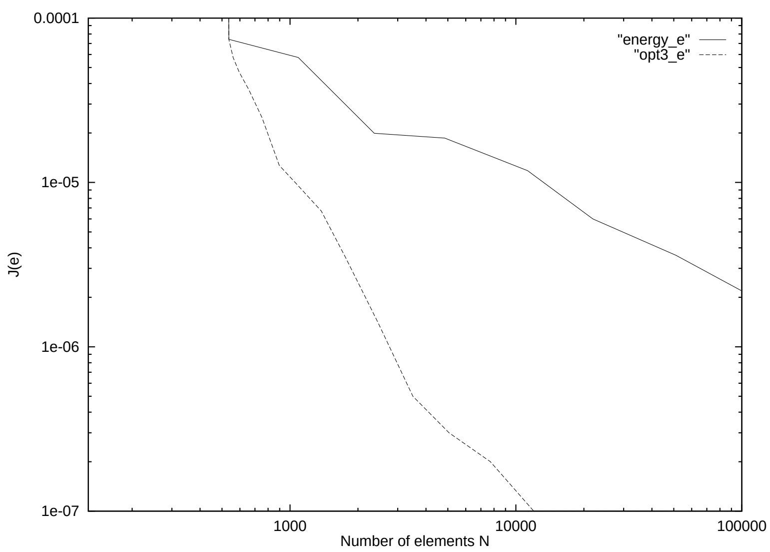
$$\begin{aligned} (u_h - c_0, \psi_h)_{\Gamma_O} + (\nabla \psi_h, \nabla z_h)_\Omega + (\psi_h, z_h)_\Omega &= 0 \quad \forall \psi_h \in V_h \\ \alpha(q_h,\chi_h)_{\Gamma_C} - (z_h,\chi_h)_{\Gamma_C} &= 0 \quad \forall \chi_h \in Q_h \\ (\nabla u_h, \nabla \varphi_h)_\Omega + (u_h,\varphi_h)_\Omega - (f,\varphi_h)_\Omega - (q_h,\varphi_h)_{\Gamma_C} &= 0 \quad \forall \varphi_h \in V_h \end{aligned}$$

Test case 1:

N	596	1616	5084	8648	15512
E_h	2.56e-04	2.38e-04	8.22e-05	4.21e-05	3.99e-05
I_{eff}	0.34	0.81	0.46	0.29	0.43

