Optimal control and Maximum principle

Daniel Wachsmuth, RICAM Linz

Content

1. **Model problem**
2. Adjoint equation
3. Maximum principle
4. Numerical algorithms
5. Flow control

Objective: Reach target as fast as possible.

Control: Acceleration.

Constraints: Stop at target, control constraints.

THE model

Quantities: t .. time, $x(t) \in \mathbb{R}$ position, $u(t) \in [-1, +1]$ control, $m > 0$ mass, $x_0 \neq 0$ initial position

Equation of motion:

$$m\, x''(t) = u(t)$$

$$x(0) = x_0$$

Quantities: t .. time, $x(t) \in \mathbb{R}$ position, $u(t) \in [-1, +1]$ control, $m > 0$ mass, $x_0 \neq 0$ initial position

Equation of motion:

$$m\, x''(t) = u(t)$$

$$x(0) = x_0$$

Optimal control problem:

Minimize T subject to

- Equation of motion
- $x(T) = 0$, $x'(T) = 0$
- $|u(t)| \leq 1$

THE model

Quantities: t .. time, $x(t) \in \mathbb{R}$ position, $u(t) \in [-1, +1]$ control, $m > 0$ mass, $x_0 \neq 0$ initial position

Equation of motion:

$$m\,x''(t) = u(t)$$

$$x(0) = x_0$$

Optimal control problem:

Minimize T subject to

- Equation of motion
- $x(T) = 0$, $x'(T) = 0$
- $|u(t)| \leq 1$

Problem: Find a solution!

Features:

- optimization problem
- some optimization variable are functions $\rightarrow$ infinite-dimensional optimization
- differential equations (ode / pde)
- inequalities

Minimize functional J given by

$$J(x,u,T) := \int_0^T f_0(x(t),u(t),t)\ dt$$

subject to the ODE

$$x'(t) = f(x(t),u(t)) \text{ a.e. on } (0,T),$$

initial and terminal conditions

$$x(0) = x_0, \quad x(T) = z$$

control constraints

$$u(t) \subset U.$$

Setting $f_0 = 1$ the case of the time-optimal problem is contained as special case.

Unknowns: measurable functions u, x with $u(t) \in \mathbb{R}^m$, $x(t) \in \mathbb{R}^n$.

Problem: Find a solution!

Consider the simpler problem of minimizing $f : \mathbb{R} \to \mathbb{R}$,

$$\min f(x), \quad x \in \mathbb{R}.$$

If f is differentiable, then every solution $\bar{x}$ satisfies

$$f'(x) = 0.$$

Solve this equation to find candidates for solutions!

Consider the simpler problem of minimizing $f : \mathbb{R} \to \mathbb{R}$,

$$\min f(x), \quad x \in \mathbb{R}.$$

If f is differentiable, then every solution $\bar{x}$ satisfies

$$f'(x) = 0.$$

Solve this equation to find candidates for solutions!

What is f' in the optimal control problem??

Content

1. Model problem
2. **Adjoint equation**
3. Maximum principle
4. Numerical algorithms
5. Flow control

Assumption: For every control function u there exists a uniquely determined solution of the ODE $x = x(u)$.

Simplification: No terminal constraint.

Task: Compute directional derivative Given control u^0, x^0, direction h

$$J'(x^0, u^0)h \approx \frac{1}{\epsilon}(J(x(u^0 + \epsilon h), u^0 + \epsilon h) - J'(x^0, u^0))$$

Disadvantages:

- Choice of ϵ
- Each evaluation of one difference quotient requires one (nonlinear) ODE solve

Chain rule:

$$\frac{d}{du}J(x^0,u^0) = J_x(x^0,u^0)\frac{d}{du}x(u^0) + J_u(x^0,u^0)$$

(Total) Directional derivative

$$\frac{d}{du}J(x^0,u^0)h = J_x(x^0,u^0)\frac{d}{du}x(u^0)h + J_u(x^0,u^0)h$$

The quantity $z := \frac{d}{du}x(u^0)h$ is the solution of the linearized ode

$$z' = f_x(x^0,u^0)z + f_u(x^0,u^0)h, \quad z(0) = 0$$

The quantity $J_x(x^0,u^0)\frac{d}{du}x(u^0)h$ is a dual product:

$$J_x(x^0,u^0)\frac{d}{du}x(u^0)h = \langle J_x(x^0,u^0),\ \frac{d}{du}x(u^0)h\rangle$$

