1 Optimization Problems and Techniques

Abstraction is selective ignorance.
— Andrew Koenig —

Due to the importance of optimization tasks, numerical optimization has a long tradition and many algorithms have been proposed. In this chapter we formulate the optimal control problem considered throughout this thesis and give a necessarily brief and incomplete overview of the methods used to solve this type of optimal control problems. Additionally, we will give a historical overview of the development of interior point methods.

1.1 Optimal Control Problems

The type of optimal control problem considered in this thesis is

$$\int_0^1 \mathfrak{J}(y(t), u(t))\, dt \to \min \tag{1.1}$$

subject to the equality constraints

$$\dot{y}(t) = f(y(t), u(t), t) \tag{1.2}$$
$$r_0(y(0)) = 0 \tag{1.3}$$
$$r_1(y(1)) = 0 \tag{1.4}$$

and the control inequality constraints

$$\mathfrak{g}^u(u(t)) \geq 0 \tag{1.5}$$

and the state (or path) inequality constraints

$$\mathfrak{g}^y(y(t)) \geq 0 \tag{1.6}$$

for $t \in [0, 1]$. The functions $f, r_0, r_1, \mathfrak{g}^y$ and $\mathfrak{g}^u$ are assumed to be at least twice continuously differentiable.

Besides the Lagrange type cost functional (1.1), the Mayer cost function

$$J_1(y(1)) \to \min$$

and the Bolza cost function

$$J_1(y(1)) + \int_0^1 \mathfrak{J}(y(t), u(t))\, dt \to \min$$

are widely used. It is well known that all three formulations are mutually equivalent.

1.2 Methods for Constrained Optimization

1.2.1 Indirect Methods

The classical approach to solving optimal control problems as formulated above is based on PONTRJAGIN's minimum principle [45]. The minimum principle states the existence of *adjoint variables* λ (also called co-states) that satisfy *adjoint differential equations* and *transversality conditions*, and characterizes the optimal control u^* as implicit function of the states y and the adjoint variables λ.

Theorem 1.2.1. (Pontrjagin's Minimum Principle) *Let (y, u) be an optimal solution of (1.1–1.5). Then there exist adjoint variables λ satisfying the adjoint equations*

$$\dot{\lambda}(t) = -\partial_y f(y(t), u(t))^T \lambda(t) \tag{1.7}$$

and Lagrange multipliers α_0 and α_1 satisfying the transversality conditions

$$\lambda(0) = r_0'(y(0))^T \alpha_0 \tag{1.8}$$
$$\lambda(1) = r_1'(y(1))^T \alpha_1\,. \tag{1.9}$$

Furthermore, the Hamiltonian $\mathcal{H}(y, u, \lambda) := \mathfrak{J}(y, u) + \lambda^T f(y, u)$ is pointwise minimized by the optimal control u in the set of all admissible controls:

$$\mathcal{H}(y(t), u(t), \lambda(t)) = \min_{\mathfrak{g}^u(\tilde{u}) \geq 0} \mathcal{H}(y(t), \tilde{u}, \lambda(t)) \tag{1.10}$$

Often a partitioning $u = (u_1, u_2)$ such that $\mathcal{H}(y, u_1, u_2, \lambda) = \mathcal{H}^l(y, u_1, \lambda) + S(y, \lambda)^T u_2$ can be used to determine u as a function of y and λ in the case of box control constraints $\beta \leq u_2 \leq \gamma$. The nonlinearly occurring part u_1 is determined implicitly by $\mathcal{H}^l_{u_1} = 0$, and u_2 by

$$u_2 = \begin{cases} \beta & S > 0 \\ \gamma & S < 0 \end{cases}\,. \tag{1.11}$$

Once the switching structure (1.11) of the optimal solution is known, the necessary conditions (1.7) up to (1.9) together with the original state equation (1.2) and boundary conditions (1.3,1.4) lead to a boundary value problem.

Introduction of the switching times τ_i as additional variables and the switching conditions $S(y(\tau_i), \lambda(\tau_i)) = 0$ then leads to a multipoint boundary value problem in the simplest case (ignoring singular subarcs with $S \equiv 0$). These multipoint boundary value problems can be solved by multiple shooting or collocation methods, yielding the solution of the optimization problem in most cases of real life applications.

In the presence of state constraints (1.6), additional differential algebraic equations of possibly higher order may occur, and additional entry and exit conditions associated with state constrained arcs must be satisfied.

Unfortunately, the switching structure is in general not known a-priori and may be difficult to obtain for real life problems. Thus, an interactive iterative process involving the solution of several multipoint boundary value problems and analytic calculations is often necessary to determine the correct switching structure. Both insight into the actual problem and understanding of the mathematical background are typically required for this task. Consequently, indirect methods are most often applied when the expected improvement is significant and enough time for obtaining the solution is available, e.g. in the aerospace problem domain.

1.2.2 Direct Methods

In contrast to aerospace applications, industrial applications often require less accurate solutions but impose restrictions on solution time and expense. Therefore, the increasing demand for optimization in industrial applications led to the development of *direct methods*.

In a first step, the control functions u are parameterized by, e.g. piecewise constant or piecewise linear functions on a suitably chosen grid $0 = t_1 < t_2 < \cdots < t_n = 1$, substituting u with $u_i = u(t_i)$. The state equation and accordingly the state variables y are discretized by either a multiple shooting approach or a collocation method. The inequality constraints are typically considered only on the grid points in the form $g^u(u_i) \geq 0$ and $g^y(y_i) \geq 0$, where $1 \leq i \leq n$, accepting state inequality constraint violations of the order of the discretization error. This discretization turns the optimal control problem into a finite dimensional nonlinear program

$$J(u, y) \rightarrow \min$$

subject to

$$\begin{aligned} c(u, y) &= 0 \\ g(u, y) &\geq 0 \,. \end{aligned}$$

In a second step, standard methods are used to solve the nonlinear program. Most often, sequential quadratic programming methods are used, but occasionally augmented Lagrangian, projected gradient and penalty methods are applied. They involve the solution of a sequence of linear equation systems or linear-quadratic programs, which in turn can be solved by some standard method or interior point method. See for example the textbooks by Fletcher [27], Luenberger [36], or Polak [44] for an overview. Interior point methods will be discussed in Section 1.3.

The Kuhn-Tucker conditions and the collocation or multiple shooting discretization lead to a special sparse structure of the arising linear equation systems. This can be utilized by standard band solvers or special recursive KKT solvers [47] for an efficient solution process.

Direct methods do not need a-priori information about the switching structure. It is determined automatically in the course of the solution of the nonlinear programs.

The direct method with multiple shooting and collocation approach has been realized by Bock et. al. for parameter estimation problems in the codes PARFIT and COLFIT [6] and by Bock and Plitt in the optimal control code MUSCOD [7].

The respective advantages of direct and indirect methods can be favorably combined in a two-stage approach proposed by Bulirsch, Nerz, Pesch, and von Stryk [12], who compute the initial guess for the indirect multiple shooting method using a direct collocation method. In particular, the Lagrange multipliers of the solution of the discrete nonlinear programming problem can be used to generate good estimates for the adjoint variables of the indirect method, which are difficult to obtain otherwise.

1.3 Interior Point Methods

In this section we will give a necessarily short survey of the development of interior point methods using the same notation as above instead of the *standard form* popular in linear programming.

For finite-dimensional programs the requirement of the iterates to stay in the feasible set $M := \{x : g(x) \geq 0\}$ can be achieved by the addition of a penalty term to the original cost functional:

$$\tilde{J}(x) := J(x) - \mu \sum_i B(g_i(x))$$

subject to $c(x) = 0$. If $B(\xi) \to -\infty$ for $\xi \to 0$, the minimum of $\tilde{J}$ lies in the interior of the feasible region M. Under suitable assumptions, following the path $x(\mu)$ of minimizers to $\mu \to 0$ then leads to the solution x^*. Such *barrier* methods, among which the logarithmic barrier method with $B(\xi) = \log \xi$ is particularly popular, have first been analyzed by Fiacco and McCormick [25] in 1968. Unfortunately, the resulting equality constrained minimization problems become more and more ill-conditioned for $\mu \to 0$, due to the dominating rank-1 terms in the Hessian

$$\tilde{J}''(x) = J''(x) - \mu \sum_i \frac{g_i''(x) g_i(x) - g_i'(x) g_i'(x)^T}{g_i(x)^2}$$

orthogonal to the nearly active constraints with $g_i(x) \approx 0$. Due to the numerical problems introduced by the ill-conditioned subproblems and the necessity of finding a feasible initial point, barrier methods (as penalty methods in general) have been somewhat unpopular.

On the other hand, *active set* methods, frequently used in linear optimization in the form of the well-known *simplex* method, exhibit an exponential worst case complexity. Although not encountered in practice, the bad worst case behavior of active set methods motivated the search for different algorithms with polynomial complexity. The first of them, the ellipsoid method, has been presented in 1979 by Khachiyan [33], but turned out to be much slower for practical problems than the simplex method.

The breakthrough has been achieved by Karmarkar [31] in 1984 who proved polynomial complexity for his algorithm, that achieved competitive performance and turned out to be a *logarithmic barrier* method afterwards. Since then, a remarkable revival of barrier methods has attracted a lot of research in this direction, resulting in a large variety of highly efficient *interior point* methods. By introduction of Lagrange multipliers

$$\eta := \frac{\mu}{g(x)} \tag{1.12}$$

and reformulation of the multiplier equation (1.12) in the form $\eta \cdot g(x) = \mu$, the ill-conditioning can even be overcome. Here, quotient and product $/, \cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ are to be taken component wise. The resulting *primal-dual* interior point approach based on

$$\begin{aligned} J'(x) - c'(x)^* \lambda - g'(x)^* \eta &= 0 \\ c(x) &= 0 \\ \eta \cdot g(x) &= \mu \end{aligned}$$

is generally assumed to be most efficient.

Nesterov and Nemirovskii [43] established a general convergence theory for interior point methods applied to convex programming problems using *self-concordant* barrier functions, where they can show superlinear convergence. At present, primal-dual interior point methods seem to be more efficient than the simplex method, especially for large linear programs.

Currently, interior point methods are extended into four main directions. Much attention is gained by the application of interior point methods to *semi-definite programming* problems which arise as relaxations of discrete optimization problems. Here, *matrix-valued* variables occur, and the inequalitiy $A \succeq 0$ denotes the positive semidefiniteness of A.

A second branch of research activity is targeted at extending interior point methods to *nonlinear* and *nonconvex* programming problems [24], though not without difficulties [52].

The application of interior point techniques to the more general class of nonlinear complementarity problems

$$\begin{aligned} \langle x, c(x) \rangle &= 0 \\ x \geq 0\,,\; c(x) &\geq 0 \end{aligned} \tag{1.13}$$

has been studied in the last decade. In this context, very similar *noninterior point* or *complementarity* methods allowing iterates to leave the feasible region have become popular [15].

The fourth direction aims at solving *semi-infinite programming* problems involving an infinite number of inequality constraints. See the survey paper by Reemtsen et. al. [46]. Although interior point methods are highly efficient especially for large scale problems, most straightforward extensions to (semi-)infinite problems seem to fail for two reasons: First, the sum over logarithmic barrier functions is replaced by an integral, which no longer forms a self-concordant barrier function. In general, it does not even form a barrier function at all [29]. Second, the convergence theory of interior point methods is in general formulated in terms of the number of unknowns and constraints, and currently the best known convergence rate estimates are of order $\mathcal{O}(1 - \frac{1}{\sqrt{n}})$. The decrease of convergence speed with increasing number of constraints seems to be caused by the incompatibility of the norms in L_2 and L_∞. Nevertheless, some progress has been made in showing convergence of some interior point algorithms applied to a restricted class of infinite-dimensional problems, see e.g. Todd [48], Todd and Tunçel [49] and Vanderbei [50].

