



ISNM

International Series of Numerical Mathematics

Vol. 135

Managing Editors:

K.-H. Hoffmann, München

D. Mittelmann, Tempe

Associate Editors:

R. E. Bank, La Jolla

H. Kawarada, Chiba

R. J. LeVeque, Seattle

C. Verdi, Milano

Honorary Editor:

J. Todd, Pasadena

Nonlinear Multiobjective Optimization

A Generalized Homotopy Approach

Claus Hillermeier

Springer Basel AG

Author:

Claus Hillermeier
Siemens AG
ZT PP2
81730 München (Perlach)
Germany

until August 2001:
Chair of Applied Mathematics II
University of Erlangen-Nürnberg
Martensstr. 3
91058 Erlangen
Germany

2000 Mathematics Subject Classification 74P20, 58E17, 90C29, 65H20

A CIP catalogue record for this book is available from the Library of Congress, Washington D.C., USA

Deutsche Bibliothek Cataloging-in-Publication Data

Hillermeier, Claus:

Nonlinear multiobjective optimization : a generalized homotopy approach / Claus Hillermeier.

- Basel ; Boston ; Berlin : Birkhäuser, 2001

(International series of numerical mathematics ; Vol. 135)

ISBN 978-3-0348-9501-9 ISBN 978-3-0348-8280-4 (eBook)

DOI 10.1007/978-3-0348-8280-4

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, broadcasting, reproduction on microfilms or in other ways, and storage in data banks. For any kind of use whatsoever, permission from the copyright owner must be obtained.

© 2001 Springer Basel AG

Originally published by Birkhäuser Verlag, in 2001

Softcover reprint of the hardcover 1st edition 2001

Printed on acid-free paper produced of chlorine-free pulp. TCF ∞

ISBN 978-3-0348-9501-9

9 8 7 6 5 4 3 2 1

Dedicated to my parents

Preface

Real industrial systems are usually assessed by setting several objectives which are often competing with each other. Good compromise solutions are then looked for. The task of multiobjective optimization is to determine so-called efficient (or Pareto optimal) solutions which cannot be improved simultaneously with regard to all objectives.

The present book first gives a survey of the principles and classical methods of multiobjective optimization. Afterwards, the set of Pareto candidates is considered as a differentiable manifold, and a local chart is constructed which is fitted to the local geometry of this Pareto manifold. This opens up the possibility of generating new Pareto candidates by evaluating that local chart numerically. The generalized homotopy method thus developed has important advantages. It is capable of solving multiobjective optimization problems with an arbitrary number k of objectives, enables the generation of all types of Pareto optimal solutions and is able to produce a homogeneous discretization of the Pareto set. In the theoretical part of the book, the homotopy method is put on a sound mathematical basis by providing a necessary and sufficient condition for the set of Pareto candidates to form a $(k - 1)$ -dimensional differentiable manifold. The theoretical discussion is followed by a description of the numerical details of the proposed homotopy algorithm. Finally, by solving three multiobjective sample problems we demonstrate how this algorithm works in practice. Two of these problems originate in optimization applications within the configuration of industrial systems.

Acknowledgements

First of all I wish to express my gratitude to Prof. Dr. Dr. h. c. Karl-Heinz Hoffmann for encouraging and supporting the piece of research presented here. I would like to thank Prof. Dr. Klaus Ritter and Prof. DDr. Stefan Schäffler for several fruitful discussions which were a pleasure and a great help. Special thanks also go to my colleagues at Siemens Corporate Technology and to our coach Prof. Dr. Albert Gilg for creating an enjoyable and stimulating working atmosphere. With gratitude I would like to mention the successful and pleasant collaboration with my colleagues at Siemens KWU. I wish to express my appreciation to Prof. Dr. Johannes Jahn for revising parts of the manuscript and providing valuable comments. Last, but not least, I am indebted to Rudolf Knop for his help with the English translation and to Dr. Michael Greiner for generously providing his TEX-expertise.

The work presented here has been supported by the German “Bundesministerium für Bildung und Forschung” in the framework of the project LEONET. This support is gratefully acknowledged.

Contents

1	Introduction	3
2	Vector Optimization in Industrial Applications	9
2.1	The Design of a Combined-Cycle Power Plant	10
2.2	The Optimal Operating Point of a Recovery-Boiler	12
3	Principles and Methods of Vector Optimization	15
3.1	The Concept of Pareto Optimality	15
3.2	Survey of Methods	19
3.3	A New Stochastic Method for Unconstrained Vector Optimization	30
3.3.1	A Curve of Dominated Points	31
3.3.2	Notions from Probability Theory	37
3.3.3	A Special Stochastic Differential Equation	39
3.3.4	A Stochastic Algorithm for Vector Optimization	42
4	The Connection with Scalar-Valued Optimization	45
4.1	The Karush-Kuhn-Tucker(KKT) Condition for Pareto Optimality	45
4.2	Differential-Topological Notations	47
4.3	The Geometrical Meaning of the Weight Vector	53
4.4	Classification of Efficient Points	59
5	The Manifold of Stationary Points	65
5.1	Karush-Kuhn-Tucker Points as a Differentiable Manifold M	66
5.2	Criteria for the Rank Condition	68
5.2.1	A Necessary and Sufficient Criterion	68
5.2.2	Interpretation in View of Optimization	71
5.2.3	Variability of the Weight Vector	75
5.3	A Special Class of Local Charts	79
6	Homotopy Strategies	87
6.1	Method I: Local Exploration of M	88
6.1.1	Method Principle	88

6.1.2	Comparison with the Classical Homotopy Method	89
6.1.3	Homogeneous Discretization of the Efficient Set	93
6.1.4	Numerical Algorithm	95
6.2	Method II: Purposeful Change of the Weights	99
6.2.1	Significance of the Weight Vector for the User	99
6.2.2	Principle of the Procedure	101
6.2.3	Numerical Algorithm	104
7	Numerical Results	109
7.1	Example 1 (academic)	109
7.2	Example 2: Design of a Combined-Cycle Power Plant	115
7.3	Example 3: The Optimal Operating Point of a Recovery-Boiler . .	123
	Bibliography	129
	Index	133

Chapter 1

Introduction

The behavior of industrial plants and systems is in general determined by a number of physical and technical parameters. The values of these system parameters — in the following called variables¹ — are fixed more and more with the help of optimization calculations. Based on a mathematical model of the system an optimization algorithm searches for those values of the variables, for which certain objective quantities take on a minimum (or a maximum).

In order to be able to apply the methods of classical optimization, the quality of the system behavior has to be measured with a single objective, i.e. with a scalar-valued quantity. However, real industrial systems are mostly assessed by more than one criterion. All corresponding objectives should possibly be optimized simultaneously. Often the criteria are contradictory in such a way that optimization of one criterion entails a poor value of another criterion. Therefore, one requires good compromise solutions which are suboptimal, but acceptable as to the individual objectives. The problem of finding good compromise solutions is called a multiobjective optimization problem (MOP), a vector optimization problem (VOP) or a multicriteria optimization problem. We shall use the terms vector optimization and multiobjective optimization synonymously in this book. In almost all fields of the science of engineering original vector optimization problems turn up. They range from designs of electrical switching circuits, machine parts, airplanes and weight-bearing structures (bridges, pylons etc.) to planning and controlling of water-supply systems. One can find detailed bibliographical references to application problems of vector optimization in [STADLER, 1988] and [DAS, 1997]. Chapter 2 describes some characteristic applications of multiobjective optimization arising within the configuration of industrial systems at Siemens company.

At an early stage of decision-making one often has no precise information about the relative importance of the individual objectives. In these cases a whole family of variable values have to be considered as being of equal importance,

¹ Following the linguistic usage common in the field of optimization, we will from now on call the technical system parameters 'variables'. The word 'parameter' shall be used in the present book as a mathematical denomination (and will be specified more precisely, if need be, as chart parameter or homotopy parameter).

namely such solutions — denoted as **efficient** —, in which no objective can be further improved without impairing at least one other objective. At this early stage of decision-making the purpose of mathematical vector optimization is therefore to give a survey of efficient solution alternatives to the user (also called decision-maker) or, in the ideal case, to determine the entire set of efficient solutions. To solve this mathematical problem, a number of methods has been evolved (see e.g. [JAHN, 1986], [GÖPFERT & NEHSE, 1990] and [DAS, 1997]). Most of them are based on the idea of transforming the vector optimization problem into a problem of scalar-valued optimization or of breaking it down into partial problems which can be solved with methods of scalar-valued optimization. A survey of the most important classical methods of multiobjective optimization can be found in Section 3.2 of this book. Apart from that, Section 3.3 presents a recent and completely different approach to vector optimization based on stochastic concepts (see [SCHÄFFLER ET AL., 1999]).

One of the most common approaches to multiobjective optimization is the so-called **weighting method** (see e.g. [GÖPFERT & NEHSE, 1990] and [DAS & DENNIS, 1996A]). It interprets a convex linear combination of the individual objectives as a (now scalar-valued) objective function and searches for a minimizer of this objective function. Global minimizers of such a convex combination are necessarily efficient solutions of the initial vector optimization problem. By variation of the coefficients in the convex combination, i.e. by variation of the relative weights of the individual objectives, various efficient solutions can be generated. The weights are thus parameters of a family of scalar-valued optimization problems. The weighting method therefore treats the multiobjective optimization problem as one of classical parametric optimization.

In general, a parametric optimization problem has a family of minimizers, of which each one is a stationary point of the objective function — or, if the search space is restricted by equality constraints, of the Lagrangian function — and thus is necessarily also a *zero of a parametrized function* (namely of the gradient of the parametrized objective or Lagrangian function). Consequently, the parameter of the optimization problem can be interpreted as a *homotopy parameter*. In contrast to the (common) case, in which such a homotopy parameter is artificially introduced to build a bridge – by variation of this parameter – between a system of equations, the solution of which is known, and a system of equations with unknown solution, in parametric optimization problems the homotopy parameter is given in a natural way. If a solution (i.e. a minimizer) is known for a special parameter value, homotopy methods can be applied to find solutions for different parameter values (see e.g. [SCHWETLICK, 1979], [RHEINBOLDT, 1986], and [ALLGOWER & GEORG, 1990]). Indeed, homotopy methods — also known as continuation methods — can be utilized successfully for parametric optimization problems (see [RAO & PAPALAMBROS, 1989]).

Therefore it seems reasonable to interpret the vector optimization problem, following the approach of the weighting method, as a parametric optimization

problem and to employ the homotopy technique for its solution. In fact, such an approach was proposed by Rakowska et al. [RAKOWSKA ET AL., 1991].

Contrary to classical parametric optimization problems, the vector optimization problem (VOP) has two peculiarities, which have to be taken into account, if one intends to establish the homotopy method as a theoretically founded and generally applicable solution method for multiobjective optimization problems.

- (a) If k denotes the number of objectives of the VOP to be minimized, the weight vector has $(k - 1)$ components which can be chosen freely — the k -th component results from normalizing the sum of the components to 1. The VOP therefore has a natural $(k - 1)$ -dimensional homotopy parameter. The classical homotopy techniques presuppose a one-dimensional homotopy parameter (which is, as we mentioned earlier, in most cases introduced artificially).
- (b) The interpretation of the VOP as a parametric optimization problem has its theoretical grounds in a theorem of Kuhn and Tucker [KUHN & TUCKER, 1951]. It says that to every efficient solution of the VOP necessarily there exists a convex combination of the objectives, i.e. a scalar-valued function, such that the efficient point (in the variable space) is a Karush-Kuhn-Tucker point of this scalar-valued objective function. (Remember that in the case of unconstrained optimization a Karush-Kuhn-Tucker point is just a stationary point.) However there is no necessary optimality condition of second order in the VOP – in contrast to the scalar-valued optimization. The link between vector and scalar-valued optimization does therefore not extend to second order optimality conditions. Consequently an efficient point does *not necessarily* have to be a *minimum* of the corresponding convex combination of the individual objectives.

The homotopy approach which has been proposed by Rakowska et al. [RAKOWSKA ET AL., 1991] does not take these peculiarities of the vector optimization problem into consideration.

On the one hand it is limited a priori to the special case of bicriterial optimization problems (i.e. $k = 2$). In this special case the (weight) homotopy parameter is one-dimensional, a property, on which Rakowska's homotopy method is based²: A homotopy curve is determined numerically by means of a predictor-corrector technique. Both curve points calculated last are interpolation nodes of a cubic Hermite-interpolant, which itself is a predictor of the curve point to be calculated. On the other hand Rakowska's approach is limited to the determination of such efficient points, which are minima of a convex combination of the objectives.

² From this conceptional limitation of Rakowska's homotopy approach the following generalization is erroneously inferred in current articles on vector optimization (see [DAS & DENNIS, 1996b] and [DAS, 1997]): 'A continuation/homotopy based strategy for tracing out the Pareto curve ... cannot be applied to problems with more than two objectives in general'. (Pareto-points correspond to efficient solutions.)

On the way towards a homotopy method which enables us to solve genuine multicriterial vector optimization problems (i.e. cases with $k > 2$ as well) on good theoretical grounds, we have to ask the following questions:

- (A) What part do saddle points of convex combinations of the objectives play within the totality of efficient solutions?
- (B) Under what circumstances is the zero manifold, which consists of stationary points of the objective function (or Lagrangian function) parametrized by the weight vector, suitable for some kind of homotopy method?
- (C) How can a homotopy method be constructed which enables us to examine freely in all directions (i.e. in all dimensions) the generally multidimensional zero manifold of (potentially) efficient solutions, starting from a point of this manifold, instead of restricting us — like in common homotopy methods — to one-dimensional submanifolds (curves) of this zero manifold?

The purpose of the present book is finding an answer to these questions. The key to the answer lies in a thorough examination of the set of efficient points (or of the mentioned zero manifold which contains all efficient points) from the viewpoint of **differential topology**. Depending on whether one looks at this set in the variable space — more precisely: in the product space of variables, Lagrange multipliers and weight parameters — or at the image of this set (under the mapping of the objective function) in the k -dimensional objective space, one gains different insights and results.

The differential-topological look at the solution set in the **objective space** makes it possible to extend the comprehension of the interrelation, discovered by Kuhn and Tucker, between scalar-valued optimization and vector optimization: First, one can show what geometric significance the weight vector has with respect to the manifold of efficient points in the objective space (see Section 4.3). From this geometric significance follows in turn that the weight vector contains important information for the user, by means of which he is able to distinguish and interpret the calculated efficient solutions (see Paragraph 6.2.1). Furthermore, a connection can be established between the local curvature of the solution manifold in the objective space and the question, what sort of stationary points (i.e. minima or saddle points of a convex combination of the objectives) the corresponding efficient solutions represent (see Section 4.4). Automatically the important part which saddle points play within the totality of efficient solutions will then have been clarified.

If one looks at the solution set in the (extended) **variable space** from the standpoint of differential topology, one first has to ask the question, whether or under which premises the zero manifold — which consists of stationary points of convex combinations of the objectives and therefore of candidates for efficient solutions — is a differentiable manifold of the dimension $(k - 1)$ (= the number

of components of the weight vector that can be chosen freely). In Section 5.2 we will show that (sufficiently small) neighborhoods of minima as well as of saddle points (with the additional property of having a regular Hessian matrix of the Lagrangian function) are automatically $(k - 1)$ -dimensional differentiable manifolds. Furthermore, we will indicate a weakly restrictive condition which is sufficient for neighborhoods of border points between minimum and saddle point regions to be $(k - 1)$ -dimensional differentiable manifolds. (We refer to border points between a region of the zero manifold, in which the stationary points are minima of a convex combination of the objectives, and a region, in which the stationary points are saddle points of a convex combination of the objectives.) By virtue of this important assertion in principle it is possible to reach minima regions from saddle point regions and vice versa by homotopy. Hence, by means of the differential-topological way of looking at things it is possible to gain theoretical assertions which safeguard the use of homotopy methods for vector optimization.

Moreover, the differential-topological look at the solution set in the extended variable space provides constructive guidelines for a generalized homotopy method³, which takes into account the dimensionality of the natural homotopy parameters in the case of multiobjective optimization (see Section 5.3). Every homotopy step is interpreted as a numerical evaluation of a chart (= parametrization for a local description of the solution manifold) which is fitted to the local geometry of the solution manifold. The homotopy method based on this central idea is formulated in Chapter 6 as a numerical algorithm. Besides its main property of completely exhausting the natural multidimensionality of the solution set this homotopy method will provide the user with important advantages:

- (1) The method is capable of generating a homogeneous distribution of efficient solution points in the objective space or, if need be, of controlling this distribution in a simple way (see Paragraph 6.1.3). The decision-maker thereby obtains sufficient information in all areas of the solution space about the mutual competition of the different objectives.
- (2) Alternatively the user can *either* become acquainted with the efficient points situated in the neighborhood of a known efficient solution in all directions and thus gain a local survey of efficient alternative solutions (method variant I, described in Section 6.1) *or* vary the relative weight of the individual objectives in a purposeful way (variant II, described in Section 6.2).
- (3) The homotopy method determines the weight vector which is associated to each calculated efficient solution. This vector contains the relative valences

³ Strictly speaking, the developed method is not a homotopy method in the narrow sense, since it does not utilize the natural homotopy parameters (i.e. the components of the weight vector), but constructs in each step own homotopy parameters which are fitted to the local geometry of the solution manifold. (One can find a comparison with classical homotopy methods in Section 6.1.2.) For the sake of brevity we will not speak, however, of a 'generalized homotopy method', but simply of a homotopy method.

of the individual objectives in this solution point and provides the decision-maker with valuable information for interpreting the solution point (see Paragraph 6.2.1).

Chapter 7 describes the use of the method by solving two industrial problems of vector optimization. These problems come from the field of power plant construction and the field of operating point optimization of industrial plants and are presented in the following Chapter 2.

Let us still emphasize two points:

- For the homotopy method to be applicable to a given vector optimization problem, both the (vector-valued) objective function and the functions which define the restrictions must be twice continuously differentiable. Because of its universality this assumption will no longer be stated explicitly in many places of the present book. It is a sufficient prerequisite for the results of Section 4.3 (geometric significance of the weight vector) that the objective function and the restrictions are once continuously differentiable.
- The homotopy method developed here is applicable outside vector optimization as well, when solutions of systems of equations are searched for which depend on several parameters in a natural way.

Vector Optimization in Industrial Applications

Application problems of vector optimization that arise in the science of engineering are documented in literature in great numbers (see e.g. [STADLER, 1988] and [DAS, 1997]). Instead of listing these quotations here again, we will present the multiobjective problems which originate in optimization applications within the configuration of industrial systems. Subsequently we will discuss in detail two multiobjective problems which arise in the concrete practice of the plant manufacturer SIEMENS.

Manufacturers of (industrial, power, telecommunication etc.) plants and, more generally, technical systems are mostly confronted with the following types of optimization problems:

The **design** phase of plants or systems involves the optimization of physical and technical design variables of a plant or its components.

In the phase of putting a plant into service its **operating point** has to be determined, i.e. those values of the control variables have to be found which result from the viewpoint of the plant operator in an optimal system behavior.

Design and operating point optimization are each based on a **model** of the plant behavior. Such a model consists of the physical and technical correlations between the system parameters and contains mostly several quantities which have to be determined by comparing model predictions with the results of measurements. Since the aim is to minimize the deviation of the model predictions from the measurements, model optimization is another industrial field of applied optimization.

All three application fields of optimization are in many cases characterized by several contradictory objectives, for which good compromise solutions have to be found.

An illustrative example of a multicriterial plant design is the optimization of variables characterizing the geometry of a vacuum pump. Such a pump has to have simultaneously maximum suction capacity, minimal power demand and minimal demand for operating liquid.

Typical conflicting objectives within industrial system design are the maximiza-

tion of efficiency (or plant productivity), the minimization of failure and the minimization of the investment funds to be raised for the acquisition of the plant. Another class of multicriterial design problems originates from the fact that in long-term plant investments the later operation conditions of the plant (e.g. in a power plant: full load or sub-load running) are not predictable with certainty at the time of the plant design. Since, however, the values of the objectives (e.g. the efficiency of the power plant) depend on the operation conditions, the following way of acting is adequate: From the set of possible operation scenarios a few prototypic representatives are chosen (e.g. full load plus a sub-load scenario). The value of the original objective (e.g. power plant efficiency), which is obtained within a prototypic operation scenario, is now an objective component of the new, henceforth vector-valued, optimization problem. The dimension of the objective space is given by the number of prototypic operation scenarios¹. One has an essential competitive advantage when making an offer, if one is able to show efficient design alternatives for this multiobjective problem. Out of the quantity of efficient design alternatives the management of the potential purchaser and future user of the plant can choose the one which is best integrated with the overall strategy of his enterprise.

When optimizing the operating point, the vector of objectives in general consists of the quantities of the single desired plant products (to be maximized each) and the quantities of the unwanted by-products or pollutants (to be minimized each).

Model optimization is often also a multicriterial problem. In this case, the vector of objectives is spanned by the discrepancies between the single measured quantities or measurement points within the real plant and the corresponding model predictions.

To fill these general assertions with life, in the sequel two examples from the concrete SIEMENS-practice are discussed in detail. Both multiobjective optimization problems were solved numerically by means of the homotopy method developed in this book. The results can be found in Chapter 7

2.1 The Design of a Combined-Cycle Power Plant

The type of power plant in which the highest efficiencies in electricity production can be achieved are the so-called combined-cycle power plants (in short: CC-power plants). In these plants two thermodynamic processes are coupled for the purpose of efficiency improvement (see [STRAUSS, 1994]). The liquid or gaseous fuel (generally natural gas) is injected into a combustion chamber filled with compressed air. In a gas turbine the combustion gas expands to a low pressure,

¹ If the original objective is already vector-valued, the dimension of the objective space is the product of the number of the operation scenarios and the number of the original objectives.

thus powering a generator and producing electricity. The residual heat contained in the (up to 600 degrees centigrade) hot exhaust gas of the gas turbine is used in a so-called heat recovery boiler to drive a second thermodynamic process, namely a water/steam cycle. In the heat recovery boiler water is transformed into overheated steam (so-called live steam) which for its part powers a steam turbine and thus contributes to the electricity production. Since the hot exhaust gas cools off when flowing through the heat recovery boiler, residual heat on different temperature levels can be disposed of. In order to utilize the residual heat of each level in an optimal way, live steam is generated in different thermodynamic states adapted to the relative temperature level of the exhaust gas. State of the art are so-called triple-pressure cycles with a high pressure(hp)-stage, a medium pressure(mp)-stage and a low pressure(lp)-stage. The hot exhaust gas flowing out of the gas turbine generates first high pressure steam, cools down, then generates medium pressure steam, and the residual heat is used for generating low pressure steam. Since the steam turbine is also divided into different areas, the steam of each pressure stage can be introduced at a suitable point into the steam turbine and can thus be used for electricity production.

To what degree heat is transferred from exhaust gas to water (or steam) within each pressure stage, is characterized by the so-called pinch-point, a quantity which is specific for each pressure stage. It represents the smallest temperature difference between the exhaust gas and the steam, i.e. between the heat-emitting and the heat-absorbing medium. Since heat transfers are caused by temperature differences, small pinch-points can be obtained only with large – and thus expensive – heat exchanger surfaces. On the other hand small temperature differences between heat-emitting and heat-absorbing media imply a thermodynamically effective exploitation of the residual heat and consequently an increase of efficiency.

As the purchaser (and future operator) of a power plant wants to keep both his fuel and his investment costs as low as possible, the design of the three pinch-points of a triple-pressure combined-cycle power plant is characterized by two contradictory objectives: the maximization of the thermodynamical efficiency (or, equivalent to it, the minimization of the negative efficiency) and the minimization of the investment costs connected with the pinch-point design, i.e. the costs of the heat recovery boiler and the cooling system.

Thus, the optimum pinch-point design constitutes a problem of bicriterial optimization². Its solution can be found by means of the homotopy method developed and will be presented in Section 7.2.