Chain rule:

$$\frac{d}{du}J(x^0,u^0) = J_x(x^0,u^0)\frac{d}{du}x(u^0) + J_u(x^0,u^0)$$

(Total) Directional derivative

$$\frac{d}{du}J(x^0,u^0)h = J_x(x^0,u^0)\frac{d}{du}x(u^0)h + J_u(x^0,u^0)h$$

The quantity $z := \frac{d}{du}x(u^0)h$ is the solution of the linearized ode

$$z' = f_x(x^0,u^0)z + f_u(x^0,u^0)h, \quad z(0) = 0$$

The quantity $J_x(x^0,u^0)\frac{d}{du}x(u^0)h$ is a dual product:

$$J_x(x^0,u^0)\frac{d}{du}x(u^0)h = \langle J_x(x^0,u^0),\ \frac{d}{du}x(u^0)h\rangle = \left\langle \left(\frac{d}{du}x(u^0)\right)^* J_x(x^0,u^0),\ h\right.$$

Adjoint equation

How can we characterize

$$q := \left(\frac{d}{du}x(u^0)\right)^* J_x(x^0, u^0)?$$

It turns out that

$$q = f_u(x^0, u^0)^T p,$$

where p solves the linear ODE [Recall $J = \int f_0(x, u)$]

$$-p'(t) = f_x(x^0, u^0)^T p + f_{0,x}(x^0, u^0)^T, \quad p(T) = 0.$$

Conclusion: Given (x^0, u^0) and p^0. Then

$$\frac{d}{du}J(x^0, u^0)h = \int_0^T p^T f_u(x^0, u^0)h + f_{0,u}(x^0, u^0)h$$

Advantage: One linear ODE needed to evaluate many directional derivatives.

Universal applicability

Adjoint equation:

$$-p'(t) = f_x(x^0, u^0)^T p + f_{0,x}(x^0, u^0)^T, \quad p(T) = 0.$$

Properties:

- Linear in p
- Operator in the equation is the linearized and transposed (adjoint) operator of the state equation
- Inhomogeneities originate from objective functional
- Nonzero data only where observation takes place

Advantage:

- One linear ODE solve to obtain all derivative information
- Works for many problems: PDE, shape optimization, etc

We only wanted to evaluate f' ...

Content

1. Model problem
2. Adjoint equation
3. **Maximum principle**
4. Numerical algorithms
5. Flow control

Recall optimal control problem

Minimize functional J given by

$$J(x,u,T) := \int_0^T f_0(x(t),u(t),t)\;dt$$

subject to the ODE

$$x'(t) = f(x(t),u(t)) \text{ a.e. on } (0,T),$$

initial and terminal conditions

$$x(0) = x_0, \quad x(T) = z$$

control constraints

$$u(t) \subset U.$$

Optimal control: (x^*,u^*,T^*) is optimal if

$$J(x^*,u^*,T^*) \le J(x,u,T)$$

for all admissible (x,u,T).

Define Hamilton-Function:

$$H(t, x, u, p, \lambda_0) = p^T f(x, u) - \lambda_0 f_0(x, u, t)$$

Theorem: Let the functions f, f_0 be continuous wrt (x, u, t) and continuously differentiable wrt (t, u). Let $U \subset \mathbb{R}^m$ be given.

Let (x^*, u^*, T^*) be optimal.

Then there exists $p_{T^*} \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}$ with $(\lambda_0, p_{T^*}) \neq 0$, $\lambda_0 \geq 0$ such that the following conditions are satisfied

- Adjoint equation

$$-p'(t) = f_x(x^*(t), u^*(t), t)^T p(t) - \lambda_0 f_{0,x}(x^*(t), u^*(t), t)^T$$
$$p(T^*) = p_{T^*}$$

- Maximum condition

$$H(t, x^*(t), u^*(t), p(t), \lambda_0) = \max_{v \in U} H(t, x^*(t), v, p(t), \lambda_0) \qquad \text{a.e. on } (0, T^*$$

- Maximum function

$$\max_{v\in U} H(t, x^*(t), v, p(t), \lambda_0)$$

is continuous on $[0, T^*]$ and satisfies at T^*

$$\max_{v\in U} H(T^*, x^*(T^*), v, p(T^*), \lambda_0) = 0.$$

In particular, the maximum condition is satisfied in all points of left/right-continuity of u^*.

Message: The maximum principle generalizes the equation $f'(x) = 0$. Solve the system given by PMP to obtain solution candidates.

- PMP is a necessary optimality condition: sometimes sufficient (convex problems)
- Comparison to Kuhn-Tucker-type optimality conditions: Here no derivatives wrt u needed!
- Role of λ_0: Indicates (non-)degeneracy of constraints.

  If one knows $\lambda_0 > 0$ a-priori, the PMP-system can be scaled such that $\lambda_0 = 1$.

  The point $(\lambda_0, p_{T^*}) = 0$ is a solution of the PMP-system.

  There is no equation to determine λ_0 -in computations set $\lambda_0 = 1$.