Furthermore, interior point methods have been applied successfully in the context of direct methods for solving optimal control problems [47] and topology optimization [37]. This indicates the applicability of interior point techniques to *infinite dimensional* optimization problems, because even if only discretized problems are solved, the properties of the underlying infinite dimensional problem tend to govern the solution process once the discretization is sufficiently fine.

In the following chapter, we will address the difficulties that have to be overcome when extending interior point methods to optimal control problems.

2 Complementarity Methods for Optimal Control

Life does not consist mainly,
or even largely, of facts and happenings.
It consists mainly of the storm of thought
that is forever flowing through one's head.
— Mark Twain —

After stating the problem we will analyze the difficulties encountered by complementarity methods applied to infinite dimensional optimization problems. Subsequently, we will propose a way to overcome these difficulties.

2.1 Statement of the Problem

Let $X := X_u \times X_y$ be a function space over the domain $\Omega := [0, 1]$. Assume X_y contains functions of higher regularity than X_u such that $\dot{y} \in X_u$ for $y \in X_y$. We will always use $x = (u, y)$ interchangeably for elements of X. Assume $\mathfrak{J} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$, $\mathfrak{c} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_c}$, $\mathfrak{c}^r : \mathbb{R}^{n_y} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_r}$, and $\mathfrak{g}^u : \mathbb{R}^{n_u} \to \mathbb{R}^{m_u}$ are at least twice, $\mathfrak{g}^y : \mathbb{R}^{n_y} \to \mathbb{R}^{m_y}$ at least three times Lipschitz-continuously differentiable. We will consider the optimization problem

$$\int_\Omega \mathfrak{J}(u(t), y(t))\, dt \to \min$$

subject to

$$\begin{aligned} \mathfrak{c}(u(t), y(t), \dot{y}(t)) &= 0 \quad \text{a.e.} \\ \mathfrak{c}^r(y(0), y(1)) &= 0 \\ \mathfrak{g}^u(u(t)) &\geq 0 \quad \text{a.e.} \\ \mathfrak{g}^y(y(t)) &\geq 0 \quad \text{a.e.} \end{aligned}$$

In order to achieve a more compact functional analytic notation, we define the functional

$$J(x) := \int_\Omega \mathfrak{J}(u(t), y(t))\, dt$$

and the equality constraints operator c on X by

$$c(x) := \begin{bmatrix} \tilde{c}(x) \\ \mathfrak{c}^r(y(0), y(1)) \end{bmatrix} ,$$

where

$$\tilde{c}(x)(t) := \mathfrak{c}(u(t), y(t), \dot{y}(t)) \text{ for } t \in \Omega \, .$$

The inequality constraints operator g on X is defined in a similar way by

$$g(x) := \begin{bmatrix} \mathbf{g}^{\mathbf{u}}(u) \\ \mathbf{g}^{\mathbf{y}}(y) \end{bmatrix} ,$$

where

$$\mathbf{g}^{\mathbf{u}}(u)(t) := \mathfrak{g}^u(u(t)) \quad \text{and} \quad \mathbf{g}^{\mathbf{y}}(y)(t) := \mathfrak{g}^y(y(t)) \quad \text{for } t \in \Omega \, .$$

The optimal control problem above can then be written as

$$J(x) \to \min$$

subject to

$$\begin{aligned} c(x) &= 0 \\ g(x) &\geq 0 \, . \end{aligned}$$

Remark 2.1.1. The restriction to the domain $\Omega = [0, 1]$ is primarily for convenience and no limitation at all, since any domain $[a, b]$ is easily mapped to Ω by an affine transformation. This also holds for optimal control problems with unknown end time, such as minimal time problems, where we map $[T_0, T]$ to Ω and introduce T as a new scalar control variable. ◁

2.2 Necessary and Sufficient Optimality Conditions

Necessary optimality conditions have first been given by Karush [32] and Kuhn and Tucker [34] in the context of (non)linear programming. Necessary as well as sufficient optimality conditions have then been given in various forms by several researchers, cf. the survey article by Kurcyusz [35] or the generalized conditions by Maurer and Zowe [40] for infinite dimensional problems. Compared to the finite dimensional setting, stronger assumptions are necessary to guarantee the existence of Lagrange multipliers for infinite dimensional problems.

In the most abstract setting, an optimization problem can be stated as

$$\begin{aligned} \text{minimize} \quad & J(x) \\ \text{subject to} \quad & g(x) \in K, \end{aligned} \tag{2.1}$$

where J is a functional defined on a real Banach space X, g a map from X into a real Banach space Z and K a closed convex cone in Z. Since any closed convex cone $K \subset Z$ defines a partial ordering on Z, we will write $z_1 \leq z_2$ and $z_2 - z_1 \in K$ interchangeably. Let $K^+ := \{l \in Z^* : \langle l, k \rangle \geq 0 \text{ f.a. } k \in K\}$ denote the dual (or polar) cone of K. J and g are assumed to be twice continuously Fréchet differentiable. The following theorems have been proved in [40].

Definition 2.2.1. A point $x \in X$ is called *regular*, if

$$0 \in \operatorname{int}(g(x) + g'(x)X - K),$$

where int denotes the topological interior.

Theorem 2.2.2. (Necessary conditions) *Let x be a regular solution of problem* (2.1)*. Then there is some $l \in K^+$ such that*

$$\begin{aligned} J'(x) - g'(x)^* l &= 0 \\ \langle l, g(x) \rangle &= 0. \end{aligned} \tag{2.2}$$

Theorem 2.2.3. (Sufficient conditions) *Suppose x is a regular point of problem* (2.1)*. Assume there is a Lagrange multiplier $l \in K^+$ such that $J'(x) - g'(x)^* l = 0$ and $\langle l, g(x) \rangle = 0$. Let $L(x) = J(x) - \langle l, g(x) \rangle$. Suppose that there are $\delta > 0$ and $\beta > 0$ such that*

$$\langle L''(x)h, h \rangle \geq \delta \|h\|^2 \tag{2.3}$$

for all h with $g'(x)h \in K + \mathbb{R}g(x)$ and $\langle l, g'(x)h \rangle \leq \beta \|h\|$. Then there exists a neighborhood U of x such that x is the unique local solution of (2.1) *in U.*

As has been observed by Maurer [39], the assumptions of Theorem 2.2.3 are in general not satisfied for optimal control problems, since most often the solution x^* is regular only if $X \subset L_\infty$, but the Hessian $L''(x^*)$ is coercive only if $L_2 \subset X$. This is illustrated in the following example.

Example 2.2.4. Consider the trivial optimization problem

$$\int_\Omega u^2 \, dx \to \min$$

subject to

$$-1 \leq u \leq 1 \text{ a.e.}$$

The optimal solution is trivially $u^* = 0$ with completely inactive constraints and therefore vanishing Lagrange multipliers $\eta_1 = \eta_2 = 0$.

Theorem 2.2.2 assumes regularity

$$0 \in \operatorname{int}(g(u^*) + g'(u^*)X_0 - K)$$

with $g(u) = (u+1, 1-u)^T$ and $K := \{(a,b)^T : a \geq 0, b \geq 0\}$ in order to guarantee the existence of Lagrange multipliers satisfying the Kuhn-Tucker conditions (2.2). Here,

$$\begin{aligned} U &:= \operatorname{int}(g(0) + g'(0)X_u - K) \\ &= \left\{(1+\nu-a, 1-\nu-b)^T : \nu \in X_u, a \geq 0, b \geq 0\right\}. \end{aligned}$$

For 0 to be in the interior of U there must be $\nu \in X_0$ and $a, b \in K$ such that

$$\begin{matrix} 1+\nu-a = \epsilon_a \\ 1-\nu-b = \epsilon_b \end{matrix} \Rightarrow a+b = 2-\epsilon_a-\epsilon_b \text{ a.e.}$$

for every $\epsilon_a, \epsilon_b \in B(0,\rho)$ with sufficiently small ρ. Therefore u^* is regular if and only if $X_u \subset L_\infty$.

On the other hand, Theorem 2.2.3 requires $L''(u^*)$ to be positive definite on the whole space X_u in this case. Since $L''(0)(h,h) = 2\|h\|_2^2$, this is only satisfied if $L_2 \subset X_u$, in contradiction to the requirement $X_u \subset L_\infty$ for regularity of u^*. Therefore, Theorem 2.2.3 is not applicable. ◁

This *two-norm-discrepancy* can be addressed using two different norms for the space itself and the coercivity condition (2.3):

Theorem 2.2.5. (Sufficient conditions) *Suppose x is a regular point for problem* (2.1)*. Assume that J and g are defined and twice continuously Fréchet differentiable on some larger space $X_p \supset X$. Assume* (2.2) *holds and let $L(x) = J(x) - \langle l, g(x)\rangle$. Suppose that there are $\delta > 0$ and $\beta > 0$ such that*

$$\langle L''(x)h, h\rangle \geq \delta\|h\|_p^2$$

for all h with $g'(x)h \in K + \mathbb{R}g(x)$ and $\langle l, g'(x)h\rangle \leq \beta\|h\|_p$. Then there exists a neighborhood U of x in X_p such that x is the unique local solution of (2.1) *in U.*

For recent developments in the area of sufficient conditions in the presence of two-norm discrepancy we refer to Malanowski [38].

2.2.1 The Role of Lagrange Multipliers

Besides allowing a characterization of the solution, the Lagrange multipliers l associated with an optimal point x by the Kuhn-Tucker conditions (2.2) have an interpretation as *sensitivities* of the cost functional with respect to violations of the corresponding constraints (cf. LUENBERGER [36]).

More technically, under suitable assumptions one can show that if x^* is a solution of (2.1) with associated Lagrange multiplier l, then there is a mapping $z \mapsto x(z)$ for sufficiently small perturbations $z \in Z$, such that $x(z)$ solves

$$\min J(x) \quad \text{subject to} \quad g(x) - z \in K\,. \tag{2.4}$$

Furthermore, the derivative of the optimal value of the perturbed problem (2.4) is

$$\partial_z J(x(z))|_{z=0} = l\,.$$

In particular, this equality constitutes the basis for the construction of weighted error estimators for adaptive methods tailored for optimization problems [4].

Moreover, jointly vanishing Lagrange multipliers of (weakly) active constraints can indicate a nonunique solution.

Example 2.2.6. Consider the artificial optimization problem

$$\begin{aligned} \min -y(1) \quad \text{subject to} \quad y(0) &= 0 \quad & u &\le 1 \\ \dot y &= u & y &\le \max(1-3t, 0, 3t-2) \end{aligned}$$

with an obvious solution

$$y = \max\left(0, t - \frac{2}{3}\right) \quad \text{and} \quad u = \begin{cases} 0, & t < \frac{2}{3} \\ 1, & t > \frac{2}{3} \end{cases}$$

and the corresponding Lagrange multipliers

$$\begin{aligned} \lambda &= \begin{cases} 0, & t < \frac{2}{3} \\ 1, & t > \frac{2}{3} \end{cases} & \lambda^r &= 0 \\ \eta^u &= \lambda & \eta^y &= \delta_{\frac{2}{3}} \end{aligned}$$

satisfying the first order necessary conditions (2.2). The fact that the Lagrange multipliers vanish on the whole interval $[0, 2/3)$ already indicates that the solution is not unique and the control u can be changed without altering the functional's value. In fact, every feasible control $\tilde u$ with

$$\int_0^{\frac{2}{3}} \tilde u \, dt = 0$$

generates an optimal solution (cf. Figure 2.1). A similar situation is encountered in the abort landing problem treated in Section 4.3.

