An augmented Lagrangian algorithm for optimization with equality constraints in Hilbert spaces

Jan H. Maruhn

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Mathematics

Dr. Ekkehard W. Sachs
Dr. Belinda B. King
Dr. Robert C. Rogers
Dr. James R. Lang

April 30, 2001
Blacksburg, Virginia

Keywords: Nonlinear optimization, equality constraints, optimization in Hilbert spaces, augmented Lagrangian methods, discrete approximations

An augmented Lagrangian algorithm for optimization with equality constraints in Hilbert spaces

Jan H. Maruhn

(ABSTRACT)

Since augmented Lagrangian methods were introduced by Powell and Hestenes, this class of methods has been investigated very intensively. While the finite dimensional case has been treated in a satisfactory manner, the infinite dimensional case is studied much less.

The general approach to solve an infinite dimensional optimization problem subject to equality constraints is as follows: First one proves convergence for a basic algorithm in the Hilbert space setting. Then one discretizes the given spaces and operators in order to make numerical computations possible. Finally, one constructs a discretized version of the infinite dimensional method and tries to transfer the convergence results to the finite dimensional version of the basic algorithm.

In this thesis we discuss a globally convergent augmented Lagrangian algorithm and discretize it in terms of functional analytic restriction operators. Given this setting, we prove global convergence of the discretized version of this algorithm to a stationary point of the infinite dimensional optimization problem. The proposed algorithm includes an explicit rule of how to update the discretization level and the penalty parameter from one iteration to the next one - questions that had been unanswered so far. In particular the latter update rule guarantees that the penalty parameters stay bounded away from zero which prevents the Hessian of the discretized augmented Lagrangian functional from becoming more and more ill conditioned.

Acknowledgments

I wish to express my gratitude to Professor Dr. Ekkehard W. Sachs for suggesting the problem and for being my supervisor. Further I would like to thank Professor Dr. Belinda B. King for presenting various applications of the abstract theory in this thesis to me. I am also deeply indebted to Professor Dr. Robert C. Rogers for various stimulating discussions after his functional analysis lectures. Moreover, I want to thank Professor Dr. James R. Lang for introducing me to the way business is taught in the United States - maybe one of the reasons for the strong U.S. economy.

Further I gratefully acknowledge the financial support of the Mathematics Department of Virginia Polytechnic Institute and State University, without which this research and my whole stay in the United States would not have been possible. Finally, I would like to thank the German National Scholarship Foundation for continuous backing.

Contents

List of Figures

CHAPTER 1

Introduction and Motivation

Having its roots in the calculus of variations and the theory of differential equations, today control theory is a well established field of mathematics in its own right. Historically, the first control problems were studied in a physical or engineering context. Besides those classical applications of optimal control, the theory has constantly gained importance in economics and finance (see for example [20]).

During the seventies a number of economists showed the need to formulate macroeconomic questions in a dynamic and stochastic setup. Agents in an economy (consumers, firms or even governments) can be viewed to make their decisions by optimizing specific objective functions, taking into account that the consequences of their actions

- are uncertain a priori
- will be noticed through a number of periods
- influence some other variables, that will feedback into the economy

As shown in [24], complex stochastic models can simulate whole economies and thus they can be used to study effects of certain controls imposed by the government like taxes or pollution charges. In fact these models are used as decision making tools of governmental institutions in most industrialized countries. An example is the Federal Reserve Banks' Board of Governors' large-scale rational expectations macroeconomic model of the U.S. economy (FRB/US) which is used in [10] to solve for optimal control rules. But also smaller countries simulate their economy in order to predict effects of governmental policy. For example the Forecasting and Policy System Model (FPS) of the Reserve Bank of New Zealand is used in [9] to examine the implication of uncertainty about potential output for simple efficient policy rules.

We will now give a brief summary of some of the features of the FRB/US model as presented in [7] and [5].

1.1 The FRB/US model

The FRB/US model of the U.S. economy is maintained at the Federal Reserve Board for use in policy analysis and forecasting. With FRB/US, the Board's staff can gauge the likely consequences of specific events through simulation analysis - computational "what-if" exercises in which the model is used to predict the outcomes from alternative assumptions regarding fiscal and monetary policy, international conditions, and so forth. In a similar manner, the staff can use model simulations to assess possible implications for economic performance of the full range of disturbances likely to be experienced over extended periods of time.

The equations of FRB/US are specified in accordance with standard economic theory. In particular, households, businesses, and investors are assumed to be forward-looking in their decision making as they seek to optimize their welfare:

Individuals choose a path for current and future consumption that maximizes their lifetime utility, subject to a budget constraint. This assumption implies that consumer spending today is related to the present value of expected future earnings and the current value of assets. Similarly, firms maximize expected profits in hiring workers, investing in capital goods, and setting prices. In financial markets, investors equate expected rates of return on different assets, subject to premiums that compensate borrowers and lenders for differences in risk and liquidity.

FRB/US is what is often called a New Keynesian model because of the assumptions it incorporates. As mentioned before, households and firms are forward-looking - that is, they base their decisions on the income and sales, financial conditions, and prices that they expect for the future. However, rather than being instantaneous, the response to changes in these fundamental factors is gradual because capital installation costs, contracts, and other considerations create significant frictions that slow the process. For this reason, the failure of markets to clear quickly after disturbances to the economy can result in periods of over- or under-utilization of labor and capital resources.

According to the viewpoint embedded in the model, monetary policy can mitigate these swings in aggregate resource utilization by altering financial market conditions and thereby exerting an indirect influence on output and employment in the short term and on inflation over the longer term. In FRB/US, policymakers alter financial conditions by changing the short-term interest rate under the control of the Federal Reserve: the federal funds rate. Current and anticipated changes in this rate influence prices and rates of return on various financial assets, including bonds and corporate equities, and on foreign exchange. Changes in these financial conditions in turn influence spending by households and firms and, by altering resource utilization in labor and product markets, affect the rate of inflation.

Due to the complexity of the processes and interactions in the economy, FRB/US is a large-scale model, containing some 300 equations and identities. However, the number of stochastic "core" equations or estimated descriptions of the economic behavior of firms, households, and investors is much smaller, around 50 equations.

Clearly, a presentation of such a model is beyond the scope of an introduction. In order to illustrate the concept of economic modeling, we rather introduce a simple model used to determine an optimal tax policy for a government - optimal in the sense of maximizing social welfare subject to budget constraints. More precisely, we introduce a model for optimal redistributive capital taxation in a neoclassical growth model as stated in [17].

1.2 A Model for Optimal Capital Taxation

The model economy consists of a government, identical competitive firms, and two types of infinitely-lived, price taking agents called workers and capitalists. For simplicity, the number of each type in the population is normalized to one. Further the model abstracts from uncertainty, technological progress, and population growth. Workers supply one unit of labor inelastically and incur a transaction cost for saving or borrowing small amounts. As a result, all wealth is concentrated in the hands of the capitalists, who do not work.

Instead of discussing the whole model as stated in [17], we restrict ourselves to the treatise of the capitalists. We will point out the main ideas and briefly sketch the rest of the model.

At a given time t capitalists can either consume $c(t)$ or invest $i(t)$ of their given income. They derive this income by renting their capital stock $k(t)$ to competitive firms at the given rate $r(t)$. The gross rental income is then taxed at rate $\tau_k(t)$. Hence the capitalists' consumption and investment satisfies

$$c(t) + i(t) = [1 - \tau_k(t)]\, r(t)k(t) + \tau_k(t)\delta k(t)$$

where the term $\tau_k(t)\delta k(t)$ is a depreciation allowance with a fixed depreciation rate $\delta \in [0, 1]$. The capital stock $k(t)$ evolves according to

$$\dot{k}(t) = i(t) - \delta k(t).$$

In making their decisions, capitalists take $r(t)$ and $\tau_k(t)$ as given and want to maximize their lifetime utility

$$\int_0^\infty e^{-\rho t} u\left(c(t)\right) dt \qquad , \rho > 0 \tag{1.1}$$

where the utility function $u = u(c)$ is assumed to be increasing and strictly concave with $\lim_{c \to 0} u'(c) = \infty$ and $\lim_{c \to \infty} u'(c) = 0$. The following figure illustrates the shape of such a function:

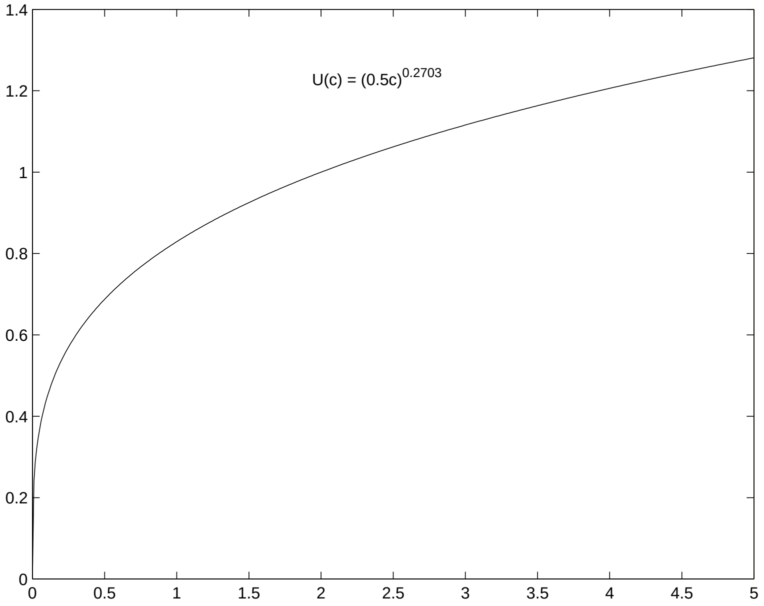

Figure 1.1: An Example of a Utility Function

The typical form of utility functions is a standard assumption in economics and represents the fact that each additional unit of a product we consume provides less utility (better known as *the principle of decreasing marginal utility*[1]).

The factor $e^{-\rho t}$ in (1.1) and thus ρ represents the time preference of the capitalist, because one usually ascribes more utility to instantaneous consumption than to consumption in 10 years. To summarize, the decision problem faced by capitalists is:

[1] As an example of this principle one usually considers the thirsty man in the desert, who suddenly faces a refrigerator filled with cool water. Clearly the first gulps eliminate most of the thirst. However, if he drinks more than ten bottles, he will clearly have no additional utility from drinking more water.

$$\begin{aligned} &\max_{c(\cdot)} \int_0^\infty e^{-\rho t} u\left(c(t)\right) dt \\ \text{subject to}\quad & c(t) + i(t) = \left[1 - \tau_k(t)\right] r(t)k(t) + \tau_k(t)\delta k(t) \\ & \dot{k}(t) = i(t) - \delta k(t) \qquad , k(0) \text{ given} \end{aligned} \tag{1.2}$$

In analogy we can set up a model for the workers which have their own utility function and corresponding lifetime utility. Firms hire the workers, pay wages and thus influence the distribution of capital.

Given this dynamic neoclassical system, the government tries to find an optimal tax policy which maximizes social welfare consisting of the lifetime utility of workers and capitalists (the society) subject to budget constraints. In fact, as shown in [17], the resulting problem has the standard form of an optimal control problem.

1.3 The General Problem

Instead of considering the model for the whole economy, we focus our attention on the decision problem for the capitalist (1.2). If we solve the first equation in (1.2) for $i(t)$, substitute it into the second one and assume the tax rate $\tau_k(t)$ to be given, then (1.2) describes an optimal control problem itself. It can easily be seen, that this class of problems can also be formulated as a more general optimization problem of the form

$$\min f(x) \quad \text{subject to} \quad c(x) = 0, \quad x \in X \tag{1.3}$$

where $f : X \to I\!R$ and $c : X \to Y$ are given mappings with Hilbert spaces X and Y over the reals. Problems of this type arise for example in the context of optimal control or parameter identification. In these cases X and Y are appropriate function spaces and the constraint $c : X \to Y$ describes a system of differential equations.

A classical technique to solve (1.3) is to compute a Karush Kuhn Tucker point $(x_*, \lambda_*) \in X \times Y$ via an augmented Lagrangian method. In this method, instead of fixing the Lagrange multiplier λ and solving for a Karush Kuhn Tucker point directly, one considers a suitable sequence of subproblems (see Chapter 3 for a further motivation)

$$\min_{x \in X} \|\nabla_x \Phi(x, \lambda_k, \mu_k)\|_X \tag{1.4}$$

where $\Phi : X \times Y \times (0, \infty) \to I\!R$ is the augmented Lagrangian functional defined by

$$\Phi(x, \lambda, \mu) := f(x) + < \lambda, c(x) >_Y + \frac{1}{2\mu}\|c(x)\|_Y^2$$

In this context the vector λ is known as the Lagrange multiplier estimate and μ is known as the penalty parameter. $< \cdot, \cdot >_Z$ denotes the inner product in some Hilbert space Z, and $\| \cdot \|_Z := < \cdot, \cdot >_Z^{1/2}$ is the corresponding norm.

1.3.1 Current Status of Research

Since augmented Lagrangian methods were introduced by Powell [22] and Hestenes [12] for the finite dimensional case $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, this class of methods has been investigated very intensively. Conn, Gould and Toint prove in [3] local and global convergence results for an augmented Lagrangian algorithm for optimization with equality constraints and simple bounds. These results were extended by Conn, Gould, Sartenaer and Toint in [1] to linear inequality constraints. In particular those papers gave an explicit rule of how to adapt the penalty parameter μ and the Lagrange multiplier estimate λ in each iteration such that μ is bounded away from zero.

While the finite dimensional case has been treated in a satisfactory manner, the infinite dimensional case is studied much less. In case of equality constraints, Polyak and Tretyakov [21] give an elegant treatise of the augmented Lagrangian method. Ito and Kunisch prove in [13] local convergence results for an augmented Lagrangian method for the minimization of a nonlinear functional in presence of equality and affine inequality constraints. In [11] Hager applies the augmented Lagrangian method to optimal control problems and points out some difficulties that arise from the discretization of the infinite dimensional setting. Ito and Kunisch analyze in [14] the augmented Lagrangian-SQP-technique with second order update of the Lagrange multiplier and prove quadratical local convergence. Volkwein applies this method in [26] to an optimal control setting, discretizes it, and proves local convergence for the finite dimensional approximation of the Hilbert space algorithm.

The methods mentioned above provide very strong results for the infinite dimensional setting. In order to implement these algorithms numerically, one must discretize the corresponding spaces and operators. Volkwein took a very general approach in [26] and approximated the infinite dimensional method by means of functional analytic restriction and prolongation operators. Using this general theory, he could derive quadratic convergence for the augmented Lagrangian-SQP-method.

However, up to this point it is not known how one should adjust the penalty parameter and the discretization level from one iteration to the next one. In this research, we will introduce an augmented Lagrangian algorithm which gives an explicit rule for the update of these parameters. Moreover, we will prove global convergence of the discretized method to a Kuhn Tucker point under appropriate assumptions.

1.3.2 Approach and Outline

The general approach to solve our task is to construct an algorithm for the Hilbert space setting and prove convergence results for this infinite dimensional method. Then one considers a discretized version of this algorithm and tries to transfer the convergence results to the finite dimensional case.

In this thesis, the infinite dimensional method used to solve (1.3) is a class of augmented Lagrangian algorithms introduced by E. W. Sachs and Annick Sartenaer in [25]. The proposed method ALINF computes in each iteration an iterate x_k which approximately solves subproblem (1.4) for fixed Lagrange multiplier estimate λ_k and penalty parameter μ_k. Here approximately is to be understood in the sense that

$$\|\nabla_x \Phi(x_k, \lambda_k, \mu_k)\|_X \leq w_k \tag{1.5}$$

where w_k is a suitable tolerance at iteration k. This tolerance will decrease from one iteration to the next one, so that the subproblem will be solved more and more precisely.

Once the iterate x_k is computed, we will test it for constraint violation. Depending on this result, ALINF adapts the penalty parameter and the Lagrange multiplier estimate in a suitable manner. One of the features of ALINF is, that it allows a certain tolerance considering the test of the constraint violation and the update of the Lagrange multiplier estimate. These properties are crucial for the discretization of the algorithm.

In Chapter 3 we introduce ALINF and analyze its convergence properties. In fact we can prove global convergence of the infinite dimensional method in the sense, that any convergent subsequence of the iterates $(x_k)_{k \in I\!N}$ converges to Kuhn Tucker point $(x_*, \lambda(x_*))$, i.e.:

$$\nabla f(x_*) + c'(x_*)^* \lambda(x_*) = 0 \ , \qquad c(x_*) = 0$$

In case of convergence of $(x_k)_{k \in I\!N}$ to a single limit point we can show further, that the penalty parameters $(\mu_k)_{k \in I\!N}$ are bounded away from zero and that the sequences of iterates $(x_k)_{k \in I\!N}$ and Lagrange multiplier estimates $(\lambda_k)_{k \in I\!N}$ converge at least R-linearly to x_* and $\lambda(x_*)$, respectively.

As mentioned before, we must approximate the infinite dimensional method in order to allow numerical computations. In this thesis the discretization is carried out in terms of functional analytic restriction and prolongation operators, similar to Volkwein [26] and Aubin [4]. This general approach allows the treatise of a wide class of problems in a rigorous way. The approximation theory needed for the discretization of the spaces and operators is derived in Chapter 4.

In Chapter 5 we introduce the discretized algorithm ALDISCR which is based on ALINF. In this algorithm, subproblem (1.5) is replaced by

$$\left\| \frac{d}{dx} \Phi_{n_k, m_k}(x_k, \lambda_k, \mu_k) \cdot \right\|_{\mathcal{L}(X_{n_k}, I\!R)} \leq c w_k \tag{1.6}$$

where $\Phi_{n_k,m_k} : X_{n_k} \times Y_{m_k} \times (0, \infty) \rightarrow I\!R$ is the discretized augmented Lagrangian functional, and X_{n_k}, Y_{m_k} are finite dimensional approximations of the Hilbert spaces X, Y with discretization levels n_k and m_k in iteration k, respectively.

By approximating the infinite dimensional method, discretization errors arise in various forms. The question is whether one can gain enough control over these errors (without imposing too restrictive assumptions) such that the convergence results of the Hilbert space algorithm can be transferred to the finite dimensional method.

It turns out, that the approximation of the derivative of the augmented Lagrangian functional with respect to x is the crucial part in the discretization process of algorithm ALINF. This is why a major part of Chapter 4 is devoted to the analysis of this problem. In fact, we can prove the pointwise convergence of the derivative of the discretized augmented Lagrangian functional with respect to x to the infinite dimensional counterpart in operator norm based on the assumption of converging approximations of the dual spaces X^* and Y^*. In order to justify this assumption, we prove that one can always find a “nice” discretization of a Hilbert space with countable basis such that the dual approximations converge.

Based on these results, we construct the nested algorithm ALDISCR. It incorporates an explicit rule of how to adjust the discretization levels, the penalty parameter and the Lagrange multiplier estimate from one iteration to the next one. In fact we can prove, that all iterates of ALDISCR satisfy the requirements of ALINF. Hence the convergence results derived for the infinite dimensional method transfer to ALDISCR. In particular the algorithm is globally convergent, converges locally at least R-linearly, and the penalty parameter is bounded away from zero. The last feature is of particular importance for the numerical realization, because it prevents the Hessian of the discretized augmented Lagrangian functional with respect to x from becoming more and more ill-conditioned. Finally, we briefly introduce a relaxation of algorithm ALDISCR in order to improve numerical efficiency.

In the next chapter we present the necessary functional analytic background and general concepts which will be used intensively in later chapters.

CHAPTER 2

Theoretical Background

In this chapter we first introduce some basic concepts in Hilbert spaces, similar to [18] and [15]. However, we extend these concepts by several theorems and examples which will be used intensively in Chapters 3 and 4. This approach facilitates the readability of proofs in those chapters.

Throughout this thesis we will consider two Hilbert spaces X and Y over the reals with inner products $< \cdot, \cdot >_X$ and $< \cdot, \cdot >_Y$ and the corresponding induced norms $\| \cdot \|_X$ and $\| \cdot \|_Y$. In most cases the assumption of Hilbert spaces can be relaxed to Banach spaces or even normed spaces, but this is clear from the context. Further let $\mathcal{L}(X,Y)$ be the set of all bounded (i.e. continuous) linear maps $g : X \to Y$. We know from functional analysis that $\mathcal{L}(X,Y)$ is a normed linear space with norm $\|g\|_{\mathcal{L}(X,Y)} = \sup_{\|h\|_X=1} \|g(h)\|_Y$.

2.1 General Derivatives and Gradient

In advanced calculus courses one introduces the directional derivative for points in some open subset of $I\!R^n$. This derivative can be generalized to Banach spaces in the following way:

Definition 2.1.1. Let $S \subset X$ be nonempty and open, let $x \in S$. Further let $f : S \to Y$ be a given mapping and h be arbitrary in X. If the limit

$$\lim_{\alpha \to 0} \frac{f(x + \alpha h) - f(x)}{\alpha}$$

exists, it is called the *Gateaux derivative of f at x in the direction of h.* If the limit exists for every $h \in X$, f is said to be *Gateaux differentiable at x.*

The analogy to the directional derivative is obvious. However, remember that the existence of the directional derivative in $I\!R^n$ does not imply continuity for $n > 1$. This implication is only true for $n = 1$. In the general case one has to demand more in order to achieve continuity as an implication. In advanced calculus one learns that the total derivative has the desired property. The appropriate generalization of this total derivative is as follows:

Definition 2.1.2. Let $S \subset X$ be nonempty and open. Further let $f : S \to Y$ be a given mapping.

i) Let $x \in S$. f is said to be *Fréchet differentiable at x* if and only if there exists a continuous linear mapping $f'(x)(\cdot) : X \to Y$ such that

$$\|f(x+h) - f(x) - f'(x)(h)\|_Y \leq \|h\|_X \varepsilon(\|h\|_X)$$

for all $h \in X$ such that $x + h \in S$ where $\varepsilon : I\!R \to I\!R$ and $lim_{r\to 0}\varepsilon(r) = 0$. Then $f'(x)(\cdot)$ is called *Fréchet-derivative of f at x* and $f'(x)(h)$ is called *Fréchet-derivative of f at x in the direction h.*

ii) f is said to be *Fréchet differentiable on $D \in S$* if and only if f is Fréchet differentiable for all $x \in D$. Then we call $f'(\cdot) : D \to \mathcal{L}(X, Y)$ defined by $x \mapsto f'(x)(\cdot)$ the *Fréchet derivative of f on D.*

Remark:

i) Note that $f'(x)$ is itself a mapping, namely:

$$f'(x) : X \to Y$$

$$X \ni h \mapsto f'(x)(h).$$

It is continuous and linear with respect to its argument h. Therefore we will switch from now onwards to the operator notation:

$$f'(x)h := f'(x)(h).$$

ii) Note, that the inequality in the definition is equivalent to

$$\sup_{\|h\|\leq r} \frac{\|f(x+h) - f(x) - f'(x)h\|}{\|h\|} \to 0, \quad r \to 0.$$

What we demand is a “uniform convergence”. This will guarantee continuity in case of existence of the derivative.

iii) We did not show yet that this new derivative is well defined. This will be proved in upcoming theorems and remarks.

As mentioned above this new derivative is the generalization of the total derivative in advanced calculus. This comparison will be clarified in the following example:

Example 2.1.3. Let now $X = I\!\!R^n$ and $Y = I\!\!R^m$ with the Euclidian norm and the standard inner product. Further let $A \in I\!\!R^{m,n}$ and $f : I\!\!R^n \to I\!\!R^m$ be defined by $f(x) := Ax$. Then

$$\|f(x+h) - f(x) - Ah\| = \|A(x+h) - Ax - Ah\| = 0 \leq 0 \cdot \|h\|$$

$$\overset{\text{Def.2.1.2}}{\Longrightarrow} \quad f'(x)h = Ah$$

That means $f'(x)$ is the operator that maps h to Ah. This differs from what one usually learns in advanced calculus where one defines $f'(x)$ to be A, an object in $I\!\!R^{m,n}$. Mathematically $f'(x)(\cdot)$ and A are the same objects, because we know from linear algebra that $I\!\!R^{m,n}$ is isomorphic to $L(I\!\!R^n, I\!\!R^m)$. However, in the general case there is no isomorphism and thus we have to be more precise with our notation.

The next theorem shows that the simplicity of the Fréchet derivative of linear mappings also transfers to the general setting. The proof is completely analogous to the last example. However, in order to meet the requirements of the definition, we have to request now that the linear operator is bounded.

Theorem 2.1.4. *Let $A : X \to Y$ be a bounded and linear operator. Then A is Fréchet differentiable on $X$and $\quad \forall x,\ h \in X \quad A'(x)h = Ah$.*

We will now prove some basic theorems.

Theorem 2.1.5. *Let $S \subset X$ be nonempty and open. Further let $f : S \to Y$ be a given mapping and $x \in S$. If the Fréchet derivative of f at x exists, it is unique.*

Proof. Suppose both $f'(x)(\cdot)$ and $\tilde{f}'(x)(\cdot)$ satisfy the requirements of the Fréchet derivative. For $h \in X$ define $l(h) := f'(x)(h) - \tilde{f}'(x)(h)$. Thus $l(\cdot)$ is linear and

$$\begin{aligned} \|l(h)\| &= \|f'(x)(h) - \tilde{f}'(x)(h)\| \\ &\leq \|f(x+h) - f(x) - f'(x)(h)\| + \|f(x+h) - f(x) - \tilde{f}'(x)(h)\| \\ &\leq \|h\|\varepsilon(\|h\|) + \|h\|\tilde{\varepsilon}(\|h\|) \\ &=: \|h\|\bar{\varepsilon}(\|h\|) \end{aligned}$$

where $\bar{\varepsilon}(r) \to 0, \quad r \to 0$.

Pick a sequence $(r_n)_{n=1}^{\infty} \subset I\!R$ such that $\lim_{n\to\infty} r_n = \infty$. $l(\cdot)$ is linear, thus

$$\|l(h)\| = \left\| r_n l(\frac{h}{r_n}) \right\| \leq r_n \left\| \frac{h}{r_n} \right\| \bar{\varepsilon}\left(\frac{\|h\|}{r_n} \right) = \|h\| \bar{\varepsilon}\left(\frac{\|h\|}{r_n} \right) \to 0, \quad n \to \infty$$

Thus $l(h) = 0$ for all $h \in X$ which implies $f'(x)(\cdot) = \tilde{f}'(x)(\cdot)$. □

Theorem 2.1.6. *Let $S \subset X$ be nonempty and open. Further let $f : S \to Y$ be a given mapping and $x \in S$. If the Fréchet derivative of f at x exists, then f is continuous at x.*

Proof. In the definition of the Fréchet derivative we requested "uniform convergence" (see the remark after the definition). This uniformity assures now, that there exists a δ such that for $\|h\| < \delta$

$$\|f(x+h) - f(x) - f'(x)h\| \leq M\|h\|$$

for some $M > 0$. Hence we obtain by the triangle inequality

$$\|f(x+h) - f(x)\| \leq M\|h\| + \|f'(x)h\| \leq M\|h\| + \|f'(x)(\cdot)\|_{\mathcal{L}(X,Y)}\|h\| =: \tilde{M}\|h\|$$

which implies continuity. □

Theorem 2.1.7. *Let $S \subset X$ be nonempty and open. Further let $f : S \to Y$ be a given mapping and $x \in S$. If the Fréchet derivative of f at x exists, then f is Gateaux differentiable at x and both derivatives are equal.*

Proof. Let $h \in X$ be arbitrary. By the definition of the Fréchet derivative we have for small $\alpha \in I\!R$

$$\|f(x+\alpha h) - f(x) - f'(x)\alpha h\| \leq \|\alpha h\| \varepsilon(\|\alpha h\|)$$

and thus

$$\left\| \frac{f(x+\alpha h) - f(x)}{\alpha} - f'(x)h \right\| = \frac{1}{\alpha} \|f(x+\alpha h) - f(x) - f'(x)\alpha h\| \leq \|h\|\varepsilon(\|\alpha h\|) \to 0, \quad \alpha \to 0$$

which proves the theorem. □

Remark: This theorem shows us how to calculate the Fréchet derivative. Recall that we do not know its form or even whether it exists in the general case. The calculation has two steps:

1. Write down the definition of the Gateaux derivative and try to calculate the limit. If this limit exists, it is our candidate for the Fréchet derivative.

2. Plug in this limit in the definition of the Fréchet derivative and try to fulfill the requirements.

This procedure will be illustrated in the following example:

Example 2.1.8. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $L^p(\Omega, \mu)$ denote the set of all equivalence classes of $\mathcal{A}$-measurable and to the p-power μ-integrable functions $f : \Omega \to I\!R$. We know from measure and probability theory, that $L^p(\Omega, \mu)$ is a normed linear space with norm

$$\|f\|_{L^p} := (\int_\Omega |f|^p d\mu)^{1/p}$$

Now let $(X, \|\cdot\|_X) := (L^2, \|\cdot\|_{L^2})$ *and* $(Y, \|\cdot\|_Y) := (L^1, \|\cdot\|_{L^1})$. Define $g : L^2 \to L^1$ by

$$L^2 \ni f \mapsto g(f) := f^2 \in L^1$$

Our goal is to calculate the Fréchet derivative of g.

i) Calculate Gateaux-derivative:

Let $x(\cdot)$, $h(\cdot) \in L^2$. Then

$$\begin{aligned}\left\|\frac{g(x(\cdot)+\alpha h(\cdot))-g(x(\cdot))}{\alpha}\right\|_{L^1} &= \int_\Omega \left|\frac{(x(\cdot)+\alpha h(\cdot))^2-x(\cdot)^2}{\alpha}\right| d\mu \\ &= \int_\Omega \left|\frac{x(\cdot)^2+2\alpha x(\cdot)h(\cdot)+\alpha^2 h(\cdot)^2-x(\cdot)^2}{\alpha}\right| d\mu \\ &= \int_\Omega \left|2x(\cdot)h(\cdot)+\alpha h(\cdot)^2\right| d\mu\end{aligned}$$

We want to calculate the limit of this expression as $\alpha \to 0$. Note that this is a limit of an integral, hence we should try to apply Lebesgue's theorem of majorized convergence. For this purpose let $(\alpha_n)_{n=1}^\infty \subset I\!R$ be a sequence such that $\alpha_n \to 0, \quad n \to \infty$.

$$\Longrightarrow \exists \sup_{n \in I\!N} |\alpha_n| =: \bar{\alpha}$$

Now look at

$$c_n := \int_\Omega |2x(\cdot)h(\cdot) + \alpha_n h(\cdot)^2| d\mu$$

Clearly, $(2x(\cdot)h(\cdot) + \alpha_n h(\cdot)^2)_n$ converges pointwise almost everywhere. Further

$$|2x(\cdot)h(\cdot) + \alpha_n h(\cdot)^2| \leq 2|x(\cdot)h(\cdot)| + \bar{\alpha} h(\cdot)^2 \in L^1$$

due to the Hölder-inequality (see Appendix, Theorem A.1). Thus we can apply Lebesgue's theorem of majorized convergence (see Appendix, Theorem A.2) and obtain

$$\begin{aligned}\lim_{n\to\infty} c_n &= \lim_{n\to\infty}\int_\Omega |2x(\cdot)h(\cdot)+\alpha_n h(\cdot)^2|d\mu\\ &= \int_\Omega \lim_{n\to\infty}|2x(\cdot)h(\cdot)+\alpha_n h(\cdot)^2|d\mu\\ &= \int_\Omega |2x(\cdot)h(\cdot)|d\mu\end{aligned}$$

Thus $2x(\cdot)h(\cdot)+\alpha_n h(\cdot)^2 \to_{n\to\infty} 2x(\cdot)h(\cdot)$ in $\|\cdot\|_{L^1}\ \forall\,(\alpha_n)_n\subset I\!R\ s.t.\ \alpha_n\to_{n\to\infty} 0$

$$\Longrightarrow \lim_{\alpha\to 0}\frac{g(x(\cdot)+\alpha h(\cdot))-g(x(\cdot))}{\alpha} = \lim_{\alpha\to 0}[2x(\cdot)h(\cdot)+\alpha h(\cdot)^2] = 2x(\cdot)h(\cdot)$$

which is by definition the Gateaux-derivative and thus $2x(\cdot)h(\cdot)$ is our candidate for the Fréchet derivative.

ii) Now we check whether this candidate fulfills the definition of the Fréchet derivative. Let $x(\cdot)$, $h(\cdot)\in L^2$. Then

$$\begin{aligned}&\|g(x(\cdot)+h(\cdot))-g(x(\cdot))-2x(\cdot)h(\cdot)\|_{L^1}\\ &= \int_\Omega \left|x(\cdot)^2+2x(\cdot)h(\cdot)-h(\cdot)^2-x(\cdot)^2-2x(\cdot)h(\cdot)\right| d\mu\\ &= \int_\Omega |h(\cdot)|^2 d\mu = \|h(\cdot)h(\cdot)\|_{L^1}\\ &\overset{\text{Th. A.1}}{\leq} \|h(\cdot)\|_{L^2}\|h(\cdot)\|_{L^2} =: \|h(\cdot)\|_{L^2}\cdot\varepsilon(\|h(\cdot)\|_{L^2})\end{aligned}$$

$$\Longrightarrow \forall x(\cdot),\ h(\cdot)\in L^2\quad f'(x(\cdot))h(\cdot) = 2x(\cdot)h(\cdot)$$

Note, that this means $f'(\cdot): L^2\to\mathcal{L}(L^2,L^1)$ and $x\mapsto f'(x)(\cdot)$ where $f'(x)(\cdot)$ is given by $h\mapsto f'(x)(h)=2xh$.

We see that (if $2x$ is interpreted as the operator $f'(x)(\cdot)$) our old formula from the real line still holds: $\frac{d}{dx}x^2 = 2x$. The difference is, that $2x$ is now an operator!

This example shows that the Fréchet derivative is really the natural generalization of the total derivative in $I\!R^n$. However, one has to use caution. In applications one should verify predicted results by hand.

Several other definitions and results can be derived in analogy to the classical derivative for mappings $f : I\!\!R \to I\!\!R$. For example it is easy to prove that the linearity of the differentiation operator also holds for the Fréchet derivative. Further it is obvious how one defines n-times Fréchet differentiable and continuous Fréchet differentiable. But one has to keep in mind what this continuity means in Hilbert or Banach spaces: Namely, that

$$x_n \to_{\|\cdot\|_X} x \quad \Longrightarrow \quad f'(x_n)(\cdot) \to_{\|\cdot\|_{\mathcal{L}(X,Y)}} f'(x)(\cdot)$$

i.e., convergence in the operator norm.

We will now derive the chain rule for the Fréchet derivative. Note that we will not obtain a product rule in the general Banach space setting as a product of vectors is not defined. Often we will even write the composition of operators as a product, because no misunderstanding can result from this notation.

