# Optimization in Function Space

Lecture notes

HANNES MEINLSCHMIDT

Winter term 2020/21
Preliminary version
Last update: January 18, 2021

# Preface

These lecture notes cover topics suitable for an introductory "5 credit point" (2+1 weekly hours) course about the theory of optimization in function spaces. Students are assumed to be familiar with general principles of functional analysis and nonlinear optimization.

The goal of the course is to give an abstract treatment of optimization problems posed in infinite-dimensional spaces, which is—as far as time allows it—self-contained in its presentation, giving students a theoretical background for further studies in the field of infinite-dimensional optimization, in particular in connection with optimal control or generally PDE-constrained problems. After completion of this course, students should be able to identify the underlying machinery for practical problems and follow their optimization setup or execute it themselves.

We refer to the excellent textbooks of Brezis [Br10] and Nocedal and Wright [NW06] as well as Ulbrich and Ulbrich [UU12] for references to functional analysis and basic PDE theory and theory of nonlinear optimization, respectively. Topics of this course are (partially) covered in, for example, the books of Bonnans and Shapiro [BS00], Schirotzek [Sc07] or Hinze, Pinnau, Ulbrich and Ulbrich [HPUU09].

The notes are based on lecture notes of Michael Ulbrich (TU München), Stefan Ulbrich (TU Darmstadt), and Christian Meyer (TU Dortmund). Comments, suggestions, and notification of errors are welcome by eMail at `hannes.meinlschmidt@ricam.oeaw.ac.at`.

# 1 Motivation and problem setting

## 1.1 Motivation

Modeling, optimization and numerical simulation of complex systems plays an important role in physics, engineering, mechanics, chemistry, biology, medicine, finance, and in many other disciplines. These models often use ordinary or partial differential equations to describe the system, which means that the resulting optimization problems involve *generically infinite-dimensional* objects such as functions, or in general quantities that live in Banach spaces. (Thus, most often, *function spaces* are involved.) There is a rising interest to solve optimization problems involving such models and there are many fascinating applications. We give a (totally incomplete) selection of examples:

- Optimal space mission design,
- optimal control of robot movements,
- identification of model parameters by means of measurements [$\rightarrow$ *inverse problems*], e.g.:
  - geological material properties from seismic measurements,
  - data assimilation: initial conditions for weather and climate models from scattered observations of all kinds (weather stations, satellite data, etc.),
  - tomography
  - derivation of parameters (e.g., volatility) in models for option pricing (*Black Scholes* PDE) from market prices ($\rightarrow$ mathematical fincance),
- optimal control of actuators, built on the surface of an aircraft, to avoid noise generation, material fatigue, or vortex generation,
- optimal radiation therapy planning,
- optimal design of the body of a ship,
- optimization of the shape of a wing, a turbine blade, etc.,
- optimal control of laser hardening of steel,
- optimal control of crystal pulling (heating, pulling speed).

The aim of this course is to develop a rigorous theory and establish methods for infinite dimensional optimization problems. This provides the appropriate framework for investigating and solving the above problem classes.

An exemplary, PDE-constrained problem could look like this:

**Example.** The following is an *optimal boundary control problem* for a *semilinear elliptic* equation:

$$\min_{y,u} \quad \frac{1}{2}\int_\Omega \big(y(x) - y_d(x)\big)^2 \,\mathrm{d}x + \frac{\alpha}{2}\int_{\partial\Omega} u(x)^2 \,\mathrm{d}\mathcal{H}_{n-1}(x)$$

$$\text{s.t.} \quad \begin{cases} -\Delta y + y^3 = 0 & \text{on } \Omega, \\ \dfrac{\partial y}{\partial \nu} + y = u & \text{on } \partial\Omega, \\ a \le u \le b & \text{on } \partial\Omega, \end{cases}$$

possibly arising from (stationary) nonlinear heat transfer. The goal is to find an ambient temperature $u$ such that the temperature profile $y$ in $\Omega \subset \mathbb{R}^n$ (e.g., a work piece) is as close as possible to a given desired temperature profile $y_d$. Thereby, the ambient temperature must stay between bound functions $a$ and $b$ which represent e.g. physical or process-related limitations, and a trade-off between optimization accuracy and heating cost is possible by the parameter $\alpha$.

We will come back to this example later. The fundamental (theoretical) questions which we will consider in this lecture with respect to infinite-dimensional optimization problems are as follows:

- Is the problem well posed to begin with, that is: does the problem admit an optimal solution at all?
- If so, can we characterize optimal solutions? This means, under which condition are there *necessary* and *sufficient* optimality conditions of e.g. first and second order?

The latter optimality conditions are also a very useful starting point for numerical algorithms to find an optimal solution.

## 1.2 Problem setting

The fundamental general problem class is given by

$$\min_{x \in X} \ f(x) \quad \text{s.t.} \quad G(x) \in K, \tag{P}$$

with the *feasible set* (`zulässige Menge`)

$$\mathcal{F} := \Big\{ x \in X \colon G(x) \in K \Big\} := G^{-1}\big[K\big].$$

We pose the following fundamental assumptions on the data in (P):

**Assumption 1.1** (Problem data)**.** The following properties of the data in (P) are valid:

(a) $G: X \to Z$ is a continuous mapping between real Banach spaces,

(b) $K \subseteq Z$ is a closed convex set,

(c) $f: X \to \mathbb{R}$ is a lower semicontinuous objective function,

(d) the feasible set $\mathcal{F}$ is nonempty.

Note that $\mathcal{F}$ is closed in $X$ under Assumption 1.1 since $G$ is continuous and $K$ is closed. Let us moreover briefly recall the notion of lower semicontinuity.

**Reminder:** We say that the function $f: X \to \mathbb{R}$ is *lower semicontinuous* (`unterhalbstetig`) if for all $x \in X$ and all sequences $(x_k) \subset X$ there holds

$$x_k \to x \quad \Longrightarrow \quad \liminf_{k\to\infty} f(x_k) \geq f(x), \tag{1.1}$$

or, equivalently, if the level set $N_f(\alpha) := \{x \in X : f(x) \leq \alpha\}$ (`Niveaumenge`) is sequentially closed for all $\alpha \in \mathbb{R}$.

Lower semicontinuity can also be defined for functions $f: S \to \mathbb{R}$ defined on a proper subset $S \subsetneq X$. One then considers sequences $S \ni x_k \to x \in S$. The equivalent definition via level sets however requires that $S$ is closed.

The compact notation (P) includes essentially all relevant situations which occur in particular problems. The two variations most often encountered are the following:

1. The *total problem*

$$\min_{x \in X} f(x) \quad \text{s.t.} \quad E(x) = 0, \quad F(x) \in C, \quad x \in M,$$

where the continuous mappings $E: X \to Z_1$ and $F: X \to Z_2$ take their values in Banach spaces $Z_1$ and $Z_2$, the set $C \subseteq Z_2$ is a closed convex cone, and $M \subseteq X$ is a closed convex set. This is a problem of type (P) for $Z = Z_1 \times Z_2 \times X$ and $K = \{0_{Z_1}\} \times C \times M$ with $G(x) := (E(x), F(x), x)$.

**Reminder:** Recall that a set $C$ is called *cone* if from $x \in C$ it follows that $\lambda x \in C$ for all $\lambda > 0$.

2. Often, the variable $x \in X = Y \times U$ has two components $x = (y, u)$, with $u$ being a *control* (`Steuerung`), design or parameter within the system, and $y$ being the associated *state* (`Zustand`) of the system. They are linked by the system dynamics described abstractly by $E(y, u) = 0$. Problems of this particular structure are called *optimal control* problems. In this context, one often encounters problems in *reduced form* (as opposed to the *total* problem above), that is,

$$\min_{u \in U} j(u) \quad \text{s.t.} \quad H(u) \in C, \quad u \in U_{\text{ad}}.$$

These are obtained from the total optimal control problem by eliminating the state equation $E(y, u) = 0$. This is done by showing that for each $u$ there exists a unique state $y = y(u)$ such that $E(y(u), u) = 0$, and inserting the so-called *control-to-state operator* (`Steuerungs-Zustands-Operator`) $y(u)$ for $y$, that is,

$$j(u) := f(y(u), u), \quad H(u) := F(y(u), u), \quad U_{\text{ad}} := M_u \cap y^{-1}[M_y]$$

where $M = M_y \times M_u \subseteq Y \times U$. Of course, the state equation $E(y(u), u) = 0$ is then superfluous because it is always satisfied by construction. Such a reduced problem is of formally easier nature, but with (much) more involved functions $j, H$ due to the control-to-state operator $u \mapsto y(u)$.

**Remark 1.2.** The conditions $F(x) \in C$ or $H(u) \in C$ with $C$ denoting a closed convex cone can be interpreted to describe *abstract inequality constraints*. If for instance $Z_2 = \mathbb{R}^m$ and $C = (-\infty, 0]^m$, then the constraint $F(x) \in C$ is the same as $F(x) \leq 0$ (component-wise), which corresponds to the standard inequality constraints in nonlinear programming. Note that there is an order structure on $C$ in this case. This additional structure is quite helpful in several situations. We will come back to this later.

**Remark 1.3.** (a) As in Assumption 1.1, we will *always* assume all occurring Banach spaces to be real in the following, even if it is not mentioned explicitly.

(b) Over the course of this lecture, we will also consider first order necessary optimality conditions for problems of type (P). Differentiability assumptions on $f$ and $G$ will then be stated as required.

## 2 Existence of Solutions

We first discuss the question of existence of solutions to the problem (P). The easiest case is that of a compact feasible set.

**Theorem 2.1.** *Consider* (P) *and let Assumption 1.1 be satisfied. Assume further that there exist a sequentially precompact set* $C \subseteq \mathcal{F}$ *and* $x_0 \in C$ *such that* $f(x) \geq f(x_0)$ *for all* $x \in \mathcal{F} \setminus C$. *Then* (P) *possesses a global solution.*

*Proof.* By assumption, no point in $\mathcal{F} \setminus C$ provides a lower objective function value than $x_0 \in C$. Hence, we can restrict our search for the minimum on the set $C$, which is sequentially precompact and nonempty. Now, consider a minimizing sequence $(x_k) \subset C$ satisfying $f(x_k) \to f^* := \inf_{x \in \mathcal{F}} f(x)$. Due to sequential precompactness of $C$, there exists a subsequence, again denoted by $(x_k)$, that converges to a limit $\bar{x}$. The feasible set $\mathcal{F}$ is closed, hence there holds $\bar{x} \in \mathcal{F}$. Now, since $f$ is lower semicontinuous, we have

$$f^* = \lim_{k \to \infty} f(x_k) = \liminf_{k \to \infty} f(x_k) \geq f(\bar{x})$$

and thus $\bar{x} \in \mathcal{F}$ is a solution of (P). □

**Remark 2.2.** Note that if $\mathcal{F}$ is already compact and nonempty itself, then the choice $C = \mathcal{F}$ is perfectly valid in the foregoing theorem.

In infinite-dimensional spaces, compactness is a strong requirement, often: *too* strong. Recall that the closed unit ball $\overline{B_X(0)}$ in a Banach space $X$ is compact *if and only if* $X$ is finite-dimensional! On the other hand, *weak* sequential compactness is often verifiable.

**Reminder:** We recall that $(x_k)$ *converges weakly* to $x$ in $X$, written $x_k \rightharpoonup x$, iff

$$\langle x', x_k \rangle \to \langle x', x \rangle \quad \text{for all } x' \in X^*.$$

A set $M$ is weakly sequentially compact (`schwach folgenkompakt`) if every sequence $(x_k) \subset M$ admits a weakly convergent subsequence whose limit is again in $M$, and a function $f\colon X \to \mathbb{R}$ is weakly lower semicontinuous if $x_k \rightharpoonup x$ implies $\liminf_{k \to \infty} f(x_k) \geq f(x)$.

We now want to replace the compactness requirement in Theorem 2.1 by weak sequential compactness. This works particularly well in *reflexive* Banach spaces, since these are precisely the ones for which the closed unit ball $\overline{B_X(0)}$ is weakly sequentially compact.

**Reminder:** A Banach space $X$ is said to be *reflexive* if the canonical injection

$$J\colon X \to (X^*)^* := X^{**}, \quad \langle Jx, x' \rangle := \langle x', x \rangle \quad \text{for all } x' \in X^*$$

is surjective (and thus a bijection).

The most commonly encountered reflexive Banach space class are Hilbert spaces. Also important, but, for $p \neq 2$ non-Hilbert, are the Lebesgue spaces $L^p$ for $1 < p < \infty$—$L^1$ and $L^\infty$ are *not* reflexive on nontrivial measure spaces—and the associated Sobolev spaces $W^{k,p}$, where $k \in \mathbb{N}$. The fact that these spaces are reflexive plays a huge role in their prevalence in functional analytic methods for PDEs.

We need the following facts from functional analysis:

**Proposition 2.3** (See [Br10, Ch. 3]). *Let $X$ be a Banach space.*

*(a) Every closed convex subset of $X$ is weakly sequentially closed.*

*(b) The space $X$ is reflexive if and only if every bounded sequence contains a weakly convergent subsequence.*

*(c) If $f\colon X \to \mathbb{R}$ is convex and lower semicontinuous, then $f$ is also weakly lower semicontinuous.*

*(d) Any weakly convergent sequence in $X$ is bounded.*

The second statement in Proposition 2.3 in particular implies that the closed unit ball $\overline{B_X(0)}$ in a reflexive Banach space $X$ is weakly sequentially compact.

Note moreover that the third assertion is derived from the first since every level set $N_f(\alpha)$ of $f$ is convex (convexity of $f$) and closed (lower semicontinuity) and thus also weakly closed, hence weakly sequentially closed. Therefore, $f$ is weakly lower semicontinuous.

We obtain the following existence result:

**Theorem 2.4.** *Consider (P) and let the following assumptions be satisfied:*

*(a) $G\colon X \to Z$ is weakly sequentially continuous from the reflexive Banach space $X$ to the Banach space $Z$, i.e., we have*

$$x_k \rightharpoonup x \quad \Longrightarrow \quad G(x_k) \rightharpoonup G(x)$$

*for all sequences $(x_k) \subset X$,*

*(b) $f\colon X \to \mathbb{R}$ is weakly lower semicontinuous,*

*(c) $K \subseteq Z$ is closed and convex,*

*(d) there exist a bounded set $C \subset \mathcal{F}$ and $x_0 \in C$ such that $f(x) \geq f(x_0)$ for all $x \in \mathcal{F} \setminus C$.*

*Then the problem (P) possesses a global solution.*

*Proof.* By assumption, analogously to the proof of Theorem 2.1, there exists a minimizing sequence $(x_k) \subset C$ with $f(x_k) \to f^* := \inf_{x \in \mathcal{F}} f(x)$. Since $C$ is bounded and $X$ is reflexive, there exists a subsequence, again denoted by $(x_k)$, that converges weakly to a limit $\bar{x} \in X$.

Therefore, by weak sequential continuity, $G(x_k) \rightharpoonup G(\bar{x})$ in $Z$. The closed convex set $K$ is weakly closed in $Z$, and thus $K \ni G(x_k) \rightharpoonup G(\bar{x}) \in K$. We conclude $\bar{x} \in \mathcal{F}$. Now, since $f\colon X \to \mathbb{R}$ is weakly lower semicontinuous, we obtain

$$f^* = \lim_{k\to\infty} f(x_k) = \liminf_{k\to\infty} f(x_k) \geq f(\bar{x}).$$

and hence $\bar{x} \in \mathcal{F}$ is a solution of (P). □

Of course, Remark 2.2 applies again for Theorem 2.4. We note that weak continuity, as assumed for $G$ in Theorem 2.4, is a delicate topic when nonlinear functions are involved. Generally, it is only to be expected to hold if the mapping under consideration is in fact (affine-) linear and continuous (see the exercises), or if there is *compactness* involved.

**Reminder:** A continuous linear operator $A \in \mathcal{L}(X;Y)$ between Banach spaces $X$ and $Y$ is said to be *compact* if it maps bounded sets in $X$ to precompact sets in $Y$.

Compactness is a very useful technique to transfer properties for the weak topology to properties for the norm topology, as the following lemma demonstrates.

**Lemma 2.5.** *Let $A \in \mathcal{L}(X;Y)$ be a compact linear operator between the Banach spaces $X$ and $Y$. If $(x_k) \subset X$ converges weakly to $x$ in $X$, then $(Ax_k) \subset Y$ converges strongly to $Ax$ in $Y$.*

*Proof.* Note first that weak convergence $x_k \rightharpoonup x$ in $X$ also implies $Ax_k \rightharpoonup Ax$ in $Y$ (why?).

Assume that $(Ax_k)$ does not converge in norm to $Ax$ in $Y$. Then there is a sufficiently small neighborhood $\mathcal{U}$ of $Ax$ in $Y$ and a subsequence $(Ax_{k_\ell})_\ell$ such that $Ax_{k_\ell} \notin \mathcal{U}$ for all $\ell \in \mathbb{N}$.

On the other hand, the weakly convergent sequence $(x_k)$ is bounded in $X$, hence $(Ax_k)$ is a precompact set in $Y$. In particular, the sequence $(Ax_{k_\ell})_\ell$ must admit a norm convergent subsequence, still denoted by $(Ax_{k_\ell})_\ell$. Denoting the limit of that subsequence by $y$, we must have $y \neq Ax$ by construction. But then the sequence $(Ax_{k_\ell})_\ell$ also converges weakly to $y \neq Ax$. Since $(Ax_{k_\ell})_\ell$ is a subsequence of $(Ax_k)$, this is a contradiction to $Ax_k \rightharpoonup Ax$. □

The ansatz to transfer properties for the weak topology to the norm topology is most often used for embeddings.

**Reminder:** A Banach space $X$ is said to be *embedded* (`eingebettet`) into another Banach space $X_0$ if there exists a continuous linear injective operator $i\colon X \to X_0$, the *embedding*, and we write $X \hookrightarrow X_0$ in this case.