² From the viewpoint of pure business management a power plant design can be assessed by a single objective quantity, namely the electricity production costs caused by this design (i.e. the costs which arise for the power plant operator when generating one kWh of electricity). Both efficiency and investment costs enter into this objective quantity:

$$\text{Electricity production costs} = \frac{\text{investment costs} \cdot \text{annuity}}{\text{electrical power} \cdot \text{working hours}} + \frac{\text{fuel price}}{\text{efficiency}} \quad (2.1)$$

2.2 The Optimal Operating Point of a Recovery-Boiler

In paper production wooden shavings are boiled in a chemical solution for breaking down cellulose. The chemicals used and most of the heat energy required for the pulping process can be recovered from the concentrated spent liquor (so-called black liquor) of the process by means of a recovery-boiler. The degree to which chemicals and heat are recovered is of decisive significance for the economy of the entire plant (see [BÖWE & FURUMOTO, 1992]).

Figure 2.1 represents the schematic structure of a recovery-boiler. The waste liquor, already concentrated, is injected into the furnace of the boiler by means of liquor guns. Waste liquor drops are formed during spraying and are dried while falling through the rising hot stack gas. The dried alkaline particles fall on the char bed. Here reactions take place which are important for the recovery: predominantly chemical reduction processes because of lack of oxygen; the remaining organic parts of the waste liquor are incinerated. As a result of the reactions one obtains alkaline ashes in the char bed, which can be removed from the boiler and from which the chemicals used for boiling wood can be recovered easily. Volatile components and reaction products are swept away by the stack gas and reach an oxidation zone. There is a surplus of oxygen and the combustion process is concluded by oxidizing reactions. The heat of the combustion gases is used to generate overheated steam and to produce electricity.

The air required for the combustion is introduced into the burning chamber in three different stages (primary, secondary and tertiary air). These three streams of air are the control variables of the system. By supplying the air and dividing it between three feeds the plant operator can control the reaction conditions in the recovery-boiler (in particular, the proportion of oxidation and reduction processes).

Constant economical operation of the recovery-boiler is the purpose of the plant control. A boiler operating economically is characterized by well-balanced reaction conditions in the char bed which are appropriate for the recovery of the chemicals, by a large steam production and by a low portion of pollutants in the waste gas outlet. As a given constraint, a certain quantity of black liquor has to be processed and incinerated by the recovery-boiler.

The economic factors 'annuity' and 'fuel price', the values of which are required for the entire operating duration in order to be inserted in the above formula, as well as the marketable electricity quantity per annum (*electrical power · working hours*) can only be roughly forecasted at the moment of the power plant design. Since unpredicted changes of these economic factors alter the relative importance of the investment costs and the efficiency within the total electricity production costs, it is of highest interest for the power plant manufacturer to know the set of efficient (alternative) solutions, which describes the 'trade-off' between efficiency and investment costs.

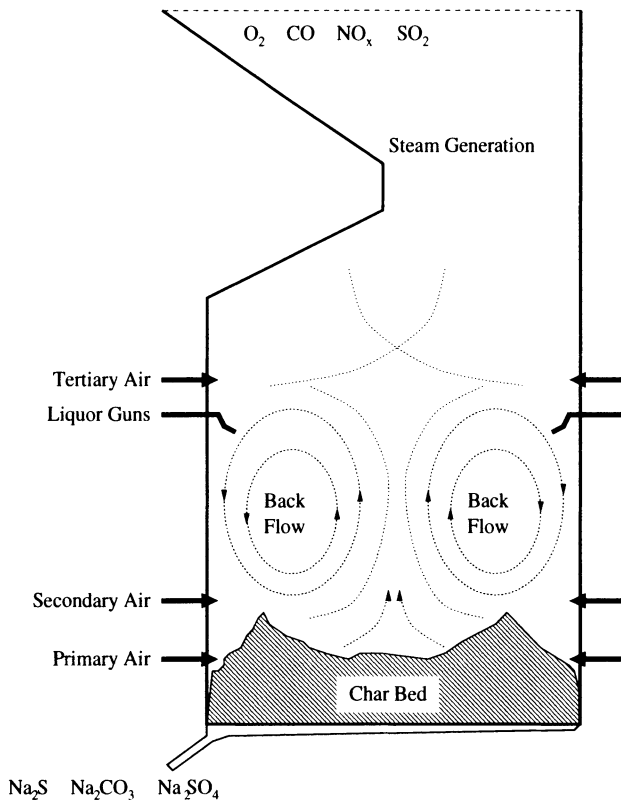


Figure 2.1: Schematic representation of a recovery-boiler

Mainly four measured quantities indicate to the plant operator whether the above requirements of the boiler operating point are met: the O_2 -concentration in the waste gas, the SO_2 -concentration in the waste gas, the mass flow of the generated steam and the temperature of the char bed. Since because of the complexity of the chemical and hydrodynamical processes no detailed physical model of the plant behavior is available, the control of the recovery-boiler is based essentially on the experience of the plant operator. According to the quantity of black liquor to be incinerated, he sets four desired values for the four above-mentioned measured quantities, which should guarantee an economical operation of the boiler. The single desired (ideal) values each take into account one of the different operation objectives, which are partially competing with each other. Therefore, in general there is no realizable operating point which complies with the desired combination of the four values given by the plant operator. More likely, an operating point has to be found for which the four measured quantities are close to the values given by the operator.

Balancing the three air supplies of a recovery-boiler is therefore a multicriteria optimization problem. The four individual objectives are constructed by the

quadratic deviations of the four measured quantities from the relative value desired by the plant operator. If a set of efficient operating points (with regard to the vector-valued objective function constructed out of these four individual objectives) has been calculated as a solution of this vector optimization problem, the plant operator can choose the most appropriate adjustment from his experience and based on his knowledge of the current urgency of the individual objectives.

Section 7.3 will present the solution of this multiobjective optimization problem.

Principles and Methods of Vector Optimization

3.1 The Concept of Pareto Optimality

Let an operation point or a plant design be characterized by n real-valued variables $x_1, \dots, x_n$. The variables can be combined to a vector $\mathbf{x} := (x_1, \dots, x_n)^T \in \mathbb{R}^n$ and are supposed to vary freely within a feasible set $R \subseteq \mathbb{R}^n$.

Quantitative criteria for the assessment of a variable vector $\mathbf{x}$ are k objectives $f_1, \dots, f_k$, which are functions of $\mathbf{x}$ and which can be combined to a vector-valued objective function $\mathbf{f}$:

$$\mathbf{f} : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^k \\ \mathbf{x} & \mapsto \mathbf{f}(\mathbf{x}) := (f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))^T. \end{cases} \quad (3.1)$$

Let us formulate the application problem in such a way as to minimize all objectives f_i at the same time¹. In general, however, individual objectives are in contradiction to each other, i.e. an improvement with regard to one objective causes the deterioration of another. The requirement of minimizing all objectives f_i simultaneously has to be interpreted in a suitable way in order to obtain a meaningful type of problem.

Since minimization presupposes in principle that various objective function values be *compared* with each other, an ordering concept in the $\mathbb{R}^k$, appropriate to the problem, is required. The definition of a total order which allows us to compare any two arbitrary elements of the considered space with each other meets with difficulties in the $\mathbb{R}^k$. If there does not exist a given hierarchy of the k objectives, it is, for instance, not possible to indicate an order relation between the two vectors (of values of a two-dimensional objective function) $\mathbf{y}^1 = (4, 2)$ and $\mathbf{y}^2 = (2, 4)$, without implying a (possibly local) weighting of the objectives. Instead of a total

¹ If the original requirement is maximizing an objective f_i , then it will be transformed into the equivalent requirement of minimizing $-f_i$.

order we therefore define only a weaker order relation in the $\mathbb{R}^k$ which is denoted by $\leq$ and which is illustrated in Figure 3.1 for the special case of the $\mathbb{R}^2$.

Definition 3.1: (Order relation $\leq$ in the $\mathbb{R}^k$)

Let $\leq$ denote an order relation in the $\mathbb{R}^k$, i.e. a special subset of the set $\mathbb{R}^k \times \mathbb{R}^k$ of all ordered pairs of elements of the $\mathbb{R}^k$. Instead of $(\mathbf{y}^1, \mathbf{y}^2) \in \leq$ one customarily uses the infix notation $\mathbf{y}^1 \leq \mathbf{y}^2$. Let the order relation be defined as follows:

$\mathbf{y}^1 \leq \mathbf{y}^2 \iff \mathbf{y}^2 - \mathbf{y}^1 \in \mathbb{R}_+^k$, where $\mathbb{R}_+^k := \{\mathbf{y} \in \mathbb{R}^k \mid y_i \geq 0 \ \forall i \in \{1, \dots, k\}\}$ denotes the non-negative orthant of the $\mathbb{R}^k$. $\square$

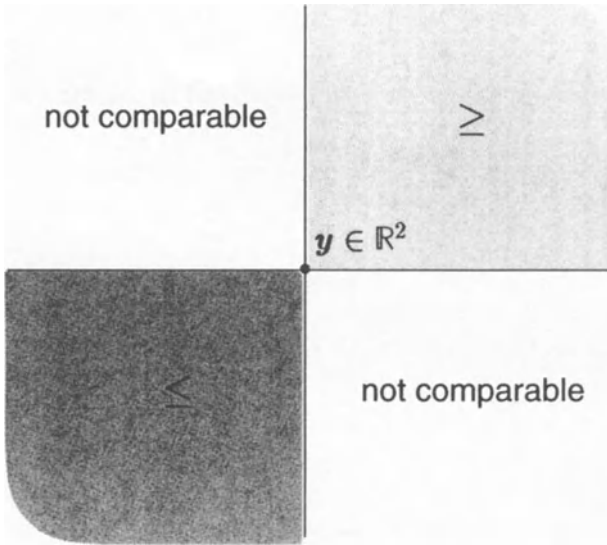


Figure 3.1 Vectors of the $\mathbb{R}^2$ as compared to some vector $\mathbf{y}$ according to the order relation $\leq$ defined above. The assertion $\mathbf{z} \geq \mathbf{y}$ is (defined as being) equivalent to $\mathbf{y} \leq \mathbf{z}$.

For the coordinates of a vector $\mathbf{y}^1$ which is unequal to $\mathbf{y}^2$ and which is smaller than $\mathbf{y}^2$ in the sense of ' $\leq$ ' we have: $\forall i \in \{1, \dots, k\} : y_i^1 \leq y_i^2$ and $\exists j \in \{1, \dots, k\}$, such that $y_j^1 < y_j^2$. If $\mathbf{y}^1$ and $\mathbf{y}^2$ represent two values of a vector-valued objective function, this means: $\mathbf{y}^1$ is at least as small (i.e. as good) as $\mathbf{y}^2$ with regard to all objectives and is strictly smaller (i.e. better) with regard to at least one objective. This ordering concept is the suitable formalization when comparing two technical solutions which are being assessed with regard to more than one criterion.

Essential properties of the order relation ' $\leq$ ' are:

- (a) There are vector pairs $\{\mathbf{y}^1, \mathbf{y}^2\}$ in the $\mathbb{R}^k$, which cannot be compared with regard to $\leq$, i.e. for which neither $\mathbf{y}^1 \leq \mathbf{y}^2$ nor $\mathbf{y}^2 \leq \mathbf{y}^1$ is true (see Figure 3.1). One example are the above-mentioned vectors $\mathbf{y}^1 = (4, 2)$ and

$$\mathbf{y}^2 = (2, 4).$$

This **partial non-comparability** reflects the fact that different objectives are of equal significance. This is why there is an essential difference between vector optimization problems and scalar-valued optimization problems; the objective space $\mathbb{R}$ of the latter possesses a total order induced by the natural numbers. The concrete meaning of total order is that for any two numbers $y^1, y^2 \in \mathbb{R}$ always $y^1 \leq y^2$ or $y^2 \leq y^1$ holds true.

(b) The order relation $\leq$ is a **partial order** in the $\mathbb{R}^k$, because:

- $\mathbf{y} \leq \mathbf{y} \quad \forall \mathbf{y} \in \mathbb{R}^k$ (reflexivity)
- $\mathbf{y}^1 \leq \mathbf{y}^2$ and $\mathbf{y}^2 \leq \mathbf{y}^3 \implies \mathbf{y}^1 \leq \mathbf{y}^3$ (transitivity)
- $\mathbf{y}^1 \leq \mathbf{y}^2$ and $\mathbf{y}^2 \leq \mathbf{y}^1 \implies \mathbf{y}^1 = \mathbf{y}^2$ (antisymmetry)

(c) Since the non-negative orthant $\mathbb{R}_+^k$ is a special case of a convex cone, $\leq$ is a **conic partial order**. Therefore, the compatibility of $\leq$ with the linear structure of the $\mathbb{R}^k$ is guaranteed:

- $\mathbf{y}^1, \mathbf{y}^2 \in \mathbb{R}^k, \mathbf{y}^1 \leq \mathbf{y}^2, \lambda \in \mathbb{R}, \lambda \geq 0 \implies \lambda \mathbf{y}^1 \leq \lambda \mathbf{y}^2$
- $\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3 \in \mathbb{R}^k, \mathbf{y}^1 \leq \mathbf{y}^2 \implies \mathbf{y}^1 + \mathbf{y}^3 \leq \mathbf{y}^2 + \mathbf{y}^3$

On the basis of this ordering concept the task of vector optimization can now be defined (see also [SAWARAGI ET AL., 1985] and [GÖPFERT & NEHSE, 1990]): It consists of finding those points $\mathbf{x}^* \in R$ the objective vectors $\mathbf{f}(\mathbf{x}^*)$ of which are ‘minimal’ with regard to the order relation $\leq$. Minimality with regard to $\leq$ is stated more precisely by defining an efficient point $\mathbf{y}^* \in \mathbb{R}^k$.

Definition 3.2: (Efficient point, Pareto optimal point, dominating point)

Let $\mathbf{f}(R)$ be the image set of the feasible set $R \subseteq \mathbb{R}^n$ under the vector-valued objective function $\mathbf{f}$. A point $\mathbf{y}^* \in \mathbf{f}(R) \subseteq \mathbb{R}^k$ is called (globally) **efficient** with regard to the order relation $\leq$ defined in $\mathbb{R}^k$, if and only if there exists no $\mathbf{y} \in \mathbf{f}(R)$, $\mathbf{y} \neq \mathbf{y}^*$, with $\mathbf{y} \leq \mathbf{y}^*$. A point $\mathbf{x}^* \in R$ with $\mathbf{y}^* = \mathbf{f}(\mathbf{x}^*)$ is called (globally) **Pareto² optimal³**, if and only if $\mathbf{y}^*$ is efficient.

A point $\mathbf{x}^1 \in R$ is said to **dominate** a point $\mathbf{x}^2 \in R$ if (and only if) $\mathbf{f}(\mathbf{x}^1) \neq \mathbf{f}(\mathbf{x}^2)$ and $\mathbf{f}(\mathbf{x}^1) \leq \mathbf{f}(\mathbf{x}^2)$. $\square$

² In some papers $\mathbf{x}^*$ is also called an Edgeworth-Pareto optimal point. In fact, Edgeworth published already in 1881 important contributions to the commonly called Pareto optimality concept (see [EDGEWORTH, 1881]). Pareto presented his results later in his book [PARETO, 1906], [PARETO, 1971]. For a historical discussion see [STADLER, 1987].

³ Some authors do not distinguish between ‘efficient points’ (which according to the above definition are elements of the objective space) and ‘Pareto optimal points’ (which according to the above definition are elements of the variable space). When the meaning becomes clear by the context, we will also sometimes use both terms synonymously in this book.

Hence, the aim of vector optimization is to find efficient points $\mathbf{y}^* \in \mathbf{f}(R)$ along with the Pareto optimal points $\mathbf{x}^*$ pertaining to them. If the different objectives are of equal importance, no efficient solution is distinguished a priori from any other efficient solution. The best what mathematics can do at this stage of the description of the problem is calculating **all efficient points** $\mathbf{y}^* \in \mathbf{f}(R)$ (along with all pertaining Pareto optimal points $\mathbf{x}^* \in R$). From this so-called efficient set $[\subset \mathbf{f}(R)]$ or Pareto set $[\subset R]$ the decision-maker can choose that particular solution which he thinks should be realized, using additional criteria not as yet taken into account in the description of the problem.

The homotopy method for generating Pareto optimal points which is developed in the present book is based on local properties of the objective function $\mathbf{f}$. In particular, when examining the Pareto optimality of the point $\mathbf{x}$ only points of a neighborhood of $\mathbf{x}$ are considered. Stating this ‘local comparison concept’ more precisely provides the term of local Pareto optimality:

Definition 3.3: (Locally Pareto optimal point)

A point $\mathbf{x}^* \in R$ is called **locally Pareto optimal**, if and only if there exists a neighborhood $U(\mathbf{x}^*)$ of $\mathbf{x}^*$ such that $\mathbf{y}^* := \mathbf{f}(\mathbf{x}^*)$ is efficient with regard to the (local) image set $\mathbf{f}(R \cap U(\mathbf{x}^*))$. Accordingly, $\mathbf{y}^*$ is called locally efficient. $\square$

Since globally Pareto optimal points are necessarily also locally Pareto optimal, our method provides locally Pareto optimal points as candidates for the (finally wanted) globally Pareto optimal points. The problem that optimization methods lead to local optimum values, whereas the application problem requires in many cases global optimum values, arises in the vector optimization in the same way as in the scalar-valued (‘ordinary’) optimization. A satisfactory solution to this problem both in the scalar-valued and in the vector-valued case is offered only by stochastic search methods because of their ability to escape from local minima by a purposeful use of stochastic effects (see [SCHÄFFLER, 1995]). In Section 3.3 we will present a stochastic method for searching for globally efficient solutions of unconstrained VOPs which has recently been developed by Schäffler et al. [SCHÄFFLER ET AL., 1999].

A disadvantage of stochastic search methods is the large number of required evaluations of the objective function. Especially in industrial cases of application, in which an evaluation of the objective function is based on a simulation of the system behavior which requires long computing times, it is therefore often sensible to apply fast local search methods.

Quite a few applications demand explicitly that the system variable be varied only locally. For example, the physical-mathematical system model has frequently only a local range of validity which one must not leave. When searching for an optimal operation point one often has knowledge about the system behavior

and its compliance with security requirements only for values of variables in the neighborhood of a tried and tested operation point. Within design problems excessive variations of the design variables are mostly unwelcome, because their effect on the costs is difficult to estimate and, therefore, not included in the objective function.

In all these aforesaid application cases the generation of locally Pareto optimal points is not only a first step (in the sense of obtaining candidates for the globally Pareto optimal points actually wanted), but already the completion of what the user can expect from mathematics.

In the following chapters — with the exception of Sections 3.2 and 3.3 where we survey the ‘state of the art’ in multiobjective optimization — we will develop a method for generating candidates for locally Pareto optimal points. These points are, of course, also candidates for globally Pareto optimal points. In most cases we will no longer distinguish between local and global Pareto optimality. The term efficiency will be treated analogously.

3.2 Survey of Methods

In the course of this section we will sketch briefly the most important existing methods which allow us to generate the set of efficient solutions of a given multiobjective optimization problem (or mostly: a subset of it). In this context we will moreover list the essential advantages of the generalized homotopy method as well as its limitations.

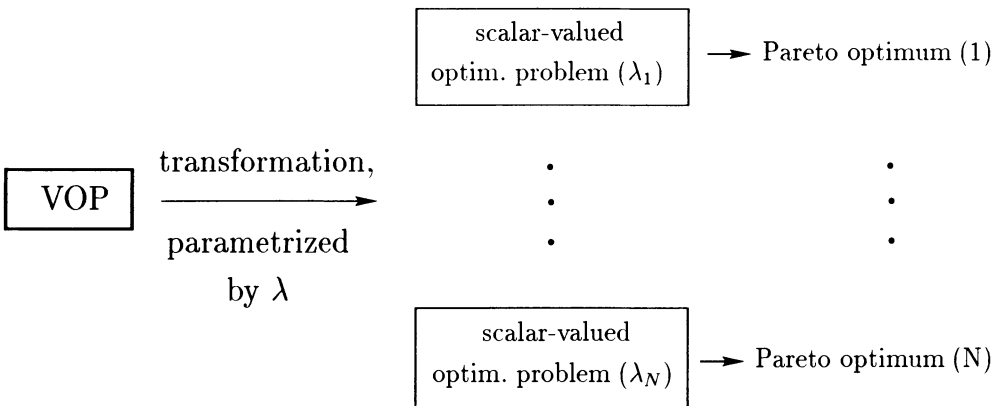


Figure 3.2 Basic idea of scalarization: The original VOP is transformed into scalar-valued optimization problems in a parametrizable way. By varying the transformation parameter λ and solving the resulting scalar-valued optimization problems one tries to generate different efficient points of the VOP.

First we will turn to the **deterministic methods**. These are (almost always) based on a ‘scalarization’ of the vector optimization problem, i.e. on the principle of transforming the vector optimization problem into a problem of scalar-valued optimization. Since by solving this scalar-valued optimization problem usually only a single efficient solution can be found, the transformation process (i.e. the scalarization) is formulated in a *parametrizable* way. By varying the (transformation) parameter different scalar-valued optimization problems and – in the form of their solutions – in general several efficient points can be generated. Figure 3.2 illustrates this basic idea schematically.

In particular, the following different approaches have to be mentioned:

(a) **Weighting method**

This method, which was first introduced by Zadeh [ZADEH, 1963], is still probably the most widely used vector optimization method. Its fundamental principle (see also Chapter 1) is to assign to each of the k (individual) objective functions a weight $\alpha_i \geq 0$ [normalized by $\sum_{i=1}^k \alpha_i = 1$] and to solve the substituting scalar problem

$$\min_{\mathbf{x} \in R} \sum_{i=1}^k \alpha_i f_i(\mathbf{x}). \quad (3.2)$$

The transformation parameter (for generating multiple scalar substituting problems) is the weight vector $\boldsymbol{\alpha} := (\alpha_1, \dots, \alpha_k)^T$, reflecting the ‘significance’ of the individual objectives. By varying $\boldsymbol{\alpha}$ one can obtain a subset of the total set of efficient solutions. Global minimizers of the scalar-valued substituting function $\sum_{i=1}^k \alpha_i f_i(\mathbf{x}) = \boldsymbol{\alpha}^T \cdot \mathbf{f}(\mathbf{x})$ are necessarily⁴ globally Pareto optimal solutions, local minimizers correspond necessarily to locally Pareto optimal points.

A geometric interpretation of the weighting method can easily be gained by considering the scalar-valued substituting function $g_{\boldsymbol{\alpha}}(\mathbf{x}) := \boldsymbol{\alpha}^T \cdot \mathbf{f}(\mathbf{x})$. Since $\boldsymbol{\alpha}^T \cdot \mathbf{f}(\mathbf{x}) = \text{constant}$ defines a plane in the objective space characterized by its normal vector $\boldsymbol{\alpha}$, each choice of the transformation parameter $\boldsymbol{\alpha}$ induces a partition of the objective space into planes of identical $g_{\boldsymbol{\alpha}}$ -values as shown in Figure 3.3.

Disadvantages of this method are (see also [DAS & DENNIS, 1996A]):

- In general vector optimization problems the image $\mathbf{f}(R)$ of the feasible set R need not be a convex subset of the objective space $\mathbb{R}^k$. For VOPs with non-convex $\mathbf{f}(R)$ there is an important class of efficient points, which are not minimizers of a scalar-valued substituting function of the type $\boldsymbol{\alpha}^T \cdot \mathbf{f}(\mathbf{x})$. Such points cannot be found with the weighting

⁴ If the value 0 is also permissible for individual weights α_i , the (global or local) Pareto optimality is guaranteed only for unique (global or local) minima.

