Constraint Satisfaction Using High-Order Taylor Models and Applications

Johannes Grote, Martin Berz and Kyoko Makino

Dept. of Physics and Astronomy, Michigan State University,
East Lansing, MI 48824, USA

Third International Workshop on Taylor Methods
Miami Beach, December 18, 2004

Assume you look for solutions of a given problem in a set D. The solutions often need to satisfy additional conditions.

$$\begin{aligned} g_j(x) &\leq 0 \text{ for } j = 1, ..., c_{\leq}, \; x \in D \\ h_j(x) &= 0 \text{ for } j = 1, ..., c_{=}, \; x \in D \end{aligned} \tag{1}$$

Typical examples are

- in optimization: find the global minimum of a function in a domain box D under aditional constraints (1)
- in Hamiltonian dynamics: solutions to Hamiltonian systems need to conserve certain invariants, e.g. energy

In particular it is hard to find *validated* solutions solving the problem *and* satisfying the constraint conditions.

A traditional approach is to subdivide D, throw away boxes which can be shown to violate the constraints (1) and attempt the minimization of f only on the remaining boxes. The problems are:

- subdivision is always expensive
- box rejection might be extra-slow because of overestimation in interval arithmetic

The general idea using Taylor Model methods:

- first represent the set F of feasible points in D which satisfy the conditions (1) by a high-order Taylor Model
$\Longrightarrow$ typically F is very small, usually a low-dimensional submanifolds
- then look for solutions only on these feasible boxes
$\Longrightarrow$ a much faster enclosure of the solutions is expected

Generally we formulate the problem as follows: Let

$$f_i(x_1, ..., x_n) = 0 \text{ for } i \in \{1, ..., m\} \tag{2}$$

be m constraint conditions, where the $f_i : \mathbb{R}^n \supset D \longrightarrow \mathbb{R}^n$ are sufficiently smooth. Assume nom that there is a point $\mathbf{x}^0$ in D satisfying the constraints (2),

$$f_i(x_1^0, ..., x_n^0) = 0 \; \forall \; i \in \{1, ..., m\}$$

We want to represent the feasible points in D by a Taylor Model expanded around $\mathbf{x}^0$. In the following we assome w.l.o.g. that $\mathbf{x}^0 = 0$.

m constraints $\Longrightarrow$ the feasible set can be parameterized by $n-m$ variables.
Assume the first $n-m$ variables can be used to do this.
Then we construct the function $\psi : D \longrightarrow \mathbb{R}^n$ by

$$\psi(x_1, \ldots, x_n) := \begin{pmatrix} x_1 \\ \vdots \\ x_{n-m} \\ f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix}$$

Note that for a feasible point $\mathbf{x}^f \in F$ we have

$$\psi(\mathbf{x}^f) = (x_1^f, \ldots, x_{n-m}^f, 0, \ldots, 0)^T$$

If ψ is invertible on D, then conversely we have

$$\psi^{-1}(x_1, \ldots, x_{n-m}, 0, \ldots, 0) = \psi^{-1}(\psi(x_1, \ldots, x_n)) = \mathcal{I} \qquad (3)$$

It is natural that the identities in the first $n - m$ components of ψ are also preserved by ψ^{-1}.

BUT the last m components in eq. (3) contain the dependency of the constrained variables on the unconstrained variables $x_k, k \in \{1, \ldots, n - m\}$.

Then one can replace x_k by $\psi_K^{-1}(x_1, \ldots, x_{n-m}, 0, \ldots, 0)$ for $k \in \{1, \ldots, n - m\}$ and search for solutions of the original problem in terms of only the unconstrained variables.