Theorem 2.1.9. *Let in addition to X and Y now $(Z, \|\cdot\|_Z)$ be a normed linear space. Let $D \subset X$, $E \subset Y$ be open sets. Let $f : D \to E$, let $g : E \to Z$. Further let f be Fréchet differentiable at $x \in D$ and g be Fréchet differentiable at $y := f(x) \in E$. Then $T := g \circ f$ is Fréchet differentiable at x and $T'(x) = g'(y)f'(x)$.*

Proof. Let $h \in X$ such that $x + h \in D$. Then

$$\begin{aligned} T(x+h) - T(x) &= g(f(x+h)) - g(f(x)) \\ &= g(f(x) + f(x+h) - f(x)) - g(f(x)) \\ &= g(y+\tilde{y}) - g(y) \qquad \text{(define } y := f(x)\text{ , } \tilde{y} := f(x+h) - f(x)\text{)} \end{aligned}$$

$$\Longrightarrow \|T(x+h) - T(x) - g'(y)\tilde{y}\| = \|g(y+\tilde{y}) - g(y) - g'(y)\tilde{y}\| \le \|\tilde{y}\|\varepsilon_g(\|\tilde{y}\|)$$

Besides this the following inequality holds:

$$\|\tilde{y} - f'(x)h\| = \|f(x+h) - f(x) - f'(x)h\| \le \|h\|\varepsilon_f(\|h\|)$$

However, f is also continuous at x, so

$$\begin{aligned} \|\tilde{y}\| &= \|f(x+h) - f(x)\| \quad \le \quad \|f(x+h) - f(x) - f'(x)h\| + \|f'(x)h\| \\ &\le \varepsilon_f(\|h\|)\|h\| + \|h\|\|f'(x)\|_{\mathcal{L}(X,Y)} \;\to\; 0, \quad \|h\| \to 0 \end{aligned}$$

$$\Longrightarrow \quad \|\tilde{y}\| = \tilde{\varepsilon}(\|h\|) \to 0, \quad \|h\| \to 0$$

So altogether:

$$\begin{aligned}
&\|T(x+h)-T(x)-g'(y)f'(x)h\| \\
&\quad\leq\ \|T(x+h)-T(x)-g'(y)\tilde{y}\|+\|g'(y)\tilde{y}-g'(y)f'(x)h\| \\
&\quad\leq\ \|\tilde{y}\|\varepsilon_g(\|\tilde{y}\|)+\|g'(y)\|\|\tilde{y}-f'(x)h\| \\
&\quad\leq\ \|\tilde{y}\|\varepsilon_g(\|\tilde{y}\|)+\|g'(y)\|\|h\|\varepsilon_f(\|h\|) \\
&\quad\leq\ [\varepsilon_f(\|h\|)\|h\|+\|h\|\|f'(x)\|]\,\varepsilon_g\left(\tilde{\varepsilon}(\|h\|)\right)+\|g'(y)\|\|h\|\varepsilon_f(\|h\|) \\
&\quad=:\ \|h\|\hat{\varepsilon}(\|h\|)
\end{aligned}$$

□

The next theorem is a simple application of the chain rule and can be proved easily with Theorem 2.1.4:

Theorem 2.1.10. *In addition to X and Y let $(Z, \|\cdot\|_Z)$ be a normed space. Let $A : Y \to Z$ be a bounded and linear operator. Further let $T : X \to Y$ be Fréchet differentiable. Then AT is Fréchet differentiable and $\forall x \in X \quad (AT)'(x) = AT'(x)$.*

Now we will introduce the concept of gradients. This concept is based on the special properties of Hilbert spaces. As defined at the beginning of chapter 2, we assume X to be a Hilbert space with scalar product $< \cdot, \cdot >_X$.

Let now $f : X \to I\!R$ be a given mapping and assume existence of the Fréchet derivative at some $x \in X$. Note that $f'(x)(\cdot) : X \to I\!R$ and thus (due to linearity and continuity) $f'(x) \in X^* = \mathcal{L}(X, I\!R)$. Thus the Riesz representation theorem (see Appendix, Theorem A.8) assures the existence of a unique $h_x \in X$ such that $f'(x)(\cdot) =< h_x, \cdot >$. This result forms the basis for the following definition:

Definition 2.1.11. Let $S \subset X$ be nonempty and open, let $x \in S$. Further let $f : S \to I\!R$ be a mapping such that the Fréchet derivative $f'(x)$ exists. Let $h_x \in X$ be the Riesz representation of $f'(x)(\cdot)$, i.e.

$$f'(x)(\cdot) =< h_x, \cdot >$$

Then we define $\nabla f(x) := h_x$ and call $\nabla f(x)$ the *gradient of f at x*.

Remark:

i) If $D \subset S$ and $\forall x \in D\ \exists\ f'(x)$, then $\nabla f(\cdot) : D \to X$ defines a mapping by $x \mapsto \nabla f(x)$.

ii) In an earlier remark we mentioned that the Fréchet differential operator is linear. This implies together with the linearity of the scalar product and the uniqueness of the Riesz representation, that $\nabla(\cdot)$ is also a linear operator. That is, assume that for two mappings $f,\ g : S \to I\!R \quad \exists\ \nabla f(x)$ and $\nabla g(x)$. Then for $\alpha,\ \beta \in I\!R$

$$\nabla(\alpha f + \beta g)(x) = \alpha\nabla f(x) + \beta\nabla g(x)$$

iii) The Riesz representation theorem (see Appendix, Theorem A.8) also implies that $\|f'(x)(\cdot)\|_{L(X,I\!R)} = \|\nabla f(x)\|_X$. This result can be used to show that $\nabla f(\cdot)$ is continuous if f is continuously Fréchet differentiable.

iv) In case of $X = I\!R^n$ this gradient definition is equivalent to what one usually learns in advanced calculus.

We will clarify this definition by the following simple example:

Example 2.1.12. Let $X = I\!R^n$ and define $f : I\!R^n \to I\!R$ by $f(x) := c^T x$ where $c \in I\!R^n$. Then Example 2.1.3 implies that the Fréchet derivative of f is given by

$$f'(x)h = c^T h =< c, h >_{I\!R^n}$$

The Riesz representation is unique, thus $\nabla f(x) = c$.

2.2 Higher Derivatives and Taylor's Theorem

In this section we will derive Taylor's Theorem for general normed linear spaces and use it to prove certain results that are needed later.

We already mentioned in Section 2.1 that one has to use caution when dealing with Fréchet derivatives of higher order. The corresponding domains and ranges of these mappings are not trivial and require a precise treatment and notation. Consider now a mapping $f : X \to Y$ which is n times Fréchet differentiable for given $n \in I\!N$.

As we already pointed out in the last section, the first derivative $f'(x)(\cdot)$ is a mapping from X to Y, whereas $f'(\cdot)$ is a mapping from X to $\mathcal{L}(X, Y)$. In case of the second derivative it gets more complicated: $f''(\cdot)$ maps from X to $\mathcal{L}(X, \mathcal{L}(X, Y))$. For given x, h_1 and $h_2 \in X$ this means that $f''(x) \in \mathcal{L}(X, \mathcal{L}(X, Y))$, $f''(x)h_2 \in \mathcal{L}(X, Y)$ and finally $f''(x)h_2h_1 \in Y$. For the n^{th} derivative $f^{(n)}(\cdot)$ we obtain that $f^{(n)}(x)h_nh_{n-1}\cdots h_1 \in Y$. Note that $f^{(n)}(x)h_nh_{n-1}\cdots h_1$ is by definition of the Fréchet derivative bounded and linear with respect to every h_ν.

However, it turns out that it is sometimes useful to suppress the arguments $h_1, \ldots, h_n$. We will write $f^{(n)}(x)$ as $f^{(n)}(x) \cdot \ldots \cdot$ with n dots after the closing parenthesis. For example $x \mapsto f''(x) \cdot\cdot$ is the mapping from X to $\mathcal{L}(X, \mathcal{L}(X, Y))$. If one calculates Fréchet derivatives of higher order, the dots do not always appear in this nice order, thus it will be important to link indices to these dots. We write:

$$f^{(n)}(x) = f^{(n)}(x) \cdot_n \cdot_{n-1} \cdots \cdot_2 \cdot_1$$

The following examples show the advantage of the new notation:

Example 2.2.1. For given $f : X \to I\!R$ and $c : X \to Y$ consider the Lagrange functional $L : X \times Y \to I\!R$ defined by

$$L(x,\lambda) := f(x) + < \lambda, c(x) >_Y$$

Assume that the mappings f and c are twice continuous Fréchet differentiable. Our goal is to calculate the first and second Fréchet derivative of $L(x,\lambda)$ with respect to its argument x. We will denote these derivatives by $\frac{d}{dx}L(x,\lambda)$ and $\frac{d^2}{dx^2}L(x,\lambda)$. By Theorem 2.1.10 and the linearity of the Fréchet differential operator we obtain:

$$\frac{d}{dx}L(x,\lambda)\cdot = f'(x)\cdot + < \lambda, c'(x)\cdot >_Y$$

and hence by the same argumentation

$$\frac{d^2}{dx^2}L(x,\lambda)\cdot\cdot = f''(x)\cdot\cdot + < \lambda, c''(x)\cdot\cdot >_Y$$

Example 2.2.2. Define for the mapping $c(\cdot)$ from the previous example $g : X \to I\!R$ by

$$g(x) := \frac{1}{2}\|c(x)\|_Y^2 = \frac{1}{2} < c(x), c(x) >_Y$$

Our goal is to calculate the second Fréchet derivative of g. Hence, we start with the first derivative and in particular with the Gateaux derivative. Let $h \in X$ and $r \in I\!R$. Then

$$\begin{aligned}
&\lim_{r\to 0}\frac{g(x+rh)-g_2(x)}{r} \\
= \;& \lim_{r\to 0}\frac{1}{2r}\Big[< c(x+rh), c(x+rh) > - < c(x), c(x) > \Big] \\
= \;& \lim_{r\to 0}\frac{1}{2r}\Big[< c(x+rh), c(x+rh) > - < c(x), c(x+rh) > \\
& \qquad + < c(x), c(x+rh) > - < c(x), c(x) > \Big] \\
= \;& \frac{1}{2}\lim_{r\to 0} < \frac{c(x+rh)-c(x)}{r}, c(x+rh) > + \frac{1}{2}\lim_{r\to 0} < c(x), \frac{c(x+rh)-c(x)}{r} > \\
= \;& \frac{1}{2} < c'(x)h, c(x) > + \frac{1}{2} < c(x), c'(x)h >
\end{aligned}$$

where the last equality holds due to continuity of the scalar product and $c(\cdot)$. The spaces we consider are Hilbert spaces over the reals and thus

$$\lim_{r\to 0}\frac{g(x+rh)-g_2(x)}{r} =< c(x), c'(x)h >$$

It is now easy to show (by using the trick of adding a neutral zero), that the Gateaux derivative fulfills the requirements of the Fréchet derivative, that is

$$g'(x)h =< c(x), c'(x)h >$$

or equivalently $g'(x)\cdot =< c(x), c'(x)\cdot >$. We continue to calculate the second Fréchet derivative of g by using the chain rule derived in Section 2.1. To apply this rule we decompose $g'(x)$. Define $T : Y \times \mathcal{L}(X, Y) \to \mathcal{L}(X, I\!R)$ by

$$T \begin{pmatrix} u \\ v \end{pmatrix} \cdot :=< u, v\cdot >_Y$$

Note that $T(u, v)$ has a dot, because it is a mapping itself. To obtain a candidate for the Fréchet derivative, we calculate the Gateaux-derivative:

$$\begin{aligned} &\lim_{r\to 0} \frac{1}{r} (< u + ru_h, (v + rv_h)\cdot > - < u, v\cdot >) = \\ &\quad = \lim_{r\to 0} \frac{1}{r} (< u, rv_h\cdot > + < ru_h, v\cdot > + < ru_h, rv_h\cdot >) \\ &\quad = < u, v_h\cdot > + < u_h, v\cdot > \end{aligned}$$

Now we check whether this candidate satisfies the definition of the Fréchet derivative:

$$\begin{aligned} &\| < u + u_h, (v + v_h)\cdot > - < u, v\cdot > - < u, v_h\cdot > - < u_h, v\cdot > \| = \\ &\quad = \| < u_h, v_h\cdot > \|_{\mathcal{L}(X,I\!R)} = \sup_{\|x\|=1} | < u_h, v_h x > | \\ &\quad \overset{C.S.I.}{\leq} \|u_h\|_X \sup_{\|x\|=1} \|v_h x\|_Y \leq \|u_h\|_X \|v_h\|_{\mathcal{L}(X,Y)} \\ &\quad \leq \|u_h\|^2 + 2\|u_h\|\|v_h\| + \|v_h\|^2 = \left(\|u_h\|_X + \|v_h\|_{\mathcal{L}(X,Y)}\right)^2 \\ &\quad = \left(\left\| \begin{pmatrix} u_h \\ v_h \end{pmatrix} \right\|_{X\times\mathcal{L}(X,Y)} \right)^2 \end{aligned}$$

Hence the Fréchet derivative exists and

$$T' \begin{pmatrix} u \\ v \end{pmatrix} \begin{pmatrix} u_h \\ v_h \end{pmatrix} \cdot =< u, v_h\cdot >_Y + < u_h, v\cdot >_Y$$

Clearly, in order to decompose $g'(x)$ and use the chain rule, we must now define a mapping $S : X \to Y \times \mathcal{L}(X, Y)$ by

$$S(x)\cdot_1 := \begin{pmatrix} c(x) \\ c'(x)\cdot_1 \end{pmatrix}$$

By definition of the norm on the product space $Y \times \mathcal{L}(X,Y)$ it is obvious that S is differentiable and

$$S'(x) \cdot_2 \cdot_1 = \begin{pmatrix} c'(x)\cdot_2 \\ c''(x) \cdot_2 \cdot_1 \end{pmatrix}$$

Hence, as $g'(x)\cdot_1 = T(S(x))\cdot_1$, we obtain by the chain rule

$$\begin{aligned} g''(x) \cdot_2 \cdot_1 &= T'(S(x)\cdot_1)S'(x) \cdot_2 \cdot_1 = h' \begin{pmatrix} c(x) \\ c'(x)\cdot_1 \end{pmatrix} \begin{pmatrix} c'(x)\cdot_2 \\ c''(x) \cdot_2 \cdot_1 \end{pmatrix} \\ &= < c(x), c''(x) \cdot_2 \cdot_1 > + < c'(x)\cdot_2, c'(x)\cdot_1 > \end{aligned}$$

In this final equation, we can see that it is very important to number the dots. Hence our notation makes sense. However, with even higher derivatives it might be more convenient to make the calculations with the corresponding h_ν and set up the dot-notation in the end.

Consider again the notation $f^{(n)}(x) = f^{(n)}(x)h_n h_{n-1} \cdots h_1$. For our purpose the case $h_n = h_{n-1} = \ldots = h_1$ is of particular interest. Clearly, a product of vectors is not defined in general normed linear spaces. However, the following notation turns out to be extremely useful and intuitive for higher derivatives:

Define for $x_0,\ x \in X$ and for $\nu = 0, \ldots, n$

$$f^{(\nu)}(x_0)x^\nu := f^{(\nu)}(x_0) \underbrace{xx \cdots x}_{\nu-\text{times}}$$

We are now ready to introduce the n^{th} Taylor polynomial of f:

Definition 2.2.3. Let $D \subset X$ be open and let $f : D \to Y$ be n-times Fréchet differentiable at $x_0 \in D$. Then for $x \in X$ the n^{th} *Taylor polynomial* is defined by

$$T_n(x) := \sum_{\nu=0}^{n} \frac{1}{\nu!} f^{(\nu)}(x_0)(x - x_0)^\nu$$

Now we derive a remarkable generalization of Taylor's Theorem for mappings from one normed linear space to another. The theorem we state here is an extension of the corresponding theorem in [6, page 281]:

Theorem 2.2.4 (Taylor's Theorem with Remainder Estimate). *Let X and Y be normed linear spaces. Further let $D \subset X$ be open and $f : D \to Y$ be n-times Fréchet differentiable on D. Let $x, x_0 \in D$ such that $x_0 + t(x - x_0) \in D\ \forall\ 0 \le t \le 1$.*

i) Assume that f is $n+1$-times Fréchet differentiable on D. Then the following inequality holds:

$$\|f(x) - T_n(x)\| \leq \frac{1}{(n+1)!}\|x - x_0\|^{n+1} \sup_{0<t<1} \|f^{(n+1)}(x_0 + t(x - x_0))\|$$

ii) If f is as assumed n-times Fréchet differentiable, then $\exists\ \theta = \theta(x, x_0, n) \in (0,1)$ such that

$$\|f(x) - T_n(x)\| \leq \frac{1}{n!}\|x - x_0\|^n \|f^{(n)}(x_0 + \theta(x - x_0)) - f^{(n)}(x_0)\|$$

Further, if f is n-times continuous Fréchet differentiable, then

$$f(x) = T_n(x) + o(\|x - x_0\|^n)$$

Proof. First we prove part i) of the theorem. According to Theorem A.5 (see Appendix A, Corollary of Hahn Banach Theorem), there exists a nonzero y^* in Y^*, the dual space of Y, such that

$$y^*(f(x) - T_n(x)) = \|y^*\|_{Y^*}\|f(x) - T_n(x)\|_Y \tag{I}$$

Now define $\Phi : [0,1] \to I\!\!R$ by

$$\Phi(\alpha) := y^*[f(x_0 + \alpha(x - x_0))]$$

But then we obtain by Theorem 2.1.10 and Theorem 2.1.9

$$\Phi'(\alpha) = y^*[f'(x_0 + \alpha(x - x_0))(x - x_0)]$$

Using the same arguments, we can derive by induction for $\nu = 0, \ldots, n+1$

$$\Phi^{(\nu)}(\alpha) = y^*[f^{(\nu)}(x_0 + \alpha(x - x_0))(x - x_0)^\nu]$$

Hence, by Taylor's Theorem for functions of a real variable (see Appendix, Theorem A.3) $\exists\ \theta \in (0,1)$ such that

$$\Phi(1) - \sum_{\nu=0}^{n} \frac{\Phi^{(\nu)}(0)}{\nu!}(1-0)^\nu = \frac{1}{(n+1)!}\Phi^{(n+1)}(\theta)$$

which implies that

$$\begin{aligned} y^* \left[f(x) - \sum_{\nu=0}^{n} \frac{1}{\nu!} f^{(\nu)}(x_0)(x-x_0)^\nu \right] &= \\ &= y^* \left[\frac{1}{(n+1)!} f^{(n+1)}(x_0 + \theta(x-x_0))(x-x_0)^{(n+1)} \right] \qquad \text{(II)} \end{aligned}$$

Thus, by the boundedness of the involved linear mappings and equation (I)

$$\begin{aligned} \|y^*\| \|f(x) - T_n(x)\| &\leq \|y^*\| \left\| \frac{1}{(n+1)!} f^{(n+1)}(x_0 + \theta(x-x_0)) \right\| \|x-x_0\|^{n+1} \\ &\leq \|y^*\| \frac{1}{(n+1)!} \|x-x_0\|^{n+1} \sup_{0<t<1} \|f^{(n+1)}(x_0 + t(x-x_0))\| \end{aligned}$$

But $y^* \neq 0$, hence part i) of the theorem is proved. In order to prove part ii) we will use the proof of part i). Define $g(x) := f(x) - T_n(x)$. Denote the Taylor polynomial of order $n-1$ of g by T^g_{n-1}. It is easy to show that $T^g_{n-1} = 0$. Hence, by equation (II), applied to g and T^g_{n-1}

$$\begin{aligned} y^* \left[g(x) - T^g_{n-1} \right] = y^* \left[f(x) - T^f_n(x) \right] &= \\ &= y^* \left[\frac{1}{n!} (f^{(n)}(x_0 + \theta(x-x_0)) - f^{(n)}(x_0))(x-x_0)^n \right] \end{aligned}$$

Now, by using the same arguments as above at the end of part i), we obtain the inequality in part ii) of the theorem:

$$\|f(x) - T_n(x)\| \leq \frac{1}{n!} \|x-x_0\|^n \|f^{(n)}(x_0 + \theta(x-x_0)) - f^{(n)}(x_0)\|$$

Clearly, if f is n-times continuous Fréchet differentiable $f^{(n)}(x_0 + \theta(x-x_0))$ converges to $f^{(n)}(x_0)$ for $x \to x_0$. Hence also the last claim in part ii) of the theorem is proven. □

If we set in part i) of Taylor's theorem $n = 0$, we obtain the generalized mean value theorem for normed linear spaces:

Corollary 2.2.5 (Mean Value Theorem). *Let $D \subset X$ be open and $f : D \to Y$ be Fréchet differentiable on D. Let $x \in D$, $h \in X$ and suppose that $x + th \in D \quad \forall\ 0 \leq t \leq 1$. Then*

$$\|f(x+h) - f(x)\| \leq \|h\| \sup_{0<t<1} \|f'(x+th)\|$$

This theorem helps us to prove the next result:

Theorem 2.2.6. *Let $f : X \to Y$ be continuous Fréchet differentiable on X. Let $x_* \in X$. Then f is Lipschitz-continuous in a neighborhood of x_*.*

Proof. Let $M > 0$. The continuity of $f'(\cdot)$ implies:

$$\exists\ \delta > 0 \quad \text{s.t.} \quad \|x - x_*\| < \delta \quad \text{implies} \quad \|f'(x)(\cdot) - f'(x_*)(\cdot)\|_{\mathcal{L}(X,Y)} < M$$

Now let $x \in B_\delta(x_*)$.

$$\Longrightarrow \quad \|f'(x)\|_{\mathcal{L}(X,Y)} - \|f'(x_*)\|_{\mathcal{L}(X,Y)} \leq \|f'(x) - f'(x_*)\|_{\mathcal{L}(X,Y)} < M$$

$$\Longrightarrow \quad \|f'(x)\|_{\mathcal{L}(X,Y)} < M + \|f'(x_*)\|_{\mathcal{L}(X,Y)} =: L$$

Thus $f'(x)$ is bounded on $B_\delta(x_*)$. Let $x \in B_\delta(x_*)$ and $h \in X$ such that $x + h \in B_\delta(x_*)$. $B_\delta(x_*)$ is convex, thus $x + \alpha h \in B_\delta(x_*) \quad \forall\ 0 \leq \alpha \leq 1$.

$$\Longrightarrow \quad \|f'(x + \alpha h)\| \leq L \quad \forall\ 0 \leq \alpha \leq 1$$

Hence the generalized mean value theorem implies $\|f(x + h) - f(x)\| \leq L\|h\|$. □

2.3 Hilbert-adjoint Operator

In this section we will introduce Hilbert-adjoint operators and apply this concept to functions needed in later chapters. As the name already indicates, the specific structure of Hilbert spaces will be used.

Definition 2.3.1. Let $T : X \to Y$ be a continuous linear mapping. Then the *Hilbert-adjoint operator* T^* of T is the operator $T^* : Y \to X$ such that

$$< y, Tx >_Y = < T^* y, x >_X \quad \forall\ x \in X,\ \forall y \in Y$$

We will show now, that the Hilbert-adjoint operator T^* is well defined:

Theorem 2.3.2. *Let $T : X \to Y$ be a continuous linear mapping. Then the Hilbert-adjoint operator T^* of T exists, is unique and is a bounded linear operator with norm*

$$\|T^*\|_{\mathcal{L}(Y,X)} = \|T\|_{\mathcal{L}(X,Y)}$$

Proof. Define $B : Y \times X \to I\!R$ by $B(y,x) :=< y, Tx >$. Our goal is to apply a corollary of the Riesz representation theorem (see Appendix, Corollary A.9). Let $x,\ x_1,\ x_2 \in X$, let $y,\ y_1,\ y_2 \in Y$. Further let $\alpha,\ \beta \in I\!R$. Then

$$\begin{aligned} B(y, \alpha x_1 + \beta x_2) &= < y, T(\alpha x_1 + \beta x_2) > \ = \ < y, \alpha T x_1 > + < y, \beta T x_2 > \\ &= \alpha < y, Tx_1 > + \beta < y, Tx_2 > \ = \ \alpha B(y, x_1) + \beta B(y, x_2) \end{aligned}$$

In analogy we can show $B(\alpha y_1 + \beta y_2, x) = \alpha B(y_1, x) + \beta B(y_2, x)$. Further we obtain by applying the Cauchy-Schwarz Inequality and using the boundedness of T:

$$|B(y,x)| \ = \ |< y, Tx >| \ \leq \ \|y\| \|Tx\| \ \leq \ \|y\|_Y \|T\|_{\mathcal{L}(X,Y)} \|x\|_X$$

This also implies:

$$\|B\| := \sup_{x \in X \setminus \{0\},\ y \in Y \setminus \{0\}} \frac{|B(y,x)|}{\|y\|_Y \|x\|_X} \leq \|T\|_{\mathcal{L}(X,Y)}$$

But on the other hand

$$\|B\| = \sup_{x \in X \setminus \{0\},\ y \in Y \setminus \{0\}} \frac{|< y, Tx >|}{\|y\|_Y \|x\|_X} \geq \sup_{x \in X \setminus \{0\},\ Tx \in Y \setminus \{0\}} \frac{|< Tx, Tx >|}{\|Tx\|_Y \|x\|_X} = \|T\|_{\mathcal{L}(X,Y)}$$

$$\Longrightarrow \quad \|B\| = \|T\|_{\mathcal{L}(X,Y)}$$

Now Corollary A.9 gives us a Riesz representation for $B(\cdot,\cdot)$, i.e. $B(y,x) =< T^* y, x >$, where $T^*(\cdot) : Y \to X$ is unique, bounded and linear with $\|T^*\|_{\mathcal{L}(Y,X)} = \|B\|$.

$$\Longrightarrow \quad B(y,x) =< T^* y, x >=< y, Tx > \quad \text{and} \quad \|T\|_{\mathcal{L}(X,Y)} = \|T^*\|_{\mathcal{L}(Y,X)}$$

□

The concept of Hilbert adjoint operators is in fact the natural generalization of the "transpose" in $I\!R^{m,n}$. This is shown by the following example:

Example 2.3.3. Let $X = I\!R^n$ and $Y = I\!R^m$. Let $A \in I\!R^{m,n}$ and define $T : I\!R^n \to I\!R^m$ by $T(x) := Ax$. Let $x \in I\!R^n,\ y \in I\!R^m$. Then

$$< y, Tx >= y^T A x = (A^T y)^T x =< A^T y, x >$$

The Hilbert-adjoint operator $T^*(\cdot)$ is unique, thus $T^*(y) = A^T y$.

The next example provides the calculation of some gradients that are needed in later chapters:

Example 2.3.4. Let $f : X \to I\!\!R$ and $c : X \to Y$ be continuously Fréchet differentiable mappings. Let $\lambda \in Y$. Our goal is to calculate the gradients of various mappings of future interest.

i) Define $g_1 : X \to I\!\!R$ by $g_1(x) :=< \lambda, c(x) >_Y$. We already showed in Example 2.2.1, that the Fréchet derivative of $g_1(\cdot)$ is given by

$$g_1'(x)\cdot =< \lambda, c'(x)\cdot >_Y$$

But $c'(x)(\cdot)$ is bounded and linear. Thus, by the last theorem, its Hilbert-adjoint operator exists and

$$g_1'(x)\cdot =< \lambda, c'(x)\cdot >=< c'(x)^*\lambda, \cdot >$$

As $\nabla g_1(x)$ is unique, we can just read from the last equation that $\nabla g_1(x) = c'(x)^*\lambda$.

ii) Now define $g_2 : X \to I\!\!R$ by

$$g_2(x) := \frac{1}{2}\|c(x)\|_Y^2 = \frac{1}{2} < c(x), c(x) >_Y$$

In example 2.2.2 we already showed that

$$g_2'(x)h =< c(x), c'(x)h >=< c'(x)^*c(x), h >$$

which implies $\nabla g_2(x) = c'(x)^*c(x)$.

iii) Finally, define $\Phi : X \times Y \times I\!\!R \to I\!\!R$ by

$$\Phi(x, \lambda, \mu) := f(x)+ < \lambda, c(x) >_Y + \frac{1}{2\mu}\|c(x)\|_Y^2$$

As ∇_x is a linear operator, we can easily compute $\nabla_x\Phi(x, \lambda, \mu)$ by using parts i) and ii):

$$\begin{aligned}\nabla_x\Phi(x, \lambda, \mu) &= \nabla_x f(x) + \nabla_x < \lambda, c(x) >_Y + \frac{1}{\mu}\nabla_x \frac{1}{2}\|c(x)\|_Y^2 \\ &= \nabla f(x) + c'(x)^*\lambda + \frac{1}{\mu}c'(x)^*c(x)\end{aligned}$$

Now we will state some basic properties of adjoints which will be referred to in the next two sections. The proof of these properties is straightforward and can be found in [16] and [23].

Theorem 2.3.5. *In addition to X and Y, let also $(Z, < \cdot, \cdot >_Z)$ be a Hilbert space.*

i) Let $T,\ S \in \mathcal{L}(X,Y)$. Then $(T+S)^* = T^* + S^*$.

ii) Let $T \in \mathcal{L}(X,Y)$. Then $(T^*)^* = T$.

iii) Let $T \in L(X,X)$ be bijective. Then T^* is also bijective and $(T^{-1})^* = (T^*)^{-1}$.

iv) Let $T \in \mathcal{L}(X,Y),\ S \in L(Y,Z)$. Then $(ST)^* = T^*S^*$.

The relations of ranges and null spaces of an operator and its adjoint are of fundamental importance for later proofs. Therefore, we introduce the following notation: For a given mapping $f : X \to Y$ let $\mathcal{R}(f)$ denote the range of f and $\mathcal{N}(f)$ the null space of f.

Theorem 2.3.6. *Let $T \in \mathcal{L}(X,Y)$ and let T^* be the corresponding Hilbert-adjoint operator. Then*

$$\mathcal{R}(T)^\perp = \mathcal{N}(T^*)$$

Proof. We prove the theorem by showing inclusions. First, let $y^* \in \mathcal{N}(T^*)$ and $y \in \mathcal{R}(T)$. Then $y = Tx$ for some $x \in X$ and

$$< y^*, y >_Y = < y^*, Tx >_Y = < T^*y^*, x >_X = < 0, x >_X = 0$$

As y was arbitrary, this implies that $y^* \in \mathcal{R}(T)^\perp$ and thus $\mathcal{N}(T^*) \subset \mathcal{R}(T)^\perp$.
Now assume $y^* \in \mathcal{R}(T)^\perp$. For every $x \in X :\ Tx \in \mathcal{R}(T)$ and therefore

$$0 = < y^*, Tx >_Y = < T^*y^*, x >_X$$

By choosing $x := T^*y^*$ we obtain $T^*y^* = 0$ which implies $\mathcal{R}(T)^\perp \subset \mathcal{N}(T^*)$. □

Lemma 2.3.7. *Let $T \in \mathcal{L}(X,Y)$ and assume that $\mathcal{R}(T)$ is closed. Then there is a constant K such that for every $y \in \mathcal{R}(T)$ there is an x satisfying $Tx = y$ and $\|x\|_X \le K\|y\|_Y$.*

Proof. T is continuous, thus $\mathcal{N}(T)$ is closed. Let now $X/\mathcal{N}(T)$ denote the quotient space consisting of all equivalence classes $[x]$ modulo $\mathcal{N}(T)$, i.e. for $x,\ y \in X$

$$x \sim y \quad :\Longleftrightarrow \quad x - y \in \mathcal{N}(T)$$

This space is a linear space with $\alpha[x] + \beta[y] := [\alpha x + \beta y]$. $\mathcal{N}(T)$ is closed, hence we know from Functional Analysis, that $X/\mathcal{N}(T)$ is a Banach space with norm

$$\|[x]\|_{X/\mathcal{N}(T)} := \inf_{n \in \mathcal{N}(T)} \|x + n\|_X$$

Now define $\bar{T} : X/\mathcal{N}(T) \to \mathcal{R}(T)$ by $\bar{T}[x] := Tx$. The specific structure of this quotient space implies that $\bar{T}$ is well defined, injective and surjective. Further

$$\begin{aligned} \bar{T}(\alpha[x] + \beta[y]) &= \bar{T}([\alpha x + \beta y]) = T(\alpha x + \beta y) \\ &= \alpha Tx + \beta Ty = \alpha\bar{T}[x] + \beta\bar{T}[y] \end{aligned}$$

Therefore $\bar{T}$ is linear. Further the definition of $\|\cdot\|_{X/\mathcal{N}(T)}$ implies that there is for every $x \in X$ an $\tilde{x} \in [x]$ such that $\frac{1}{2}\|\tilde{x}\|_X \leq \|[x]\|_{X/\mathcal{N}(T)}$. Hence,

$$\|\bar{T}[x]\|_Y = \|Tx\|_Y = \|T\tilde{x}\|_Y \leq \|T\|_{\mathcal{L}(X,Y)}\|\tilde{x}\|_X \leq \|T\|_{\mathcal{L}(X,Y)}2\|[x]\|_{X/\mathcal{N}(T)}$$

and thus $\bar{T}$ is also bounded. By assumption $\mathcal{R}(T)$ is closed. Thus $(\mathcal{R}(T), \|\cdot\|_Y)$ is a Banach space. Now the Inverse Mapping Theorem (see Appendix, Theorem A.11) implies that $\bar{T}$ has a continuous inverse $\bar{T}^{-1}$.

Finally, let $y \in \mathcal{R}(T)$ and define $[x] := \bar{T}^{-1}(y)$. Then

$$\|[x]\|_{X/\mathcal{N}(T)} = \|\bar{T}^{-1}y\|_{X/\mathcal{N}(T)} \leq \|\bar{T}^{-1}\|_{L(Y,X/\mathcal{N}(T))}\|y\|_Y$$

But $\quad \|[x]\|_{X/\mathcal{N}(T)} = \inf_{n\in\mathcal{N}(T)} \|x+n\|_X \geq \frac{\|\tilde{x}\|_X}{2} \quad$ for some $\tilde{x} \in [x]$. Thus

$$\|\tilde{x}\| \leq 2\|\bar{T}^{-1}\|_{L(Y,X/\mathcal{N}(T))}\|y\|_Y =: K\|y\|_Y$$

which proves the theorem (observe that K is independent of y). □

Theorem 2.3.8. *Let $T \in \mathcal{L}(X,Y)$ and assume that $\mathcal{R}(T)$ is closed. Then*

$$\mathcal{R}(T^*) = \big[\mathcal{N}(T)\big]^{\perp}$$

Proof. Again, we prove the theorem by showing both inclusions. First, let $x^* \in \mathcal{R}(T^*)$. Then $x^* = T^*y^*$ for some $y^* \in Y$. Now let $x \in \mathcal{N}(T)$ be arbitrary. We obtain

$$< x^*, x >_X = < T^*y^*, x >_X = < y^*, Tx >_Y = < y^*, 0 > = 0$$

Therefore, $x^* \in \big[\mathcal{N}(T)\big]^{\perp}$ and thus $\mathcal{R}(T^*) \subset \big[\mathcal{N}(T)\big]^{\perp}$.