Note, however, that Lagrange multipliers can vanish for unique solutions, too, and thus do not imply nonuniqueness. If the cost functional in this example is augmented with $\epsilon \|u\|_{L_2}^2$, where ϵ is sufficiently small, then the solution is unchanged (except for η^u on $[2/3, 1]$), but unique. ◁

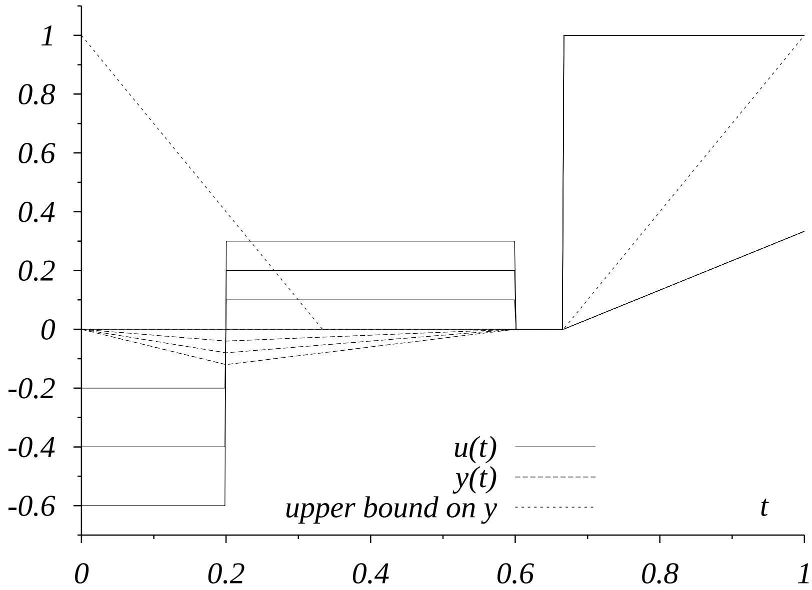

Figure 2.1: Some optimal solutions from Example 2.2.6. Recall that on $[0, 2/3)$ both Lagrange multipliers λ and η vanish, and that there is a whole continuum of optimal solutions.

2.3 Infinite dimensional Complementarity Methods

A quite common strategy for solving optimization problems of type (2.1) is to compute Karush-Kuhn-Tucker points and Lagrange multipliers satisfying (2.2) which are promising candidates for local solutions. This is a difficult task because of the *complementarity condition*

$$\begin{aligned} \langle g(x), \eta \rangle &= 0 \\ g(x) \geq 0,\ \eta &\geq 0\,. \end{aligned} \tag{2.5}$$

Formally, the complementarity approach is to substitute (2.5) by the equivalent conditions

$$\begin{aligned} \Psi(w, \eta) &= 0 \\ w - g(x) &= 0\,, \end{aligned} \tag{2.6}$$

where Ψ results from the pointwise application of a *complementarity function* ψ with

$$\psi(a, b) = 0 \Leftrightarrow \begin{matrix} ab = 0 \\ a \geq 0,\ b \geq 0\,, \end{matrix}$$

and the introduction of *slack variables* w is as usual supposed to reduce non-linear coupling between g and Ψ. In the optimal control context, the complementarity function Ψ can be divided into separate complementarity parts for the control and state constraints, respectively:

$$\Psi = \begin{bmatrix} \Psi^u \\ \Psi^y \end{bmatrix} .$$

Since no such complementarity function can be continuously differentiable, a suitable smoothing of the complementarity function, in terms of a parameter μ, is used to define a homotopy path that leads to the solution for $\mu \to 0$. Such complementarity methods have been studied recently by Kanzow [30], Burke and Xu [13], and Chen and Xiu [14] for finite dimensional complementarity problems.

Several smoothed complementarity functions have been suggested, e.g. a family of smoothings of the plus function $(a)_+ := \max(0, a)$ by Chen and Mangasarian [15] or the smoothed Fischer-Burmeister function [26]

$$\psi_{\mathrm{FB}}(a, b; \mu) := a + b - \sqrt{a^2 + b^2 + 2\mu}\,, \tag{2.7}$$

which generates exactly the same homotopy as the usual interior point method with $ab = \mu$.

Remark 2.3.1. The use of complementarity functions has two main advantages over primal-dual interior point methods:

- The iterates are not required to stay in the feasible region. Therefore, infeasible starting points can be used, which may be easier to generate than feasible starting points.

- If a step leaves the feasible region, no step size reduction is required for the method to be well defined. With interior point methods, the additional algorithmic decision of how far away from the boundary to stay is necessary. ◁

The homotopy problem defined by the complementarity formulation (2.6) is usually tackled by Newton continuation algorithms, which require the involved functions to be continuously differentiable. As demonstrated below, the complementarity functions violate this requirement when applied to functions from spaces less regular than L_∞. Note that this is a distinct feature of the infinite dimensional setting.

Consider the sequence

$$w_i := \begin{cases} 1, & t < \frac{1}{i} \\ 0, & \text{otherwise} \end{cases}$$

with $w_i \to 0$ in L_p for all $p < \infty$. Using the Fischer-Burmeister function (2.7) we have

$$\Psi_w(w_i, 0; \mu) - \Psi_w(0, 0; \mu) = \frac{w_i}{\sqrt{w_i^2 + \mu}} \,,$$

such that

$$\|\Psi_w(w_i, 0; \mu) - \Psi_w(0, 0; \mu)\| \geq \frac{1}{\sqrt{1+\mu}} \not\to 0 \,.$$

Thus, the differentiability requirement imposes the regularity requirement

$$w \in L_\infty^m \quad \text{and} \quad \eta \in L_\infty^m \tag{2.8}$$

on slacks and Lagrange multipliers.

Note that the nondifferentiability is not specific to the complementarity functions. The constraints $\mathfrak{g}^u$ and $\mathfrak{c}$ are also affected, if u appears nonlinearly.

Remark 2.3.2. Interior point methods applied to infinite dimensional problems are affected in a similar pathologic way. Since the interior of the positive cone is empty in L_p for $p < \infty$, there may occur continuation directions that violate the positivity condition for every positive step size. Although this will never be the case in actual computation with a reasonable discretization, the performance of the method may deteriorate significantly in the course of adaptive refinement.

The fact that all estimations of convergence rates of IP methods for linear programming get worse for more unknowns may have exactly that reason. For an increasing number of unknowns, the hidden L_2-structure of the problem becomes visible and the possible step sizes get smaller. $\triangleleft$

Remark 2.3.3. As pointed out recently by Wächter and Biegler [52], interior point methods applied directly to nonlinear problems may fail to converge for one more reason. They constructed a simple well posed example where for a large set of starting points and every $\mu > 0$ the Newton path for the interior point corrector leads to the boundary of the feasible set. There the iteration terminates prematurely with a singular derivative. $\triangleleft$

On the other hand, the first order necessary conditions of Theorem 2.2.2 guarantee only $\eta_u \in X_u^* \supset X_u$ and $\eta_y \in X_y^* \supset X_y$. In fact, measure valued Lagrange multipliers occur frequently in state constrained optimal control problems. Furthermore, convergence of the central path to the solution cannot be expected in L_∞, as the trivial example in the appendix shows.

To overcome this gap, we notice that with the additional assumption

$$x(\mu) \geq \epsilon(\mu) > 0 \quad \text{a.e.}$$

for points $x(\mu)$ *on the central path*, the Lagrange multipliers are bounded by $\eta(\mu) \le \frac{\mu}{\epsilon(\mu)}$. Hence, the regularity requirement (2.8) can be satisfied for all $\mu > 0$.

Consequently, the choice of spaces is

$$\begin{aligned} X &:= X_u \times X_y := L_\infty^{n_u} \times (W^1_\infty)^{n_y} \\ \Lambda &:= L_\infty^{n_c} \times \mathbb{R}^{n_r} \\ W &:= W_u \times W_y := L_\infty^{m_u} \times (W^1_\infty)^{m_y} \end{aligned}$$

with

$$V := X \times \Lambda \times W \times W$$

being the space of all unknowns $v = (x, \lambda, \eta, w)^T$. Convergence will be discussed in the space $V_p := X_p \times \Lambda_p \times W_p \times W_p$ for $p < \infty$ with $X_p := L_p^{n_u} \times (W_p^1)^{n_y}$, $\Lambda_p := L_p^{n_c} \times \mathbb{R}^{n_r}$, and $W_p := L_p^{m_u} \times (W_p^1)^{m_y}$.

Combining the KKT equations (2.2) and the complementarity formulation (2.6), the system defining the central path $v(\mu)$ implicitly by $F(v(\mu); \mu) = 0$ is

$$F(v;\mu) := \begin{bmatrix} J'(x) - c'(x)^*\lambda - g'(x)^*\eta \\ -c(x) \\ w - g(x) \\ \Psi(w, \eta; \mu) \end{bmatrix}. \tag{2.9}$$

This is the formulation to be attacked algorithmically.

In the following, we will investigate existence and convergence of the homotopy path defined by the complementarity formulation above.

2.4 Nemyckii operators in L_∞ and W^1_∞

Since the optimization problem stated in Section 2.1 is made up of Nemyckii operators, we first study basic properties of these operators in L_∞ and W^1_∞. The overall notation is borrowed from [54], Chapter 4.

Definition 2.4.1. For a domain $\Omega \subset \mathbb{R}^l$ and a mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ the corresponding *Nemyckii operator* on $L_1(\Omega)$ is defined by $\mathbf{f}(u)(t) := f(u(t))$.

Lemma 2.4.2. *For k functions $f_j \in L_{pk}$ with $k \ge 1$ and $1 \le p \le \infty$ the following extended Hölder inequality holds:*

$$\Big\| \prod_{j=1}^k f_j \Big\|_{L_p} \le \prod_{j=1}^k \|f_j\|_{L_{pk}}$$

Proof. Let $k' := \frac{k}{k-1}$. Then, by Hölder's inequality and by induction,

$$\begin{aligned}\Big\|\prod_{j=1}^{k} f_j\Big\|_{L_p} &= \Big\|\prod_{j=1}^{k} |f_j|^p\Big\|_{L_1}^{\frac{1}{p}} \leq \||f_k|^p\|_{L_k}^{\frac{1}{p}}\Big\|\prod_{j=1}^{k-1} |f_j|^p\Big\|_{L_{k'}}^{\frac{1}{p}} \\ &= \|f_k\|_{L_{pk}}\Big\|\prod_{j=1}^{k-1} f_j\Big\|_{L_{pk'}} \leq \|f_k\|_{L_{pk}}\prod_{j=1}^{k-1}\|f_j\|_{L_{(k-1)pk'}} = \prod_{j=1}^{k}\|f_j\|_{L_{pk}}\,. \quad \square\end{aligned}$$

Theorem 2.4.3. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is k times differentiable and its k-th derivative satisfies the Lipschitz condition*

$$|f^{(k)}(x) - f^{(k)}(y)| \leq \kappa|x-y|\,, \tag{2.10}$$

the corresponding Nemyckii operator $\mathbf{f}$ maps L_∞^n into L_∞^m and is k times differentiable. For $p \geq 1$ its k-th derivative can be continuously extended to an operator $\mathbf{f}^{(k)}(u) : (\prod_{j=1}^k L_{pk}^n) \to L_p^m$ that inherits boundedness and Lipschitz continuity from $f^{(k)}$:

$$\|\mathbf{f}^{(k)}(u)\|_{(\prod_{j=1}^k L_{pk}^n)\to L_p^m} \leq \sup_{|x|\leq\|u\|_{L_\infty^n}} |f^{(k)}(x)| \tag{2.11}$$

$$\|\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u)\|_{(\prod_{j=1}^k L_{pk}^n)\to L_p^m} \leq \kappa\|\delta u\|_{L_\infty^n} \tag{2.12}$$