Since an embedding $i\colon X \to X_0$ is by definition injective, it is often most useful to identify $X$ with its image $i(X) \subset X_0$. In most cases, the embedding operator $i$ is in fact given by the identity mapping $i = \mathrm{id}\colon X \to X_0$. In this case, we usually do not refer to $i$ explicitly. (It is sometimes reasonable to do so, however.)

**Definition 2.6.** A Banach space $X$ is *compactly* embedded in the Banach space $X_0$ if $X \hookrightarrow X_0$ and every bounded sequence in $X$ contains a (strongly) convergent subsequence in $X_0$. Equivalently, the embedding operator $i$ is a compact linear operator from $X$ to $X_0$.

**Corollary 2.7.** *If the Banach space $X$ is compactly embedded in the Banach space $X_0$, then every weakly convergent sequence $(x_k)$ with limit $x$ in $X$ is norm convergent to the same limit in $X_0$ up to identification of $X$ with $i(X) \subset X_0$.*

*Proof.* Apply Lemma 2.5 to the embedding operator $i\colon X \to X_0$. □

Using the foregoing corollary, we observe the following:

- If the embedding $X \hookrightarrow X_0$ is compact and $f_0\colon X_0 \to \mathbb{R}$ is lower semicontinuous, then $f := (f_0 \circ i)\colon X \to \mathbb{R}$ is weakly lower semicontinuous.

  In fact, $x_k \rightharpoonup x$ in $X$ implies $i(x_k) \to i(x)$ in $X_0$ and $\liminf_{k\to\infty} f_0(i(x_k)) \geq f_0(i(x))$.

- If the embedding $X \hookrightarrow X_0$ is compact and $G_0\colon X_0 \to Z$ is sequentially continuous from the norm topology of $X_0$ to the weak topology of $Z$ (*strong-weak continuous*), then $G := (G_0 \circ i)$ is weakly sequentially continuous from $X$ to $Z$.

  This follows from

$$x_k \rightharpoonup x \text{ in } X \quad \Longrightarrow \quad i(x_k) \to i(x) \text{ in } X_0 \quad \Longrightarrow \quad G_0(i(x_k)) \rightharpoonup G_0(i(x)) \text{ in } Z.$$

We give an example for an optimal control problem to which we apply the above existence theory.

**Example 2.8.** Consider the following optimal boundary control problem for a semilinear elliptic equation:

$$\begin{aligned} \min_{y,u} J(y,u) := \;& \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\partial\Omega)} \\ \text{s.t.}\;& \begin{cases} -\Delta y + y^3 = 0 & \text{on } \Omega, \\ \dfrac{\partial y}{\partial \nu} + y = u & \text{on } \partial\Omega, \\ a \leq u \leq b & \text{on } \partial\Omega, \end{cases} \end{aligned}$$

where $\Omega \subset \mathbb{R}^n$, for $n = 2$ or $n = 3$, is an open and bounded Lipschitz domain, $y_d \in L^2(\Omega)$, $\alpha \geq 0$, and, with the surface measure $\omega$ on $\partial\Omega$,

$$a, b \in L^2(\partial\Omega) \quad \text{with} \quad a \leq b \quad \omega\text{-a.e. in } \partial\Omega.$$

Let $U = L^2(\partial\Omega)$ and $Y = H^1(\Omega)$, and

$$U_{\text{ad}} = \Big\{ u \in U \colon a \leq u \leq b \ \omega\text{-a.e. in } \partial\Omega \Big\}.$$

We consider weak solutions of the state equation encoded in $E(y,u) = 0$, i.e.,

$$\begin{aligned} &E \colon Y \times U \to Y^*, \\ &\langle E(y,u), v \rangle_{Y^*,Y} := (\nabla y, \nabla v)_{L^2(\Omega)^n} + (y^3, v)_{L^2(\Omega)} + (y - u, v)_{L^2(\partial\Omega)} \text{ for all } v \in Y. \end{aligned}$$

This formulation is obtained by testing (that is, multiplying and integrating) the PDE with $v \in H^1(\Omega) = Y$ and integrating by parts. Note that, by Sobolev embeddings, we have $H^1(\Omega) \hookrightarrow L^6(\Omega)$ for $n \leq 3$ and thus $y^3 \in L^2(\Omega)$. Moreover, for $w \in H^1(\Omega)$ there is a well defined notion of the restriction $w|_{\partial\Omega} \in L^2(\partial\Omega)$ and $w \mapsto w|_{\partial\Omega}$ is continuous between these spaces. (Why is this mapping not an embedding?)

One can show that for each $u \in U$, the state equation possesses a unique solution $y(u) \in Y$. This defines the control-to-state operator $U \ni u \mapsto y(u) \in Y$. Furthermore, there holds

$$\|y(u)\|_Y \leq C \|u\|_U. \tag{2.1}$$

with a constant $C > 0$ independent of $u$. The estimate (2.1) follows by choosing $v = y(u)$ in the weak formulation, which gives:

$$\begin{aligned} \|\nabla y\|^2_{L^2(\Omega)^n} + \|y\|^4_{L^4(\Omega)} + \|y\|^2_{L^2(\partial\Omega)} &= (u, y)_{L^2(\partial\Omega)} \leq \|u\|_{L^2(\partial\Omega)} \|y\|_{L^2(\partial\Omega)} \\ &\leq \frac{1}{2} \|u\|^2_{L^2(\partial\Omega)} + \frac{1}{2} \|y\|^2_{L^2(\partial\Omega)}. \end{aligned}$$

By the generalized Friedrichs' inequality, there holds

$$\|y\|^2_{L^2(\Omega)} \leq C(\Omega) \left[ \|\nabla y\|^2_{L^2(\Omega)^n} + \|y\|^2_{L^2(\partial\Omega)} \right]$$

with a constant $C(\Omega)$ independent of $y$, from which now (2.1) follows. (Work this out!)

In the terminology of (P), we put

$$X = Y \times U, \quad Z = Y^* \times U, \quad K = \{0_{Y^*}\} \times U_{\text{ad}},$$

with $G(y,u) := (E(y,u), u)$ and $f := J$. We verify the requirements of Theorem 2.4. There holds:

(a) The mapping $E\colon Y \times U \to Y^*$ is weakly sequentially continuous:

In fact, $Y = H^1(\Omega)$ is compactly embedded into $L^5(\Omega)$ for $n = 2, 3$. Further, the function $\varphi(t) := t^3$ clearly satisfies

$$|\varphi(t)| = |t|^3 = |t|^{\frac{5}{(\frac{5}{3})}},$$

such that the superposition operator $y \mapsto y^3$ is continuous from $L^5(\Omega)$ to $L^{\frac{5}{3}}(\Omega)$. (Exercises!) Further, since $Y = H^1(\Omega) \hookrightarrow L^{\frac{5}{2}}(\Omega)$ densely, we also have the adjoint embedding $(L^{\frac{5}{2}}(\Omega))^* = L^{\frac{5}{3}}(\Omega) \hookrightarrow Y^*$. By Corollary 2.7 and the following factorization diagram, this means that the mapping $Y \ni y \mapsto y^3 \in Y^*$ is sequentially continuous from the weak topology on $Y$ to the norm topology on $Y^*$:

$$Y \xrightarrow[\text{compact}]{\text{id}} L^5(\Omega) \xrightarrow[\text{continuous}]{y \mapsto y^3} L^{\frac{5}{3}}(\Omega) \xrightarrow[\text{continuous}]{\hookrightarrow} Y^*$$

The operator $(y, u) \mapsto E(y, u) - y^3$ is linear and continuous from $Y \times U$ to $Y^*$, hence weakly sequentially continuous. Therefore, $E$ is weakly sequentially continuous.

The second component of $G$ is the identity id on $U$ and as such obviously weakly sequentially continuous. Overall, $G$ is weakly sequentially continuous from $Y \times U$ to $Y^* \times U$.

(b) The function $J$ is convex and continuous on $Y \times U$, hence weakly lower semicontinuous by Proposition 2.3 (c). (Here note that $Y = H^1(\Omega) \hookrightarrow L^2(\Omega)$ and that the norm on a Banach space is convex and by definition continuous.)

(c) $U_{\text{ad}} \subset U$ is bounded, closed and convex.

(d) The feasible set $\mathcal{F} = \{(y, u) \in Y \times U\colon E(y, u) = 0,\ u \in U_{\text{ad}}\}$ is nonempty and bounded:

Since $U_{\text{ad}}$ is bounded, the set $\mathcal{F}$ is also bounded by (2.1). Further, $(y(a), a) \in \mathcal{F}$.

Thus, the assumptions of Theorem 2.4 are verified and there exists a globally optimal solution to the semilinear optimal boundary control problem.

# 3 Optimality conditions

We now turn to optimality conditions for the general problem (P), both of first order necessary– and second order sufficient type. These will be conditions which characterize optimal solutions of (P). Since (P) is in general a nonconvex problem, there will be possibly multiple local solutions which are not necessarily also global ones. (The techniques in the previous section yield existence of a *globally* optimal solution, though!) We can in general only expect characterizations of *local* solutions of (P).

Before we begin with the actual definitions and statements leading to optimality conditions, we recall and define some concepts which will be needed in the following. Firstly, from the name *first* or *second order* optimality conditions, it is clear that derivatives will play a decisive role from now on.

**Definition 3.1** (Differentiability)**.** Let $X, Y$ be Banach spaces and let $F\colon X \supseteq U \to Y$ be an operator between $X$ and $Y$ defined on an open subset $U \neq \emptyset$ of $X$.

(a) We say that $F$ is *directionally differentiable* (`richtungsdifferenzierbar`) *at* $x \in U$ if the limit

$$\mathrm{d}F(x,h) = \lim_{t \searrow 0} \frac{F(x+th) - F(x)}{t} \in Y$$

exists for all $h \in X$. In this case, $\mathrm{d}F(x,h)$ is called *directional derivative* of $F$ at $x$ in the direction $h$.

(b) Moreover, $F$ is called *Gâteaux differentiable* (G-differentiable) *at* $x \in U$ if $F$ is directionally differentiable at $x$ and the directional derivative as a function in $h$, so $X \ni h \mapsto \mathrm{d}F(x,h) \in Y$, is bounded and linear, i.e., it is given by a linear operator $A \in \mathcal{L}(X;Y)$ for which we set $F'(x) := A$.

(c) Finally, we say that $F$ is *Fréchet differentiable* (F-differentiable) *at* $x \in U$ if $F$ is Gâteaux differentiable at $x$ and the following approximation condition holds:

$$\left\|F(x+h) - F(x) - F'(x)h\right\|_Y = o\big(\|h\|_X\big) \quad \text{for } h \to 0 \text{ in } X.$$

(d) If $F$ is directionally-/G-/F-differentiable at every $x \in V$ for $V \subseteq U$ open, then $F$ is called *directionally-/G-/F-differentiable on* $V$.

(e) If $F$ is G- or F-differentiable on a neighborhood of $x \in U$ and the derivative mapping $X \ni x \mapsto F'(x) \in \mathcal{L}(X;Y)$ is continuous, then we say that $F$ is *continuously G-/F-differentiable*.

**Reminder:** The $o$ notation in the definition of F-differentiability means that

$$\lim_{h \to 0} \frac{\left\|F(x+h) - F(x) - F'(x)h\right\|_Y}{\|h\|_X} = 0,$$

so, when $h \to 0$ in $X$, then the linear approximation error $\left\|F(x+h) - F(x) - F'(x)h\right\|_Y$ goes to zero *much faster* than $\|h\|_X$.

**Remark 3.2.** We note that there are concepts to deal with functions where the directional derivative is a nonlinear function in the direction $h$, but an approximation condition analogous to the one posed for F-differentiability holds true; this is so-called *Bouligand differentiability*. We will however not need these concepts for this lecture.

**Example 3.3.** Classical examples for differentiable mappings in Banach spaces include (see the exercises):

(a) Every bounded linear operator $A \in \mathcal{L}(X;Y)$ is continuously F-differentiable and its derivative in every point $x \in X$ is given by $A$.

(b) The quadratic form $X \ni u \mapsto \frac{1}{2}a(u,u) \in \mathbb{R}$ induced by a (symmetric) continuous bilinear form $a\colon X \times X \to \mathbb{R}$ is continuously $F$-differentiable and its derivative in $u \in X$ is given by $h \mapsto a(u,h)$.

(c) The superposition operator given by $\sin\colon \mathbb{R} \to \mathbb{R}$ is continuously F-differentiable as a mapping from $L^\infty(0,1)$ into itself with the derivative $h \mapsto \cos(f)h$ for every $f \in L^\infty(0,1)$. The operator is *not* F-differentiable as a mapping from $L^p(0,1)$ into itself for any $1 \le p < \infty$.

By imitating the proof for the finite-dimensional case, many classical formulas as the following are easily established. (Work this out!)

**Proposition 3.4.** *Let $X, Y$ and $Z$ be Banach spaces.*

*(a) If $F, G\colon X \to Y$ are (continuously) F-differentiable, then so are $\alpha F + \beta G$ for every $\alpha, \beta \in \mathbb{R}$ and $(\alpha F + \beta G)' = \alpha F' + \beta G'$.*

*(b) If $F\colon X \to Y$ and $G\colon Y \to Z$ are (continuously) F-differentiable in $x \in X$ and $F(x) \in Y$, respectively, then $G \circ F$ is (continuously) F-differentiable in $x$ and its derivative is given by*

$$\big(G \circ F\big)'(x) = G'\big(F(x)\big) \circ F'(x) \in \mathcal{L}(X;Z),$$

*where the latter means that $(G \circ F)'(x)h$ is given by the application of the operator $G'(F(x)) \in \mathcal{L}(Y;Z)$ to $F'(x)h \in Y$ for every $h \in X$.*

The first main goal of this chapter is to ultimately derive conditions under which, given a local solution $\bar{x}$ of (P), a so-called *Lagrange multiplier* (associated to $\bar{x}$) $\bar{\lambda} \in Z^*$ exists. The generic tools to prove *existence* of functionals in a dual space such as $Z^*$ are *separation theorems* which we will use quite frequently. (Such arguments are also used in the finite-dimensional case for Nonlinear Optimization, or even already in Linear Programming, usually in so-called *Farkas lemmas*.) We thus recall the geometric, or separation, versions of the fundamental *Hahn-Banach theorem* [Br10, Ch. 1.2].

**Reminder:** For two subset $A, B \subset X$ of a Banach space $X$, we say that the hyperplane

$$H = [f = \alpha] := \Big\{x \in X \colon \langle f, x\rangle = \alpha\Big\}$$

induced by $0 \neq f \in X^*$ and $\alpha \in \mathbb{R}$ *separates* (`trennt`) $A$ and $B$ if $\langle f, x\rangle \leq \alpha$ for all $x \in A$ and $\langle f, x\rangle \geq \alpha$ for all $x \in B$. We say the hyperplane $H$ *strictly separates* $A$ and $B$ if there exists $\varepsilon > 0$ such that $\langle f, x\rangle \leq \alpha - \varepsilon$ for all $x \in A$ and $\langle f, x\rangle \geq \alpha + \varepsilon$ for all $x \in B$.

**Proposition 3.5** (Hahn-Banach, geometric form)**.** *Let $X$ be a Banach space and let $A, B \subset X$ be two nonempty convex subsets such that $A \cap B = \emptyset$.*

*(a) Assume that $A$ or $B$ is open. Then there exists a hyperplane separating $A$ and $B$.*

*(b) Assume that $A$ is closed and that $B$ is compact. Then there exists a hyperplane strictly separating $A$ and $B$.*

A general class of first order necessary optimality conditions is obtained by observing that the Fréchet derivative at a local solution $\bar{x}$ is nonnegative along all directions that are tangential to the feasible set or point into the feasible set. These directions are characterized by the *tangent cone* (`Tangentialkegel`) of the feasible set at the solution.

**Definition 3.6** (Tangent cone)**.** The *Bouligand tangent cone* (or *contingent cone*) of a set $M \subseteq X$, where $X$ is a Banach space, at $x \in M$ is defined by

$$T(M, x) = \Big\{d \in X \colon \exists\, (x_k) \subseteq M,\ x_k \to x,\ (\eta_k) > 0 \colon\ \eta_k(x_k - x) \to d\Big\}.$$

The set $T(M, x)$ is a closed cone for every $x \in M$.

See the exercises for the proof that $T(M, x)$ is closed. In the case when $M$ is *convex*, a—possibly—more intuitive description of $T(M, x)$ can be derived using the conical hull. This description relies on the idea that we are interested in all directions which point into the feasible set. We need two more notions, which are however of independent interest and will be used frequently in the following.

**Reminder:** Recall that the *Minkowski sum* of two sets $A, B \subset X$ is given by

$$A + B := \Big\{a + b \colon a \in A,\ b \in B\Big\}$$

with the convention $a + B := \{a\} + B$, allowing to write the *Minkowski difference* as

$$A - B := \Big\{c \in X \colon c + B \subseteq A\Big\}.$$

**Definition 3.7** (Conical hull, cone of radial directions)**.** Let $X$ be a vector space and let $\emptyset \neq A \subseteq X$ be convex. Then the *conical hull* (`konische Hülle`) of $A$ is given by

$$\operatorname{cone}(A) := \left\{ \lambda x \colon x \in A,\, \lambda > 0 \right\}.$$

It is the smallest cone which includes the set $A$. More generally, the *cone of radial directions* of $A$ in a point $y \in X$ is given by

$$\operatorname{cone}(A, y) := \operatorname{cone}(A - y) = \left\{ z \in X \colon z = \lambda(x - y),\, x \in A,\, \lambda > 0 \right\}.$$

It is the smallest cone $C$ such that $A \subseteq y + C$. See also Fig. 1.