Now let $x^* \in \big[\mathcal{N}(T)\big]^{\perp}$. Further let $y \in \mathcal{R}(T)$. We know from linear algebra, that

$$\{x \in X \ s.t. \ Tx = y\} = x_y + \mathcal{N}(T)$$

where x_y is one specific solution of $Tx = y$. Now define $f(x) := < x^*, x >_X$. Due to the orthogonality of x^* to all vectors in $\mathcal{N}(T)$, the functional $f(\cdot)$ is constant on $\{x \in X \ s.t. \ Tx = y\}$. Therefore $g : \mathcal{R}(T) \to I\!R, \ y \mapsto < x^*, x_y >_X$ where $Tx_y = y$ is well defined on $\mathcal{R}(T)$. Observe, that

$$g(y) = < x^*, x_y >_X = < x^*, x >_X \qquad \forall \ x \in \{x \in X \ s.t. \ Tx = y\}$$

According to the previous lemma, there is an $\bar{x} \in X$ satisfying $T\bar{x} = y$ and $\|\bar{x}\|_X \leq K\|y\|_Y$. Therefore,

$$|g(y)| = | < x^*, x_y >_X | = | < x^*, \bar{x} >_X | \leq \|x^*\|_X\|\bar{x}\|_X \leq \|x^*\|_X K\|y\|_Y =: c\|y\|_Y$$

Thus $g(\cdot)$ is a bounded linear functional on $\mathcal{R}(T)$. Now we can use the Hahn-Banach Theorem (see Appendix, Theorem A.4) to extend $g(\cdot)$ from $\mathcal{R}(T)$ to Y, i.e. there exists a linear functional $G : Y \to I\!R$ such that

$$G(y) = g(y) \quad \forall y \in \mathcal{R}(T), \qquad |G(y)| \leq c\|y\|_Y \quad \forall y \in Y$$

Thus $G(\cdot) \in Y^*$, hence it has a Riesz representation $y_G \in Y$ (see Appendix, Theorem A.8):

$$G(y) =< y_G, y > \quad \forall y \in Y$$

Finally, let $\tilde{x} \in X$ be arbitrary. Then $T\tilde{x} \in \mathcal{R}(T)$ and thus

$$< T^* y_G, \tilde{x} >_X =< y_G, T\tilde{x} >_Y = G(T\tilde{x}) = g(T\tilde{x}) =< x^*, x_{T\tilde{x}} >_X =< x^*, \tilde{x} >_X$$

This implies $< T^* y_G - x^*, \tilde{x} >_X = 0 \quad \forall \tilde{x} \in X$ and therefore $T^* y_G = x^*$. But this means $x^* \in \mathcal{R}(T^*)$ or $\big[\mathcal{N}(T)\big]^\perp \subset \mathcal{R}(T^*)$. □

2.4 Operator Pseudoinverse

The pseudoinverse is an important concept in optimization theory. Assume that $A \in \mathcal{L}(X, Y)$ and $y \in Y$ and consider the approximation problem

$$\min_{x \in X} \|Ax - y\|_Y$$

Define $S := \{\tilde{x} \in X \ s.t. \ \tilde{x} = argmin_{x \in X}\|Ax - y\|_Y)\}$. Then the pseudoinverse is defined as the mapping $A^+ : Y \to X$ which maps a given $y \in Y$ to the minimum-norm solution $argmin_{x \in S}\|x\|_X$ of this approximation problem. In linear algebra one usually covers the matrix theory of the pseudoinverse, i.e. in the case $X = I\!R^n$, $Y = I\!R^m$. There one can easily show, that the pseudoinverse is well defined. The question is how this theory of the matrix pseudoinverse transfers to the general Hilbert space case.

As Hilbert spaces itself are no more than a generalization of finite dimensional linear vector spaces with Euclidean norm, one can hope that the theory carries over in a satisfactory manner. In fact, especially in the case of a bounded linear operator with closed range, similar theorems can be derived with the best approximation theorem [8]. There are genuine difficulties, however, when the operator has non-closed range [19].

For our purpose, the analysis of a linear and surjective operator $A \in \mathcal{L}(X, Y)$ is sufficient. The range $\mathcal{R}(A) = Y$ is closed and one can show in this special case that (AA^*) is invertible (see next theorem) and that the pseudoinverse is given by $A^+ = A^*(AA^*)^{-1}$.

In Chapter 3 the analysis of the following problem is of particular interest: As mentioned above, let $A \in \mathcal{L}(X,Y)$ be surjective. Then consider for given $x \in X$

$$\min_{y\in Y} \|A^*y - x\|_X \tag{2.1}$$

$\mathcal{R}(A)$ is closed, therefore by Theorem 2.3.8 $\mathcal{R}(A^*) = [\mathcal{N}(A)]^\perp$ is closed. Hence the pseudoinverse theory carries over to the Hilbert space case and $(A^*)^+$ exists. Furthermore, one can easily verify with the projection theorem that the solution of (2.1) has to satisfy the normal equation

$$AA^*y = Ax$$

As mentioned above, we will show in the next theorem that in this special case AA^* is invertible and therefore by Theorem 2.3.5

$$\begin{aligned} y &= (A^*)^+ x = (AA^*)^{-1}Ax \\ &= \left[A^*\left((AA^*)^{-1}\right)^*\right]^* x = \left[A^*\left((AA^*)^*\right)^{-1}\right]^* x \\ &= \left[A^*(AA^*)^{-1}\right]^* x = (A^+)^*x \end{aligned} \tag{2.2}$$

As x was arbitrary we obtain

$$(A^*)^+ = (A^+)^* \tag{2.3}$$

This relation is of particular interest for results in later chapters. As we will see, one has to assure that $(A^*)^+$ exists even in case of a perturbed, surjective operator A. The necessary theory will be derived in this section.

Theorem 2.4.1. *Let $A \in \mathcal{L}(X,Y)$ be surjective. Then AA^* is bijective.*

Proof. We show first, that AA^* is injective. Due to linearity it is sufficient to show $\mathcal{N}(AA^*) = \{0\}$. Therefore assume that $AA^*y = 0$ for some $y \in Y$. Then

$$0 =< y, AA^*y >_Y =< A^*y, A^*y >_X = \|A^*y\|_X^2$$

and consequently $A^*y = 0$. But due to Theorem 2.3.6 $\mathcal{N}(A^*) = \mathcal{R}(A)^\perp = Y^\perp = \{0\}$ which implies $y = 0$.

Now we show that AA^* is onto, i.e. $\mathcal{R}(AA^*) = Y$. Clearly, $\mathcal{R}(AA^*) \subset \mathcal{R}(A) = Y$. Hence, it is sufficient to show $\mathcal{R}(A) \subset \mathcal{R}(AA^*)$. Therefore, let $y \in \mathcal{R}(A)$. Then $y = Ax$ for some $x \in X$. $A(\cdot)$ is continuous, thus $\mathcal{N}(A)$ is closed. By the projection theorem (see Appendix, Theorem A.7) we can decompose x in $x = x_N + x_{N^\perp}$ where $x_N \in \mathcal{N}(A)$ and $x_{N^\perp} \in \mathcal{N}(A)^\perp \overset{\text{Th. 2.3.8}}{=} \mathcal{R}(A^*)$. Thus $\exists\, \tilde{y} \in Y$ such that $A^*\tilde{y} = x_{N^\perp}$. This implies

$$AA^*\tilde{y} = Ax_{N^\perp} = A(x_N + x_{N^\perp}) = Ax = y$$

Hence, $y \in \mathcal{R}(AA^*)$ and thus $\mathcal{R}(AA^*) = \mathcal{R}(A) = Y$. □

In later chapters we will consider a mapping $c : X \to Y$ which is continuous Fréchet differentiable. Further we assume $c'(x_*)(\cdot)$ to be surjective for some $x_* \in X$. We have to assure for further analysis, that $c'(x)(\cdot)$ is surjective in a neighborhood of x_*, too. It turns out, that $c'(x)(\cdot)$ can be represented as a perturbation of $c'(x_*)(\cdot)$. The next theorem is an appropriate perturbation theorem for bounded linear operators. It is a modification of Theorem 5.11 in [27].

Theorem 2.4.2. *Let $T \in \mathcal{L}(X,Y)$ be bijective. Further let $S \in \mathcal{L}(X,Y)$ s.t. $\|ST^{-1}\|_{\mathcal{L}(Y,Y)} < 1$. Then $T+S$ is also bijective and*

$$(T+S)^{-1} = \sum_{n=0}^{\infty} (-1)^n T^{-1}(ST^{-1})^n$$

Proof. First we show that $T+S$ is injective. Let $x \in X$. Then

$$\|Sx\|_Y = \|ST^{-1}Tx\|_Y \le \|ST^{-1}\|_{\mathcal{L}(Y,Y)}\|Tx\|_Y$$

This implies

$$\begin{aligned} \|(T+S)x\|_Y &\ge \|Tx\|_Y - \|Sx\|_Y \ge \|Tx\|_Y - \|ST^{-1}\|_{\mathcal{L}(Y,Y)}\|Tx\|_Y \\ &= \|Tx\|_Y(1 - \|ST^{-1}\|_{\mathcal{L}(Y,Y)}) > 0 \quad \text{for } x \neq 0 \end{aligned}$$

Hence, $T+S$ is injective. In order to show surjectivity, we define for $p \in I\!N$

$$A_p := \sum_{n=0}^{p} (-1)^n T^{-1}(ST^{-1})^n$$

Thus $A_p : Y \to X$ and by the Inverse Mapping Theorem $A_p \in L(Y,X)$. Now we show that A_p is Cauchy. Let $m \in I\!N$

$$\begin{aligned} \|A_{p+m} - A_p\|_{\mathcal{L}(Y,X)} &= \|\sum_{n=p+1}^{p+m} (-1)^n T^{-1}(ST^{-1})^n\|_{\mathcal{L}(Y,X)} \\ &\le \sum_{n=p+1}^{p+m} \|T^{-1}\|_{\mathcal{L}(Y,X)}\|(ST^{-1})^n\|_{\mathcal{L}(Y,Y)} \\ &\le \|T^{-1}\|_{\mathcal{L}(Y,X)}\|ST^{-1}\|^{p+1}_{\mathcal{L}(Y,Y)} \sum_{n=0}^{m-1} \|ST^{-1}\|^n_{\mathcal{L}(Y,Y)} \\ &\le \|T^{-1}\|_{\mathcal{L}(Y,X)}\|ST^{-1}\|^{p+1}_{\mathcal{L}(Y,Y)} \sum_{n=0}^{\infty} \|ST^{-1}\|_{\mathcal{L}(Y,Y)} \\ &=: c\|ST^{-1}\|^{p+1}_{\mathcal{L}(Y,Y)} \to 0,\ p \to \infty \end{aligned}$$

$\mathcal{L}(Y,X)$ is a Banach space, so there exists an $A \in \mathcal{L}(Y,X)$ such that $\lim_{p\to\infty} A_p = A$. Consider now

$$\begin{aligned}(T+S)A_p &= \sum_{n=0}^{p}(-1)^n(T+S)T^{-1}(ST^{-1})^n \\ &= \sum_{n=0}^{p}(-1)^n I(ST^{-1})^n + \sum_{n=0}^{p}(-1)^n(ST^{-1})^{n+1} \\ &= I - \sum_{n=1}^{p}(-1)^{n-1}(ST^{-1})^n + \sum_{n=0}^{p}(-1)^n(ST^{-1})^{n+1} \\ &= I + (-1)^p(ST^{-1})^{p+1}\end{aligned}$$

where $I : X \to X$ denotes the identity operator. Thus we obtain

$$\|(T+S)A_p - I\|_{\mathcal{L}(Y,Y)} = \|(-1)^p(ST^{-1})^{p+1}\|_{\mathcal{L}(Y,Y)} \le \|ST^{-1}\|^{p+1}_{\mathcal{L}(Y,Y)} \to 0,\ p\to\infty$$

Hence $(T+S)A_p \to I,\ p \to \infty$. Now let $y \in Y$. Then

$$A_p y \to_{\|\cdot\|_X} Ay \ \text{ and } \ (T+S)A_p y \to_{\|\cdot\|_Y} Iy = y$$

The transformations T and S are both bounded. Thus, by the closed graph theorem (see Appendix, Theorem A.12), the graph of $T+S$ is closed. Hence $(Ay, y) \in Graph(T+S)$, i.e. $(T+S)Ay = y$. Therefore $(T+S)A = I$. In particular $\mathcal{R}(T+S) = Y$ and thus $T+S$ is bijective. As an additional result we obtain $A = (T+S)^{-1}$. □

This perturbation theorem will now be applied to the perturbed surjective operator mentioned above.

Theorem 2.4.3. *Let $D \subset X$ be open and nonempty. Let $c : D \to Y$ be continuous Fréchet differentiable on D. Further let $x_0 \in D$ and assume $c'(x_0)(\cdot)$ is surjective. Then $c'(x)(\cdot)$ is surjective in some δ-neighborhood of x_0.*

Proof. We obtain from Theorem 2.4.1, that $c'(x_0)c'(x_0)^*(\cdot)$ is bijective. We will show that $c'(x)c'(x)^*(\cdot)$ is bijective in some δ-neighborhood of x_0, which implies in particular the surjectivity of $c'(x)(\cdot)$. For $x \in D$ define $T,\ S_x : Y \to Y$ by

$$T(\cdot) := c'(x_0)c'(x_0)^*(\cdot)$$

$$S_x(\cdot) := \left[c'(x)c'(x)^* - c'(x_0)c'(x_0)^*\right](\cdot)$$

Thus $c'(x)c'(x)^* = T + S_x$, so $c'(x)c'(x)^*$ can be interpreted as a perturbation of T. Now

$$\begin{aligned}
\|S_x\|_{\mathcal{L}(Y,Y)} &= \|c'(x)c'(x)^* - c'(x_0)c'(x_0)^*\|_{\mathcal{L}(Y,Y)} \\
&= \| \left[c'(x) - c'(x_0)\right] c'(x)^* + c'(x_0) \left[c'(x)^* - c'(x_0)^*\right] \|_{\mathcal{L}(Y,Y)} \\
&\leq \|c'(x) - c'(x_0)\|_{\mathcal{L}(X,Y)} \|c'(x)^*\|_{\mathcal{L}(Y,X)} + \|c'(x_0)\|_{\mathcal{L}(X,Y)} \|c'(x)^* - c'(x_0)^*\|_{\mathcal{L}(Y,X)} \\
&\overset{\text{Th.2.3.5}}{=} \|c'(x) - c'(x_0)\|_{\mathcal{L}(X,Y)} \|c'(x)^*\|_{\mathcal{L}(Y,X)} + \|c'(x_0)\|_{\mathcal{L}(X,Y)} \| \left[c'(x) - c'(x_0)\right]^* \|_{\mathcal{L}(Y,X)} \\
&\overset{\text{Th.2.3.2}}{=} \|c'(x) - c'(x_0)\|_{\mathcal{L}(X,Y)} \|c'(x)\|_{\mathcal{L}(X,Y)} + \|c'(x_0)\|_{\mathcal{L}(X,Y)} \|c'(x) - c'(x_0)\|_{\mathcal{L}(X,Y)} \\
&\leq M \|c'(x) - c'(x_0)\|_{\mathcal{L}(X,Y)}
\end{aligned}$$

in some δ_1-neighborhood of x_0. Hence

$$\begin{aligned}
\|S_x T^{-1}\|_{\mathcal{L}(Y,Y)} &\leq \|S_x\|_{\mathcal{L}(X,Y)} \|T^{-1}\|_{\mathcal{L}(Y,Y)} \\
&\leq M \|c'(x) - c'(x_0)\|_{\mathcal{L}(X,Y)} \|T^{-1}\|_{\mathcal{L}(Y,Y)} < 1
\end{aligned}$$

for x in some δ-neighborhood, $\delta < \delta_1$. Now Theorem 2.4.2 implies, that $T + S_x = c'(x)c'(x)^*$ is bijective in this δ-neighborhood. □

From the last theorem we directly obtain the following Corollary:

Corollary 2.4.4. *Let $D \subset X$ be open and nonempty. Let $c : D \to Y$ be continuous Fréchet differentiable on D. Further let $x_0 \in D$ and assume $c'(x_0)(\cdot)$ is surjective. Then the operator-pseudoinverse $c'(x)^+(\cdot)$ exists in some δ-neighborhood of x_0 and*

$$c'(x)^+(\cdot) = c'(x)^* \left[c'(x)c'(x)^*\right]^{-1}(\cdot)$$

Finally, we show that the pseudoinverse is Lipschitz-continuous in some neighborhood:

Theorem 2.4.5. *Let $D \subset X$ be open and nonempty. Let $c : D \to Y$ be twice continuous Fréchet differentiable on D. Further let $x_0 \in D$ and assume $c'(x_0)(\cdot)$ is surjective. Then $c'(\cdot)^+$ is Lipschitz-continuous in a neighborhood of x_0.*

Proof. We know from Theorem 2.4.3, that $c'(x)(\cdot)$ is surjective in some δ-neighborhood of x_0. If we define for $x \in D$ the operators T and S_x as in the proof of Theorem 2.4.3, we obtain $\|S_x T^{-1}\|_{L(Y,Y)} < 1$ in the same neighborhood. Now define $A(x) : Y \to X$ and $B(x) : Y \to Y$ by

$$A(x)(\cdot) := c'(x)^*(\cdot) \quad \text{and} \quad B(x)(\cdot) := c'(x)c'(x)^*(\cdot)$$

Then, by Corollary 2.4.4, $c'(x)^+ = A(x)B(x)^{-1}$. The mappings $A(x)(\cdot)$ and $B(x)(\cdot)$ are both continuous and linear. Further $B(x)(\cdot)$ is bijective, so the inverse mapping theorem (see Appendix, Theorem A.11) implies that $B(x)^{-1}(\cdot)$ is also a continuous and linear operator. In order to improve the readability of this proof, we will split it up into several parts:

Part i): First we show, that $A(\cdot)$ and $B(\cdot)^{-1})$ are bounded in some neighborhood of x_0. The mapping $c(\cdot)$ is twice continuous Fréchet differentiable. This implies:

$$\|A(x)\|_{\mathcal{L}(Y,X)} = \|c'(x)^*\|_{\mathcal{L}(Y,X)} \overset{\text{Th.2.3.2}}{=} \|c'(x)\|_{\mathcal{L}(X,Y)} \leq M$$

in some neighborhood of x_0 and for some $M > 0$. Further $B(x) = T + S_x$, thus Theorem 2.4.2 implies

$$\begin{aligned}
\|B(x)^{-1}\|_{\mathcal{L}(Y,Y)} &= \|(T+S_x)^{-1}\|_{\mathcal{L}(Y,Y)} \\
&= \left\| \sum_{n=0}^{\infty} (-1)^n T^{-1} (S_x T^{-1})^n \right\|_{\mathcal{L}(Y,Y)} \\
&\leq \sum_{n=0}^{\infty} \|T^{-1}\|_{\mathcal{L}(Y,Y)} \|S_x T^{-1}\|^n_{\mathcal{L}(Y,Y)} \\
&= \|T^{-1}\|_{\mathcal{L}(Y,Y)} \cdot \frac{1}{1 - \|S_x T^{-1}\|_{\mathcal{L}(Y,Y)}} \leq \tilde{M}
\end{aligned}$$

in some neighborhood of x_0 and for some $\tilde{M} > 0$. Note that the sum converges due to $\|S_x T^{-1}\|_{\mathcal{L}(Y,Y)} < 1$.

Part ii): We will show now, that the operators $A(\cdot)$, $B(\cdot)$ and $B(\cdot)^{-1}$ are Lipschitz-continuous in a neighborhood of x_0. Due to Theorem 2.2.6 $c'(\cdot)$ is Lipschitz-continuous in some neighborhood N, thus for $x,\ y \in N$

$$\begin{aligned}
\|A(x) - A(y)\|_{\mathcal{L}(Y,X)} &= \|c'(x)^* - c'(y)^*\|_{\mathcal{L}(Y,X)} = \|\,[c'(x) - c'(y)]^*\,\|_{\mathcal{L}(Y,X)} \\
&\overset{\text{Th.2.3.2}}{=} \|c'(x) - c'(y)\|_{\mathcal{L}(X,Y)} \leq L\|x-y\|_X
\end{aligned}$$

where L denotes the Lipschitz-constant of $c'(\cdot)$. In analogy to the proof of Theorem 2.4.3 we can show that

$$\begin{aligned}
\|B(x) - B(y)\|_{\mathcal{L}(Y,Y)} &= \|c'(x)c'(x)^* - c'(y)c'(y)^*\|_{\mathcal{L}(Y,Y)} \\
&\leq \|c'(x) - c'(y)\|_{\mathcal{L}(X,Y)} \|c'(x)\|_{\mathcal{L}(X,Y)} + \|c'(y)\|_{\mathcal{L}(X,Y)} \|c'(x) - c'(y)\|_{\mathcal{L}(X,Y)} \\
&\leq \hat{M} \|c'(x) - c'(y)\|_{\mathcal{L}(X,Y)} \\
&\leq \hat{M} L \|x - y\|_X
\end{aligned}$$

in some neighborhood of x_0 and for some $\hat{M} > 0$. By part i) $B(\cdot)^{-1}$ is bounded and thus

$$\begin{aligned}
\|B(x)^{-1} - B(y)^{-1}\|_{\mathcal{L}(Y,Y)} &= \|B(x)^{-1}\,(B(y) - B(x))\,B(y)^{-1}\|_{\mathcal{L}(Y,Y)} \\
&\leq \|B(x)^{-1}\|_{\mathcal{L}(Y,Y)} \|B(y) - B(x)\|_{\mathcal{L}(Y,Y)} \|B(y)^{-1}\|_{\mathcal{L}(Y,Y)} \\
&\leq \tilde{M}^2 \hat{M} L \|x - y\|_X
\end{aligned}$$

in some neighborhood of x_0.

Part iii): Finally we show the Lipschitz-continuity of the Pseudoinverse:

$$\begin{aligned}\|c'(x)^+ - c'(y)^+\|_{L(Y,X)} = \|A(x)B(x)^{-1} &- A(y)B(y)^{-1}\|_{L(Y,X)}\\ &= \quad \left\|A(x)\left[B(x)^{-1} - B(y)^{-1}\right] + \left[A(x) - A(y)\right]B(y)^{-1}\right\|_{L(Y,X)}\\ &\leq \quad \|A(x)\|\|B(x)^{-1} - B(y)^{-1}\| + \|A(x) - A(y)\|\|B(y)^{-1}\|\\ &\overset{\text{Parts i),ii)}}{\leq} \quad \tilde{L}\|x - y\|_X\end{aligned}$$

for some $\tilde{L} > 0$ and in some neighborhood of x_0. □

2.5 Necessary Conditions

In this section we will generalize necessary and sufficient conditions for mappings $f : X \longrightarrow I\!R$ where X is as before a Hilbert space over the reals. We start with a simple necessary condition for unconstrained optimization of a Fréchet differentiable function. However, as we will see later, the extension of necessary conditions for constrained optimization to Hilbert or Banach spaces is not that simple.

Theorem 2.5.1. *Let $D \subset X$ be open and $f : D \longrightarrow I\!R$ be Fréchet differentiable at $x_* \in D$. Assume further that f has a local extremum at x_*. Then $\nabla f(x_*) = 0$.*

Proof. Let $h \in X$. Then the function $g_h : I\!R \longrightarrow I\!R$ defined by $g(\alpha) := f(x_* + \alpha h)$ has a local extremum at $\alpha = 0$. The mapping f is Fréchet-differentiable. Hence, by Theorem 2.1.7, the Gateaux-derivative of f at x_* in direction h exists and is equal to $g_h'(0)$. But g_h has a local extremum at $\alpha = 0$, therefore $g_h'(0) = 0$ by ordinary calculus. As h was arbitrary, we obtain

$$g_h'(0) = \frac{d}{d\alpha} f(x_* + \alpha h)\Big|_{\alpha=0} = f'(x_*)h = 0 \qquad \forall h \in X$$

But then the Riesz-representation of $f'(x_*)(\cdot)$ must be zero, too, that is $\nabla f(x_*) = 0$. □

Our next goal is to develop a necessary condition for an extremum of a function $f : X \longrightarrow I\!R$ subject to equality constraints $c(x) = 0, \quad c : X \longrightarrow Y$. For $X = I\!R^n$ and $Y = I\!R^m$ we already know a solution: the Lagrange Multiplier Theorem. The question is whether this idea can be transferred to arbitrary Hilbert spaces X and Y. In fact this transfer is possible - but we must generalize certain techniques we learned in advanced calculus.

If we recall how the Lagrange Multiplier Theorem was derived for subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ we see, that the Inverse Function Theorem was crucial for the proof. That is why we first aim at obtaining a Generalized Inverse Function Theorem.

Theorem 2.5.2 (Generalized Inverse Function Theorem). *Let $D \subset X$ be nonempty and open. Let $T : D \to Y$ be continuously Fréchet differentiable on D. Further let $x_0 \in D$ and assume that $T'(x_0)(\cdot) : X \to Y$ is surjective. Then there is a neighborhood $N(y_0)$ of the point $y_0 := T(x_0)$ and a constant K such that the equation $T(x) = y$ has a solution for every $y \in N(y_0)$ and the solution satisfies $\|x - x_0\|_X \leq K\|y - y_0\|_Y$.*

Proof. First we will (in analogy to the proof of Lemma 2.3.7) construct a quotient space and then a bijective and bounded linear mapping. The operator $T'(x_0)(\cdot)$ is continuous, thus $L_0 := \mathcal{N}(T'(x_0)(\cdot))$ is closed. Hence, the quotient space X/L_0, consisting of equivalence classes $[x]$ induced by the equivalence relation

$$x \sim y \quad :\Longleftrightarrow \quad x - y \in L_0$$

is a Banach space with norm

$$\|[x]\|_{X/L_0} := \inf_{m \in L_0} \|x + m\|_X$$

Now define an operator $A : X/L_0 \to Y$ by

$$A[x] := T'(x_0)x$$

This operator is well defined since $x \sim y$ implies $Ax = Ay$. In analogy to the proof of Lemma 2.3.7 we can show, that $A(\cdot)$ is linear, continuous and bijective. Thus, by the Inverse Mapping Theorem (see Appendix, Theorem A.11), $A(\cdot)$ has a linear and continuous inverse $A^{-1}(\cdot) : Y \to X/L_0$.

The idea of the proof is now to construct, for a given $y \in Y$ in some neighborhood of $y_0 := T(x_0)$, a sequence of equivalence classes $(L_n)_{n=0}^{\infty} \subset X/L_0$ and a corresponding sequence $(g_n)_{n=0}^{\infty} \subset X$ with $g_n \in L_n$ such that $x_n + g_n$ converges to a solution of $T(x) = y$. The proof is quite long, so we will split it up into several parts in order to improve readability.

Part i): To start, let $y \in Y$ be fixed and choose $g_0 := 0 \in L_0$. Then define the sequences $(L_n)_{n=0}^{\infty}$ and $(g_n)_{n=0}^{\infty}$ recursively by

$$L_n - L_{n-1} := A^{-1}\big(y - T(x_0 + g_{n-1})\big) \tag{2.4}$$

and from L_n we select g_n such that

$$\|g_n - g_{n-1}\|_X \leq 2\|L_n - L_{n-1}\|_{X/L_0} \tag{2.5}$$

This is possible since for some $\tilde{g}_n \in L_n$

$$\|L_n - L_{n-1}\|_{X/L_0} = \inf_{m \in L_0} \|\tilde{g}_n - g_{n-1} + m\|_X = \inf_{g \in L_n} \|g - g_{n-1}\|_X$$

By construction, $L_{n-1} = A^{-1}T'(x_0)g_{n-1}$ and hence equation (2.4) and the linearity of $A^{-1}(\cdot)$ imply:

$$L_n = A^{-1}\big(y - T(x_0 + g_{n-1}) + T'(x_0)g_{n-1}\big) \tag{2.6}$$

and similarly

$$L_{n-1} = A^{-1}\big(y - T(x_0 + g_{n-2}) + T'(x_0)g_{n-2}\big)$$

Therefore we obtain

$$L_n - L_{n-1} = -A^{-1}\big(T(x_0 + g_{n-1}) - T(x_0 + g_{n-2}) - T'(x_0)(g_{n-1} - g_{n-2})\big) \tag{2.7}$$

Part ii): Now define $\Gamma : X \to X/L_0$ by $\Gamma(x) := -A^{-1}\big(T(x) - T'(x_0)x\big)$. The linear mappings $A^{-1}(\cdot)$ and $T'(x_0)(\cdot)$ are Fréchet differentiable. Hence, by the chain rule (Theorem 2.1.9), $\Gamma(\cdot)$ is also Fréchet differentiable and

$$\Gamma'(x)(\cdot) = -A^{-1}\big(T'(x) - T'(x_0)\big)(\cdot)$$

By applying the generalized mean value theorem (Corollary 2.2.5) we obtain:

$$\begin{aligned}
\|L_n - L_{n-1}\| &= \|\Gamma(x_0 + g_{n-1}) - \Gamma(x_0 + g_{n-2})\|_{X/L_0} \\
&= \|\Gamma(x_0 + g_{n-2} + (g_{n-1} - g_{n-2})) - \Gamma(x_0 + g_{n-2})\|_{X/L_0} \\
&\leq \|g_{n-1} - g_{n-2}\|_X \sup_{0<\alpha<1} \|\Gamma'\big(x_0 + g_{n-2} + \alpha(g_{n-1} - g_{n-2})\big)(\cdot)\|_{\mathcal{L}(X,X/L_0)} \\
&= \|g_{n-1} - g_{n-2}\|_X \cdot \\
&\quad \cdot \sup_{0<\alpha<1} \| - A^{-1}\left(T'(x_0 + g_{n-2} + \alpha(g_{n-1} - g_{n-2})) - T'(x_0)\right)(\cdot)\|_{\mathcal{L}(X,X/L_0)} \\
&\leq \|g_{n-1} - g_{n-2}\|_X \|A^{-1}\|_{\mathcal{L}(Y,X/L_0)} \cdot \\
&\quad \cdot \sup_{0<\alpha<1} \|T'(x_0 + \alpha g_{n-1} + (1-\alpha)g_{n-2})(\cdot) - T'(x_0)(\cdot)\|_{\mathcal{L}(X,X/L_0)}
\end{aligned} \tag{2.8}$$

Part iii): Now choose $r := \frac{1}{4\|A^{-1}\|}$. The continuity of $T'(\cdot)$ implies, that there is a $\delta_r > 0$ such that $\|T'(x) - T'(x_0)\|_{\mathcal{L}(X,Y)} < r \quad \forall x \; s.t. \; \|x - x_0\|_X < \delta_r$. Now assume, that $\|g_{n-1}\|_X$, $\|g_{n-2}\|_X < \delta_r$. Then $\|\alpha g_{n-1} + (1-\alpha)g_{n-2}\|_X < \delta_r$ for $0 < \alpha < 1$ and therefore (2.8) implies:

$$\|L_n - L_{n-1}\|_{X/L_0} \leq r\|A^{-1}\|_{\mathcal{L}(Y,X/L_0)}\|g_{n-1} - g_{n-2}\|_X$$

Using (2.5) we obtain

$$\begin{aligned}
\|g_n - g_{n-1}\|_X &\leq 2\|L_n - L_{n-1}\|_{X/L_0} \\
&\leq 2r\|A^{-1}\|_{\mathcal{L}(Y,X/L_0)}\|g_{n-1} - g_{n-1}\|_X \\
&= \frac{1}{2}\|g_{n-1} - g_{n-2}\|_X
\end{aligned} \tag{2.9}$$

Part iv): Now pick $\varepsilon := \frac{1}{4\|A^{-1}\|}\delta_r$ and let $y \in Y$ such that $\|y - y_0\|_Y < \varepsilon$ We will show by induction, that $\|g_n\|_X < \delta_r \; \forall n$:

$$\begin{aligned}
\|g_1\|_X &= \|g_1 - 0\|_X \;=\; \|g_1 - g_0\|_X \\
&\overset{(2.5)}{\leq} 2\|L_1 - L_0\|_{X/L_0} \overset{\text{def. norm}}{=} 2\|L_1\|_{X/L_0} \\
&\overset{(2.6)}{=} 2\|A^{-1}(y - T(x_0 + 0) + T'(x_0)0\|_{X/L_0} \\
&= 2\|A^{-1}(y - y_0)\|_{X/L_0} \;\leq\; 2\|y - y_0\|_Y\|A^{-1}\|_{\mathcal{L}(Y,X/L_0)} \\
&< 2\varepsilon\|A^{-1}\|_{\mathcal{L}(Y,X/L_0)} \;=\; \frac{1}{2}\delta_r
\end{aligned} \tag{2.10}$$

Thus $\|g_0\|_X,\; \|g_1\|_X < \delta_r$. Now assume $\|g_0\|_X, ..., \; \|g_{n-1}\|_X < \delta_r$. We must show $\|g_n\|_X < \delta_r$:

$$\begin{aligned}
\|g_n\|_X &= \|g_1 + \sum_{i=2}^{n}(g_i - g_{i-1})\|_X \;\leq\; \|g_1\|_X + \sum_{i=2}^{n}\|g_i - g_{i-1}\|_X \\
&\overset{(2.9)}{\leq} \|g_1\|_X + \sum_{i=1}^{n-1}\|g_1 - g_0\|_X\left(\frac{1}{2}\right)^i \\
&\overset{g_0=0}{=} \|g_1\|_X\sum_{i=0}^{n-1}\left(\frac{1}{2}\right)^i \;<\; \|g_1\|_X\sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^i \\
&= 2\|g_1\|_X \overset{(2.10)}{<} \delta_r
\end{aligned} \tag{2.11}$$

Thus $\|g_n\|_X < \delta_r$ for all $n \in \mathbb{N}$.

Part v): Finally we show that the constructed sequence from Part i) converges. By Part iv) and iii) we obtain for $\|y - y_0\|_Y < \varepsilon$

$$\|g_n - g_{n-1}\|_X \leq \frac{1}{2}\|g_{n-1} - g_{n-2}\|_X \quad \forall n \in \mathbb{N}$$

Let now $m,\; k \in \mathbb{N}$. Then

$$\begin{aligned}
\|g_{m+k} - g_m\|_X &\leq \| \sum_{i=m}^{m+k-1}\|g_{i+1} - g_i\|_X \;<\; \|g_{m+1} - g_m\|_X\sum_{i=0}^{k-1}\left(\frac{1}{2}\right)^i \\
&< 2\|g_{m+1} - g_m\|_X \;\leq\; 2\left(\frac{1}{2}\right)^m\|g_1\|_X \to 0, \quad m \to \infty
\end{aligned}$$

Hence $(g_n)_{n=0}^{\infty}$ is Cauchy and thus convergent, i.e. $\exists\; g \in X$ such that $g_n \to g,\; n \to \infty$. In addition, $\|g\| \leq \delta_r$. Further (2.6) implies that L_n converges to some $L \in X/L_0$ and due to

(2.4) we obtain

$$\begin{array}{ll} & L = L + A^{-1}(y - T(x_0 + g)) \\ \Longleftrightarrow & A^{-1}y = A^{-1}T(x_0 + g) \\ \Longleftrightarrow & y = T(x_0 + g) \end{array}$$

Finally, (2.10) and (2.11) imply

$$\begin{array}{lll} \|g\|_X & \leq & 2\|g_1\|_X \ \leq \ 2 \cdot 2\|y - y_0\|_Y \|A^{-1}\|_{\mathcal{L}(Y,X/L_0)} \\ & =: & K\|y - y_0\|_Y \end{array}$$

□

Remark: If one compares this generalized inverse function theorem to the one for mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$, one realizes that our $T : X \to Y$ is not locally invertible any more. We only have the existence of a solution of the equation $T(x) = y$ in some neighborhood of y_0. But this is not surprising if one recalls that the existence of this local inverse was based on the same finite dimensions of the domain and the range. In fact, the local inverse does (in general) not exist any longer for mappings $T : \mathbb{R}^n \to \mathbb{R}^m$ where $m \neq n$.

Now, that we derived a Generalized Inverse Function Theorem, we will aim at obtaining necessary conditions for an extremum of a function $f : X \to \mathbb{R}$ subject to constraints $c(x) = 0$ where $c : X \to Y$:

Lemma 2.5.3. *Let $D \subset X$ be nonempty and open. Let $f : D \to \mathbb{R}$ and $c : D \to Y$ be continuously Fréchet differentiable on D. Further let $x_0 \in D$ and assume that $c'(x_0)(\cdot) : X \to Y$ is surjective. If f achieves a local extremum at x_0 subject to constraints $c(x) = 0$, then*

$$f'(x_0)h = 0 \quad \forall h \ \ s.t. \ \ c'(x_0)h = 0$$

Proof. Without loss of generality, let f achieve a local minimum subject to $c(x) = 0$. We will assume that there is an $h \in X$ such that $c'(x_0)h = 0$ and $f'(x_0)h \neq 0$ and lead this to a contradiction.