Proof. For $u \in L_\infty^n$ let $M := \|u\|_\infty$. Since f is continuous by assumption, f is bounded on the closed ball $\overline{B(0,M)}$. Therefore,

$$\|\mathbf{f}(u)\|_\infty = \operatorname*{ess\,sup}_{t\in\Omega} |f(u(t))| < \infty\,.$$

To begin with, Definition 2.4.1 can be written as $\mathbf{f}^{(0)}(u)(t) = f^{(0)}(u(t))$. By induction, assume $i < k$ and the i-th derivative of $\mathbf{f}$ is given by

$$(\mathbf{f}(u)h_i \dots h_1)(t) = f^{(i)}(u(t))h_i(t)\dots h_1(t) \quad \text{for all } t \in \Omega.$$

Then we have

$$\begin{aligned}&((\mathbf{f}^{(i)}(u+h_{i+1}) - \mathbf{f}^{(i)}(u))h_i \dots h_1)(t) \\ &\qquad = \left(f^{(i)}(u(t)+h_{i+1}(t)) - f^{(i)}(u(t))\right) h_i(t)\dots h_1(t) \\ &\qquad = \int_0^1 f^{(i+1)}(u(t)+sh_{i+1}(t))h_{i+1}(t)h_i(t)\dots h_1(t)\,ds \\ &\qquad = f^{(i+1)}(u(t))h_{i+1}(t)h_i(t)\dots h_1(t) + r(t)\end{aligned}$$

with

$$r(t) = \int_0^1 \left(f^{(i+1)}(u(t)+sh_{i+1}(t)) - f^{(i+1)}(u(t))\right) h_{i+1}(t)h_i(t)\dots h_1(t)\,ds\,,$$

which yields

$$\|r\|_{L^m_\infty} \le \int_0^1 s\kappa_{i+1}\|h_{i+1}\|^2_{L^n_\infty}\prod_{j=1}^{i}\|h_j\|_{L^n_\infty}\,ds$$
$$\le \frac{1}{2}\kappa_{i+1}\|h_{i+1}\|^2_{L^n_\infty}\prod_{j=1}^{i}\|h_j\|_{L^n_\infty} = \mathcal{O}(\|h_{i+1}\|^2_{L^n_\infty})$$

for sufficiently small h_{i+1}. Here, κ_{i+1} is the Lipschitz constant of $f^{(i+1)}$ on the ball $B(0, 1+\|u\|_{L^n_\infty})$. Consequently, $\mathbf{f}$ has a $i+1$-st derivative defined by

$$(\mathbf{f}^{(i+1)}(u)h_{i+1}\dots h_1)(t) = f^{(i+1)}(u(t))h_{i+1}(t)\dots h_1(t)\,.$$

$f^{(k)}$ is bounded on bounded sets due to its continuity, and hence

$$\|\mathbf{f}^{(k)}(u)h_k\dots h_1\|_{L^m_p} \le \Big\| \sup_{|x|\le\|u\|_{L^n_\infty}} |f^{(k)}(u)|\prod_{j=1}^{k}|h_j|\Big\|_{L_p}$$
$$\le \sup_{|x|\le\|u\|_{L^n_\infty}} |f^{(k)}(u)|\prod_{j=1}^{k}\||h_j|\|_{L_{pk}} = \sup_{|x|\le\|u\|_{L^n_\infty}} |f^{(k)}(u)|\prod_{j=1}^{k}\|h_j\|_{L^n_{pk}}$$

for all $h_j \in L^n_\infty$ by Lemma 2.4.2. Thus, $\mathbf{f}^{(k)}(u)$ can be extended continuously to an operator $(\prod_{j=1}^k L^n_{pk}) \to L^m_p$ satisfying (2.11).

Furthermore, using (2.10) and Lemma 2.4.2 we have

$$\|\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u)\|_{(\prod_{j=1}^k L^n_{pk})\to L^m_p}$$
$$= \sup_{h_j\in L^n_{pk}} \frac{\|(\mathbf{f}^{(k)}(u+\delta u)-\mathbf{f}^{(k)}(u))h_k\dots h_1\|_{L^m_p}}{\prod_{j=1}^k\|h_j\|_{L^n_{pk}}} \le \sup_{h_j\in L^n_{pk}} \frac{\Big\|\kappa\|\delta u\|_{L^n_\infty}\prod_{j=1}^k|h_j|\Big\|_{L_p}}{\prod_{j=1}^k\|h_j\|_{L^n_{pk}}}$$
$$\le \sup_{h_j\in L^n_{pk}} \frac{\kappa\|\delta u\|_{L^m_\infty}\prod_{j=1}^k\|h_j\|_{L^n_{pk}}}{\prod_{j=1}^k\|h_j\|_{L^n_{pk}}} = \kappa\|\delta u\|_{L^m_\infty}\,. \quad \square$$

Theorem 2.4.4. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is $k+1$ times differentiable and the k-th and $k+1$-st derivatives satisfy the Lipschitz conditions*

$$|f^{(k)}(x) - f^{(k)}(y)| \le \kappa|x-y| \quad \textit{and} \quad |f^{(k+1)}(x) - f^{(k+1)}(y)| \le \kappa|x-y|\,, \tag{2.13}$$

the corresponding Nemyckii operator $\mathbf{f}$ maps $(W^1_\infty)^n$ into $(W^1_\infty)^m$ and is k times differentiable. For $p \ge 1$ its k-th derivative can be continuously extended to *an operator* $\mathbf{f}^{(k)}(u) : (\prod_{j=1}^{k}(W_{pk}^1)^n) \to (W_p^1)^m$ *that inherits boundedness and Lipschitz continuity from* $f^{(k)}$ *and* $f^{(k+1)}$*:*

$$\|\mathbf{f}^{(k)}(u)\|_{(\prod_{j=1}^{k}(W_{pk}^1)^n)\to(W_p^1)^m} \le \sup_{|x|\le\|u\|_{L_\infty^n}} (k+1)|f^{(k)}(x)| + |f^{(k+1)}(x)| \tag{2.14}$$

$$\|\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u)\|_{(\prod_{j=1}^{k}(W_{pk}^1)^n)\to(W_p^1)^m} \le (k+2)\kappa\|\delta u\|_{(W_\infty^1)^n}$$

Proof. By Theorem 2.4.3, $\mathbf{f}$ maps $(W_\infty^1)^n$ into L_∞^n. In addition, the weak derivative $(\mathbf{f}(u))_t(t) = f'(u(t))\dot{u}(t)$ is bounded almost everywhere, and thus $\mathbf{f}(u) \in (W_\infty^1)^m$.

For establishing the differentiability, we proceed as in the proof of Theorem 2.4.3, but consider additionally the derivative of the residual term, which turns out to be

$$\begin{aligned}
\dot{r}(t) = &\int_0^1 \left(f^{(i+2)}(u(t)+sh_{i+1}(t)) - f^{(i+2)}(u(t))\right)\dot{u}(t)h_{i+1}(t)h_i(t)\dots h_1(t)\,ds \\
&+ \int_0^1 sf^{(i+2)}(u(t)+sh_{i+1}(t))\dot{h}_{i+1}(t)h_{i+1}(t)h_i(t)\dots h_1(t)\,ds \\
&+ \int_0^1 \left(f^{(i+1)}(u(t)+sh_{i+1}(t)) - f^{(i+1)}(u(t))\right)\dot{h}_{i+1}(t)h_i(t)\dots h_1(t)\,ds \\
&+ \int_0^1 \sum_{j=1}^{i} \left(f^{(i+1)}(u(t)+sh_{i+1}(t)) - f^{(i+1)}(u(t))\right)h_{i+1}(t) \\
&\qquad\qquad h_i(t)\dots h_{j+1}(t)\dot{h}_j(t)h_{j-1}(t)\dots h_1(t)\,ds\,.
\end{aligned}$$

This gives rise to

$$\begin{aligned}
\|\dot{r}\|_{L_\infty^n} \le\ &\kappa_{i+2}\|h_{i+1}\|_{L_\infty^n}\|\dot{u}\|_{L_\infty^n}\|h_{i+1}\|_{L_\infty^n}\prod_{j=1}^{i}\|h_j\|_{L_\infty} \\
&+ \frac{1}{2}\|f^{(i+2)}\|\,\|\dot{h}_{i+1}\|_{L_\infty^n}\|h_{i+1}\|_{L_\infty^n}\prod_{j=1}^{i}\|h_j\|_{L_\infty} \\
&+ \kappa_{i+1}\|h_{i+1}\|_{L_\infty^n}\|\dot{h}_{i+1}\|_{L_\infty^n}\prod_{j=1}^{i}\|h_j\|_{L_\infty} \\
&+ \sum_{j=1}^{i}\kappa_{i+1}\|h_{i+1}\|_{L_\infty^n}^2\|\dot{h}_j\|_{L_\infty^n}\prod_{l\neq j}\|h_l\|_{L_\infty^n} \\
=\ &\mathcal{O}(\|h_{i+1}\|_{W_\infty^1}^2)\,,
\end{aligned}$$

where κ_{i+1} and κ_{i+2} are the Lipschitz constants of $f^{(i+1)}$ and $f^{(i+2)}$, respectively, on $B(0, 1+\|u\|_{L_\infty^n})$. Thus, $\|r\|_{W_\infty^1} = \mathcal{O}(\|h_{i+1}\|_{W_\infty^1})$, such that $\mathbf{f} : (W_\infty^1)^n \to (W_\infty^1)^m$ has a $i+1$-th differential.

Let $M_j := \sup_{|x|\le\|u\|_{L^n_\infty}} |f^{(j)}(x)|$. Utilizing Lemma 2.4.2,

$$
\begin{aligned}
&\|(\mathbf{f}^{(k)}h_k\dots h_1)_t\|_{L^m_p} \\
&\qquad\le \|\mathbf{f}^{(k+1)}(u)\dot{u}h_k\dots h_1\|_{L^m_p} + \sum_{j=1}^{k} \|\mathbf{f}^{(k)}(u)h_k\dots h_{j+1}\dot{h}_j h_{j-1}\dots h_1\|_{L^m_p} \\
&\qquad\le M_{k+1}\left\| |\dot{u}| \prod_{j=1}^{k} |h_j| \right\|_{L_p} + \sum_{j=1}^{k} M_k \left\| |\dot{h}_j| \prod_{i\ne j} |h_i| \right\|_{L^n_p} \\
&\qquad\le M_{k+1}\|\dot{u}\|_{L^n_\infty} \prod_{j=1}^{k} \|h_j\|_{L^n_{pk}} + M_k \sum_{j=1}^{k} \|\dot{h}_j\|_{L^n_{pk}} \prod_{i\ne j} \|h_i\|_{L^n_{pk}} \\
&\qquad\le M_{k+1} + kM_k \prod_{j=1}^{k} \|h_j\|_{(W^1_{pk})^n} ,
\end{aligned}
$$

for $h_j \in (W^1_{pk})^n$, where the constant depends on u and k. Together with the corresponding result of Theorem 2.4.1, this confirms the result (2.14), such that $\mathbf{f}^{(k)}(u)$ can be continuously extended to an operator $(\prod_{j=1}^k (W^1_{pk})^n) \to (W^1_p)^m$.