Taylor inversion of ψ

We obtain the Taylor polynomial of ψ^{-1} from the Taylor polynomial of ψ:

- write $\psi(\mathbf{x}) = L\,\mathbf{x} + N(\mathbf{x})$, where L is linear and N is purely nonlinear
- note that $\psi(\mathbf{0}) = \mathbf{0}$
- $$\begin{aligned} \psi \circ \psi^{-1} &= \mathcal{I} \\ \Longrightarrow \quad L \circ \psi^{-1} &= \mathcal{I} - N \circ \psi^{-1} \\ \Longrightarrow \quad \psi^{-1} &= L^{-1} \circ (\mathcal{I} - N \circ \psi^{-1}) \end{aligned}$$
- define the contracting operator $\mathcal{O}(.) := L^{-1} \circ (\mathcal{I} - N \circ .)$ and apply a fixed point theorem: repeated iteration (beginning with the identity $\mathcal{I}$) yields ψ^{-1} in finitely many steps

After the polynomial part of ψ^{-1} has been obtained, we can immediately extract the polynomial parts of the parametrization of the constrained variables $x_{n-m+1}, \ldots, x_n$ in terms of the unconstrained variables $x_1, \ldots, x_{n-m}$.

Let P_i be the polynomial part of the constrained x_i, then outfit P_i heuristically with a remainder bound

$$I_i := [-\epsilon_l^i, \epsilon_u^i] \ \forall \ i \in \{1, \ldots, m\}$$

An educated guess for the magnitude of the $\epsilon_{l,u}^i$ would be to evaluate $f_i(x_1 + [0,0], \ldots, x_{n-m} + [0,0], P_1 + [0,0], \ldots, P_m + [0,0])$.

The validity of the attempted remainder bound I_i can be checked by scanning the complement of $P_i(D) + I_i$ in D and for a violation of the constraint f_i.

We can inplement this scanning procedure by expressing the $(n - m) + i$-th variable $x_{(n-m)+i}$ by the auxiliary variables $y_i \in [-1, 1]$

$$\begin{aligned} x^+_{n-m+i} &= (P_i + \epsilon^i_u)(1 + y_i)/2 + x^{max}_{(n-m)+i}(1 - y_i)/2 \\ x^-_{n-m+i} &= (P_i - \epsilon^i_l)(1 + y_i)/2 + x^{min}_{(n-m)+1}(1 - y_i)/2 \end{aligned}$$

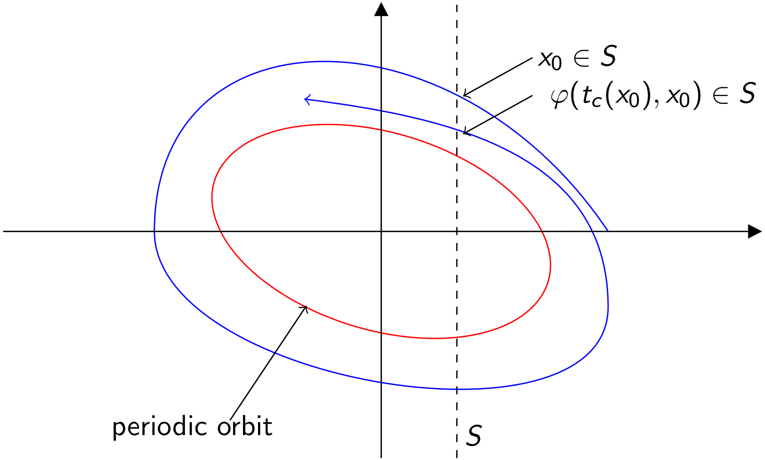
If it can be ascertained that

$$\begin{aligned} 0 &\notin f_i(x_1, ..., x^+_{n-m+i}(y_i), ...x_n) \text{ for } x_j, y_i \in [-1, 1] \\ 0 &\notin f_i(x_1, ..., x^-_{n-m+i}(y_i), ...x_n) \text{ for } x_j, y_i \in [-1, 1] \end{aligned}$$

using a rigorous range bounding scheme, then it is shown that the i-th constraint is violated everywhere outside $P_i + I_i$

$\Longrightarrow P_i + I_i$ forms a rigorous enclosure of the feasible set satisfying f_i.

Review: Definition of a Poincaré map



The construction of the Poincaré map $\mathcal{P}(x_0)$ can thus be reduced the the construction of the crossing time t_c.