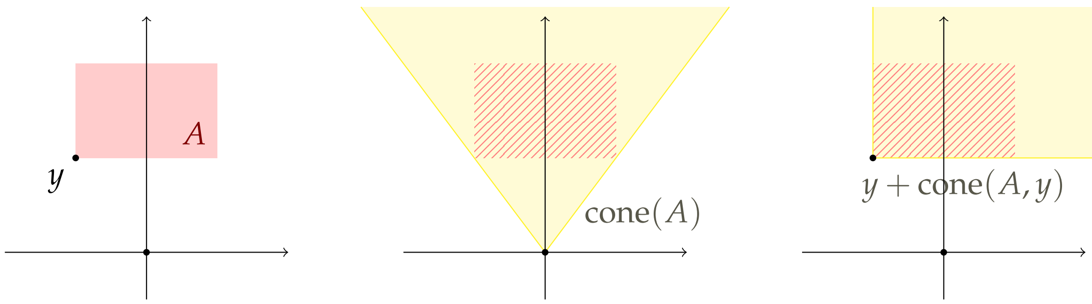


Figure 1: Conical hull of $A$ and cone of radial directions of $A$ at $y$

**Lemma 3.8.** *Let $M \subseteq X$ be convex and $x \in M$, then*

$$T(M, x) = \overline{\operatorname{cone}(M, x)}.$$

*Moreover,* $\operatorname{cone}(M, x)$ *and thus* $T(M, x)$ *are convex.*

*Proof.* Let $d \in T(M, x)$. Then there exist $(x_k) \subset M$ with $x_k \to x$ and $(\eta_k) > 0$ with $d_k := \eta_k(x_k - x) \to d$. By definition, $(d_k) \subset \operatorname{cone}(M, x)$. (Set $y = x_k$ and $\lambda = \eta_k$ for each $k$.) Hence, $\lim_{k\to\infty} d_k = d \in \overline{\operatorname{cone}(M, x)}$.

Conversely, to prove $\overline{\operatorname{cone}(M, x)} \subset T(M, x)$ it is sufficient to prove $\operatorname{cone}(M, x) \subset T(M, x)$ since $T(M, x)$ is closed. To this end, consider any $d \in \operatorname{cone}(M, x)$. Then there exist $\lambda > 0$ and $y \in M$ with $d = \lambda(y - x)$. Now set $\eta_k := k\lambda$ and $x_k := x + (y - x)/k$ for $k \geq 1$. Then $x_k \in M$ by convexity of $M$ and $x_k \to x$. Further, $\eta_k(x_k - x) = \lambda(y - x) = d$. This shows $d \in T(M, x)$.

We lastly show that if $M$ is convex, then so is $\operatorname{cone}(M, x)$. Since the closure of convex sets is convex (why?), this implies that $T(M, x)$ is convex, too. Let $d_1, d_2 \in \operatorname{cone}(M, x)$, i.e., for

$i = 1,2$, there are $\lambda_i > 0$ and $y_i \in M$ such that $d_i = \lambda_i(y_i - x)$. Then, for $\alpha \in (0,1)$, we have

$$(1-\alpha)d_1 + \alpha d_2 = \big((1-\alpha)\lambda_1 + \alpha\lambda_2\big)\Big(\tfrac{(1-\alpha)\lambda_1}{(1-\alpha)\lambda_1+\alpha\lambda_2}y_1 + \tfrac{\alpha\lambda_2}{(1-\alpha)\lambda_1+\alpha\lambda_2}y_2 - x\Big).$$

Since $M$ is assumed to be convex, $(1-\alpha)d_1 + \alpha d_2 \in \operatorname{cone}(M,x)$ follows. □

As apparent from the proof, the inclusion $T(M,x) \subseteq \overline{\operatorname{cone}(M,x)}$ is always true, independent of convexity of $M$. (Think of nonconvex set $M$ where $\overline{\operatorname{cone}(M,x)} \not\subseteq T(M,x)$!)

We now can state a first order optimality condition.

**Theorem 3.9.** *Let $f\colon U \to \mathbb{R}$ be defined on an open neighborhood $U$ of the set $M \subseteq X$, where $X$ is a Banach space. Let $\bar{x} \in M$ be a local minimizer of $f$ on $M$, i.e., a local solution of $\min_{x\in M} f(x)$ and assume that $f$ is F-differentiable at $\bar{x}$. Then there holds*

$$\langle f'(\bar{x}), d\rangle_{X^*,X} \geq 0 \quad \text{for all } d \in T(M,\bar{x}). \tag{3.1}$$

*Proof.* For all $d \in T(M,\bar{x})$ there exist sequences $(x_k) \subseteq M$ and $(\eta_k) > 0$ such that $x_k \to \bar{x}$ and $\eta_k(x_k - \bar{x}) \to d$.

We then have $\eta_k o\big(\|x_k - \bar{x}\|_X\big) \to 0$. Now, for sufficiently large $k$, there holds $f(x_k) \geq f(\bar{x})$ since $\bar{x}$ was a local minimum, and thus

$$\begin{aligned}\langle f'(\bar{x}), d\rangle &= \lim_{k\to\infty} \eta_k \langle f'(\bar{x}), x_k - \bar{x}\rangle = \lim_{k\to\infty}\Big[\eta_k\big(f(x_k) - f(\bar{x})\big) + \eta_k o\big(\|x_k - \bar{x}\|_X\big)\Big] \\ &\geq \lim_{k\to\infty} \eta_k o\big(\|x_k - \bar{x}\|_X\big) = 0.\end{aligned}$$

□

We are interested in applying this result to the problem (P), i.e., with $M = \mathcal{F}$. However, the cone $T(\mathcal{F},\bar{x})$ is difficult and often impossible to compute in practice. Hence, we approximate it by linearization and give conditions (so-called *constraint qualifications*) under which the linearizing cone and the contingent cone coincide. We recall that the set $K$ is convex as in the basic assumptions in Assumption 1.1.

**Definition 3.10** (Linearizing cone)**.** Let $G$ be F-differentiable at $x \in \mathcal{F} = G^{-1}[K]$. The *linearizing cone* at $x$ is given by

$$\begin{aligned}T_\ell(G,K,x) &= \Big\{d \in X\colon G'(x)d \in T(K,G(x))\Big\} \\ &= \Big\{d \in X\colon G'(x)d \in \overline{\operatorname{cone}(K,G(x))}\Big\}.\end{aligned}$$

**Remark 3.11.** Note that due to convexity of $K$, the linearizing cone $T_\ell(G,K,x)$ is always convex. (See Lemma 3.8.)

**Remark 3.12.** For the classical NLP

$$\min f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0$$

with $f\colon \mathbb{R}^n \to \mathbb{R}$, $g\colon \mathbb{R}^n \to \mathbb{R}^m$ and $h\colon \mathbb{R}^n \to \mathbb{R}^p$ continuously differentiable, given in our setting for (P) by $X = \mathbb{R}^n$, $Z = \mathbb{R}^m \times \mathbb{R}^p$, $G(x) = \binom{g(x)}{h(x)}$ and $K = (-\infty,0]^m \times \{0\}^p$, we indeed have (see the exercises)

$$T_\ell(G,K,x) = \Big\{d \in \mathbb{R}^n \colon \nabla h(x)^T d = 0,\ \nabla g_i(x)^T d \leq 0 \ \text{ for } i \in \mathcal{A}(x)\Big\},$$

where $\mathcal{A}(x) = \{i\colon g_i(x) = 0\}$ is the set of active inequality constraints. This is the standard linearizing cone from Nonlinear Optimization.

Our aim now is to infer from (3.1) that the optimality condition holds not only for directions $d$ from the cone $T(\mathcal{F},\bar{x})$, using $M = \mathcal{F}$, but also for directions from the cone $T_\ell(G,K,\bar{x})$. This, however, is in general only true if $T_\ell(G,K,\bar{x}) \subseteq T(\mathcal{F},\bar{x})$.

**Definition 3.13** (Abadie Constraint Qualification)**.** The *Abadie Constraint Qualification* (ACQ) holds true at a point $\bar{x} \in \mathcal{F}$ if the condition

$$T_\ell(G,K,\bar{x}) \subseteq T(\mathcal{F},\bar{x})$$

is satisfied. We call any condition which implies (ACQ) a *constraint qualification* (CQ).

Note that there are sets $\mathcal{F}$ where (ACQ) is *never* satisfied. For example, this is always the case when $T(\mathcal{F},x)$ is not convex, since we have noted in Remark 3.11 that $T_\ell(G,K,x)$ is always convex. However, the reverse inclusion is always true.

**Lemma 3.14.** *Let $G$ be F-differentiable at $x \in \mathcal{F} = G^{-1}[K]$. Then $T(\mathcal{F},x) \subseteq T_\ell(G,K,x)$.*

*Proof.* For $d \in T(\mathcal{F},x)$ there exist sequences $(x_k) \subseteq \mathcal{F}$ and $(\eta_k) > 0$ such that $x_k \to x$ and $\eta_k(x_k - x) \to d$. Without loss of generality, we can assume $\eta_k \to \infty$ (why?). We need to show that $G'(x)d \in T(K,G(x))$. As in the proof of Theorem 3.9, we find $\eta_k o\big(\|x_k - \bar{x}\|_X\big) \to 0$ and thus

$$\eta_k\big[G(x_k) - G(x)\big] = G'(x)\big[\eta_k(x_k - x)\big] + \eta_k o\big(\|x_k - x\|_X\big) \to G'(x)d.$$

Due to $\eta_k \to \infty$, or, alternatively, continuity of $G$, we have $G(x_k) \to G(x)$ and, of course, $G(x_k) \in K$ for every $k$. We conclude $G'(x)d \in T(K,G(x))$. □

## 3.1 Robinson's constraint qualification

The first CQ we consider is an algebraic-topological condition due to Robinson [Ro76] and will be of great relevance. It will be used to prove the *Karush-Kuhn-Tucker conditions* for (P), which are the most common form of first order optimality conditions.

**Definition 3.15** (Robinson's constraint qualification, regularity)**.** We say that *Robinson's constraint qualification* (RCQ) is satisfied for the problem (P) at $\bar{x} \in \mathcal{F}$ if there holds

$$0 \in \operatorname{int}\big(G(\bar{x}) + G'(\bar{x})X - K\big). \tag{3.2}$$

In this case, we also say that $\bar{x} \in \mathcal{F}$ is *regular*.

In this section we will show that (3.2) implies the ACQ. (So it *is* a CQ.) First, we establish the connection to a well known CQ in the finite-dimensional case.

**Example 3.16.** Consider the case of an NLP

$$\min f(x) \quad \text{s.t.} \quad g(x) \le 0, \quad h(x) = 0$$

as in Remark 3.12. We will show that in this case (3.2) is equivalent to the *Mangasarian Fromovitz constraint qualification* (MFCQ):

$$\operatorname{rank} \nabla h(\bar{x}) = p, \quad \exists\, d \in \mathbb{R}^n: \quad \nabla h(\bar{x})^T d = 0, \quad \nabla g_i(\bar{x})^T d < 0 \ \text{ for } i \in \mathcal{A}(\bar{x}). \tag{3.3}$$

We now prove this equivalence: Let $\bar{x}$ be a feasible point for the NLP, so $h(\bar{x}) = 0$ and $g(\bar{x}) \le 0$.

Let RCQ (3.2) be true for $\bar{x}$ satisfying $g(\bar{x}) \le 0$ and $h(\bar{x}) = 0$, i.e.,

$$0 \in \operatorname{int}\left\{ \begin{pmatrix} g(\bar{x}) + \nabla g(\bar{x})^T s - v \\ \nabla h(\bar{x})^T s \end{pmatrix} : s \in \mathbb{R}^n,\ v \in (-\infty, 0]^m \right\}.$$

The lower block requires that $\nabla h(\bar{x})^T$ is surjective, which implies $\operatorname{rank} \nabla h(\bar{x}) = p$. Now, let $\delta > 0$ and set $w \in \mathbb{R}^m$ by $w_i = -\delta$ if $g_i(\bar{x}) = 0$ and $w_i = 0$ if $g_i(\bar{x}) < 0$. Then

$$\begin{pmatrix} w \\ 0 \end{pmatrix} \in \left\{ \begin{pmatrix} g(\bar{x}) + \nabla g(\bar{x})^T s - v \\ \nabla h(\bar{x})^T s \end{pmatrix} : s \in \mathbb{R}^n,\ v \in (-\infty, 0]^m \right\}$$

if we choose $\delta$ sufficiently small. This means that there exist $s \in \mathbb{R}^n$ and $v \in (-\infty, 0]^m$ with

$$\begin{pmatrix} w \\ 0 \end{pmatrix} = \begin{pmatrix} g(\bar{x}) + \nabla g(\bar{x})^T s - v \\ \nabla h(\bar{x})^T s \end{pmatrix}.$$

Hence, $\nabla h(\bar{x})^T s = 0$ and, for all $i$ with $g_i(\bar{x}) = 0$:

$$\nabla g_i(\bar{x})^T s = w_i + v_i = -\delta + v_i \leq -\delta < 0.$$

Thus, the MFCQ (3.3) is satisfied for $d := s$.

Conversely, let the MFCQ (3.3) hold true and let $d \in \mathbb{R}^n$ be the corresponding vector. We show that there exists $\varepsilon > 0$ such that

$$B_{\varepsilon,\mathbb{R}^{m+p}} \subseteq \left\{ \begin{pmatrix} g(\bar{x}) + \nabla g(\bar{x})^T s - v \\ \nabla h(\bar{x})^T s \end{pmatrix} : s \in \mathbb{R}^n,\ v \in (-\infty, 0]^m \right\}.$$

Firstly, from $g_i(\bar{x}) < 0$ for $i \notin \mathcal{A}(\bar{x})$, we can find numbers $\delta_1, t > 0$ such that

$$g_i(\bar{x}) + \nabla g_i(\bar{x})^T td < -2\delta_1 \quad \text{for } i \notin \mathcal{A}(\bar{x}).$$

Secondly, for $i \in \mathcal{A}(\bar{x})$ we know that $\nabla g_i(\bar{x})^T d < 0$, thus we can find another number $\delta_2 > 0$ such that $\nabla g_i(\bar{x})^T td < -2\delta_2$ for all $i \in \mathcal{A}(\bar{x})$. But then we have

$$g_i(\bar{x}) + \nabla g_i(\bar{x})^T td < -2\delta \quad \text{for } i = 1, \dots, m,$$

where $\delta := \min(\delta_1, \delta_2)$, and there exists $\rho > 0$ such that

$$g_i(\bar{x}) + \nabla g_i(\bar{x})^T \left(td + s_0\right) < -\delta \quad \text{for } i = 1, \dots, m,$$

for all $s_0 \in B_{\rho,\mathbb{R}^n}(0)$. From the rank assumption on $\nabla h(\bar{x})$ in (3.3), we finally obtain a number $\varepsilon_1$ such that $B_{\varepsilon_1,\mathbb{R}^p}(0) \subseteq \nabla h(\bar{x})^T B_{\rho,\mathbb{R}^n}(0)$ (why?) and set $\varepsilon := \min(\delta, \varepsilon_1)$.

Now consider an arbitrary vector $w = \binom{w_1}{w_2} \in \mathbb{R}^{m+p}$ with $w \in B_{\varepsilon,\mathbb{R}^{m+p}}(0)$. Then we also have $\|w_i\|_2 < \varepsilon$ for $i = 1, 2$, and by construction of $\rho$ and $\varepsilon$ there exists a vector $s_0 \in B_{\rho,\mathbb{R}^n}(0) \subset \mathbb{R}^n$ with $w_2 = \nabla h(\bar{x})^T s_0 = \nabla h(\bar{x})^T (td + s_0)$. Choosing $s := td + s_0$ and

$$v := -w_1 + g(\bar{x}) + \nabla g(\bar{x})^T s,$$

we find $v_i < \varepsilon - \delta \leq 0$ and thus $v \in (-\infty, 0]^m$. But this means that

$$w = \begin{pmatrix} g(\bar{x}) + \nabla g(\bar{x})^T s - v \\ \nabla h(\bar{x})^T s \end{pmatrix}$$

and, since $w \in B_{\varepsilon,\mathbb{R}^{m+p}}(0)$ was arbitrary,

$$B_{\varepsilon,\mathbb{R}^{m+p}}(0) \subset \left\{ \begin{pmatrix} g(\bar{x}) + \nabla g(\bar{x})^T s - v \\ \nabla h(\bar{x})^T s \end{pmatrix} : s \in \mathbb{R}^n,\ v \in (-\infty, 0]^m \right\},$$

which is exactly Robinson's CQ (3.2) for $\bar{x}$. □

A particularly interesting case for RCQ is the one where $\operatorname{int} K \neq \emptyset$, which allows to obtain a slightly more direct formulation.

**Lemma 3.17.** *If* $\operatorname{int} K \neq \emptyset$*, then Robinson's CQ for* $\bar{x} \in \mathcal{F} = G^{-1}[K]$ *is equivalent to the existence of* $h \in X$ *with*

$$G(\bar{x}) + G'(\bar{x})h \in \operatorname{int} K, \tag{3.4}$$

*which is also called the* Linarized Slater constraint qualification *(LSCQ).*

*Proof.* From (3.4) it follows that for $\delta > 0$ sufficiently small there holds $G(\bar{x}) + G'(\bar{x})h + B_{Z,\delta}(0) \subset K$. Hence,

$$B_{Z,\delta}(0) \subset G(\bar{x}) + G'(\bar{x})h - K \subset G(\bar{x}) + G'(\bar{x})X - K$$

and RCQ is satisfied. We next show that if the LSCQ is not satisfied, then the RCQ also fails to hold. So, let (3.4) be violated. This means that the *convex* sets $G(\bar{x}) + G'(\bar{x})X$ and $\operatorname{int} K$ have an empty intersection. (Why is $\operatorname{int} K$ convex?) By the Hahn-Banach theorem as in Proposition 3.5, the sets can thus be separated by a hyperplane $[z' = \alpha]$, i.e., there exist $z' \in Z^* \setminus \{0\}$ and $\alpha \in \mathbb{R}$ with

$$\left\langle z', G(\bar{x}) + G'(\bar{x})h \right\rangle \geq \alpha \geq \langle z', z\rangle \quad \text{for all } h \in X,\ z \in K.$$

Now choose a vector $v \in Z$ with $\langle z', v\rangle_{Z^*,Z} < 0$. Then for all $t > 0$ there holds

$$\left\langle z', G(\bar{x}) + G'(\bar{x})h - z \right\rangle \geq 0 > \langle z', tv\rangle \quad \text{for all } h \in X,\ z \in K,$$

which shows that $tv \notin G(\bar{x}) + G'(\bar{x})X - K$ for all $t > 0$. But this means that Robinson's CQ (3.2) cannot hold. □

There are also more sufficient conditions and equivalencies to RCQ, in particular for the often-occurring case of multiple "blocks" of constraints with different structure. We collect some useful cases in the following proposition. For the proofs we refer to the exercises.