Efficiency of the weighted error estimator for Configuration 1



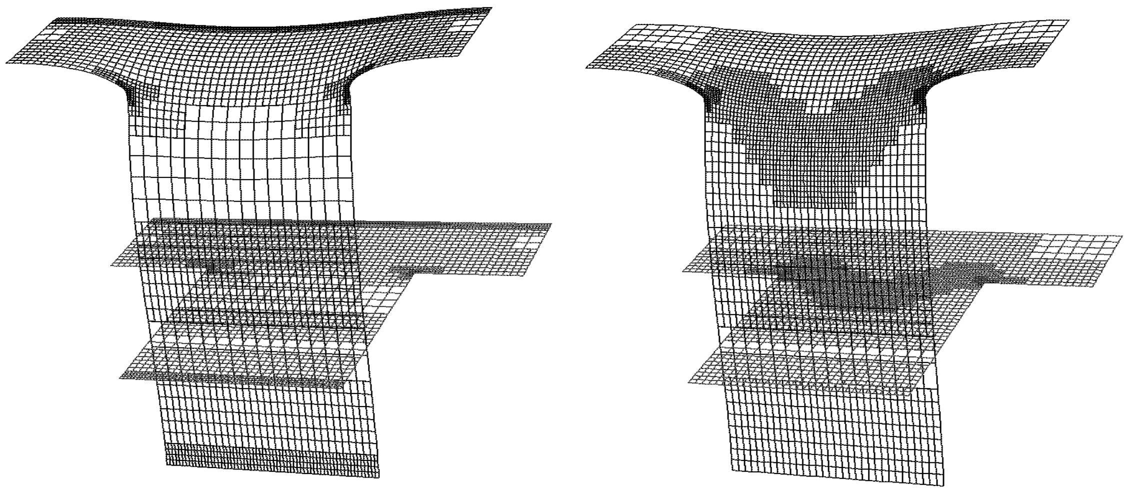


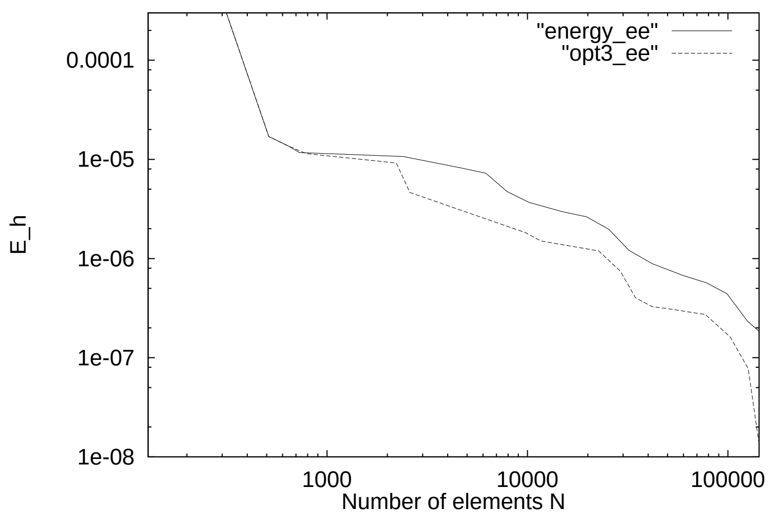
Efficiency of the meshes generated by the weighted error estimator (□) and the energy-error estimator (×) in log / log scale

Test case 2:

N	512	15368	27800	57632	197408
E_h	9.29e-05	8.14e-07	4.86e-07	2.31e-07	4.58e-08
I_{eff}	1.32	0.56	0.35	0.42	0.32

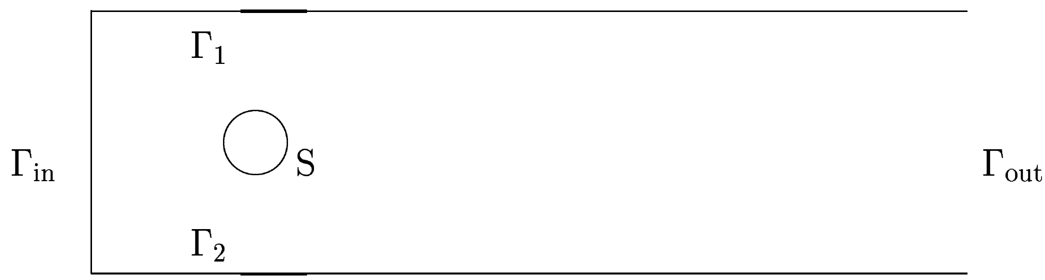
Efficiency of the weighted error estimator for Configuration 1





Efficiency of the meshes generated by the weighted error estimator (□) and the energy-error estimator (×) in log / log scale

7.2 Minimization of drag (R. Becker 1999)



$u \in u^{in} + V$ state variable
q 'boundary control' (piecewise constant at $\Gamma_Q := \Gamma_1 \cup \Gamma_2$)

$$J(u,q) := J_{drag} \;\to\; \min, \qquad \mathcal{A}(u) + \mathcal{B}q = 0$$

Variational formulation:

$$a_\delta(u)(\varphi) + b(q,\varphi) = 0 \qquad \forall \varphi \in V$$

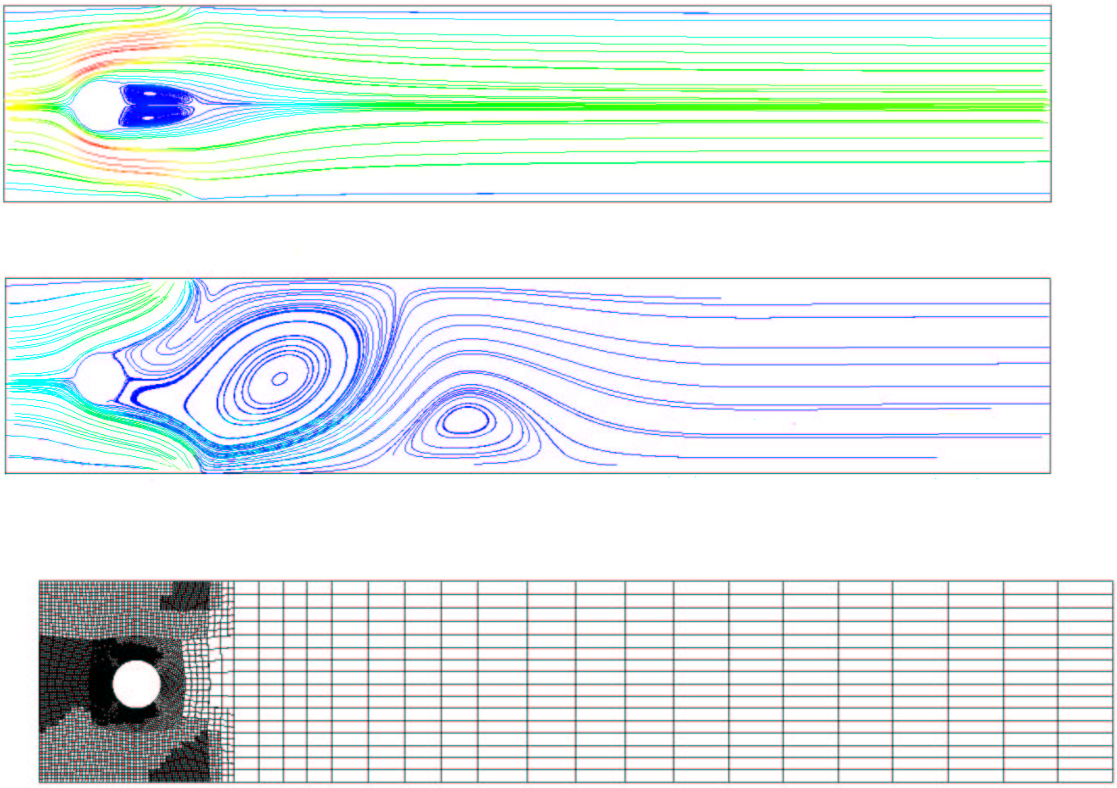
'control form' $b(q,\varphi) := -(q, n \cdot \varphi^v)_{\Gamma_Q}$

Computational example

Uniform refinement		Adaptive refinement	
N	J_{drag}	N	J_{drag}
10512	3.31321	1572	3.28625
41504	3.21096	4264	3.16723
164928	3.11800	11146	3.11972

Table. Uniform refinement versus adaptive refinement for Re $= 40$

Extension to nonstationary control (e.g., by rotation of the cylinder).



Velocity of the uncontrolled flow (top)
controlled flow (middle)
corresponding adapted mesh (bottom)

Problem: Stability of 'optimal' flow (computed by Newton method)?

7.3 Stability of flows (V. Heuveline 2001)

Stability of base solution $\hat{u} = \{\hat{v}, \hat{q}\}$

a) Linear stability theory ('spectral argument'):

Non-symmetric eigenvalue problem for $u := \{v, p\} \in V$ and $\lambda \in \mathbb{C}$:

$$\mathcal{A}'(\hat{u})u := -\nu\Delta v + \hat{v}\cdot\nabla v + v\cdot\nabla\hat{v} + \nabla p = \lambda v, \quad \nabla\cdot v = 0$$

$$\operatorname{Re}\lambda \ge 0 \quad \Rightarrow \quad \hat{u} \text{ stable (?)}$$

b) Nonlinear stability theory ('energy argument'):

Symmetric eigenvalue problem for $u := \{v, p\} \in V$ and $\lambda \in \mathbb{R}$:

$$-\nu\Delta v + \tfrac{1}{2}(\nabla\hat{v} + \nabla\hat{u}^T)v + \nabla p = \lambda\, v, \quad \nabla\cdot v = 0$$

$$\operatorname{Re}\lambda \ge 0 \quad \Rightarrow \quad \hat{u} \text{ stable (!)}$$

Variational formulation:

$$\begin{aligned} a'(\hat{u})(\psi,\varphi) &:= \nu(\nabla\psi^v, \nabla\varphi^v) + (\hat{v}\cdot\nabla\psi^v, \varphi^v) + (\psi^v\cdot\nabla\hat{v}, \varphi^v) \\ &\quad - (\psi^p, \nabla\cdot\varphi^v) + (\varphi^p, \nabla\cdot\psi^v), \\ m(\psi,\varphi) &:= (\psi^v, \varphi^v). \end{aligned}$$