Furthermore, using (2.13) and Lemma 2.4.2 we have

$$
\begin{aligned}
&\|\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u)\|_{(\prod_{j=1}^k (W^1_{pk})^n)\to(W^1_p)^m} \\
&\qquad\le \sup_{h_1,\dots,h_k\in(W^1_{pk})^n} \frac{\|(\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u))h_k\dots h_1\|_{L^n_p}}{\prod_{j=1}^k \|h_j\|_{(W^1_{pk})^n}} \\
&\qquad\quad + \sup_{h_1,\dots,h_k\in(W^1_{pk})^n} \frac{\|((\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u))h_k\dots h_1)_t\|_{L^n_p}}{\prod_{j=1}^k \|h_j\|_{(W^1_{pk})^n}} \\
&\qquad\le \kappa\|\delta u\|_{L^n_\infty} + \sup_{h_1,\dots,h_k\in(W^1_{pk})^n} \frac{\kappa\|\delta u\|_{(W^1_\infty)^n}(k+1)\prod_{j=1}^k \|h_j\|_{(W^1_{pk})^n}}{\prod_{j=1}^k \|h_j\|_{(W^1_{pk})^n}} \\
&\qquad\le \kappa(k+2)\|\delta u\|_{(W^1_\infty)^n}
\end{aligned}
$$

□

Theorem 2.4.5. *Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is k times continuously differentiable. If $k \ge 1$, $f'(x)$ is symmetric, and coercive in the sense that*

$$\xi^T f'(x)\xi \ge \alpha|\xi|^2$$

uniformly for all $\xi \in \mathbb{R}^n$ and $|x| \le R$, then $\mathbf{f}'(u) : L^n_2 \to L^n_2$ is symmetric positive (semi-)definite for all $u \in L^n_\infty$ with $\|u\|_\infty \le R$. Furthermore, $\mathbf{f}'(u) : L^n_\infty \to L^n_\infty$ has an inverse bounded by α^{-1}.

If $k \ge 1$ and the derivatives of $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^n$ commute, then so do the derivatives of the corresponding Nemyckii operators $\mathbf{f}'$ and $\mathbf{g}'$.

Proof. If $f'(u(t))$ is positive (semi-)definite for $t \in \Omega$, then

$$(\mathbf{f}'(u)\xi, \xi)_2 = \int_\Omega \xi(t)^T f'(u(t))\xi(t)\, dt \geq \int_\Omega \alpha|\xi(t)|^2\, dt = \alpha\|\xi\|_2^2$$

for all $\xi \in L_2^n$. Furthermore, for every $\zeta \in L_\infty^n$ the function ξ defined by

$$\xi(t) := f'(u(t))^{-1}\zeta(t)$$

with $\|\xi\|_\infty \leq \alpha^{-1}\|\zeta\|_\infty$ satisfies $\mathbf{f}'(u)\xi = \zeta$.

If $f'(x)$ and $g'(x)$ commute, then

$$\begin{aligned}
\langle \mathbf{f}'(u)\mathbf{g}'(w)\xi, \zeta\rangle &= \int_\Omega \zeta(t)^T(\mathbf{f}'(u)\mathbf{g}'(w)\xi)(t)\, dt \\
&= \int_\Omega \zeta(t)^T f'(u(t))(\mathbf{g}'(w)\xi)(t)\, dt = \int_\Omega \zeta(t)^T f'(u(t))g'(w(t))\xi(t)\, dt \\
&= \int_\Omega \zeta(t)^T g'(w(t))f'(u(t))\xi(t)\, dt = \int_\Omega \zeta(t)^T(\mathbf{g}'(w)\mathbf{f}'(u)\xi)(t)\, dt \\
&= \langle \mathbf{g}'(w)\mathbf{f}'(u)\xi, \zeta\rangle
\end{aligned}$$

for all ξ, ζ. □

2.5 Existence of the Central Path

Since the central path is defined by the homotopy (2.9) given by the complementarity function Ψ, we first collect some results concerning boundedness, continuity, and invertibility of the operators Ψ_w and Ψ_η occurring in the derivative F_v. Afterwards, the same properties of F_v are addressed.

Lemma 2.5.1. *Using the Fischer-Burmeister complementarity function* (2.6) *with* $\mu \leq 1/2$, *the derivative* $\Psi_w(w, \eta; \mu)$ *is symmetric positive semidefinite, bounded by*

$$\|\Psi_w^u\|_{W_u \to W_u} \leq 2\,, \tag{2.15}$$

$$\|(\Psi_w^u)^{-1}\|_{W_u \to W_u} \leq \max\left(4, 2\frac{\|w\|_{W_u}^2}{\mu}\right)\,, \tag{2.16}$$

$$\|\Psi_w^y\|_{W_y \to W_y} \leq 2 + \frac{(\|w\|_{W_y} + \|\eta\|_{W_y} + 1)^2}{2\mu}\,, \tag{2.17}$$

$$\|(\Psi_w^y)^{-1}\|_{W_y \to W_y} \leq 5\frac{(2\|w\|_{W_y} + \|\eta\|_{W_y} + 1)^2}{\mu}\,, \tag{2.18}$$

and Lipschitz continuous with a Lipschitz constant of $27\mu^{-1}$. *The corresponding holds for* $\Psi_\eta(w, \eta; \mu)$. *Furthermore, the derivatives commute.*

Proof. From $(1+\phi)^{-1/2} \le \max(1-\phi/4, 3/4)$ for $\phi > 0$ we infer

$$\begin{aligned}
\min\left(\frac{\mu}{2a^2}, \frac{1}{4}\right) &= 1 - \max\left(1 - \frac{\mu}{2a^2}, \frac{3}{4}\right) \le 1 - \frac{1}{\sqrt{1+\frac{2\mu}{a^2}}} \\
&\le 1 - \frac{1}{\sqrt{1+\frac{b^2}{a^2}+\frac{2\mu}{a^2}}} = 1 - \frac{|a|}{\sqrt{a^2+b^2+2\mu}} \\
&\le \psi_a(a,b;\mu) \\
&\le 1 + \frac{|a|}{\sqrt{a^2+b^2+2\mu}} \le 2\,.
\end{aligned}$$

Thus, ψ_a is uniformly positive definite. Due to Theorem 2.4.5, the derivative $\Psi^u_w(w,\eta;\mu)$ of the Nemyckii operator Ψ^u is bounded by (2.15) and has an inverse which is bounded by (2.16).

For estimating the operator norms of the derivative in the finer norm of W_y for the state constraints, we first note that

$$\begin{aligned}
\left|\left(\sqrt{w(t)^2+\eta(t)^2+2\mu}\right)_t\right| &= \left|\frac{w(t)\dot{w}(t)+\eta(t)\dot{\eta}(t)}{\sqrt{w(t)^2+\eta(t)^2+2\mu}}\right| \\
&\le |\dot{w}(t)| + |\dot{\eta}(t)| \le \|w\|_{W_y} + \|\eta\|_{W_y} \quad \text{a.e.}
\end{aligned}$$

and therefore

$$\begin{aligned}
&|(\psi_w(w(t),\eta(t);\mu))_t| \\
&\quad= \left|\frac{\dot{w}(t)\sqrt{w(t)^2+\eta(t)^2+2\mu} - w(t)\left(\sqrt{w(t)^2+\eta(t)^2+2\mu}\right)_t}{w(t)^2+\eta(t)^2+2\mu}\right| \\
&\quad\le (2\mu)^{-1}\left(\|w\|_{W_y}(\|w\|_{W_y}+\|\eta\|_{W_y}+1) + \|w\|_{W_y}(\|w\|_{W_y}+\|\eta\|_{W_y})\right) \\
&\quad\le (2\mu)^{-1}(\|w\|_{W_y}+\|\eta\|_{W_y}+1)^2\,.
\end{aligned}$$

Together with (2.15) this confirms the result (2.17). Because of

$$\psi_a(a,b;\mu)^{-1} = \frac{\sqrt{a^2+b^2+2\mu}}{\sqrt{a^2+b^2+2\mu}-a}$$

and

$$\begin{aligned}
\sqrt{a^2+b^2+2\mu} - a &\ge |a|\left(\sqrt{1+\frac{2\mu}{a^2}} - 1\right) \\
&\ge |a|\min\left(\frac{\mu}{2a^2}, 1\right) = \min\left(\frac{\mu}{2|a|}, |a|\right) \ge \sqrt{\frac{\mu}{2}}
\end{aligned}$$

we have

$$
\begin{aligned}
&|(\psi_w(w(t),\eta(t);\mu)^{-1})_t| \\
&\qquad \le \left|\frac{(\sqrt{w^2+\eta^2+2\mu})_t(\sqrt{w^2+\eta^2+2\mu}-w)}{(\sqrt{w^2+\eta^2+2\mu}-w)^2}\right| \\
&\qquad\quad + \left|\frac{\sqrt{w^2+\eta^2+2\mu}(\sqrt{w^2+\eta^2+2\mu}-w)_t}{(\sqrt{w^2+\eta^2+2\mu}-w)^2}\right| \\
&\qquad \le \frac{(\|w\|_{W_y}+\|\eta\|_{W_y})(2\|w\|_{W_y}+\|\eta\|_{W_y}+1)}{\frac{\mu}{2}} \\
&\qquad\quad + \frac{(\|w\|_{W_y}+\|\eta\|_{W_y}+1)(\|w\|_{W_y}+\|\eta\|_{W_y}+\|w\|_{W_y})}{\frac{\mu}{2}} \\
&\qquad \le 4\frac{(2\|w\|_{W_y}+\|\eta\|_{W_y}+1)^2}{\mu}\,.
\end{aligned}
$$

Together with (2.16) this confirms the result (2.18). As for the Lipschitz continuity, we estimate

$$
\begin{aligned}
|\psi_{aa}| &= \left|\frac{\sqrt{a^2+b^2+2\mu}-\frac{a}{\sqrt{a^2+b^2+2\mu}}a}{a^2+b^2+2\mu}\right| \\
&\le \left|\frac{1}{\sqrt{a^2+b^2+2\mu}}\right| + \left|\frac{1}{\sqrt{a^2+b^2+2\mu}}\right| \le \sqrt{\frac{2}{\mu}}
\end{aligned}
$$

and

$$
|\psi_{ab}| = \left|\frac{a\frac{b}{\sqrt{a^2+b^2+2\mu}}}{a^2+b^2+2\mu}\right| \le \left|\frac{1}{\sqrt{a^2+b^2+2\mu}}\right| \le \sqrt{\frac{1}{2\mu}}
$$

to achieve a Lipschitz constant of $3(2\mu)^{-\frac{1}{2}}$ for ψ_a and hence for Ψ_w^u by Theorem 2.4.3. Straightforward computation verifies the bounds $|\psi_{aaa}| \le 3\mu^{-1}$, $|\psi_{aab}| \le 2\mu^{-1}$, $|\psi_{aba}| \le 2\mu^{-1}$, and $|\psi_{abb}| \le 2\mu^{-1}$, from which we conclude a Lipschitz constant of $3 \cdot 9\mu^{-1}$ for Ψ_w^y by Theorem 2.4.4.