Define $T : X \to \mathbb{R} \times Y$ by $T(x) := (f(x), c(x))$. It is easy to prove, that $\mathbb{R} \times Y$ is a Banach space with norm $\|(r, y)\|_{\mathbb{R}\times Y} := |r| + \|y\|_Y$. The Fréchet derivative of $T(\cdot)$ at x_0 with respect to this norm is $T'(x_0)(\cdot) = (f'(x_0)(\cdot), c'(x_0)(\cdot))$. Since $f'(x_0)h \neq 0$, $f'(x_0)(\cdot)$ is linear and $c'(x_0)(\cdot)$ is onto, the mapping $T'(x_0)(\cdot)$ is surjective. Hence we can apply the generalized inverse function theorem to the mapping $T(\cdot)$ in order to show that in every δ-neighborhood of x_0 there exists an $\tilde{x}$ such that $f(\tilde{x}) < f(x_0)$:

Let $\delta > 0$ be arbitrary. Let $N(T(x_0))$ be the neighborhood given by the inverse function theorem and let K be the corresponding constant (see Theorem 2.5.2). Now choose $0 < \theta <$

$\frac{\delta}{K}$ such that $(f(x_0)-\theta, 0) \in N(T(x_0))$. Then we obtain by Theorem 2.5.2 that there exists an $\tilde{x} \in X$ such that $T(\tilde{x}) = (f(x_0)-\theta, 0)$ and

$$\|\tilde{x}-x_0\|_X \leq K \left\| \begin{pmatrix} f(x_0)-\theta \\ 0 \end{pmatrix} - \begin{pmatrix} f(x_0) \\ 0 \end{pmatrix} \right\|_{I\!R\times Y} = K|\theta| < \delta$$

Thus $\tilde{x}$ is in the δ-neighborhood of x_0, $c(\tilde{x}) = 0$ and $f(\tilde{x}) = f(x_0) - \theta < f(x_0)$ which is a contradiction to our assumption that f achieves a local minimum at x_0 subject to $c(x) = 0$. □

Finally, we state the generalized version of the Lagrange Multiplier Theorem:

Theorem 2.5.4 (Lagrange Multiplier Theorem). *Let $D \subset X$ be nonempty and open and let $f : D \to I\!R$ and $c : X \to Y$ be continuously Fréchet differentiable on D. Further let $x_0 \in D$ and assume that $c'(x_0)(\cdot) : X \to Y$ is surjective. If f achieves a local extremum at x_0 subject to constraints $c(x) = 0$, then there exists a unique $\lambda_0 \in Y$ such that the Lagrangian functional $L : X \times Y \to I\!R$ defined by $L(x,\lambda) := f(x) + <\lambda, c(x)>_Y$ is for $\lambda = \lambda_0$ stationary with respect to x at x_0, i.e.:*

$$\nabla_x L(x_0,\lambda_0) = \nabla f(x_0) + c'(x_0)^*\lambda_0 = 0$$

Proof. By applying Lemma 2.5.3 we obtain

$$f'(x_0)h = <\nabla f(x_0), h> = 0 \quad \forall h \in \mathcal{N}(c'(x_0)(\cdot))$$

Hence $\nabla f(x_0) \perp \mathcal{N}(c'(x_0)(\cdot))$. By assumption, $c'(x_0)(\cdot)$ is surjective, thus $\mathcal{R}(c'(x_0)(\cdot))$ is closed. Now Theorem 2.3.8 implies, that

$$\mathcal{R}(c'(x_0)^*(\cdot)) = [\mathcal{N}(c'(x_0)(\cdot))]^\perp$$

Therefore, $\nabla f(x_0) \in \mathcal{R}(c'(x_0)^*(\cdot))$. But $c'(x_0)^*(\cdot)$ is linear, thus $-\nabla f(x_0) \in \mathcal{R}(c'(x_0)^*(\cdot))$, i.e.

$$\begin{aligned} & \exists\, \lambda_0 \in Y \quad s.t. \quad -\nabla f(x_0) = c'(x_0)^*\lambda_0 \\ \Longleftrightarrow \quad & \nabla f(x_0) + c'(x_0)^*\lambda_0 = 0 \end{aligned}$$

We showed in the proof of Theorem 2.4.1, that $c'(x_0)^*(\cdot)$ is injective. Hence the Lagrange multiplier λ_0 is unique. □

The Lagrange Multiplier Theorem states that candidates for local minima of functionals subject to equality-constraints can be found by solving the equation

$$\nabla_x L(x,\lambda) = \nabla f(x) + c'(x)^*\lambda = 0$$

This is the concept this thesis is based on. The next chapter will be devoted to the derivation of an algorithm which can solve this equation for the unknowns x and λ.

CHAPTER 3

Basic Algorithm for Hilbert Spaces

In the next chapters we will consider the following problem: Let X and Y be Hilbert spaces over the reals. Let $f : X \to I\!R$ and $c : X \to Y$ be twice continuously Fréchet-differentiable. Then we want to find candidate minimizers for

$$\min f(x) \qquad \text{subject to} \qquad c(x) = 0, \qquad x \in X \tag{3.1}$$

Assume that $x_* \in X$ solves (3.1) and that $c'(x_*)(\cdot)$ is surjective. Then Theorem 2.5.4 states, that there exists a $\lambda_* \in Y$ such that

$$\nabla f(x_*) + c'(x_*)^* \lambda_* = 0$$

But by Theorem 2.5.1 and Example 2.3.4, part i), this is a necessary condition for

$$\min_{x \in X} L(x, \lambda_*) \tag{3.2}$$

where $L : X \times Y \to I\!R$ is the *Lagrange functional* defined by

$$L(x, \lambda) := f(x) + < \lambda, c(x) >_Y \tag{3.3}$$

In order to solve (3.2), the Lagrange multiplier $\lambda_* \in Y$ must be known. Therefore, an idea to solve (3.1) is to predict a good λ, and then solve (3.2) with the Newton-method. However, this has the disadvantage that the constraint $c(x) = 0$ is not considered if $\lambda \neq \lambda_*$.

A classical technique to improve this simple idea is the augmented Lagrange method which penalizes a violation of the constraint by augmenting the Lagrange functional (3.3). Hence we obtain a new objective function $\Phi : X \times Y \times (0, \infty) \to I\!R$ defined by

$$\Phi(x, \lambda, \mu) := f(x) + < \lambda, c(x) >_Y + \frac{1}{2\mu} \|c(x)\|_Y^2 \tag{3.4}$$

Therefore, Φ is known as the *augmented Lagrange function.* Further, for the reasons described above, λ is known as the *Lagrange multiplier estimate.* Obviously, the degree of penalty for violating the constraint is regulated by μ, which is therefore called the *penalty-parameter.*

The idea for an algorithmic concept is now to minimize a suitable sequence of augmented Lagrangian functions $\Phi(\cdot, \lambda, \mu)$. Suitable is to be understood in the sense that the Lagrange multiplier λ and the penalty parameter μ are subsequently updated depending on how much the constraint is violated.

Thus, a rough sketch of this algorithm is: Given λ_k and μ_k, solve the subproblem

$$\min_{x \in X} \Phi(x, \lambda_k, \mu_k) \tag{3.5}$$

and obtain the iterate x_k. Then adjust λ_k and μ_k.

How does one adjust the parameters λ_k and μ_k? By intention one would say that the penalty parameter must be decreased if the constraint violation is too large. But what is a good update for the Lagrange multiplier?

Our solution $(x_*, \lambda_*) \in X \times Y$ shall satisfy

$$\nabla f(x_*) + c'(x_*)^* \lambda_* = 0$$

At such a point the algorithm should terminate. However, in general

$$\|\nabla f(x_k) + c'(x_k)^* \lambda_k\|_X > 0$$

Thus our goal should be to make this error as small as possible. An appropriate rule for the update of the Lagrange multiplier would be: Choose λ_{k+1} such that it satisfies

$$\min_{\lambda \in Y} \|\nabla f(x_k) + c'(x_k)^* \lambda\|_X = \|\nabla f(x_k) + c'(x_k)^* \lambda_{k+1}\|_X$$

As pointed out on page 29 this minimization is solved by

$$\lambda(x_k) := -[c'(x_k)^*]^+ \nabla f(x_k) \overset{(2.3)}{=} -[c'(x_k)^+]^* \nabla f(x_k)$$

However, the computation of this operator pseudoinverse cannot be carried out in every iteration. As we will prove later, the first order Lagrange multiplier estimate

$$\bar{\lambda}(x_k, \lambda_k, \mu_k) := \lambda_k + \frac{1}{\mu_k} c(x_k)$$

is a good approximation of $\lambda(x_k)$ and by far easier to compute. Hence a good update λ_{k+1} for our Lagrange multiplier λ_k is $\bar{\lambda}(x_k, \lambda_k, \mu_k)$. We introduce the following notation which will be used throughout the following chapters: Define for $x \in X,\ \lambda \in Y$ and $\mu \in (0, \infty)$

$$\begin{aligned}\bar{\lambda}(x, \lambda, \mu) &:= \lambda + \frac{1}{\mu} c(x) \\ \lambda(x) &:= -[c'(x)^+]^* \nabla f(x) \end{aligned} \tag{3.6}$$

where $c'(x)^+ = c'(x)^*[c'(x)c'(x)^*]^{-1}$ as pointed out on page 29. In later sections we will further use the following identity:

$$\begin{aligned}\nabla_x \Phi(x, \lambda, \mu) &\overset{Ex.2.3.4}{=} \nabla f(x) + c'(x)^* \lambda + \frac{1}{\mu} c'(x)^* c(x) \\ &= \nabla f(x) + c'(x)^*[\lambda + \frac{1}{\mu} c(x)] \\ &= \nabla f(x) + c'(x)^* \bar{\lambda}(x, \lambda, \mu) \\ &= \nabla_x L\left(x, \bar{\lambda}(x, \lambda, \mu)\right) \end{aligned} \tag{3.7}$$

This chapter is devoted to the derivation of an augmented Lagrangian algorithm which solves (3.1) in the general Hilbert space setting. The proposed algorithm ALINF was developed by E.W. Sachs and Annick Sartenaer in [25]. As we will see in the next sections, this algorithm gives an explicit rule of how to adapt the penalty parameter from one iteration to the next one. Further it provides an approximative rule of how to choose the Lagrange multiplier update.

Algorithm ALINF solves (3.1) by generating arbitrary iterates in the Hilbert space X. This is not satisfying, because numerical computations are in general not possible in infinite-dimensional spaces. It is desirable to choose certain elements of the Hilbert space as iterates which are representable by a finite set of numbers. However, algorithm ALINF allows a certain degree of freedom for the choice of elements in the Hilbert space. This degree of freedom will be used by the second algorithm ALDISCR which we present in chapter 5. ALDISCR has ALINF as a basis and generates "nice" iterates with the properties mentioned above. Moreover, the convergence theory we derive for ALINF will also apply for ALDISCR.

In section 3.1 we will introduce algorithm ALINF. Subsequent sections are devoted to local and global convergence results for this algorithm as stated in [25].

3.1 Description of the Algorithm

The algorithm as described in [25] computes in every iteration an iterate x_k which approximately solves the subproblem

$$\min_{x \in X} \Phi(x, \lambda_k, \mu_k)$$

where the Lagrange multiplier estimate λ_k and the penalty parameter μ_k are fixed in this inner iteration. Here approximately is to be understood in the sense that

$$\|\nabla_x \Phi(x_k, \lambda_k, \mu_k)\|_X \leq w_k$$

where w_k is a suitable tolerance at iteration k. This tolerance will decrease from one iteration to the next one, so that the subproblem will be solved more and more precisely. However, the inexact solution of the subproblems will be of particular importance in later chapters when we approximate those subproblems by discretizations.

Once the iterate x_k is computed, we will test it for constraint violation. Depending on this result we will adapt the penalty parameter and the Lagrange multiplier in a suitable manner. This update is designed in such a way that the multiplier updates take over in a neighborhood of a stationary point.

The degree of freedom mentioned above is as follows: Both the test of the constraint violation and the update of the Lagrange multiplier allow a certain tolerance. These relaxations are needed in later chapters when we choose certain "nice" iterates which are representable by a finite set of numbers.

The algorithm E.W. Sachs and Annick Sartenaer propose is as follows:

Algorithm ALINF

0. Initialization: Pick an initial Lagrange multiplier estimate λ_0 and an initial penalty parameter $0 < \mu_0 < 1$. Further let $w_* \ll 1$, $\eta_* \ll 1$, $\gamma_1 < 1$, $\gamma_2 > 1$, $\tau < 1$, $\alpha_\eta < 1$ and $\beta_\eta < 1$ be strictly positive constants. Set $w_0 := \mu_0$, $\eta_0 := \mu_0^{\alpha_\eta}$ and $k := 0$.

1. Inner Iteration: Find $x_k \in X$ such that

$$\|\nabla_x \Phi(x_k, \lambda_k, \mu_k)\|_X \leq w_k \tag{3.8}$$

2. Test for convergence: If $\|\nabla_x \Phi(x_k, \lambda_k, \mu_k)\|_X \leq w_*$ and $\|c(x_k)\|_Y \leq \eta_*$, stop.

3. Updates: If

$$\|c(x_k)\|_Y \leq \gamma_1 \eta_k \tag{3.9}$$

execute step 3a. If

$$\|c(x_k)\|_Y \geq \gamma_2\eta_k \tag{3.10}$$

execute step 3b. Otherwise if

$$\gamma_1\eta_k < \|c(x_k)\|_Y < \gamma_2\eta_k \tag{3.11}$$

execute Step 3a or Step 3b. Increment k by one and go to Step 1.

3a. Update Lagrange multiplier estimate: Choose λ_{k+1} that satisfies

$$\|\lambda_{k+1} - \bar{\lambda}(x_k, \lambda_k, \mu_k)\|_Y \leq w_k \tag{3.12}$$

and set

$$\begin{aligned} \mu_{k+1} &:= \mu_k \\ w_{k+1} &:= w_k\mu_{k+1} \\ \eta_{k+1} &:= \eta_k\mu_{k+1}^{\beta_\eta} \end{aligned} \tag{3.13}$$

3b. Reduce penalty parameter: Set

$$\begin{aligned} \lambda_{k+1} &:= \lambda_k \\ \mu_{k+1} &:= \tau\mu_k \\ w_{k+1} &:= \mu_{k+1} \\ \eta_{k+1} &:= \mu_{k+1}^{\alpha_\eta} \end{aligned} \tag{3.14}$$

■

Note, that the solutions of the subproblems and the tests on the constraints get more and more precise as k tends to infinity, i.e.:

Lemma 3.1.1. $\lim_{k\to\infty} w_k = \lim_{k\to\infty} \eta_k = 0$

Proof. The adaption of the penalty parameter in step 3a and step 3b implies:

$$0 < \mu_{k+1} \leq \mu_k \leq \ldots \leq \mu_0 < 1$$

Therefore, by the sandwich theorem, $\exists \lim_{k\to\infty} \mu_k =: \mu_*$. We show that w_k converges to zero as k tends to infinity. As $0 < \mu_k < 1$ we further know by the choice of w_k that $0 < w_k < 1 \; \forall k$. In detail, the update of w_{k+1} is defined as follows:

In step 3a): $w_{k+1} \overset{(3.13)}{=} w_k\mu_{k+1} \overset{(3.13)}{=} w_k\mu_k \overset{0<w_k<1}{<} \mu_k$

In step 3b): $w_{k+1} \overset{(3.14)}{=} \mu_{k+1} \overset{(3.14)}{=} \tau\mu_k \overset{\tau<1}{<} \mu_k$

So altogether $0 < w_{k+1} < \mu_k$. Now we show the desired result by separating two cases:

Case 1: $\mu_* = 0$
Then the sandwich theorem implies $\lim_{k\to\infty} w_k = 0$.

Case 2: $\mu_* > 0$
i.e.: $\lim_{k\to\infty} \mu_k = \mu_* > 0$. We will show by contradiction, that

$$\exists\ k_0 \in I\!N \qquad \text{s.t.} \qquad \mu_{k+1} = \mu_{k_0} \qquad \forall\ k \geq k_0$$

i.e. step 3a is executed for all $k \geq k_0$. Now assume this is not the case, i.e. $\exists\ (n_l)_{l\in I\!N} \subset I\!N$ such that step 3b is executed for all iterations n_l. Without loss of generality assume that $(n_l)_{l\in I\!N}$ is the sequence of all iterates for which step 3b is executed. This implies $\mu_{n_l+1} = \tau\mu_{n_l} = \tau^{l+1}\mu_{n_0}$. But $\tau < 1$, thus $lim_{l\to\infty}\mu_{n_l+1} = 0$ which is a contradiction to $\mu_* > 0$.

So we have $\forall\ k \geq k_0 \quad w_{k+1} = w_k\mu_{k+1} = w_k\mu_{k_0}$ for some $k_0 \in I\!N$. We can show by induction that $\forall\ k \geq k_0$ and $\forall\ l \in I\!N \quad w_{k+l} = (\mu_{k_0})^l w_k$. But $0 < \mu_{k_0} < 1$ and thus $\lim_{k\to\infty} w_k = 0$.

In analogy one can show $\lim_{k\to\infty} \eta_k = 0$. The only thing that needs to be changed is the corresponding inequality: one obtains now $0 < \eta_{k+1} < \mu_k^{min\{\alpha_\eta,\beta_\eta\}}$. □

3.2 Global Convergence Analysis

We will make the following general assumptions to show that algorithm ALINF is globally convergent:

AS 1. The mappings $f : X \to I\!R$ and $c : X \to Y$ are twice continuously Fréchet differentiable.

AS 2. The iterates $(x_k)_{k\in I\!N}$ lie within a compact set $\Omega \subset X$.

AS 3. At any limit point x_* of $(x_k)_{k\in I\!N}$ the operator $c'(x_*)(\cdot)$ is surjective.

AS 4. The convergence tolerances are $w_* = \eta_* = 0$.

Remark:

i) Assumption AS2 implies that there exists at least a convergent subsequence of iterates $(x_{k_l})_{l\in I\!N}$. We will denote this subsequence from now onwards briefly with $(x_k)_{k\in\mathcal{K}}$ where $\mathcal{K} \subset I\!N$.

ii) As we want to show global convergence in this section, i.e. convergence of iterates to a stationary point, we cannot terminate algorithm ALINF when the convergence tolerance $w_* > 0$ is reached. Thus we assume for theoretical results $w_* = \eta_* = 0$ (AS4).

The following Lemma prepares our convergence results. We make the following remarks considering the assumptions of the Lemma: Note that the sequence $(\lambda_k)_{k\in\mathcal{K}}$ is unspecified. This is necessary, because algorithm ALINF allows a certain tolerance in the choice of the Lagrange multiplier updates and because it is unclear which step is executed in a specific iteration.

In addition the Lemma allows some freedom of choice for the sequences $(\mu_k)_{k\in\mathcal{K}}$ and $(w_k)_{k\in\mathcal{K}}$, because those sequences - if generated by algorithm ALINF - depend on the specific initialization and the convergent subsequence. However, note that the sequences defined in the algorithm satisfy the requirements of the Lemma (as shown in Lemma 3.1.1 and its proof).

Lemma 3.2.1. *Assume AS1 - AS4 to be valid. Let $x_* \in X$ be a limit point of $(x_k)_{k\in\mathbb{N}}$ and let $(x_k)_{k\in\mathcal{K}}$ be a subsequence which converges to x_*. Further let $\lambda(x_*)$ be defined as on page 42. Assume that $(\lambda_k)_{k\in\mathcal{K}} \subset Y$ is any sequence of vectors and that $(\mu_k)_{k\in\mathcal{K}}$ form a non-increasing sequence of scalars. Suppose further, that the iterates $(x_k)_{k\in\mathcal{K}}$ satisfy (3.8) where the w_k are positive scalar parameters which converge to zero as $k \in \mathcal{K}$ increases.*

Then there are positive constants κ_1 and κ_2 such that

$$\begin{aligned} \|\bar{\lambda}(x_k,\lambda_k,\mu_k) - \lambda(x_*)\| &\leq \kappa_1 w_k + \kappa_2\|x_k - x_*\| \\ \|\lambda(x_k) - \lambda(x_*)\| &\leq \kappa_2\|x_k - x_*\| \\ \|c(x_k)\| &\leq \kappa_1 w_k\mu_k + \mu_k\|\lambda_k - \lambda(x_*)\| + \kappa_2\mu_k\|x_k - x_*\| \end{aligned}$$

for all $k \in \mathcal{K}$ sufficiently large. Hence, the sequences $(\bar{\lambda}(x_k,\lambda_k,\mu_k))_{k\in\mathcal{K}}$ and $(\lambda(x_k))_{k\in\mathcal{K}}$ converge to $\lambda(x_*)$. Further we obtain

$$\lim_{k\in\mathcal{K}} \nabla_x\Phi(x_k,\lambda_k,\mu_k) = \nabla_x L(x_*,\lambda(x_*)) = 0$$

Proof. AS1 and AS3 imply by Theorem 2.4.3 that $c'(x)(\cdot)$ is surjective in a neighborhood of x_*. As $(x_k)_{k\in\mathcal{K}}$ converges to x_*, we have that $c'(x_k)(\cdot)$ is surjective for $k \in \mathcal{K}$ sufficiently large. Consequently, by Corollary 2.4.4, $c'(x_k)^+(\cdot)$ exists for $k \in \mathcal{K}$, k larger than some k_0.

By Theorem 2.4.5 $c'(\cdot)^+$ is Lipschitz-continuous and hence continuous in some neighborhood of x_*. Therefore, $(c'(x_k)^+)_{k\in\mathcal{K}}$ is bounded and converges to $c'(x_*)^+$. Thus

$$\|\left(c'(x_k)^+\right)^*\| = \|c'(x_k)^+\| \leq \kappa_1$$

for some $\kappa_1 > 0$. The inner iteration is terminated as soon as

$$\|\nabla_x\Phi(x_k,\lambda_k,\mu_k)\| \overset{(3.7)}{=} \|\nabla f(x_k) + c'(x_k)^*\bar{\lambda}(x_k,\lambda_k,\mu_k)\| \leq w_k \tag{3.15}$$

Combining these two inequalities we obtain

$$\begin{aligned}\|\bar\lambda(x_k,\lambda_k,\mu_k)-\lambda(x_k)\| &\overset{(3.6)}{=} \|\bar\lambda(x_k,\lambda_k,\mu_k)+\left[c'(x_k)^+\right]^*\nabla f(x_k)\| \\ &\overset{\text{see below}}{=} \|\left(c'(x_k)^+\right)^*\left[c'(x_k)^*\bar\lambda(x_k,\lambda_k,\mu_k)+\nabla f(x_k)\right]\| \\ &\leq \left\|\left(c'(x_k)^+\right)^*\right\|\cdot w_k \leq \kappa_1 w_k \qquad (3.16)\end{aligned}$$

where the second equality holds due to

$$(c'(x_k)^+)^*c'(x_k)^* \overset{(2.2)}{=} [c'(x_k)c'(x_k)^*]^{-1}c'(x_k)c'(x_k)^* = I$$

As a result of AS1 $\nabla f(\cdot)$ is continuous (see remark iii after the definition of the gradient). Hence $\|\nabla f(x_k)\|$ is bounded for $k\in\mathcal{K}$. Further Theorem 2.2.6 implies that $f'(\cdot)$ is Lipschitz continuous in a neighborhood of x_*. Combining all these facts, we obtain by the Lipschitz-continuity of $c'(\cdot)^+$:

$$\begin{aligned}&\|\lambda(x_k)-\lambda(x_*)\| = \\ &\overset{(3.6)}{=} \|[c'(x_k)^+]^*\nabla f(x_k)-[c'(x_*)^+]^*\nabla f(x_*)\| \\ &\leq \|[c'(x_k)^+]^*-[c'(x_*)^+]^*\|\|\nabla f(x_k)\|+\|[c'(x_*)^+]^*\|\|\nabla f(x_k)-\nabla f(x_*)\| \\ &\leq \kappa_2\|x_k-x_*\| \qquad (3.17)\end{aligned}$$

for $k\in\mathcal{K}$ sufficiently large and for some $\kappa_2>0$. By combining (3.16) and (3.17) we obtain

$$\begin{aligned}\|\bar\lambda(x_k,\lambda_k,\mu_k)-\lambda(x_*)\| &\leq \|\bar\lambda(x_k,\lambda_k,\mu_k)-\lambda(x_k)\|+\|\lambda(x_k)-\lambda(x_*)\| \\ &\leq \kappa_1 w_k+\kappa_2\|x_k-x_*\| \qquad (3.18)\end{aligned}$$

Hence, by (3.17), (3.18) and the assumption that w_k converges to zero for $k\in\mathcal{K}$, we have that $(\lambda(x_k))_{k\in\mathcal{K}}$ and $(\bar\lambda(x_k,\lambda_k,\mu_k))_{k\in\mathcal{K}}$ converge to $\lambda(x_*)$. Thus we can conclude due to the continuity of $\nabla f(\cdot)$

$$\begin{aligned}\nabla_x\Phi(x_k,\lambda_k,\mu_k) &= \nabla f(x_k)+c'(x_k)^*\bar\lambda(x_k,\lambda_k,\mu_k)\to \\ &\to_{k\in\mathcal{K}} \nabla f(x_*)+c'(x_*)^*\lambda(x_*)=\nabla_x L(x_*,\lambda(x_*))\end{aligned}$$

where $L(\cdot,\cdot)$ is the Lagrange functional as defined in (3.3). Then (3.15) implies due to $\lim_{k\in\mathcal{K}} w_k=0$

$$0=\lim_{k\in\mathcal{K}}\nabla_x\Phi(x_k,\lambda_k,\mu_k)=\nabla_x L(x_*,\lambda(x_*))=\nabla f(x_*)+c'(x_*)^*\lambda(x_*)$$

By definition $\bar\lambda(x_k,\lambda_k,\mu_k)=\lambda_k+\frac{1}{\mu_k}c(x_k)$ which is equivalent to

$$c(x_k) = \mu_k\left(\bar\lambda(x_k,\lambda_k,\mu_k)-\lambda_k\right)=\mu_k\left\{\left(\bar\lambda(x_k,\lambda_k,\mu_k)-\lambda(x_*)\right)+(\lambda(x_*)-\lambda_k)\right\}$$

Taking norms and using (3.18) we obtain:

$$\|c(x_k)\| \leq \mu_k \kappa_1 w_k + \mu_k \kappa_2 \|x_k - x_*\| + \mu_k \|\lambda(x_*) - \lambda_k\|$$

□

If $c(x_*) = 0$, then the lemma would imply that x_* is a Kuhn Tucker point for problem (3.1). Note that, although $w_k \to 0$ and $x_k \to x_*$ for $k \in \mathcal{K}$, we cannot deduce $c(x_*) = 0$ by using the inequality for $\|c(x_k)\|$ yet. We know that $\bar{\lambda}(x_k, \lambda_k, \mu_k)$ converges to $\lambda(x_*)$ for $k \in \mathcal{K}$, but we do not know how λ_k itself behaves.

Although λ_{k+1} satisfies (3.12) in step 3a, we cannot make any direct conclusions for the sequence $(\lambda_k)_{k\in\mathcal{K}}$. The reason for this is that we do not know when step 3a is executed, if it is executed infinitely often and how the choice of λ_{k+1} in step 3b affects the convergence of the sequence. In fact, we will need some additional assumptions to show that the Lagrange multiplier estimate λ_k converges to $\lambda(x_*)$, too. This will be one of our goals in the next section.

However, one can hope that λ_k does not behave too badly so that we can derive $c(x_*) = 0$. This would be sufficient to show that our iterates $(x_k)_{k\in\mathcal{K}}$ converge to the Kuhn Tucker point x_* which would prove the global convergence of algorithm ALINF. That λ_k does in fact not behave too badly is shown by the following lemma whose proof is based on the specific structure of algorithm ALINF.

Lemma 3.2.2. *Let $(x_k)_{k\in I\!N}$, $(\lambda_k)_{k\in I\!N}$, $(\eta_k)_{k\in I\!N}$ and $(w_k)_{k\in I\!N}$ be sequences generated by algorithm ALINF. Further assume that step 3b is executed infinitely often ($\lim_{k\to\infty} \mu_k = 0$). Then the product $\mu_k \|\lambda_k\|$ converges to zero.*

Proof. Let $\mathcal{S} := \{k_0,\ k_1,\ k_2, \dots\}$ be the set of indices of the iterations in which step 3b is executed and for which

$$\mu_k \leq \left(\frac{1}{2}\right)^{1/\beta_\eta} \leq \frac{1}{2} \tag{3.19}$$

This is possible since $(\mu_k)_{k\in I\!N}$ is a non-increasing sequence with limit zero. Now we want to express the Lagrange multiplier estimates between two successive iterations indexed in the set $\mathcal{S}$. Algorithm ALINF does not give an explicit rule how to choose λ_{k+1} in step 3a, so we introduce the following notation:

$$e_k := \lambda_{k+1} - \bar{\lambda}(x_k, \lambda_k, \mu_k)$$

Using this notation for iteration $k_i + j$, where $k_i < k_i + j \leq k_{i+1}$ and $k_i,\ k_{i+1} \in \mathcal{S}$, one can easily verify by induction that

$$\lambda_{k_i+j} = \lambda_{k_i} + \sum_{l=1}^{j-1} \left(\frac{c(x_{k_i+l})}{\mu_{k_i+l}} + e_{k_i+l} \right) \tag{3.20}$$

Further we obtain for these iterations

$$\mu_{k_i+j} = \mu_{k_i+1} = \tau\mu_{k_i} \tag{3.21}$$

We distinguish the following cases:

Case 1: $k_i + 1 < k_{i+1}$

Let j be such that $k_i < k_i + j \leq k_{i+1}$. Then for $k_i + l,\ 1 \leq l < j$ either (3.9) or (3.11) is fulfilled, so $\|c(x_{k_i+l})\| < \gamma_2\eta_{k_i+l}$ must hold. Thus by the recursive definition of η_k in (3.13)

$$\|c(x_{k_i+l})\| < \gamma_2\eta_{k_i+l} = \gamma_2\eta_{k_i+1}\mu_{k_i+1}^{\beta_\eta(l-1)} \overset{\text{step 3b}}{=} \gamma_2\mu_{k_i+1}^{\beta_\eta(l-1)+\alpha_\eta} = \gamma_2\mu_{k_{i+1}}^{\beta_\eta(l-1)+\alpha_\eta} \tag{3.22}$$

Further we obtain from (3.12) and the recursive definition of w_k in (3.13)

$$\|e_{k_i+l}\| \leq w_{k_i+l} = w_{k_i+1}\mu_{k_i+1}^{l-1} \overset{\text{step 3b}}{=} \mu_{k_i+1}^{l} = \mu_{k_{i+1}}^{l} \tag{3.23}$$

Thus we get by (3.20) and inequalities (3.22) and (3.23):

$$\begin{aligned}
\|\lambda_{k_i+j}\| &\leq \|\lambda_{k_i}\| + \sum_{l=1}^{j-1}\left(\frac{\|c(x_{k_i+l})\|}{\mu_{k_i+l}} + \|e_{k_i+l}\|\right) \\
&\overset{\mu_{k_i+l}=\mu_{k_{i+1}}}{\leq} \|\lambda_{k_i}\| + \gamma_2\mu_{k_{i+1}}^{\alpha_\eta-1}\sum_{l=1}^{j-1}\mu_{k_{i+1}}^{\beta_\eta(l-1)} + \sum_{l=1}^{j-1}\mu_{k_{i+1}}^{l} \\
&\overset{\text{geom. series}}{\leq} \|\lambda_{k_i}\| + \gamma_2\mu_{k_{i+1}}^{\alpha_\eta-1}\frac{1}{1-\mu_{k_{i+1}}^{\beta_\eta}} + \mu_{k_{i+1}}\frac{1}{1-\mu_{k_{i+1}}} \\
&\overset{(3.19)}{\leq} \|\lambda_{k_i}\| + 2\gamma_2\mu_{k_{i+1}}^{\alpha_\eta-1} + 2\mu_{k_{i+1}}
\end{aligned}$$

Thus we obtain by (3.21):

$$\begin{aligned}
\mu_{k_i+j}\|\lambda_{k_i+j}\| &\leq \tau\mu_{k_i}\|\lambda_{k_i}\| + 2\gamma_2\mu_{k_{i+1}}^{\alpha_\eta} + 2\mu_{k_{i+1}}^{2} \\
&\overset{0<\alpha_\eta<1}{\leq} \tau\mu_{k_i}\|\lambda_{k_i}\| + 2(\gamma_2+1)\mu_{k_{i+1}}^{\alpha_\eta}
\end{aligned} \tag{3.24}$$

Case 2: $k_i + 1 = k_{i+1}$. Then $\mu_{k_i+1} = \mu_{k_{i+1}} = \tau\mu_{k_i}$ and $\lambda_{k_i+1} = \lambda_{k_i}$ and hence

$$\mu_{k_{i+1}}\|\lambda_{k_{i+1}}\| = \tau\mu_{k_i}\|\lambda_{k_i}\| \leq \tau\mu_{k_i}\|\lambda_{k_i}\| + 2(\gamma_2+1)\mu_{k_{i+1}}^{\alpha_\eta}$$

Thus, if we pick in (3.24) $k_i + j = k_{i+1}$, then due to case 1 and 2 the inequality

$$\mu_{k_{i+1}} \|\lambda_{k_{i+1}}\| \leq \tau \mu_{k_i} \|\lambda_{k_i}\| + 2(\gamma_2 + 1)\mu_{k_{i+1}}^{\alpha_\eta} \tag{3.25}$$

holds for all $i \in I\!N$. We will show now that $\mu_{k_i}\|\lambda_{k_i}\|$ converges to zero as i increases. Define

$$\delta_i := \mu_{k_i}\|\lambda_{k_i}\| \qquad \text{and} \qquad \rho_i := 2(\gamma_2 + 1)\mu_{k_i}^{\alpha_\eta}$$

Then by (3.21) $\rho_{i+1} = \tau^{\alpha_\eta}\rho_i$ and hence by (3.25) $\delta_{i+1} \leq \tau\delta_i + \tau^{\alpha_\eta}\rho_i$. Therefore, one can show by induction

$$\begin{aligned} 0 \leq \quad \delta_i \quad & \leq \quad \tau^i \delta_0 + \rho_0 (\tau^{\alpha_\eta})^i \sum_{l=0}^{i-1} (\tau^{1-\alpha_\eta})^l \\ & \overset{0<\alpha_\eta<1}{\leq} \quad \tau^i \delta_0 + \rho_0 (\tau^{\alpha_\eta})^i \frac{1}{1 - \tau^{1-\alpha_\eta}} \end{aligned}$$

Thus we obtain by the sandwich theorem

$$0 = \lim_{i\to\infty} \delta_i = \lim_{i\to\infty} \mu_{k_i}\|\lambda_{k_i}\|$$

or equivalently $(\mu_k\|\lambda_k\|) \to_{k\in\mathcal{S}} 0$. To show now that the whole sequence $(\mu_k\|\lambda_k\|)_{k\in I\!N}$ converges to zero, we pick $m \in I\!N$ large enough such that $\mu_{k_i}\|\lambda_{k_i}\| < \varepsilon \quad \forall\ i \geq m$. Then $\forall\ k > k_m, k \notin \mathcal{S}$ we can uniquely represent k by $k = k_i + j$ with $i \geq m$ and $k_i < k_i + j < k_{i+1}$. Hence (3.24) holds and due to $\mu_k \to 0$ we have

$$\mu_k\|\lambda_k\| = \mu_{k_i+j}\|\lambda_{k_i+j}\| < \tilde{\varepsilon}$$

□

This result will be used now to prove the global convergence property of algorithm ALINF:

Theorem 3.2.3 (Global Convergence Theorem). *Let AS1 - AS4 be valid. Further let x_* be any limit point of the sequence $(x_k)_{k\in I\!N}$ generated by algorithm ALINF and let $(x_k)_{k\in\mathcal{K}}$ be a subsequence whose limit is x_*. Define $\lambda(x_*)$ as on page 42. Then x_* is a Kuhn Tucker point (first order stationary point), and $\lambda(x_*)$ is the corresponding Lagrange multiplier, i.e.:*

$$\nabla_x L(x_*, \lambda(x_*)) = \nabla f(x_*) + c'(x_*)^*\lambda(x_*) = 0 \ , \qquad c(x_*) = 0$$

Further the sequences $(\bar{\lambda}(x_k, \lambda_k, \mu_k))_{k\in\mathcal{K}}$ and $(\lambda(x_k))_{k\in\mathcal{K}}$ converge to $\lambda(x_*)$ and the gradients $\nabla_x\Phi(x_k, \lambda_k, \mu_k)$ converge to $\nabla_x L(x_*, \lambda(x_*)) = 0$ for $k \in \mathcal{K}$.