Primal and dual eigenvalue problems: $u,\, u^* \in V$:

$$\begin{aligned} a'(\hat{u})(u,\varphi) &= \lambda\, m(u,\varphi) \quad \forall\varphi \in V \\ a'(\hat{u})(\varphi,u^*) &= \lambda\, m(\varphi,u^*) \quad \forall\varphi \in V \end{aligned}$$

Normalization: $m(u,u) = m(u,u^*) = 1\,.$

If $m(u,u^*) = 0$, the boundary value problem

$$a'(\hat{u})(\tilde{u},\varphi) - \lambda\, m(\tilde{u},\varphi) = m(u,\varphi) \quad \forall\varphi \in V$$

has a solution $\tilde{u} \in V$ ('generalized eigenfunction') $\Rightarrow$ $\operatorname{defect}(\lambda) > 0$

Discretization

Stabilized sesquilinear form

$$\tilde{a}'_\delta(\hat{v}_h)(u_h,\varphi_h) := a'(\hat{u})(u_h,\varphi) + (\mathcal{A}'(\hat{u})u - \lambda_h v, \mathcal{S}(\hat{u})\varphi)_\delta$$

Discrete primal and dual eigenvalue problems $u_h,\ u_h^* \in V_h,\ \lambda_h \in \mathbb{C}$:

$$\begin{aligned} \tilde{a}'_\delta(\hat{u}_h)(u_h,\varphi_h) &= \lambda_h\, m(u_h,\varphi_h) \quad \forall \varphi_h \in V_h \\ \tilde{a}'_\delta(\hat{u}_h)(\varphi_h,u_h^*) &= \lambda_h\, m(\varphi_h,u_h^*) \quad \forall \varphi_h \in V_h \end{aligned}$$

Stabilization $\quad m(u_h,u_h) = m(u_h,u_h^*) = 1$

Blow-up criterion:

$$m(u_h^*,u_h^*) \to \infty \quad (h \to 0) \quad \Rightarrow \quad \text{defect}(\lambda) > 0$$

A posteriori error estimation

Embedding into the general framework of variational equations:

$$\begin{aligned} &\mathcal{V} := V \times V \times \mathbb{C}, \quad \mathcal{V}_h := V_h \times V_h \times \mathbb{C} \\ &U := \{\hat{u}, u, \lambda\}, \quad U_h := \{\hat{u}_h, u_h, \lambda_h\}, \quad \Phi = \{\hat{\varphi}, \varphi, \mu\} \in \mathcal{V} \end{aligned}$$

Semi-linear form:

$$\begin{aligned} A(U)(\Phi) := \ & \underbrace{-a_\delta(\hat{u})(\hat{\varphi})}_{\text{base solution}} \\ & + \underbrace{\lambda\, m(u,\varphi) - \tilde{a}'_\delta(\hat{u})(u,\varphi)}_{\text{eigenvalue equation}} + \underbrace{\overline{\mu}\{m(u,u) - 1\}}_{\text{normalization}} \end{aligned}$$

Compact variational formulation:

$$\begin{aligned} A(U)(\Phi) &= 0 \quad \forall \Phi \in \mathcal{V} \\ A(U_h)(\Phi_h) &= 0 \quad \forall \Phi_h \in \mathcal{V}_h \end{aligned}$$

Error control functional:

$$J(\Phi) := \mu\, m(\varphi,\varphi) \quad \Rightarrow \quad J(U) = \lambda m(u,u) = \lambda.$$

Dual solutions $Z = \{\hat{z}, z, \pi\} \in \mathcal{V}$, $Z_h = \{\hat{z}_h, z_h, \pi_h\} \in \mathcal{V}_h$:

$$\begin{aligned} A'(U)(\Phi, Z) &= J'(U)(\Phi) \qquad \forall \Phi \in \mathcal{V} \\ A'(U_h)(\Phi_h, z_h) &= J'(U_h)(\Phi_h) \quad \forall \Phi_h \in \mathcal{V}_h \end{aligned}$$

Observation: $\quad \hat{z} = \hat{u}^*, \quad z = u^*, \quad \pi = \lambda$

$$a'(\hat{u})(\psi, \hat{u}^*) = \; - a''(\hat{u})(\psi, u, u^*) \quad \forall \psi \in V$$