**Proposition 3.18.** *Let* $\bar{x} \in \mathcal{F} = G^{-1}[K]$ *be given.*

*(a) If* $G'(\bar{x}) \in \mathcal{L}(X;Z)$ *is surjective, then RCQ is satisfied for* $\bar{x}$.

*(b)* Ljusternik's theorem*: Let* $K = \{0_Z\}$*. Then RCQ for* $\bar{x}$ *is satisfied if and only if* $G'(\bar{x})\colon X \to Z$ *is surjective.*

*Now let G be of the form*

$$G = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} : X \to Z_1 \times Z_2 = Z$$

*and let accordingly* $K = K_1 \times K_2$ *where* $K_i \subseteq Z_i$ *for* $i = 1, 2$.

*(c) If* $G_1'(\bar{x}) \in \mathcal{L}(X;Z_1)$ *is surjective, then RCQ for* $\bar{x}$ *is equivalent to*

$$0 \in \operatorname{int}\Big(G_2(\bar{x}) + G_2'(\bar{x})\big(G_1'(\bar{x})^{-1}[K_1 - G_1(\bar{x})]\big) - K_2\Big).$$

(d) *If $G_1'(\bar{x}) \in \mathcal{L}(X; Z_1)$ is surjective and* $\operatorname{int} K_2 \neq \emptyset$ *in $Z_2$, then RCQ for $\bar{x}$ is equivalent to the existence of $h \in X$ such that*

$$\begin{aligned} G_1(\bar{x}) + G_1'(\bar{x})h &\in K_1, \\ G_2(\bar{x}) + G_2'(\bar{x})h &\in \operatorname{int} K_2. \end{aligned} \tag{3.5}$$

*In particular, if $K_1 = \{0_{Z_1}\}$, then the first of the two conditions in (3.5) collapses to* $h \in \ker G_1'(\bar{x})$.

We want to show that Robinson's CQ (3.2) implies the ACQ. To do so, we will need an estimate that allows us to bound the distance of a point $x$ from $G^{-1}[K]$ by means of the distance of $G(x)$ from $K$.

**Reminder:** The *distance* of a point $b$ to a set $A$ in a normed vector space is given by

$$\operatorname{dist}(b, A) := \inf_{a \in A} \|a - b\|.$$

A special case of a celebrated result by Robinson provides such an estimate, cf. [BS00, Thm. 2.87], see also [Ro76, Cor. 1] and [Ro76b]:

**Theorem 3.19.** *Assume that Robinson's CQ is satisfied at $\bar{x} \in \mathcal{F} = G^{-1}[K]$ and let $G\colon X \to Z$ be continuously F-differentiable near $\bar{x}$. Then there exist constants $c > 0$ and $\delta > 0$ such that*

$$\operatorname{dist}\big(x, G^{-1}[K - z]\big) \le c \operatorname{dist}\big(G(x) + z, K\big) \tag{3.6}$$

*for all $x \in B_{\delta,X}(\bar{x})$ and $z \in B_{\delta,Z}(0)$, where*

$$G^{-1}[K - z] = \Big\{x \in X \colon G(x) + z \in K\Big\}.$$

**Remark 3.20.**

(a) The condition (3.6) is called *metric regularity of $G$ at $\bar{x}$ with respect to $K$*. This is also the origin of calling $\bar{x}$ *regular* if it satisfies RCQ.

(b) For the special choice $z = 0$, we obtain

$$\operatorname{dist}(x, \mathcal{F}) \le c \operatorname{dist}(G(x), K)$$

for all $x \in B_{\delta,X}(\bar{x})$.

(c) In the case $K = \{0\}$, the metric regularity condition (3.6) is equivalent to

$$\operatorname{dist}\big(x, \{y \in X \colon G(y) = z\}\big) \le c\,\|G(x) - z\|_Z$$

for all $x \in B_{\delta,X}(\bar{x})$ and $z \in B_{\delta,Z}(0)$.

Using Theorem 3.19, we finally prove that Robinson's constraint qualification implies the ACQ.

**Theorem 3.21.** *Let $G$ be continuously F-differentiable near $\bar{x} \in \mathcal{F}$ and assume that Robinson's constraint qualification (3.2) holds at $\bar{x}$. Then the ACQ is satisfied at $\bar{x}$, i.e., $T(\mathcal{F}, \bar{x}) = T_\ell(G, K, \bar{x})$.*

*Proof.* We only need to show $T_\ell(G, K, \bar{x}) \subseteq T(\mathcal{F}; \bar{x})$, so consider an arbitrary direction $h \in T_\ell(G, K, \bar{x})$. Then, by definition, $G'(\bar{x})h = v \in T(K, G(\bar{x}))$ and there exist sequences $(z_k) \subseteq K$, $(\eta_k) > 0$ such that $z_k \to G(\bar{x})$ and $v_k := \eta_k(z_k - G(\bar{x})) \to v$ as $k \to \infty$. We can always choose $(z_k)$ and $(\eta_k)$ such that $\eta_k \to \infty$ (why?) and we suppose that this is the case from now on.

By Taylor expansion, or the definition of F-differentiability, we obtain

$$\begin{aligned} G\big(\bar{x} + \eta_k^{-1} h\big) &= G(\bar{x}) + G'(\bar{x})\eta_k^{-1} h + r_k(h) = G(\bar{x}) + \eta_k^{-1} v + r_k(h) \\ &= z_k + \eta_k^{-1}(v - v_k) + r_k(h), \end{aligned}$$

where $r_k(h) \in Z$ with $\|r_k(h)\|_Z = o\big(\eta_k^{-1}\|h\|_X\big)$ as $k \to \infty$. Hence, using Theorem 3.19 with $z = 0$ (cf. Remark 3.20) and $\eta_k^{-1} \to 0$, there exists $c > 0$ and $\ell > 0$ such that for all $k \ge \ell$ we have

$$\begin{aligned} \operatorname{dist}\big(\bar{x} + \eta_k^{-1} h, \mathcal{F}\big) &\le c \operatorname{dist}\big(G(\bar{x} + \eta_k^{-1} h), K\big) \\ &\le c\,\big\|G(\bar{x} + \eta_k^{-1} h) - z_k\big\|_Z = c\,\big\|\eta_k^{-1}(v - v_k) + r_k(h)\big\|_Z. \end{aligned}$$

Now, for each $k \ge \ell$, there exists an infimal sequence $(x_m^k)_m \subseteq \mathcal{F}$ for $\operatorname{dist}(\bar{x} + \eta_k^{-1} h, \mathcal{F})$. Thus there is $m_0(k)$ such that

$$\big\|\bar{x} + \eta_k^{-1} h - x_m^k\big\|_X \le \operatorname{dist}(\bar{x} + \eta_k^{-1} h, \mathcal{F}) + \frac{1}{k\eta_k} \quad \text{for } m \ge m_0(k).$$

Accordingly, the sequence $(x_k) := (x_{m_0(k)}^k)_k \subseteq \mathcal{F}$ satisfies

$$\big\|\bar{x} + \eta_k^{-1} h - x_k\big\|_X \le c\,\big\|\eta_k^{-1}(v - v_k) + r_k(h)\big\|_Z + \frac{1}{k\eta_k} \quad \text{for } k \ge \ell.$$

But this implies

$$\big\|\eta_k(x_k - \bar{x}) - h\big\|_X \le c\,\|v - v_k\|_Z + c\,\eta_k o\big(\eta_k^{-1}\|h\|_X\big) + \frac{1}{k} \longrightarrow 0 \quad \text{as } k \to \infty,$$

so $\eta_k(x_k - \bar{x}) \to h$ and also $x_k \to \bar{x}$ (recall that $\eta_k \to \infty$). This proves $h \in T(\mathcal{F}, \bar{x})$. □

## 3.2 The Karush-Kuhn-Tucker conditions

We will take advantage of the easier structure of $T_\ell(G, K, \bar{x})$ to derive optimality conditions of Karush-Kuhn-Tucker type under a constraint qualification. This will be achieved by applying a separation theorem, cf. Proposition 3.5. To prove the required interior point condition, Robinson's CQ (3.2) is used once again in a form that will be derived now.

We give the first result in an abstract formulation to make the idea more transparent. It builds upon the following extension of the open mapping theorem for multi-valued functions, again due to Robinson ([Ro72], [Ro76b, Thm. 1]).

**Reminder:** The fundamental *open mapping theorem* says that if $A \in \mathcal{L}(X; Y)$ is a continuous linear *surjective* operator between Banach spaces $X$ and $Y$, then it is an open mapping, i.e., $0 \in \operatorname{int}(AB_{r,X}(0))$ for all $r > 0$. The Banach space property for $X$ and $Y$ is crucial here.

Note that the graph of a continuous linear operator $A \in \mathcal{L}(X; Y)$, so the set $\{(x, y) \in X \times Y \colon Ax = y\}$, is closed and convex, and if $A$ is surjective, then $0 \in \operatorname{int}(AB_{r,X}(0))$ for some (and thus all) $r > 0$. Moreover, $A$ is surjective if and only if $0 \in \operatorname{int}(AX)$. In this sense, the following result is a true generalization of the classical open mapping theorem.

**Proposition 3.22** (Generalized open mapping theorem). *Let $\Psi \colon X \rightrightarrows Z$ be a set-valued function (i.e., $\Psi(x) \subseteq Z$ for all $x \in X$) between Banach spaces $X$ and $Z$ whose graph*

$$\operatorname{graph} \Psi = \Big\{(x, z) \in X \times Z \colon z \in \Psi(x)\Big\}.$$

*is a closed convex set, and let $z \in \operatorname{int} \Psi(X)$. Then $z \in \operatorname{int} \Psi(B_{r,X}(x))$ for all $r > 0$ and all $x \in \Psi^{-1}[\{z\}]$.*

**Lemma 3.23.** *Let $A \in \mathcal{L}(X; Z)$ with Banach spaces $X$ and $Z$ and let $C \subseteq Z$ be a closed convex set with $0 \in C$. Then the following assertions are equivalent:*

*(i)* $AX + \operatorname{cone}(C) = Z$, *and*

*(ii)* $0 \in \operatorname{int}\Big(A\overline{B_X(0)} + \big(C \cap \overline{B_Z(0)}\big)\Big)$.

*Proof.* We start with (ii) $\Longrightarrow$ (i). Let $z \in Z$. Choosing $\varepsilon > 0$ sufficiently small, (ii) means that

$$\varepsilon z \in A\overline{B_X(0)} + \big(C \cap \overline{B_Z(0)}\big) \qquad \Longrightarrow \qquad z \in A\overline{B_{\varepsilon^{-1},X}(0)} + \varepsilon^{-1}\big(C \cap \overline{B_Z(0)}\big)$$

and thus $z \in AX + \operatorname{cone}(C)$. Since $z$ was arbitrary, this implies (i).

Now we turn to (i) $\Longrightarrow$ (ii). We employ the generalized open mapping theorem, setting

$$\Psi\colon X\times\mathbb{R}\rightrightarrows Z,\quad \Psi(x,t):=\begin{cases}Ax+t\big(C\cap\overline{B_Z(0)}\big) & \text{if } t\geq 0,\\ \emptyset & \text{otherwise.}\end{cases}$$

Then by (i), $\Psi(X,\mathbb{R})=AX+\operatorname{cone}(C)=Z$ and thus clearly $\Psi(0,0)=0\in\operatorname{int}\Psi(X,\mathbb{R})$. Moreover, the graph of $\Psi$ is closed and convex and, due to Proposition 3.22, we obtain

$$\begin{aligned}0&\in\operatorname{int}\Psi\big(B_X(0),B(0)\big)=\operatorname{int}\Big(AB_X(0)+[0,1)\big(C\cap\overline{B_Z(0)}\big)\Big)\\ &\subset\operatorname{int}\Big(A\overline{B_X(0)}+\big(C\cap\overline{B_Z(0)}\big)\Big),\end{aligned}$$

where we have used that $0\in C$. □

With Lemma 3.23, we are now able to give an alternative form of the RCQ: the *Zowe-Kurcyusz constraint qualification* (ZKCQ) [ZK79] in a feasible point $\bar{x}\in\mathcal{F}=G^{-1}[K]$ given by

$$Z=G'(\bar{x})X-\operatorname{cone}(K,G(\bar{x})). \tag{3.7}$$

It is indeed equivalent to (RCQ), and even to the restricted version

$$0\in\operatorname{int}\Big(G'(\bar{x})\overline{B_X(0)}-\big((K-G(\bar{x}))\cap\overline{B_Z(0)}\big)\Big), \tag{3.8}$$

as the following lemma shows:

**Lemma 3.24.** *Robinson's constraint qualification* (3.2) *in $\bar{x}\in\mathcal{F}$ is equivalent to both forms of the Zowe-Kurcyusz constraint qualification, that is*

$$(3.2)\quad\Longleftrightarrow\quad(3.7)\quad\Longleftrightarrow\quad(3.8).$$

*Proof.* The equivalence of (3.7) and (3.8) is exactly the statement of Lemma 3.23 for the choices $A=G'(\bar{x})$ and $C=K-G(\bar{x})$, the latter being a closed convex set satisfying $0\in C$ thanks to $G(\bar{x})\in K$.

Robinson's CQ (3.2) follows from (3.8) immediately due to the existence of $\varepsilon>0$ such that

$$B_{\varepsilon,Z}(0)\subset G'(\bar{x})\overline{B_X(0)}-\big((K-G(\bar{x}))\cap\overline{B_Z(0)}\big)\subset G(\bar{x})+G'(\bar{x})X-K.$$

Lastly, from Robinson's CQ (3.2) we infer the (ZKCQ) (3.7) by considering $z\in Z$ and observing that for $\varepsilon>0$ sufficiently small we find $\varepsilon z\in G(\bar{x})+G'(\bar{x})X-K$ and thus

$$z\in G'(\bar{x})\varepsilon^{-1}X-\varepsilon^{-1}\big(K-G(\bar{x})\big)\subset G'(\bar{x})X+\operatorname{cone}(K,G(\bar{x})).$$

Since $z\in Z$ was arbitrary, this implies (3.7). □

For stating the main result of this section, we need the notion of the polar cone.

**Definition 3.25** (Polar cone)**.** Let $\emptyset \neq C \subseteq Z$. Then the set

$$C^\circ := \left\{z' \in Z^*\colon \langle z', z\rangle_{Z^*,Z} \leq 0 \text{ for all } z \in C\right\} \subseteq Z^*$$

denotes the *polar cone* of $C$, which is a closed convex cone.

**Theorem 3.26** (First-order necessary optimality conditions)**.** *Let $X$ and $Z$ be Banach spaces and $K \subseteq Z$ be closed and convex. Further, let $\bar{x}$ be a local solution of* (P) *at which $f\colon X \to \mathbb{R}$ and $G\colon X \to Z$ are continuously F-differentiable. Assume that Robinson's constraint qualification* (3.2) *is satisfied at $\bar{x}$.*

*Then there exists a* Lagrange multiplier $\bar{\lambda} \in Z^*$ *such that the* Karush-Kuhn-Tucker conditions *for* (P)

$$f'(\bar{x}) + G'(\bar{x})^*\bar{\lambda} = 0, \tag{3.9}$$

$$G(\bar{x}) \in K, \quad \bar{\lambda} \in T(K, G(\bar{x}))^\circ, \tag{3.10}$$

*are satisfied.*

*Proof.* To show the existence of a Lagrange multiplier, we define the set $M \subseteq \mathbb{R} \times Z$ as follows:

$$M = \left\{\left(\left\langle f'(\bar{x}), h\right\rangle_{X^*,X} + \sigma, G'(\bar{x})h - v\right)\colon h \in X,\, \sigma \geq 0,\, v + G(\bar{x}) \in K\right\}$$

The idea is to separate $\operatorname{int} M$ from the origin $(0,0)$ in $\mathbb{R} \times Z$ and to derive a Lagrange multiplier from the separating hyperplane. Due to the linear appearances of $h, \sigma$ and $v$ and the convexity of $K$, the set $M$ is quite obviously convex. Accordingly, $\operatorname{int} M$ is also convex (why?) and clearly open. In order to apply the Hahn Banach separation theorem as stated in Proposition 3.5, we need to show that $(0,0) \notin M$ and that $\operatorname{int} M$ is in fact nonempty.

We first claim that $(0,0)$ is a boundary point of $M$, so $(0,0) \notin \operatorname{int} M$. It is evident that $(0,0) \in M$, but every open neighborhood of $(0,0)$ in $\mathbb{R} \times Z$ must contain elements which are not in $M$, namely at least those of the form $(-\tau, 0)$ for $\tau > 0$. In fact, assume that $(-\tau, 0) \in M$ for some $\tau > 0$. We show that this contradicts the local optimality of $\bar{x}$ expressed by Theorem 3.9, which was

$$\left\langle f'(\bar{x}), d\right\rangle_{X^*,X} \geq 0 \quad \text{for all } d \in T(\mathcal{F}, \bar{x}) = T_\ell(G, K, \bar{x}),$$

where we have already used that Robinson's CQ implies the ACQ (Theorem 3.21). Indeed, if $(-\tau, 0) \in M$ for some $\tau > 0$, then there exist $h \in X$ and $\sigma \geq 0$ with $\langle f'(\bar{x}), h\rangle_{X^*,X} + \sigma = -\tau < 0$ and $G'(\bar{x})h = v \in K - G(\bar{x})$. But then $h \in T_\ell(G, K, \bar{x})$ and $\langle f'(\bar{x}), h\rangle_{X^*,X} < 0$. This is the contradiction.