Rocket car

Here: non-degenerate case $\lambda_0 > 0$ if $x_0 \neq 0$.

Control u^* is bang-bang

$$u^*(t) = \begin{cases} -\operatorname{sign}(x_0) & \text{if } t \in (0, T^*/2) \\ \operatorname{sign}(x_0) & \text{if } t \in (T^*/2, T^*) \end{cases}$$

Content

1. Model problem
2. Adjoint equation
3. Maximum principle
4. **Numerical algorithms**
5. Flow control

Suppose that f is continuously differentiable wrt u, and $U = \mathbb{R}^m$ (no control constraints). Then the maximum condition implies

$$H_u(t, x^*(t), u^*(t), p(t), \lambda_0) = 0,$$

which gives

$$-p(t)^T f_u(x^*(t), u^*(t), t) + f_{0,u}(x^*(t), u^*(t), t) = 0$$

If we can solve this equation for $u^* = u^*(p)$, then we can replace the control u by the function $u(p)$ in the state equation, and obtain a boundary value problem for (x^*, p).

Solve with ODE-integration methods.

Discretize-then-optimize

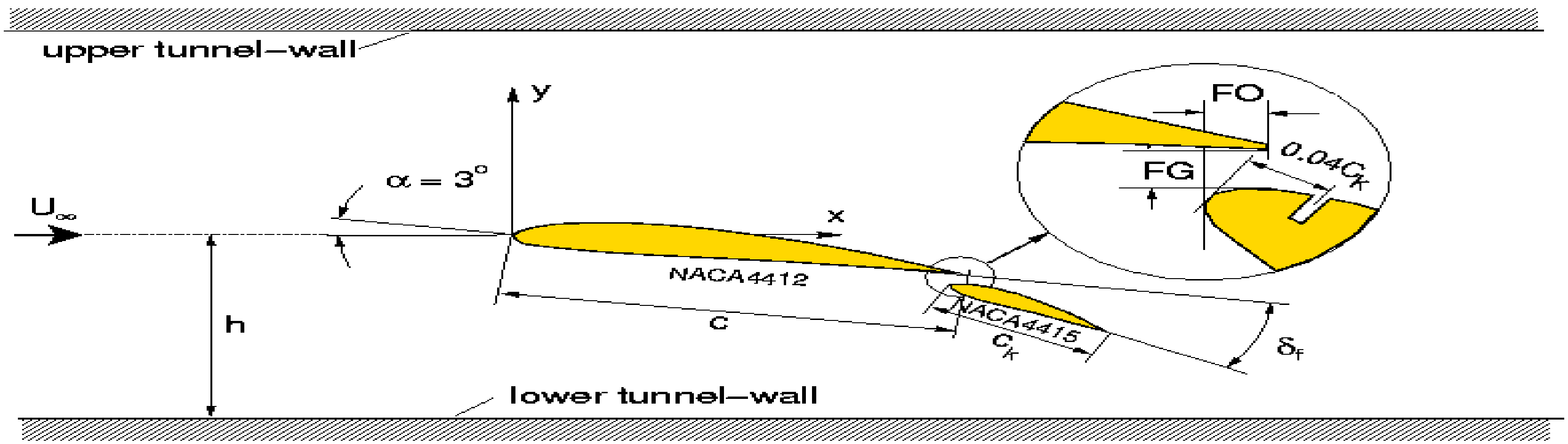
1. Discretize ODE by some discretization method (e.g. finite differences)
2. Obtain finite-dimensional optimization problem
3. Use optimization software

Optimize-then-discretize

1. Derive formulas for derivatives (gradient, Hessian) of optimal control problem
2. Use (infinite-dimensional) optimization algorithm
3. Discretize and run the algorithm in finite-dimensional space

Content

1. Model problem
2. Adjoint equation
3. Maximum principle
4. Numerical algorithms
5. **Flow control**



Maximize Lift

under the constraints

$$\begin{cases} \text{Navier-Stokes equations} \\ \text{Maximal drag} \\ \text{Control constraints} \end{cases}$$

Navier-Stokes equations:

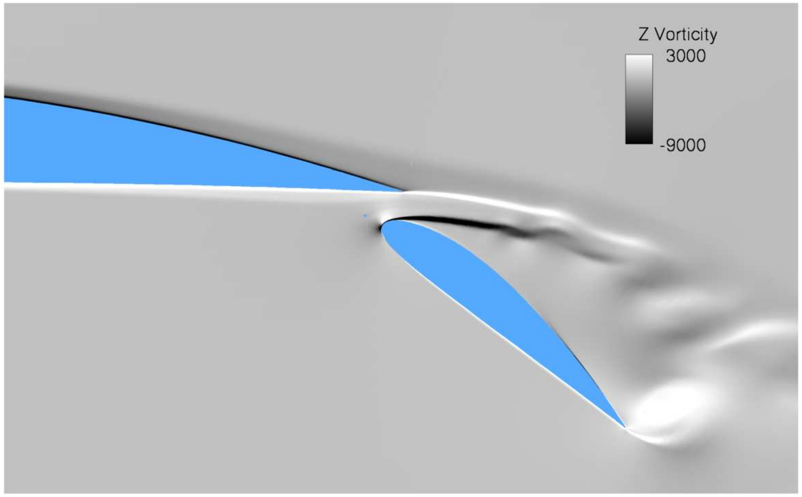
$$
\begin{aligned}
y_t - \nu \Delta y + (y \cdot \nabla) y + \nabla p &= 0 \\
\operatorname{div} y &= 0 \\
y|_\Gamma &= u \\
y(0) &= y_0 .
\end{aligned}
$$

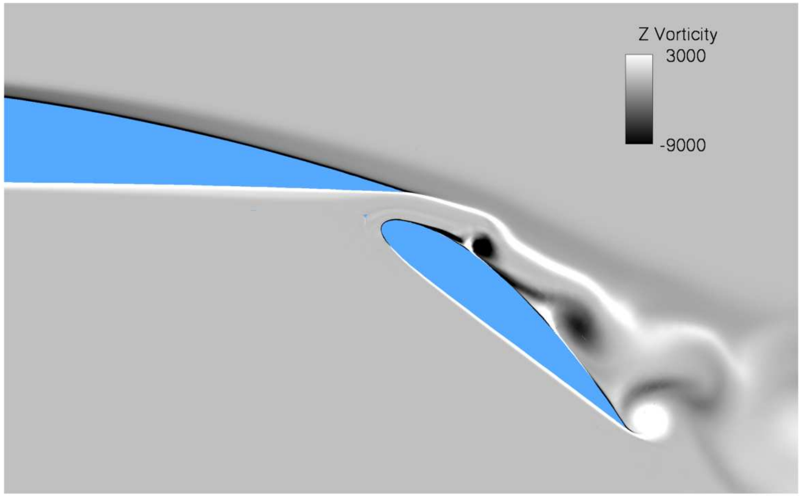
Adjoint equations:

$$
\begin{aligned}
-\lambda_t - \nu \Delta \lambda + (\nabla y)^T \lambda - (y \cdot \nabla)\lambda + \nabla \pi &= 0 \\
\operatorname{div} \lambda &= 0 \\
\lambda|_\Gamma &= \vec{e} \\
\lambda(T) &= 0.
\end{aligned}
$$

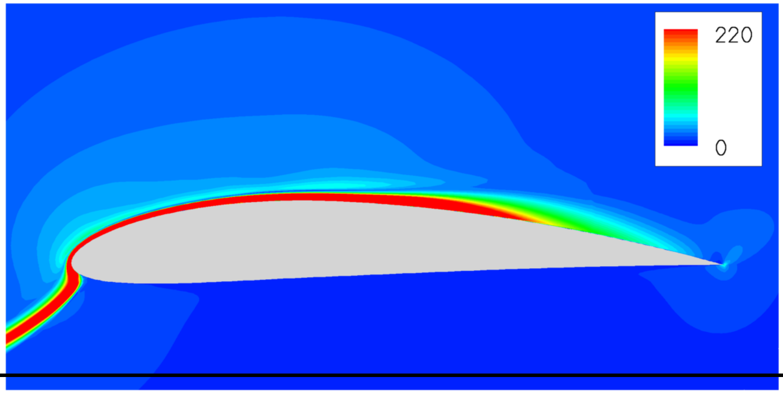
$$
\frac{d}{du} J(y^0, p^0, u^0) h = \int_{(0,T)\times\Gamma} -(\nu \frac{\partial \lambda}{\partial n} - \pi n) h
$$

Snapshots of vorticity: uncontrolled / controlled





Adjoint velocity field: Large near wing and near stagnation point streamline



References

Optimal control of ODEs

- L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. The Macmillan Co., New York, 1964.
- A. D. Ioffe and V. M. Tihomirov. *Theory of extremal problems*. North-Holland Publishing Co., Amsterdam, 1979.
- Jack W. Macki and Aaron Strauss. *Introduction to optimal control theory*. Springer-Verlag, New York, 1982.

Numerical methods

- J. Stoer and R. Bulirsch. *Introduction to numerical analysis*. Springer-Verlag, New York, third edition, 2002.

Optimal control of PDEs

- M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*. Springer, New York, 2009.
- F. Tröltzsch. *Optimal control of partial differential equations*. AMS, Providence, RI, 2010.