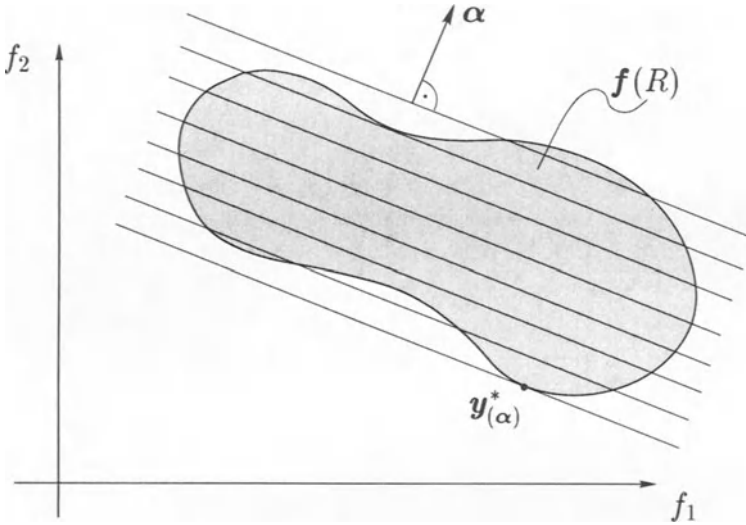


Figure 3.3 Scalarization of the vector-valued objective space as induced by the weighting method (see text).

method.

Section 4.4 provides a detailed discussion of this class of efficient points.

- Since any numerical method solving a VOP will only be able to compute a limited number of efficient points (or candidates), it becomes crucial to have these points be spread in the objective space as uniformly as possible, so that a good approximation of the whole efficient set is obtained. The weighting method fails to meet this requirement and generates an irregular discretization of the set of efficient points. The distances between the points generated in the objective space cannot be controlled directly.

(b) **Weighted L_p -metric method**

Methods of this kind choose a ‘desired point’ $\bar{\mathbf{y}}$ in the objective space $\mathbb{R}^k$ and search for efficient solutions which come ‘as close as possible’ to this point $\bar{\mathbf{y}}$. Both the desired point $\bar{\mathbf{y}}$ and the metric, by means of which the deviation of a solution $\mathbf{f}(\hat{\mathbf{x}})$ from the desired point $\bar{\mathbf{y}}$ is quantified, can be varied. If the objective space $\mathbb{R}^k$ is metrized by means of a weighted vector norm ρ_p of the form $\rho_p(\mathbf{y}) = \left(\sum_{i=1}^k \omega_i |y_i|^p \right)^{\frac{1}{p}}$, with $p \in [1, \infty) \cup \{\infty\}$, $\omega_i > 0$ and $\sum_{i=1}^k \omega_i = 1$,⁵ one obtains the following scalar substituting problem:

$$\min_{\mathbf{x} \in R} \sum_{i=1}^k \omega_i |f_i(\mathbf{x}) - \bar{y}_i|^p. \quad (3.3)$$

⁵ ρ_∞ is defined as $\rho_\infty(\mathbf{y}) := \max\{\omega_1|y_1|, \dots, \omega_k|y_k|\}$, i.e. as a weighted maximum norm.

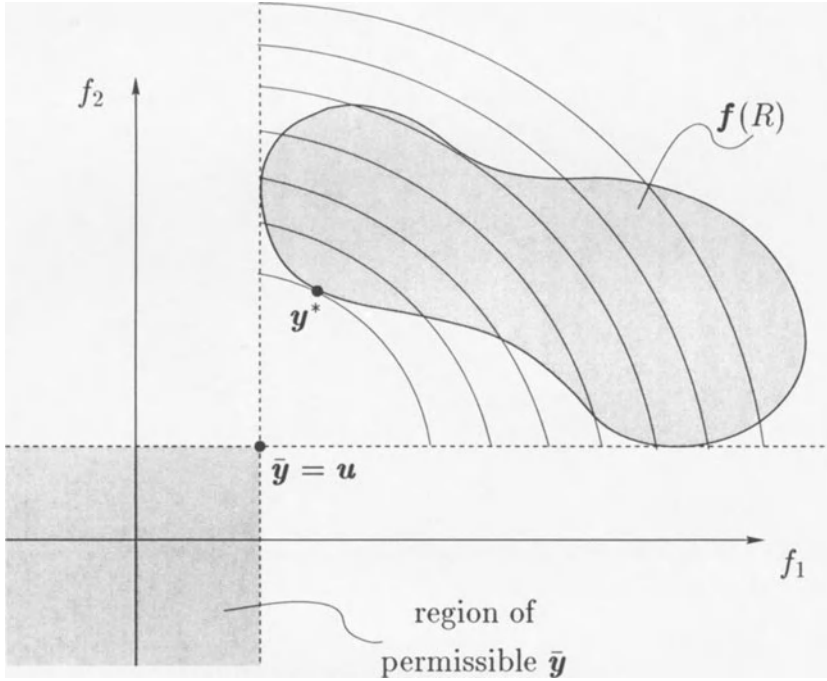


Figure 3.4 Scalarization of the vector-valued objective space as induced by the weighted L_p -metric method. The quarter-circles are the curves (in the objective space) along which the scalar-valued substituting function of (3.3) is constant. Here, the transformation parameters have been chosen as $\bar{\mathbf{y}} = \mathbf{u}$, $\omega = (1, 1)^T$ and $p = 2$.

The scalarization achieved in that way is shown by means of the resulting contours in Figure 3.4.

In the case of $p \in [1, \infty)$ the solutions of the substituting problems thus obtained are necessarily efficient. The desired point $\bar{\mathbf{y}}$, the weighting vector ω (see weighting method) and the exponent p of the vector norm are the transformation parameters of the scalarization. $\bar{\mathbf{y}}$ has to meet the requirement $\bar{y}_i \leq f_i(\mathbf{x}) \ \forall \mathbf{x} \in R, i \in \{1, \dots, k\}$. For instance, this is true for the so-called ideal objective vector (also called utopia point) $\mathbf{u}$, the i -th component of which is the global minimum of the individual objective function f_i .

When including the $(p = \infty)$ -norm, in principle all efficient solutions can be generated by varying $(\bar{\mathbf{y}}, \omega, p)$. However, it is just the strategy of a *meaningful* variation of the transformation parameters which is the main problem of this class of methods. In particular, one cannot make a universal statement about the conditions, under which it is actually possible to generate different efficient solutions by controlling the parameters $(\bar{\mathbf{y}}, \omega, p)$ (see [GÖPFERT & NEHSE, 1990]).

(c) ϵ -constraint method

This method goes back to Marglin [MARGLIN, 1967] and Haimes [HAIMES, 1973]. It chooses one individual objective f_j , $j \in \{1, \dots, k\}$ to be minimized. For each of the other objectives an upper level is fixed which must not be exceeded. Hence, the scalar substituting problem has the following appearance

$$\begin{aligned} & \min_{\mathbf{x} \in (R \cap C)} f_j(\mathbf{x}) \text{ for a } j \in \{1, \dots, k\} \\ & C := \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) \leq \epsilon_i \text{ } \forall i \in \{1, \dots, k\} \text{ with } i \neq j\}. \end{aligned} \quad (3.4)$$

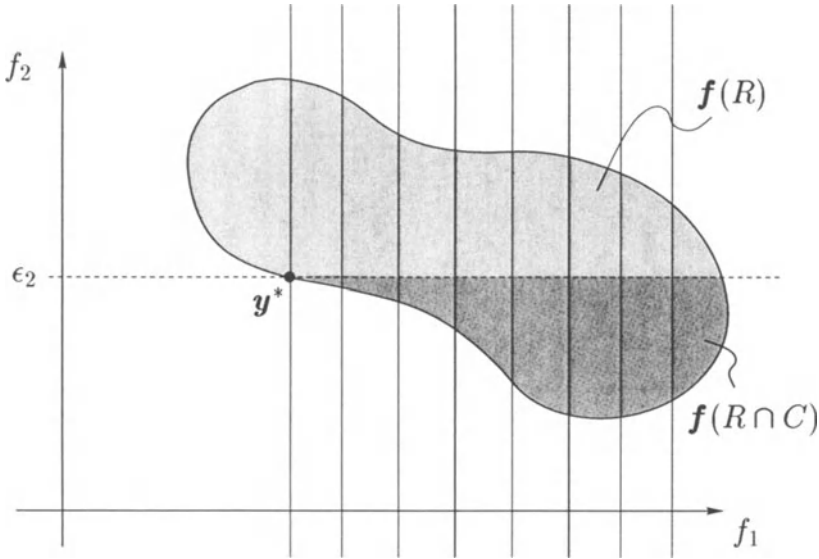


Figure 3.5 Scalarization of the objective space as induced by the ϵ -constraint method. f_1 has been chosen as the reference objective to be minimized.

A unique global minimizer of problem (3.4) is necessarily a globally Pareto optimal solution of the original VOP, a unique local minimizer of (3.4) is necessarily a locally Pareto optimal point. Figure 3.5 sketches the contours introduced into the objective space according to this transformation of the VOP. The index j of the chosen reference objective f_j and the upper levels ϵ_i for the other objectives play the part of transformation parameters. By varying these parameters all efficient points are in principle attainable.

The main difficulty of the ϵ -constraint method consists in finding the range of reasonable values for the upper levels. By choosing too low (i.e. too restrictive) levels ϵ_i one frequently generates scalar substituting problems which do not possess a feasible solution. If the upper levels of certain objectives, on the other hand, are set too high, these objectives cease to play

a part in the substituting problem (3.4), so that by local variation of the upper levels in question no new solution points are generated.

(d) **Method of equality constraints**

The method of equality constraints proposed by Lin [LIN, 1976] also chooses an objective f_j to be minimized and therefore is closely related to the ϵ -constraint method. The remaining objectives enter the scalar-valued substituting problem in the form of equality constraints:

$$\begin{aligned} \min_{\mathbf{x} \in (R \cap D)} f_j(\mathbf{x}) \quad \text{for a } j \in \{1, \dots, k\} \\ D := \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) - \epsilon_i = 0 \quad \forall i \in \{1, \dots, k\} \text{ with } i \neq j\} \end{aligned} \quad (3.5)$$

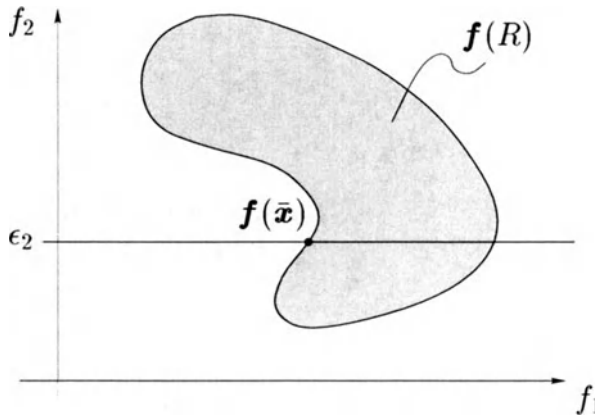


Figure 3.6 Illustration of a situation where a global minimizer $\bar{\mathbf{x}}$ of the scalar-valued problem (3.5) [with f_1 chosen as the reference objective to be minimized and ϵ_2 set as the equality constraint value for the second objective f_2] is not Pareto optimal with respect to the original VOP.

By varying the transformation parameters, i.e. the index j of the reference objective f_j and the constraint values ϵ_i for the remaining objectives, in principle all efficient points can be attained. However, as can be seen from Figure 3.6, the Pareto optimality of a solution of the scalar-valued substituting problem (3.5) is not automatically guaranteed, but has to be verified – e.g. by making use of a necessary and sufficient condition as indicated by [LIN, 1976].

Similarly to the ϵ -constraint method, the method of equality constraints has the disadvantage that many of the generated scalar substituting problems do not have a feasible solution.

(e) Normal-Boundary Intersection

A technique which has been devised recently by Das and Dennis [DAS & DENNIS, 1996B] and is called Normal-Boundary Intersection (NBI) scalarizes the multiobjective problem in a geometrically motivated way.

At first, for all individual objectives $f_i, i \in \{1, \dots, k\}$ the respective global minimizers $\mathbf{x}_i^* \in R$ are required. The convex hull of the individual minima in the objective space, called (CHIM) [DAS & DENNIS, 1996B], i.e. the convex hull of the vectors $\{f(\mathbf{x}_1^*), \dots, f(\mathbf{x}_k^*)\} \subset \mathbb{R}^k$, can be expressed by means of the matrix $\Phi := (f(\mathbf{x}_1^*) \dots f(\mathbf{x}_k^*)) \in \mathbb{R}^{k \times k}$ as $\{\Phi\beta \mid \beta \in \mathbb{R}^k, \sum_{i=1}^k \beta_i = 1, \beta_i \geq 0\}$. It represents a simplex, the points of which are characterized (or parametrized) by the weight vector β . Figure 3.7 illustrates the CHIM-simplex in the bicriterial case.

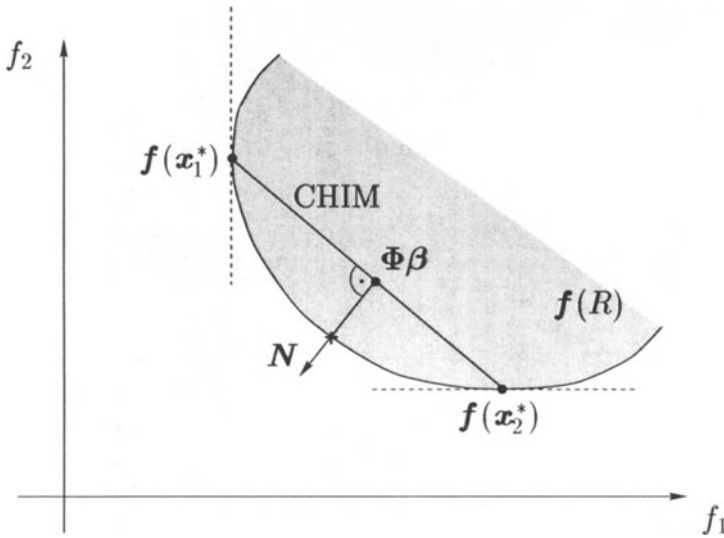


Figure 3.7 Illustration of the CHIM-simplex in a bicriterial example with a convex image set $f(R)$. The solution of the substituting NBI-problem (with the starting point $\Phi\beta$ on the CHIM-set) is marked by +.

The NBI-approach is based on the following observation, from which also the name of the method is derived: If to an arbitrary point $\Phi\beta$ of the CHIM one attaches the unit normal vector N to the CHIM-simplex (oriented towards the negative orthant), under certain circumstances the half-line generated by N intersects the boundary of the image set $f(R)$ in an efficient point. In Figure 3.7 this point of intersection is indicated with a '+'. The problem of finding the point of intersection can be expressed as a scalar

optimization problem (NBI substituting problem):

$$\begin{aligned} \min_{\substack{x \in R \\ t \in \mathbb{R}}} -t \quad & \text{with the additional constraint} \\ \Phi\beta + t \cdot N = f(x), \end{aligned} \quad (3.6)$$

where $\Phi\beta$ denotes the starting point on the CHIM-simplex. By varying the weight vector β , i.e. by varying the starting point on the CHIM-simplex and by solving the resulting NBI substituting problems a subset of the efficient set can be generated.

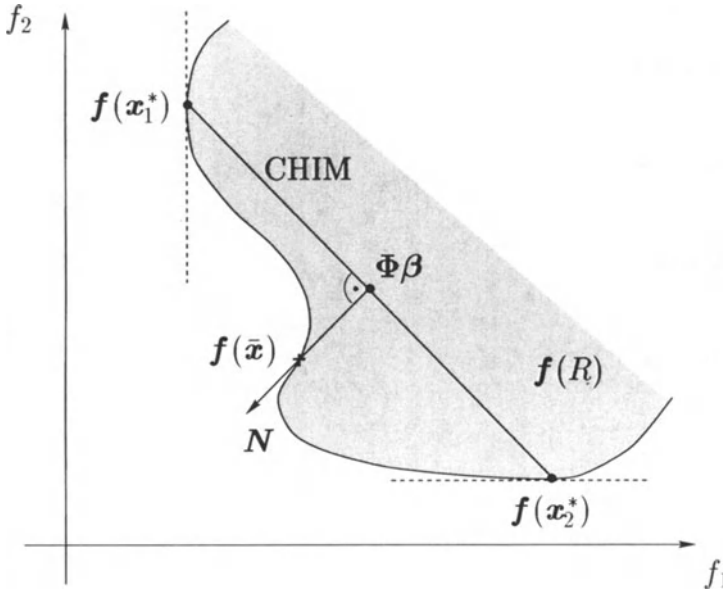


Figure 3.8 A situation [with non-convex image $f(R)$] where a global minimizer $\bar{x}$ of an NBI subproblem is not Pareto optimal with respect to the original VOP.

In bicriterial vector optimization problems, for every Pareto optimal point x^* there exists a corresponding NBI substituting problem of which x^* is the solution. For more than two objectives (i.e. for $k \geq 3$), however, this assertion is no longer true (see also [DAS & DENNIS, 1996B]). A simple counterexample for the case of $k = 3$ is a VOP the image $f(R)$ of the feasible set R of which is a sphere in $\mathbb{R}_+^3$ touching the coordinate axes. Then the CHIM-simplex is the triangle formed by joining the three points where the sphere touches the axes. We extend the boundary of the CHIM-simplex — while remaining on the plane containing the CHIM-simplex — until it touches the boundary of the sphere $f(R)$ and denote the extended CHIM by CHIM_+ . Now it can be stated that there exist points in $\text{CHIM}_+ \setminus \text{CHIM}$

underneath which there are efficient points on the sphere. Those efficient points are not solutions of an NBI substituting problem.

As a further drawback, in VOPs with a non-convex image $f(R)$ of the feasible set R a solution of an NBI substituting problem is not necessarily a Pareto optimal point (not even necessarily locally Pareto optimal) as can be seen in Figure 3.8.

(f) Homotopy approach

The only proposal known by the author to use the homotopy method for solving multiobjective problems was made by Rakowska et al. [RAKOWSKA ET AL., 1991] and has already been introduced in Chapter 1. We discussed there that this proposed method has two serious disadvantages: namely, on the one hand, the method limits itself by construction to bicriterial problems and, on the other hand, only those Pareto optimal points are determined which are minima of a convex combination of the individual objectives.

The generalized homotopy method which is developed in this book eliminates both shortcomings. In order to judge the newly developed homotopy method in comparison to the multiobjective optimization methods surveyed above and in order to furnish the user with criteria under which circumstances the application of this method is possible or useful, we will now list its assets and limitations.

The following advantages must be emphasized (see also Chapter 1):

- (+) The method makes also such Pareto optimal points accessible which are not minima of a convex combination of the objectives (cf. the weighting method).
- (+) The method attains a high numerical efficiency by making extensive use of the linearized information about the zero manifold of Pareto candidates.
- (+) The discretization density of the efficient set in the objective space can be controlled in a simple way. In particular, the method is capable of generating a homogeneous discretization of the efficient set.
- (+) For each of the calculated solutions the relative valences of the individual objectives (in this solution point) are provided, so that the decision-maker obtains valuable information for the interpretation of this solution point.

For these assets, however, one has to pay with some limitations or potential disadvantages:

- The applicability of the method presupposes that both the vector-valued objective function and the functions defining the constraints

have to be twice continuously differentiable. Moreover, the method requires an explicit calculation of the Hessian matrix of the Lagrangian function.

- Since the method is based by construction on local properties of the objective and restriction functions, the global Pareto optimality of a generated point is not automatically guaranteed. If in a generated point the Hessian matrix of the Lagrangian function — which has been calculated in the course of the methodical procedure — is positive definite in an appropriate linear subspace, then its local Pareto optimality is thereby guaranteed (which in many cases of application is sufficient, see Section 3.1).
- If a vector optimization problem with inequality constraints has to be solved by means of the developed method, either slack variables have to be introduced or one has to adopt an active-set-strategy. As mentioned briefly in footnote 1 on page 65 both approaches are not entirely unproblematic.
- By starting from one Pareto optimal point, the whole Pareto set can be generated by means of homotopy only if this Pareto set is contained in a *connected* differentiable manifold. If, however, the manifold of Pareto candidates is composed of more than one connection components, one starting point for each connection component is required to calculate the entire Pareto set.

The second class of methods for solving multiobjective optimization problems are **stochastic methods**. Here, stochasticity is the crucial feature which makes it possible to generate not only a single efficient solution, but a whole set of efficient points without changing the instruction of the method. As already mentioned in Chapter 1, typical applications of stochastic vector optimization methods are characterized by the search for *globally* efficient solutions and by sufficiently available computing time. After surveying two well-known stochastic methods in the following, in Section 3.3 we will acquaint the reader in some more detail with a recent and promising stochastic approach for multiobjective optimization developed by Schäffler et al. [SCHÄFFLER ET AL., 1999].

(g) Stochastic search according to Timmel

Starting point of this method stated by Timmel [TIMMEL, 1980] is a set of realizations of an n -dimensional random variable (i.e. a set of n -dimensional random numbers), which is distributed equally on the feasible set⁶ $R \subseteq \mathbb{R}^n$ and which outside R has the probability density 0. From this set of realizations all points $\mathbf{x}$ which are dominated by some other point of this set, i.e. for which there exists a point $\bar{\mathbf{x}}$ of this set with $\mathbf{f}(\bar{\mathbf{x}}) \leq \mathbf{f}(\mathbf{x})$, are eliminated. The point set thinned out in this manner is a first approximation

⁶ The feasible set R in this method is assumed to be compact.

of the set of Pareto optimal points, $A_0 = \{\mathbf{x}_1^0, \dots, \mathbf{x}_{r_0}^0\} \subset R$. Here, $r_0 \in \mathbb{N}$ denotes the cardinal number of the approximation set A_0 in the 0-th iteration step.

As of now the approximation set is subject to an iteration instruction $A_l \mapsto A_{l+1}$. In this process one tries to generate out of every point $\mathbf{x}_j^l \in A_l$ a new point which is not dominated by $\mathbf{x}_j^l$. This is done by choosing stochastically a search direction (normalized to the length 1) out of the polyhedral cone which is generated by the negative gradients of the individual objectives, $-\nabla f_i(\mathbf{x}_j^l)$, $i = 1, \dots, k$. The steplength t_l is reduced between successive iterations. After having generated new points out of all points $\mathbf{x}_j^l$ according to this instruction [i.e. after having completed for each point $\mathbf{x}_j^l$ a search step for a new point not dominated by $\mathbf{x}_j^l$], one obtains the improved approximation set A_{l+1} by uniting A_l with the newly generated points and by eliminating all points that are dominated by another point of this sum-set.

For differentiable objective functions Timmel is able to show the following stochastic convergence of the approximation set A_l towards the set of Pareto optimal points: For all $\epsilon > 0$ and all efficient points $\bar{\mathbf{y}} \in \mathbf{f}(R)$ one has:

$$\lim_{l \rightarrow \infty} P(\exists \mathbf{x} \in A_l : \|\mathbf{f}(\mathbf{x}) - \bar{\mathbf{y}}\| < \epsilon) = 1, \quad (3.7)$$

where $P(\cdot)$ denotes the probability of the event described within the brackets.

(h) Evolutionary algorithms

By the term ‘evolutionary algorithms’ a class of heuristic optimization algorithms is subsumed which simulate the survival of the fittest in biological evolution by means of algorithms. Originally devised for scalar-valued optimization problems, evolutionary algorithms are appropriate also for multiobjective optimization, since they utilize an entire set (‘population’) of variable points simultaneously. Similarly to Timmel’s approach the population is interpreted as an approximation of a subset of the efficient set which is improved from iteration to iteration with regard to this approximation property.

The renewal of a population is based on the so-called ‘genetic operators’: recombination (out of two points of the population picked out at random a new point is generated, e.g. by averaging), mutation (single, randomly selected digits of the newly generated point are substituted by a realization of a random variable) and selection (out of the union of the (original) population and the newly generated points those with the best ‘fitness’ are taken over into the new population). There exists a great variety of detailed iteration instructions which follow this principle.