Next, we show that $M$ has a nonempty interior. To this end, we use that (3.2) implies (3.8) as in Lemma 3.24. By (3.8), there exists $\delta > 0$ such that for any $z \in B_{\delta,Z}(0)$ there exist $h \in \overline{B_X(0)}$ and $v \in K - G(\bar{x})$ with $G'(\bar{x})h - v = z$. Moreover, $\langle f'(\bar{x}), h\rangle_{X^*,X} \le \|f'(\bar{x})\|_{X^*}$ due to $h \in \overline{B_X(0)}$. This shows that $\big[\|f'(\bar{x})\|_{X^*}, \infty\big) \times B_{\delta,Z}(0) \subseteq M$, hence $M$ has nonempty interior.

Now Proposition 3.5 shows that $(0,0)$ and $\operatorname{int} M$ can be separated by a hyperplane $[(\alpha, z') = \beta]$, i.e., there exist $\alpha \in \mathbb{R}$ and $z' \in Z^*$ with $(\alpha, z') \neq (0,0)$ such that

$$\left\langle \begin{pmatrix} \alpha \\ z' \end{pmatrix}, \begin{pmatrix} t \\ z \end{pmatrix} \right\rangle = \alpha t + \langle z', z\rangle_{Z^*,Z} \ge \beta \ge \left\langle \begin{pmatrix} \alpha \\ z' \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\rangle = 0 \quad \text{for all } (t,z) \in M.$$

Inserting the definition of $M$, the foregoing inequality means that

$$\alpha\big(\langle f'(\bar{x}), h\rangle_{X^*,X} + \sigma\big) + \langle z', G'(\bar{x})h - v\rangle_{Z^*,Z} \ge 0 \qquad \text{for all } h \in X,\ \sigma \ge 0,\ v \in K - G(\bar{x}). \tag{3.11}$$

We will derive the KKT conditions (3.9) and (3.10) from this inequality. For $h = 0$, $v = 0$ and $\sigma > 0$ we obtain $\alpha \ge 0$. Now assume that $\alpha = 0$. Then

$$\langle z', G'(\bar{x})\lambda h - \lambda v\rangle_{Z^*,Z} \ge 0 \quad \text{for all } h \in X,\ v \in K - G(\bar{x}),\ \lambda \ge 0,$$

from which we find that for $\alpha = 0$, (3.11) implies that

$$\langle z', z\rangle_{Z^*,Z} \ge 0 \quad \text{for all } z \in G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x})).$$

From the (ZKCQ) (3.7) we obtain the contradiction $z' = 0$, so we have $\alpha > 0$. This allows to multiply (3.11) by $\alpha^{-1}$ and to obtain, setting $\bar{\lambda} = \alpha^{-1} z'$:

$$\big(\langle f'(\bar{x}), h\rangle_{X^*,X} + \sigma\big) + \langle \bar{\lambda}, G'(\bar{x})h - v\rangle_{Z^*,Z} \ge 0 \quad \text{for all } h \in X,\ \sigma \ge 0,\ v \in K - G(\bar{x}).$$

The choice $h = 0$ and $\sigma = 0$ shows that $\bar{\lambda} \in \operatorname{cone}(K, G(\bar{x}))^\circ = T(K, G(\bar{x}))^\circ$.

Further, choosing $\sigma = 0$ and $v = 0$ shows that

$$\langle f'(\bar{x}), h\rangle_{X^*,X} + \langle \bar{\lambda}, G'(\bar{x})h\rangle_{Z^*,Z} \ge 0 \quad \text{for all } h \in X,$$

which is the same as

$$\langle f'(\bar{x}) + G'(\bar{x})^* \bar{\lambda}, h\rangle_{X^*,X} \ge 0 \quad \text{for all } h \in X.$$

This implies

$$f'(\bar{x}) + G'(\bar{x})^* \bar{\lambda} = 0.$$

Thus, the existence of a Lagrange multiplier is proved. □

So far, we have derived the KKT conditions under Robinson's CQ and used the equivalences to other conditions as proven in Lemma 3.24. One might wonder how restrictive Robinson's CQ or the equivalent (ZKCQ) (3.7) actually is. The following lemma proves necessary properties for the set of Lagrange multipliers associated to (P) and shows that these properties are also *nearly* sufficient for the (ZKCQ).

**Lemma 3.27.** *Let $\bar{x} \in \mathcal{F}$ be given and consider the set of Lagrange multipliers associated to (P) given by*

$$\Lambda(\bar{x}) = \Big\{\lambda \in T(K, G(\bar{x}))^\circ \colon f'(\bar{x}) + G'(\bar{x})^*\lambda = 0\Big\}.$$

*It is a closed and convex set characterized by the following assertions:*

1. *If the (ZKCQ)*
$$Z = G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x})) \tag{3.7}$$
*or equivalently (3.8) or Robinson's CQ (3.2) are satisfied in $\bar{x}$, then the set $\Lambda(\bar{x})$ is nonempty and bounded.*
2. *Let conversely $\Lambda(\bar{x})$ be nonempty and bounded. Then*
$$Z = \overline{G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x}))},$$
*i.e., $G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x}))$ is dense in $Z$.*

*Proof.* Closedness and convexity of $\Lambda(\bar{x})$ are evident since the polar cone is closed and convex (as an infinite intersection of closed half spaces) and the defining equation, so $f'(\bar{x}) + G'(\bar{x})^*\lambda = 0$, is linear and continuous w.r.t. $\lambda \in Z^*$.

We start with (1), so let one of the equivalent constraint qualifications (3.2), (3.7), or (3.8) be satisfied. Then Theorem 3.26 shows that $\Lambda(\bar{x})$ is nonempty and the KKT conditions (3.9) and (3.10) are satisfied, i.e.,

$$f'(\bar{x}) + G'(\bar{x})^*\bar{\lambda} = 0, \tag{3.9}$$
$$G(\bar{x}) \in K, \quad \bar{\lambda} \in T(K, G(\bar{x}))^\circ. \tag{3.10}$$

Due to $\bar{\lambda} \in T(K, G(\bar{x}))^\circ = \operatorname{cone}(K, G(\bar{x}))^\circ$, we know that

$$\big\langle \bar{\lambda}, v - G(\bar{x})\big\rangle_{Z^*,Z} \le 0 \quad \text{for all } v \in K.$$

Moreover, from (3.8) there exists $\delta > 0$ such that for all $z \in B_{\delta,Z}(0)$ there are $h \in \overline{B_X(0)}$ and $v \in K$ such that $-z = G(\bar{x}) + G'(\bar{x})h - v$. Applying $\bar{\lambda}$ to $z$ yields

$$\begin{aligned}
\big\langle \bar{\lambda}, z\big\rangle_{Z^*,Z} &= \big\langle \bar{\lambda}, v - G(\bar{x}) - G'(\bar{x})h\big\rangle_{Z^*,Z} \\
&= \big\langle \bar{\lambda}, v - G(\bar{x})\big\rangle_{Z^*,Z} - \big\langle f'(\bar{x}) + G'(\bar{x})^*\bar{\lambda}, h\big\rangle_{X^*,X} + \big\langle f'(\bar{x}), h\big\rangle_{X^*,X} \\
&\le \big\|f'(\bar{x})\big\|_{X^*}\|h\|_X \le \big\|f'(\bar{x})\big\|_{X^*}.
\end{aligned}$$

But this shows that

$$\big\langle \bar{\lambda}, \bar{z}\big\rangle_{Z^*,Z} \le \delta^{-1}\big\|f'(\bar{x})\big\|_{X^*} \quad \text{for all } \bar{z} \in B_Z(0)$$

since $z \in B_{\delta,Z}(0)$ was arbitrary, and thus

$$\|\bar{\lambda}\|_{Z^*} \leq \delta^{-1}\|f'(\bar{x})\|_{X^*}.$$

Since this estimate is uniform in $\bar{\lambda}$, the set $\Lambda(\bar{x})$ is bounded.

Now assume that $\Lambda(\bar{x})$ is nonempty and bounded. We argue via contradiction, so assume that there exists

$$\bar{z} \in Z \setminus M \quad \text{with} \quad M := \overline{G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x}))}.$$

The set $M$ is clearly closed and convex and contains 0, so it is nonempty. We use Proposition 3.5 to separate $\bar{z}$ and $M$: There exists an hyperplane $[z' = \alpha]$ such that

$$\langle z', z\rangle_{Z^*,Z} \geq \alpha \geq \langle z', \bar{z}\rangle_{Z^*,Z} \quad \text{for all } z \in M.$$

(In fact, Proposition 3.5 states that there is a hyperplane $H$ which even *strictly* separates $\bar{z}$ and $M$. We will however not need the strict separation.) Since $G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x}))$ is a cone, so is its closure $M$, hence the foregoing inequality implies $\langle z', z\rangle_{Z^*,Z} \geq 0$ for all $z \in M$: The right-hand side is a fixed number and we are allowed to scale the left-hand side by an arbitrary number $\lambda > 0$ by inserting $\lambda z$ for $z \in M$. The inequality can then only be true if $\langle z', z\rangle_{Z^*,Z} \geq 0$ for all $z \in M$, which in turn implies

$$\langle z', G'(\bar{x})h + G(\bar{x}) - v\rangle_{Z^*,Z} \geq 0 \quad \text{for all } h \in X,\ v \in K. \tag{3.12}$$

We will derive that $\bar{\lambda} + \beta z' \in \Lambda(\bar{x})$ for every $\bar{\lambda} \in \Lambda(\bar{x})$ and $\beta \geq 0$ from this inequality. Choosing $v = G(\bar{x}) \in K$ in (3.12) shows that

$$\langle z', G'(\bar{x})h\rangle_{Z^*,Z} = \langle G'(\bar{x})^* z', h\rangle_{X*,X} \geq 0 \quad \text{for all } h \in X$$

and thus (insert $h \in X$ and $-h \in X$)

$$G'(\bar{x})^* z' = 0 \quad \text{in } X^*. \tag{3.13}$$

Conversely, inserting $h = 0$ in (3.12) implies

$$\langle z', v - G(\bar{x})\rangle_{Z^*,Z} \leq 0 \quad \text{for all } v \in K,$$

so

$$z' \in \operatorname{cone}(K, G(\bar{x}))^\circ = T(K, G(\bar{x}))^\circ. \tag{3.14}$$

Now finally consider $\bar{\lambda} \in \Lambda(\bar{x})$ satisfying

$$f'(\bar{x}) + G'(\bar{x})^*\bar{\lambda} = 0, \tag{3.9}$$
$$G(\bar{x}) \in K, \quad \bar{\lambda} \in T(K, G(\bar{x}))^\circ. \tag{3.10}$$

From (3.13) and (3.14) we observe that

$$f'(\bar{x}) + G'(\bar{x})^*(\bar{\lambda} + \beta z') = 0,$$
$$G(\bar{x}) \in K, \quad \bar{\lambda} + \beta z' \in T(K, G(\bar{x}))^\circ,$$

so $\bar{\lambda} + \beta z' \in \Lambda(\bar{x})$ for *every* $\beta \geq 0$. Letting $\beta \to \infty$ gives a contradiction to the boundedness of $\Lambda(\bar{x})$. □

**Remark 3.28.** A couple $(0,0) \neq (\alpha, \bar{\lambda}) \in \mathbb{R}^+ \times Z^*$ satisfying the generalized KKT conditions

$$\alpha f'(\bar{x}) + G'(\bar{x})^* \bar{\lambda} = 0, \tag{3.15}$$

$$G(\bar{x}) \in K, \quad \bar{\lambda} \in T(K, G(\bar{x}))^\circ. \tag{3.16}$$

is called *generalized Lagrange multiplier*. The functional $z'$ constructed in the second point of the foregoing proof is a particular instance of a so-called *singular Lagrange multiplier* $(\alpha, \bar{\lambda}) = (0, z') \in \mathbb{R} \times Z^*$ satisfying the generalized KKT conditions (3.15) and (3.16) for $\alpha = 0$ and $\bar{\lambda} \neq 0$. If such a singular Lagrange multiplier exists, the set $\Lambda(\bar{x})$ can never be bounded, as seen in the foregoing proof.

One can show that, vice versa, the existence of a singular Lagrange multiplier implies that

$$Z \neq \overline{G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x}))}.$$

This is seen as follows: Let $(0, \bar{\lambda}) \neq (0,0)$ be a singular Lagrange multiplier. Then we have by (3.15) and (3.16)

$$\langle \bar{\lambda}, G'(\bar{x})h - v \rangle_{Z^*,Z} \geq 0 \quad \text{for all } h \in X,\, v \in \operatorname{cone}(K, G(\bar{x})),$$

and thus $-\bar{\lambda} \in (G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x})))^\circ$ from which there would follow $\bar{\lambda} = 0$ if $G'(\bar{x})X - \operatorname{cone}(K, G(\bar{x}))$ was dense in $Z$ (why?).

**Lemma 3.29.** *If $\bar{x}$ is a KKT-point, so $\Lambda(\bar{x}) \neq \emptyset$, then*

$$\langle f'(\bar{x}), d \rangle_{X^*,X} \geq 0 \quad \textit{for all } d \in T_\ell(G, K, \bar{x}).$$

*Proof.* For $d \in T_\ell(G, K, \bar{x})$ there holds $G'(\bar{x})d \in T(K, G(\bar{x}))$. Now with $\bar{\lambda} \in \Lambda(\bar{x})$, we have

$$\langle f'(\bar{x}), d \rangle_{X^*,X} = -\langle \bar{\lambda}, G'(\bar{x})d \rangle_{Z^*,Z} \geq 0,$$

since $\bar{\lambda} \in T(K, G(\bar{x}))^\circ$. □

The KKT conditions can be written very concisely by means of the Lagrange function which we have already encountered in Nonlinear Optimization:

**Definition 3.30** (Lagrangian)**.** The *Lagrange function* or *Lagrangian* $L \colon X \times Z^* \to \mathbb{R}$ for (P) is given by

$$L(x, \lambda) = f(x) + \langle \lambda, G(x) \rangle_{Z^*,Z}.$$

Using the Lagrangian, the first KKT expression can be expressed quite comfortably by

$$L'_x(\bar{x}, \bar{\lambda}) = f'(\bar{x}) + G'(\bar{x})^* \bar{\lambda} = 0.$$

We next turn to particular cases of (P) and their associated KKT conditions.

### 3.2.1 The case of cone constraints

As mentioned earlier, in the case where $K$ is a closed convex *cone*, the condition $G(x) \in K$ can be viewed as an abstract inequality constraint. In fact, we define a relation $\leq_K$ induced by $-K$ by

$$z_1 \leq_K z_2 \quad \Longleftrightarrow \quad z_2 - z_1 \in -K.$$

If $K$ is *pointed* (`spitz`), i.e., if $z, -z \in K$ implies that $z = 0$, then this is indeed a partial ordering, see the exercises. (A non-pointed cone $K$ is also sometimes called *flat* or *blunt*; there is an unfortunate amount of different notions for properties of cones in general.)

**Example 3.31.** In a function space $X$ consisting of functions defined on some set $\Omega \subset \mathbb{R}^d$, the cone of nonpositive functions

$$K_- := \left\{ f \in X \colon f(x) \leq 0 \ \text{ for almost all } x \in \Omega \right\}$$

is pointed and induces the usual pointwise ordering of functions, so $f \leq_{K_-} g$ if and only if $f(x) \leq g(x)$ for almost all $x \in \Omega$.

**Remark 3.32.** The seemingly unnecessarily complicated definition of the ordering $\leq_K$ is tailored to the classical notion of nonlinear programs: $z \leq_K 0$ means exactly that $z \in K$; so in case of cone constraints the standard constraint $G(x) \in K$ can be written as $G(x) \leq_K 0$.

Note however that the notation can be misleading, as the partial ordering induced by the cone of non*negative* functions $K_+$ (cf. Example 3.31) is given by $f \leq_{K_+} g$ if and only if $f(x) \geq g(x)$ for almost all $x \in \Omega \ldots$

We derive an easier representation of $T(K, \bar{z})^\circ$ for $\bar{z} \in K$ when $K$ is a closed convex cone. This is done using the *annihilator*.

**Reminder:** The *annihilator* $A^\perp$ of a set $A \subseteq X$ is given by

$$A^\perp := \left\{ x' \in X^* \colon \langle x', x \rangle = 0 \ \text{ for all } x \in A \right\},$$

so the collection of all functionals $x' \in X^*$ for which $A \subseteq \ker x'$.

**Lemma 3.33.** *If $K$ is a closed convex cone and $\bar z \in K$, then*

$$T(K,\bar z)^\circ = \operatorname{cone}(K,\bar z)^\circ = K^\circ \cap \{\bar z\}^\perp.$$

*Proof.* The first equality follows from $A^\circ = \overline{A}^\circ$ for every set $A$; see the exercises.

For the second equality: For every $z' \in K^\circ \cap \{\bar z\}^\perp$, there holds for all $t > 0$ and all $z \in K$:

$$\big\langle z', t(z-\bar z)\big\rangle_{Z^*,Z} = t\big\langle z', z\big\rangle_{Z^*,Z} - t\big\langle z', \bar z\big\rangle_{Z^*,Z} = t\big\langle z', z\big\rangle_{Z^*,Z} \le 0.$$

Hence, $\operatorname{cone}(K,\bar z)^\circ \supseteq K^\circ \cap \{\bar z\}^\perp$.

Conversely, let $z' \in \operatorname{cone}(K,\bar z)^\circ$. Due to $K$ being closed, we have $0 \in K$ (why?) and thus $-\bar z = 1\cdot(0-\bar z) \in \operatorname{cone}(K,\bar z)$. This implies $\langle z', -\bar z\rangle_{Z^*,Z} \le 0$. On the other hand, as above,

$$\big\langle z', t(z-\bar z)\big\rangle_{Z^*,Z} \le 0 \quad \text{for all } z \in K,\ t > 0. \tag{3.17}$$

Since $K$ is a cone, $z = 2\bar z \in K$, so $\langle z', \bar z\rangle_{Z^*,Z} \le 0$ But this means that both $\langle z', \bar z\rangle_{Z^*,Z} \le 0$ and $\langle z', -\bar z\rangle_{Z^*,Z} \le 0$, hence $z' \in \{\bar z\}^\perp$. Then, in (3.17), $\langle z', z\rangle_{Z^*,Z} \le 0$ for all $z \in K$, which is exactly the definition of $z' \in K^\circ$. □

**Reminder:** Recall that a cone $K$ is convex *if and only if* from $x, y \in K$ it follows that $x+y \in K$. Indeed, if $K$ is convex, then $\frac{1}{2}(x+y) \in K$, and since $K$ is a cone, $2\cdot\frac{1}{2}(x+y) = x+y \in K$. Conversely, if $x, y \in K$, then also $(1-t)x \in K$ and $ty \in K$ if $t \in [0,1]$, and then by assumption also $(1-t)x + ty \in K$ and $K$ is convex.