- the crossing time for trajectories starting at $x_0 \in D$ exists and is uniquely determined
- it is contained implicitly in the vectorfield of the RHS of the ODE and the geometry of the surface S
- a polynomial approximation of $t_c(x_0)$ yields a polynomial approximation of $\mathcal{P}(x_0)$ by insertion into the flow $\varphi(t, x_0)$

$$\mathcal{P}(x_0) := \varphi(t_c(x_0), x_0)$$

- a Taylor model of $t_c(x_0) + I_{t_c}$ yields a Taylor model of $\mathcal{P}(x_0) + I_{\mathcal{P}}$ by insertion into the flow $\varphi(t, x_0) + I_\varphi$

$$\mathcal{P}(x_0) + I_{\mathcal{P}} := \varphi(t_c(x_0) + I_{t_c}, x_0 + [0, 0]) + I_\varphi$$

Treatable types of Poincaré sections

The method is able to model Poincaré sections S of the form $S := \{x \in \mathbb{R}^n : \sigma(x) = 0\}$ for some C^k-function $\sigma : \mathbb{R}^n \longrightarrow \mathbb{R}$.

- typically σ is $C^\infty \Longrightarrow S$ is smooth
- the case where S is a plane is the most common in applications: $S := \{x \in \mathbb{R}^n : x_1 = c\}$ for some $c \in \mathbb{R}$. Here $\sigma(x) := x_1 - c$

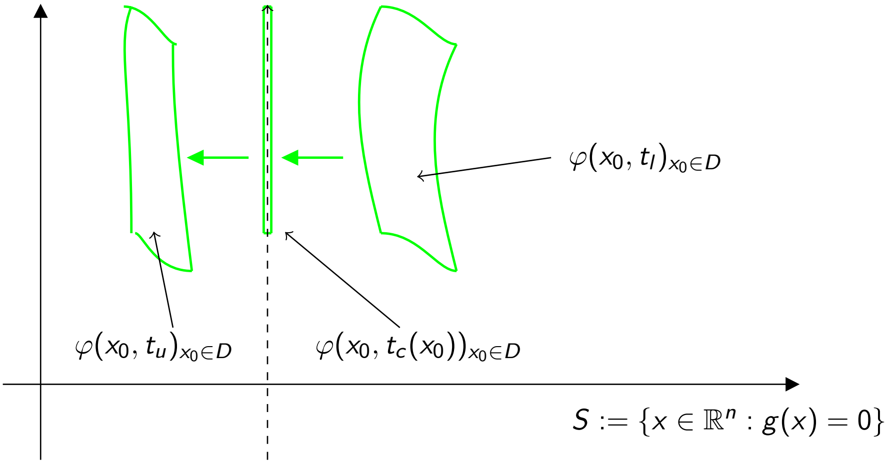
An additional condition is that the flow is 'sufficiently' transverse to the section.

Assume that we have performed a validated Taylor model integration from the initial condition $x(t=0) = X_0 + D$ until the last timestep before we cross the Poincaré section, at time t_l.

The interesting part happens in the final timestep, i.e. there is a time interval $I_t = [0, t_u]$ sucht that for the reference orbit:

$$\exists \tau \in I_t : \varphi(x_0, \tau) \in S$$

This allows us to use the machinery from the last section and formulate the finding of the Poincaré map as a constraint satisfaction problem.


$\varphi(x_0, t_l)_{x_0 \in D}$
$\varphi(x_0, t_u)_{x_0 \in D}$
$\varphi(x_0, t_c(x_0))_{x_0 \in D}$
$S := \{x \in \mathbb{R}^n : g(x) = 0\}$

Consider the domain box $D \times I_t$, where $I_t = [t_l, t_u]$. For the crossing time $t_c(x_0)$ we have to satisfy

$$\sigma(\varphi(x_0, t_c(x_0))) = 0 \,\forall\, x_0 \in D$$

whence we construct

$$\psi(x_0, t) = \begin{pmatrix} x_{0,1} \\ \vdots \\ x_{0,n} \\ \sigma(\varphi(x_0, t)) \end{pmatrix}, \; (x_0, t) \in D \times I_t$$

and after that we perform the inversion using DA-arithmetic and obtain ψ^{-1}.
Then we can evaluate $\psi^{-1}(x_{0,1}, ..., x_{0,n}, 0)$ and extract