In scalar-valued optimization problems the objective function itself serves mostly as fitness. For vector optimization Goldberg (see

[GOLDBERG, 1989]) was the first to propose a Pareto-based fitness assignment: First the whole population is considered, and all points which are not dominated by some other point of the population receive the (fitness) rank 1. Then the difference between the population and the rank-1 points is looked upon, and to all non-dominated points in this difference set the rank 2 is assigned, etc. A related scheme also assigns a rank to each point of a population. That rank is given by the number of those points within the population which dominate this point.

Since the aim is to arrive after a given number of iterations at a population which can be regarded as an approximately homogeneous discretization of the efficient set, one has to prevent the population from converging to a single point. To this end one introduces some repulsion between the points of the population. This can be accomplished by means of the so-called fitness sharing technique, where one diminishes the fitness of a point depending on the number of points of the population which are situated in its close neighborhood.

A survey of evolutionary algorithms for vector optimization can be found in [FONSECA & FLEMING, 1995].

Since evolutionary algorithms do not utilize derivative information (which of course is eminently important for accelerating the search for (Pareto) optima), their use is recommended only for non-differentiable objective functions or for problems, in which the gradients can be evaluated only numerically. A serious disadvantage of this class of methods is furthermore that in virtue of the ‘wide range hopping’ of points which is induced by recombination and because of the couplings between the points of the population entailed by selection no universally applicable assertions on stochastic convergence can be made.

The following section will be devoted to a recently developed stochastic method for multiobjective optimization. We shall present this promising and theoretically well-founded approach in greater detail in order to stimulate further research in that area.

3.3 A New Stochastic Method for Unconstrained Vector Optimization

The method has been devised by Schäffler et al. [SCHÄFFLER ET AL., 1999] for the solution of the following unconstrained vector optimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{f}(\mathbf{x}) . \quad (3.8)$$

Here, the objective function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is assumed to be twice continuously differentiable. ‘Minimization’ requires calculating all (or a large number) of the

Pareto optimal points.

The basic idea is to construct a deterministic dynamics resulting in a special curve of dominated points and to perturb this dynamics by a stochastic (Brownian) motion. Paragraph 3.3.1 introduces the (sophisticated) ordinary differential equation which defines the basic curve of dominated points. This deterministic dynamics is perturbed by a Brownian motion and thus forms the drift part of a stochastic differential equation (SDE) which is introduced and analyzed in Paragraph 3.3.3. The numerical solution of this SDE yields an algorithm for computing (a large number of) Pareto optimal points.

The discussion given here follows closely the original paper by Schäffler et al. [SCHÄFFLER ET AL., 1999].

3.3.1 A Curve of Dominated Points

The deterministic part of the method can be motivated by looking at scalar-valued optimization from the viewpoint of numerical mathematics: A large class of algorithms for the unconstrained minimization of a (twice continuously differentiable) scalar-valued objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ can be interpreted as special numerical solutions of the following initial value problem

$$\dot{\mathbf{x}}(t) = -\nabla f(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad (3.9)$$

where $\nabla f(\mathbf{x})$ denotes the gradient of f at $\mathbf{x} \in \mathbb{R}^n$. The solution $\mathbf{x} : [0, \infty[\rightarrow \mathbb{R}^n$ of this initial value problem can be considered a special temporal parametrization of the *curve of steepest descent*, which in each point follows the direction of the negative gradient. In particular, $\mathbf{x}(t)$ consists of points with decreasing function values, which means that if $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$, then

$$f(\mathbf{x}(t)) < f(\mathbf{x}(s)) \quad \text{for all } 0 \leq s < t < \infty. \quad (3.10)$$

Schäffler et al. generalize this curve-of-steepest-descent approach to unconstrained vector optimization problems of the form (3.8). The role of a curve of points with ever decreasing function values approaching a (local) minimizer is played in vector optimization by a *curve of dominated points* — i.e. a curve each point of which is dominated by all its successors — approaching a (local) Pareto optimum.

In order to realize that plan one has to formulate an initial value problem (IVP); its unique solution is a curve $\mathbf{x} : [0, \infty[\rightarrow \mathbb{R}^n$ consisting of dominated points, i.e.

$$f(\mathbf{x}(t)) \leq f(\mathbf{x}(s)) \quad \text{and} \quad f(\mathbf{x}(t)) \neq f(\mathbf{x}(s)) \quad \text{for all } 0 \leq s < t < \infty. \quad (3.11)$$

For the construction of such an IVP, we consider the following quadratic optimization problem for each $\mathbf{x} \in \mathbb{R}^n$:

$$(\text{QOP}(\mathbf{x})) \quad \min_{\alpha \in \mathbb{R}^k} \left\{ \left\| \sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}) \right\|^2 \mid \sum_{i=1}^k \alpha_i = 1 \text{ and } \alpha_i \geq 0, i = 1, \dots, k \right\}. \quad (3.12)$$

Since $\sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}) = \nabla \left(\sum_{i=1}^k \alpha_i f_i \right) (\mathbf{x})$, $\text{QOP}(\mathbf{x})$ searches for that weight vector α for which the convex combination $g_\alpha(\mathbf{x}) := \sum_{i=1}^k \alpha_i f_i(\mathbf{x})$ of the individual objectives has the smallest gradient (with respect to its Euclidean norm).

The following two properties of $\text{QOP}(\mathbf{x})$ result from convex analysis:

- (a) For each $\mathbf{x} \in \mathbb{R}^n$ there exists a global minimizer $\hat{\alpha}$ of $\text{QOP}(\mathbf{x})$, which is not unique in general. Each local minimizer of $\text{QOP}(\mathbf{x})$ is also a global minimizer.
- (b) Let $\hat{\alpha}$ and $\tilde{\alpha}$ be two global minimizers of $\text{QOP}(\mathbf{x})$ for a fixed $\mathbf{x} \in \mathbb{R}^n$, then

$$\sum_{i=1}^k \hat{\alpha}_i \nabla f_i(\mathbf{x}) = \sum_{i=1}^k \tilde{\alpha}_i \nabla f_i(\mathbf{x}) \quad (3.13)$$

Taking these properties into account we define the function

$$\mathbf{q} : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^n \\ \mathbf{x} & \mapsto \mathbf{q}(\mathbf{x}) := \sum_{i=1}^k \hat{\alpha}_i \nabla f_i(\mathbf{x}) = \nabla g_{\hat{\alpha}}(\mathbf{x}) \end{cases}, \quad (3.14)$$

where $g_{\hat{\alpha}} := \sum_{i=1}^k \hat{\alpha}_i f_i$ is the convex combination of the individual objectives f_i characterized by the weight vector $\hat{\alpha}$, and where $\hat{\alpha}$ is a global minimizer of $\text{QOP}(\mathbf{x})$. The following theorem investigates this function $\mathbf{q}$.

Theorem 3.1:

Consider $\text{QOP}(\mathbf{x})$ and let $\mathbf{q}$ be the function defined by (3.14), where $\hat{\alpha}$ is a global minimizer of $\text{QOP}(\mathbf{x})$. Then the following two assertions are true:

- (i) *Either $\mathbf{q}(\mathbf{x}) = \mathbf{0}$ holds, or $-\mathbf{q}(\mathbf{x})$ is a descent direction for all individual objective functions $f_1, \dots, f_k$ at $\mathbf{x}$.*
- (ii) *The function $\mathbf{q}$ is locally Lipschitzian, i.e. for each $\hat{\mathbf{x}} \in \mathbb{R}^n$ there exists a neighborhood $U(\hat{\mathbf{x}})$ and a constant $L_{\hat{\mathbf{x}}} \in \mathbb{R}_0^+$ such that*

$$\|\mathbf{q}(\mathbf{x}) - \mathbf{q}(\mathbf{y})\| \leq L_{\hat{\mathbf{x}}} \|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in U(\hat{\mathbf{x}}). \quad (3.15)$$

Proof. Ad (i): Define the set $K(\mathbf{x})$ of gradients for all convex combinations g_α (of the objectives) at the point $\mathbf{x}$

$$K(\mathbf{x}) := \left\{ \sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}) \mid \sum_{i=1}^k \alpha_i = 1 \text{ and } \alpha_i \geq 0, i = 1, \dots, k \right\} \quad (3.16)$$

and assume that $\mathbf{0} \notin K(\mathbf{x})$ for any fixed $\mathbf{x} \in \mathbb{R}^n$. Assume furthermore that there exists a vector $\mathbf{v}(\mathbf{x}) \in K(\mathbf{x})$ with $\mathbf{q}(\mathbf{x})^T \mathbf{v}(\mathbf{x}) \leq 0$, then we obtain the following properties of the vectors $\lambda(\mathbf{q}(\mathbf{x}) - \mathbf{v}(\mathbf{x}))$, $0 \leq \lambda \leq 1$:

(A) $(\mathbf{v}(\mathbf{x}) + \lambda(\mathbf{q}(\mathbf{x}) - \mathbf{v}(\mathbf{x}))) \in K(\mathbf{x})$ for all $0 \leq \lambda \leq 1$.

(B) $\mathbf{q}(\mathbf{x})^T (\lambda(\mathbf{q}(\mathbf{x}) - \mathbf{v}(\mathbf{x}))) > 0$ for all $0 < \lambda \leq 1$.

Let $\hat{\lambda}$ be the global minimizer of the quadratic optimization problem

$$\min_{0 \leq \lambda \leq 1} \{ \|\mathbf{v}(\mathbf{x}) + \lambda(\mathbf{q}(\mathbf{x}) - \mathbf{v}(\mathbf{x}))\|^2 \}, \quad (3.17)$$

then it is obvious that $\|\mathbf{v}(\mathbf{x}) + \hat{\lambda}(\mathbf{q}(\mathbf{x}) - \mathbf{v}(\mathbf{x}))\|^2 < \|\mathbf{q}(\mathbf{x})\|^2$, because

$$\hat{\lambda} = 1 \iff \mathbf{q}(\mathbf{x})^T (\hat{\lambda}(\mathbf{q}(\mathbf{x}) - \mathbf{v}(\mathbf{x}))) \leq 0 \text{ for all } 0 < \lambda \leq 1. \quad (3.18)$$

Since $\mathbf{v}(\mathbf{x}) + \hat{\lambda}(\mathbf{q}(\mathbf{x}) - \mathbf{v}(\mathbf{x})) \in K(\mathbf{x})$, we obtain a contradiction to the definition of $\mathbf{q}$. Hence, $\mathbf{v}(\mathbf{x})^T \mathbf{q}(\mathbf{x}) > 0$ for all $\mathbf{v}(\mathbf{x}) \in K(\mathbf{x})$. As all gradients $\nabla f_1(\mathbf{x}), \dots, \nabla f_k(\mathbf{x})$ are elements of $K(\mathbf{x})$, this implies assertion (i).

Ad (ii): Consider the following system of nonlinear equations with inequalities in $(\boldsymbol{\alpha}(\mathbf{x}), \boldsymbol{\xi}(\mathbf{x}), \lambda(\mathbf{x}), \boldsymbol{\mu}(\mathbf{x})) \in \mathbb{R}^{k+n+1+k}$, where $\mathbf{e}_i$ denotes the i -th unit vector and $(\nabla f_1(\mathbf{x}) \dots \nabla f_k(\mathbf{x})) \in \mathbb{R}^{n \times k}$:

$$\begin{aligned} (\nabla f_1(\mathbf{x}) \dots \nabla f_k(\mathbf{x})) \boldsymbol{\alpha}(\mathbf{x}) - \boldsymbol{\xi}(\mathbf{x}) &= \mathbf{0} \\ (\nabla f_1(\mathbf{x}) \dots \nabla f_k(\mathbf{x}))^T \boldsymbol{\xi}(\mathbf{x}) - \lambda(\mathbf{x}) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \sum_{i=1}^k \mu_i(\mathbf{x}) \mathbf{e}_i &= \mathbf{0} \\ \sum_{i=1}^k \alpha_i(\mathbf{x}) - 1 &= 0 \\ \mu_i(\mathbf{x}) \alpha_i(\mathbf{x}) &= 0 \quad i = 1, \dots, k \\ \mu_i(\mathbf{x}) &\leq 0 \quad i = 1, \dots, k \\ \lambda(\mathbf{x}) &\geq 0. \end{aligned} \quad (3.19)$$

The system (3.19) represents the necessary and sufficient conditions of first order for global minimizers of $\text{QOP}(\mathbf{x})$. Assuming that $(\hat{\boldsymbol{\alpha}}(\mathbf{x}), \mathbf{q}(\mathbf{x}), \hat{\lambda}(\mathbf{x}), \hat{\boldsymbol{\mu}}(\mathbf{x}))$ is a solution of (3.19) for a fixed $\mathbf{x} \in \mathbb{R}^n$, we obtain:

- (1) $\mathbf{q}(\mathbf{x})$ is unique (cf. Theorem 3.1 (i)).
- (2) $\hat{\lambda}(\mathbf{x})$ and $\hat{\boldsymbol{\mu}}(\mathbf{x})$ are unique.

Let $\{\mathbf{x}_i\}_{i \in \mathbb{N}}$ be a sequence of vectors $\mathbf{x}_i \in \mathbb{R}^n$ which converges to a point $\bar{\mathbf{x}} \in \mathbb{R}^n$. Then the sequences $\{\mathbf{q}(\mathbf{x}_i)\}_{i \in \mathbb{N}}$ and $\{\hat{\boldsymbol{\alpha}}(\mathbf{x}_i)\}_{i \in \mathbb{N}}$ are bounded, and there exist convergent subsequences $\{\mathbf{q}(\mathbf{x}_j)\}$ with limit $\bar{\mathbf{q}}$ and $\{\hat{\boldsymbol{\alpha}}(\mathbf{x}_j)\}$ with limit $\bar{\boldsymbol{\alpha}}$. Therefore we obtain a vector $(\bar{\boldsymbol{\alpha}}, \bar{\mathbf{q}}, \bar{\lambda}, \bar{\boldsymbol{\mu}})$ that solves (3.19) at $\mathbf{x} = \bar{\mathbf{x}}$.

Because of Theorem 3.1 $\tilde{q}$ is equal to $q(\bar{x})$ and q , $\tilde{\lambda}$ and $\tilde{\mu}$ are continuous functions. If $\alpha_i(\bar{x})$ is unique and greater than zero for all $i \in \{1, \dots, k\}$ with $\hat{\mu}_i(\bar{x}) = 0$, then q is continuously differentiable in a neighborhood of $\bar{x}$. Otherwise, there exist a finite number of points $x_1, \dots, x_l$ and closed neighborhoods $U(x_1), \dots, U(x_l)$ of these points such that

- (1) x_i is an inner point of $U(x_i)$ for $i = 1, \dots, l$.
- (2) $\bar{x} \in U(x_i)$ for $i = 1, \dots, l$.
- (3) $\bar{x}$ is an inner point of $U(x_1) \cup \dots \cup U(x_l)$.
- (4) The function q restricted to $U(x_i)$ is for all $i = 1, \dots, l$ a continuously differentiable rational function in some components of the first order derivatives of the objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ of (3.8) (see system (3.19)).

Hence, q is locally Lipschitzian, because $f \in C^2$. ■

Inspecting claim (i) of Theorem 3.1 one may ask for an interpretation of the case $q(x) = 0$.

As will be discussed in detail in Section 4.1, for a Pareto optimal solution x^* of the unconstrained vector optimization problem (3.8) there necessarily exists a weight vector α^* (i.e. $\sum_{i=1}^k \alpha_i^* = 1$ and $\alpha_i^* \geq 0$, $i = 1, \dots, k$) such that x^* is a stationary point of the corresponding convex combination $g_{\alpha^*} := \sum_{i=1}^k \alpha_i^* f_i$ of the individual objectives, i.e. $\nabla g_{\alpha^*}(x^*) = 0$. Since $q(\bar{x}) = 0$ implies the existence of such a weight vector, the feature $q(\bar{x}) = 0$ qualifies the point $\bar{x}$ to meet the (first order) necessary condition for a Pareto point and thus to be a candidate for a Pareto optimal solution of (3.8).

The properties of the function q enable us to generalize the curve-of-steepest-descent approach to scalar optimization problems to the following initial value problem for unconstrained vector optimization problems of the form (3.8):

$$\dot{x}(t) = -q(x(t)), \quad x(0) = x_0, \quad (3.20)$$

where $q : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined in (3.14). Assuming that the set of variable points $x \in \mathbb{R}^n$ dominating the starting point x_0 is bounded, the following theorem proves the existence of a curve of dominated points which is the unique solution of the initial value problem (3.20).

Theorem 3.2:

Consider the vector optimization problem (3.8) and the corresponding initial value problem (3.20) with $q(x_0) \neq 0$. Define the set $R_{\leq}$ of points $x \in \mathbb{R}^n$ dominating x_0 ,

$$R_{\leq} := \{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\}, \quad (3.21)$$

and assume that $R_{\leq}$ is bounded. Then there exists a unique solution $\mathbf{x} : [0, \infty[\rightarrow \mathbb{R}^n$ of (3.20) with the following dominance property:

$$\mathbf{f}(\mathbf{x}(t)) \leq \mathbf{f}(\mathbf{x}(s)) \text{ and } \mathbf{f}(\mathbf{x}(t)) \neq \mathbf{f}(\mathbf{x}(s)) \text{ for all } 0 \leq s < t < \infty. \quad (3.22)$$

Proof. Since $\mathbf{q}$ is locally Lipschitzian (see Theorem 3.1), there exist a real number $T > 0$ and a unique solution $\mathbf{x} : [0, T[\rightarrow \mathbb{R}^n$ of (3.20) which, because of $\mathbf{q}(\mathbf{x}_0) \neq \mathbf{0}$ and the continuity of $\mathbf{q}$, has the property $\mathbf{q}(\mathbf{x}(t)) \neq \mathbf{0}$ for all $t \in [0, T[$. Using Theorem 3.1(i) we get for all $i \in \{1, \dots, k\}$:

$$\frac{d}{dt} f_i(\mathbf{x}(t)) = \nabla f_i(\mathbf{x}(t))^T \dot{\mathbf{x}}(t) = -\nabla f_i(\mathbf{x}(t))^T \mathbf{q}(\mathbf{x}(t)) < 0 \quad (3.23)$$

for all $t \in [0, T[$. Therefore, $f_i(\mathbf{x}(\bullet)) : [0, T[\rightarrow \mathbb{R}$ is a strictly monotonously decreasing function for each $i \in \{1, \dots, k\}$. It follows that

$$\mathbf{f}(\mathbf{x}(t)) \leq \mathbf{f}(\mathbf{x}(s)) \text{ and } \mathbf{f}(\mathbf{x}(t)) \neq \mathbf{f}(\mathbf{x}(s)) \text{ for all } 0 \leq s < t < T. \quad (3.24)$$

Now, let us assume that T is the largest real number such that $\mathbf{x} : [0, T[\rightarrow \mathbb{R}^n$ is a solution of (3.20) with the property (3.24). Since $\mathbf{x}(t) \in R_{\leq}$ for all $t \in [0, T[$ and since $R_{\leq}$ is bounded, the finiteness of T must be due to $\mathbf{q}(\mathbf{x}(T)) = \mathbf{0}$. For the same reasons, this solution $\mathbf{x}$ can be extended continuously to $\mathbf{x}(T)$ at $t = T$ with $\mathbf{q}(\mathbf{x}(T)) = \mathbf{0}$. For the following initial value problem

$$\dot{\mathbf{y}}(t) = \mathbf{q}(\mathbf{y}(t)), \quad \mathbf{y}(0) = \mathbf{x}(T) \quad (3.25)$$

we know two solutions, namely $\mathbf{y}(t) \equiv \mathbf{x}(t)$ and $\mathbf{y}(t) = \mathbf{x}(T - t)$ for all $t \in [0, T[$. This is a contradiction to the uniqueness of a solution of (3.25) which is a consequence of the local Lipschitz-property of $\mathbf{q}$. Therefore, the existence of a largest number T cannot be true, and the solution $\mathbf{x}$ of (3.20) with dominance property (3.22) is defined in $[0, \infty[$. ■

For $t \rightarrow \infty$ the curve $\mathbf{x}(t)$ solving the initial value problem (3.20) approaches a candidate point for a Pareto optimal solution of the unconstrained vector optimization problem (3.8). This property is formulated in the following theorem.

Theorem 3.3:

Consider an arbitrary starting point $\mathbf{x}_0 \in \mathbb{R}^n$ for which $R_{\leq}$ is bounded and the (unique) curve $\mathbf{x}(t)$ solving the initial value problem (3.20). Then for $t \rightarrow \infty$ the curve $\mathbf{x}(t)$ comes arbitrarily close to a point $\mathbf{x}^* \in \mathbb{R}^n$ with $\mathbf{q}(\mathbf{x}^*) = \mathbf{0}$.

Proof. Since $R_{\leq}$ is supposed to be bounded and since $\mathbf{x}(t) \in R_{\leq}$ for all $t \in [0, \infty[$, the whole curve $\mathbf{x}(t)$ is contained within a compact subset of $\mathbb{R}^n$. Therefore, any discretization $0 = t_0 < t_1 < \dots$ of the time half-line $[0, \infty[$ will yield a sequence $\{\mathbf{x}_n := \mathbf{x}(t_n)\}_{n=1}^{\infty}$ which has a subsequence $\{\hat{\mathbf{x}}_n := \mathbf{x}(\hat{t}_n)\}_{n=1}^{\infty}$

converging towards some point $\mathbf{x}^*$. Because of the continuity of $\mathbf{q}$ it follows that $\mathbf{q}(\tilde{\mathbf{x}}_n) \xrightarrow[n \rightarrow \infty]{} \mathbf{q}(\mathbf{x}^*)$.

Let us assume that $\mathbf{q}(\mathbf{x}^*) \neq \mathbf{0}$. According to Theorem 3.1 (i) this implies that

$$-\nabla f_i(\mathbf{x}^*)^T \mathbf{q}(\mathbf{x}^*) < 0 \quad \text{for all } i \in \{1, \dots, k\}. \quad (3.26)$$

Now we will prove that each time the curve $\mathbf{x}(t)$ approaches $\mathbf{x}^*$, the value of f_i (where $i \in \{1, \dots, k\}$ is arbitrarily chosen), considered via $f_i(\mathbf{x}(t))$ as a function of t , decreases at least by some minimum amount.

Since ∇f_i and $\mathbf{q}$ are continuous, there exists an ϵ -neighborhood $U_\epsilon(\mathbf{x}^*)$ of $\mathbf{x}^*$ with

$$-\nabla f_i(\mathbf{x})^T \mathbf{q}(\mathbf{x}) < -\frac{1}{2} \nabla f_i(\mathbf{x}^*)^T \mathbf{q}(\mathbf{x}^*) \quad \text{for all } \mathbf{x} \in U_\epsilon(\mathbf{x}^*). \quad (3.27)$$

Furthermore, there is a δ -neighborhood $U_\delta(\mathbf{x}^*)$ of $\mathbf{x}^*$ with

$$\|\mathbf{q}(\mathbf{x})\| < 2 \|\mathbf{q}(\mathbf{x}^*)\| \quad \text{for all } \mathbf{x} \in U_\delta(\mathbf{x}^*). \quad (3.28)$$

As there exists an $N_0 \in \mathbb{N}$ such that $\|\tilde{\mathbf{x}}_n - \mathbf{x}^*\| \leq \frac{1}{2} \min(\epsilon, \delta)$ for all $n \geq N_0$, each time interval (around some time $\tilde{t}_n$, $n \geq N_0$), during which the curve $\mathbf{x}(t)$ stays in $U_{\min(\epsilon, \delta)}(\mathbf{x}^*)$, lasts at least $t_{\text{minimum}} = \frac{\min(\epsilon, \delta)}{2\|\mathbf{q}(\mathbf{x}^*)\|}$. [If $\mathbf{x}(t)$ does not leave $U_{\min(\epsilon, \delta)}(\mathbf{x}^*)$ between $\tilde{t}_n$ and $\tilde{t}_{n+1}$, we consider $\frac{1}{2}(\tilde{t}_{n+1} - \tilde{t}_n)$ as t_{minimum} and revise (3.29) accordingly.]