Hence, in the case of a closed convex cone $K$, the KKT condition

$$G(\bar x) \in K, \quad \bar\lambda \in T(K, G(\bar x))^\circ \tag{3.10}$$

can be written equivalently as a cone *complementarity condition*

$$G(\bar x) \in K, \quad \bar\lambda \in K^\circ, \quad \big\langle \bar\lambda, G(\bar x)\big\rangle_{Z^*,Z} = 0.$$

Using the cone ordering notation from above, we can rewrite the latter yet again to

$$G(\bar x) \le_K 0, \quad \bar\lambda \ge_{K^+} 0, \quad \big\langle \bar\lambda, G(\bar x)\big\rangle_{Z^*,Z} = 0,$$

where $K^+ := -K^\circ$ is the *dual cone* to $K$. The last representation quite exactly resembles the classical KKT conditions from Nonlinear Optimization.

### 3.2.2 The Slater condition

We now consider the case of convex problem, for which we first give a notion of convexity in function spaces.

**Definition 3.34** (Generalized convexity). Let $K \subseteq Z$ be a closed convex cone. We say that $G\colon X \to Z$ is *convex with respect to* $-K$ (or $\leq_K$), if

$$G\big((1-t)x + ty\big) \leq_K (1-t)G(x) + tG(y) \quad \text{for all } x, y \in X,\ t \in [0,1],$$

or equivalently

$$(1-t)G(x) + tG(y) - G\big((1-t)x + ty\big) \in -K \quad \text{for all } x, y \in X,\ t \in [0,1].$$

In case of F-differentiable and convex $G$, the *Slater constraint qualification* (or *Slater condition*), well-known from Nonlinear Optimization,

$$\exists\, \hat{x} \in \mathcal{F}\colon \qquad G(\hat{x}) \in \operatorname{int} K \tag{3.18}$$

implies the Robinson CQ:

**Lemma 3.35.** *Let $K \subseteq Z$ is a closed convex cone and let $G$ be F-differentiable and convex with respect to $\leq_K$. Assume that the Slater condition (3.18) is satisfied. Then RCQ (3.2) in the form of the Linearized Slater CQ (3.4) is satisfied in every feasible point $\bar{x} \in \mathcal{F} = G^{-1}[K]$.*

*Proof.* Using convexity of $G$ and the fact that $K$ is a cone, we observe that for every $t \in (0,1]$ and $\bar{x} \in \mathcal{F}$ we have

$$\frac{G\big((1-t)\bar{x} + t\hat{x}\big) - (1-t)G(\bar{x}) - tG(\hat{x})}{t} \in K.$$

Taking the limit $t \searrow 0$ gives

$$G(\bar{x}) + G'(\bar{x})(\hat{x} - \bar{x}) - G(\hat{x}) \in K \qquad \Longleftrightarrow \qquad G(\bar{x}) + G'(\bar{x})(\hat{x} - \bar{x}) \in K + G(\hat{x}).$$

By the Slater condition, there exists $\varepsilon > 0$ with $G(\hat{x}) + B_{\varepsilon,Z}(0) \subseteq K$. Therefore,

$$G(\bar{x}) + G'(\bar{x})(\hat{x} - \bar{x}) + B_{\varepsilon,Z}(0) \subseteq K + G(\hat{x}) + B_{\varepsilon,Z}(0) \subseteq K + K = K.$$

Thus,

$$G(\bar{x}) + G'(\bar{x})(\hat{x} - \bar{x}) \in \operatorname{int} K,$$

which is exactly the LSCQ (3.4) and thus equivalent to the Robinson CQ by Lemma 3.17. ∎

**Remark 3.36.** In the proof of Lemma 3.35, we have shown along the way that the differentiable convex function $G$ satisfies

$$G(y) - G(x) - G'(x)(y-x) \in -K \quad \text{or} \quad G'(x)(y-x) \leq_K G(y) - G(x) \tag{3.19}$$

for all $x, y \in \mathcal{F} = G^{-1}[K]$. This is exactly the analogue of the well-known characterization of classical convex functions $f\colon \mathbb{R}^n \to \mathbb{R}$, given by

$$\nabla f(x)^T (y-x) \leq f(y) - f(x).$$

The proof that (3.19) also implies convexity works again analogously to the classical case.

### 3.2.3 Applications

We proceed by giving two examples for the KKT conditions in an optimal control setting. The first one is still of rather abstract nature and incorporates control constraints, while the second one is slightly more specialized and has state constraints.

**Reminder:** Let $H$ be a Hilbert space. The *Fréchet-Riesz representation theorem* says that there exists an isometric isomorphism $T \in \mathcal{L}(H^*; H)$—the *Riesz isomorphism*—such that

$$\langle g, v \rangle_{H^*,H} = \left(Tg, v\right)_H \quad \text{for all } g \in H^*,\ v \in H,$$

where $(\cdot,\cdot)_H$ is the inner product on the Hilbert space $H$. In particular, $\|g\|_{H^*} = \|Tg\|_H$. In this sense, we can always identify a Hilbert space $H$ with its dual $H^*$ up to the application of the Riesz isomorphism.

**Example 3.37** (Optimal control problem with control constraints)**.** We consider a control-constrained optimal control problem

$$\min_{(y,u)\in Y\times U} J(y,u) \quad \text{s.t.} \quad E(y,u) = 0, \quad u \in U_{\text{ad}} \tag{3.20}$$

governed by the state equation

$$E(y,u) = 0,$$

where $E\colon Y \times U \to W$ is continuously differentiable, $Y$ and $W$ are Banach spaces, and $U = L^2(\Omega)$ for a bounded domain $\Omega \subset \mathbb{R}^n$. The objective function $J\colon Y \times U \to \mathbb{R}$ is assumed to be F-differentiable and the control constraints are given by

$$U_{\text{ad}} = \left\{ u \in U \colon a \leq u \leq b \text{ a.e. in } \Omega \right\}$$

with $a, b \in L^2(\Omega)$ and $a \leq b$ almost everywhere on $\Omega$, so $U_{\text{ad}} \neq \emptyset$.

Let $(\bar{y}, \bar{u}) \in Y \times U$ be a local solution to (3.20). We need a constraint qualification to be satisfied in order to derive a KKT characterization for $(\bar{y}, \bar{u})$. To this end, we first transfer the problem to standard form. This is obtained by identifying

$$X = Y \times U, \quad Z = W \times U \quad \text{and} \quad K = \{0_W\} \times U_{\text{ad}} \subseteq Z,$$

as well as $\bar{x} = (\bar{y}, \bar{u})$ and $G \colon X \to Z$ given by $G(x) = \binom{E(y,u)}{u}$.

We have seen in Proposition 3.18 that $G'(\bar{x})$ being surjective is a constraint qualification by implying RCQ. To show that $G'(\bar{x})$ is indeed surjective, we need to show that for any $(w, u) \in Z = W \times U$ there exists $(h_y, h_u) \in X = Y \times U$ such that

$$G'(\bar{x})h = \begin{pmatrix} E'_y(\bar{y}, \bar{u}) & E'_u(\bar{y}, \bar{u}) \\ 0 & \text{id}_U \end{pmatrix} \begin{pmatrix} h_y \\ h_u \end{pmatrix} = \begin{pmatrix} w \\ u \end{pmatrix},$$

where we have identified $G'(\bar{x})$ with its Jacobian matrix type representation, see the exercises. Due to the upper triangular form of this Jacobian of $G'(\bar{x})$, it will turn out that it is both sufficient and necessary to assume that $E'_y(\bar{y}, \bar{u})$ is surjective in order to have $G'(\bar{x})$ surjective. In fact, looking at the second row in the foregoing equality, we immediately see that necessarily $h_u = u$. Thus, surjectivity of $G'(\bar{x})$ is equivalent to, for every $(w, u) \in W \times U$, being able to find $h_y \in Y$ such that

$$E'_y(\bar{y}, \bar{u})h_y = w - E'_u(\bar{y}, \bar{u})u.$$

But this is exactly the question of surjectivity of $E'_y(\bar{y}, \bar{u})$. (Choose any $u \in U$ and consider $w := E'_u(\bar{y}, \bar{u})u) + v$ for arbitrary $v \in W$.)

So, if we *assume* that

$$E'_y(\bar{y}, \bar{u}) \text{ is surjective,} \tag{3.21}$$

then $G'(\bar{x})$ is surjective and RCQ for (3.20) is satisfied. Hence, there exist a Lagrange multiplier (pair)

$$\bar{\lambda} = (\bar{p}, \bar{\mu}) \in Z^* = (W \times U)^* = W^* \times U^* = W^* \times L^2(\Omega)$$

such that (3.9) and (3.10) are satisfied, i.e.,

$$\begin{aligned} J_y'(\bar y,\bar u) + E_y'(\bar y,\bar u)^*\bar p &= 0 && \text{in } Y^*,\\ J_u'(\bar y,\bar u) + E_u'(\bar y,\bar u)^*\bar p + \bar\mu &= 0 && \text{in } L^2(\Omega),\\ E(\bar y,\bar u) &= 0 && \text{in } W,\\ \bar u \in U_{\text{ad}}, \quad \bar\mu \in T(U_{\text{ad}},\bar u)^\circ. \end{aligned}$$

Note that $\{0_W\}^\circ = W^*$, thus the condition $\bar p \in \{0_W\}^\circ$ is void.

We next consider $T(U_{\text{ad}},\bar u)^\circ$. Since $U_{\text{ad}}$ is convex, we have $T(U_{\text{ad}},\bar u) = \overline{\operatorname{cone}(U_{\text{ad}},\bar u)}$ (Lemma 3.8). Using this, it is easy to see that

$$T(U_{\text{ad}},\bar u) = \Big\{h \in L^2(\Omega) \colon h|_{[\bar u = a]} \ge 0,\ h|_{[\bar u = b]} \le 0\Big\},$$

where $[\bar u = a] = \{x \in \Omega \colon \bar u(x) = a(x)\}$ and analogously for $[\bar u = b]$. Thus, $T(U_{\text{ad}},\bar u)^\circ$ consists exactly of all functions $s \in L^2(\Omega)$ such that

$$\int_\Omega h(x)s(x)\,\mathrm{d}x \le 0 \quad \text{for all } h \in L^2(\Omega) \text{ with } h|_{[\bar u = a]} \ge 0,\ h|_{[\bar u = b]} \le 0.$$

So let $s \in T(U_{\text{ad}},\bar u)^\circ$. Consider $M \subseteq [\bar u = a]$. Then $\chi_M \in T(U_{\text{ad}},\bar u)$, where

$$\chi_M(x) = \begin{cases} 1 & \text{if } x \in M,\\ 0 & \text{otherwise.}\end{cases}$$

Suppose that $s|_M > 0$ if the measure $|M| > 0$. We obtain

$$\int_\Omega \chi_M(x)s(x)\,\mathrm{d}x = \int_M s(x)\,\mathrm{d}x > 0,$$

which is a contradiction to $s \in T(U_{\text{ad}},\bar u)^\circ$. Thus $s|_{[\bar u = a]} \le 0$ almost everywhere. By the analogous argument with $-\chi_M \in T(U_{\text{ad}},\bar u)$ for $M \subseteq [\bar u = b]$, we find that then $s|_{[\bar u = b]} \ge 0$ almost everywhere. Finally, consider $M \subseteq [a < \bar u < b]$. Then $\pm\chi_M \in T(U_{\text{ad}},\bar u)$, and it follows that $s_{[a<\bar u<b]} = 0$. It is easily seen that these conditions together are also sufficient for $s \in T(U_{\text{ad}},\bar u)^\circ$. Hence,

$$T(U_{\text{ad}},\bar u)^\circ = \Big\{s \in L^2(\Omega) \colon s|_{[\bar u = a]} \le 0,\ s|_{[\bar u = b]} \ge 0,\ s|_{[a<\bar u<b]} = 0\Big\}.$$

Hence, the multiplier $\bar\mu$ for the control constraints is an $L^2(\Omega)$ function satisfying

$$\bar\mu|_{[\bar u = a]} \le 0, \quad \bar\mu|_{[\bar u = b]} \ge 0, \quad \bar\mu|_{[a<\bar u<b]} = 0,$$

which is usually interpreted as a complimentarity condition for the control constraints $a \le \bar u \le b$.

**Remark 3.38.** Note that $\operatorname{int} U_{\mathrm{ad}} = \emptyset$ in $U = L^2(\Omega)$ as seen in the exercises, so there is no other (practical) characterization of RCQ than surjectivity of $G'(\bar{x})$ available in the foregoing example.

**Example 3.39** (Elliptic optimal control problem with state contraints)**.** We consider an optimal control problem with an abstract elliptic state equation, control on the right-hand side and pointwise state constraints, that is, the constraints are of the form

$$Ay = Bu + b \quad \text{and} \quad y \leq \psi.$$

The state equation is to be seen as an abstract form of an elliptic partial differential equation in weak formulation by the following assumptions:

- $\Omega \subset \mathbb{R}^n$, where $1 \leq n \leq 3$, is a bounded Lipschitz domain,
- $B \in \mathcal{L}(U; L^2(\Omega))$, where the control space $U$ is a Banach space, and $b \in L^2(\Omega)$,
- $A \in \mathcal{L}(H_0^1(\Omega); H^{-1}(\Omega))$ with $A^{-1} \in \mathcal{L}(H^{-1}(\Omega); H_0^1(\Omega))$, and additionally, the mapping $y \mapsto Ay$ defines a bounded, injective and surjective operator from the state space $Y = H_0^1(\Omega) \cap H^2(\Omega)$ to $L^2(\Omega)$. As a consequence, $v \mapsto A^{-1}v$ defines the *solution operator* $S \in \mathcal{L}(L^2(\Omega); Y)$. See also Remark 3.40 below.

For the upper bound $\psi$ in the state constraint we suppose $\psi \in C(\overline{\Omega})$, and we describe the state constraint by

$$\mathcal{E}y - \psi \in K_-$$

with the well-known cone of nonnegative functions in $C(\overline{\Omega})$

$$K_- = \left\{ q \in C(\overline{\Omega}) \colon q(x) \leq 0 \text{ for all } x \in \overline{\Omega} \right\}$$

and the embedding $\mathcal{E} \in \mathcal{L}(Y; C(\overline{\Omega}))$ granted by the Sobolev embedding theorem. Together with the state equation $E \colon Y \times U \to L^2(\Omega)$ given by $E(y,u) = Ay - Bu - b$, we collect all constraints in the function $G \colon X \to Z$ by setting $X = Y \times U$ and $Z = L^2(\Omega) \times C(\overline{\Omega})$ and $G(x) = \binom{E(y,u)}{\mathcal{E}y - \psi} \in K$ with $K = \{0_{L^2(\Omega)}\} \times K_-$.

Since $\operatorname{int} K_- \neq \emptyset$ in $C(\overline{\Omega})$, the following linearized Slater-type assumption will be sufficient for RCQ at any feasible point $(\bar{y}, \bar{u})$ for this problem: There exist $\hat{u} \in U$ and $\hat{y} \in Y$ with $A\hat{y} = B\hat{u} + b$ and $\mathcal{E}\hat{y} - \psi < 0$ on $\overline{\Omega}$, so $(\hat{y}, \hat{u})$ is feasible for the equality constraint and satisfies the inequality constraint strictly.

Indeed, since $A\colon Y \to L^2(\Omega)$ is continuously invertible, it is in particular surjective, and so is $E'_y(y,u) = A$ and thus $E'(y,u)$ for all $(y,u) \in Y \times U$. Let $(\bar y, \bar u) \in Y \times U$ be feasible. Then $(h_y, h_u) := (\hat y - \bar y, \hat u - \bar u)$ satisfies

$$E'(\bar y,\bar u)(h_y,h_u) = Ah_y - Bh_u = A\hat y - B\hat u - \big(A\bar y - B\bar u\big) = b - b = 0.$$

Hence, we have $(h_y, h_u) \in \ker E'(\bar y, \bar u)$. Moreover, from $\mathcal{E}\hat y - \psi \in \operatorname{int} K_-$ we infer

$$\mathcal{E}\bar y - \psi + \mathcal{E}h_y = \mathcal{E}\hat y - \psi \in \operatorname{int} K_-,$$

and in (3.5) in Proposition 3.18 we have seen that surjectivity of $E'(\bar y, \bar u)$ and the existence of $(h_y, h_u) \in Y \times U$ with the mentioned properties is even equivalent to RCQ for this problem.

Hence, under the assumption that $(\hat y, \hat u)$ as above exists and letting $J\colon Y \times U \to \mathbb{R}$ be F-differentiable, we obtain a KKT characterization of a locally optimal control $(\bar y, \bar u)$ for the problem

$$\min_{(y,u)\in Y\times U} J(y,u) \quad \text{s.t.} \quad Ay = Bu + b, \quad y \le \psi$$

as follows:

There exist $\bar p \in L^2(\Omega)^* = L^2(\Omega)$ and $\bar\mu \in C(\overline{\Omega})^*$ such that

$$\begin{aligned} J'_y(\bar y,\bar u) + A^*\bar p + \mathcal{E}^*\bar\mu &= 0 && \text{in } Y^*,\\ J'_u(\bar y,\bar u) - B^*\bar p &= 0 && \text{in } U^*,\\ A\bar y &= B\bar u + b && \text{in } L^2(\Omega),\\ \mathcal{E}\bar y - \psi &\le 0 && \text{in } C(\overline{\Omega}),\\ \bar\mu \in T(K_-, \mathcal{E}\bar y - \psi)^\circ &= K_-^\circ \cap (\mathcal{E}\bar y - \psi)^\perp && \text{in } C(\overline{\Omega})^*, \end{aligned}$$

where the last equality follows from Lemma 3.33, since $K_-$ is a closed convex cone.