Proof. The assumptions of the theorem fulfill those of Lemma 3.2.1. In order to prove the theorem, we only need to show that $c(x_*) = 0$. We distinguish two cases:

Case 1: Step 3b is executed only finitely often.
Hence step 3a is executed in every iteration for k sufficiently large. Thus by (3.9) and (3.11) $\|c(x_k)\| < \gamma_2 \eta_k$ for all $k \in \mathcal{K}$ greater some k_0. But then Lemma 3.1.1 implies that $c(x_k)$ converges to zero for $k \in \mathcal{K}$ which means by the continuity of $c(\cdot)$ that $c(x_*) = 0$.

Case 2: Step 3b is executed infinitely often ($\lim_{k\to\infty} \mu_k = 0$). Then by Lemma 3.2.1

$$\|c(x_k)\| \le \kappa_1 w_k \mu_k + \mu_k \|\lambda_k\| + \mu_k \|\lambda(x_*)\| + \kappa_2 \mu_k \|x_k - x_*\|$$

But now, by Lemma 3.2.2 and Lemma 3.1.1, all summands on the right hand side converge to zero for $k \in \mathcal{K}$ which establishes the desired result. □

3.3 Local Convergence Analysis

In this section we will analyze the local convergence behavior of algorithm ALINF. For the derivation of these results we need the following additional assumptions.

AS 5. The second Fréchet derivatives of the functions f and c are Lipschitz-continuous at any limit point x_* of the sequence of iterates $(x_k)_{k\in I\!N}$.

AS 6. Suppose that (x_*, λ_*) is a Kuhn-Tucker point for problem (3.1). Then we assume that the operator $A : X \times Y \to \mathcal{L}(X, I\!R) \times Y$ defined by

$$A \begin{pmatrix} x \\ \lambda \end{pmatrix} := \begin{pmatrix} \left[\frac{d^2}{dx^2} L(x_*, \lambda_*)\right] x \cdot + < \lambda, c'(x_*) \cdot > \\ c'(x_*) x \end{pmatrix}$$

has a continuous inverse.

Remark: The operator $A(\cdot)$ in AS 6 is clearly a linear operator. Thus $A^{-1}(\cdot)$ is also linear and as $A^{-1}(\cdot)$ is continuous its norm $\|A^{-1}\| =: M$ is finite. Further note that $\mathcal{L}(X, I\!R) \times Y$ is a normed linear space with norm

$$\left\| \begin{pmatrix} g \\ y \end{pmatrix} \right\|_{\mathcal{L}(X,I\!R)\times Y} := \|g\|_{\mathcal{L}(X,I\!R)} + \|y\|_Y$$

In Theorem 3.2.3 we proved under appropriate assumptions, that $\lambda(x_k)$ converges to $\lambda(x_*)$ and that then $(x_*, \lambda(x_*))$ is a Karush Kuhn Tucker point for problem (3.1). Hence we will speak from now onwards of the limit point x_* and its corresponding Lagrange multiplier λ_*.

For brevity we will also occasionally use a different notation for the first order Lagrange multiplier update $\bar{\lambda}(x_k, \lambda_k, \mu_k)$. We will write $\bar{\lambda}_k := \bar{\lambda}(x_k, \lambda_k, \mu_k)$.

Our first goal in this section is to show that the penalty parameter μ_k is bounded away from zero. We will prove this later by contradiction. The next two lemmas prepare this contradiction proof.

Lemma 3.3.1. *Assume that AS 1 - AS 4 hold. Let $(x_k)_{k\in\mathbb{N}}$ be the sequence of iterates generated by algorithm ALINF. Further let $(x_k)_{k\in\mathcal{K}}$ be a subsequence which converges to the limit point x_* with corresponding Lagrange multiplier λ_* at which AS 5 and AS 6 hold. Assume furthermore that $\lim_{k\to\infty}\mu_k = 0$. Then there are positive constants $\bar{\mu} < 1$, κ_3, κ_4, κ_5, κ_6 and an integer k_1 such that $\forall k \geq k_1,\ k \in \mathcal{K}$*

$$\begin{aligned}
\|x_k - x_*\| &\leq \kappa_3 w_k + \kappa_4 \mu_k \|\lambda_k - \lambda_*\| \\
\|\bar{\lambda}(x_k, \lambda_k, \mu_k) - \lambda(x_*)\| &\leq \kappa_5 w_k + \kappa_6 \mu_k \|\lambda_k - \lambda_*\| \\
\|c(x_k)\| &\leq \kappa_5 w_k \mu_k + \mu_k(1 + \kappa_6 \mu_k)\|\lambda_k - \lambda_*\| \\
\mu_k &\leq \bar{\mu} < 1
\end{aligned}$$

Proof. Observe that our assumptions include those of Theorem 3.2.3. In order to improve readability, we split the proof up into several parts.

Part i): In this first part we will transform some expressions for further reference. Using (3.7) in operator notation, we obtain:

$$\begin{aligned}
\left(\frac{d}{dx}\Phi(x_k, \lambda_k, \mu_k)\right)\cdot &= f'(x_k)\cdot + < c'(x_k)^*\lambda_k, \cdot > + \frac{1}{\mu_k} < c'(x_k)^* c(x_k), \cdot > \\
&= f'(x_k)\cdot + < c'(x_k)^*(\lambda_k + \frac{1}{\mu_k}c(x_k)), \cdot > \\
&= f'(x_k)\cdot + < c'(x_k)^*\bar{\lambda}_k, \cdot > \\
&= f'(x_k)\cdot + < \bar{\lambda}_k, c'(x_k)\cdot > \\
&= \frac{d}{dx}L(x_k, \bar{\lambda}_k)\cdot \qquad (3.26)
\end{aligned}$$

Applying Taylor's theorem 2.2.4 around x_* we obtain

$$\frac{d}{dx}L(x_k, \bar{\lambda}_k)\cdot = \frac{d}{dx}L(x_*, \bar{\lambda}_k)\cdot + \left[\frac{d^2}{dx^2}L(x_*, \bar{\lambda}_k)\right](x_k - x_*)\cdot + r_1(x_k, x_*, \bar{\lambda}_k) \qquad (3.27)$$

where

$$\|r_1(x_k, x_*, \bar{\lambda}_k)\| \leq \|\frac{d^2}{dx^2}L(x_* + \theta(x_k - x_*), \bar{\lambda}_k)\cdot\cdot - \frac{d^2}{dx^2}L(x_*, \bar{\lambda}_k)\cdot\cdot\|\|x_k - x_*\| \qquad (3.28)$$

for some $\theta \in (0,1)$. By Example 2.2.1 we can rewrite this as

$$\begin{aligned} &\| \frac{d^2}{dx^2} L(x_* + \theta(x_k - x_*), \bar{\lambda}_k) \cdot \cdot - \frac{d^2}{dx^2} L(x_*, \bar{\lambda}_k) \cdot \cdot \| \\ &\quad = \quad \| f''(x_* + \theta(x_k - x_*)) \cdot \cdot + < \bar{\lambda}_k, c''(x_* + \theta(x_k - x_*)) \cdot \cdot > \\ &\qquad\qquad - f''(x_*) \cdot \cdot - < \bar{\lambda}_k, c''(x_*) \cdot \cdot > \| \\ &\quad \le \quad \| f''(x_* + \theta(x_k - x_*)) \cdot \cdot - f''(x_*) \cdot \cdot \| \\ &\qquad\qquad + \| \bar{\lambda}_k \| \| c''(x_* + \theta(x_k - x_*)) \cdot \cdot - c''(x_*) \cdot \cdot \| \end{aligned}$$

But by the global convergence theorem in the last section $\bar{\lambda}_k$ converges to $\lambda(x_*) = \lambda_*$ and hence we obtain by the Lipschitz continuity of $f''(\cdot)$ and $c''(\cdot)$ at x_*

$$\| \frac{d^2}{dx^2} L(x_* + \theta(x_k - x_*), \bar{\lambda}_k) \cdot \cdot - \frac{d^2}{dx^2} L(x_*, \bar{\lambda}_k) \cdot \cdot \| \le \kappa_7 \| x_k - x_* \|$$

for some $\kappa_7 > 0$. Therefore, by (3.28)

$$\| r_1(x_k, x_*, \bar{\lambda}_k) \| \le \kappa_7 \| x_k - x_* \|^2 \tag{3.29}$$

We continue to modify expression (3.27). First note that

$$\begin{aligned} \frac{d}{dx} L(x_*, \bar{\lambda}_k) \cdot \quad &= \quad f'(x_*) \cdot + < \bar{\lambda}_k, c'(x_*) \cdot > \\ &= \quad \frac{d}{dx} L(x_*, \lambda_*) + < \bar{\lambda}_k - \lambda_*, c'(x_*) \cdot > \end{aligned} \tag{3.30}$$

and second

$$\begin{aligned} &\left[\frac{d^2}{dx^2} L(x_*, \bar{\lambda}_k) \right] (x_k - x_*) \cdot \\ &\qquad \overset{\text{ex.2.2.1}}{=} \quad f''(x_*)(x_k - x_*) \cdot + < \bar{\lambda}_k, c''(x_*)(x_k - x_*) \cdot > \\ &\qquad = \quad \left[\frac{d^2}{dx^2} L(x_*, \lambda_*) \right] (x_k - x_*) \cdot + < \bar{\lambda}_k - \lambda_*, c''(x_*)(x_k - x_*) \cdot > \end{aligned} \tag{3.31}$$

Inserting (3.30) and (3.31) in (3.27) we obtain

$$\begin{aligned} \frac{d}{dx} L(x_k, \bar{\lambda}_k) \cdot \quad &= \quad \frac{d}{dx} L(x_*, \lambda_*) \cdot + \left[\frac{d^2}{dx^2} L(x_*, \lambda_*) \right] (x_k - x_*) \cdot + \\ &\qquad + < \bar{\lambda}_k - \lambda_*, c'(x_*) \cdot > + r_1(x_k, x_*, \bar{\lambda}_k) + r_2(x_k, x_*, \bar{\lambda}_k, \lambda_*) \end{aligned} \tag{3.32}$$

where $r_2(x_k, x_*, \bar{\lambda}_k, \lambda_*) := < \bar{\lambda}_k - \lambda_*, c''(x_*)(x_k - x_*) \cdot >$ and

$$\| r_2(x_k, x_*, \bar{\lambda}_k, \lambda_*) \| \le \| \bar{\lambda}_k - \lambda_* \| \cdot \| x_k - x_* \| \kappa_8 \tag{3.33}$$

for some $\kappa_8 > 0$. Moreover, using Taylor's theorem again, along with the fact that theorem 3.2.3 ensures $c(x_*) = 0$, we obtain

$$c(x_k) = c'(x_*)(x_k - x_*) + r_3(x_k, x_*) \tag{3.34}$$

where we can use the same arguments as earlier in this proof to show that

$$\|r_3(x_k, x_*)\| \leq \kappa_9 \|x_k - x_*\|^2 \tag{3.35}$$

for some $\kappa_9 > 0$.

Part ii): Combining (3.26), (3.32) and (3.34), we obtain by using the notation of AS 6:

$$A \begin{pmatrix} x_k - x_* \\ \bar{\lambda}_k - \lambda_* \end{pmatrix} = \begin{pmatrix} \frac{d}{dx}\Phi(x_k, \lambda_k, \mu_k) \cdot - \frac{d}{dx} L(x_*, \lambda_*)\cdot \\ c(x_k) \end{pmatrix} - \begin{pmatrix} r_1 + r_2 \\ r_3 \end{pmatrix} \tag{3.36}$$

Note that by Theorem 3.2.3 $\frac{d}{dx} L(x_*, \lambda_*)\cdot = 0$ and hence

$$A \begin{pmatrix} x_k - x_* \\ \bar{\lambda}_k - \lambda_* \end{pmatrix} = \begin{pmatrix} \frac{d}{dx}\Phi(x_k, \lambda_k, \mu_k)\cdot \\ c(x_k) \end{pmatrix} - \begin{pmatrix} r_1 + r_2 \\ r_3 \end{pmatrix} \tag{3.37}$$

Now we conclude by (3.29), (3.33), (3.35) and Lemma 3.2.1 that

$$\begin{aligned} \left\| \begin{pmatrix} r_1 + r_2 \\ r_3 \end{pmatrix} \right\| &\leq \kappa_7 \|x_k - x_*\|^2 + \kappa_8 \|\bar{\lambda}_k - \lambda_*\| \cdot \|x_k - x_*\| + \kappa_9 \|x_k - x_*\|^2 \qquad (3.38) \\ &\leq (\kappa_7 + \kappa_9)\|x_k - x_*\|^2 + \kappa_8(\kappa_1 w_k + \kappa_2 \|x_k - x_*\|)\|x_k - x_*\| \\ &= (\kappa_7 + \kappa_9 + \kappa_8\kappa_2)\|x_k - x_*\|^2 + \kappa_8\kappa_1 w_k \|x_k - x_*\| \\ &=: \kappa_{10}\|x_k - x_*\|^2 + \kappa_{11} w_k \|x_k - x_*\| \end{aligned}$$

Further we have by Lemma 3.2.1

$$\|c(x_k)\| \leq \kappa_1 w_k \mu_k + \mu_k \|\lambda_k - \lambda_*\| + \kappa_2 \mu_k \|x_k - x_*\|$$

Hence the inner iteration termination criterion (3.8) implies

$$\begin{aligned} \left\| \begin{pmatrix} \frac{d}{dx}\Phi(x_k, \lambda_k, \mu_k)\cdot \\ c(x_k) \end{pmatrix} \right\|_{\mathcal{L}(X, I\!R) \times Y} &= \left\| \begin{pmatrix} \nabla_x \Phi(x_k, \lambda_k, \mu_k) \\ c(x_k) \end{pmatrix} \right\|_{X \times Y} = \\ &\leq w_k + \kappa_1 w_k \mu_k + \mu_k \|\lambda_k - \lambda_*\| + \kappa_2 \mu_k \|x_k - x_*\| \qquad (3.39) \end{aligned}$$

By AS 6, the operator $A(\cdot)$ in (3.37) is invertible with $\|A^{-1}\| =: M$ and hence we obtain by combining (3.37), (3.38) and (3.39)

$$\begin{aligned}\left\|\begin{pmatrix} x_k - x_* \\ \bar{\lambda}_k - \lambda_* \end{pmatrix}\right\| \;\leq\;& M(w_k + \kappa_1 w_k \mu_k + \mu_k\|\lambda_k - \lambda_*\| + \kappa_2\mu_k\|x_k - x_*\| \\ & +\kappa_{10}\|x_k - x_*\|^2 + \kappa_{11}w_k\|x_k - x_*\|) \qquad (3.40)\end{aligned}$$

Note that by Lemma 3.1.1 $\lim_{k\to\infty} w_k = 0$ and by assumption $\lim_{k\to\infty}\mu_k = 0$. Hence we may choose $k_0 \in \mathcal{K}$ such that $\forall k \geq k_0$

$$w_k \leq \frac{1}{4M\kappa_{11}} \qquad \text{and} \qquad \mu_k \leq \bar{\mu} := min\left\{\mu_0, \frac{1}{4M\kappa_2}\right\}$$

Recall that $\mu_0 < 1$ and thus $\bar{\mu} < 1$. Then $\forall k \geq k_0,\ k \in \mathcal{K}$ we have $\mu_k < 1$ and by (3.40)

$$\begin{aligned}\|x_k - x_*\| \;&\leq\; \left\|\begin{pmatrix} x_k - x_* \\ \bar{\lambda}_k - \lambda_* \end{pmatrix}\right\| \\ &\leq\; \frac{1}{4}\|x_k - x_*\| + \frac{1}{4}\|x_k - x_*\| + M\big[\kappa_{10}\|x_k - x_*\|^2 + \\ &\qquad +\mu_k\|\lambda_k - \lambda_*\| + w_k(1+\kappa_1\mu_k)\big] \\ &\overset{\kappa_{12}:=1+\kappa_1}{\leq}\; \frac{1}{2}\|x_k - x_*\| + M\left[\kappa_{10}\|x_k - x_*\|^2 + \mu_k\|\lambda_k - \lambda_*\| + \kappa_{12}w_k\right] \qquad (3.41)\end{aligned}$$

Now, due to $x_k \to_{k\in\mathcal{K}} x_*$ we can choose k large enough such that $\|x_k - x_*\| \leq 1/(4M\kappa_{10})$. Hence we obtain by (3.41)

$$\begin{aligned}\|x_k - x_*\| \;&\leq\; 2M\left[\kappa_{10}\|x_k - x_*\|^2 + \mu_k\|\lambda_k - \lambda_*\| + \kappa_{12}w_k\right] \\ &\leq\; \frac{1}{2}\|x_k - x_*\| + 2M\left[\mu_k\|\lambda_k - \lambda_*\| + \kappa_{12}w_k\right] \\ \Longleftrightarrow \quad \|x_k - x_*\| \;&\leq\; 4M\left[\mu_k\|\lambda_k - \lambda_*\| + \kappa_{12}w_k\right]\end{aligned}$$

which proves our first claim in the lemma with $\kappa_3 := 4M\kappa_{12}$ and $\kappa_4 := 4M$. Next we use the last inequality and Lemma 3.2.1 to show that

$$\|\bar{\lambda}_k - \lambda_*\| \;\leq\; \kappa_1 w_k + \kappa_2\|x_k - x_*\| \;\leq\; \kappa_1 w_k + \kappa_2(\kappa_3 w_k + \kappa_4\mu_k\|\lambda_k - \lambda_*\|)$$

which is our second inequality in the lemma with $\kappa_5 := \kappa_1 + \kappa_2\kappa_3$ and $\kappa_6 := \kappa_2\kappa_4$. Finally, we obtain by the definition of $\bar{\lambda}_k = \bar{\lambda}(x_k, \lambda_k, \mu_k)$ on page 42

$$\begin{aligned}\|c(x_k)\| \;&=\; \mu_k\|\bar{\lambda}_k - \lambda_k\| \\ &\leq\; \mu_k\left(\|\bar{\lambda}_k - \lambda_*\| + \|\lambda_k - \lambda_*\|\right) \\ &\leq\; \mu_k\kappa_5 w_k + \mu_k^2\kappa_6\|\lambda_k - \lambda_*\| + \mu_k\|\lambda_k - \lambda_*\|\end{aligned}$$

Note that all derived inequalities hold $\forall k \geq k_1,\ k \in \mathcal{K}$ with some $k_1 \in I\!N$. □

For the rest of this section we will assume that the sequence of iterates generated by algorithm ALINF converges to a single limit point x_*. This assumption makes AS 2 unnecessary. We will not define the new assumption as AS 7 but rather spell it out in each of the following theorems in order to emphasize its usage. As mentioned earlier, the following lemma shall also prepare an upcoming contradiction proof.

Lemma 3.3.2. *Assume that AS 1 and AS 4 hold. Further assume that $(x_k)_{k\in I\!N}$, the sequence of iterates generated by algorithm ALINF, converges to a single limit point x_* at which AS 3 holds. Let λ_* be the corresponding Lagrange multiplier and suppose that* $\lim_{k\to\infty} \mu_k = 0$. *Then the Lagrange multiplier estimates also converge to λ_*, i.e.* $\lim_{k\to\infty} \lambda_k = \lambda_*$.

Proof. Note that our assumptions are sufficient to apply the global convergence theorem 3.2.3. In the following we distinguish two cases:

Case 1: Step 3a is executed infinitely often

Define $S := \{k \in I\!N$ s.t. Step 3a is executed in iteration k$\}$. By the triangle inequality we obtain for $k \in S$

$$\begin{aligned} \|\lambda_{k+1} - \lambda_*\| &\leq \|\lambda_{k+1} - \bar{\lambda}(x_k, \lambda_k, \mu_k)\| + \|\bar{\lambda}(x_k, \lambda_k, \mu_k) - \lambda_*\| \\ &\overset{(3.12)}{\leq} w_k + \|\bar{\lambda}(x_k, \lambda_k, \mu_k) - \lambda_*\| \end{aligned}$$

Hence, by Theorem 3.2.3 the right hand side converges to zero for $k \in S$, which implies $\lambda_{k+1} \to_{k\in S} \lambda_*$. But for $k \notin S$ we have by (3.14) $\lambda_{k+1} = \lambda_k$. Hence the whole sequence $(\lambda_k)_{k\in I\!N}$ converges to λ_*.

Case 2: Step 3a is executed only finitely often

Hence Step 3b is executed $\forall k \geq k_2$ for some $k_2 \in I\!N$. By (3.14) λ_k will then remain constant and thus $\|\lambda_k - \lambda_*\| = \textit{const}\ \forall k \geq k_2$. But then we obtain by Lemma 3.2.1 and Lemma 3.1.1

$$\|c(x_k)\| \ \leq\ \kappa_1 w_k \mu_k + \mu_k \|\lambda_k - \lambda_*\| + \kappa_2 \mu_k \|x_k - x_*\| \ \leq\ \kappa_{13}\mu_k$$

for some $\kappa_{13} > 0$ and $\forall k \geq k_3 \geq k_2$. Now the assumption that $\lim_{k\to\infty} \mu_k = 0$, together with $\alpha_\eta < 1$ implies, that there exists a $k_4 \in I\!N$ such that $\kappa_{13}\mu_k < \gamma_1 \mu_k^{\alpha_\eta}\ \forall k \geq k_4$. But then $\forall k \geq \max\{k_3, k_4\}$

$$\|c(x_k)\| \ \leq\ \kappa_{13}\mu_k \ <\ \gamma_1 \mu_k^{\alpha_\eta} \ \overset{\text{Step 3b}}{=}\ \gamma_1 \eta_k$$

Hence inequality (3.9) is satisfied which is impossible since this would imply that Step 3a is again executed.

Thus Step 3a must be executed infinitely often and λ_k converges to λ_*. □

Now we are ready to prove that the penalty parameter μ_k is bounded away from zero. This result will be of particular importance in later chapters, because it prevents the Hessian matrix of the discretized augmented Lagrange functional from becoming more and more ill conditioned. In addition, the next theorem will enable us to derive results for the local convergence behavior of algorithm ALINF.

Theorem 3.3.3. *Assume that AS 1 and AS 4 hold. Suppose that the sequence of iterates $(x_k)_{k\in I\!N}$ generated by algorithm ALINF converges to a single limit point x_* at which AS 3 holds. Let λ_* be the corresponding Lagrange multiplier and suppose that AS 5 and AS 6 hold at (x_*, λ_*). Then the sequence of penalty parameters $(\mu_k)_{k\in I\!N}$ is bounded away from zero, i.e. $\exists\ \mu_{min} \in (0,1)$ such that $\mu_k \geq \mu_{min}\ \forall k \in I\!N$.*

Proof. We will prove this theorem by contradiction. As $(\mu_k)_{k\in I\!N}$ is by definition a non-increasing sequence, it is sufficient to assume that μ_k converges to zero and lead this to a contradiction.

By definition of the algorithm, if μ_k tends to zero, step 3b must be executed infinitely often. Note that the assumptions in the theorem are sufficient to apply Theorem 3.2.3 and Lemma 3.3.1. Furthermore, as $(x_k)_{k\in I\!N}$ converges to a single limit point, we may apply Lemma 3.3.1 to all $x_k,\ k \in I\!N$, if k is large enough.

Choose k_1 from Lemma 3.3.1 such that we can apply the inequalities in the lemma for all k greater or equal than k_1. In particular we have

$$\mu_k \leq \bar{\mu} < 1 \qquad \forall k \geq k_1 \tag{3.42}$$

where $\bar{\mu}$ is as stated in Lemma 3.3.1. Note that $\forall k \in I\!N\ w_k < 1$ and hence by definition of step 3a and step 3b in the algorithm

$$w_k \leq \mu_k \tag{3.43}$$

Now let k_4 be the smallest integer k such that

$$\mu_k^{1-\alpha_\eta} \leq \frac{\gamma_1}{2+\kappa_5} \tag{3.44}$$

$$\mu_k^{1-\beta_\eta} \leq \min\left\{\frac{1}{\kappa_{14}}, \frac{\gamma_1}{2\kappa_{14}+\kappa_5}\right\} \tag{3.45}$$

where $\kappa_{14} := 1 + \kappa_5 + \kappa_6$. By (3.42), (3.45) and using $0 < 1 - \beta_\eta < 1$ we obtain for all $k \geq \max\{k_1, k_4\}$:

$$\mu_k \ \leq\ \mu_k^{1-\beta_\eta} \ \leq\ \frac{1}{\kappa_{14}} \ \leq\ \frac{1}{\kappa_6} \tag{3.46}$$

Due to Lemma 3.3.2 we may choose $k_5 \in I\!N$ such that

$$\|\lambda_k - \lambda_*\| \leq 1 \qquad \forall k \geq k_5 \tag{3.47}$$

Now define $k_6 := \max\{k_1, k_4, k_5\}$ and

$$\Gamma := \{k \in I\!N \mid \text{step 3b is executed at iteration } k-1 \text{ and } k \geq k_6\}$$

and let k_0 be the smallest element of Γ. By assumption μ_k tends to zero, hence Γ has an infinite number of elements.

We will show now by induction, that step 3a is executed $\forall k \geq k_0$ which contradicts the fact that $|\Gamma|$ is infinite. We start by showing that step 3a is executed at iteration k_0. By definition of Γ, we have for iteration k_0

$$w_{k_0} = \mu_{k_0} \qquad \text{and} \qquad \eta_{k_0} = \mu_{k_0}^{\alpha_\eta} \tag{3.48}$$

Then we obtain by Lemma 3.3.1

$$\begin{aligned}
\|c(x_{k_0})\| &\leq \kappa_5 w_{k_0}\mu_{k_0} + \mu_{k_0}(1+\kappa_6\mu_{k_0})\|\lambda_{k_0} - \lambda_*\| \\
&\overset{(3.46)}{\leq} \kappa_5 w_{k_0}\mu_{k_0} + 2\mu_{k_0}\|\lambda_{k_0} - \lambda_*\| \\
&\overset{(3.43),(3.47)}{\leq} (2+\kappa_5\mu_{k_0})\mu_{k_0} \\
&\overset{(3.42)}{\leq} (2+\kappa_5)\mu_{k_0} \;=\; (2+\kappa_5)\mu_{k_0}^{1-\alpha_\eta}\mu_{k_0}^{\alpha_\eta} \\
&\overset{(3.44)}{\leq} \gamma_1\mu_{k_0}^{\alpha_\eta} \;=\; \gamma_1\eta_{k_0}
\end{aligned} \tag{3.49}$$

As a consequence of this inequality, step 3a will be executed with $\lambda_{k_0+1} = \bar{\lambda}(x_{k_0}, \lambda_{k_0}, \mu_{k_0}) + e_{k_0}$, where $\|e_{k_0}\| \leq w_{k_0}$ by (3.12). Then we obtain by Lemma 3.3.1, (3.43), (3.47) and a simple application of the triangle inequality that

$$\|\lambda_{k_0+1} - \lambda_*\| \;\leq\; (\kappa_5+1)w_{k_0} + \kappa_6\mu_{k_0}\|\lambda_{k_0} - \lambda_*\| \;\leq\; \kappa_{14}\mu_{k_0} \tag{3.50}$$

Now, to fulfill our inductive proof, assume that step 3a is executed for iterations $k_0+i,\ 0 \leq i \leq t$ and that

$$\|\lambda_{k_0+i+1} - \lambda_*\| \;\leq\; \kappa_{14}\mu_{k_0}^{1+\beta_\eta i} \qquad ,0 \leq i \leq t \tag{3.51}$$

Inequalities (3.49) and (3.50) show that this is true for $t = 0$. Now we aim at showing that this is also true for $i = t+1$. By assumption step 3a is executed in iteration k_0+t. Hence

$$\mu_{k_0+t+1} = \mu_{k_0}\ , \qquad w_{k_0+t+1} \overset{(3.48)}{=} \mu_{k_0}^{t+2}\ , \qquad \eta_{k_0+t+1} \overset{(3.48)}{=} \mu_{k_0}^{\beta_\eta(t+1)+\alpha_\eta} \tag{3.52}$$

Then, by Lemma 3.3.1

$$\begin{aligned}
\|c(x_{k_0+t+1})\| &\leq \kappa_5 w_{k_0+t+1}\mu_{k_0+t+1} + \mu_{k_0+t+1}(1+\kappa_6\mu_{k_0+t+1})\|\lambda_{k_0+t+1}-\lambda_*\| \\
&\overset{(3.46)}{\leq} \kappa_5 w_{k_0+t+1}\mu_{k_0+t+1} + 2\mu_{k_0+t+1}\|\lambda_{k_0+t+1}-\lambda_*\| \\
&\overset{(3.51),(3.52)}{\leq} \kappa_5\mu_{k_0}^{t+3} + 2\kappa_{14}\mu_{k_0}\mu_{k_0}^{1+\beta_\eta t} \\
&\overset{\alpha_\eta,\beta_\eta<1,\ \mu_{k_0}<1}{\leq} \kappa_5\mu_{k_0}^{\alpha_\eta+\beta_\eta(t+1)+1} + 2\kappa_{14}\mu_{k_0}\mu_{k_0}^{\alpha_\eta+\beta_\eta t} \\
&= \kappa_5\mu_{k_0}\mu_{k_0}^{\alpha_\eta+\beta_\eta(t+1)} + 2\kappa_{14}\mu_{k_0}^{\beta_\eta(t+1)+\alpha_\eta}\mu_{k_0}^{1-\beta_\eta} \\
&\overset{(3.46)}{\leq} (2\kappa_{14}+\kappa_5)\mu_{k_0}^{1-\beta_\eta}\mu_{k_0}^{\beta_\eta(t+1)+\alpha_\eta} \\
&\overset{(3.45)}{\leq} \gamma_1\mu_{k_0}^{\beta_\eta(t+1)+\alpha_\eta} \overset{(3.52)}{=} \gamma_1\eta_{k_0+t+1}
\end{aligned}$$

Thus step 3a will be executed again and hence

$$\lambda_{k_0+t+2} = \bar{\lambda}(x_{k_0+t+1}, \lambda_{k_0+t+1}, \mu_{k_0+t+1}) + e_{k_0+t+1}$$

where $\|e_{k_0+t+1}\| \leq w_{k_0+t+1}$ by (3.12). Then, by Lemma 3.3.1 and the triangle inequality

$$\begin{aligned}
\|\lambda_{k_0+t+2}-\lambda_*\| &\leq (\kappa_5+1)w_{k_0+t+1} + \kappa_6\mu_{k_0+t+1}\|\lambda_{k_0+t+1}-\lambda_*\| \\
&\overset{(3.51),(3.52)}{\leq} (\kappa_5+1)\mu_{k_0}^{t+2} + \kappa_6\kappa_{14}\mu_{k_0}\mu_{k_0}^{1+\beta_\eta t} \\
&\overset{\beta_\eta<1}{\leq} (\kappa_5+1)\mu_{k_0}^{1+\beta_\eta(t+1)} + \kappa_6\kappa_{14}\mu_{k_0}\mu_{k_0}^{1+\beta_\eta t} \\
&= (\kappa_5+1+\kappa_6\kappa_{14}\mu_{k_0}^{1-\beta_\eta})\mu_{k_0}^{1+\beta_\eta(t+1)} \\
&\overset{(3.45)}{\leq} (\kappa_5+1+\kappa_6)\mu_{k_0}^{1+\beta_\eta(t+1)} \\
&= \kappa_{14}\mu_{k_0}^{1+\beta_\eta(t+1)}
\end{aligned}$$

which establishes (3.51) for $i = t+1$. Thus step 3a is executed for all iterations $k \geq k_0$ which contradicts that Γ consists of infinitely many elements. □

Now we are ready to prove local convergence properties of algorithm ALINF. In addition, the last theorem enables us for the first time to show that the Lagrange multiplier estimates λ_k converge to the Lagrange multiplier $\lambda_* = \lambda(x_*)$:

Theorem 3.3.4 (Local Convergence Theorem). *Assume that AS 1 and AS 4 are valid and that the sequence of iterates $(x_k)_{k\in\mathbb{N}}$ generated by algorithm ALINF converges to a single limit point x_* at which AS 3 holds. Let λ_* be the corresponding Lagrange multiplier and suppose that AS 5 and AS 6 hold at (x_*, λ_*). Then the Lagrange multiplier estimates $(\lambda_k)_{k\in\mathbb{N}}$ converge to λ_*. Further the sequences $(x_k)_{k\in\mathbb{N}}$, $(\bar{\lambda}(x_k, \lambda_k, \mu_k))_{k\in\mathbb{N}}$ and $(\lambda_k)_{k\in\mathbb{N}}$ are* *at least R-linearly convergent with R-factor at most $\mu_{min}^{\beta_\eta}$ where μ_{min} is the smallest value of the penalty parameter generated by algorithm ALINF.*

Proof. By the last theorem, the penalty parameter μ_k is bounded away from zero. Thus $\mu_k = \mu_{min} > 0 \;\; \forall k \geq k_{max}$ for some $k_{max} \in I\!N$. Hence step 3a is executed $\forall k \geq k_{max}$ and we obtain for those k by (3.13)

$$w_{k+1} = \mu_{min} w_k \qquad \text{and} \qquad \eta_{k+1} = \mu_{min}^{\beta_\eta} \eta_k \tag{3.53}$$

Moreover, we must have by (3.9), (3.10) and (3.11) that $\|c(x_k)\| < \gamma_2 \eta_k \;\; \forall k \geq k_{max}$. We will now use some results we derived in the proof of Lemma 3.3.1. Note that we only use results of that proof which do not depend on the assumption of Lemma 3.3.1 that the penalty parameter μ_k converges to zero. Following the second part of that proof we can replace the bound on the right side of (3.39) by $w_k + \gamma_2 \eta_k$ (here we used $\|c(x_k)\| < \gamma_2 \eta_k$) and hence we obtain instead of (3.40)

$$\begin{array}{rcl} \|x_k - x_*\| & \leq & \left\| \begin{pmatrix} x_k - x_* \\ \bar{\lambda}_k - \lambda_* \end{pmatrix} \right\| \\ & \leq & M(w_k + \gamma_2 \eta_k + \kappa_{10} \|x_k - x_*\|^2 + \kappa_{11} w_k \|x_k - x_*\|) \end{array} \tag{3.54}$$

Therefore, if $k \geq k_{max}$ is sufficiently large such that

$$w_k \leq \frac{1}{2M\kappa_{11}}\;, \qquad \|x_k - x_*\| \leq \frac{1}{4M\kappa_{10}} \tag{3.55}$$

we use inequalities (3.54) and (3.55) to obtain

$$\begin{array}{rrcl} & \|x_k - x_*\| & \leq & M(w_k + \gamma_2 \eta_k) + \frac{1}{4}\|x_k - x_*\| + \frac{1}{2}\|x_k - x_*\| \\ \Longleftrightarrow & \|x_k - x_*\| & \leq & 4M(w_k + \gamma_2 \eta_k) = 4M w_k + 4M\gamma_2 \eta_k \\ & & =: & \kappa_{15} w_k + \kappa_{16} \eta_k \end{array} \tag{3.56}$$

Using this inequality, (3.53), $\beta_\eta < 1$ and $\mu_{min} < 1$ we obtain

$$\begin{array}{rcl} \|x_k - x_*\| & \leq & \kappa_{15} w_k + \kappa_{16} \eta_k \;\; = \;\; \kappa_{15} \mu_{min} w_{k-1} + \kappa_{16} \mu_{min}^{\beta_\eta} \eta_{k-1} \\ & \leq & \mu_{min}^{\beta_\eta} (\kappa_{15} w_{k-1} + \kappa_{16} \eta_{k-1}) \end{array}$$

which shows that x_k converges to x_* at least R-linearly with R-factor $\mu_{min}^{\beta_\eta}$. Then inequality (3.56) shows, applied to Lemma 3.2.1, that the same result holds for $\bar{\lambda}(x_k, \lambda_k, \mu_k)$. As step 3a is executed for $k \geq k_{max}$, (3.12) holds for all successive iterations, which guarantees that $(\lambda_k)_{k \in I\!N}$ also converges R-linearly to λ_* with R-factor $\mu_{min}^{\beta_\eta}$. □

CHAPTER 4

Approximation Theory

In the last chapter we derived an algorithm which generates a sequence $(x_k)_{k\in I\!\!N} \subset X$ that converges under appropriate assumptions to a Karush Kuhn Tucker point x_* with corresponding Lagrange multiplier λ_*.