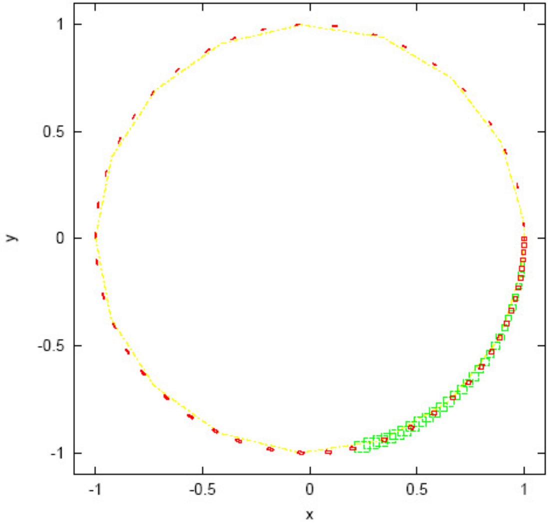
$$t_c(x_0) := \psi^{-1}_{n+1}(x_{0,1}, ..., x_{0,n}, 0)$$

An example: a muon cooling ring

$$\begin{aligned}
\dot{x_1} &= x_3 \\
\dot{x_2} &= x_4 - \frac{\alpha}{\sqrt{x_3^2 + x_4^2}} \cdot x_3 + \frac{\alpha}{\sqrt{x_1^2 + x_2^2}} \cdot x_2 \\
\dot{x_3} &= x_4 \\
\dot{x_4} &= -x_3 - \frac{\alpha}{\sqrt{x_3^2 + x_4^2}} \cdot x_4 - \frac{\alpha}{\sqrt{x_1^2 + x_2^2}} \cdot x_1
\end{aligned}$$

- $\alpha \in [0, 1]$ is the damping parameter, $\alpha = 1$ being the strongest damping

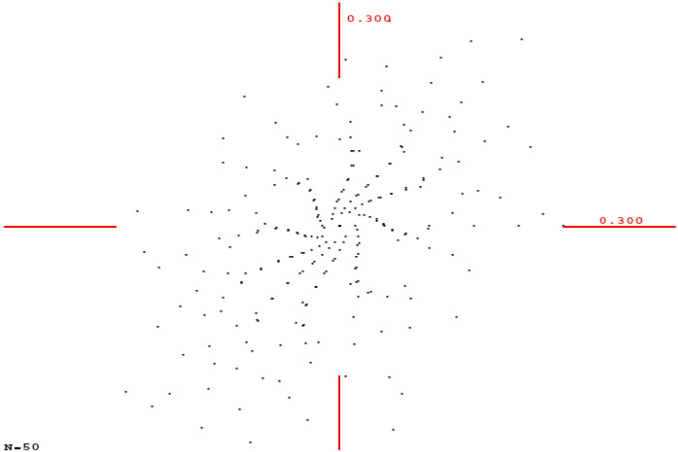
Muon cooler: action in x_1-x_2-plane



Muon cooler: example settings

- we consider the initial conditions $x(0) = X_0 + x_0$ where $X_0 = (0, 1, 1, 0)^T$ and $D = [-10^{-4}, 10^{-4}]$
- trajectories close to $\varphi(0, t)$ are 'pulled' towards the reference orbit $\Longrightarrow$ 'cooling' in transverse coordinates
- we consider the surface where $x_1 = 0$
- an 16th order computation was performed using COSY Infinity [3]

Muon cooler:Tracking picture in x_2-x_4- coordinates



References

1. M. Berz, *Modern Map Methods in Particle Beam Physics*, Academic Press, San Diego, 1999.
2. M. Berz and K. Makino, Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models, *Reliable Computing*, 4(4): 361-369, 1998.
3. M. Berz, J. Hoefkens and K. Makino, COSY Infinity Version 8.1 - programming manual. Technical Report MSUHEP-20703, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2001.