The last conditions can be rewritten as a complimentarity condition

$$\mathcal{E}\bar y - \psi \le 0, \quad \bar\mu \in K_-^\circ, \quad \langle \bar\mu, \mathcal{E}\bar y - \psi\rangle_{C(\overline{\Omega})^*, C(\overline{\Omega})} = 0.$$

Here, one can show [La93, Ch. IX, Thm. 4.2] that there indeed holds $C(\overline{\Omega})^* = M(\overline{\Omega})$, where $M(\overline{\Omega})$ is the space of (real) regular Borel measures on $\overline{\Omega}$ with the *total variation* norm

$$\|\mu\|_{M(\overline{\Omega})} := |\mu|(\overline{\Omega}) = \sup_{\|v\|_{C(\overline{\Omega})}\le 1} \int_{\overline{\Omega}} v \, \mathrm{d}\mu.$$

Thereby, the dual pairing of $\mu \in M(\overline{\Omega}) = C(\overline{\Omega})^*$ with $f \in C(\overline{\Omega})$ is given by

$$\langle \mu, f \rangle_{M(\overline{\Omega}),C(\overline{\Omega})} = \int_{\overline{\Omega}} f \,\mathrm{d}\mu.$$

Further, $\bar{\mu} \in K_-^\circ$ means that $\bar{\mu} \geq 0$ in the sense of $\langle \bar{\mu}, q \rangle \leq 0$ for all functions $q \in K_-$.

From these properties it follows that the above complementarity condition can be written as

$$\mathcal{E}\bar{y} - \psi \leq 0, \quad \bar{\mu} \geq 0, \quad \bar{\mu}\big([\mathcal{E}\bar{y} - \psi < 0]\big) = 0$$

or, in other words

$$T(K_-, \mathcal{E}\bar{y} - \psi)^\circ = K_-^\circ \cap (\mathcal{E}\bar{y} - \psi)^\perp = \Big\{\bar{\mu} \in M(\overline{\Omega}) \colon \bar{\mu} \geq 0,\, \bar{\mu}\big([\mathcal{E}\bar{y} - \psi < 0]\big) = 0\Big\}.$$

**Remark 3.40.** The regularity assumptions on $A$ in Example 3.39 are to be understood as follows: The basic assumption says that for every $f \in H^{-1}(\Omega)$, there exist a unique $y \in H_0^1(\Omega)$ such that $Ay = f$ and a constant $C > 0$ independent of $f$ such that $\|y\|_{H_0^1(\Omega} \leq C\|f\|_{H^{-1}(\Omega)}$.

This is the classical result obtained by the Lax-Milgram lemma for large classes of elliptic partial differential operators [Br10, Ch. 5.3], such as exemplarily the divergence-gradient operators $y \mapsto -\operatorname{div}(\mu \nabla y)$, complemented with homogeneous Dirichlet boundary conditions, in their weak form given by

$$\langle Ay, \varphi \rangle = \int_\Omega (\mu \nabla y) \cdot \nabla \varphi \,\mathrm{d}x \quad \text{for } \varphi \in H_0^1(\Omega),$$

where $\mu \in L^\infty(\Omega; \mathbb{S}_n)$ takes its values in the space $\mathbb{S}_n$ of symmetric $n \times n$-matrices and satisfies the *coercivity-* or *ellipticity condition*: There exists $\mu_0 > 0$ such that

$$v^T \mu(x) v \geq \mu_0 \|v\|_2^2 \quad \text{for all } v \in \mathbb{R}^n \quad \text{for almost all } x \in \Omega.$$

(Compare also with the weak form of the negative Laplacian $-\Delta$, so the divergence-gradient operator with $\mu$ being the $n \times n$-identity matrix, in Example 2.8.) See the exercises.

Since the underlying partial differential operators are of order two, the assumption $A^{-1} \in \mathcal{L}(H^{-1}(\Omega); H_0^1(\Omega))$ can be seen as a *maximal Sobolev regularity* result in the sense that $A$ and $A^{-1}$ operate exactly between Sobolev spaces with a gap of differentiability of order two, namely $H_0^1(\Omega)$ and $H^{-1}(\Omega)$.

The additional assumption on $A$ then requires the following: If the right-hand side $f$ in the equation $Ay = f$ is actually from the better space $L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$, then this additional $L^2$-regularity for the data implies additional $H^2$-regularity for the state $y$, so $y \in H_0^1(\Omega) \cap H^2(\Omega)$ with a continuous dependence of $y$ on $f$, in the sense of the existence of a constant $C > 0$ independent of $f$ such that

$$\|y\|_{H_0^1(\Omega)} + \|y\|_{H^2(\Omega)} \le C\|f\|_{L^2(\Omega)}.$$

Of course, this additional assumption is again a maximal Sobolev regularity property, but unfortunately, proving such a property is quite difficult, even for the Laplacian. It can be shown, e.g., if $\Omega$ is of class $C^{1,1}$ and the coefficients $\mu_{ij}$ of the operator $A$ are uniformly continuous on $\overline{\Omega}$, cf. [GT01, Thm. 9.15].

**Remark 3.41.** In the foregoing Example 3.39, consider again the KKT condition or *adjoint equation*

$$A^*\bar{p} + \mathcal{E}^*\bar{\mu} = -J_y'(\bar{y}, \bar{u}).$$

The operator $A^* \in \mathcal{L}(H_0^1(\Omega); H^{-1}(\Omega)) \cap \mathcal{L}(L^2(\Omega); Y^*)$ can also be considered as an elliptic partial differential operator. (Also, since $A$ is continuously invertible, so is $A^*$.) Thus, the foregoing adjoint equation is also an elliptic PDE for the *adjoint state* $\bar{p}$ and one can make use of smoothing effects of a PDE solution operator.

For the right hand side in the adjoint equation, we only have $J_y'(\bar{y}, \bar{u}) \in Y^*$ in general. But if in fact $J$ is already F-differentiable from $Y_0 \times U$ to $\mathbb{R}$, where $Y$ is densely embedded into $Y_0$—for instance, if $Y_0 = H_0^1(\Omega)$ or $Y_0 = L^2(\Omega)$—, then $J_y'(\bar{y}, \bar{u}) \in Y_0^* \hookrightarrow Y^*$ and additional regularity for the adjoint state $\bar{p}$ may be derived from the adjoint equation. (In the present case, the regularity of $\bar{p}$ is however limited by the presence of the measure $\bar{\mu}$, so we will not be able to obtain full regularity in general.)

Note that a similar effect can also occur for the optimal control $\bar{u}$ from the KKT condition $J_u'(\bar{y}, \bar{u}) = B^*\bar{p}$, since often $J_u'(\bar{y}, \bar{u}) = \beta\bar{u}$ for some $\beta > 0$. (See e.g. Example 2.8.) Then

$$J_u'(\bar{y}, \bar{u}) = B^*\bar{p} \quad \Longleftrightarrow \quad \bar{u} = \frac{1}{\beta}B^*\bar{p}.$$

See the exercises.

## 3.3 Sufficient optimality conditions

The KKT conditions give a satisfying characterization of (first-order) *necessary* optimality conditions. Of course, we are also interested in *sufficient* conditions. Such conditions are of particular interest in view of numerical algorithms. Also, they often include that the optimal point in question is an isolated optimum, which is an extremely useful property in the numerical analysis of optimal control problems.

We will however see that the situation is quite more delicate than in the finite-dimensional case and that the standard second-order sufficient condition from Nonlinear Optimization will in general not be sufficient in infinite-dimensional settings any more. We begin with the particular and important case of a *convex* problem, where everything still works out just fine.

### 3.3.1 The convex case

In the convex case, the KKT conditions alone already turn out to be also sufficient for global optimality:

**Theorem 3.42.** *Let $X$ and $Z$ be Banach spaces. Suppose that $f\colon X \to \mathbb{R}$ is convex and F-differentiable, $G\colon X \to Z$ is F-differentiable and convex w.r.t. $-K$, where $K$ is a closed convex cone. Further, let $\bar{x} \in \mathcal{F} = G^{-1}[K]$ be a KKT-point of (P), so $\Lambda(\bar{x}) \neq \emptyset$. Then $\bar{x}$ is a global solution of (P).*

*Proof.* Recall from (3.19) that since $f\colon X \to \mathbb{R}$ and $G\colon X \to Z$ are convex—the latter w.r.t. $\leq_K$—, we have the inequality

$$f(x) - f(\bar{x}) \geq \langle f'(\bar{x}), x - \bar{x}\rangle_{X^*,X} \quad \text{for all } x \in X$$

and the inclusion

$$G(\bar{x}) - G(x) + G'(\bar{x})(x - \bar{x}) \in K \quad \text{for all } x \in \mathcal{F}$$

at hand. Now, for $x \in \mathcal{F}$, we have $K + G(x) \subset K$ since $K$ is a *convex* cone, and hence

$$G'(\bar{x})(x-\bar{x}) \in \big(K + G(x)\big) - G(\bar{x}) \subset K - G(\bar{x}) \subset \operatorname{cone}(K, G(\bar{x})) \subset T(K, G(\bar{x}))$$

for all $x \in \mathcal{F}$. (Recall Lemma 3.8.) But then optimality of $\bar{x}$ is immediate from the KKT conditions (3.15) and (3.16) for all $x \in \mathcal{F}$ as follows:

$$\begin{aligned} f(x) - f(\bar{x}) \geq \langle f'(\bar{x}), x - \bar{x}\rangle_{X^*,X} &= -\big\langle G'(\bar{x})^*\bar{\lambda}, x - \bar{x}\big\rangle_{X^*,X} \\ &= -\big\langle \bar{\lambda}, G'(\bar{x})(x-\bar{x})\big\rangle_{Z^*,Z} \geq 0, \end{aligned}$$

since $G'(\bar{x})(x - \bar{x}) \in T(K, G(\bar{x}))$ and $\bar{\lambda} \in T(K, G(\bar{x}))^\circ$ as seen above. □

**Remark 3.43.** Note that we have only supposed $\Lambda(\bar{x}) \neq \emptyset$ in Theorem 3.42 instead of $\bar{x}$ regular, i.e., we have *not* needed a constraint qualification.

### 3.3.2 Second-order sufficient conditions

We next plan to use second-order information to augment the KKT-conditions such that they become sufficient, also in the case of a *nonconvex* problem.

**Reminder:** The second derivative $h''(\bar{x})$ of a twice F-differentiable function $h\colon X \to Z$ is given by the F-derivative of the mapping $h'\colon X \to \mathcal{L}(X;Z)$, and is thus a mapping

$$h''\colon X \to \mathcal{L}\big(X;\mathcal{L}(X;Z)\big) \cong \mathcal{L}^2(X \times X; Z),$$

where $\mathcal{L}^2(X \times X; Z)$ denotes the space of continuous bilinear forms on $X \times X$ mapping into $Z$.

As a first step, we define the *critical cone* as a subset of the linearizing cone.

**Definition 3.44** (Critical cone)**.** The *critical cone* at $\bar{x} \in \mathcal{F}$ is defined by

$$C(\bar{x}) = \Big\{ d \in T_\ell(G,K,\bar{x})\colon \langle f'(\bar{x}), d\rangle_{X^*,X} \le 0 \Big\}.$$

Suppose that $\bar{x} \in \mathcal{F}$ is a KKT-point of $f$ on $\mathcal{F}$, so $\Lambda(\bar{x}) \neq \emptyset$. Then we have already seen in Lemma 3.29 that there are no directions $d \in C(\bar{x})$ with $\langle f'(\bar{x}), d\rangle_{X^*,X} < 0$. Accordingly, in this case, the critical cone becomes

$$C(\bar{x}) = \Big\{ d \in T_\ell(G,K,\bar{x})\colon \langle f'(\bar{x}), d\rangle_{X^*,X} = 0 \Big\},$$

and from the KKT condition (3.9) we have

$$\big\langle G'(\bar{x})^*\bar{\lambda}, d\big\rangle_{X^*,X} = \big\langle \bar{\lambda}, G'(\bar{x})^* d\big\rangle_{Z^*,Z} = 0 \quad \text{for all } \bar{\lambda} \in \Lambda(\bar{x}),\ d \in C(\bar{x}).$$

Recall that the first part (3.9) of the KKT conditions could be rewritten to $L'_x(\bar{x},\bar{\lambda}) = 0$ in $X^*$. Now, in order to formulate a sufficient condition for a given KKT point $\bar{x}$ to be a local solution, so a *minimizer*, it is natural—from variational principles, see [BS00, Ch. 3.1.1]—to require that there are no directions of nonpositive curvature w.r.t. $x$ of the Lagrangian function in $C(\bar{x})$; that is, there are no directions $d \in C(\bar{x}) \setminus \{0\}$ with $L''_{xx}(\bar{x},\bar{\lambda})(d,d) \le 0$:

$$L''_{xx}(\bar{x},\bar{\lambda})(d,d) > 0 \quad \text{for all } d \in C(\bar{x}) \setminus \{0\}. \tag{3.22}$$

This is the classical second-order sufficient condition (SOSC) from Nonlinear Optimization, where we also have seen that posing the curvature condition solely on $f$ will not work even there [UU12, P. 103]. Unfortunately, the next example shows that the classical SOSC (3.22) is *not* strong enough in infinite dimensions to indeed obtain a sufficient optimality condition.

**Example 3.45.** Let $X = Z = \ell^2$, where $\ell^2 = \ell^2(\mathbb{R})$ is the Hilbert space of $\mathbb{R}$-valued square summable sequences whose inner product and norm are given by

$$(x,y)_{\ell^2} := \sum_{i=1}^{\infty} x_i y_i, \quad \text{so} \quad \|x\|_{\ell^2} = \left(\sum_{i=1}^{\infty} x_i^2\right)^{1/2}.$$

The problem under consideration is

$$\min_{x\in\ell^2} (c,x)_{\ell^2} - (x,x)_{\ell^2} \quad \text{s.t.} \quad x_i \geq 0 \text{ for all } i \in \mathbb{N}$$

with the objective function $f(x) = (c,x)_{\ell^2} - (x,x)_{\ell^2}$, where $c \in \ell^2$ satisfies $c_i > 0$ for all $i \in \mathbb{N}$. The constraints are given by $G(x) \in K$ with $K = \{x \in \ell^2 \colon x_i \geq 0 \text{ for all } i \in \mathbb{N}\}$ and $G(x) = x$, the identity mapping.

Since $G'(x) = \mathrm{id}_{\ell^2}$ is clearly surjective for all $x \in \ell^2$, RCQ holds in every feasible $x \in K$, see Proposition 3.18.

We consider $\bar{x} = 0$ and claim that it is a KKT point with $-c$ being the unique Lagrange multiplier, so $\{-c\} = \Lambda(\bar{x}) \neq \emptyset$. First, concerning the multiplier rule (3.9) in the KKT conditions, observe that $f'(\bar{x}) \in (\ell^2)^*$ can be identified with $c - 2\bar{x} \in \ell^2$. (This is the Fréchet-Riesz representation theorem; in fact, we can regard $c - 2\bar{x}$ as the gradient $\nabla f(\bar{x})$.) Hence, for $\bar{\lambda}$ to be a Lagrange multiplier for $\bar{x} = 0$, we calculate

$$f'(\bar{x}) + G'(\bar{x})^*\bar{\lambda} = c - 2\bar{x} + \bar{\lambda} = c + \bar{\lambda} \overset{!}{=} 0 \quad \Longleftrightarrow \quad \bar{\lambda} = -c,$$

so $-c$ is the unique candidate for an element of $\Lambda(\bar{x})$. Moreover, thanks to RCQ, we have (Theorem 3.21)

$$T_\ell(G,K,\bar{x}) = T(K,G(\bar{x})) = T(K,0) = \overline{\mathrm{cone}(K,0)} = K.$$

Clearly, $(c,d)_{\ell^2} \geq 0$ for all $d \in K$, so we obtain $-c \in K^\circ = T(K,G(\bar{x}))^\circ$. Hence, in fact $\{-c\} = \Lambda(\bar{x}) \neq \emptyset$.

Next we investigate the proposed sufficient condition. The critical cone in $\bar{x}$ is given by

$$C(\bar{x}) = \left\{d \in K \colon (f'(\bar{x}),d)_{\ell^2} = 0\right\},$$

but for all $d \in K \setminus \{0\}$, we obtain

$$(f'(\bar{x}),d)_{\ell^2} = (c,d)_{\ell^2} = \sum_{i=1}^{\infty} c_i d_i > 0,$$

since $c_i > 0$ and $0 \neq d \geq 0$. Therefore, $C(\bar{x}) = \{0\}$, so $C(\bar{x}) \setminus \{0\} = \emptyset$, and the classical SOSC (3.22) is indeed satisfied although

$$L''_{xx}(\bar{x}, \bar{\lambda})(d,d) = f''(\bar{x})(d,d) = -2\|d\|^2_{\ell^2} \leq 0 \quad \text{for all } d \in \ell^2$$

(verify this!). In fact, we have even seen that all directions $d \in K \setminus \{0\} = T(K, G(\bar{x})) \setminus \{0\}$ are *ascent directions* (`Aufstiegsrichtungen`) for $f$!

But nevertheless, $\bar{x}$ is not a local minimum of $f$ on $\mathcal{F} = K$: In fact, define a sequence $(x^k) \subset \ell^2$ by $(x^k) := (2\delta_{ik} c_i)_{i \in \mathbb{N}} \subset K$, so

$$(x^k)_i := \begin{cases} 2c_i & \text{if } i = k, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\|x^k - \bar{x}\|_{\ell^2} = \|x^k\|_{\ell^2} = 2c_k \to 0$ as $k \to \infty$, but

$$f(x^k) = 2c_k^2 - 4c_k^2 = -2c_k^2 < 0 = f(\bar{x}),$$

so $\bar{x}$ cannot be a local minimum of $f$ on $K$.