However, this result cannot be used in practice yet, because a computer cannot handle arbitrary objects in Hilbert spaces. In order to make numerical computations possible, one must approximate those arbitrary objects by objects which can be represented by a finite set of numbers. Further one must discretize the mappings $f(\cdot)$ and $c(\cdot)$.

As we will see in this chapter, the appropriate functional analytic tool to solve these tasks are restrictions. We will pick up some basic ideas that for example Volkwein presented in [26] and modify and extend them for our purpose.

4.1 Discretization of the Hilbert Spaces

Consider the Hilbert spaces X and Y as introduced in earlier chapters. Our goal in this section is to generate sequences of finite dimensional Hilbert spaces $(X_n)_{n\in I\!\!N}$ with $X_n \subset X$ and $(Y_n)_{n\in I\!\!N}$ with $Y_n \subset Y$ which approximate the original spaces X and Y in a suitable manner.

As the task is essentially the same for the spaces X and Y, we will rather discuss the procedure for a given Hilbert space Z with scalar product $< \cdot, \cdot >_Z$ and induced norm $\| \cdot \|_Z$. The major tool for the discretization is introduced in the following definition:

Definition 4.1.1. Let $(Z, < \cdot, \cdot >_Z)$ be a Hilbert space and let Z_n be a finite dimensional subspace of Z. Then a surjective mapping $r^{Z_n} \in \mathcal{L}(Z, Z_n)$ is called a *restriction.*

Restrictions provide the finite dimensional interpretation of the infinite dimensional vector space. We make the following remark on the definition:

Remark: The restriction operator maps an element of the space Z to a finite dimensional subspace Z_n of Z. Thus, as we will also see in examples, r^{Z_n} is more or less described by a specific basis of Z_n. In fact, in applications one usually chooses "nice" basis-elements $\phi_1, ..., \phi_n \in Z$ and constructs the subspace Z_n by the span of those elements denoted by $span(\phi_1, ..., \phi_n)$. The restriction operator r^{Z_n} is then chosen based on the given basis elements.

Note that the subspace Z_n of Z is requested to be finite dimensional. Thus it is also a Hilbert space with scalar product $< \cdot, \cdot >_{Z_n} := < \cdot, \cdot >_Z$ and induced norm $\| \cdot \|_{Z_n} := \| \cdot \|_Z$.

The following example will clarify the definition of restrictions and the procedure of how to construct them:

Example 4.1.2. Let $Z := H^1(0,1) := \{u \in L^2(0,1) : Du \in L^2(0,1)\}$ be the Sobolev space $W^{1,2}(0,1)$ where Du denotes the weak derivative of u defined by the equation

$$\int_0^1 u(x)\phi'(x)dx = -\int_0^1 \phi(x)Du(x)dx \qquad \forall \phi \in C_0^\infty(0,1)$$

where $C_0^\infty(0,1)$ denotes the set of all $C^\infty(0,1)$-functions with compact support, the test functions. It is well known that $H^1(0,1)$ is a Hilbert space with inner product

$$< u, v >_{H^1(0,1)} := \int_0^1 u(x)v(x)dx + \int_0^1 Du(x)Dv(x)dx$$

Now we construct the basis elements which will span our finite dimensional subspace. Define for fixed $n \in I\!N$ the mesh-size $h := 1/n$ and for $j = 0, ..., n$ the mesh points $a_j := jh$. Then we construct the *linear finite elements* as follows

$$\phi_0(x) := \left\{ \begin{array}{rcl} \frac{1}{h}(a_1 - x) & , & x \in [a_0, a_1] \\ 0 & , & \text{otherwise} \end{array} \right.$$

$$\phi_j(x) := \left\{ \begin{array}{rcl} \frac{1}{h}(x - a_{j-1}) & , & x \in (a_{j-1}, a_j] \\ \frac{1}{h}(a_{j+1} - x) & , & x \in (a_j, a_{j+1}] \\ 0 & , & \text{otherwise} \end{array} \right\} \qquad , 1 \leq j \leq n-1$$

$$\phi_n(x) := \left\{ \begin{array}{rcl} \frac{1}{h}(x - a_{n-1}) & , & x \in (a_{n-1}, a_n] \\ 0 & , & \text{otherwise} \end{array} \right.$$

The linear finite elements for $n = 4$ are illustrated in Figure 4.1.

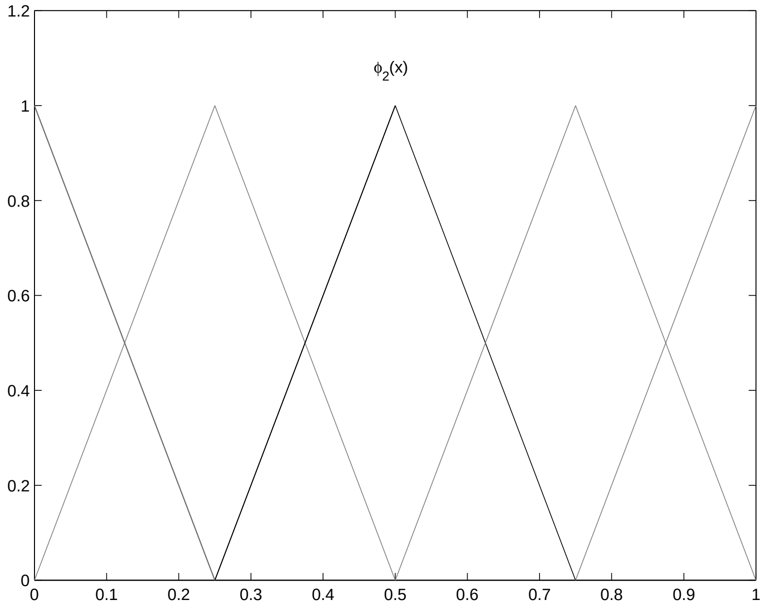

Figure 4.1: Linear Finite Elements for $n = 4$

It is easy to verify that $\phi_0, ..., \phi_n$ are linearly independent. Using these elements we define the finite dimensional subspace of Z as $Z_n := span(\phi_0, ..., \phi_n)$ and the restriction $r^{Z_n} : Z \to Z_n$ by

$$r^{Z_n} u := \sum_0^n u(a_i)\phi_i$$

In order to show that r^{Z_n} is well defined, we must prove that $u(a_i)$ is meaningful. This is not obvious, because $u \in H^1(0,1)$ is only measurable on first sight. However, one can show that $H^1(0,1)$ is compactly embedded in $C[0,1]$ (see for example [2, page 144]) so that $u \in H^1(0,1)$ is in fact a continuous function.

Finally we will show that r^{Z_n} satisfies the requirements of a restriction operator. Clearly, r^{Z_n} is linear. In order to show that it is bounded, we use again the fact that $H^1(0,1)$ is compactly embedded in $C[0,1]$. In particular, the identity operator $I : H^1(0,1) \to C[0,1]$ is bounded. Now let $(u_m)_m \subset H^1(0,1)$ such that $u_m \to u$ in $\| \cdot \|_{H^1(0,1)}$. Then

$$\|u_m - u\|_{C[0,1]} \leq c\|u_m - u\|_{H^1(0,1)} \to_{m\to\infty} 0$$

where $c := \|I\|_{\mathcal{L}(H^1(0,1),C[0,1])}$. But $\| \cdot \|_{C[0,1]}$ is the supremum-norm, thus $u_m(a_i)$ converges to

$u(a_i)$ for $i = 0, ..., n$. Therefore

$$\begin{aligned} \|r^{Z_n}u_m - r^{Z_n}u\|_{H^1(0,1)} &= \left\| \sum_{i=0}^{n} u_m(a_i)\phi_i - \sum_{i=0}^{n} u(a_i)\phi_i \right\|_{H^1(0,1)} \\ &\leq \sum_{i=0}^{n} \|\phi_i\|_{H^1(0,1)} |u_m(a_i) - u(a_i)| \longrightarrow_{m\to\infty} 0 \end{aligned}$$

Hence $r^{Z_n}(\cdot)$ is continuous and thus $r^{Z_n} \in \mathcal{L}(Z, Z_n)$.

In general there exist several restrictions $r^{Z_n} : Z \to Z_n$ for a given finite dimensional subspace Z_n of Z. However, there only exists one optimal restriction, where optimality is defined as follows:

Definition 4.1.3. Let Z be a Hilbert space and $Z_n \subset Z$ a finite dimensional subspace. We say that a surjective mapping $\widehat{r^{Z_n}} \in \mathcal{L}(Z, Z_n)$ is the *optimal restriction* to Z_n in Z if

$$\|z - \widehat{r^{Z_n}}z\|_Z = \inf_{\tilde{z}\in Z_n} \|z - \tilde{z}\|_Z \qquad \forall z \in Z$$

The next theorem will show, that the optimal restriction exists and is well defined:

Lemma 4.1.4. *Let Z be a Hilbert space and $Z_n \subset Z$ a finite dimensional subspace. Then the optimal restriction $\widehat{r^{Z_n}} \in \mathcal{L}(Z, Z_n)$ exists, is unique and equal to the projection operator given by the projection theorem (see appendix, Theorem A.7).*

Proof. As Z_n is a finite dimensional subspace of Z, it is closed. Thus we can deduce from the best approximation theorem (see Appendix, Theorem A.6), that $\forall z \in Z$ there exists a unique $\widehat{z} \in Z_n$ such that

$$\|z - \widehat{z}\|_Z \leq \|z - \tilde{z}\|_Z \qquad \forall \tilde{z} \in Z_n$$

Define $\widehat{r^{Z_n}} : Z \to Z_n$ by $\widehat{r^{Z_n}}z := \widehat{z}$. Clearly, $\widehat{r^{Z_n}}$ is well defined and surjective. In order to prove the theorem we must only show, that $\widehat{r^{Z_n}} \in \mathcal{L}(Z, Z_n)$.

Let $z_1, z_2 \in Z$. Then we know by the projection theorem (see Appendix, Theorem A.7), that $z_1 - \widehat{z_1}, z_2 - \widehat{z_2} \in Z_n^{\perp}$ and $\widehat{z_1},\ \widehat{z_2} \in Z_n$. Then

$$z_1 + z_2 = \underbrace{\widehat{z_1} + \widehat{z_2}}_{\in Z_n} + \underbrace{(z_1 - \widehat{z_1}) + (z_2 - \widehat{z_2})}_{\in Z_n^{\perp}}$$

which implies that

$$\|z_1 + z_2 - (\widehat{z_1} + \widehat{z_2})\| \leq \|z_1 + z_2 - \tilde{z}\| \qquad \forall \tilde{z} \in Z_n$$

Hence $\widehat{z_1+z_2} = \widehat{z_1} + \widehat{z_2}$ or equivalently $\widehat{r^{Z_n}}(z_1+z_2) = \widehat{r^{Z_n}} z_1 + \widehat{r^{Z_n}} z_2$. Obviously, we have $\widehat{\lambda z} = \lambda \widehat{z}$ for all $\lambda \in I\!R$. Thus $\widehat{r^{Z_n}}$ is linear. Further we obtain by the Pythagorean theorem

$$\|\widehat{r^{Z_n}} z\|^2 = \|\widehat{z}\|^2 \leq \|\widehat{z}\|^2 + \|z - \widehat{z}\|^2 = \|\widehat{z} + z - \widehat{z}\|^2 = \|z\|^2$$

Hence $\widehat{r^{Z_n}}$ is bounded and the theorem is proved. □

The question remains whether we can choose a "nice" basis such that the projection operator can be expressed in a convenient way. The following theorem shows, that this is possible for every Hilbert space with a countable basis:

Theorem 4.1.5. *Let Z be a Hilbert space with a countable orthonormal basis $(\phi_j)_{j \in I\!N}$. Define a finite dimensional subspace $Z_n := span(\phi_1, ..., \phi_n)$ and $r^{Z_n} : Z \to Z_n$ by*

$$r^{Z_n} z := \sum_{i=1}^{n} < \phi_i, z >_Z \phi_i$$

Then r^{Z_n} is the optimal restriction to Z_n in Z.

Proof. Let $z \in Z$ be arbitrary. Then z can be written as (see for example [23, page 45])

$$z = \sum_{i=1}^{\infty} < \phi_i, z >_Z \phi_i$$

But then we obtain by the Pythagorean theorem $\forall \lambda = (\lambda_1, ..., \lambda_n)^T \in I\!R^n$

$$\begin{aligned} \|z - r^{Z_n} z\|^2 &= \left\| \sum_{i=1}^{\infty} < \phi_i, z >_Z \phi_i - \sum_{i=1}^{n} < \phi_i, z >_Z \phi_i \right\|^2 = \left\| \sum_{i=n+1}^{\infty} < \phi_i, z >_Z \phi_i \right\|^2 \\ &\leq \left\| \sum_{i=n+1}^{\infty} < \phi_i, z >_Z \phi_i \right\|^2 + \left\| \sum_{i=1}^{n} (< \phi_i, z >_Z - \lambda_i) \phi_i \right\|^2 \\ &= \left\| \sum_{i=1}^{\infty} < \phi_i, z >_Z \phi_i - \sum_{i=1}^{n} \lambda_i \phi_i \right\|^2 \end{aligned}$$

which is equivalent to

$$\|z - r^{Z_n} z\| \leq \|z - \tilde{z}\| \qquad \forall \tilde{z} \in Z_n$$

But this means, that r^{Z_n} is the projection operator which we proved to be the optimal restriction in Lemma 4.1.4. Hence the claim follows. □

Note that the last theorem is a theoretical result. Its usefulness in practice depends on how easy the Fourier coefficients $< \phi_i, z >_Z$ can be computed for a given orthonormal basis and how nice the basis elements themselves are.

Example 4.1.6. We consider again $Z := H^1(0,1)$ and define the following basis elements for $j \in I\!N$:

$$\phi_j(x) := \left(\frac{2}{1+(j-1)^2\pi^2}\right)^{\frac{1}{2}} cos((j-1)\pi x)$$

It can be verified that the $(\phi_j)_{j\in I\!N}$ form an orthonormal basis of $H^1(0,1)$. Hence r^{Z_n} as defined in the last theorem is the optimal restriction to the subspace $Z_n := span(\phi_0, ..., \phi_n)$ in Z. However, the Fourier coefficients

$$\begin{aligned} < \phi_j, z >_{H^1(0,1)} &= \int_0^1 \phi_j(x)z(x)dx + \int_0^1 \phi_j'(x)Dz(x)dx \\ &= \sqrt{2+2(j-1)^2\pi^2}\int_0^1 cos((j-1)\pi x)z(x)dx \end{aligned}$$

are not necessarily easy to compute for arbitrary $z \in Z$. But often the vectors $z \in Z$ which we need to approximate have “nice” properties, so that this integral might be relatively easy to calculate.

Up to this point, we constructed a restriction r^{Z_n} which generates the finite dimensional subspace Z_n of Z. In applications it is not sufficient to approximate the infinite dimensional problem by one finite dimensional subspace, but rather by a sequence of subspaces or restrictions which “converge” to the subspace Z.

In order to define this convergence, we consider a sequence of restrictions $(r^{Z_n})_{n\in I\!N}$ where $\forall n \in I\!N\ r^{Z_n} : Z \to Z_n$ with subspaces $(Z_n)_{n\in I\!N}$ generated by $(r^{Z_n})_{n\in I\!N}$. Then we will call the pairs $(Z_n, r^{Z_n})_{n\in I\!N}$ *approximations* of the space Z.

By intention we will request that the subspaces $(Z_n)_{n\in I\!N}$ become dense in Z as n tends to infinity. However, in order to define the convergence mentioned above, it is more appropriate to formulate the desired properties in terms of the restrictions $(r^{Z_n})_{n\in I\!N}$:

Definition 4.1.7. Let Z be a Hilbert space. Further let $(r^{Z_n})_{n\in I\!N}$ be a sequence of restrictions which generates the subspaces $(Z_n)_{n\in I\!N}$. Then the restrictions $(r^{Z_n})_{n\in I\!N}$ are said to be *convergent* if

$$\lim_{n\to\infty} \|z - r^{Z_n}z\|_Z = 0 \qquad \forall z \in Z$$

Equivalently we say that the approximations $(Z_n, r^{Z_n})_{n\in I\!N}$ converge.

The convergence property of the restrictions $(r^{Z_n})_{n\in I\!N}$ will allow us to approximate any vector $z \in Z$ arbitrarily close by its discretization $r^{Z_n} z \in Z_n$. We make the following remark on the definition:

Remark: Note that the definition above means nothing else than that the sequence of mappings $(r^{Z_n})_{n\in I\!N}$ converges *pointwise* to the identity operator. In arbitrary spaces we will not be able to request more than this. In fact it is hard enough to verify this pointwise convergence in specific examples. The associated proofs are often technical and lengthy.

However, if we restrict our attention to bounded sets of specific elements in Z, then it might be possible to prove a "uniform convergence" of the restriction operators on these sets.

The following example will state the convergence property of the restriction operators defined in example 4.1.2.

Example 4.1.8. Let $Z := H^1(0,1)$ and define the approximations $(Z_n, r^{Z_n})_{n\in I\!N}$ as in Example 4.1.2. Then the restrictions $(r^{Z_n})_{n\in I\!N}$ converge, i.e.

$$\lim_{n\to\infty} \|u - r^{Z_n} u\|_Z = 0 \qquad \forall u \in H^1(0,1)$$

For the proof we refer the reader to [4, pages 40-42].

The next theorem will show that the optimal restrictions defined in Theorem 4.1.5 are necessarily convergent.

Theorem 4.1.9. *Let Z be a Hilbert space with a countable orthonormal basis $(\phi_j)_{j\in I\!N}$. Then the approximations $(Z_n, r^{Z_n})_{n\in I\!N}$ as defined in Theorem 4.1.5 are convergent.*

Proof. Let $z \in Z$ be arbitrary. Then z can be written as $z = \sum_{i=1}^{\infty} < \phi_i, z >_Z \phi_i$. Further we obtain by the Parseval equality (see [23, page 46])

$$\begin{aligned}
\|z - r^{Z_n} z\|^2 &= \left\| z - \sum_{i=1}^{n} < \phi_i, z >_Z \phi_i \right\|^2 = \|z\|^2 - \sum_{i=1}^{n} | < \phi_i, z >_Z |^2 \\
&= \sum_{i=1}^{\infty} | < \phi_i, z >_Z |^2 - \sum_{i=1}^{n} | < \phi_i, z >_Z |^2 \\
&= \sum_{i=n+1}^{\infty} | < \phi_i, z >_Z |^2 \longrightarrow_{n\to\infty} 0
\end{aligned}$$

□

Of course it has certain advantages for the discretization to choose optimal restrictions as defined in Theorem 4.1.5. Besides the optimality of those operators one can readily apply the last theorem to obtain the convergence property. However, in order to construct the optimal restrictions we have to prove that the generating basis elements $(\phi_i)_{i\in I\!N}$ form an orthonormal basis of the Hilbert space Z. These proofs are in general also very technical and lengthy.

The following example is a simple application of the last theorem:

Example 4.1.10. The cosine approximations from Example 4.1.6 are convergent.

In the next chapter we will assume for simplicity that the subspaces $(Z_n)_{n\in I\!N}$ generated by the convergent restrictions $(r^{Z_n})_{n\in I\!N}$ are nested, i.e.:

$$Z_1 \subset Z_2 \subset \ldots \subset Z_n \subset Z_{n+1} \subset \ldots \subset Z$$

This is usually equivalent to the assumption, that the mesh size of the discretization is of the form $(1/2)^n$ for $n \in I\!N$. The construction of nested subspaces is shown in the following example:

Example 4.1.11. Let $Z := H^1(0,1)$ and define the approximations $(Z_n, r^{Z_n})_{n\in I\!N}$ as in Example 4.1.2. In particular $Z_n := span(\phi_0, ..., \phi_n)$. Now define $\widetilde{Z_n} := Z_{2^n} = span(\phi_0, ..., \phi_{2^n})$ and $r^{\widetilde{Z_n}} := r^{Z_{2^n}}$. We claim that $(\widetilde{Z_n})_{n\in I\!N}$ is a sequence of nested subspaces of Z. In order to prove this let for $n > 1$

$$\begin{aligned} \widetilde{Z_n} &= span(\bar{\phi}_0, ..., \bar{\phi}_{2^n}) \\ \widetilde{Z_{n+1}} &= span(\phi_0, ..., \phi_{2^{n+1}}) \end{aligned}$$

Then it is easy to show that

$$\begin{aligned} 2\bar{\phi}_0 &= 2\phi_0 + \phi_1 \\ 2\bar{\phi}_i &= \phi_{2i-1} + 2\phi_{2i} + \phi_{2i+1} \qquad , i = 1, ..., 2^n - 1 \\ 2\bar{\phi}_{2^n} &= \phi_{2^{n+1}-1} + 2\phi_{2^{n+1}} \end{aligned}$$

Hence $span(\bar{\phi}_0, ..., \bar{\phi}_{2^n}) \subset span(\phi_0, ..., \phi_{2^{n+1}})$ or equivalently $\widetilde{Z_n} \subset \widetilde{Z_{n+1}}$. As mentioned above, the mesh size of the discretization is $(1/2)^n$. Finally note that the approximations $(\widetilde{Z_n}, r^{\widetilde{Z_n}})_{n\in I\!N}$ are convergent due to the convergence of the restrictions $(r^{Z_n})_{n\in I\!N}$.

We conclude this section with the following property of convergent restrictions:

Theorem 4.1.12. *Let Z be a Hilbert space and assume that $(Z_n, r^{Z_n})_{n\in I\!N}$ are convergent approximations. Further let $(y_n)_{n\in I\!N} \subset Z$ be a sequence such that $y_n \rightarrow_{n\rightarrow\infty} y_*$ in Z. Then*

$$\lim_{n\rightarrow\infty} \|r^{Z_n} y_n - y_*\| = 0$$

Proof. A simple application of the triangle inequality implies

$$\begin{aligned}\|r^{Z_n}y_n - y_*\| &\leq \|r^{Z_n}y_n - r^{Z_n}y_*\| + \|r^{Z_n}y_* - y_*\| \\ &\leq \|r^{Z_n}\|\|y_n - y_*\| + \|r^{Z_n}y_* - y_*\| \end{aligned} \tag{4.1}$$

But for all $y \in Z$ the sequence $(r^{Z_n}y)_{n\in I\!N}$ converges to y, hence it is bounded. Therefore, by the uniform boundedness principle (see Appendix, Theorem A.10), $\|r^{Z_n}\|_{\mathcal{L}(Z,Z_n)}$ is bounded. Thus the right hand side of (4.1) converges to zero. □

4.2 Approximation of the Dual Spaces

While the last section equipped us with the tools to discretize vectors in the Hilbert spaces X and Y, we will now analyze how we can approximate linear functionals in the dual spaces X^* and Y^* of X and Y, respectively. As in Section 4.1, we will discuss these issues for a given Hilbert space Z with scalar product $< \cdot, \cdot >_Z$ and induced norm $\| \cdot \|_Z$.

Assume that Z_n is a finite dimensional subset of Z. Then it is obvious, that Z_n^* somehow approximates Z^*. However, we must be able to compare elements of Z_n^* with those of Z^* in order to make a statement about the approximation error. Hence it is of particular interest to find prolongations of linear functionals $l \in Z_n^*$ to Z^*. The basis for our discussion are the restriction operators $(r^{Z_n})_{n\in I\!N}$ introduced in the last section:

Definition 4.2.1. Let $(Z, < \cdot, \cdot >_Z)$ be a Hilbert space and assume that $(Z_n, r^{Z_n})_{n\in I\!N}$ are approximations of Z. For given $l \in Z_n^*$ the *dual prolongation operator* $q^{Z_n} : Z_n^* \to Z^*$ is defined by $q^{Z_n}l := lr^{Z_n}$. The pairs $(Z_n^*, q^{Z_n})_{n\in I\!N}$ are called *dual approximations* of Z^*.

We make the following remark on this definition:

Remark: As $r^{Z_n} \in \mathcal{L}(Z, Z_n)$ and $l \in \mathcal{L}(Z_n, I\!R)$ it is clear that the prolongation $q^{Z_n}l$ is an element of Z^*. Thus the prolongation operator is well defined. Further it is easy to verify that q^{Z_n} is bounded and linear.

Example 4.2.2. We choose $Z := H^1(0,1)$ and r^{Z_n} as in Example 4.1.2. There we defined the restriction for $u \in H^1(0,1)$ by $r^{Z_n}u := \sum_{i=0}^n u(a_i)\phi_i$ where $\phi_0, \dots, \phi_n$ are the linear finite elements introduced in example 4.1.2. Let now $l \in Z_n^*$ and $u \in Z$. Then

$$(q^{Z_n}l)u = lr^{Z_n}u = l\sum_{i=0}^n u(a_i)\phi_i = \sum_{i=0}^n u(a_i)l(\phi_i)$$

Now we will define convergence of the dual approximations:

Definition 4.2.3. Let $(Z, < \cdot, \cdot >_Z)$ be a Hilbert space with approximations $(Z_n, r^{Z_n})_{n \in I\!N}$. Then the dual approximations $(Z_n^*, q^{Z_n})_{n \in I\!N}$ are said to be *convergent* if and only if

$$\|l - q^{Z_n} l|_{Z_n}\|_{Z^*} = \|l - l r^{Z_n}\|_{Z^*} = \sup_{\|h\|=1, h \in Z} |lh - l r^{Z_n} h| \rightarrow_{n \to \infty} 0 \qquad \forall l \in Z^*$$

Note that the convergence of the approximations $(Z_n, r^{Z_n})_{n \in I\!N}$ of Z does in general not imply the convergence of the dual approximations $(Z_n^*, q^{Z_n})_{n \in I\!N}$ as the operator norm demands a "uniform convergence". However, if we choose optimal restrictions $(r^{Z_n})_{n \in I\!N}$, then the conclusion holds true:

Theorem 4.2.4. *Let Z be a Hilbert space with inner product $< \cdot, \cdot >_Z$ and countable orthonormal basis. Let the approximations $(Z_n, \widehat{r^{Z_n}})_{n \in I\!N}$ be given and assume that the restrictions $(\widehat{r^{Z_n}})_{n \in I\!N}$ are optimal in the sense of definition 4.1.3. Then the associated dual approximations $(Z_n^*, \widehat{q^{Z_n}})_{n \in I\!N}$ are convergent.*

Proof. By Lemma 4.1.4, $\widehat{r^{Z_n}}$ is nothing else than the projection on the closed subspace Z_n of Z. Thus we have for all $z \in Z$

$$\widehat{r^{Z_n}} z \in Z_n \quad , \qquad z - \widehat{r^{Z_n}} z \in Z_n^{\perp}$$

Further any $l \in Z^*$ has a Riesz-representation $z_l \in Z$. Hence we can conclude

$$\begin{aligned} \|l - \widehat{q^{Z_n}} l\|_{Z^*} &= \sup_{\|z\|=1} |lz - l\widehat{r^{Z_n}} z| = \sup_{\|z\|=1} | < z_l, z - \widehat{r^{Z_n}} z > | \\ &= \sup_{\|z\|=1} | < \widehat{r^{Z_n}} z_l + z_l - \widehat{r^{Z_n}} z_l, z - \widehat{r^{Z_n}} z > | \\ &= \sup_{\|z\|=1} | < \widehat{r^{Z_n}} z_l, z - \widehat{r^{Z_n}} z > + < z_l - \widehat{r^{Z_n}} z_l, z - \widehat{r^{Z_n}} z > | \\ &= \sup_{\|z\|=1} | < z_l - \widehat{r^{Z_n}} z_l, z > - < z_l - \widehat{r^{Z_n}} z_l, \widehat{r^{Z_n}} z > | \\ &\overset{\text{C.S.I.}}{\leq} \sup_{\|z\|=1} \|z_l - \widehat{r^{Z_n}} z_l\| \|z\| = \|z_l - \widehat{r^{Z_n}} z_l\| \rightarrow_{n \to \infty} 0 \end{aligned}$$

□

Using Theorems 4.1.5 and 4.1.9, we immediately obtain the following corollary:

Corollary 4.2.5. *For every Hilbert space with countable basis one can find convergent restrictions such that the dual approximations converge.*

The last corollary justifies that we may in fact assume that the dual approximations are convergent.

Example 4.2.6. The dual approximations associated with the cosine approximations in Example 4.1.6 are convergent.

We finally note, that it is often sufficient in applications to request the convergence of the dual approximations for specific elements of Z^*. Thus it may not be necessary to consider the optimal restrictions $(r^{Z_n})_{n\in I\!N}$.

4.3 Approximation of the Operators

In the last sections we discretized the Hilbert spaces X and Y and their duals by means of restrictions. Now we will focus on the discretizations of the mappings involved in our constrained optimization problem

$$\min f(x) \quad \text{s.t.} \quad c(x) = 0 \quad , x \in X$$

where $f : X \to I\!R$ and $c : X \to Y$ are twice continuous Fréchet differentiable mappings. The basis for the discretization is a sequence of convergent restrictions.

Let $(X_n, r^{X_n})_{n\in I\!N}$ and $(Y_m, r^{Y_m})_{m\in I\!N}$ be convergent approximations of the spaces X and Y. Intuitively, we will want the discretizations of the mappings f and c to have the following discretized domains and ranges

$$c_{n,m} : X_n \to Y_m \ , \qquad f_n : X_n \to I\!R$$

In addition the mappings $c_{n,m}$ and f_n should be "close" to the original mappings c and f. These requirements are met by the following approximations which are defined for $n \in I\!N$, $m \in I\!N$ and $x \in X_n$

$$\begin{aligned} c_{n,m}(x) &:= r^{Y_m} c(x)|_{X_n} = r^{Y_m} c(x) \\ f_n(x) &:= f(x)|_{X_n} = f(x) \end{aligned} \tag{4.2}$$

Clearly, we have by definition of r^{Y_m}, c and f that $c_{n,m} : X_n \to Y_m$ and $f_n : X_n \to I\!R$. We make the following important observations:

Remark:

i) Although the restriction operator r^{X_n} does not explicitely appear on the right hand side of (4.2), it is extremely important to keep in mind that the domain of these operators is the subspace X_n of X. As we will see below, this causes severe difficulties in the approximation of the derivative of $c(\cdot)$.

ii) By definition of $f_n(\cdot)$ in (4.2), we formally do not have discretization errors for the function $f(\cdot)$ itself. A different situation applies for $c_{n,m}(\cdot)$, because in this case an approximation takes place in terms of the restrictions r^{Y_m}. However, note that the convergence of the restrictions $(r^{Y_m})_{m\in I\!N}$ implies the pointwise convergence of the discretization $c_{n,m}$ to c as m tends to infinity, i.e.:

$$\|c(x) - c_{n,m}(x)\| = \|c(x) - r^{Y_m}c(x)\| \rightarrow_{m\to\infty} 0 \qquad \forall x \in X_n$$

Recall that we cannot request more than pointwise convergence of the restriction operators in general. However, note that the vectors $x \in X_n$ are linear combinations of the "nice" basis elements of X_n. Thus we can hope that the mapping $c(\cdot)$ transfers X_n to a nice subset $c(X_n)$ of Y. Using the specific structure of a given mapping $c(\cdot)$ one can often show that the discretization $c_{n,m}(\cdot)$ converges uniformly to $c(\cdot)$ on a bounded subset of X_n as $m \in I\!N$ approaches infinity.

The discretization of $c(\cdot)$ and $f(\cdot)$ in terms of linear and bounded restriction operators also implies the differentiability of the approximations $c_{n,m}(\cdot)$ and $f_n(\cdot)$:

Theorem 4.3.1. *Let $f : X \to I\!R$ and $c : X \to Y$ be twice continuous Fréchet differentiable. Then the discretized operators $f_n : X_n \to I\!R$ and $c_{n,m} : X_n \to Y_m$ are also twice continuous Fréchet differentiable and*

$$f_n'(x)h = f'(x)h \ , \qquad c_{n,m}'(x)h = r^{Y_m}c'(x)h \qquad \forall x, h \in X_n$$

Proof. The result is obvious for the mapping $f_n : X_n \to I\!R$ as Definition 2.1.2 holds in particular for the subspace $X_n \subset X$. In order to show the same for $c_{n,m} : X_n \to Y_m$ we observe, that $r^{Y_m} : Y \to Y_m$ is linear and bounded. Thus we can apply Theorem 2.1.10 and obtain the desired formula for the first derivative. In analogy we can show the remaining properties stated in the theorem. □

Remark: The last theorem shows that the Fréchet derivatives of the discretized operators can be computed easily. However, as mentioned earlier, one has to use caution considering the domains of the mappings. Note that $f_n'(x) = f'(x)|_{X_n}$. Due to this restriction $f_n'(x)(\cdot)$ is an element of $\mathcal{L}(X_n, I\!R)$ and not $\mathcal{L}(X, I\!R)$. In particular we only know that

$$\|f_n'(x)(\cdot)\|_{\mathcal{L}(X_n,I\!R)} = \sup_{\|h\|=1, h\in X_n} |f_n'(x)h| \leq \sup_{\|h\|=1, h\in X} |f'(x)h| = \|f'(x)(\cdot)\|_{\mathcal{L}(X,I\!R)}$$

but in general equality does not hold. Moreover, as the Riesz-representation of $f_n'(x)(\cdot)$ is now chosen with respect to the Hilbert space X_n, we even have in general that $\nabla f_n(x) \neq \nabla f(x)$. Similar problems occur for the mapping $c_{n,m}(\cdot)$. Thus we have to be very cautious when dealing with derivatives of $f_n(\cdot)$ and $c_{n,m}(\cdot)$.

Now we will analyze the approximation of the Fréchet derivatives $c'(x)(\cdot)$ and $f'(x)(\cdot)$ by their discretized counterparts. As the operator norms $\|\cdot\|_{\mathcal{L}(X,I\!\!R)}$ and $\|\cdot\|_{\mathcal{L}(X,Y)}$ request a uniform convergence, we cannot evade assuming an additional property of the involved restrictions. As it turns out, it is sufficient to assume that the dual approximations $(X_n^*, q^{X_n})_n$ and $(Y_m^*, q^{Y_m})_m$ induced by $(X_n, r^{X_n})_n$ and $(Y_m, r^{Y_m})_m$, respectively, are convergent.