Now we can estimate the decrease Δf_i of f_i during a stay of $\mathbf{x}(t)$ in $U_{\min(\epsilon, \delta)}(\mathbf{x}^*)$:

$$\begin{aligned} \Delta f_i &= \int_{t_{\text{enter}}}^{t_{\text{leave}}} \frac{d}{dt} f_i(\mathbf{x}(t)) dt = \int_{t_{\text{enter}}}^{t_{\text{leave}}} -\nabla f_i(\mathbf{x}(t))^T \mathbf{q}(\mathbf{x}(t)) dt \leq \\ &\leq -\frac{1}{2} \nabla f_i(\mathbf{x}^*)^T \mathbf{q}(\mathbf{x}^*) \cdot \frac{\min(\epsilon, \delta)}{2\|\mathbf{q}(\mathbf{x}^*)\|} \end{aligned} \quad (3.29)$$

Since according to the proof of Theorem 3.2 f_i is strictly decreasing along $\mathbf{x}(t)$, the value of $f_i(\mathbf{x}(t))$, considered as a function of t , decreases due to (3.29) below any (potential) lower bound for $t \rightarrow \infty$. This is a contradiction to the convergence of $\{\tilde{\mathbf{x}}_n\}_{n=1}^\infty$, and the assumption $\mathbf{q}(\mathbf{x}^*) \neq \mathbf{0}$ cannot be true. ■

Theorem 3.3 implies that solving the initial value problem (3.20) numerically results in a candidate for a Pareto optimal solution. A numerical treatment of (3.20) should rely on *explicit* numerical schemes, as the function $\mathbf{q}$ is not continuously differentiable. The dominance property (3.22) can be utilized for a suitable stepsize control.

Now we have shown that — for a given starting point $\mathbf{x}_0$ — the initial value problem (3.20) can be used for the computation of a *single* Pareto candidate. The application of a special stochastic perturbation to (3.20) will lead to a method for the numerical computation of a *large number of* Pareto optima.

As a preparation, the next paragraph provides some stochastic preliminaries.

3.3.2 Notions from Probability Theory

In the following we will list some stochastic notions which will be used in Paragraphs 3.3.3 and 3.3.4. For a detailed discussion we refer to standard textbooks on probability theory and stochastic processes, e.g. [ASH, 2000] or [BAUER, 1991].

On our way to introduce the Brownian motion process we start by defining a special sample space Ω : Let Ω be the set of functions $v : [0, \infty[\rightarrow \mathbb{R}^n$, $n \in \mathbb{N}$. Ω is endowed with a metric d defined as

$$d(\omega_1, \omega_2) := \sum_{m=1}^{\infty} \frac{1}{2^m} \min \left(\left(\max_{0 \leq t \leq m} \|\omega_1(t) - \omega_2(t)\| \right); 1 \right). \quad (3.30)$$

By $\mathcal{B}(\Omega)$ — the so-called Borel σ -field of Ω — we mean the smallest σ -field containing all open sets of Ω in the topology induced by the metric d .

Let further $\bar{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ be the compactification of $\mathbb{R}$ with $0 \cdot (\infty) = 0 \cdot (-\infty) = (\infty) \cdot 0 = (-\infty) \cdot 0 := 0$ and let $\mathcal{B}(\bar{\mathbb{R}})$ be the Borel σ -field of $\bar{\mathbb{R}}$ given by $B \in \mathcal{B}(\bar{\mathbb{R}}) \iff (B \cap \mathbb{R}) \in \mathcal{B}(\mathbb{R})$.

Definition 3.4: (Numerical function)

A function $g : \Omega \rightarrow \bar{\mathbb{R}}$ is called a numerical function. □

Definition 3.5: (Stochastic process)

A family $\{X_t\}$ of n -dimensional real random variables X_t , $t \geq 0$, defined on Ω is called a stochastic process. □

Definition 3.6: (Path of a stochastic process)

Let $\{X_t\}$ be a stochastic process. For each fixed $\omega \in \Omega$ the function $X_\omega : [0, \infty[\rightarrow \mathbb{R}^n$, $t \mapsto X_t(\omega)$ is called a path of $\{X_t\}$. □

Definition 3.7: (Continuous stochastic process)

Let $\{X_t\}$ be a stochastic process. $\{X_t\}$ is called continuous, if each path of $\{X_t\}$ is a continuous function. □

The functions v which the sample space Ω consists of can be made the paths of a stochastic process $\{B_t\}$ when we define $\{B_t\}$ by $B_t(\omega) := \omega(t)$ for all $t \geq 0$. Using $\{B_t\}$, for each $n \in \mathbb{N}$ the so-called Wiener measure $\mathcal{W}$ is uniquely defined by the following conditions (for a proof see e.g. [BAUER, 1991]):

- (i) The process $\{B_t\}$ starts almost surely at 0:
 $B_0(\omega) = 0 \quad \mathcal{W}\text{-almost surely.}$

- (ii) The increments of $\{B_t\}$ on disjoint intervals are independent:
For every $0 = t_0 < t_1 < \dots < t_m$, $m \in \mathbb{N}$, the random variables $B_0, B_{t_1} - B_{t_0}, \dots, B_{t_m} - B_{t_{m-1}}$ are stochastically independent.
- (iii) The increment $B_t - B_s$ is normally distributed with mean 0 and variance $t - s$:
For every $0 \leq s < t$ the random variable $B_t - B_s$ is $N(0, (t - s)\mathbf{I}_n)$ Gaussian distributed, where $\mathbf{I}_n$ denotes the n -dimensional identity matrix.

Throughout this section we will exclusively consider the probability space $(\Omega, \mathcal{B}(\Omega), \mathcal{W})$. In that probability space the Brownian motion process is defined in a very natural way.

Definition 3.8: (Brownian motion)

The stochastic process $\{B_t\}$ defined by $B_t(\omega) := \omega(t)$ for all $t \geq 0$ is called n -dimensional Brownian motion. $\square$

It should be noted that the concept of Brownian motion has been designed as a mathematical model of the random movement of particles of pollen in water (see the above properties (i) to (iii) and the continuity of the paths).

The following definition is important for utilizing specially constructed stochastic processes for the solution of optimization tasks.

Definition 3.9: (Random time)

A $[\mathcal{B}(\Omega) - \mathcal{B}(\bar{\mathbb{R}})]$ -measurable numerical function g is called a random time. $\square$

The class of random times which is most useful for our purposes is related to some stochastic process $\{X_t\}$ and indicates for each sample point ω the shortest time at which the path X_ω hits some Borel set A . The following theorem whose proof can e.g. be read in [PROTTER, 1990] states that this first-hit time is indeed a random time.

Theorem 3.4:

Let $\{X_t\}$ be a continuous stochastic process, then the function defined as

$$g : \begin{cases} \Omega & \rightarrow \bar{\mathbb{R}} \\ \omega & \mapsto \begin{cases} \inf\{t \geq 0; X_\omega(t) \in A\} & \text{if } \bigcup_{t \geq 0} (X_\omega(t) \cap A) \neq \emptyset \\ \infty & \text{else} \end{cases} \end{cases} \quad (3.31)$$

is a random time for each open or closed set $A \in \mathcal{B}(\mathbb{R}^n)$, $n \in \mathbb{N}$. $\blacksquare$

3.3.3 A Special Stochastic Differential Equation

An important class of stochastic processes is defined as the solution of some stochastic Itô-differential equation (SDE). We construct a special SDE by perturbing the deterministic initial value problem (3.20) by a Brownian motion. In the resulting SDE the function $-\mathbf{q}$, where $\mathbf{q}$ is defined in (3.14), will play the role of the (deterministic) drift part.

In order to obtain the desired properties of that SDE we have to make the following assumption concerning $\mathbf{q}$. It describes a special behavior of $\mathbf{q}$ and therefore of $\mathbf{f}$ outside a ball with radius r (for which only the existence is postulated).

Assumption (A): There exists an $\epsilon > 0$ such that

$$\mathbf{x}^T \mathbf{q}(\mathbf{x}) \geq \frac{1 + n\epsilon^2}{2} \max(1, \|\mathbf{q}(\mathbf{x})\|) \quad (3.32)$$

for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq r\}$ with some $r > 0$.

Speaking in a casual way, Assumption (A) guarantees that outside some ball with radius r each drift vector $-\mathbf{q}(\mathbf{x})$ has a component along the direction $-\mathbf{x}$ (i.e. directed towards the origin), which is sufficiently large to prevent the escape of paths (besides a set of measure zero) to infinity. Since each solution $\mathbf{x}(t)$ of an initial value problem (3.20), for which Assumption (A) is fulfilled, enters the ball after some finite time, an argument similar to the proof of Theorem 3.3 shows that there exists a point $\mathbf{x}^*$ inside that ball with $\mathbf{q}(\mathbf{x}^*) = \mathbf{0}$, i.e. a candidate point for a Pareto optimal solution.

Perturbing the initial value problem (3.20) by a Brownian motion $\{B_t\}$ results in the following SDE

$$dX_t = -\mathbf{q}(X_t) dt + \epsilon dB_t, \quad X_0 = \mathbf{x}_0, \quad (3.33)$$

with $\epsilon > 0$, $\mathbf{x}_0 \in \mathbb{R}^n$. Concerning existence, uniqueness and regularity of the solution of this SDE we can make the following statement.

Theorem 3.5:

Consider the stochastic differential equation (3.33). For all $\mathbf{x}_0 \in \mathbb{R}^n$ and for all ϵ , for which Assumption (A) is fulfilled, we obtain the following.

- (i) *There exists a unique stochastic process $\{X_t\}$ that solves (3.33).*
- (ii) *All paths of $\{X_t\}$ are continuous.*
- (iii) *$X_0 = \mathbf{x}_0$.*

Proof. Let $l \in \mathbb{N}$ and consider

$$\mathbf{q}_{\mathbf{x}_0}^{(l)} : \begin{cases} \mathbb{R}^n \rightarrow \mathbb{R}^n \\ \mathbf{x} \mapsto \begin{cases} \mathbf{q}(\mathbf{x}) & \text{if } \|\mathbf{x} - \mathbf{x}_0\| \leq l \\ \mathbf{q}\left(\frac{l(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|}\right) & \text{if } \|\mathbf{x} - \mathbf{x}_0\| > l \end{cases} \end{cases} \quad (3.34)$$

Since $\mathbf{q}$ is satisfying a local Lipschitz condition, $\mathbf{q}_{\mathbf{x}_0}^{(l)}$ fulfills for each $l \in \mathbb{N}$, $\mathbf{x}_0 \in \mathbb{R}^n$ a global Lipschitz condition with Lipschitz constant $L_{\mathbf{x}_0}^{(l)}$. Now we investigate for each $l \in \mathbb{N}$ and $\mathbf{x}_0 \in \mathbb{R}^n$ the following class of stochastic Itô-differential equations, where $\{B_t\}$ is an n -dimensional Brownian motion.

$$dX_t^{(l)} = -\mathbf{q}_{\mathbf{x}_0}^{(l)}(X_t^{(l)}) dt + \epsilon dB_t, \quad X_0 = \mathbf{x}_0, \quad (3.35)$$

with $\epsilon > 0$ such that Assumption (A) is fulfilled. Integration of (3.35) leads to the equivalent formulation

$$X_t^{(l)}(\omega) = \mathbf{x}_0 + \epsilon (B_t(\omega) - B_0(\omega)) - \int_0^t \mathbf{q}_{\mathbf{x}_0}^{(l)}(X_\tau^{(l)}(\omega)) d\tau \quad (3.36)$$

for each $\omega \in \Omega$. Define for each $\bar{t} \in \mathbb{R}^+$ and each $\omega \in \Omega$ the operator $T_{\omega, \bar{t}} : C([0, \bar{t}], \mathbb{R}^n) \rightarrow C([0, \bar{t}], \mathbb{R}^n)$,

$$(T_{\omega, \bar{t}} g)(t) = \mathbf{x}_0 + \epsilon (B_t(\omega) - B_0(\omega)) - \int_0^t \mathbf{q}_{\mathbf{x}_0}^{(l)}(g(\tau)) d\tau, \quad 0 \leq t \leq \bar{t}, \quad (3.37)$$

where $C([0, \bar{t}], \mathbb{R}^n)$ represents the set of all continuous functions $g : [0, \bar{t}] \rightarrow \mathbb{R}^n$. Using the Banach space $(C([0, \bar{t}], \mathbb{R}^n), \|\cdot\|_C)$ with the norm

$$\|\cdot\|_C : \begin{cases} C([0, \bar{t}], \mathbb{R}^n) \rightarrow \mathbb{R} \\ g \mapsto \max_{0 \leq t \leq \bar{t}} \left(\exp(-2L_{\mathbf{x}_0}^{(l)}) \|g(t)\| \right) \end{cases}, \quad (3.38)$$

we compute for $g, h \in C([0, \bar{t}], \mathbb{R}^n)$:

$$\begin{aligned} \|(T_{\omega, \bar{t}} g)(t) - (T_{\omega, \bar{t}} h)(t)\| &= \left\| \int_0^t \mathbf{q}_{\mathbf{x}_0}^{(l)}(g(\tau)) d\tau - \int_0^t \mathbf{q}_{\mathbf{x}_0}^{(l)}(h(\tau)) d\tau \right\| \leq \\ &\leq \int_0^t \|\mathbf{q}_{\mathbf{x}_0}^{(l)}(g(\tau)) - \mathbf{q}_{\mathbf{x}_0}^{(l)}(h(\tau))\| d\tau \leq \\ &\leq L_{\mathbf{x}_0}^{(l)} \int_0^t \|g(\tau) - h(\tau)\| d\tau = \\ &= L_{\mathbf{x}_0}^{(l)} \int_0^t \|g(\tau) - h(\tau)\| \exp(-2L_{\mathbf{x}_0}^{(l)}\tau) \exp(2L_{\mathbf{x}_0}^{(l)}\tau) d\tau \leq \\ &\leq L_{\mathbf{x}_0}^{(l)} \int_0^t \max_{0 \leq s \leq \bar{t}} (\|g(s) - h(s)\| \exp(-2L_{\mathbf{x}_0}^{(l)}s)) \exp(2L_{\mathbf{x}_0}^{(l)}\tau) d\tau = \\ &= L_{\mathbf{x}_0}^{(l)} \|g - h\|_C \int_0^t \exp(2L_{\mathbf{x}_0}^{(l)}\tau) d\tau \leq \\ &\leq \frac{1}{2} \|g - h\|_C \exp(2L_{\mathbf{x}_0}^{(l)}t) \end{aligned} \quad (3.39)$$

for all $t \in [0, \bar{t}]$. Hence we can write

$$\|T_{\omega, \bar{t}} g - T_{\omega, \bar{t}} h\|_C \leq \frac{1}{2} \|g - h\|_C. \quad (3.40)$$

Using Banach's fixed point theorem we obtain a unique function g with $T_{\omega, \bar{t}} g = g$, which is the unique solution of (3.36) for $t \in [0, \bar{t}]$ and $\omega \in \Omega$. Therefore, the fixed points of $T_{\omega, \bar{t}}$ depending on $\omega \in \Omega$ define the paths of $\{X_t^{(l)}\}$ for $t \in [0, \bar{t}]$.

Now we have to consider how to come from the solution $\{X_t^{(l)}\}$ of (3.36) to the solution $\{X_t\}$ of (3.33). For that purpose, we define the following random time

$$s_l : \begin{cases} \Omega & \rightarrow \bar{\mathbb{R}}^n \\ \omega & \mapsto \begin{cases} \inf\{t \geq 0; \|X_t^{(l)}(\omega) - \mathbf{x}_0\| > l\} & \text{if } \{t \geq 0; \|X_t^{(l)}(\omega) - \mathbf{x}_0\| > l\} \neq \emptyset \\ \infty & \text{else} \end{cases} \end{cases} \quad (3.41)$$

Using the fact that $X_t^{(l)}(\omega) = X_t(\omega)$ for all $0 \leq t \leq s_l(\omega)$ we have only to show that

$$\lim_{l \rightarrow \infty} s_l(\omega) = \infty \quad \text{for each } \omega \in \Omega. \quad (3.42)$$

Assume the existence of $\tilde{\omega} \in \Omega$ with

$$\lim_{l \rightarrow \infty} s_l(\tilde{\omega}) = \xi < \infty \quad (3.43)$$

and consider the function

$$u : \begin{cases} [0, \xi[& \rightarrow \mathbb{R} \\ t & \mapsto \frac{1}{2} \|X_t(\tilde{\omega}) + \epsilon(B_\xi(\tilde{\omega}) - B_t(\tilde{\omega}))\|^2 \end{cases} \quad (3.44)$$

With Assumption (A), we obtain the existence of a real number $\rho \in [0, \xi[$ such that

$$\dot{u}(t) \leq -\frac{1 + n\epsilon^2}{4} \quad \text{for all } t \in]\rho, \xi[\text{ and all } t \text{ with } \|X_t(\tilde{\omega})\| > r. \quad (3.45)$$

This is a contradiction to the existence of $\tilde{\omega}$ with $\lim_{t \rightarrow \xi} \|X_t(\tilde{\omega})\| = \infty$, which is implied by (3.43). ■

Let $\bar{\mathbf{x}} \in \mathbb{R}^n$ and consider the ball centered at $\bar{\mathbf{x}}$ with radius p , $S(\bar{\mathbf{x}}, p) := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \bar{\mathbf{x}}\| \leq p\}$. For the investigation of the relations between the solution $\{X_t\}$ of the stochastic differential equation (3.33) and the vector optimization problem (3.8) we need the random time $s_{\bar{\mathbf{x}}, p}$ for the first hit of $S(\bar{\mathbf{x}}, p)$. In compliance with (3.31) $s_{\bar{\mathbf{x}}, p}$ is defined as

$$s_{\bar{\mathbf{x}}, p} : \begin{cases} \Omega & \rightarrow \bar{\mathbb{R}} \\ \omega & \mapsto \begin{cases} \inf\{t \geq 0; \|X_\omega(t) - \bar{\mathbf{x}}\| \leq p\} & \text{if } \{t \geq 0; \|X_\omega(t) - \bar{\mathbf{x}}\| \leq p\} \neq \emptyset \\ \infty & \text{else} \end{cases} \end{cases} \quad (3.46)$$

Now we are able to formulate a theorem stating that for each Pareto optimal solution $\bar{\mathbf{x}}$ of the unconstrained VOP (3.8) $\mathcal{W}$ -almost all paths of $\{X_t\}$ hit any ball $S(\bar{\mathbf{x}}, p)$ centered at $\bar{\mathbf{x}}$ (for an arbitrarily chosen radius $p > 0$) after a **finite** time for all starting points $\mathbf{x}_0 \in \mathbb{R}^n$. Moreover, the **expectation** of the random time $s_{\bar{\mathbf{x}}, p}$ is **finite**.

This theorem is a direct application of some important results from Lyapunov's stability calculus for stochastic differential equations (see [HASMINSKIJ, 1980]).

Theorem 3.6:

Consider the stochastic differential equation (3.33) with $\epsilon > 0$ such that Assumption (A) is fulfilled. Then one can state the following for each starting point $\mathbf{x}_0 \in \mathbb{R}^n$ and for each Pareto optimal solution $\bar{\mathbf{x}} \in \mathbb{R}^n$ (and for each radius $p > 0$):

- (i) $\mathcal{W}(\{\omega \in \Omega \mid s_{\bar{\mathbf{x}}, p}(\omega) < \infty\}) = 1$.
- (ii) $E(s_{\bar{\mathbf{x}}, p}) < \infty$, where E denotes the expectation.
- (iii) *The stochastic process $\{X_t\}$ defined in Theorem 3.5 converges in distribution to a random variable $X : \Omega \rightarrow \mathbb{R}^n$ with $E(\mathbf{q}(X)) = 0$.*

■

Claim (iii) of the above theorem states that the stochastic process $\{X_t\}$ asymptotically approaches a random variable X . The first moment (= expectation) of $\mathbf{q}(X)$ is identical to the first moment of $\mathbf{q}(Y)$, where Y is assumed to be a random variable concentrated (with measure 1) at candidate points for Pareto optimality, i.e. at points $\mathbf{x}^*$ with $\mathbf{q}(\mathbf{x}^*) = 0$.

3.3.4 A Stochastic Algorithm for Vector Optimization

Theorem 3.6 suggests the following method for solving the vector optimization problem (3.8):

Approximate numerically an arbitrary path X_ω of the stochastic process $\{X_t\}$ which solves the SDE (3.33). This path comes arbitrarily close to any Pareto optimal solution of (3.8).

With regard to the practical application of that method, two questions are still to be answered:

- (a) What is the right choice of the parameter ϵ ?
- (b) What is an adequate numerical scheme for the approximation of X_ω ?

In order to answer question (a), we consider the SDE (3.33). The parameter ϵ is a measure for the balance between the curve of dominated points

$$X_t^{x_0} = x_0 - \int_0^t \mathbf{q}(X_\tau^{x_0}) d\tau, \quad (3.47)$$

and a random search utilizing realizations of Gauss distributed random vectors with increasing variance

$$X_t^{x_0}(\omega) = x_0 + B_t(\omega) - B_0(\omega). \quad (3.48)$$

If we choose ϵ for a fixed ω such that (3.47) dominates, then the chosen path of $\{X_t^{x_0}\}$ spends a long time close to any (local) Pareto optimal solution of (3.8). If we choose ϵ such that the random search (3.48) dominates, then the Pareto optimal solutions of (3.8) play no significant role along this path of $\{X_t^{x_0}\}$.

The optimal balance between (3.47) and (3.48) and, therefore, the optimal choice of ϵ depends on the objective function $\mathbf{f}$ and the scale used. If one observes during the numerical computations of a path of $\{X_t^{x_0}\}$ that this path spends a very long time close to any (local) Pareto optimal solution of (3.8), then ϵ is too small. If on the other hand the (local) Pareto optimal solutions of (3.8) play no significant role along this path, then ϵ is too large.

In response to question (b), i.e. for the numerical computation of a path of $\{X_t\}$, we consider the following iteration scheme. It results from the Euler method, a standard approach in the numerical analysis of ordinary differential equations. For a fixed steplength σ set

$$\mathbf{x}_{j+1}^1 := \mathbf{x}_j - \sigma \mathbf{q}(\mathbf{x}_j) - \epsilon \mathbf{n}_3 \left(\frac{\sigma}{2} \right)^{\frac{1}{2}}, \quad (3.49)$$

$$\mathbf{x} \left(\frac{\sigma}{2} \right) := \mathbf{x}_j - \left(\frac{\sigma}{2} \right) \mathbf{q}(\mathbf{x}_j) - \epsilon \mathbf{n}_1 \left(\frac{\sigma}{2} \right)^{\frac{1}{2}}, \quad (3.50)$$

$$\mathbf{x}_{j+1}^2 := \mathbf{x} \left(\frac{\sigma}{2} \right) - \left(\frac{\sigma}{2} \right) \mathbf{q} \left(\mathbf{x} \left(\frac{\sigma}{2} \right) \right) - \epsilon \mathbf{n}_2 \left(\frac{\sigma}{2} \right)^{\frac{1}{2}}, \quad (3.51)$$

where $\mathbf{n}_1$ and $\mathbf{n}_2$ are realizations of independent $N(0, \mathbf{I}_n)$ normally distributed random vectors, which are computed by pseudo-random numbers, and $\mathbf{n}_3 = \mathbf{n}_1 + \mathbf{n}_2$. The scheme calculates the next iteration point $\mathbf{x}_{j+1}$ of the discretized path X_ω in two ways. $\mathbf{x}_{j+1}^1$ results from performing one Euler step with the steplength σ . Performing successively two Euler steps with steplength $(\frac{\sigma}{2})$ yields the point $\mathbf{x}_{j+1}^2$. Relating the random vectors $\mathbf{n}_3$ and $\{\mathbf{n}_1, \mathbf{n}_2\}$ by $\mathbf{n}_3 = \mathbf{n}_1 + \mathbf{n}_2$ ensures that in both calculations the same path X_ω is approximated.