Because of symmetry, the analogous holds for Ψ_η, which commutes with Ψ_w by Theorem 2.4.5. □

Theorem 2.5.2. *The complementarity system F defined in* (2.9) *is a continuously differentiable mapping from $V \times]0,\mu_0]$ to $Z := \widehat{X^*} \times \Lambda \times W \times W$ for all $\mu > 0$, where $\widehat{X^*} := X_u \times (W_\infty^{-1})^{n_y}$. Moreover, the derivative F_v satisfies the Lipschitz condition*

$$
\|(F_v(v+\delta v;\mu) - F_v(v;\mu))\delta v\|_{Z_2} \le \operatorname{const}\mu^{-1}\|\delta v\|_{V_2}\,\|\delta v\|_V\,. \tag{2.19}
$$

Proof. First we note that the operators J, c, g, and Ψ can be written as Nemyckii operators of $\mathfrak{J}$, $\mathfrak{c}$, $\mathfrak{g}$, and ψ either directly or as $J(x) = \langle \mathbf{J}(x), \mathbb{1} \rangle$ and $\tilde{c}(x) = \mathbf{c}(Tx)$ with

$$T = \begin{bmatrix} I & \\ & I \\ & \partial_t \end{bmatrix} : L_p^{n_u} \times (W_p^1)^{n_y} \to L_p^{n_u} \times (W_p^1)^{n_y} \times L_p^{n_y} \quad \text{for} \quad 1 \le p \le \infty \,.$$

Thus, Theorem 2.4.3 can be applied and yields

$$\begin{aligned} J'(x) &= \mathbf{J}'(x)^* \mathbb{1} \in L_\infty^{n_u + n_w} \,, \\ c'(x)^* \lambda &= \left[T^* \mathbf{c}'(Tx)^* \quad (\mathfrak{c}^r)'(y(0), y(1))^* \right] \lambda \in L_\infty^{n_u} \times (W_\infty^{-1})^{n_y} \,, \end{aligned}$$

and

$$g'(x)^* \eta \in L_\infty^{n_u + n_y} \,.$$

Therefore, the adjoint equation component of F is contained in $\widehat{X^*}$. Furthermore,

$$\begin{aligned} c(x) &\in L_\infty^{n_c} \times \mathbb{R}^{n_r} = \Lambda \,, \\ g(x) &\in L_\infty^{m_u} \times (W_\infty^1)^{m_y} = W \,, \end{aligned}$$

and

$$\Psi(w, \eta; \mu) \in L_\infty^{m_u} \times (W_\infty^1)^{m_y} = W$$

for every fixed $\mu > 0$. Thus, $F : V \to L_\infty^{m_u} \times (W_\infty^1)^{m_y} = W$.

Since the operators J, c, g, and Ψ are twice Lipschitz-continuously differentiable by Theorem 2.4.3, the derivative

$$F_v(v, \mu) = \begin{bmatrix} L & -c'(x)^* & -g'(x)^* & \\ -c'(x) & & & \\ -g'(x) & & & I \\ & & \Psi_\eta(w, \eta; \mu) & \Psi_w(w, \eta; \mu) \end{bmatrix}$$

with

$$L = J''(x) - c''(x)^* \lambda - g''(x)^* \eta$$

is well defined and Lipschitz continuous. Because only Ψ_w and Ψ_η depend on μ, the Lipschitz constant satisfies (2.19) due to Lemma 2.5.1. □

The assumptions under which $F_v(v; \mu)$ is also invertible are very similar to the sufficient second order conditions 2.2.3. Informally, the active or binding constraints have to be linearly independent, and the Hessian of the Lagrangian must be positive definite on the nullspace of the linearized active constraints. We will derive a bound for the inverse which is later on used to establish the existence of the central path, and hence has not to be sharp. We start with a Lemma that will be used in the proof of the main theorem.

Lemma 2.5.3. *Let X and Z be Banach spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Z$, respectively. Suppose there are inner products $\langle x, x\rangle_{\bar{X}} \leq \|x\|_X^2$ and $\langle z, z\rangle_{\bar{Z}} \leq \|z\|_Z^2$ defined on X and Z, respectively, such that the closures $\bar{X} := \overline{X}^{\|\cdot\|_{\bar{X}}}$ and $\bar{Z} := \overline{Z}^{\|\cdot\|_{\bar{Z}}}$ are Hilbert spaces. Let $\hat{X} \subset \bar{X}^*$ and $\hat{Z} \subset \bar{Z}^*$ be additional Banach spaces. Assume the following conditions hold:*

1. *The linear operator $C : \bar{X} \to \bar{Z}$ satisfies the inf-sup-condition: There exists a constant $\beta > 0$ such that*
$$\inf_{\xi \in \bar{X}} \sup_{\zeta \in \bar{Z}^*} \frac{\langle \zeta, C\xi\rangle}{\|\xi\|_{\bar{X}} \|\zeta\|_{\bar{Z}^*}} \geq \beta$$
2. *The linear operator $A : \bar{X} \to \bar{X}^*$ is symmetric positive definite on* $\ker C$*: There exists a constant $\alpha > 0$ such that*
$$\langle x, Ax\rangle \geq \alpha \|x\|_{\bar{X}}^2$$
holds for all $x \in \ker C$.
3. *$A : X \to \hat{X}$ and $C : X \to Z$.*
4. *There exist splittings $X = U \oplus Y$ and $\hat{X} = \hat{U} \oplus \hat{Y}$ inducing corresponding operator splittings*
$$A = \begin{bmatrix} A_{uu} & A_{uy} \\ A_{yu} & A_{yy} \end{bmatrix} \text{ and } C = \begin{bmatrix} C_u & C_y \end{bmatrix} .$$
5. *$C_y : Y \to Z$ is an isomorphism.*
6. *$C_y^* : \hat{Z} \to \hat{Y}$ and $C_u^* : \hat{Z} \to \hat{U}$ are bounded and C_y^* is an isomorphism.*
7. *$A_{yy} : \bar{Y} \to \hat{Y}$ and $A_{yu} : \bar{U} \to \hat{Y}$ are bounded.*
8. *There is some $\gamma > 0$ such that $\|A_{uu}u + A_{uy}y\|_{\hat{U}} \geq \gamma \|u\|_U$ for all $(u, y)^T \in \ker C$.*

Then the operator
$$G := \begin{bmatrix} A & C^* \\ C & \end{bmatrix} \tag{2.20}$$
is an isomorphism $X \times \hat{Z} \to \hat{X} \times Z$. Its inverse is bounded by
$$\|G^{-1}\|_{\hat{X}\times Z \to X \times \hat{Z}} \leq \text{const}\,(1 + \|A\|_{X\to\hat{X}})(1 + \|A\|_{\bar{X}\to\bar{X}^*}^2) \cdot (1 + \|A_{yu}\|_{\bar{U}\to\hat{Y}} + \|A_{yy}\|_{\bar{Y}\to\hat{Y}}) ,$$
where the constant does not depend on any further norm of A.

Proof. Because of assumptions 3 and 6 the operator G maps $X \times \hat{Z}$ into $\hat{X} \times Z$, such that we only need to establish injectivity and the boundedness of the inverse. By conditions 1 and 2, G is an isomorphism $\bar{X} \times \bar{Z}^* \to \bar{X}^* \times \bar{Z}$, such that for every system

$$\begin{bmatrix} A_{uu} & A_{uy} & C_u^* \\ A_{yu} & A_{yy} & C_y^* \\ C_u & C_y & \end{bmatrix} \begin{bmatrix} u \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} a_u \\ a_y \\ b \end{bmatrix}$$

with $a_u \in \hat{U}$, $a_y \in \hat{Y}$ and $b \in Z$ there is a solution $u \in \bar{U}$, $y \in \bar{Y}$ and $\lambda \in \bar{Z}^*$ with

$$\|u\|_{\bar{U}} + \|y\|_{\bar{Y}} + \|\lambda\|_{\bar{Z}} \le \text{const}\,(1 + \|A\|^2_{\bar{X}\to\bar{X}^*})(\|a_u\|_{\bar{U}^*} + \|a_y\|_{\bar{Y}^*} + \|b\|_{\bar{Z}})$$

(cf. Braess [8]).

With assumptions 6 and 7,

$$\underbrace{A_{yu}u + A_{yy}y}_{\in\hat{Y}} + C_y^*\lambda = \underbrace{a_y}_{\in\hat{Y}} \Rightarrow \lambda \in \hat{Z}$$

with

$$\begin{aligned} \|\lambda\|_{\hat{Z}} &\le \text{const}\,\|a_y - A_{yu}u - A_{yy}y\|_{\hat{Y}} \\ &\le \text{const}\,(\|a_y\|_{\hat{Y}} + \|A_{yu}\|_{\bar{U}\to\hat{Y}}\|u\|_{\bar{U}} + \|A_{yy}\|_{\bar{Y}\to\hat{Y}}\|y\|_{\bar{Y}}) \\ &\le \text{const}\,(1 + \|A_{yu}\|_{\bar{U}\to\hat{Y}} + \|A_{yy}\|_{\bar{Y}\to\hat{Y}}) \\ &\quad\cdot (\|a_y\|_{\hat{Y}} + (1 + \|A\|^2_{\bar{X}\to\bar{X}^*})(\|a\|_{\bar{X}^*} + \|b\|_{\bar{Z}})) \\ &\le \text{const}\,\kappa(\|a_y\|_{\hat{Y}} + \|a\|_{\bar{X}^*} + \|b\|_{\bar{Z}})\,, \end{aligned}$$

where $\kappa := (1 + \|A\|^2_{\bar{X}\to\bar{X}^*})(1 + \|A_{yu}\|_{\bar{U}\to\hat{Y}} + \|A_{yy}\|_{\bar{Y}\to\hat{Y}})$. By condition 5 there exists $y_0 \in Y$ with $\|y_0\|_Y \le \text{const}\,\|b\|_Z$ such that

$$\begin{bmatrix} A_{uu} & A_{uy} & C_u^* \\ A_{yu} & A_{yy} & C_y^* \\ C_u & C_y & \end{bmatrix} \begin{bmatrix} u \\ y - y_0 \\ \lambda \end{bmatrix} = \begin{bmatrix} a_u - A_{uy}y_0 \\ a_y - A_{yy}y_0 \\ 0 \end{bmatrix},$$

in particular $C_u u + C_y(y - y_0) = 0$. Consequently, assumptions 5 and 3 lead to

$$\frac{\|y - y_0\|_Y}{\|C_y^{-1}\|_{Z\to Y}\|C_u\|_{U\to Z}} \le \|u\|_U\,.$$

Let $x_0 = (0, y_0)$. Then

$$\begin{aligned} \|x - x_0\|_X &\le \text{const}\,\|u\|_U \\ &\le \text{const}\,\|A_{uu}u + A_{uy}(y - y_0)\|_{\hat{U}} \\ &\le \text{const}\,(\|a\|_{\hat{X}} + \|A\|_{X\to\hat{X}}\|y_0\|_Y + \|\lambda\|_{\hat{Z}}) \\ &\le \text{const}\,(1 + \|A\|_{X\to\hat{X}})\kappa(\|a\|_{\hat{X}} + \|b\|_Z) \end{aligned}$$

because of assumptions 6 and 8. □

Theorem 2.5.4. *Suppose the assumptions of Theorem 2.5.2 are satisfied. Assume there is a bounded set $D \subset V$ and a constant $0 < \mu_0 \leq 1/2$ such that the following conditions hold uniformly for $v \in D$ and $\mu \leq \mu_0$.*

1. *The linearized state equation is solvable: There exists a constant $\beta > 0$ such that for every constraints variation $\delta \in \Lambda_p$ there is a state variation $\delta y \in X_p^y$ with*
$$c_y(x)\delta y = \delta \quad \text{and} \quad \|\delta y\|_{X_p^y} \leq \beta \|\delta\|_{\Lambda_p}$$
for $p = 1, 2, \infty$.
2. *the state equation $\mathfrak{c}(u(t), y(t), \dot{y}(t)) = 0$ is linear in $\dot{y}(t)$.*
3. *The strengthened Legendre-Clebsch condition holds: There exists a constant $\gamma > 0$ such that*
$$\xi(t)^T H(t)\xi(t) \geq \gamma |\xi(t)|^2$$
for almost all $t \in \Omega$ and $x \in \ker c'$. Here,
$$\begin{aligned} H(t) := \mathfrak{J}''(x(t)) - \mathfrak{c}''(u(t), y(t), \dot{y}(t))^T \lambda(t) - \mathfrak{g}''(x(t))^T \eta(t) \\ + \mathfrak{g}'(x(t))^T \psi_\eta(w(t), \eta(t); \mu)^{-1} \psi_w(w(t), \eta(t); \mu) \mathfrak{g}'(u(t)) \,. \end{aligned}$$

Then $F_v(v; \mu)$ has an inverse which is bounded by
$$\|F_v(v; \mu)^{-1}\|_{Z \to V} \leq \text{const}\, \mu^{-12}$$
uniformly for $v \in D$.