Note that $f : X \to I\!\!R$ and thus $f'(x)(\cdot) \in X^*$. Thus the convergence of the dual approximations immediately implies the pointwise convergence of $f_n'(x)$ to $f'(x)$ in operator norm, i.e.:

$$\|f'(x) - q^{X_n} f_n'(x)\|_{X^*} = \|f'(x) - f'(x) r^{X_n}\|_{X^*} \to_{n\to\infty} 0$$

But this result does not hold true for $c'(x)$, because neither the convergence of the dual approximations $(X_n^*, q^{X_n})_n$ nor of $(Y_m^*, q^{Y_m})_m$ implies the convergence of $c_{n,m}'(x) r^{X_n}(\cdot)$ to $c'(x)(\cdot)$ in the operator norm $\|\cdot\|_{\mathcal{L}(X,Y)}$. However, we do not have to approximate $c'(x)(\cdot)$ itself in order to discretize algorithm ALINF. In fact it is sufficient to approximate the augmented Lagrangian functional $\Phi : X \times Y \times (0,\infty) \to I\!\!R$:

Recall that the inner iteration termination criterion (3.8) of algorithm ALINF was defined as follows

$$\left\| \frac{d}{dx}\Phi(x_k, \lambda_k, \mu_k)\cdot \right\|_{\mathcal{L}(X,I\!\!R)} \le w_k$$

Using the discretized mappings $f_n(\cdot)$ and $c_{n,m}(\cdot)$ it is straightforward to define the discretized augmented Lagrangian functional $\Phi_{n,m} : X_n \times Y_m \times (0,\infty) \to I\!\!R$ by

$$\begin{aligned} \Phi_{n,m}(x,\lambda,\mu) &:= f_n(x) + < \lambda, c_{n,m}(x) >_{Y_m} + \frac{1}{2\mu}\|c_{n,m}(x)\|_{Y_m}^2 \\ &= f(x) + < \lambda, r^{Y_m} c(x) >_Y + \frac{1}{2\mu}\|r^{Y_m} c(x)\|_Y^2 \end{aligned} \tag{4.3}$$

In order to transfer convergence results from algorithm ALINF to a discretized version we will have to assure that the new inner iteration termination criterion

$$\left\| \frac{d}{dx}\Phi_{n,m}(x_k, \lambda_k, \mu_k)\cdot \right\|_{\mathcal{L}(X_n,I\!\!R)} \le w_k$$

implies that criterion (3.8) is approximately fulfilled. Thus it is of particular importance how well $\frac{d}{dx}\Phi_{n,m}(x,\lambda,\mu)$ approximates $\frac{d}{dx}\Phi(x,\lambda,\mu)$.

Applying Theorem 4.3.1 one can prove in analogy to Examples 2.2.1 and 2.2.2, that the Fréchet derivative of the discretized augmented Lagrangian functional is given by:

$$\frac{d}{dx}\Phi_{n,m}(x,\lambda,\mu)\cdot = f'(x)\cdot + < \lambda, r^{Y_m} c'(x)\cdot >_Y + \frac{1}{\mu} < r^{Y_m} c(x), r^{Y_m} c'(x)\cdot >_Y \tag{4.4}$$

Note that the operator $c'(x)(\cdot)$ does neither appear solely in (4.4) nor in the derivative of the augmented Lagrangian functional. In fact this operator always appears as one of the arguments of the scalar product $< \cdot, \cdot >_Y$. This circumstance allows us to approximate $\Phi(x, \lambda, \mu)$ arbitrarily close by using the convergence of the dual approximations $(X_n^*, q^{X_n})_n$. We will prove this now in several steps, first without and then with the assumption that the dual approximations $(Y_m^*, q^{Y_m})_m$ are convergent:

Lemma 4.3.2. *Let the approximations $(X_n, r^{X_n})_{n\in I\!N}$ and the corresponding dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$ be convergent. Further assume that the discretization of the Fréchet differentiable operator $f(\cdot)$ is defined as in (4.2). Then $\forall x \in X$*

$$\|f'(x) - q^{X_n} f_n'(x)\|_{X^*} = \|f'(x) - f'(x) r^{X_n}\|_{X^*} \quad \rightarrow_{n\rightarrow\infty} \quad 0$$

Further we can express the error of the approximation in the ∇-notation by

$$\|f'(x) - q^{X_n} f_n'(x)\|_{X^*} = \|\nabla f(x) - (r^{X_n})^* \nabla f_n(x)\|_X$$

where $\nabla f_n(x)$ denotes the Riesz representation of $f_n'(x)$ with respect to the domain X_n.

Proof. While the first property is a direct implication of the convergence of the dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$, the error representation follows from the following simple calculation:

$$\begin{aligned} \|f'(x) - q^{X_n} f_n'(x)\|_{X^*} &= \sup_{\|h\|=1, h\in X} |f'(x)h - f_n'(x) r^{X_n} h| \\ &= \sup_{\|h\|=1, h\in X} | < \nabla f(x), h > - < \nabla f_n(x), r^{X_n} h > | \\ &= \sup_{\|h\|=1, h\in X} | < \nabla f(x) - (r^{X_n})^* \nabla f_n(x), h > | \\ &= \|\nabla f(x) - (r^{X_n})^* \nabla f_n(x)\|_X \end{aligned}$$

□

Now we will focus on those summands in (4.4) that involve derivatives of the operator $c(\cdot)$. In the next lemma we will not include the convergence of the dual approximations $(Y_m^*, q^{Y_m})_{m\in I\!N}$:

Lemma 4.3.3. *Let the approximations $(X_n, r^{X_n})_{n\in I\!N}$ and $(Y_m, r^{Y_m})_{m\in I\!N}$ be convergent. Further assume that the dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$ converge and that the discretization of the Fréchet differentiable operator $c(\cdot)$ is defined as in (4.2). Then*

$$\sup_{\|h\|=1, h\in X} | < y, c'(x)h - c'(x) r^{X_n} h > | \rightarrow_{n\rightarrow\infty} 0 \qquad \forall y \in Y, \quad \forall x \in X$$

$$\|c'(x) - c_{n,m}'(x)\|_{\mathcal{L}(X_n, Y)} \longrightarrow_{m\rightarrow\infty} 0 \qquad \forall n \in I\!N, \quad \forall x \in X_n$$

Proof. We start with the first property. Note that

$$\sup_{\|h\|=1, h\in X} | < y, c'(x)h - c'(x)r^{X_n}h > | \quad = \quad \sup_{\|h\|=1, h\in X} | < c'(x)^* y, h - r^{X_n}h > |$$

But x and y are fixed, thus $< c'(x)^*y, \cdot > \in X^*$. Hence the first claim follows from the convergence of the dual approximations. Now we will focus on the second claim:

For fixed discretization n of X we have that X_n is spanned by finitely many basis elements $\phi_1, ..., \phi_{dim(X_n)}$. Hence we have for all $h \in X_n$ with $\|h\| = 1$ that $h = \sum_{i=1}^{dim(X_n)} \lambda_i \phi_i$ with $\lambda_i \in I\!R$ and $|\lambda_i| \leq \lambda_{max}$ for $i = 1, ..., dim(X_n) =: \hat{n}$ and some $\lambda_{max} > 0$. But then we obtain by the linearity of $c'(x)(\cdot)$

$$\begin{aligned}
\|c'(x) - c'_{n,m}(x)\|_{\mathcal{L}(X_n,Y)} &= \sup_{h\in X_n, \|h\|=1} \|c'(x)h - c'_{n,m}(x)h\|_Y \\
&\leq \sup_{|\lambda_i|\leq\lambda_{max}} \|c'(x)\sum_{i=1}^{\hat{n}} \lambda_i\phi_i - c'_{n,m}(x)\sum_{i=1}^{\hat{n}}\lambda_i\phi_i\|_Y \\
&\leq \lambda_{max}\sum_{i=1}^{\hat{n}} \|c'(x)\phi_i - c'_{n,m}(x)\phi_i\|_Y \\
&\overset{\text{Th.4.3.1}}{=} \lambda_{max}\sum_{i=1}^{\hat{n}} \|c'(x)\phi_i - r^{Y_m}c'(x)\phi_i\|_Y
\end{aligned}$$

But due to the convergence of the restrictions $(r^{Y_m})_{m\in I\!N}$, the right hand side of this inequality converges for fixed $n \in I\!N$ and thus fixed $\hat{n}$ to zero as m tends to infinity. □

We use this lemma to prove that the augmented Lagrangian functional $\Phi(x, \lambda, \mu)$ can in fact be approximated arbitrarily close by its discretized counterpart $\Phi_{n,m}(x, \lambda, \mu)$:

Theorem 4.3.4. *Let the approximations $(X_n, r^{X_n})_{n\in I\!N}$ and $(Y_m, r^{Y_m})_{m\in I\!N}$ be convergent. Further assume that the dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$ converge and that the discretizations of the Fréchet differentiable operators $f(\cdot)$ and $c(\cdot)$ are defined as in (4.2). For fixed n_k, $m_k \in I\!N$ let $x \in X_{n_k}$, $\lambda \in Y_{m_k}$ and $\mu \in (0, \infty)$ be given. Then $\forall\varepsilon > 0$ there exist n, $m \in I\!N$, $n \geq n_k$, $m \geq m_k$, with m depending on n such that*

$$\left\| \frac{d}{dx}\Phi(x, \lambda, \mu) \cdot - \frac{d}{dx}\Phi_{n,m}(x, \lambda, \mu) r^{X_n} \cdot \right\|_{L(X, I\!R)} < \varepsilon$$

Proof. Let $n \geq n_k$ and $m \geq m_k$. Then we obtain

$$
\begin{aligned}
&\| \frac{d}{dx}\Phi(x,\lambda,\mu)\cdot - \frac{d}{dx}\Phi_{n,m}(x,\lambda,\mu) r^{X_n}\cdot \|_{L(X,I\!R)} \\
&\quad = \sup_{\|h\|=1, h\in X} |f'(x)h + <\lambda, c'(x)h> + \frac{1}{\mu} < c(x), c'(x)h> - f'(x)r^{X_n}h \\
&\qquad\qquad - <\lambda, r^{Y_m}c'(x)r^{X_n}h> - \frac{1}{\mu} < r^{Y_m}c(x), r^{Y_m}c'(x)r^{X_n}h> | \\
&\quad \leq \sup_{\|h\|=1, h\in X} |f'(x)h - f'(x)r^{X_n}h| + \sup_{\|h\|=1, h\in X} | <\lambda, c'(x)h - r^{Y_m}c'(x)r^{X_n}h> | \\
&\qquad\qquad + \sup_{\|h\|=1, h\in X} \frac{1}{\mu} | < c(x), c'(x)h> - < r^{Y_m}c(x), r^{Y_m}c'(x)r^{X_n}h> |
\end{aligned} \tag{4.5}
$$

Now we will show that we can choose n and m such that all summands on the right hand side of this inequality are arbitrarily small. This is trivial in case of the first summand, see Lemma 4.3.2. The second summand can be estimated by

$$
\begin{aligned}
&\sup_{\|h\|=1, h\in X} | <\lambda, c'(x)h - r^{Y_m}c'(x)r^{X_n}h> | \\
&\quad \leq \sup_{\|h\|=1, h\in X} | <\lambda, c'(x)h - c'(x)r^{X_n}h> | \\
&\qquad\qquad + \sup_{\|h\|=1, h\in X} | <\lambda, c'(x)r^{X_n}h - r^{Y_m}c'(x)r^{X_n}h> | \\
&\quad \leq \sup_{\|h\|=1, h\in X} | < c'(x)^*\lambda, h - r^{X_n}h> | \\
&\qquad\qquad + \|\lambda\| \|r^{X_n}\| \sup_{\|h\|=1, h\in X_n} \|c'(x)h - r^{Y_m}c'(x)h\|_Y
\end{aligned} \tag{4.6}
$$

But now we can choose n according to Lemma 4.3.3 to make the left summand in (4.6) small. Since $\|r^{X_n}\|$ is bounded by the uniform boundedness principle, we can then choose m according to the same Lemma in order to make the right summand in (4.6) small enough. Note that m depends on n, because the supremum of the right summand is with respect to $h \in X_n$. Finally, we estimate the last summand in (4.5):

$$
\begin{aligned}
&\sup_{\|h\|=1, h\in X} \frac{1}{\mu} | < c(x), c'(x)h> - < r^{Y_m}c(x), r^{Y_m}c'(x)r^{X_n}h> | \\
&\quad \leq \frac{1}{\mu} \sup_{\|h\|=1, h\in X} | < c(x), c'(x)h - r^{Y_m}c'(x)r^{X_n}h> | + \\
&\qquad\qquad + \frac{1}{\mu} \sup_{\|h\|=1, h\in X} | < c(x) - r^{Y_m}c(x), r^{Y_m}c'(x)r^{X_n}h> |
\end{aligned}
$$

$$
\begin{aligned}
\leq\ & \frac{1}{\mu} \sup_{\|h\|=1, h\in X} |< c(x), c'(x)h - c'(x)r^{X_n}h > | \\
& \qquad + \frac{1}{\mu} \sup_{\|h\|=1, h\in X} |< c(x), c'(x)r^{X_n}h - r^{Y_m}c'(x)r^{X_n}h > | \\
& \qquad + \frac{1}{\mu} \|c(x) - r^{Y_m}c(x)\|_Y \|r^{Y_m}\| \|c'(x)\| \|r^{X_n}\|
\end{aligned} \tag{4.7}
$$

The first summand in (4.7) can be made arbitrarily small due to Lemma 4.3.3, while the last summand converges to zero, because the restrictions $(r^{Y_m})_m$ are convergent. Thus we only have to estimate the second summand in (4.7):

$$
\begin{aligned}
& \frac{1}{\mu} \sup_{\|h\|=1, h\in X} |< c(x), c'(x)r^{X_n}h - r^{Y_m}c'(x)r^{X_n}h > | \\
& \qquad \leq \frac{1}{\mu} \|c(x)\| \|r^{X_n}\| \sup_{\|h\|=1, h\in X_n} \|c'(x)h - r^{Y_m}c'(x)h\|
\end{aligned} \tag{4.8}
$$

Again we can apply Lemma 4.3.3 to choose m such that the right hand side is as small as desired. □

Remark: The last theorem and its proof make apparent why we treat the discretizations of X and Y separately. More precisely: If we would not do this, the second property in Lemma 4.3.3 would not hold in general without assuming the convergence of the dual approximations $(Y_m^*, q^{Y_m})_m$. In order to motivate this, assume that the discretizations of X and Y are not treated separately, i.e. $n = m$. Then the discretization of $c(\cdot)$ is given by $\tilde{c}_n(\cdot) := c_{n,n}(\cdot) = r^{Y_n}c(\cdot)|_{X_n}$.

While the convergence of the restrictions $(r^{Y_n})_{n\in I\!N}$ still implies the pointwise convergence of the mappings $\tilde{c}_n(\cdot)$ to $c(\cdot)$, we cannot deduce in general that

$$
\|c'(x) - \tilde{c}_n{}'(x)\|_{\mathcal{L}(X_n,Y)} = \sup_{\|h\|=1,\ h\in X_n} \|c'(x)h - r^{Y_n}c'(x)h\| \rightarrow_{n\to\infty} 0 \tag{4.9}
$$

The reason for the problem is, that we can only approximate finitely many vectors in Y arbitrarily close with the restriction operators $(r^{Y_n})_{n\in I\!N}$. But note that the spaces X_n and the norms $\|\cdot\|_{\mathcal{L}(X_n,Y)}$ also change as n tends to infinity. For a general proof of (4.9) we would thus have to take all $h \in X$ with $\|h\| = 1$ into account. But this resulting uniform convergence cannot be proved for pointwise convergent linear operators. In fact, one can easily find counterexamples where this does not work (see for example [27, page 76]).

As proven, we can approximate the derivative of the augmented Lagrangian functional arbitrarily close without assuming the convergence of the dual approximations $(Y_m^*, q^{Y_m})_m$ by treating the discretizations separately. This can be an advantage in cases where the space Y has a complicated structure which does not allow a nice formula for convergent dual approximations. For example, if $c : X \to Y$ is a differential operator, the elements of the space

X are “smoother” than their images. Thus X describes a “nicer” class of functions than Y does.

The problems described in the last remark are not relevant if we also assume the convergence of the dual approximations $(Y_m^*, q^{Y_m})_m$. In this case we obtain the following theorem which shows the independence of the discretization levels n and m:

Theorem 4.3.5. *Let the approximations $(X_n, r^{X_n})_{n\in I\!N}$ and $(Y_m, r^{Y_m})_{m\in I\!N}$ be convergent. Further assume that the dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$ and $(Y_m^*, q^{Y_m})_{m\in I\!N}$ converge and that the discretizations of the Fréchet differentiable operators $f(\cdot)$ and $c(\cdot)$ are defined as in (4.2). For fixed $n_k,\ m_k \in I\!N$ let $x \in X_{n_k},\ \lambda \in Y_{m_k}$ and $\mu \in (0,\infty)$ be given. Then*

$$\lim_{(n,m)\to(\infty,\infty)} \left\| \frac{d}{dx}\Phi(x,\lambda,\mu)\cdot - \frac{d}{dx}\Phi_{n,m}(x,\lambda,\mu) r^{X_n}\cdot \right\|_{L(X,I\!R)} = 0$$

Proof. We can use nearly all results derived in the proof of Theorem 4.3.4. As indicated in the last remark, we only need to change those parts of the proof where we explicitely used the second property of Lemma 4.3.3. We will now analyze these parts and derive alternative estimates which do not impose a relation of the discretization levels n and m. We estimate instead of (4.6):

$$\begin{aligned} &\sup_{\|h\|=1, h\in X} | < \lambda, c'(x) r^{X_n} h - r^{Y_m} c'(x) r^{X_n} h > | \\ &\qquad \le \ \|c'(x)\| \|r^{X_n}\| \sup_{\|y\|=1, y\in Y} | < \lambda, y - r^{Y_m} y > | \end{aligned} \tag{4.10}$$

But $\|r^{X_n}\|$ is bounded and thus the right hand side converges to zero due to the convergence of the dual approximations $(Y_m^*, q^{Y_m})_m$. With a similar argumentation we replace (4.8) by

$$\begin{aligned} &\frac{1}{\mu} \sup_{\|h\|=1, h\in X} | < c(x), c'(x) r^{X_n} h - r^{Y_m} c'(x) r^{X_n} h > | \\ &\qquad \le \ \frac{1}{\mu} \|c'(x)\| \|r^{X_n}\| \sup_{\|y\|=1, y\in Y} | < c(x), y - r^{Y_m} y > | \to_{m\to\infty} 0 \end{aligned} \tag{4.11}$$

□

In this section we discretized the mappings involved in our optimization problem and introduced the concepts which are necessary to approximate the augmented Lagrangian functional. In the next section we will briefly introduce the final step in the discretization process: The representation in terms of $I\!R^n$ and $I\!R^m$.

4.4 Interpretation by Isomorphisms

By now we have interpreted all major elements of the infinite dimensional method in terms of finite dimensional Hilbert spaces X_n and Y_m of X and Y. Our goal in this section is to represent elements of those spaces and the involved mappings $f_n(\cdot)$ and $c_{n,m}(\cdot)$ in terms of a finite set of numbers.

It is a standard approach to make this final step by linear isomorphisms called *prolongations*

$$p^{X_n} : I\!R^{dim(X_n)} \to X_n, \quad p^{X_n} \in \mathcal{L}(I\!R^{dim(X_n)}, X_n)$$

$$p^{Y_m} : I\!R^{dim(Y_m)} \to Y_m, \quad p^{Y_m} \in \mathcal{L}(I\!R^{dim(Y_m)}, Y_m)$$

which represent the Hilbert space elements in terms of their basis coefficients. Define $\hat{n} := dim(X_n)$ and assume that X_n is spanned by the basis elements $\phi_1, ..., \phi_{\hat{n}} \in X$. Then we define $p^{X_n} : I\!R^{\hat{n}} \to X_n$ by

$$p^{X_n}\alpha := \sum_{i=1}^{\hat{n}} \alpha_i \phi_i$$

for $\alpha \in I\!R^{\hat{n}}$. In analogy we define p^{Y_m} with respect to the basis elements spanning Y_m. It is a well known fact that the mappings $p^{X_n}(\cdot)$ and $p^{Y_m}(\cdot)$ so defined are continuous and linear bijections. Note that $I\!R^{\hat{n}}$ is a Hilbert space with scalar product

$$< \alpha, \beta >_{I\!R^{\hat{n}}} := < p^{X_n}\alpha, p^{X_n}\beta >_{X_n} = < p^{X_n}\alpha, p^{X_n}\beta >_X$$

and induced norm

$$\| \cdot \|_{I\!R^{\hat{n}}} := \sqrt{< \cdot, \cdot >_{I\!R^{\hat{n}}}} = \sqrt{< p^{X_n}\cdot, p^{X_n}\cdot >_X}$$

In analogy $I\!R^{\hat{m}}$, $\hat{m} := dim(Y_m)$, is a Hilbert space with a similar defined scalar product and induced norm. Then one defines the final discretized mappings $\widehat{f_n} : I\!R^{\hat{n}} \to I\!R$ and $\widehat{c_{n,m}} : I\!R^{\hat{n}} \to I\!R^{\hat{m}}$ by

$$\widehat{f_n} := f_n p^{X_n} = f\, p^{X_n}$$

$$\widehat{c_{n,m}} := (p^{Y_m})^{-1} c_{n,m} p^{X_n} = (p^{Y_m})^{-1} r^{Y_m} c\, p^{X_n}$$

Clearly, these final discretizations are again twice continuous Fréchet differentiable with respect to the norms mentioned above, if the approximated operators $f(\cdot)$ and $c(\cdot)$ are twice continuous Fréchet differentiable.

As no additional error is introduced by the isomorphisms, it does not make a difference if we discuss algorithms in terms of the mappings $f_n(\cdot)$ and $c_{n,m}(\cdot)$ or $\widehat{f_n}(\cdot)$ and $\widehat{c_{n,m}}(\cdot)$. However, with the intention to keep notation simple, we will take the approach without isomorphisms. They can easily be implemented after the theory is derived in the next chapter.

Now we are ready to use the introduced approximation theory in order to discuss a discrete algorithm which has algorithm ALINF as a basis. The next chapter will be devoted to this task.

CHAPTER 5

Discretized Algorithm

In Chapter 3 we proved very strong theoretical results for algorithm ALINF. As mentioned before, these theoretical results cannot be used in practice if we are not able to provide a discretized version of this algorithm which generates iterates and Lagrange multiplier estimates in finite dimensional subspaces of X and Y. The derivation of such a "discretized" algorithm is the goal of this chapter.

The tools for the discretization were derived in the last chapter. Given problem (3.1) we make the following general assumption:

AS 7. Let $(X_n, r^{X_n})_{n\in I\!N}$ and $(Y_m, r^{Y_m})_{m\in I\!N}$ be convergent approximations of the Hilbert spaces X and Y such that the finite dimensional subspaces $(X_n)_{n\in I\!N}$ and $(Y_m)_{m\in I\!N}$ are nested. Further assume that the dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$ and $(Y_m^*, q^{Y_m})_{m\in I\!N}$ also converge. Given this setting, let the discretizations of $f(\cdot)$ and $c(\cdot)$ be as in (4.2) and define the approximation of the augmented Lagrangian functional as in (4.3).

In particular this assumption implies the boundedness of the sequences $(\|r^{X_n}\|_{\mathcal{L}(\mathcal{X},\mathcal{X})})_{n\in I\!N}$ and $(\|r^{Y_m}\|_{\mathcal{L}(\mathcal{Y},\mathcal{Y})})_{m\in I\!N}$. We denote their upper bounds by

$$M_X := \sup_{n\in I\!N} \|r^{X_n}\|_{\mathcal{L}(\mathcal{X},\mathcal{X})} \ , \qquad M_Y := \sup_{m\in I\!N} \|r^{Y_m}\|_{\mathcal{L}(\mathcal{Y},\mathcal{Y})} \tag{5.1}$$

Besides AS7 we also assume that AS1 holds.

The most desirable discretized version of an algorithm in an infinite dimensional setting is such that each iterate of the discretized algorithm satisfies all requirements of the original algorithm. In this case all convergence results of the infinite dimensional method can be transferred to the new setting.

In terms of approximation theory, such a situation could be provided by assuming that we have absolute control over the approximation error. But such an assumption is not very realistic, because one would have to request for example, that the dual approximations converge uniformly on compact subsets of the dual space. We do not make such an assumption because it would restrict the applicability of the algorithm to very special cases. Instead of this, we propose a nested algorithm consisting of two parts, where the inner algorithm terminates when we have enough control over the approximation error.

5.1 Description of the Algorithm

The basis for the new algorithm ALDISCR is algorithm ALINF. One of the advantages of ALINF is that it provides an explicit rule of how to adapt the penalty parameter. Now it is our concern to incorporate the discretization aspect and to give an explicit formula for the update of the Lagrange multiplier.

As the new algorithm is a finite dimensional approximation of the infinite dimensional method, the subproblems $\min_{x\in X} \Phi(x, \lambda_k, \mu_k)$ in algorithm ALINF are replaced by the sequence of discretized subproblems

$$\min_{x\in X_{n_k}} \Phi_{n_k,m_k}(x, \lambda_k, \mu_k) \qquad , k \in I\!N \tag{5.2}$$

where $\Phi_{n_k,m_k} : X_{n_k} \times Y_{m_k} \times (0, \infty) \to I\!R$ is defined as in (4.3) and n_k, m_k are the discretization levels of X and Y in iteration k, respectively. Of course these subproblems do not need to be solved precisely during an early iteration. We will adapt the convergence tolerance of ALINF in (3.8) in a suitable manner.

As mentioned before the proposed algorithm shall be “close” to algorithm ALINF in every iteration. The crucial part in this approximation is the derivative of the augmented Lagrangian functional $\Phi(x, \lambda, \mu)$ with respect to x. We proved in the last chapter that this functional can be approximated arbitrarily close by its discretized counterpart $\frac{d}{dx}\Phi_{n,m}(x, \lambda, \mu)$. The following algorithm will include a mesh refinement based on the discretization error.

Algorithm ALDISCR

0. Initialization: Choose initial discretization levels n_0 of X and m_0 of Y. Let the Lagrange multiplier estimate $\lambda_0 \in Y_{m_0}$ and the penalty parameter $0 < \mu_0 < 1$ be given. Further let $w_* \ll 1$, $\eta_* \ll 1$, $\tau < 1$, $\alpha < 1$, $\alpha_\eta < 1$ and $\beta_\eta < 1$ be strictly positive constants. Set $w_0 := \mu_0$, $\eta_0 := \mu_0^{\alpha_\eta}$ and $k := 0$.

1. Inner Algorithm: Set $i := 1$, $n_{k_1} := n_k$ and $m_{k_1} := m_k$.

1a. Inner iteration: Find $x_{k_i} \in X_{n_{k_i}}$ such that

$$\left\| \frac{d}{dx}\Phi_{n_{k_i},m_{k_i}}(x_{k_i},\lambda_k,\mu_k)\cdot \right\|_{\mathcal{L}(X_{n_{k_i}},I\!R)} \leq \frac{w_k}{2iM_X} \tag{5.3}$$

1b. Refinement: Choose $n_{k_{i+1}} \geq n_{k_i},\ m_{k_{i+1}} \geq m_{k_i}$ such that

$$\left\| \frac{d}{dx}\Phi(x_{k_i},\lambda_k,\mu_k)\cdot - \frac{d}{dx}\Phi_{n_{k_{i+1}},m_{k_{i+1}}}(x_{k_i},\lambda_k,\mu_k) r^{X_{n_{k_{i+1}}}}\cdot \right\|_{\mathcal{L}(X,I\!R)} \leq \frac{w_k}{4} \tag{5.4}$$

$$\|c(x_{k_i}) - r^{Y_{m_{k_{i+1}}}}c(x_{k_i})\|_Y < \min\left\{\frac{\eta_*}{2}, \mu_k w_k, \alpha\eta_k\right\} \tag{5.5}$$

1c. Inner algorithm termination criterion: If

$$\left\| \frac{d}{dx}\Phi_{n_{k_{i+1}},m_{k_{i+1}}}(x_{k_i},\lambda_k,\mu_k) r^{X_{n_{k_{i+1}}}}\cdot - \frac{d}{dx}\Phi_{n_{k_i},m_{k_i}}(x_{k_i},\lambda_k,\mu_k) r^{X_{n_{k_i}}}\cdot \right\|_{X^*} \leq \frac{w_k}{4} \tag{5.6}$$

then set $n_k := n_{k_i},\ n_{k+1} := n_{k_{i+1}},\ m_k := m_{k_i},\ m_{k+1} := m_{k_{i+1}},\ x_k := x_{k_i}$ and go to step 2. Otherwise, if (5.6) does not hold, set $i = i + 1$ and go to step 1a.

2. Test for convergence: If $w_k \leq w_*$ and $\|c_{n_{k+1},m_{k+1}}(x_k)\|_Y \leq \eta_*/2$, stop.

3. Updates: If

$$\|c_{n_{k+1},m_{k+1}}(x_k)\|_Y \leq \eta_k \tag{5.7}$$

execute step 3a. If

$$\|c_{n_{k+1},m_{k+1}}(x_k)\|_Y > \eta_k \tag{5.8}$$

execute step 3b.

3a. Update Lagrange multiplier estimate: Choose

$$\lambda_{k+1} := \lambda_k + \frac{1}{\mu_k} c_{n_{k+1},m_{k+1}}(x_k) \tag{5.9}$$

and set

$$\begin{array}{rcl} \mu_{k+1} & := & \mu_k \\ w_{k+1} & := & w_k \mu_{k+1} \\ \eta_{k+1} & := & \eta_k \mu_{k+1}^{\beta_\eta} \end{array} \tag{5.10}$$

3b. Reduce penalty parameter: Set

$$\begin{aligned} \lambda_{k+1} &:= \lambda_k \\ \mu_{k+1} &:= \tau \mu_k \\ w_{k+1} &:= \mu_{k+1} \\ \eta_{k+1} &:= \mu_{k+1}^{\alpha_\eta} \end{aligned} \tag{5.11}$$

Increment k by one and go to Step 1.

■

We will show now that algorithm ALDISCR is well defined. Note first, that the iterates x_k belong to the domains of the discretized operators, because the subspaces $(X_n)_{n\in I\!N}$ are nested. Further the nestedness of $(Y_m)_{m\in I\!N}$ implies that the update of the Lagrange multiplier estimate is well defined.

The refinement conditions (5.4) and (5.5) can be fulfilled due to AS7 and Theorem 4.3.5. However, in order to show that the inner algorithm terminates, we must impose the following additional assumption:

AS 8. Let $(X_i, r^{X_i})_{i\in I\!N}$ be approximations of X as described in AS7. Let $g : X \to I\!R$ and $g_i : X_i \to I\!R$, $i \in I\!N$ be twice continuous Fréchet differentiable mappings satisfying $\forall x \in X_i$

$$|g(x) - g_i(x)| \to_{i\to\infty} 0 \ , \qquad \|g'(x) \cdot - g_i'(x) r^{X_i} \cdot \|_{X^*} \to_{i\to\infty} 0$$

Further let $(e_i)_{i\in I\!N} \subset I\!R$ be a decreasing sequence such that $e_i \downarrow 0$ for $i \to \infty$. Then we assume that the algorithm used to solve (5.3) generates for the problems

$$\|g_i'(x)\|_{\mathcal{L}(X_i, I\!R)} \le e_i$$

a converging sequence $(x_i)_{i\in I\!N}$, given that the algorithm starts problem $i+1$ with the solution x_i of the previous i-th problem.