The difference between $\mathbf{x}_{j+1}^1$ and $\mathbf{x}_{j+1}^2$ is used for controlling the steplength σ . We choose a tolerated error limit $\delta > 0$ and take $\mathbf{x}_{j+1} = \mathbf{x}_{j+1}^2$ if $\|\mathbf{x}_{j+1}^1 - \mathbf{x}_{j+1}^2\| \leq \delta$. Otherwise, the steps (3.49) and (3.51) have to be repeated with $\frac{\sigma}{2}$ instead of σ .

The Connection with Scalar-Valued Optimization

A necessary condition for Pareto optimality given by Kuhn and Tucker builds the bridge between vector optimization and scalar-valued optimization: On the assumption that the constraints meet a certain constraint qualification, necessarily for a Pareto optimal point $\mathbf{x}^*$ there exists a convex combination of the objectives $g_{\alpha}(\mathbf{x}) := \sum_{i=1}^k \alpha_i f_i(\mathbf{x})$, so that $\mathbf{x}^*$ is a Karush-Kuhn-Tucker point of the scalar-valued function g_{α} .

In the following chapter this connection between vector and scalar-valued optimization shall be enlarged. We will briefly compile the required differential-topological terms in Section 4.2. Section 4.3 demonstrates that the weight vector α has a geometrical meaning in the objective space $\mathbb{R}^k$: Let R denote the feasible set and $\mathbf{f}(R)$ its image under the mapping $\mathbf{f}$ (the vector-valued objective function). Then α is a normal vector to the tangent plane of the border $\partial \mathbf{f}(R)$ of $\mathbf{f}(R)$.

Subsequently, in Section 4.4 we will derive a relation between the curvature of $\partial \mathbf{f}(R)$ in the point $\mathbf{f}(\mathbf{x}^*)$ and the type of the stationary point $\mathbf{x}^*$ (i.e. minimum or saddle point) of g_{α} .

4.1 The Karush-Kuhn-Tucker(KKT) Condition for Pareto Optimality

Simultaneously to their optimality conditions for scalar-valued optimization problems Kuhn and Tucker [KUHN & TUCKER, 1951] put forward a necessary condition for Pareto optimality in problems of vector optimization. This condition presupposes that the feasible set R is given in the form of equality and inequality constraints. The present chapter therefore deals with the following vector optimization problem:

Definition 4.1: (Vector optimization with equality and inequality constraints)

Find Pareto optimal points of the objective function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^k$, where the

feasible set $R \subseteq \mathbb{R}^n$ is given in the form of

$$R := \left\{ \mathbf{x} \in \mathbb{R}^n \mid \begin{array}{l} h_i(\mathbf{x}) = 0 \quad \forall i = 1, \dots, m \\ h_j(\mathbf{x}) \leq 0 \quad \forall j = m+1, \dots, m+q \end{array} \right\}. \quad (4.1)$$

The functions $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ and $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are assumed to be continuously differentiable. [Beginning from Section 4.4, this assumption will be tightened to twice continuous differentiability.] $\square$

The theorem of Kuhn and Tucker says (cf. [GÖPFERT & NEHSE, 1990]):

Theorem 4.1: (Necessary condition for Pareto optimality [KUHN & TUCKER, 1951])

Consider the vector optimization problem 4.1 and a point $\mathbf{x}^*$ where the following constraint qualification is fulfilled: The vectors $\{\nabla h_i(\mathbf{x}^*) \mid i \text{ is an index of an active constraint}\}$ are linearly independent.

If $\mathbf{x}^*$ is Pareto optimal, then there exist vectors

$$\boldsymbol{\alpha} \in \mathbb{R}^k \text{ with } \alpha_i \geq 0 \text{ and } \sum_{i=1}^k \alpha_i = 1 \quad (4.2)$$

$$\boldsymbol{\lambda} \in \mathbb{R}^{m+q} \quad (4.3)$$

such that:

$$\sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m+q} \lambda_j \nabla h_j(\mathbf{x}^*) = \mathbf{0} \quad (4.4)$$

$$h_i(\mathbf{x}^*) = 0, \quad i = 1, \dots, m \quad (4.5)$$

$$\lambda_j \geq 0, \quad h_j(\mathbf{x}^*) \leq 0, \quad \lambda_j \cdot h_j(\mathbf{x}^*) = 0, \quad j = m+1, \dots, m+q \quad (4.6)$$

■

We introduce the scalar-valued function

$$g_{\boldsymbol{\alpha}}(\mathbf{x}) := \sum_{i=1}^k \alpha_i f_i(\mathbf{x}) \quad (4.7)$$

and note that $\sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}) = \nabla g_{\boldsymbol{\alpha}}(\mathbf{x})$. Obviously, the equations (4.4) to (4.6) are equivalent to the claim that $\mathbf{x}^*$ is a Karush-Kuhn-Tucker¹ point of the corresponding scalar-valued optimization problem with the objective function $g_{\boldsymbol{\alpha}}$.

¹ The classical Karush-Kuhn-Tucker conditions of scalar-valued optimization were given by Karush [KARUSH, 1939] and Kuhn & Tucker [KUHN & TUCKER, 1951]. For a thorough discussion of optimality conditions see [JAHN, 1999].

Due to (4.2) and (4.7) g_α constitutes a convex linear combination of the individual objective functions f_i , where each coefficient α_i indicates the relative weight, with which the individual objective f_i is part of the linear combination g_α . In a certain way the weighting method (see Section 3.2) is based on the result of Kuhn and Tucker, by looking for minimizers – i.e. a special form of stationary points – of convex combinations g_α .

In general, however, that approach does not yield the complete Pareto optimal set, because the second-order conditions which are necessary for a point $\mathbf{x}^*$ to be a local minimizer of the scalar-valued function g_α are **not** necessary for $\mathbf{x}^*$ to be a Pareto optimal point of the vector optimizing problem 4.1. The missing (necessary) second-order optimality condition theoretically distinguishes multi-objective optimization from scalar-valued optimization and can be considered the price one has to pay for the attenuation of the ordering concept (partial order in the vector-valued objective space versus total order of the scalar-valued objective space). Because of the missing second-order condition, in principle saddle points of a convex combination g_α can also be Pareto optimal. The important role which, as a matter of fact, saddle-points (of g_α) play in the Pareto optimal set shall be discussed in Section 4.4 on the base of differential-topological arguments.

4.2 Differential-Topological Notations

The following section provides terms and notations of differential topology which will be required for a further analysis of the vector optimization problem. The notation follows largely the textbooks [FORSTER, 1984], [JÄNICH, 1992] and [CARMO, 1976]. The compilation does not claim to be complete and uses a rather casual language in its definitions, omitting – on account of brevity and better legibility – several technical details.

(a) [Chart, change of charts, atlas, differentiable manifold]

Let M be a topological space. A homeomorphism $h : U \rightarrow T$ of an open subset $U \subseteq M$ (chart area) upon an open subset $T \subseteq \mathbb{R}^l$ is called an *l-dimensional chart* for M . By stating explicitly the chart area, one can also write (U, h) .

If (U, h) and (V, k) are two *l-dimensional charts* for M , the homeomorphism [(C^r)-diffeomorphism] $k \circ (h^{-1}|_{h(U \cap V)})$ of $h(U \cap V) \subseteq \mathbb{R}^l$ upon $k(U \cap V) \subseteq \mathbb{R}^l$ is called the [(C^r)-differentiable] *change of charts* of h upon k (see Figure 4.1).

A set of *l-dimensional charts*, the areas of which cover all M and the changes of which are all differentiable, is called an *l-dimensional differentiable atlas* for M . If the topological space M is supplied with a maximum *l-dimensional differentiable atlas* [this means, adding a further chart would destroy the property that all changes of charts are differentiable], M is called an *l-dimensional differentiable manifold*.

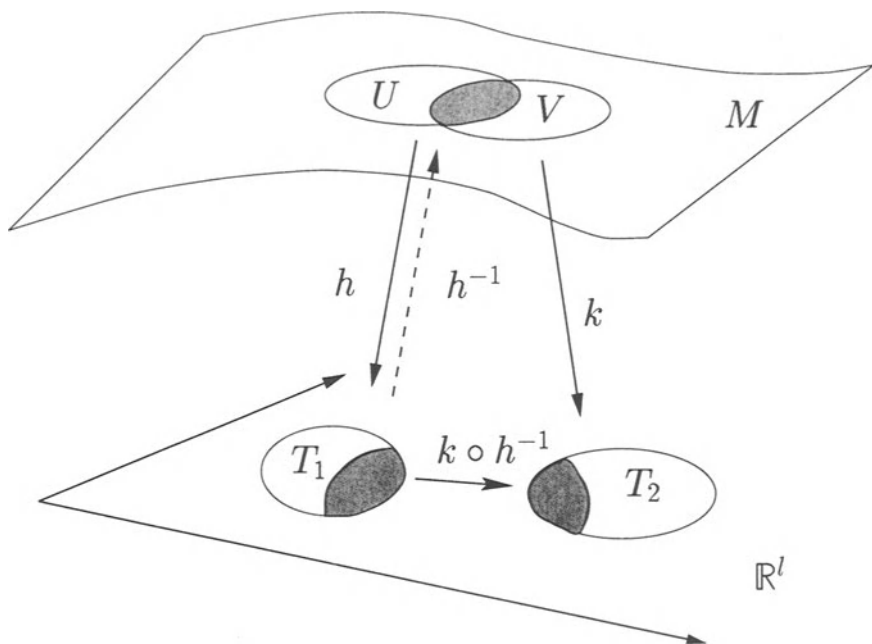


Figure 4.1: The change of charts $k \circ h^{-1}$ as a differentiable mapping from $\mathbb{R}^l$ to $\mathbb{R}^l$.

(b) [Submanifold of the $\mathbb{R}^n$, chart parameter]

A highly important class of differentiable manifolds are the zero manifolds of systems of non-linear equations. A subset $M \subset \mathbb{R}^n$ is an l -dimensional (C^r) -differentiable manifold and is called an l -dimensional (C^r) -differentiable submanifold of the $\mathbb{R}^n$, if for each point $\mathbf{p} \in M$ there is an open neighborhood $U \subset \mathbb{R}^n$ and an $(r$ -fold) continuously differentiable function $\mathbf{F} : U \rightarrow \mathbb{R}^{n-l}$, such that:

- $M \cap U = \{\mathbf{x} \in U \mid \mathbf{F}(\mathbf{x}) = \mathbf{0}\}$
- $\text{rank } \mathbf{F}'(\mathbf{p}) = n - l$, where $\mathbf{F}'$ is the Jacobian matrix of $\mathbf{F}$.

Since zero manifolds play a central role in this paper, the manifolds examined here are in most cases² submanifolds of the $\mathbb{R}^n$ in the above sense. When examining l -dimensional submanifolds of the $\mathbb{R}^n$ we will call the inversion of the chart mapping defined under (a), i.e. a homeomorphism $\varphi : T \rightarrow V \subset M$ [T open in $\mathbb{R}^l$ and V open in M] *local parametrization*. Such a mapping φ we will also denote (nota bene: within the context of

² Exceptions are limited to some special cases in the present Chapter 4, in which the function $\mathbf{F}$, which defines the zero manifold, is not indicated.

l -dimensional submanifolds of the $\mathbb{R}^n$) as a *chart* of M , its arguments as local parameters or *chart parameters* (see [FORSTER, 1984]).

(c) [*Tangent space (geometrical definition)*]

A continuous mapping $f: M \rightarrow N$ between differentiable manifolds is called differentiable at $p \in M$, if it is with respect to charts. Let M be an l -dimensional differentiable manifold and denote by $\Lambda_p(M)$ the set of those differentiable curves in M which pass through p at $t=0$, i.e. $\Lambda_p(M) := \{\beta: (-\epsilon, \epsilon) \rightarrow M \mid \beta \text{ differentiable, } \epsilon > 0 \text{ and } \beta(0) = p\}$. Two such curves $\beta, \gamma \in \Lambda_p(M)$ we will call *tangentially equivalent*, if for one (then every) chart (U, h) around p we have $(h \circ \beta)'(0) = (h \circ \gamma)'(0) \in \mathbb{R}^l$, i.e. if the velocity vectors (brought to the $\mathbb{R}^l$ by a chart) coincide in the point p . The tangential equivalence defines an equivalence relation. The equivalence classes generated thereby are called the (geometrically defined) *tangent vectors* of M in p , the vector space generated by the tangent vectors is called *tangent space* to M in p , briefly $T_p M$.

Now the tangent space in the special case of submanifolds of the $\mathbb{R}^n$ shall be examined more thoroughly. Let M be an l -dimensional differentiable submanifold of the $\mathbb{R}^n$, defined as the zero manifold of the function $\mathbf{F}: \mathbb{R}^n \rightarrow \mathbb{R}^{n-l}$. Then every tangent vector to M in $\mathbf{p}$ can be characterized by the velocity vector $\mathbf{v} \in \mathbb{R}^n$ of a curve γ on M (representing the respective equivalence class), which is considered a curve in the $\mathbb{R}^n$ (by means of the embedding of M in $\mathbb{R}^n$): $\gamma: (-\epsilon, \epsilon) \rightarrow M \subset \mathbb{R}^n$ with $\gamma'(0) = \mathbf{v}$. The tangent space $T_p M$ thus generated is often also denoted by *tangent plane* to M in the point $\mathbf{p}$. It has the following properties, of which we are going to make extensive use in the further course of this book:

- $T_p M$ is an l -dimensional vector subspace of the $\mathbb{R}^n$. It can be imagined as a local approximation of the submanifold M by a linear space.
- Let $\varphi: T \rightarrow V \subset M$ [T open in $\mathbb{R}^l$, V an open neighborhood of $\mathbf{p}$ in M] be a chart of M [according to the chart concept for submanifolds], let $\mathbf{t} \equiv (t_1, \dots, t_l)$ be the vector of the chart parameters and let the chart parameter point $\mathbf{t}_0$ denote the inverse image of $\mathbf{p}$, i.e. $\varphi(\mathbf{t}_0) = \mathbf{p}$. Then the vectors $\frac{\partial \varphi}{\partial t_1}(\mathbf{t}_0), \dots, \frac{\partial \varphi}{\partial t_l}(\mathbf{t}_0)$ constitute a basis of $T_p M$.
- Let $F_i, i = 1, \dots, n-l$ be the components of the function $\mathbf{F}$ which defines the zero manifold M . Then $T_p M$ is the orthogonal complement of the subspace spanned by the gradients $\nabla F_i(\mathbf{p})$, i.e. $T_p M = \text{span}\{\nabla F_1(\mathbf{p}), \dots, \nabla F_{n-l}(\mathbf{p})\}^\perp$. For the case of a twodimensional submanifold of the $\mathbb{R}^3$ this relation is illustrated in Figure 4.2.

(d) [*Bordered manifold*]

Manifolds (in their general definition, see (a)) are characterized by the $\mathbb{R}^l$ as a local model. Bordered manifolds are a generalization in the sense that also the closed half-space (e.g. $\mathbb{R}^{l+}_+ := \{\mathbf{x} \in \mathbb{R}^l \mid x_1 \geq 0\}$) is permitted as a

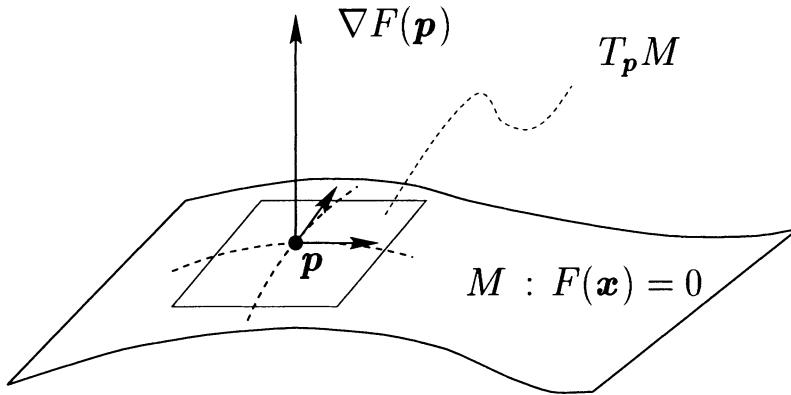


Figure 4.2 The gradient $\nabla F(\mathbf{p})$ of a function $F: \mathbb{R}^3 \rightarrow \mathbb{R}$ is orthogonal to the tangent plane $T_p M$ of the submanifold M defined as $M := \{\mathbf{x} \in \mathbb{R}^3 \mid F(\mathbf{x}) = 0\}$.

local model. Accordingly one understands by a bordered l -dimensional differentiable chart for the bordered manifold M a homeomorphism h of an open subset $U \subseteq M$ onto a set T which is an open subset of $\mathbb{R}^{l,+}$ or of $\mathbb{R}^l$. By defining differentiability also for functions defined on the half-space $\mathbb{R}^{l,+}$ — namely existence of a differentiable continuation in $\mathbb{R}^l$ (or an open subset of it) —, it is possible to describe diffeomorphic changes of charts also for bordered charts. Thus, one arrives at the definition of a *bordered l -dimensional differentiable manifold* M as a topological space M equipped with a maximum bordered, l -dimensional and differentiable atlas. A point $p \in M$ is called *border point* of M , if it is mapped by one (then every) chart onto a border point of the $\mathbb{R}^{l,+}$. The set ∂M of border points is called the *border* of the bordered manifold M and constitutes an ordinary (non-bordered) $(l-1)$ -dimensional manifold.

The tangent space $T_p M$ is well-defined as a full vector space also for bordered manifolds and, there, also for border points $p \in \partial M$ [one has to revert, however, to the algebraic definition of a tangent vector, which in ordinary manifolds is equivalent to the present geometric definition]. For border points $p \in \partial M$ one can define additionally the half-spaces $T_p^\pm M := (dh_p)^{-1}(\mathbb{R}^{l,\pm})$, where dh_p denotes the differential, i.e. the linear approximation, of the chart mapping h in the point p . $T_p^+ M \setminus T_p \partial M$ is called the *inward directed tangent space* and accordingly $T_p^- M \setminus T_p \partial M$ the *outward directed tangent space*. A vector $\mathbf{v} \in T_p M$ is directed outwards exactly in those cases, when with respect to one (then every) chart the first component v_1 of $\mathbf{v}$ is smaller than zero [nota bene: the chart maps a neighborhood $U(p) \in M$ into the $\mathbb{R}^{l,+}$].

Let us, moreover, note that the *compact set with a smooth border* is an important special case of a bordered manifold. It is defined as a special subset of the $\mathbb{R}^l$, namely as the solution manifold of a system of inequalities.

(e) [Curvatures of hypersurfaces]

For Section 4.4 we still require some terms to describe the curvature behavior of $(n - 1)$ -dimensional submanifolds of the $\mathbb{R}^n$, so-called hypersurfaces. In order to examine the (local) curvature of hypersurfaces in the embedding space, one assesses curvatures of surface curves (i.e. curves the image sets of which lie in the hypersurface).

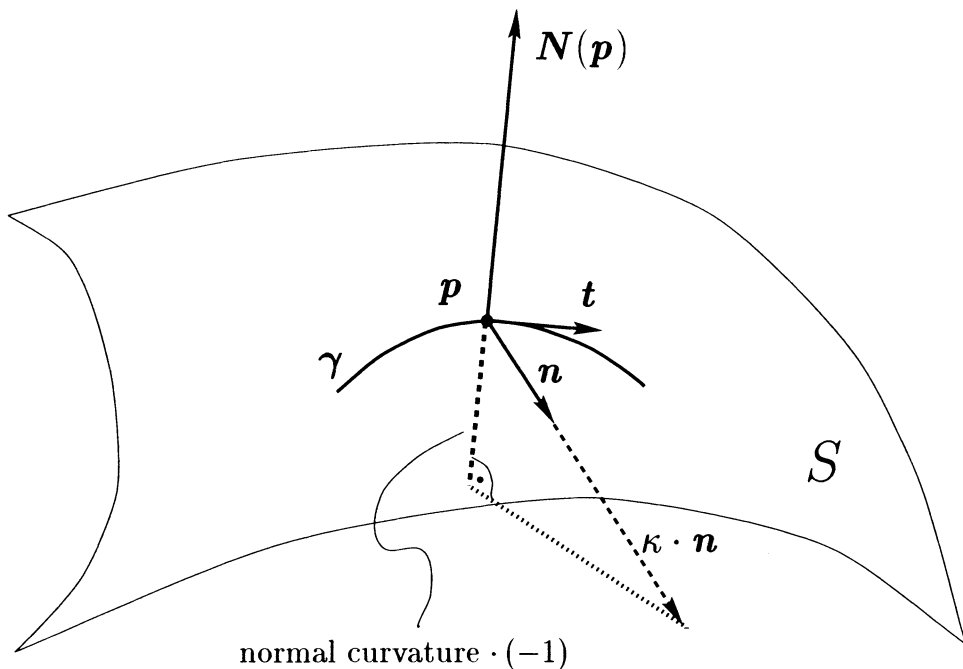


Figure 4.3 Normal curvature of a curve γ in S . Since in this example the angle between $N(p)$ and the curve normal n is greater than $\pi/2$, the normal curvature is negative. All curves in S (passing through p) with the same tangent vector t have the same normal curvature.

Let S be a hypersurface, p a point of S and N a differentiable field (defined at least in a neighborhood of p) of unit normal vectors on S [i.e. $N(p)$ is a vector orthogonal to the tangent space $T_p S$ (considered as an $(n - 1)$ -dimensional vector subspace of the $\mathbb{R}^n$)]. Let furthermore γ be a curve in S passing through p with the curvature κ (in the point p) and the normal vector n . Then the projection of the curvature vector $\kappa \cdot n$ onto the surface normal $N(p)$ is called the *normal curvature* of γ (regarded as a surface curve in S) in p . Figure 4.3 illustrates that definition. Tangentially equivalent curves have the same normal curvature, i.e. the normal curvature can be considered a function of the tangent vector.