Residuals:

$$\begin{aligned} \rho(\hat{u}_h)(\cdot) &:= -a_\delta(\hat{u}_h)(\cdot) \\ \rho^*(\hat{u}_h^*)(\cdot) &:= -a''_\delta(\hat{u})(\cdot, u_h, u_h^*) - \tilde{a}'_\delta(\hat{u}_h)(\cdot, \hat{u}_h^*) \end{aligned}$$

$$\begin{aligned} \rho(u_h, \lambda_h)(\cdot) &:= \lambda_h\, m(u_h, \cdot) - \tilde{a}'_\delta(\hat{u}_h)(u_h, \cdot) \\ \rho^*(u_h^*, \lambda_h)(\cdot) &:= \lambda_h\, m(\cdot, u_h^*) - \tilde{a}'_\delta(\hat{u}_h)(\cdot, u_h^*) \end{aligned}$$

Proposition. *We have the error representation*

$$\lambda - \lambda_h = \tfrac{1}{2}\, \underbrace{\rho(\hat{u}_h)(\hat{u}^* - \hat{\psi}_h) + \tfrac{1}{2}\rho^*(\hat{u}_h^*)(\hat{u} - \hat{\varphi}_h)}_{\textit{base solution residuals}} \Big\}$$

$$+ \; \tfrac{1}{2}\, \underbrace{\rho(u_h, \lambda_h)(u^* - \psi_h) + \tfrac{1}{2}\rho^*(u_h^*, \lambda_h)(u - \varphi_h)}_{\textit{eigenvalue residuals}} + R_h$$

for arbitrary $\hat{\psi}_h$, ψ_h, $\hat{\varphi}_h$, $\varphi_h \in V_h$. *The remainder* R_h *is cubic in the errors* $\hat{e}^v := \hat{v} - \hat{v}_h$, $\hat{e}^{v*} := \hat{v}^* - \hat{v}_h^*$ *and* $e^\lambda := \lambda - \lambda_h$, $e^v := v - v_h$, $e^{v*} := v^* - v_h^*$.

Computational example

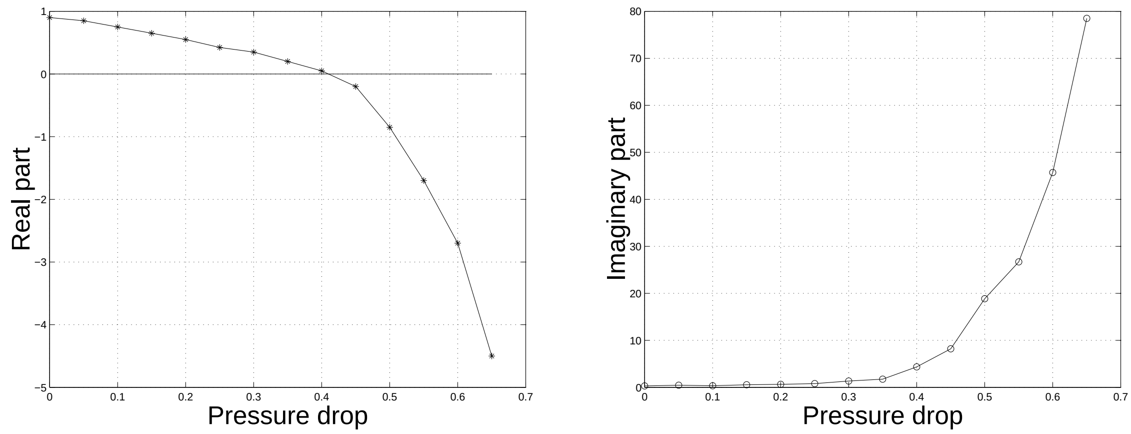


Figure. Real part and imaginary part of critical eigenvalue as function of imposed pressure