Proof. We show that there is a unique solution of $F_v(v; \mu)\Delta v = z$ with $\|v\|_{V_\infty} \leq \|z\|_{Z_\infty}$. The proof is performed in three steps. First the system is reduced to a smaller one that meets the requirements of Lemma 2.5.3. In the second step, the assumptions are checked and the Lemma is applied. Finally, the reduction of the first step is reversed.

Step 1. For readability reasons we will omit the arguments x, λ, η, w, and μ of the operators L_{xx}, c', g', Ψ_η, and Ψ_w in the following. For any system
$$F_v(v; \mu)\Delta v = \begin{bmatrix} L_{xx} & -(c')^* & -(g')^* & \\ -c' & & & \\ -g' & & & I \\ & & \Psi_\eta & \Psi_w \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\eta \\ \Delta w \end{bmatrix} = \begin{bmatrix} z_a \\ z_c \\ z_s \\ z_p \end{bmatrix},$$
elimination of the slack variables $\Delta w = z_s + g'\Delta x$ yields the equivalent system
$$\begin{bmatrix} L_{xx} & -(c')^* & -(g')^* \\ -c' & & \\ \Psi_w g' & & \Psi_\eta \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\eta \end{bmatrix} = \begin{bmatrix} z_a \\ z_c \\ z_p - \Psi_w z_s \end{bmatrix}. \tag{2.21}$$

Space in	
Theorem 2.5.4	Lemma 2.5.3
X	X
Λ	Z
X_2	$\bar{X}$
Λ_2	$\bar{Z}$
$\widehat{X^*}$	$\hat{X}$
Λ	$\hat{Z}$

Table 2.1: Matching of spaces in Theorem 2.5.4 and Lemma 2.5.3.

Due to Lemma 2.5.1, Ψ_η is invertible, such that by elimination of $\Delta\eta = \Psi_\eta^{-1}(z_p - \Psi_w(z_s + g'\Delta x))$ (2.21) can be written as

$$\begin{bmatrix} H & -(c')^* \\ -c' & \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} z_a - (g')^* \Psi_\eta^{-1}(z_p - \Psi_w z_s) \\ z_c \end{bmatrix} \tag{2.22}$$

where $H := L_{xx}(x, \lambda, \eta) + g'(x)^* \Psi_\eta(w, \eta; \mu)^{-1} \Psi_w(w, \eta; \mu) g'(x)$.

Step 2. Under the assumptions of the theorem, Lemma 2.5.3 can be applied to the system (2.22). To begin with, the involved spaces can be identified as shown in Table 2.1. Then the conditions of Lemma 2.5.3 can be verified: $-c'$ satisfies the inf-sup-condition and H is symmetric positive definite on the nullspace of c' due to assumptions 1 and 3. Condition 3 is satisfied due to Theorem 2.5.2 and the Nemyckii structure of H. The space splitting is given by $X = X^u \times X^y$. Condition 5 is satisfied because of assumption 1 with $p = \infty$. Condition 6 is satisfied due to assumption 1 with $p = 1$, Theorem 2.4.3, and the Nemyckii structure of c_u. Condition 7 is satisfied because of assumption 2 and the resulting Nemyckii structure of H. Condition 8 is satisfied due to assumptions 2 and 3.

Therefore, Lemma 2.5.3 is applicable and there exists a unique solution $\Delta x \in X$ and $\Delta\lambda \in \Lambda$. As for the dependency on μ, any norm of H is affected only via the term $\Psi_w(w, \eta; \mu)^{-1} \Psi_\eta(w, \eta; \mu)$. Since control and state constraints are decoupled, the mixed term H_{yu} is independent of μ. By Lemma 2.5.1 and the boundedness of D we have

$$\|(\Psi_\eta^y)^{-1} \Psi_w^y\|_{W_2^y \to \widehat{(X_2^y)^*}} \le \|(\Psi_\eta^y)^{-1} \Psi_w^y\|_{W_2^y \to W_2^y} \le \text{const}\, \mu^{-2},$$
$$\|\Psi_\eta^{-1} \Psi_w\|_{W \to \widehat{X^*}} \le \|\Psi_\eta^{-1} \Psi_w\|_{W \to W} \le \text{const}\, \mu^{-2}, \quad \text{and}$$
$$\|\Psi_\eta^{-1} \Psi_w\|_{W_2 \to X_2^*} \le \|\Psi_\eta^{-1} \Psi_w\|_{W_2 \to W_2} \le \text{const}\, \mu^{-2}$$

such that

$$\|H\|_{X \to \widehat{X}^*} \leq \text{const}\,\mu^{-2}$$
$$\|H\|_{X_2 \to X_2^*} \leq \text{const}\,\mu^{-2}$$
$$\|H_{yu}\|_{X_{u,2} \to \widehat{X}_y^*} + \|H_y y\|_{X_{y,2} \to \widehat{X}_y^*} \leq \text{const}\,\mu^{-2}\,,$$

and the inverse of the saddle point operator from (2.22) is bounded by const μ^{-8}. Thus, by Lemmas 2.5.3 and 2.5.1,

$$\begin{aligned}\|\Delta x\|_X + \|\Delta\lambda\|_\Lambda &\leq \text{const}\,\mu^{-8}(\|z_a - (g')^*\Psi_\eta^{-1}(z_p - \Psi_w z_s)\|_{\widehat{X}^*} + \|z_c\|_\Lambda) \\ &\leq \text{const}\,\mu^{-10}(\|z_a\|_{\widehat{X}^*} + \|z_p\|_W + \|z_s\|_W + \|z_c\|_\Lambda)\,.\end{aligned}$$

Note that the constant is independent of μ.

Step 3. Tracing back the elimination chain from Step 1 yields

$$\begin{aligned}\|\Delta\eta\|_W &= \|\Psi_\eta^{-1}(z_p - \Psi_w(z_s + g'\Delta x))\|_W \\ &\leq \text{const}\,\mu^{-2}(\|z_p\|_W + \|z_s\|_W + \|\Delta x\|_X \\ &\leq \text{const}\,\mu^{-12}(\|z_p\|_W + \|z_s\|_W + \|z_a\|_{\widehat{X}^*} + \|z_c\|_\Lambda)\end{aligned}$$

and

$$\begin{aligned}\|\Delta w\|_W &= \|z_s + g'\Delta x\|_W \\ &\leq \text{const}\,\mu^{-10}(\|z_p\|_W + \|z_s\|_W + \|z_a\|_{\widehat{X}^*} + \|z_c\|_\Lambda)\,.\end{aligned}$$ □

Theorem 2.5.5. *Suppose the assumptions of Theorem 2.5.4 are satisfied. Let $R > 0$ and $M := \{v \in D : B(v, R) \subset D\}$. If there are $v_0 \in M$ and $\mu_0 > 0$ with $F(v_0; \mu_0) = 0$, then there exists a central path $v(\mu) \in M$. Either the central path leaves M for some $\mu_{\mathit{final}} > 0$ or it is defined on the whole interval $]0, \mu_0]$.*

Proof. By Theorem 2.5.4, $F_v(v;\mu)^{-1}$ is bounded uniformly on $D \times [\epsilon, \mu_0]$ for every $\epsilon > 0$. Furthermore, both F and F_v are uniformly continuous on $D \times [\epsilon, \mu]$ thru Theorem 2.5.2. Thus, by the implicit function theorem, the central path exists and can be followed to $v(\epsilon)$ unless it leaves M (cf. [54], Theorem 4.B and Proposition 6.10). □

2.6 Convergence of the Central Path

As pointed out in Section 2.3 and in the appendix, convergence of the central path in V_∞ cannot be expected. This is particularly the case in the presence of state constraints, where the Lagrange multipliers are often measure valued. Therefore, the central path will not be bounded, not even in V_2. In this respect, Theorem 2.5.5 is as good a result as one might expect. To prove convergence of the central path at least in V_2, we have to resort to pure control constraints.

One obstacle that has to be overcome is that $\Psi_\eta^{-1}\Psi_w$ is in general not uniformly bounded for $\mu \to 0$. Thus, the elimination of $\Delta\eta$ in the proof of Theorem 2.5.4 leads to the undesirable system (2.22), from which a bound of F_v^{-1} is difficult to obtain.

The key observation is that the usual splitting of the inequality constraints into active (or binding) and inactive constraints can be utilized to perform a more careful elimination step. In particular, active constraints can be substituted equivalently by corresponding equality constraints, whereas inactive constraints can simply be dropped. Since in general the partitioning into active and inactive constraints is not known a-priori, we need to define an approximate splitting. In this way the system modifications are approached asymptotically by elimination of $\Delta\eta$ only for the approximately inactive constraints.

Definition 2.6.1. For $\rho > 0$, the *ρ-nearly active set* for component i of the inequality constraints g is defined by

$$\Omega^\rho_{A,i}(v) := \{t \in \Omega : w(t) \le \rho\eta(t)\}\,.$$

The complete nearly active set is $\Omega^\rho_A(v) := \{\Omega^\rho_{A,i} : i = 1, \dots, m_u + m_y\}$. Analogously, the nearly inactive set is $\Omega^\rho_I(v) := \Omega \backslash \Omega^\rho_A$.

Corresponding to the component wise splitting of the domain Ω into active and inactive regions there is a function space splitting of W into active and inactive constraints and multipliers:

$$\begin{aligned} W^\mu_A &:= \{w|_{\Omega^\rho_A} : w \in W\} \\ W^\mu_I &:= \{w|_{\Omega^\rho_I} : w \in W\} \end{aligned} \tag{2.23}$$

The space splitting induces variable splittings $w \mapsto (w_A, w_I)$ and $\eta \mapsto (\eta_A, \eta_I)$ and operator splittings $g \mapsto (g_A, g_I)$ and $\Psi \mapsto (\Psi^A, \Psi^I)$.

Lemma 2.6.2. *The splitting* (2.23) *leads to the diagonal operator splittings*

$$\Psi_w(w,\eta;\mu) = \begin{bmatrix} \Psi^A_w & \\ & \Psi^I_w \end{bmatrix} \quad \text{and} \quad \Psi_\eta(w,\eta;\mu) = \begin{bmatrix} \Psi^A_\eta & \\ & \Psi^I_\eta \end{bmatrix},$$

where $\|(\Psi^A_w)^{-1}\|_{W_{2A}\to W_{2A}}$ and $\|(\Psi^I_\eta)^{-1}\|_{W_{2A}\to W_{2A}}$ are bounded independently of μ.

Proof. Due to the Nemyckii structure of Ψ, the operator splitting is diagonal by Theorem 2.4.5.