In order to calculate the normal curvature of a given tangent vector out of $T_p S$, we have to define the Weingarten-mapping. For that purpose, we first

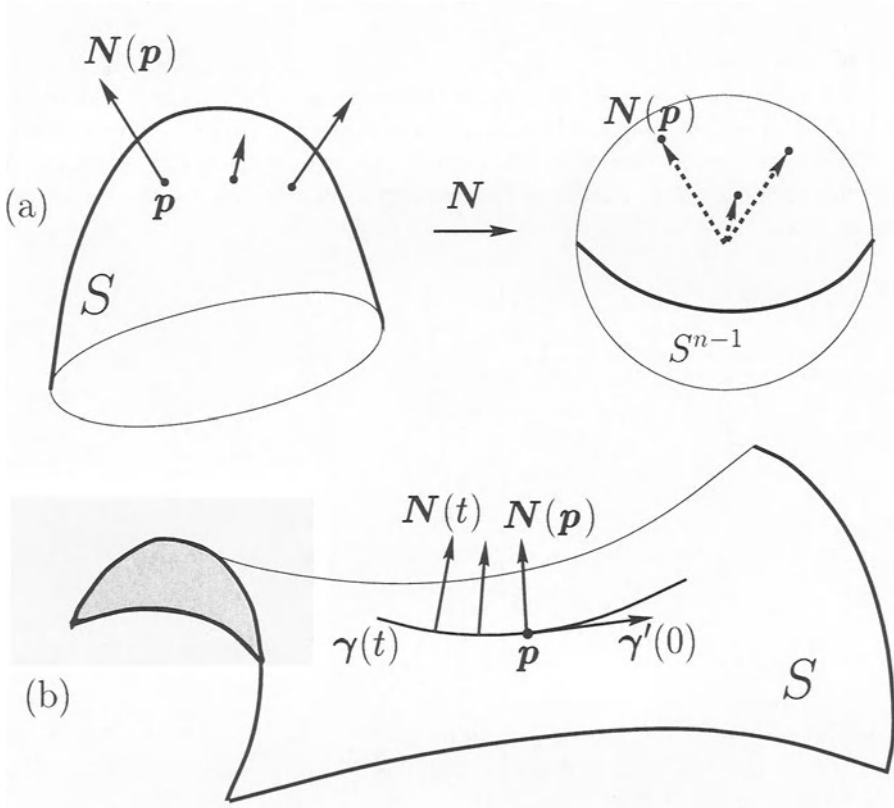


Figure 4.4 (a) The normal field N as a mapping from the hypersurface S into the $(n-1)$ -unit-sphere S^{n-1} . (b) The derivative of the parametrized curve $N(t) = N \circ \gamma(t)$ measures how N pulls away from $N(p)$ in a neighborhood of p .

note that the normal field N is a mapping from the hypersurface S into the $(n-1)$ -unit-sphere S^{n-1} (see Figure 4.4(a)) and consider the differential dN_p . By a differential we understand the local linear approximation of a differentiable mapping between manifolds³. The differential dN_p of N at $p \in S$ is a linear mapping from $T_p S$ to $T_{N(p)} S^{n-1}$. Since $T_p S$ and $T_{N(p)} S^{n-1}$ are parallel hyperplanes in the $\mathbb{R}^n$, they can be identified, and dN_p can be looked upon as a linear mapping on $T_p S$.

The linear map $dN_p : T_p S \rightarrow T_p S$ operates as follows (see Figure 4.4(b)). For each parametrized curve $\gamma(t)$ in S with $\gamma(0) = p$ we consider the parametrized curve $N \circ \gamma(t) = N(t)$ in the $(n-1)$ -sphere S^{n-1} . This amounts to restricting the normal vector N to the curve $\gamma(t)$. The tangent vector $N'(0) = dN_p(\gamma'(0))$ is a vector in $T_p S$ [via the above-mentioned identification of tangent spaces]. It measures the rate of change of the normal vector N , restricted to the curve $\gamma(t)$, at $t = 0$. Thus, dN_p measures

³ Formally the differential can be defined by means of the curve transport induced by the mapping.

how $\mathbf{N}$ ‘pulls away’ from $\mathbf{N}(\mathbf{p})$ in a neighborhood of $\mathbf{p}$. In the case of curves this measure is given by a number, the curvature. In the case of surfaces this measure is characterized by a linear map (cf. [CARMO, 1976]).

The negative differential $-d\mathbf{N}_{\mathbf{p}}$ is called Weingarten-mapping. Since $-d\mathbf{N}_{\mathbf{p}}$ is a self-adjoint linear mapping, $II_{\mathbf{p}} : T_{\mathbf{p}}S \rightarrow \mathbb{R}$, $\mathbf{v} \mapsto -\langle d\mathbf{N}_{\mathbf{p}}(\mathbf{v}), \mathbf{v} \rangle$ [with $\langle \cdot, \cdot \rangle$ as denomination of the scalar product] defines a quadratic form, the so-called *second fundamental form* of S in $\mathbf{p}$. It opens up the possibility of calculating the normal curvature: The value of the second fundamental form for a tangent vector $\mathbf{v} \in T_{\mathbf{p}}S$ is equal to the normal curvature associated with $\mathbf{v}$.

Minimum and maximum of the fundamental form – restricted to tangent vectors of the length 1 – are given by the smallest and the largest eigenvalue of the associated self-adjoint linear mapping $-d\mathbf{N}_{\mathbf{p}}$. Because of their importance for the theory of curvatures the eigenvalues of the Weingarten-mapping are called *principal curvatures* of S in $\mathbf{p}$, the corresponding eigenvectors are called principal curvature vectors (or principal directions).

4.3 The Geometrical Meaning of the Weight Vector

According to Theorem 4.1 a Pareto optimal point $\mathbf{x}^*$ is a Karush-Kuhn-Tucker-point for a convex combination g_{α} . The stationarity of g_{α} in the point $\mathbf{x}^*$ shall be used below to show that the weight vector $\alpha \in \mathbb{R}_+^k$ associated with g_{α} contains an important information about the local geometry of the efficient set.

One gets a plausible indication for the geometrical meaning of α by looking at the following special case, which is sketched in Figure 4.5. Let $k = 2$ — i.e. we examine a bicriterial vector optimization problem — and let $\mathbf{x}^*$ be a Pareto optimal point of the function $\mathbf{f}$ and a global minimizer of the scalar-valued function g_{α} with $g_{\alpha}(\mathbf{x}) = \alpha^T \cdot \mathbf{f}(\mathbf{x})$ and $\alpha \in \mathbb{R}_+^2$. Since $\mathbf{x}^*$ is a minimizer of g_{α} , there is no other point $\bar{\mathbf{x}} \in R$ with the property $(\bullet)$: $g_{\alpha}(\bar{\mathbf{x}}) < g_{\alpha}(\mathbf{x}^*) =: c$. The set of all points $\mathbf{y} \in \mathbb{R}^2$ in the objective space for which we have $\mathbf{f}(\mathbf{x}) = \mathbf{y} \Rightarrow g_{\alpha}(\mathbf{x}) = c$ forms the straight line $\alpha^T \cdot \mathbf{y} = c$ which is defined by its normal vector α (see Figure 4.5). Points in the objective space with a smaller value of g_{α} are situated to the left of (or beneath) this straight line, points with a larger value of g_{α} to the right of (or above) the straight line. $\mathbf{f}(R)$ has to be situated completely to the right of or above this straight line, so that there is no point $\bar{\mathbf{x}}$ with the property $(\bullet)$. If the border of $\mathbf{f}(R)$ is smooth (i.e., if it is a continuously differentiable curve), it follows that the tangent to this curve of efficient points must not form some non-zero angle with the straight line $\alpha^T \cdot \mathbf{y} = c$. Therefore it has to be identical with it. Consequently, α is the normal vector to the tangent of the efficient curve.

The above conclusion is well-known in the literature on vector optimization (see [DAS & DENNIS, 1996A], [DAS, 1997]), but only in the context of the spe-

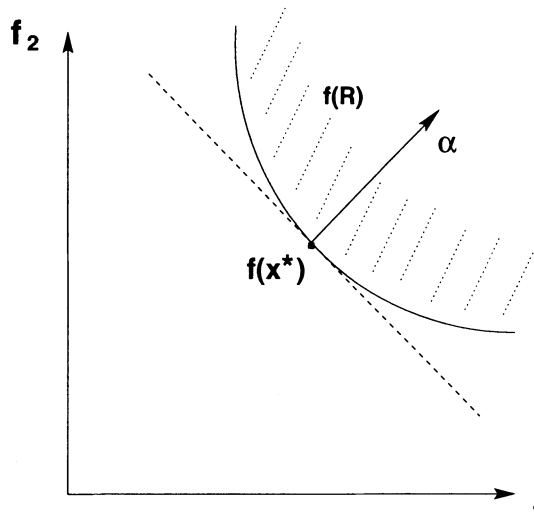


Figure 4.5 Detail of the image set $f(R)$ and the delimiting border curve (efficient curve) in the bicriterial case $k = 2$ (schematized). The dashed line is the straight line $\alpha^T \cdot y = c$ (see text).

cial case described, i.e. for *minima* of g_α in *bicriterial* problems. The geometrical meaning of the vector α is, however, not limited to this special case. In the remainder of this section we are going to demonstrate the following generalization by means of a differential-topological examination: If the image set $f(R)$ behaves in a neighborhood of $f(x^*)$ like a compact set with a smooth border, then α is the normal vector to the tangent plane of the border $\partial f(R)$ of the image set $f(R)$.

In order to keep the argumentation as transparent as possible we will first examine the special case of unconstrained vector optimization problems.

The following (auxiliary) proposition does not yet require smoothness properties of the border of the image set $f(R)$:

Theorem 4.2:

Let y^* be a (locally) efficient point and x^* a corresponding (locally) Pareto optimal point [i.e. $f(x^*) = y^*$] of an unconstrained vector optimization problem. Let g_α denote a convex combination of the objectives for which x^* is a stationary point (and therefore fulfills the Karush-Kuhn-Tucker condition).

Then for the weight vector α we have:

α is an element of the orthogonal complement to the vector subspace (image $f'(x^*)$) $\subset \mathbb{R}^k$, where $f'(x^*)$ is the Jacobian matrix of f in the point⁴ x^* .

⁴ In order not to overload the notation, in the following the evaluation point x^* will no longer be indicated in such cases where it is implied by the context.

Proof. From the first-order Karush-Kuhn-Tucker condition we get:

$$\nabla g_{\alpha}(\mathbf{x}^*) = \mathbf{0} \iff \sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}^*) = \mathbf{0} \iff \alpha^T \begin{pmatrix} \nabla f_1(\mathbf{x}^*)^T \\ \vdots \\ \nabla f_k(\mathbf{x}^*)^T \end{pmatrix} = 0. \quad (4.8)$$

Because of $\begin{pmatrix} \nabla f_1(\mathbf{x}^*)^T \\ \vdots \\ \nabla f_k(\mathbf{x}^*)^T \end{pmatrix} = \mathbf{f}'(\mathbf{x}^*)$ it follows that α is orthogonal to the columns of the Jacobian matrix $\mathbf{f}'(\mathbf{x}^*)$ and thus to the image of the linear mapping $\mathbf{f}'(\mathbf{x}^*)$. ■

As a corollary one obtains the statement $\text{rank } \mathbf{f}'(\mathbf{x}^*) < k$, i.e. the linear mapping $\mathbf{f}'(\mathbf{x}^*) : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is not surjective in a Pareto optimal point $\mathbf{x}^*$. If $\text{rank } \mathbf{f}'(\mathbf{x}^*) = k - 1$, from Theorem 4.2 one can conclude furthermore the **uniqueness** of the assignment of a weight vector α (and therefore of a scalar-valued function g_{α}) to a Pareto optimal point $\mathbf{x}^*$.

If in an appropriate neighborhood of $\mathbf{f}(\mathbf{x}^*)$ the image set $\mathbf{f}(\mathbb{R}^n)$ behaves like a bordered differentiable manifold of dimension k , the geometrical meaning of the weight vectors α can be put in even more concrete forms:

Theorem 4.3:

Let $\mathbf{y}^*$ be a globally efficient point and $\mathbf{x}^*$ an associated globally Pareto optimal point [i.e. $\mathbf{f}(\mathbf{x}^*) = \mathbf{y}^*$] of an unconstrained vector optimization problem. Let g_{α} denote a convex combination of the objectives for which $\mathbf{x}^*$ is a stationary point (and therefore fulfills the Karush-Kuhn-Tucker condition). In addition be:

- $\text{rank } \mathbf{f}'(\mathbf{x}^*) = k - 1$
- *There is an open neighborhood $U(\mathbf{y}^*)$ of $\mathbf{y}^*$, so that $\mathbf{f}(\mathbb{R}^n) \cap U(\mathbf{y}^*) =: M$ is a bordered differentiable manifold of dimension k .*

Then we have:

- (A) $\mathbf{y}^* \in \partial M$, where the $(k - 1)$ -dimensional border manifold of M is denoted by ∂M .
- (B) α is orthogonal to the tangent plane $T_{\mathbf{y}^*} \partial M$ of ∂M in $\mathbf{y}^*$.

Proof. Assertion (A) will be proved by contradiction. Assume therefore that $\mathbf{y}^*$ is not an element of ∂M . It follows that $\mathbf{y}^*$ is an inner point of M , i.e. that there is a δ -neighborhood $U(\mathbf{0}, \delta)$ of $\mathbf{0} \in \mathbb{R}^k$ with $\mathbf{y}^* + U(\mathbf{0}, \delta) \subseteq M$. Now choose a vector $\mathbf{v} \in \mathbb{R}_+^k$. Then there is a $\lambda \in \mathbb{R}, \lambda > 0$ with $(-\lambda) \cdot \mathbf{v} \in U(\mathbf{0}, \delta)$, and $\hat{\mathbf{y}} := (\mathbf{y}^* - \lambda \mathbf{v}) \in M \subset \mathbf{f}(\mathbb{R}^n)$ is true. Because of $\mathbf{y}^* - \hat{\mathbf{y}} = \lambda \mathbf{v} \in \mathbb{R}_+^k$ we

can conclude that $\tilde{\mathbf{y}} \leq \mathbf{y}^*$ [and $\tilde{\mathbf{y}} \in \mathbf{f}(\mathbb{R}^n)$], in contradiction to the global efficiency of $\mathbf{y}^*$. Note that a proof of the immediately plausible assertion that a globally efficient point cannot be an inner point of the image set can also be found in the literature (see e.g. [GÖPFERT & NEHSE, 1990]).

(B) The assertion follows from Theorem 4.2, if one can show that $T_{\mathbf{y}^*}\partial M = \text{image } \mathbf{f}'(\mathbf{x}^*)$. We will prove this by contradiction. Therefore, be $T_{\mathbf{y}^*}\partial M \neq \text{image } \mathbf{f}'(\mathbf{x}^*)$. As we assumed that both $T_{\mathbf{y}^*}\partial M$ and $\text{image } \mathbf{f}'(\mathbf{x}^*)$ are $(k-1)$ -dimensional vector subspaces of the $\mathbb{R}^k$, it follows that $\text{image } \mathbf{f}'(\mathbf{x}^*) \not\subseteq T_{\mathbf{y}^*}\partial M$, and hence:

There is a vector $\delta \mathbf{y} \in \text{image } \mathbf{f}'(\mathbf{x}^*)$, so that $\delta \mathbf{y}$ can be represented by $\delta \mathbf{y} = \boldsymbol{\xi} + \boldsymbol{\eta}$, where $\boldsymbol{\xi} \in T_{\mathbf{y}^*}\partial M$, $\boldsymbol{\eta} \in (T_{\mathbf{y}^*}\partial M)^\perp$, $\boldsymbol{\eta} \neq \mathbf{0}$. Let us denote the corresponding inverse-image vector by $\delta \mathbf{x}$, i.e. $\mathbf{f}'(\mathbf{x}^*) \cdot \delta \mathbf{x} = \delta \mathbf{y}$.

Now for a sufficiently small $a \in \mathbb{R}_+$ consider the curve

$$\tau : \begin{cases} (-a, +a) \rightarrow \mathbb{R}^k \\ t \mapsto \mathbf{f}(\mathbf{x}^* + t \cdot \delta \mathbf{x}) \end{cases} \quad (4.9)$$

By virtue of $\tau'(0) = \mathbf{f}'(\mathbf{x}^*) \cdot \delta \mathbf{x} = \delta \mathbf{y} = \boldsymbol{\xi} + \boldsymbol{\eta}$, either $+\tau'(0)$ or $-\tau'(0)$ is an element⁵ of the outward directed tangent space of the bordered manifold M in the point $\mathbf{y}^*$. For one of the two possible signs of t and for sufficiently small $|t|$ [so that $\tau(t)$ is represented sufficiently well by the linear approximation $\tau(0) + \tau'(0) \cdot t$] the image points of the curve τ are therefore situated outside M . This is a contradiction to the definition of M , so that the assumption $T_{\mathbf{y}^*}\partial M \neq \text{image } \mathbf{f}'(\mathbf{x}^*)$ must be false. ■

Analogous statements about the geometrical meaning of the weight vector $\boldsymbol{\alpha}$ can also be proved for the **constrained case**.

Scenario (*): Let $\mathbf{y}^*$ be an efficient point with $\mathbf{f}(\mathbf{x}^*) = \mathbf{y}^*$. Assume that in $\mathbf{x}^*$ the constraints $h_1, \dots, h_p$ (from the set of equality and inequality constraints) are active, i.e. $h_1(\mathbf{x}^*) = 0, \dots, h_p(\mathbf{x}^*) = 0$. Let the following constraint qualification be fulfilled: The vectors $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_p(\mathbf{x}^*)\}$ are linearly independent. Now consider the $(n-p)$ -dimensional submanifold N , which is created by transforming the p active constraints in an open neighborhood $U(\mathbf{x}^*)$ of $\mathbf{x}^*$ into equality constraints: $N := \{\mathbf{x} \in U(\mathbf{x}^*) | h_1(\mathbf{x}) = 0, \dots, h_p(\mathbf{x}) = 0\}$. Consider a local parameter representation (chart) of the submanifold N in the form of a differentiable homeomorphism $\mathbf{s} : T \rightarrow V \subset \mathbb{R}^n$, where T is an open subset of $\mathbb{R}^{n-p}$ and $V \subset N$ is an open neighborhood [with respect to N] of $\mathbf{x}^*$.

Within this scenario, analogously to Proposition 4.2 (for the unconstrained case), the following proposition is valid:

⁵ Strictly speaking, $\tau'(0)$ denotes the linear mapping $\tau'(0) : \delta t \mapsto (\boldsymbol{\xi} + \boldsymbol{\eta}) \cdot \delta t$, i.e. the multiplication by this vector.

Theorem 4.4:

Assuming that scenario $(*)$ is valid and that g_α is a convex combination of the objectives for which $\mathbf{x}^*$ fulfills the Karush-Kuhn-Tucker condition, we have for the weight vector α :

$\alpha \in \left(\text{image } \tilde{\mathbf{f}}'(\mathbf{t}^*) \right)^\perp$, where $\tilde{\mathbf{f}}(\mathbf{t}) := \mathbf{f} \circ \mathbf{s}(\mathbf{t})$ is the objective function defined on the parameters $\mathbf{t}$ of a chart $\mathbf{s}$ of the submanifold N . [For the chart be $\mathbf{s}(\mathbf{t}^*) = \mathbf{x}^*$.]

This statement is valid for any charts of the differentiable structure of N (casually speaking: for any local parametrization of N).

Proof. The rank p of the matrix $\begin{pmatrix} \nabla h_1(\mathbf{x}^*)^T \\ \vdots \\ \nabla h_p(\mathbf{x}^*)^T \end{pmatrix}$ is full. The implicit-function theorem guarantees that by rearranging the vector components x_i (i.e. by renaming the coordinate axes) one can always obtain a local parametrization of the submanifold N by $\{x_{p+1}, \dots, x_n\}$ in the form of $\mathbf{s} : (x_{p+1}, \dots, x_n)^T \mapsto (s_1(x_{p+1}, \dots, x_n), \dots, s_p(x_{p+1}, \dots, x_n), x_{p+1}, \dots, x_n)^T$. The objective function $\tilde{\mathbf{f}}$ defined on these chart parameters is

$$\tilde{\mathbf{f}}(x_{p+1}, \dots, x_n) = \mathbf{f}(s_1(x_{p+1}, \dots, x_n), \dots, s_p(x_{p+1}, \dots, x_n), x_{p+1}, \dots, x_n) \quad (4.10)$$

and has the Jacobian matrix $\tilde{\mathbf{f}}' = \mathbf{f}' \frac{\partial \mathbf{s}}{\partial x_{p+1} \dots x_n}$. Scalar multiplication of this equation by α results in

$$\alpha^T \tilde{\mathbf{f}}'(x_{p+1}^*, \dots, x_n^*) = \alpha^T \mathbf{f}'(\mathbf{x}^*) \frac{\partial \mathbf{s}}{\partial x_{p+1} \dots x_n}(x_{p+1}^*, \dots, x_n^*) . \quad (4.11)$$

By taking the active constraints $h_1, \dots, h_p$ into consideration, from the Karush-Kuhn-Tucker condition follows:

$$\nabla g_\alpha(\mathbf{x}^*)^T \equiv \alpha^T \mathbf{f}'(\mathbf{x}^*) = - \sum_{j=1}^p \lambda_j \nabla h_j(\mathbf{x}^*)^T. \quad (4.12)$$

Insertion of this relation into Equation (4.11) yields

$$\alpha^T \tilde{\mathbf{f}}'(x_{p+1}^*, \dots, x_n^*) = \left(- \sum_{j=1}^p \lambda_j \nabla h_j(\mathbf{x}^*)^T \right) \frac{\partial \mathbf{s}}{\partial x_{p+1} \dots x_n}(x_{p+1}^*, \dots, x_n^*) . \quad (4.13)$$

The system of equations $h_1(\mathbf{s}(x_{p+1}, \dots, x_n)) = 0, \dots, h_p(\mathbf{s}(x_{p+1}, \dots, x_n)) = 0$ is fulfilled for an entire neighborhood of the parameter point $(x_{p+1}^*, \dots, x_n^*)^T$, i.e. the function $\tilde{\mathbf{h}} := (h_1 \circ \mathbf{s}, \dots, h_p \circ \mathbf{s})^T$ in this neighborhood is a different

expression for the zero map. Therefore, the corresponding Jacobian matrix $\tilde{h}'(x_{p+1}^*, \dots, x_n^*)$ is also a zero map:

$$\begin{aligned} \tilde{h}'(x_{p+1}^*, \dots, x_n^*) &= h'(\mathbf{x}^*) \frac{\partial s}{\partial x_{p+1} \dots x_n}(x_{p+1}^*, \dots, x_n^*) \\ &= \begin{pmatrix} \nabla h_1(\mathbf{x}^*)^T \\ \vdots \\ \nabla h_p(\mathbf{x}^*)^T \end{pmatrix} \frac{\partial s}{\partial x_{p+1} \dots x_n}(x_{p+1}^*, \dots, x_n^*) = \mathbf{0}. \end{aligned} \quad (4.14)$$

Since consequently for each of the gradients $\nabla h_i(\mathbf{x}^*), i \in \{1, \dots, m\}$ we have: $\nabla h_i(\mathbf{x}^*)^T \frac{\partial s}{\partial x_{p+1} \dots x_n}(x_{p+1}^*, \dots, x_n^*) = 0$, from Equation (4.13) follows: $\alpha^T \tilde{f}'(x_{p+1}^*, \dots, x_n^*) = 0$. The vector α is hence orthogonal to all columns of $\tilde{f}'(x_{p+1}^*, \dots, x_n^*)$ and consequently to the vector subspace $\text{image } \tilde{f}'(x_{p+1}^*, \dots, x_n^*)$.

It still remains to be proved that the claim is valid independently of the parametrization s . Let $\mathbf{r}(t)$ be another parametrization (chart) of N in a neighborhood of the point $\mathbf{x}$ with $\mathbf{r}(t^*) = \mathbf{x}^*$. The objective function defined on the chart parameters t corresponding to that chart $\mathbf{r}$ is $\tilde{f} := f \circ \mathbf{r} = f \circ s \circ s^{-1} \circ \mathbf{r} = \tilde{f} \circ s^{-1} \circ \mathbf{r}$ and has the Jacobian matrix $\tilde{f}'(t^*) = \tilde{f}'(x_{p+1}^*, \dots, x_n^*) \cdot (s^{-1} \circ \mathbf{r})'(t^*)$ in the point t^* . Since the change of charts $s^{-1} \circ \mathbf{r}$ is a diffeomorphism, the Jacobian matrix $(s^{-1} \circ \mathbf{r})'(t^*)$ is an isomorphism. We obtain that $\text{image } \tilde{f}'(t^*) = \text{image } \tilde{f}'(x_{p+1}^*, \dots, x_n^*)$. ■

If $f(R)$ is locally a bordered k -dimensional manifold (called M), it is possible to show for the constrained case as well that α is a normal vector to the tangent plane $T_{\mathbf{y}^*} \partial M$:

Theorem 4.5:

Let $\mathbf{y}^*$ be a globally efficient point and $\mathbf{x}^*$ an associated (globally) Pareto optimal point, i.e. $f(\mathbf{x}^*) = \mathbf{y}^*$. Let the constraints $h_1, \dots, h_p$ be active in $\mathbf{x}^*$ and the constraint qualification be fulfilled, i.e. the vectors $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_p(\mathbf{x}^*)\}$ are linearly independent. Let g_α denote a convex combination of the objectives for which $\mathbf{x}^*$ fulfills the Karush-Kuhn-Tucker condition. Furthermore, let $U(\mathbf{x}^*)$ denote an open neighborhood of $\mathbf{x}^*$, let s be a chart of the $(n-p)$ -dimensional submanifold $N := \{\mathbf{x} \in U(\mathbf{x}^*) | h_1(\mathbf{x}) = 0, \dots, h_p(\mathbf{x}) = 0\}$, let $\mathbf{p}^* \in \mathbb{R}^{(n-p)}$ be the inverse image of $\mathbf{x}^*$ with respect to s [i.e. $s(\mathbf{p}^*) = \mathbf{x}^*$], and let $\tilde{f} := f \circ s$ denote the objective function defined on the chart parameters.