Using this assumption, we can in fact prove that the inner algorithm terminates and thus algorithm ALDISCR is well defined:

Theorem 5.1.1. *Let AS1, AS7 and AS8 hold. Then the inner algorithm terminates in each iteration of algorithm ALDISCR.*

Proof. Let the outer iteration $k \in I\!N$ be fixed and define $\widetilde{\Phi} : X \to I\!R$ and $\widetilde{\Phi}_i : X_{n_{k_i}} \to I\!R$ by

$$\widetilde{\Phi}(x) := \Phi(x, \lambda_k, \mu_k) \ , \qquad \widetilde{\Phi}_i(x) := \Phi_{n_{k_i}, m_{k_i}}(x, \lambda_k, \mu_k)$$

We can deduce from Theorems 4.3.1 and 4.3.5 that $\widetilde{\Phi}(\cdot)$ and $\widetilde{\Phi}_i(\cdot)$ satisfy all requirements imposed on the mappings $g(\cdot)$ and $g_i(\cdot)$ in AS8, respectively. Further, as the right hand side of (5.3) is a strictly decreasing sequence in i, AS8 implies the convergence of $(x_{k_i})_{i\in I\!N}$ to some limit point $x_k^* \in X$. In order to prove the theorem, we must show that

$$\begin{aligned}
&\left\|\frac{d}{dx}\widetilde{\Phi}_{i+1}(x_{k_i})r^{X_{n_{k_{i+1}}}}\cdot - \frac{d}{dx}\widetilde{\Phi}_i(x_{k_i})r^{X_{n_{k_i}}}\cdot\right\|_{X^*} = \\
&\quad = \left\|\frac{d}{dx}\Phi_{n_{k_{i+1}},m_{k_{i+1}}}(x_{k_i},\lambda_k,\mu_k)r^{X_{n_{k_{i+1}}}}\cdot - \frac{d}{dx}\Phi_{n_{k_i},m_{k_i}}(x_{k_i},\lambda_k,\mu_k)r^{X_{n_{k_i}}}\cdot\right\|_{X^*}
\end{aligned}$$

converges to zero for $i \to \infty$. Note that

$$\begin{aligned}
&\left\|\frac{d}{dx}\widetilde{\Phi}_{i+1}(x_{k_i})r^{X_{n_{k_{i+1}}}}\cdot - \frac{d}{dx}\widetilde{\Phi}_i(x_{k_i})r^{X_{n_{k_i}}}\cdot\right\|_{X^*} = \\
&\quad = \sup_{\|h\|=1,h\in X} |f'(x_{k_i})r^{X_{n_{k_{i+1}}}}h - f'(x_{k_i})r^{X_{n_{k_i}}}h + < \lambda_k, r^{Y_{m_{k_{i+1}}}}c'(x_{k_i})r^{X_{n_{k_{i+1}}}}h > \\
&\qquad - < \lambda_k, r^{Y_{m_{k_i}}}c'(x_{k_i})r^{X_{n_{k_i}}}h > + \frac{1}{\mu_k} < r^{Y_{m_{k_{i+1}}}}c(x_{k_i}), r^{Y_{m_{k_{i+1}}}}c'(x_{k_i})r^{X_{n_{k_{i+1}}}}h > \\
&\qquad - \frac{1}{\mu_k} < r^{Y_{m_{k_i}}}c(x_{k_i}), r^{Y_{m_{k_i}}}c'(x_{k_i})r^{X_{n_{k_i}}}h > |
\end{aligned} \tag{5.12}$$

We will split up the summands of the right hand side and show that they converge to zero:

$$\begin{aligned}
&\|f'(x_{k_i})r^{X_{n_{k_{i+1}}}}\cdot - f'(x_{k_i})r^{X_{n_{k_i}}}\cdot\|_{X^*} \leq \\
&\quad \leq \|f'(x_{k_i})r^{X_{n_{k_{i+1}}}}\cdot - f'(x_k^*)r^{X_{n_{k_{i+1}}}}\cdot\| + \|f'(x_k^*)r^{X_{n_{k_{i+1}}}}\cdot - f'(x_k^*)\cdot\| \\
&\qquad + \|f'(x_k^*)\cdot - f'(x_k^*)r^{X_{n_{k_i}}}\cdot\| + \|f'(x_k^*)r^{X_{n_{k_i}}}\cdot - f'(x_{k_i})r^{X_{n_{k_i}}}\cdot\| \\
&\quad \leq \|r^{X_{n_{k_{i+1}}}}\|\|f'(x_{k_i})\cdot - f'(x_k^*)\cdot\| + \|f'(x_k^*)r^{X_{n_{k_{i+1}}}}\cdot - f'(x_k^*)\cdot\| \\
&\qquad + \|f'(x_k^*)\cdot - f'(x_k^*)r^{X_{n_{k_i}}}\cdot\| + \|r^{X_{n_{k_i}}}\|\|f'(x_k^*)\cdot - f'(x_{k_i})\cdot\|
\end{aligned} \tag{5.13}$$

The second and third summand in (5.13) converge to zero due to the convergence of the dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$, while the continuity of $f'(\cdot)$ takes care of the first and last summand. Now we analyze the second pair in (5.12).

$$\sup_{\|h\|=1,h\in X} | < \lambda_k, r^{Y_{m_{k_{i+1}}}}c'(x_{k_i})r^{X_{n_{k_{i+1}}}}h > - < \lambda_k, r^{Y_{m_{k_i}}}c'(x_{k_i})r^{X_{n_{k_i}}}h > | =$$

$$\begin{aligned}
&= \sup_{\|h\|=1, h\in X} | < \lambda_k, r^{Y_{m_{k_{i+1}}}} c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h - c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h > \\
&\quad + < \lambda_k, c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h - c'(x_k^*) r^{X_{n_{k_{i+1}}}} h > + < \lambda_k, c'(x_k^*) r^{X_{n_{k_{i+1}}}} h - c'(x_k^*) h > \\
&\quad + < \lambda_k, c'(x_k^*) h - c'(x_k^*) r^{X_{n_{k_i}}} h > + < \lambda_k, c'(x_k^*) r^{X_{n_{k_i}}} h - c'(x_{k_i}) r^{X_{n_{k_i}}} h > \\
&\quad + < \lambda_k, c'(x_{k_i}) r^{X_{n_{k_i}}} h - r^{Y_{m_{k_i}}} c'(x_{k_i}) r^{X_{n_{k_i}}} h > | \\
&\leq \|r^{X_{n_{k_{i+1}}}}\| \|c'(x_{k_i})\| \sup_{\|y\|=1, y\in Y} | < \lambda_k, r^{Y_{m_{k_{i+1}}}} y - y > | \\
&\quad + \|\lambda_k\| \|r^{X_{n_{k_{i+1}}}}\| \|c'(x_{k_i}) - c'(x_k^*)\| + \sup_{\|h\|=1, h\in X} | < c'(x_k^*)^* \lambda_k, r^{X_{n_{k_{i+1}}}} h - h > | \\
&\quad + \sup_{\|h\|=1, h\in X} | < c'(x_k^*)^* \lambda_k, h - r^{X_{n_{k_i}}} h > + \|\lambda_k\| \|r^{X_{n_{k_i}}}\| \|c'(x_k^*) - c'(x_{k_i})\| \\
&\quad + \|r^{X_{n_{k_i}}}\| \|c'(x_{k_i})\| \sup_{\|y\|=1, y\in Y} | < \lambda_k, y - r^{Y_{m_{k_i}}} y > | \qquad (5.14)
\end{aligned}$$

But now the right hand side of (5.14) converges to zero due to the convergence of the dual approximations of X^* and Y^* and the continuity of $c'(\cdot)$. Finally, we analyze the last pair in (5.12). We do not consider the factor $1/\mu_k$, because it is independent of i:

$$\begin{aligned}
&\sup_{\|h\|=1, h\in X} | < r^{Y_{m_{k_i}}} c(x_{k_i}), r^{Y_{m_{k_i}}} c'(x_{k_i}) r^{X_{n_{k_i}}} h > - < r^{Y_{m_{k_{i+1}}}} c(x_{k_i}), r^{Y_{m_{k_{i+1}}}} c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h > | \\
&= \sup_{\|h\|=1, h\in X} | < r^{Y_{m_{k_i}}} c(x_{k_i}) - c(x_k^*), r^{Y_{m_{k_i}}} c'(x_{k_i}) r^{X_{n_{k_i}}} h > \\
&\qquad + < c(x_k^*), r^{Y_{m_{k_i}}} c'(x_{k_i}) r^{X_{n_{k_i}}} h - r^{Y_{m_{k_{i+1}}}} c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h > \\
&\qquad + < c(x_k^*) - r^{Y_{m_{k_{i+1}}}} c(x_{k_i}), r^{Y_{m_{k_{i+1}}}} c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h > | \\
&\leq \|c(x_k^*) - r^{Y_{m_{k_i}}} c(x_{k_i})\| \|r^{Y_{m_{k_i}}}\| \|c'(x_{k_i})\| \|r^{X_{n_{k_i}}}\| \\
&\quad + \|c(x_k^*) - r^{Y_{m_{k_{i+1}}}} c(x_{k_i})\| \|r^{Y_{m_{k_{i+1}}}}\| \|c'(x_{k_i})\| \|r^{X_{n_{k_{i+1}}}}\| \\
&\quad + \sup_{\|h\|=1, h\in X} | < c(x_k^*), r^{Y_{m_{k_i}}} c'(x_{k_i}) r^{X_{n_{k_i}}} h - c'(x_{k_i}) r^{X_{n_{k_i}}} h > \\
&\qquad + < c(x_k^*), c'(x_{k_i}) r^{X_{n_{k_i}}} h - c'(x_k^*) r^{X_{n_{k_i}}} h > \\
&\qquad + < c(x_k^*), c'(x_k^*) r^{X_{n_{k_i}}} h - c'(x_k^*) h > \\
&\qquad + < c(x_k^*), c'(x_k^*) h - c'(x_k^*) r^{X_{n_{k_{i+1}}}} h > \\
&\qquad + < c(x_k^*), c'(x_k^*) r^{X_{n_{k_{i+1}}}} h - c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h > \\
&\qquad + < c(x_k^*), c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h - r^{Y_{m_{k_{i+1}}}} c'(x_{k_i}) r^{X_{n_{k_{i+1}}}} h > |
\end{aligned}$$

Now we can use Theorem 4.1.12, the convergence of the approximations and the dual approximations, and the continuity of $c'(\cdot)$ to show, that the right hand side of this inequality converges to zero. □

We proved, that algorithm ALDISCR is well defined. However, the mesh refinement condition (5.4) is not formulated in a satisfactory manner. The following theorem will give a sufficient condition for (5.4):

Theorem 5.1.2. *Assume that AS1 and AS7 hold. If $n_{k_{i+1}} \geq n_{k_i}$ and $m_{k_{i+1}} \geq m_{k_i}$ are chosen such that all the following conditions are satisfied, then the mesh refinement conditions (5.4) and (5.5) in step 1b hold:*

$$\begin{gathered}
\left\| \nabla f(x_{k_i}) - (r^{X_{n_{k_{i+1}}}})^* \nabla f_{n_{k_{i+1}}}(x_{k_i}) \right\|_X \leq \frac{w_k}{20} \\
\left\| c(x_{k_i}) - r^{Y_{m_{k_{i+1}}}} c(x_{k_i}) \right\|_Y < \min \left\{ \frac{\eta_*}{2}, \alpha\eta_k, \frac{\mu_k w_k}{\max\{1, 40 M_X M_Y \|c'(x_{k_i})\|\}} \right\} \\
\sup_{\|h\|=1, h \in X} | < c'(x_{k_i})^* \lambda_k, h - r^{X_{n_{k_{i+1}}}} h > | \leq \frac{w_k}{20} \\
\sup_{\|h\|=1, h \in X} | < c'(x_{k_i})^* c(x_{k_i}), h - r^{X_{n_{k_{i+1}}}} h > | \leq \frac{w_k \mu_k}{20} \\
\sup_{\|y\|=1, y \in Y} | < \lambda_k, y - r^{Y_{m_{k_{i+1}}}} y > | \leq \frac{w_k}{20 M_X \|c'(x_{k_i})\|} \\
\sup_{\|y\|=1, y \in Y} | < c(x_{k_i}), y - r^{Y_{m_{k_{i+1}}}} y > | \leq \frac{w_k \mu_k}{40 M_X \|c'(x_{k_i})\|}
\end{gathered}$$

Proof. The constants on the right hand side of the inequalities are chosen in such a way that the crucial inequalities in the proofs of Theorems 4.3.4 and 4.3.5 can be estimated by fractions of w_k which add up to the constant $w_k/4$ for the refinement condition (5.4). As condition (5.5) is included in the list above, the claim follows. □

The unhandy conditions in the theorem above simplify in case of optimal restrictions to the following inequalities which can be verified easily:

Corollary 5.1.3. *Let AS1 and AS7 hold. Further assume that the restrictions $(r^{X_n})_{n\in\mathbb{N}}$ and $(r^{Y_m})_{m\in\mathbb{N}}$ are optimal. Then, if $n_{k_{i+1}} \geq n_{k_i}$ and $m_{k_{i+1}} \geq m_{k_i}$ are chosen such that all the following conditions are satisfied, the mesh refinement conditions (5.4) and (5.5) in step 1b hold:*

$$\begin{gathered}
\left\| \nabla f(x_{k_i}) - (r^{X_{n_{k_{i+1}}}})^* \nabla f_{n_{k_{i+1}}}(x_{k_i}) \right\|_X \leq \frac{w_k}{20} \\
\|c'(x_{k_i})^* \lambda_k - r^{X_{n_{k_{i+1}}}} c'(x_{k_i})^* \lambda_k\|_X \leq \frac{w_k}{20} \\
\|c'(x_{k_i})^* c(x_{k_i}) - r^{X_{n_{k_{i+1}}}} c'(x_{k_i})^* c(x_{k_i})\|_X \leq \frac{w_k \mu_k}{20} \\
\left\| c(x_{k_i}) - r^{Y_{m_{k_{i+1}}}} c(x_{k_i}) \right\|_Y < \min \left\{ \frac{\eta_*}{2}, \alpha\eta_k, \frac{\mu_k w_k}{\max\{1, 40 M_X M_Y \|c'(x_{k_i})\|, 40 M_X \|c'(x_{k_i})\|\}} \right\} \\
\|\lambda_k - r^{Y_{m_{k_{i+1}}}} \lambda_k\|_Y \leq \frac{w_k}{20 M_X \|c'(x_{k_i})\|}
\end{gathered}$$

Proof. The claim directly follows from Theorem 5.1.2 if one uses the major inequality from the proof of Theorem 4.2.4. □

5.2 Convergence Analysis

In order to apply the convergence results of algorithm ALINF, we need to do the following: For given initial parameters in algorithm ALDISCR, especially the new parameter α, we must show that the sequences $(x_k)_{k\in\mathbb{N}}$ and $(\lambda_k)_{k\in\mathbb{N}}$ satisfy the decision rules of algorithm ALINF in every iteration. Or in other words: we must show that those sequences could also have been generated by algorithm ALINF.

We will prove this by advancing step by step through an arbitrary iteration of algorithm ALINF and show that the decisions and updates of algorithm ALDISCR satisfy those of ALINF. We start with the inner iteration termination criterion (3.8) of ALINF in step 1, and step 2, the test for convergence:

Lemma 5.2.1. *Let AS 1, AS 7 and AS 8 hold. Let $(x_k)_{k\in\mathbb{N}}$ be the sequence of iterates generated by algorithm ALDISCR. Then*

i) Each iterate x_k also satisfies the inner iteration termination criterion of algorithm ALINF, i.e.

$$\left\|\frac{d}{dx}\Phi(x_k,\lambda_k,\mu_k)\cdot\right\|_{X^*}\leq w_k$$

ii) If algorithm ALDISCR is terminated at iteration k_0, algorithm ALINF would also terminate.

Proof. First we prove item i). Using (5.3), (5.4) and (5.6) we obtain

$$\begin{aligned}
&\left\|\frac{d}{dx}\Phi(x_k,\lambda_k,\mu_k)\cdot\right\|_{X^*}\leq\\
&\quad\leq\ \left\|\frac{d}{dx}\Phi(x_k,\lambda_k,\mu_k)\cdot-\frac{d}{dx}\Phi_{n_{k+1},m_{k+1}}(x_k,\lambda_k,\mu_k)r^{X_{n_{k+1}}}\cdot\right\|_{X^*}\\
&\qquad+\left\|\frac{d}{dx}\Phi_{n_{k+1},m_{k+1}}(x_k,\lambda_k,\mu_k)r^{X_{n_{k+1}}}\cdot-\frac{d}{dx}\Phi_{n_k,m_k}(x_k,\lambda_k,\mu_k)r^{X_{n_k}}\cdot\right\|_{X^*}\\
&\qquad+\left\|\frac{d}{dx}\Phi_{n_k,m_k}(x_k,\lambda_k,\mu_k)r^{X_{n_k}}\cdot\right\|_{X^*}\\
&\quad\leq\ \frac{w_k}{4}+\frac{w_k}{4}+\left\|r^{X_{n_k}}\right\|\left\|\frac{d}{dx}\Phi_{n_k,m_k}(x_k,\lambda_k,\mu_k)\cdot\right\|_{\mathcal{L}(X_{n_k},\mathbb{R})}\\
&\quad\leq\ \frac{w_k}{2}+M_X\frac{w_k}{2M_X}\ =\ w_k
\end{aligned}$$

Now we prove part ii) of the lemma: For iteration k_0 we obtain by step 2 in algorithm ALDISCR: $\|c_{n_{k_0+1},m_{k_0+1}}(x_{k_0})\| \leq \frac{\eta_*}{2}$. Thus we can deduce by (5.5)

$$\|c(x_{k_0})\| \ \leq \ \|c(x_{k_0}) - c_{n_{k_0+1},m_{k_0+1}}(x_{k_0})\| + \|c_{n_{k_0+1},m_{k_0+1}}(x_{k_0})\| \ \leq \ \frac{\eta_*}{2} + \frac{\eta_*}{2} \ = \ \eta_*$$

By part i) of the lemma and step 2 of algorithm ALDISCR, the test for convergence, we finally obtain

$$\|\nabla_x \Phi(x_{k_0}, \lambda_{k_0}, \mu_{k_0})\| \ \leq \ w_{k_0} \ \leq \ w_*$$

□

We will show now that the Lagrange multiplier estimate as chosen in step 3a of algorithm ALDISCR satisfies the requirements in step 3a of algorithm ALINF. The reason why we analyze this before the decision rule in step 3 will become clear in the proof of Lemma 5.2.3.

Lemma 5.2.2. *Let AS 1, AS 7 and AS 8 hold. Let $(x_k)_{k\in\mathbb{N}}$, $(\lambda_k)_{k\in\mathbb{N}}$ and $(\mu_k)_{k\in\mathbb{N}}$ be the sequences generated by algorithm ALDISCR. Then the Lagrange multiplier estimates λ_{k+1} in step 3a of algorithm ALDISCR satisfy the requirement of algorithm ALINF, i.e.:*

$$\|\lambda_{k+1} - \bar{\lambda}(x_k, \lambda_k, \mu_k)\| \leq w_k$$

Proof. By the definition of λ_{k+1} in step 3a of algorithm ALDISCR and the definition of $\bar{\lambda}(x_k, \lambda_k, \mu_k)$ on page 42, we obtain

$$\begin{aligned} \|\lambda_{k+1} - \bar{\lambda}(x_k, \lambda_k, \mu_k)\| &= \left\| \lambda_k + \frac{1}{\mu_k} c_{n_{k+1},m_{k+1}}(x_k) - \left(\lambda_k + \frac{c(x_k)}{\mu_k} \right) \right\| \\ &= \frac{1}{\mu_k} \|c_{n_{k+1},m_{k+1}}(x_k) - c(x_k)\| \\ &\overset{(5.5)}{\leq} \frac{1}{\mu_k} \mu_k w_k \ = \ w_k \end{aligned}$$

□

Finally we analyze the decision rule for the updates in step 3:

Lemma 5.2.3. *Let AS 1, AS 7 and AS 8 hold. Let $(x_k)_{k\in\mathbb{N}}$ be the sequence of iterates generated by algorithm ALDISCR. Pick α as initialized in algorithm ALDISCR and define $\gamma_1 := 1 - \alpha$, $\gamma_2 := 1 + \alpha$. Then the decision rule of algorithm ALINF is fulfilled, i.e.:*

If $\|c(x_k)\| \leq \gamma_1 \eta_k$ then step 3a of algorithm ALINF is executed.
If $\|c(x_k)\| \geq \gamma_2 \eta_k$ then step 3b of algorithm ALINF is executed.

Proof. We have by (5.5) that $\|c(x_k) - c_{n_{k+1},m_{k+1}}(x_k)\| < \alpha\eta_k$. Denote this inequality by (I). Now suppose that $\|c(x_k)\| \leq \gamma_1\eta_k$. Then

$$\begin{aligned}\|c_{n_{k+1},m_{k+1}}(x_k)\| &\leq \|c(x_k)\| + \|c(x_k) - c_{n_{k+1},m_{k+1}}(x_k)\| \\ &\overset{(I)}{<} \gamma_1\eta_k + \alpha\eta_k \ = \ \eta_k\end{aligned}$$

Hence, by algorithm ALDISCR, step 3a is executed and the Lagrange multiplier λ_k is updated according to the rule (5.9). But we showed in Lemma 5.2.2 that this multiplier also satisfies the requirement (3.12) in algorithm ALDISCR. The update formulas for μ_k, w_k and η_k are identical.

Now suppose that $\|c(x_k)\| \geq \gamma_2\eta_k$. Then

$$\begin{aligned}\|c_{n_{k+1},m_{k+1}}(x_k)\| &\geq \|c(x_k)\| - \|c(x_k) - c_{n_{k+1},m_{k+1}}(x_k)\| \\ &\overset{(I)}{>} \gamma_2\eta_k - \alpha\eta_k \ = \ \eta_k\end{aligned}$$

Therefore, by algorithm ALDISCR, step 3b is executed which is identical to step 3b of algorithm ALINF. □

The previous lemma also reveals the following fact: The constant α, initialized in step 0 of algorithm ALDISCR, specifies the parameters γ_1 and γ_2 in the decision rule of algorithm ALINF.

All these Lemmas can be summarized in the following Theorem:

Theorem 5.2.4. *Let AS 1, AS 7 and AS 8 hold. Let $(x_k)_{k\in\mathbb{N}}$, $(\lambda_k)_{k\in\mathbb{N}}$ and $(\mu_k)_{k\in\mathbb{N}}$ be the sequences generated by algorithm ALDISCR. Then these sequences satisfy all requirements of algorithm ALINF. Or in other words: The sequences could also have been generated by algorithm ALINF.*

As a direct conclusion of this theorem, all the theory we derived in Chapter 3 for algorithm ALINF carries over to algorithm ALDISCR if AS7 and AS8 are added as a prerequisite to the lemmas and theorems in that Chapter. Nevertheless, we will list the most important results:

Corollary 5.2.5 (Global Convergence Theorem). *Let AS1 - AS4, AS7 and AS8 be valid. Further let x_* be any limit point of the sequence $(x_k)_{k\in\mathbb{N}}$ generated by algorithm ALDISCR and let $(x_k)_{k\in\mathcal{K}}$ be a subsequence whose limit is x_*. Define $\lambda(x_*)$ as on page 42. Then x_* is a Kuhn Tucker point (first order stationary point), and $\lambda(x_*)$ is the corresponding Lagrange multiplier, i.e.:*

$$\nabla_x L(x_*, \lambda(x_*)) = \nabla f(x_*) + c'(x_*)^*\lambda(x_*) = 0 \ , \qquad c(x_*) = 0$$

Further the sequences $(\bar{\lambda}(x_k, \lambda_k, \mu_k))_{k\in\mathcal{K}}$ and $(\lambda(x_k))_{k\in\mathcal{K}}$ converge to $\lambda(x_*)$ and the gradients $\nabla_x\Phi(x_k, \lambda_k, \mu_k)$ converge to $\nabla_x L(x_*, \lambda(x_*)) = 0$ for $k \in \mathcal{K}$.

Corollary 5.2.6. *Assume that AS 1, AS 4, AS7 and AS8 hold. Suppose that the sequence of iterates $(x_k)_{k\in I\!N}$ generated by algorithm ALDISCR converges to a single limit point x_* at which AS 3 holds. Let λ_* be the corresponding Lagrange multiplier and suppose that AS 5 and AS 6 hold at (x_*, λ_*). Then the sequence of penalty parameters $(\mu_k)_{k\in I\!N}$ is bounded away from zero, i.e. $\exists\ \mu_{min} \in (0, 1)$ such that $\mu_k \geq \mu_{min}\ \forall k \in I\!N$.*

As the penalty parameters $(\mu_k)_{k\in I\!N}$ are bounded away from zero, one can see in Theorem 5.1.2, that the discretization levels n_k of X and m_k of Y are more or less determined by the quantity w_k. Further the discretized Hessian matrix of $\Phi_{n_k,m_k}(\cdot, \lambda_k, \mu_k)$ is prevented from becoming more and more ill conditioned.

Corollary 5.2.7 (Local Convergence Theorem). *Assume that AS1, AS4, AS7 and AS8 are valid and that the sequence of iterates $(x_k)_{k\in I\!N}$ generated by algorithm ALDISCR converges to a single limit point x_* at which AS 3 holds. Let λ_* be the corresponding Lagrange multiplier and suppose that AS 5 and AS 6 hold at (x_*, λ_*). Then the Lagrange multiplier estimates $(\lambda_k)_{k\in I\!N}$ converge to λ_*. Further the sequences $(x_k)_{k\in I\!N}$, $(\bar{\lambda}(x_k, \lambda_k, \mu_k))_{k\in I\!N}$ and $(\lambda_k)_{k\in I\!N}$ are at least R-linearly convergent with R-factor at most $\mu_{min}^{\beta_\eta}$ where μ_{min} is the smallest value of the penalty parameter generated by algorithm ALDISCR.*

5.3 Relaxations

As proven in the last section, algorithm ALDISCR and algorithm ALINF coincide under appropriate assumptions in every iteration. The question is if this is really desired: Does it make sense to "perfectly" approximate the infinite dimensional method in an early iteration? One of the nice properties of algorithm ALDISCR is, that it allows a certain flexibility to handle this question:

Theoretically, one can omit the inner algorithm termination criterion (5.6) for all iterations $k \leq k_0$ where $k_0 \in I\!N$ is an arbitrary integer. This is meant in the way, that ALDISCR proceeds for these iterations with $x_k = x_{k_1}$ and discretization level $n_{k+1} := n_{k_2}$ to step 3, neglecting criterion (5.6).

Then, if a desired accuracy in the finite dimensional subproblem is reached, one can start the inner algorithm to reduce the approximation error of the finite dimensional method. In fact, if the inner algorithm is executed for all iterations $k > k_0$, the iterates $(x_k)_{k>k_0}$ coincide again with algorithm ALINF, as proven in the last section. Thus all convergence results carry over to those iterates.

Hence it is possible to run a discretized version of ALINF parallel to the infinite dimensional method without meeting its requirements. However, whenever it is desired, we can obtain a "real" iterate of ALINF by starting the inner algorithm to reduce the discretization error.

As this statement is true for all $k_0 \in I\!\!N$, one could expand the procedure above such that we terminate algorithm ALDISCR according to the criteria in step 2, never verifying (5.6). In order to run this extreme version of the algorithm one would not even have to request the convergence of the dual approximations $(Y_m^*, q^{Y_m})_{m \in I\!\!N}$, as proven in Theorem 4.3.4.

CHAPTER 6

Conclusions

In this thesis we derived very strong convergence properties for the discretized version of an augmented Lagrangian algorithm in infinite dimensional spaces. The algorithm gives an answer to the question how one should adapt the discretization level and the penalty parameter from one iteration to the next one. While these results are theoretically satisfying, the numerical performance of algorithm ALDISCR needs to be explored. My future work will be focused on this task.

However, there exist various other possibilities to extend and improve the theoretical results presented in this research:

First of all, one needs to drop the assumption of the nestedness of the approximations $(X_n)_{n\in I\!N}$ and $(Y_m)_{m\in I\!N}$ of the Hilbert spaces X and Y which was assumed to hold for simplicity. To the knowledge of the author, the incorporation of general approximations - along with more general restriction and prolongation operators - should not cause any problems in the derivation of the theory.

A far more interesting question is whether one has to assume the convergence of the dual approximations $(X_n^*, q^{X_n})_{n\in I\!N}$ and $(Y_m^*, q^{Y_m})_{m\in I\!N}$ in order to prove convergence of the discretized method. We showed in Chapter 4, that one can partially avoid this assumption for the dual prolongation operators $(q^{Y_m})_{m\in I\!N}$ by using the "right" refinement strategy. However, my intention is that the basic algorithm ALINF or more precisely inequality (3.8) requires this assumption.

Further the theory presented in this thesis does not include inequality constraints. This restricts the applicability of the algorithm with respect to optimal control problems, as these constraints often arise naturally from the considered model. In case of $X = I\!R^n$ and $Y = I\!R^m$ Conn, Gould, Sartenaer and Toint prove in [1] convergence for an algorithm similar to ALINF in presence of affine inequality constraints. It needs to be analyzed whether these results can be extended to the infinite dimensional setting in an “inexact” way.

Finally, one could try to generalize the theory to Banach spaces X and Y in order to widen the class of possible applications.

There is enough work left to do.... Packen wir’s an!

APPENDIX A

Selected Theorems

Theorem A.1 (Hölder-Inequality). *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $p \in (1, \infty)$. Further let $L^p(\Omega, \mu)$ denote the set of all equivalence classes of $\mathcal{A}$-measurable and to the p-power μ-integrable functions $f : \Omega \to I\!R$. Let*

$$\|f\|_{L^p} := (\int_\Omega |f|^p d\mu)^{1/p}$$

denote the usual norm in L^p. Let p and q be conjugate exponents, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Further let $f \in L^p$ and $g \in L^q$. Then

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}$$

Theorem A.2 (Lebesgue's Theorem of Majorized Convergence). *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $p \in (1, \infty)$. Further let $L^p(\Omega, \mu)$ denote the set of all equivalence classes of $\mathcal{A}$-measurable and to the p-power μ-integrable functions $g : \Omega \to I\!R$. Let $(f_n)_{n=1}^\infty \subset L^p$ be an almost everywhere pointwise convergent sequence. Further assume that there exists an $h \in L^p$ such that $|f_n| \leq h$ almost everywhere. Define $f(x)$ as the pointwise limit of $(f_n(x))_n$ at all x where that limit exists and define $f(x)$ as 0 everywhere else. Then*

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu$$

Theorem A.3 (Taylor's Theorem). *Let $I \subset I\!R$ be an interval and let $f : I \to I\!R$ be $n+1$-times differentiable on I. Further let $x_0 \in I$. Then $\forall\ x \in I \quad \exists \theta = \theta(x, x_0, n) \in (0, 1)$*

such that

$$f(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!}(x-x_0)^\nu + \frac{1}{(n+1)!} f^{(n+1)}(x_0+\theta(x-x_0))(x-x_0)^{n+1}$$

Theorem A.4 (Hahn Banach Theorem). *Let X be a real vector-space. Let $p : X \to I\!R$ satisfy*

$$p(\alpha x + (1-\alpha)y) \leq \alpha p(x) + (1-\alpha)p(y) \qquad \forall x,\ y \in X,\ \alpha \in [0,1]$$

Suppose Y is a subspace of X and $\lambda : Y \to I\!R$ is a linear functional such that

$$\lambda(x) \leq p(x) \qquad \forall x \in Y$$

Then there exists a linear functional $\Lambda : X \to I\!R$ such that

$$\Lambda(x) = \lambda(x) \qquad \forall x \in Y, \qquad \Lambda(x) \leq p(x) \qquad \forall x \in X$$

Corollary A.5. *Let y be an element of a normed linear space X. Then there exists a nonzero $\Lambda \in X^*$ such that*

$$\Lambda(y) = \|\Lambda\|_{X^*}\|y\|_X$$

Theorem A.6 (Best Approximation Theorem). *Let H be a Hilbert space and $M \subset H$ be a closed subspace. Then $\forall x \in H$ there exists a unique $m_* \in M$ such that*

$$\|x - m_*\|_H \leq \|x - m\|_H \qquad \forall m \in M$$

Theorem A.7 (Projection Theorem). *Let H be a Hilbert space and $M \subset H$ be a closed subspace. Then $\forall x \in H$ there exists a unique $m \in M$ such that $m^\perp := x - m \in M^\perp$.*

Theorem A.8 (Riesz Representation Theorem). *Let H be a Hilbert space and let H^* denote its dual space. Then there exists for each $T \in H^*$ a unique $y_T \in H$ such that*

$$T(x) =< y_T, x > \quad \forall x \in H \qquad \text{and} \qquad \|T\|_{H^*} = \|y_T\|_H$$

Corollary A.9. *Let X, Y be Hilbert spaces and let $B : X \times Y$ be a bounded, sesquilinear form, i.e. $B(\cdot,\cdot)$ satisfies for $x,\ x_1,\ x_2 \in X$ and $y,\ y_1,\ y_2 \in Y$ and all scalars $\alpha,\ \beta$:*

$$B(x_1 + x_2, y) = B(x_1, y) + B(x_2, y)$$

$$B(x, y_1 + y_2) = B(x, y_1) + B(x, y_2)$$

$$B(\alpha x, y) = \alpha B(x, y)$$

$$B(x, \beta y) = \bar{\beta} B(x, y)$$

$$\exists\ c \ \ \text{such that} \ \ |B(x,y)| \leq c\|x\|_X\|y\|_Y$$

Then $B(\cdot,\cdot)$ has a representation

$$B(x,y) =< Sx, y > \qquad \forall x \in X,\ y \in Y$$

where $S : X \to Y$ is a bounded linear operator. S is uniquely determined by $B(\cdot,\cdot)$ and has norm

$$\|S\|_{L(X,Y)} = \|B\| := \sup_{x\in X\setminus\{0\},\ y\in Y\setminus\{0\}} \frac{|B(x,y)|}{\|x\|_X\|y\|_Y}$$

Theorem A.10 (Uniform Boundedness Principle). *Let X be a Banach space. Let $\mathcal{F}$ be a family of bounded linear mappings from X to some normed linear space Y. Suppose that $\forall x \in X$ the set $\{\|Tx\|_Y : T \in \mathcal{F}\}$ is bounded. Then $\{\|T\|_{L(X,Y)} : T \in \mathcal{F}\}$ is bounded.*

Theorem A.11 (Inverse Mapping Theorem). *Let $(X, \|\cdot\|)_X$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. Further let $T \in L(X,Y)$ be bijective. Then its inverse T^{-1} is continuous.*

Theorem A.12 (Closed Graph). *Let X and Y be Banach spaces. Further let $T : X \to Y$ be linear and let $X \oplus Y$ denote the direct sum of X and Y with norm $\|(x,y)\|_{X\oplus Y} := \|x\|_X + \|y\|_Y$. Then T is bounded if and only if the graph of T is closed in $X \oplus Y$.*

Bibliography

[1] N. Gould, A. Sartenaer, A. R. Conn and P. L. Toint. Convergence properties of an augmented lagrangian algorithm for optimization with a combination of general equality and linear constraints. *SIAM J. Optimization*, 6(3):674–703, 1996.

[2] R. A. Adams. *Sobolev Spaces*. Academic Press, 1975.

[3] P. L. Toint, Andrew R. Conn, Nicholas I. M. Gould. A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds. *SIAM Journal on Numerical Analysis*, 28(2):545–572, 1991.

[4] J.-P. Aubin. *Approximation of elliptic boundary-value problems*. Number 26 in Pure and Applied Mathematics. Wiley-Interscience, 1972.

[5] F. Brayton and e. T. Tinsley. A guide to FRB/US: A macroeconomic model of the United States. Technical report, Federal Reserve Board, Washington, D.C., 1996.

[6] L. Collatz. *Functional Analysis and Numerical Mathematics*. Academic Press, New York, 1966.

[7] R. Tetlow, David Reifschneider and J. Williams. Aggregate disturbances, monetary policy, and the macroeconomy: The FRB/US perspective. Technical report, Federal Reserve Board, Washington, D.C., 1999.

[8] C. Desoer and B. Whalen. A note on pseudoinverses. *J. SIAM*, 11:442–447, 1963.

[9] A. Drew and B. Hunt. The effects of potential output uncertainty on the performance of simple policy rules. Technical report, Reserve Bank of New Zealand, Economics Department, 2000.

[10] F. S. Finan and R. Tetlow. Optimal control of large, forward-looking models. Technical report, Board of Governors of the Federal Reserve System, Division of Research and Statistics, Washington, DC, 2000.

[11] W. W. Hager. Multiplier methods for nonlinear optimal control. *SIAM Journal on Numerical Analysis*, 27(4):1061–1080, 1990.

[12] M. R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 4:303–320, 1968.

[13] K. Ito and K. Kunisch. The augmented Lagrangian method for equality and inequality constraints in hilbert spaces. *Mathematical Programming*, 46(3):341–360, 1990.

[14] K. Ito and K. Kunisch. Augmented Lagrangian-SQP-methods in Hilbert spaces and application to control in the coefficients problems. *SIAM Journal on Optimization*, 6(1):96–125, 1996.

[15] J. Jahn. *Introduction to the Theory of Nonlinear Optimization.* Springer-Verlag, Berlin Heidelberg, 1994.

[16] E. Kreyszig. *Introductory Functional Analysis with Applications.* John Wiley and Sons, New York, 1978.

[17] K. J. Lansing. Optimal redistributive capital taxation in a neoclassical growth model. Technical report, Federal Reserve Bank of San Francisco, Research Department, 1998.

[18] D. G. Luenberger. *Optimization by Vector Space Methods.* Series in Decision and Control. John Wiley & Sons, Inc., New York, 1969.

[19] M. Z. Nashed, editor. *Generalized Inverses and Applications*, number 32, New York, 1976. Academic Press.

[20] A. Novales. The role of simulation methods in macroeconomics. Technical report, Universidad Complutense, Departamento de Economía Cuantitativa, Spain, 2000.

[21] V. T. Polyak and N. Tret'yakov. The method of penalty estimates for conditional extremum problems. *Zurnal Vycislitel'noi Matematiki i Matematiceskoi Fiziki*, 13:34–46, 1973.

[22] M. J. D. Powell. A method for nonlinear constraints in minimization problems. In *Optimization*, pages 283–298. R. Fletcher, ed., Academic Press, 1968.

[23] M. Reed and B. Simon. *Functional Analysis.* Number I in Methods of Modern Mathematical Physics. Academic Press, New York Boston, 1980.

[24] M. Runkel. Product durability, solid waste management, and market structure. Technical report, University of Siegen, Department of Economics IV, Germany, 1999.

[25] E. W. Sachs and A. Sartenaer. A class of augmented Lagrangian algorithms for infinite-dimensional optimization with equality constraints. Technical report, Universität Trier, Germany, 1999.

[26] S. Volkwein. *Mesh-Independence of an Augmented Lagrangian-SQP Method in Hilbert Spaces and Control Problems for the Burgers Equation.* PhD thesis, Fachbereich Mathematik der Technischen Universitaet Berlin, 1997.

[27] J. Weidmann. *Linear Operators in Hilbert Spaces.* Number 68. Springer Verlag, Heidelberg Berlin, 1980.

Vita

Jan H. Maruhn

Jan Hendrik Maruhn was born July 22, 1977 in Ludwigshafen am Rhein, Germany to Jürgen Maruhn and Gisela Maruhn. He went to High School in Trier, Germany and earned his "Abitur" in the Summer of 1996. Afterwards he joined the German Army for one year where he started to focus on the application of Mathematics in Business Administration. In October, 1997 he began his studies at the University of Trier. Starting February 1998, he worked for two years in a joint research project with Nestlé on the optimal control of industry autoclaves. Jan earned a "Vordiplom"-degree in Mathematics in the Fall of 1999, also graduating with a double major in Business Administration.

He immediately began Graduate study at the University of Trier, focusing on Numerical Analysis, Operations Research, Mathematics for Finance and Finance in general. In February 2000, he was admitted to the German National Scholarship Foundation.

Jan began Graduate study at Virginia Polytechnic Institute and State University in August, 2001, pursuing an interdisciplinary plan of study in the Mathematics Department. He continued on until the Spring of 2001 when he received a Master of Science degree in Mathematics.