For the nearly active set we infer

$$\begin{aligned} \psi^A_w(w,\eta;\mu)^{-1} &= \left(1 - \frac{w}{\sqrt{w^2+\eta^2+2\mu}}\right)^{-1} \\ &= \frac{\sqrt{w^2+\eta^2+2\mu}}{\sqrt{w^2+\eta^2+2\mu}-w} \le \frac{\sqrt{2(w^2+\eta^2)}}{\sqrt{w^2+\eta^2}-w}\,. \end{aligned}$$

From $sw = \eta$ with $s \geq \rho^{-1}$ we conclude

$$\psi_w^A(w,\eta;\mu)^{-1} \leq \frac{\sqrt{2(1+s^2)w^2}}{\sqrt{(1+s^2)w^2} - w} = \frac{\sqrt{2(1+s^2)}}{\sqrt{1+s^2} - 1} \leq \frac{\sqrt{2(1+\rho^{-2})}}{\sqrt{1+\rho^{-2}} - 1} .$$

By Theorem 2.4.3, $\|(\Psi_w^A)^{-1}\|_{W_{2A}\to W_{2A}} \leq \text{const}$, where the constant depends only on ρ. Analogously, the corresponding result is obtained for $(\Psi_\eta^I)^{-1}$. □

For the proof of the main theorem we will use the following Lemma by Braess and Blömer [9][1] on saddle point operators with penalty term.

Lemma 2.6.3. *Let X and Z be Hilbert spaces. Assume the following conditions hold:*

1. *The continuous linear operator $C : X \to Z$ satisfies the inf-sup-condition: There exists a constant $\beta > 0$ such that*

$$\inf_{x\in X} \sup_{\zeta\in Z} \frac{\langle \zeta, Cx\rangle}{\|x\|_X \|\zeta\|_Z} \geq \beta .$$

2. *The continuous linear operator $A : X \to X^*$ is symmetric positive definite on the nullspace of C and positive semidefinite on the whole space X: There exists a constant $\alpha > 0$ such that*

$$\langle x, Ax\rangle \geq \alpha \|x\|_X^2 \quad \textit{for all } x \in \ker C$$

and

$$\langle x, Ax\rangle \geq 0 \quad \textit{for all } x \in X .$$

3. *The continuous linear operator $D : Z^* \to Z$ is symmetric positive semidefinite.*

Then, the operator

$$\begin{bmatrix} A & C^* \\ C & -D \end{bmatrix}$$

is invertible. The inverse is bounded by a constant depending only on α, β, and the norms of A, C, and D.

Lemma 2.6.4. *Let X and Y be metric spaces and $f : X \to Y$ uniformly continuous. If x_n is a Cauchy sequence, then so is $f(x_n)$.*

[1]Note that there is a misprint in Lemma B.1 in the article. The assumption that A is positive semidefinite on the whole space is used in the proof without being stated.

Proof. Let $\epsilon > 0$ be arbitrary. Then there is a $\delta > 0$ such that $d(f(x), f(y)) \le \epsilon$ whenever $d(x,y) \le \delta$. There is $N \in \mathbb{N}$ such that $d(x_n, x_m) \le \delta$ for all $n, m \ge N$, thus $d(f(x_n), f(x_m)) \le \epsilon$. Therefore, $f(x_n)$ form a Cauchy sequence. □

Theorem 2.6.5. *Let $m_y = 0$. Suppose that the following conditions are satisfied uniformly for all $v = (x, \lambda, \eta, w)$ on the central path.*

1. *For the nearly active constraints $C := (c'(x), g'_A(x))^T$ the inf-sup-condition*
$$\inf_{\chi \in \Lambda_2, \zeta \in W^*_{2A}} \sup_{\xi \in X_2} \frac{\langle C\xi, (\chi, \zeta)^T \rangle}{\|\xi\|_{X_2}(\|\chi\|_{\Lambda_2} + \|\zeta\|_{W^*_{2A}})} \ge \beta$$
holds for some $\beta > 0$.

2. *The modified Hessian of the Lagrangian is positive definite on the nullspace of the nearly active constraints and positive semidefinite on the whole space: There exists a constant $\alpha > 0$ such that*
$$\langle Hx, x \rangle \ge \alpha \|x\|^2_{X_2} \quad \text{for all } x \in \ker C$$
$$\langle Hx, x \rangle \ge 0 \quad \text{for all} \quad x \in X_2$$
where $H := L_{xx}(x, \lambda, \eta) + g'_I(x)^* \Psi^I_\eta(w, \eta; \mu)^{-1} \Psi^I_w(w, \eta; \mu) g'_I(x)$.

If the central path $v(\mu)$ is defined on the whole interval $(0, \mu]$, it converges to a limit point $v(0)$ in V_2 for $\mu \to 0$. If, in addition, the control u enters linearly in J, c, and g, then the limit point $v(0)$ satisfies the first order necessary Kuhn-Tucker conditions.

Proof. As in the proof of Theorem 2.5.4 we first eliminate Δw from the system $F_v(v; \mu)\Delta v = z$ to get
$$\begin{bmatrix} L_{xx} & -(c')^* & -(g')^* \\ -c' & & \\ \Psi_w g' & & \Psi_\eta \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \eta \end{bmatrix} = \begin{bmatrix} z_a \\ z_c \\ z_p - \Psi_w z_s \end{bmatrix}. \tag{2.24}$$
Using Lemma 2.6.2, we can write
$$\begin{bmatrix} L_{xx} & -(c')^* & -(g'_A)^* & -(g'_I)^* \\ -c' & & & \\ \Psi^A_w g'_A & & \Psi^A_\eta & \\ \Psi^I_w g'_I & & & \Psi^I_\eta \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \eta_A \\ \Delta \eta_I \end{bmatrix} = \begin{bmatrix} z_a \\ z_c \\ z_{pA} - \Psi^A_w z_{sA} \\ z_{pI} - \Psi^I_w z_{sI} \end{bmatrix}.$$
Since both Ψ^I_η and Ψ^A_w are invertible, we can eliminate $\Delta\eta_I = (\Psi^I_\eta)^{-1}(z_{pI} - \Psi^I_w(z_{sI} + g'_I \Delta x)$ and multiply the third row by $-(\Psi^A_w)^{-1}$, which yields
$$\begin{bmatrix} H & -(c')^* & -(g'_A)^* \\ -c' & & \\ -g'_A & & -(\Psi^A_w)^{-1}\Psi^A_\eta \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \eta_A \end{bmatrix}$$
$$= \begin{bmatrix} z_a + (g'_I)^*(\Psi^I_\eta)^{-1}(z_{pI} - \Psi^I_w z_{sI}) \\ z_c \\ -(\Psi^A_w)^{-1} z_{pA} + z_{sA} \,. \end{bmatrix} \tag{2.25}$$

In a second step we will show that the inverse of the operator in (2.25) is bounded independently of μ. First we note that $(\Psi_w^A)^{-1}$, Ψ_η^A, $(\Psi_\eta^I)^{-1}$, and Ψ_w^I are all bounded independently of μ by Lemmas 2.5.1 and 2.6.2. Then, using Assumptions 1 and 2, we can apply Lemma 2.6.3 which guarantees the existence of a unique solution $\Delta x \in X_2$, $\Delta\lambda \in \Lambda_2$, and $\Delta\eta_A \in W_{2A}$ with

$$\begin{aligned}\|\Delta x\|_{X_2} &\le \text{const}\,\|z_a + (g_I')^*(\Psi_\eta^I)^{-1}(z_{pI} - \Psi_w^I z_{sI})\|_{X_2^*}\\ &\le \text{const}\,(\|z_a\|_{X_2^*} + \|(g_I')^*\|\|(\Psi_\eta^I)^{-1}\|(\|z_{pI}\|_{W_2} + \|\Psi_w^I\|\|z_{sI}\|_{W_2}))\\ &\le \text{const}\,(\|z_a\|_{X_2^*} + \text{const}\,(\|z_{pI}\|_{W_2} + \|z_{sI}\|_{W_2}))\\ &\le \text{const}\,\|z\|_{Z_2}\end{aligned}$$

and

$$\begin{aligned}\|\Delta\lambda\|_{\Lambda_2} + \|\Delta\eta_A\|_{W_2} &\le \text{const}\,(\|z_c\|_{\Lambda_2^*} + \|(\Psi_w^A)^{-1}z_{pA} - z_{sA}\|_{W_{2A}})\\ &\le \text{const}\,(\|z_c\|_{\Lambda_2^*} + \|(\Psi_w^A)^{-1}\|\|z_{pA}\|_{W_2} + \|z_{sI}\|_{W_2})\\ &\le \text{const}\,(\|z_c\|_{\Lambda_2^*} + \|z_{pA}\|_{W_2} + \|z_{sI}\|_{W_2})\\ &\le \text{const}\,\|z\|_{Z_2}\,,\end{aligned}$$

where the constants are independent of μ. Tracing the eliminations of Step 1 back yields

$$\begin{aligned}\|\Delta\eta_I\|_{W_2} &= \|(\Psi_\eta^I)^{-1}(z_{pI} - \Psi_w^I(z_{sI} + g_I'\Delta x))\|_{W_2}\\ &\le \|(\Psi_\eta^I)^{-1}\|(\|z_{pI}\|_{W_2} + \|\Psi_w^I\|\|z_{sI} + g_I'\Delta x)\|_{W_2})\\ &\le \text{const}\,(\|z_{pI}\|_{W_2} + \|z_{sI} + g_I'\Delta x)\|_{W_2})\\ &\le \text{const}\,\|z\|_{Z_2}\end{aligned}$$

and

$$\|\Delta w\|_{W_2} = \|z_s + g'\Delta x\|_{W_2} \le \|z_s\|_{W_2} + \text{const}\,\|\Delta_x\|_{X_2} \le \text{const}\,\|z\|_{Z_2}\,.$$

Hence, $F_v(v(\mu);\mu)^{-1}$ is bounded independently of μ.

Finally, from

$$\psi_\mu(w,\eta;\mu) = \frac{1}{\sqrt{w^2 + \eta^2 + 2\mu}} \le \mu^{-\frac{1}{2}}$$

and

$$F_\mu(v(\mu)) = \begin{bmatrix}0 & 0 & 0 & -\Psi_\mu\end{bmatrix}^T$$

we infer for the derivative of the central path $v(\mu)$

$$\begin{aligned}\|v'(\mu)\|_{V_2} &= \|F_v(v(\mu);\mu)^{-1}F_\mu(v(\mu))\|_{V_2}\\ &\le \|F_v(v(\mu);\mu)^{-1}\|_{Z_2\to V_2}\|F_\mu(v(\mu))\|_{Z_2} \le \text{const}\,\mu^{-\frac{1}{2}}\,.\end{aligned}$$

Therefore, the total length of the path,

$$\int_0^{\mu_0} \|v'(\mu)\|_{V_2} \, d\mu \,,$$

is bounded and the path itself is uniformly continuous and thus converges to a limit point $v(0)$ as $\mu \to 0$ by Lemma 2.6.4.

If the control enters linearly, both the path $v(\mu)$ and F are continuous. Therefore we have $F(v(0), 0) = 0$, such that $v(0)$ satisfies the first order optimality conditions (2.2). □

Remark 2.6.6. In general, Assumption 1 imposes an upper bound on the choice of ρ due to the monotonicity

$$\rho_1 \le \rho_2 \Rightarrow W_A^\mu(\rho_1) \subset W_A^\mu(\rho_2) \,.$$

If W_A^μ gets too large, C may become non-injective and thus no longer satisfy the inf-sup-condition. ◁

Remark 2.6.7. Assuming coercivity of $L_{xx}(v)$ on the nullspace of C is a too strong requirement. Since ρ should be small, the estimated active set A tends to be too small. Therefore the nullspace of C tends to be too large, leaving room for $L_{xx}(v)$ being indefinite on the gap even for well defined problems for which the sufficient second order condition 2.3 holds. ◁