In the foregoing example, the condition (3.22) posed is not sufficient for optimality of a KKT point because the critical cone $C(\bar{x})$ is too small. Thus, a possible (partial) solution to obtain a sufficient condition is to increase the critical cone $C(\bar{x})$ as follows.

**Definition 3.46** (Approximate critical cone)**.** For $\eta \geq 0$ we define the *$\eta$-approximate critical cone* at $\bar{x} \in \mathcal{F}$ by

$$C_\eta(\bar{x}) = \Big\{ d \in T_\ell(G, K, \bar{x}) \colon \langle f'(\bar{x}), d \rangle_{X^*, X} \leq \eta \|d\|_X \Big\}.$$

Note that $C_0(\bar{x}) = C(\bar{x})$ and $C_\eta(\bar{x}) = T_\ell(G, K, \bar{x})$ for all $\eta \geq \|f'(\bar{x})\|_{X^*}$.

We will replace (3.22) with a condition involving $C_\eta(\bar{x})$ for $\eta > 0$. For the proof that such a condition is indeed sufficient for optimality, an auxiliary result will be needed: under the Robinson CQ, the linearizing cone $T_\ell(G, K, \bar{x})$ at $\bar{x}$ is approximated by feasible directions in the sense that

$$\operatorname{dist}\big(x - \bar{x}, T_\ell(G, K, \bar{x})\big) = o\big(\|x - \bar{x}\|_X\big) \tag{3.23}$$

for $\mathcal{F} \ni x \to \bar{x}$. This already foreshadows that we will need to suppose that the KKT-point $\bar{x}$ is in fact regular.

In fact, the approximation property (3.23) is implied by the following lemma (see the exercises):

**Lemma 3.47.** *If the Robinson constraint qualification* (3.2) *holds at* $\bar{x} \in \mathcal{F}$, *then there exists a map* $h\colon \mathcal{F} \to T_\ell(G, K, \bar{x})$ *with*

$$\|h(x) - (x - \bar{x})\|_X = o\big(\|x - \bar{x}\|_X\big) \quad \textit{for } \mathcal{F} \ni x \to \bar{x}.$$

*Proof.* Let $x \in \mathcal{F}$ be arbitrary. Then the F-differentiability of $G$ implies

$$G(x) = G(\bar{x}) + G'(\bar{x})(x - \bar{x}) + r(x), \quad \text{where} \quad \|r(x)\|_Z = o\big(\|x - \bar{x}\|_X\big).$$

Then RCQ in the form of the ZKCQ (3.8) shows that there exists $\delta > 0$ such that

$$\overline{B_{\delta,Z}(0)} \subset G'(\bar{x})\overline{B_X(0)} - \big((K - G(\bar{x})) \cap \overline{B_Z(0)}\big).$$

Hence, we can find $s(x) \in X$ and $v(x) \in K - G(\bar{x})$ with

$$\|s(x)\|_X \le \frac{\|r(x)\|_Z}{\delta} = o\big(\|x - \bar{x}\|_X\big), \quad \|v(x)\|_Z \le \frac{\|r(x)\|_Z}{\delta} = o\big(\|x - \bar{x}\|_X\big)$$

satisfying $r(x) = G'(\bar{x})s(x) - v(x)$. Setting $h(x) := x - \bar{x} + s(x)$, there holds

$$\|h(x) - (x - \bar{x})\|_X = o\big(\|x - \bar{x}\|_X\big).$$

It remains to show that $h(x) \in T_\ell(G, K, \bar{x})$, that is, $G'(\bar{x})h(x) \in T(K, G(\bar{x}))$. So:

$$\begin{aligned} G'(\bar{x})h(x) &= G'(\bar{x})(x - \bar{x}) + G'(\bar{x})s(x) \\ &= G'(\bar{x})(x - \bar{x}) + r(x) + v(x) \\ &= G(x) - G(\bar{x}) + v(x) \end{aligned}$$

Writing $v(x) \in K - G(\bar{x})$ in the form $v(x) = k(x) - G(\bar{x})$ with $k(x) \in K$, we have

$$G'(\bar{x})h(x) = G(x) - G(\bar{x}) + k(x) - G(\bar{x}) = 2\Bigg(\underbrace{\frac{G(x) + k(x)}{2}}_{\in K} - G(\bar{x})\Bigg) \in \operatorname{cone}(K, G(\bar{x})).$$

Hence, $G'(\bar{x})h(x) \in \operatorname{cone}(K, G(\bar{x})) \subset T(K, G(\bar{x}))$ and so $h(x) \in T_\ell(G, K, \bar{x})$, and $h$ has all required properties. □

**Remark 3.48.** It is also possible to prove the approximation property (3.23) directly via metric regularity, Theorem 3.19, and to construct the function $h$ as in Lemma 3.47 from there. See the exercises.

We now prove a very general theorem about second-order sufficient conditions. The proof rests on Taylor expansion for the Lagrange function and Lemma 3.47.

**Theorem 3.49** (Second-order sufficient conditions). *Let $X$ and $Z$ be Banach spaces with $K \subset Z$ closed and convex. Further, let $f: X \to \mathbb{R}$ and $G: X \to Z$ be twice F-differentiable. Assume that $\bar{x} \in \mathcal{F} = G^{-1}[K]$ satisfies the RCQ (3.2) and the KKT-conditions with multiplier $\bar{\lambda}$:*

$$\begin{aligned} &f'(\bar{x}) + G'(\bar{x})^*\bar{\lambda} = 0, \\ &G(\bar{x}) \in K, \quad \bar{\lambda} \in T(K, G(\bar{x}))^\circ. \end{aligned}$$

*Let in addition the following second-order condition hold:*

$$L''_{xx}(\bar{x}, \bar{\lambda})(d, d) \geq \gamma \|d\|_X^2 \quad \text{for all } d \in C_\eta(\bar{x}), \tag{3.24}$$

*where $\gamma > 0$ and $\eta > 0$ are fixed constants. Then $\bar{x}$ is an isolated local solution of (P) and there exist $\kappa, \delta > 0$ such that the* quadratic growth condition

$$f(x) \geq f(\bar{x}) + \kappa \|x - \bar{x}\|_X^2 \quad \text{for all } x \in \mathcal{F} \cap B_{\delta,X}(\bar{x})$$

*holds true.*

*Proof.* Consider $x \in \mathcal{F} \cap B_{\delta,X}(\bar{x})$ for some $\delta > 0$ sufficiently small to be chosen later, and set $d(x) := x - \bar{x}$. Note that we do not know whether $d(x) \in T_\ell(K, G, \bar{x})$, so (3.24) cannot be used for $d(x)$ directly. However, Lemma 3.47 braces us with the approximation $d(x) = h(x) + r(x)$ where $h(x) \in T_\ell(K, G, \bar{x})$ and $\|r(x)\|_X = o(\|d\|_X)$. The plan is to use $h(x)$ as a surrogate for $d(x)$ in $T_\ell(G, K, \bar{x})$. To this end, we first establish some approximation properties of $h(x)$, starting with

$$\|h(x)\|_X = \|d(x)\|_X + o(\|d(x)\|_X). \tag{3.25}$$

We use both triangle inequalities to estimate $\|h(x)\|_X = \|d(x) - r(x)\|_X$ from above and below and observe that $\|r(x)\|_X \leq \|d(x)\|_X$ if $\delta$ is small enough, since $\|r(x)\|_X = o(\|d(x)\|_X)$, to obtain

$$-\|r(x)\|_X \leq \|h(x)\|_X - \|d(x)\|_X \leq \|r(x)\|_X$$

which clearly implies (3.25). We also find the quadratic equivalent

$$\|h(x)\|_X^2 = \|d(x)\|_X^2 + o\big(\|d(x)\|_X^2\big) \tag{3.26}$$

of (3.25) to be true by

$$\begin{aligned} \|h(x)\|_X^2 &= \big(\|h(x)\|_X - \|d(x)\|_X + \|d(x)\|_X\big)^2 \\ &= \underbrace{\big(\|h(x)\|_X - \|d(x)\|_X\big)^2}_{=o(\|d\|_X^2)} + \underbrace{2\|d(x)\|_X\big(\|h(x)\|_X - \|d(x)\|_X\big)}_{=o(\|d\|_X^2)} + \|d(x)\|_X^2. \end{aligned}$$

Now, back in the actual proof, the case where $\langle f'(\bar{x}), h(x)\rangle_{X^*,X} > \eta\|h(x)\|_X$ is quite easy: Using $\langle f'(\bar{x}), r(x)\rangle_{X^*,X} = o(\|d(x)\|_X)$, we find

$$\begin{aligned} f(x) - f(\bar{x}) &= \langle f'(\bar{x}), d(x)\rangle_{X^*,X} + o(\|d(x)\|_X) \\ &= \langle f'(\bar{x}), h(x)\rangle_{X^*,X} + o(\|d(x)\|_X) \\ &> \eta\|h(x)\|_X + o(\|d(x)\|_X) \\ &\overset{(3.25)}{=} \eta\|d(x)\|_X + o(\|d(x)\|_X) \geq \eta\|d(x)\|_X^2 = \eta\|x - \bar{x}\|_X^2, \end{aligned}$$

where the last inequality holds true for $\delta$ sufficiently small.

Next we deal with the case $\langle f'(\bar{x}), h(x)\rangle_{X^*,X} \leq \eta\|h(x)\|_X$, so $h(x) \in C_\eta(\bar{x})$. From the KKT conditions, we have $\bar{\lambda} \in T(K, G(\bar{x}))^\circ$, so

$$L(x, \bar{\lambda}) - L(\bar{x}, \bar{\lambda}) = f(x) - f(\bar{x}) + \langle \bar{\lambda}, G(x) - G(\bar{x})\rangle_{Z^*,Z} \leq f(x) - f(\bar{x}), \tag{3.27}$$

since $G(x) - G(\bar{x}) \in T(K, G(\bar{x}))$. Moreover, by (3.25) and the construction of $r(x)$, we have

$$L''_{xx}(\bar{x}, \bar{\lambda})(r(x), h(x)) + L''_{xx}(\bar{x}, \bar{\lambda})(h(x), r(x)) + L''_{xx}(\bar{x}, \bar{\lambda})(r(x), r(x)) = o(\|d\|_X^2). \tag{3.28}$$

Recalling that the first KKT condition means in fact $L'_x(\bar{x}, \bar{\lambda}) = 0$ in $X^*$, we thus obtain by Taylor expansion

$$\begin{aligned} f(x) - f(\bar{x}) &\overset{(3.27)}{\geq} L(x, \bar{\lambda}) - L(\bar{x}, \bar{\lambda}) \\ &= \langle L'_x(\bar{x}, \bar{\lambda}), d(x)\rangle_{X^*,X} + \frac{1}{2}L''_{xx}(\bar{x}, \bar{\lambda})(d(x), d(x)) + o(\|d(x)\|_X^2) \\ &= \frac{1}{2}L''_{xx}(\bar{x}, \bar{\lambda})(h(x) + r(x), h(x) + r(x)) + o(\|d(x)\|_X^2) \\ &\overset{(3.28)}{=} \frac{1}{2}L''_{xx}(\bar{x}, \bar{\lambda})(h(x), h(x)) + o(\|d(x)\|_X^2) \\ &\geq \frac{\gamma}{2}\|h(x)\|_X^2 + o(\|d(x)\|_X^2) \\ &= \frac{\gamma}{2}\|d(x)\|_X^2 + o(\|d(x)\|_X^2) \geq \frac{\gamma}{4}\|d(x)\|_X^2 = \frac{\gamma}{4}\|x - \bar{x}\|_X^2, \end{aligned}$$

the last inequality again for $\delta$ sufficiently small.

The assertion then follows with $\kappa := \min(\frac{\gamma}{4}, \eta)$. □

Theorem 3.49 works in a very general setting, which is a fortune, but also an obstacle to obtain sharper results. Especially the uniform coercivity condition (3.24) is quite harsh. On the other hand, we have posed no further assumptions on $f$ and $G$ at all, so there might be room to strengthen assumptions elsewhere and weaken condition (3.24). Indeed, it is possible to derive second-order sufficient conditions in the classical form (3.22) under more specific assumptions on $f$ and $G$:

**Theorem 3.50.** *Let $X$ and $Z$ be Banach spaces with $K \subset Z$ closed and convex and $X$ reflexive. Further, let $f\colon X \to \mathbb{R}$ be given in the form $f = f_1 + f_2$ with $f_1, f_2\colon X \to \mathbb{R}$, and let $f_1, f_2$ and $G\colon X \to Z$ be twice F-differentiable. Assume that $\bar{x} \in \mathcal{F} = G^{-1}[K]$ is a KKT-point, so $\Lambda(\bar{x}) \neq \emptyset$, and that the second derivatives of $f_1$ and $G$ exhibit the following weak continuity properties: If $d_k \rightharpoonup d$ in $X$, then*

$$f_1''(\bar{x})(d_k, d_k) \to f_1''(\bar{x})d \quad \text{in } \mathbb{R} \quad \text{and} \quad G''(\bar{x})(d_k, d_k) \rightharpoonup G''(\bar{x})d \quad \text{in } Z.$$

*Let in addition $f_2''(\bar{x})$ be coercive, so there exists $\alpha > 0$ such that*

$$f_2''(\bar{x})(d, d) \geq \alpha \|d\|_X^2 \quad \text{for all } d \in X,$$

*and let the following second-order condition hold:*

$$L''_{xx}(\bar{x}, \bar{\lambda})(d, d) > 0 \quad \text{for all } d \in C(\bar{x}) \setminus \{0\},$$

*where $\gamma > 0$ is a fixed constant. Then $\bar{x}$ is an isolated local solution of* (P) *and there exist $\kappa, \delta > 0$ such that the* quadratic growth condition

$$f(x) \geq f(\bar{x}) + \kappa \|x - \bar{x}\|_X^2 \quad \text{for all } x \in \mathcal{F} \cap B_{\delta,X}(\bar{x})$$

*holds true.*

*Proof.* Without proof for the moment. □

**Remark 3.51.**

(a) Weak continuity properties of $f_1''(\bar{x})$ and $G''(\bar{x})$ as required in Theorem 3.50 are surprisingly common and will usually originate from some underlying compactness properties of the operators, cf. Lemma 2.5.

(b) The assumption on $f_2$ is also quite common, in particular in optimal control problems, where classically we have $f(x) = J(y, u) = f_1(y) + \frac{\beta}{2}\|u\|^2_{L^2(\Omega)}$ or comparable functions. (Such an example was also in the exercises.)

Unfortunately, both Theorems 3.49 and 3.50 suffer from a generic problem: There is a quite simple class of optimization problems for which the objective functional $f$ is twice F-differentiable precisely on a given space $X$, but $f''(\bar{x})$ in a locally optimal point $\bar{x}$ does *not* admit a coercivity estimate of the form $f''(\bar{x})(d, d) \geq \gamma \|d\|_X^2$ for $d \in C_\eta(\bar{x})$. The most prominent example is the (unconstrained) problem

$$\min_{x \in X} - \int_0^1 \cos(x(t))\, \mathrm{d}t$$

for $X = L^\infty(0,1)$ (see the exercises) and the optimal point $\bar{x} \equiv 0$. Here, the lack of coercivity of $f''(\bar{x})$ on $X$ means that the second-order condition in Theorem 3.49 is never satisfied *even though the constraint $G(x) \in K$ is trivial or void*! (One can show argue analogously as to why the assumptions on $f$ in Theorem 3.50 prevent its applicability.)

Luckily, one often observes that there is another Banach space $X^+$ with $X \hookrightarrow X^+$ densely on which $f$ may not be twice F-differentiable, but to which we might extend $f''(\bar{x})$, and this extension admits the needed coercivity estimate $f''(\bar{x})(d,d) \geq \gamma \|d\|^2_{X^+}$ for $d \in X^+$. (In the example above it would be $X^+ = L^2(0,1)$.) Under additional assumptions, it is possible -to adapt Theorem 3.50 to this situation and to show that $\bar{x}$ is still an isolated local solution of (P) in this case. The price to pay is that one only obtains the weaker quadratic growth condition

$$f(x) \geq f(\bar{x}) + \kappa \|x - \bar{x}\|^2_{X^+} \quad \text{for all } x \in \mathcal{F} \cap B_{\delta,X}(\bar{x}),$$

which however is often sufficient for further purposes of this property. This whole phenomenon is known as the *two norm gap*; we refer to [CT15] for a survey and many further results.

# References


[Br10] H. Brezis: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer, New York, 2010.

[BS00] J.F. Bonnans, A. Shapiro: Perturbation Analysis of Optimization Problems. Springer Science + Business Media, 2000.

[CT15] E. Casas, F. Tröltzsch: Second Order Optimality Conditions and Their Role in PDE Control, *Jahresber. Deutsch. Math. Verein.* 117 no. 1, 3–44, 2015.

[GT01] D. Gilbarg and N.S. Trudinger: Elliptic Partial Differential Equations of Second Order. Springer Science + Business Media, 2001.

[HPUU09] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich: Optimization with PDE Constraints. Springer Netherlands, 2009.

[La93] S. Lang: Real and Functional Analysis. Springer, New York, 1993.

[NW06] J. Nocedal, S. Wright: Numerical Optimization. Springer, New York, 2006.

[Ro72] S.M Robinson: Normed convex processes, *Trans. Amer. Math. Soc.* 174, 127–140, 1972.

[Ro76] S.M Robinson: Stability theory for systems of inequalities in nonlinear programming, part II: differentiable nonlinear systems. *SIAM J. Num. Anal.* 13, 497–513, 1976.

[Ro76b] S.M Robinson: Regularity and stability for convex multivalued functions, *Math. Oper. Res.* 1, 130–143, 1976.

[Sc07] W. Schirotzek: Nonsmooth analysis. Springer, Berlin Heidelberg, 2007.

[UU12] M. Ulbrich, S. Ulbrich: Nichtlineare Optimierung. Birkhäuser, Basel, 2012.

[ZK79] J. Zowe, S. Kurcyusz: Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optimization* 5, 49–62, 1979.