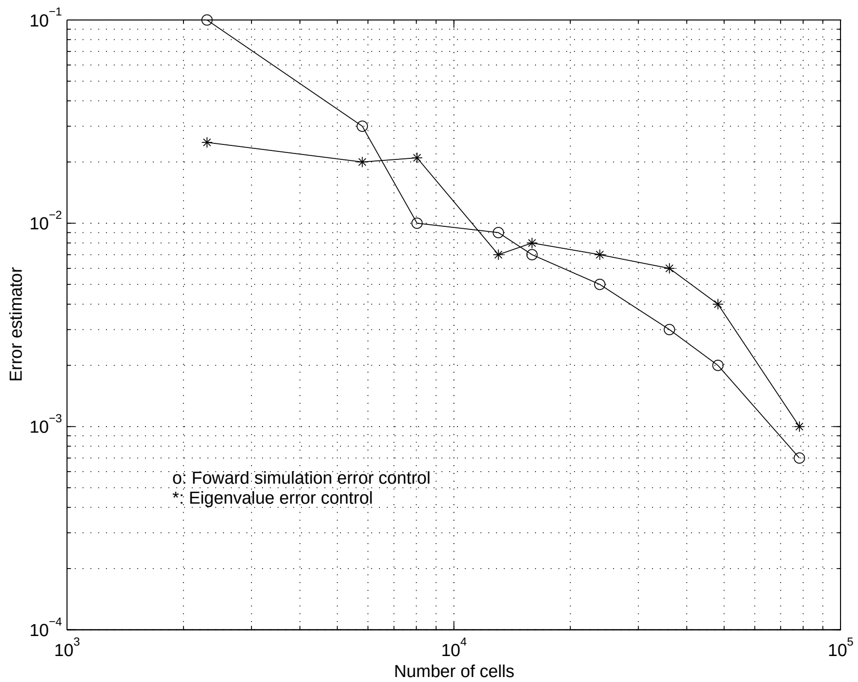


Figure. Dominance of error indicator contributions, 'base solution-part' and 'eigenvalue part'

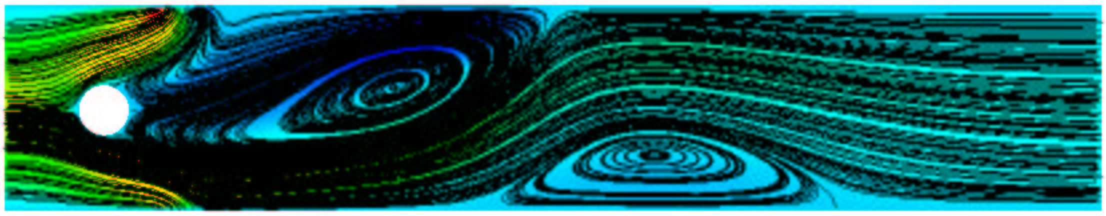


Figure. The base velocity field of the 'minimal-drag' solution (computed on a globally refined mesh

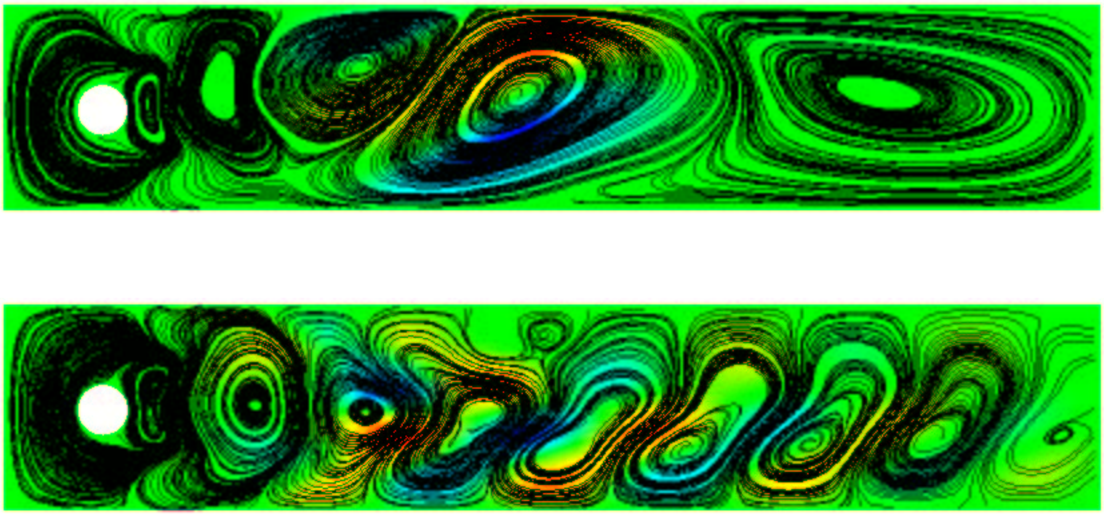


Figure. Streamline plots of real part of 'critical' eigenfunction shortly before and shortly after Hopf bifurcation