In addition, let the following assumptions be valid:

- Rank $\tilde{f}'(\mathbf{p}^*) = k - 1$ [As changes of chart are diffeomorphic, this claim about the rank is valid for all charts of the atlas, once it is fulfilled for one.]

- There exists an open neighborhood $U(\mathbf{y}^*)$ of $\mathbf{y}^*$, such that $\mathbf{f}(R) \cap U(\mathbf{y}^*) =: M$ is a bordered k -dimensional differentiable manifold.

Then we have:

(A) $\mathbf{y}^* \in \partial M$, where ∂M denotes the $(k-1)$ -dimensional border manifold of M .

(B) $\boldsymbol{\alpha}$ is orthogonal to the tangent plane $T_{\mathbf{y}^*} \partial M$ of ∂M in $\mathbf{y}^*$.

Proof. The proof of assertion (A) is identical to Theorem 4.3.

(B) In analogy to Theorem 4.3 the assertion (B) follows from Theorem 4.4, if one can show that $T_{\mathbf{y}^*} \partial M = \text{image } \tilde{\mathbf{f}}'(\mathbf{p}^*)$. This shall be proved again by contradiction.

Let $T_{\mathbf{y}^*} \partial M \neq \text{image } \tilde{\mathbf{f}}'(\mathbf{p}^*)$. It follows that $\text{image } \tilde{\mathbf{f}}'(\mathbf{p}^*) \not\subseteq T_{\mathbf{y}^*} \partial M$, and we can conclude:

There exists a vector $\delta \mathbf{y} \in \text{image } \tilde{\mathbf{f}}'(\mathbf{p}^*)$ with $\delta \mathbf{y} = \boldsymbol{\xi} + \boldsymbol{\eta}$, $\boldsymbol{\xi} \in T_{\mathbf{y}^*} \partial M$, $\boldsymbol{\eta} \in (T_{\mathbf{y}^*} \partial M)^\perp$, $\boldsymbol{\eta} \neq \mathbf{0}$, and furthermore there exists a vector $\delta \mathbf{p} \in \mathbb{R}^{n-p}$ with $\tilde{\mathbf{f}}'(\mathbf{p}^*) \delta \mathbf{p} = \delta \mathbf{y}$.

Let us now, for sufficiently small $a \in \mathbb{R}_+$, examine the curve

$$\boldsymbol{\tau} : \begin{cases} (-a, +a) & \rightarrow \mathbb{R}^k \\ t & \mapsto \mathbf{f} \circ \mathbf{s}(\mathbf{p}^* + t \cdot \delta \mathbf{p}) \equiv \tilde{\mathbf{f}}(\mathbf{p}^* + t \cdot \delta \mathbf{p}) \end{cases} \quad (4.15)$$

By appropriately reducing the neighborhood $U(\mathbf{x}^*)$ one can guarantee that apart from the constraints that are active in $\mathbf{x}^*$ there are no further constraints active in any point of $R \cap U(\mathbf{x}^*)$ [reason: continuity of the functions h_j]. The submanifold N thus takes in consideration a superset of the constraints which are active in points of $R \cap U(\mathbf{x}^*)$, so that $N \subseteq R$. As the chart $\mathbf{s}$ is defined on an open neighborhood of $\mathbf{p}^*$ and the image points of $\mathbf{s}$ are situated in N , the existence of the curve $\boldsymbol{\tau}$ is therefore guaranteed and all curve points are situated within $\mathbf{f}(R)$.

On the other hand, because of $\boldsymbol{\tau}'(0) = \tilde{\mathbf{f}}'(\mathbf{p}^*) \delta \mathbf{p} = \delta \mathbf{y} = \boldsymbol{\xi} + \boldsymbol{\eta}$, either $+\boldsymbol{\tau}'(0)$ or $-\boldsymbol{\tau}'(0)$ is an element of the outward directed tangent space of the bordered manifold M in the point $\mathbf{y}^*$. For one of the two possible signs of t and for sufficiently small $|t|$ [so that $\boldsymbol{\tau}(t)$ is represented adequately well by the linear approximation $\boldsymbol{\tau}(0) + \boldsymbol{\tau}'(0) \cdot t$] the image points of the curve $\boldsymbol{\tau}$ are therefore situated outside M . As this contradicts the definition of M , the assumption $T_{\mathbf{y}^*} \partial M \neq \text{image } \tilde{\mathbf{f}}'(\mathbf{p}^*)$ must be false. ■

4.4 Classification of Efficient Points

The theorem of Kuhn and Tucker supplies information of **first order** about a convex combination g_α of the objectives in a Pareto optimal point $\mathbf{x}^*$. This information was used in the previous section to determine partially the geometry

of the tangent plane $T_{\mathbf{y}} \cdot \partial M$ of the border ∂M of the image set $\mathbf{f}(R)$ — in the form of the normal vector to this tangent plane given by the weight vector α . In this section we will now establish a connection between the information of **second order** about g_α in the point $\mathbf{x}^*$ — i.e. the type of the stationary point $\mathbf{x}^*$ — and the information of second order about the border manifold ∂M . It will turn out that depending on the local **curvature** of ∂M the Pareto optimal point $\mathbf{x}^*$ is either a **minimizer** or a **saddle point** of the scalar-valued function g_α . In the following the objective function $\mathbf{f}$ is assumed to be twice continuously differentiable.

The principal connection can again be made clear when considering the bicriterial case.

Let us first assume that $\mathbf{x}^*$ is a Pareto optimal point and a global minimizer of g_α . A conclusion of Section 4.3 is that all points of $\mathbf{f}(R)$ must be situated above the straight line $\alpha^T \cdot \mathbf{y} = g_\alpha(\mathbf{x}^*) = c$. In the case of a smooth efficient curve this does not only imply that the tangent to the curve in the point $\mathbf{f}(\mathbf{x}^*)$ is identical with the straight line $\alpha^T \cdot \mathbf{y} = c$. Moreover, the efficient curve must also be bent ‘inwards’, i.e. like a border curve of a convex set [as shown in Figure 4.5].

If on the other hand for any small neighborhood $U(\mathbf{x}^*)$ of a Pareto optimal point $\mathbf{x}^*$ the border of the image set $\mathbf{f}(U(\mathbf{x}^*))$ is bent outwards, each of these neighborhoods contains points $\tilde{\mathbf{x}}$ with $g_\alpha(\tilde{\mathbf{x}}) < g_\alpha(\mathbf{x}^*) = c$ [Let g_α be a convex combination of the objectives to which $\mathbf{x}^*$ is a stationary point]. Therefore, such a stationary point $\mathbf{x}^*$ cannot be a minimizer of g_α , but must be a *saddle point* to g_α — provided that $\mathbf{f}(U(\mathbf{x}^*))$ possesses the full dimension k (i.e. 2).

Generalizing the above argumentation we will now show for the unconstrained vector optimization problem of arbitrary dimension k the following connection between the local curvature of the border manifold [of the image set $\mathbf{f}(\mathbb{R}^n)$] and the type of stationary point of g_α :

Theorem 4.6:

Let $\mathbf{y}^*$ be a (locally) efficient point and $\mathbf{x}^*$ an associated Pareto optimal point [i.e. $\mathbf{f}(\mathbf{x}^*) = \mathbf{y}^*$] of an unconstrained vector optimization problem. Let g_α denote a convex combination of the objectives for which $\mathbf{x}^*$ is a stationary point. Furthermore let there be an open neighborhood $V(\mathbf{y}^*)$ of $\mathbf{y}^*$ and a sequence of ϵ -neighborhoods $U_\epsilon(\mathbf{x}^*)$ of $\mathbf{x}^*$ (with $\epsilon \xrightarrow{n \rightarrow \infty} 0$), so that $\mathbf{f}(U_\epsilon(\mathbf{x}^*)) \cap V(\mathbf{y}^*) =: M_\epsilon$ is a sequence of bordered k -dimensional differentiable manifolds with the following property: The principal curvatures of the $(k-1)$ -dimensional border manifolds (hypersurfaces) ∂M_ϵ in the point $\mathbf{y}^* \in \partial M_\epsilon$ converge to the values $\mu_1, \dots, \mu_{k-1}$ for $\epsilon \rightarrow 0$, where $\mu_1 \neq 0, \dots, \mu_{k-1} \neq 0$. The curvatures refer to the normal vector α of ∂M_ϵ pointing towards the interior of M_ϵ (see Theorem 4.3).

Then the following assertions are true:

- (A) $\mu_i > 0 \quad \forall i = 1, \dots, k-1 \iff \mathbf{x}^*$ is a local minimizer of g_α .
 (B) $\exists i \in \{1, \dots, k-1\}$ with $\mu_i < 0 \iff \mathbf{x}^*$ is a saddle point of g_α .

Proof. Assertion (A), ' $\implies$ ' :

According to our assumption all principal curvatures of ∂M_ϵ [from an $\epsilon_0 > 0$ onwards in the sequence of manifolds] are larger than zero. Consequently, the normal curvatures of all surface curves γ on ∂M_ϵ passing through $\mathbf{y}^*$ are strictly positive in the point $\mathbf{y}^*$. For all such curves γ [with $\gamma(0) = \mathbf{y}^*$] the tangent to the curve (in the point $\mathbf{y}^*$) lies in the tangent plane $T_{\mathbf{y}^*} \partial M_\epsilon$ affinely shifted to $\mathbf{y}^*$, which again, according to Theorem 4.3, contains the points $\mathbf{y}$ with $\alpha^T \cdot \mathbf{y} = \alpha^T \cdot \mathbf{y}^* = g_\alpha(\mathbf{x}^*) = c$. Therefore for curve parameters t sufficiently close to zero all image points $\gamma(t)$ of these surface curves γ — and consequently all points $\mathbf{y}$ of a sufficiently small neighborhood of $\mathbf{y}^*$ on ∂M_ϵ — comply with the inequality $\alpha^T \cdot \gamma(t) \geq g_\alpha(\mathbf{x}^*)$, or $\alpha^T \cdot \mathbf{y} \geq g_\alpha(\mathbf{x}^*)$ respectively. Since for points of M_ϵ the minimum with respect to the g_α -value is realized on the border ∂M_ϵ , the above inequality is valid for all points of a (sufficiently small) neighborhood of $\mathbf{y}^*$ on M_ϵ . Due to the continuity of $\mathbf{f}$ all image points of a (sufficiently small) neighborhood $U(\mathbf{x}^*)$ are contained in this neighborhood, so that in virtue of $\alpha^T \cdot \mathbf{f}(\mathbf{x}) = g_\alpha(\mathbf{x}) \geq g_\alpha(\mathbf{x}^*) \quad \forall \mathbf{x} \in U(\mathbf{x}^*)$ the partial assertion is proved.

Assertion (A), ' $\impliedby$ ' :

For arbitrarily small neighborhoods $U_\epsilon(\mathbf{x}^*)$ of $\mathbf{x}^*$ one can assume: $g_\alpha(\mathbf{x}) \geq g_\alpha(\mathbf{x}^*) \quad \forall \mathbf{x} \in U_\epsilon(\mathbf{x}^*)$. Thus, for all points $\mathbf{y}$ of the corresponding manifolds M_ϵ , especially for all points of the border manifolds ∂M_ϵ and, consequently, also for all image points of the surface curves γ on ∂M_ϵ passing through $\mathbf{y}^*$, the following inequality (*) is true: $\alpha^T \cdot \mathbf{y} \geq c$. Since the tangents to the curves γ lie in the tangent surface defined by $\alpha^T \cdot \mathbf{y} = c$, the inequality (*) is in contradiction to the assumption that there can exist (for arbitrarily small ϵ) a negative normal curvature of ∂M_ϵ in the point $\mathbf{y}^*$. Therefore all normal curvatures and consequently all principal curvatures are positive.

Assertion (B), ' $\implies$ ' :

According to the assumption of the theorem the image-manifold M_ϵ of any arbitrarily small neighborhood $U_\epsilon(\mathbf{x}^*)$ of $\mathbf{x}^*$ has the full dimension k . In particular, M_ϵ contains points $\bar{\mathbf{y}}$, which, looked at from the border point $\mathbf{y}^* \in \partial M_\epsilon$, are situated inside M_ϵ . Thus: $\exists \bar{\mathbf{y}} \in M_\epsilon$ and $\lambda > 0$, so that $\bar{\mathbf{y}} = \mathbf{y}^* + \lambda \alpha$. For the inverse image $\bar{\mathbf{x}} \in U_\epsilon(\mathbf{x}^*)$ of such a point $\bar{\mathbf{y}}$ one can write correspondingly: $g_\alpha(\bar{\mathbf{x}}) = \alpha^T \cdot \mathbf{f}(\bar{\mathbf{x}}) = \alpha^T \cdot \bar{\mathbf{y}} = g_\alpha(\mathbf{x}^*) + \lambda \alpha^T \alpha > g_\alpha(\mathbf{x}^*)$.

On the other hand, because of the left-hand side of assertion (B) one can assume that at least one principal curvature μ_{i_0} of ∂M_ϵ in the point $\mathbf{y}^*$ is smaller than zero. Let γ denote a surface curve, the velocity vector of which is given by the principal direction corresponding to μ_{i_0} [see the definition of tangent vectors as equivalence classes of curves, Section 4.2, point (c)] and for which $\gamma(0) = \mathbf{y}^*$. For all curve parameters t sufficiently close to zero, we get for the image points $\gamma(t)$ of this curve: $\alpha^T \cdot \gamma(t) < \alpha^T \cdot \mathbf{y}^* = g_\alpha(\mathbf{x}^*) = c$,

as these points $\gamma(t)$ are situated on that side of the (affinely shifted) tangent plane $\alpha^T \cdot y = c$ which is opposite to the interior of M_ϵ . Since arbitrarily small ϵ -neighborhoods $U_\epsilon(x^*)$ contain inverse images (with respect to f) of such curve points $\gamma(t)$, one can deduce: In every arbitrarily small neighborhood $U_\epsilon(x^*)$ there is a point $\tilde{x}$ with $g_\alpha(\tilde{x}) = \alpha^T \cdot \gamma(t) < g_\alpha(x^*)$. Since in every arbitrarily small neighborhood $U_\epsilon(x^*)$ there are both a point $\tilde{x}$ with $g_\alpha(\tilde{x}) > g_\alpha(x^*)$ and a point $\tilde{x}$ with $g_\alpha(\tilde{x}) < g_\alpha(x^*)$, x^* must be a saddle point of g_α .

Assertion (B), ' $\Leftarrow$ ' :

According to assertion (A) principal curvatures of ∂M_ϵ , which are all larger than zero, imply that x^* is a minimizer of g_α . If x^* is a saddle point of g_α , at least one of the principal curvatures must be smaller than zero. ■

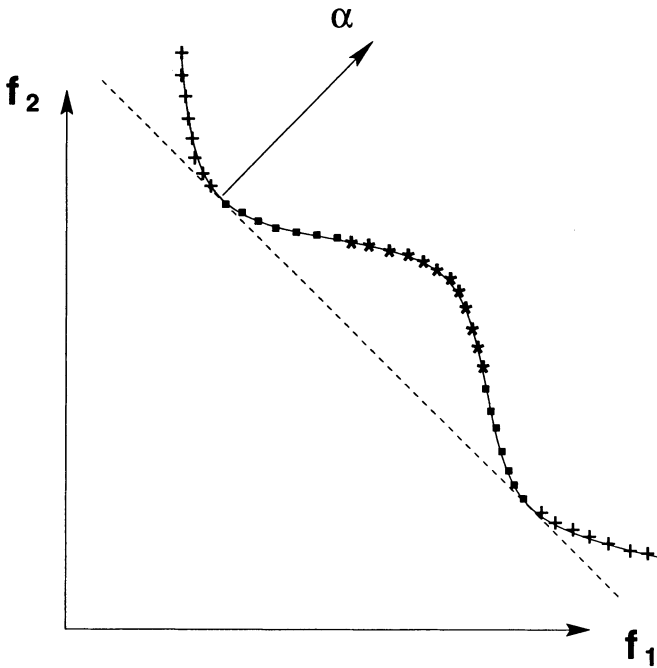


Figure 4.6 Schematic efficient curve of a bicriterial vector optimization problem. Depending on the curvature of the efficient curve the associated Pareto optimal points are either minima or saddle points (the latter marked by $*$) of the convex combinations g_α (parametrized by α). The curve parts marked by $+$ consist of global minima of g_α , while the curve parts marked by $\blacksquare$ are formed by local minima. For each of these points, in the other curve arc there exists a counterpart with an even smaller g_α -value (for the same α). [Pay attention to the fact that the points y of equal g_α -values (for a given α) lie on the straight line $\alpha^T \cdot y = c$, where c is the distance of this straight line from the coordinate origin.]

Considering the local curvature characteristics of the border of the image set of $\mathbf{f}$ one can thereby classify the stationary points of the convex combinations of the objectives.

Figure 4.6 illustrates the result of the considerations of this section taking as an example a bicriterial efficient curve. A particularly interesting point is the contact point between minimum region and saddle point region. While all other points are surrounded by points which have the same sign of curvature, so that in the neighborhood of these points the curve normal α varies in both directions — i.e. both towards larger and towards smaller values of the component α_1 —, this is not the case in the above-mentioned transition point. Since in this point the curvature changes its sign, α varies in the neighborhood of this point only in one direction, i.e. the component α_1 has an extremum here. This phenomenon can indeed be observed in the numerical example calculated in Section 7.2 (see Figure 7.9 and the last paragraph of Section 7.2).

When we look at the submanifold consisting of the stationary points of g_α , we can see the formal reflection of this behavior of α in the structure of the Jacobian matrix the full rank of which guarantees the dimensionality of the submanifold. The connection between the local variation of α and the structure of this Jacobian matrix will be discussed in Paragraph 5.2.3.

The Manifold of Stationary Points

The further examinations presuppose that the feasible set R is defined by m equality constraints¹ $h_i(\mathbf{x}) = 0, i = 1, \dots, m$. In this case the necessary condition — according to Kuhn and Tucker — for Pareto optimal points has the form of a system of equations. The set of all points which fulfill this condition can therefore be interpreted as a zero manifold in an extended variable space, the product space formed by the actual variables $\mathbf{x}$, the Lagrange multipliers $\boldsymbol{\lambda}$ and the weight vectors $\boldsymbol{\alpha}$. On certain conditions this zero manifold is a $(k - 1)$ -dimensional differentiable manifold.

This differentiable manifold will be examined more closely in the following chapter. In Section 5.1 it will be defined exactly, Section 5.2 gives a necessary and sufficient criterion for its existence and interprets this criterion in view of optimization. In Section 5.3, finally, a parametrization will be constructed which meets the special requirements of a homotopy method with several homotopy parameters.

For all statements of this chapter — like for the rest of the present book — the objective function f and the constraint function h are supposed to be twice continuously differentiable.

¹ Application problems with inequality constraints can either be put in this form by introducing slack variables or can be transformed into subproblems which have only equality constraints by means of active-set strategies (see e.g. [LUENBERGER, 1984]). If one uses slack variables, one loses the information contained in the sign of the Lagrange multipliers of the active inequality constraints. This one has to pay special attention to. On the other hand, active-set-strategies produce systems of non-linear equations of variable dimension. Since the actual dimension has to be determined by numerical calculations, rounding errors can lead to false decisions regarding the dimension.

5.1 Karush-Kuhn-Tucker Points as a Differentiable Manifold M

For every Pareto optimal point $\mathbf{x}^*$ there is, according to Theorem 4.1, a weight vector $\alpha^* \in \mathbb{R}_+^k$, so that $\mathbf{x}^*$ is a Karush-Kuhn-Tucker point (in short: KKT-point) of the corresponding scalar-valued optimization problem with the objective function g_{α^*} . If the feasible set R is given in the form of m equality constraints, this implies the following statement:

For every Pareto optimal point $\mathbf{x}^*$ (fulfilling the mentioned constraint qualification) there exists a vector² $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \alpha^*) \in \mathbb{R}^{n+m+k}$, which satisfies the condition $\alpha^* \in \mathbb{R}_+^k$ and solves the following system of equations:

$$\sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}) + \sum_{j=1}^m \lambda_j \nabla h_j(\mathbf{x}) = \mathbf{0} \quad (n \text{ equations}) \quad (5.1)$$

$$h_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \quad (m \text{ equations}) \quad (5.2)$$

$$\sum_{l=1}^k \alpha_l = 1 \quad (1 \text{ equation}) \quad (5.3)$$

By defining a function $\mathbf{F} : \mathbb{R}^{n+m+k} \rightarrow \mathbb{R}^{n+m+1}$ in the following way

$$\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) := \begin{pmatrix} \sum_{i=1}^k \alpha_i \nabla f_i(\mathbf{x}) + \sum_{j=1}^m \lambda_j \nabla h_j(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \\ \sum_{l=1}^k \alpha_l - 1 \end{pmatrix}, \quad (5.4)$$

where the vector-valued function $\mathbf{h} := (h_1, \dots, h_m)^T$ enables us to write the equality constraints (5.2) as $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, we obtain the simple form

$$\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = \mathbf{0} \quad (5.5)$$

for the system of Equations (5.1) to (5.3).

When reading Theorem 4.1 in the opposite direction, one obtains the assertion: Points $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \alpha^*) \in \mathbb{R}^{n+m+k}$, which satisfy Equation (5.5) and the condition $\alpha^* \in \mathbb{R}_+^k$, are candidates for Pareto optimal points.

In the following a subset M of this candidate set is going to be examined more closely. We obtain M by restricting the condition $\alpha^* \in \mathbb{R}_+^k$ to $\alpha^* \in \overset{\circ}{\mathbb{R}}_+^k$, where $\overset{\circ}{\mathbb{R}}_+^k$ is the symbol for the (strictly) positive orthant $\overset{\circ}{\mathbb{R}}_+^k := \{\alpha \in \mathbb{R}^k \mid \alpha_i > 0 \forall i \in \{1, \dots, k\}\}$. The following theorem clarifies, under which circumstances this zero manifold is a $(k-1)$ -dimensional differentiable manifold.

² When distinguishing clearly between row and column vectors the correct expression would be $(\mathbf{x}^{*T}, \boldsymbol{\lambda}^{*T}, \alpha^{*T})^T$. In order not to overload our notation, in such cases we shall do without the transposition symbol.

Theorem 5.1:

Let M be defined as $M := \{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in \mathbb{R}^{n+m+k} \mid \mathbf{F}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) = \mathbf{0} \wedge \boldsymbol{\alpha}^* \in \overset{\circ}{\mathbb{R}}_+^k\}$.
If for all points of M the rank condition

$$\text{rank } \mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) = n + m + 1 \quad (5.6)$$

is fulfilled, where $\mathbf{F}'$ is the Jacobian matrix of $\mathbf{F}$, then M is a $(k-1)$ -dimensional differentiable submanifold of the $\mathbb{R}^{n+m+k}$.

Proof. According to the definition of a differentiable submanifold (see Section 4.2, point (b) and [FORSTER, 1984]) the claim for M is correct, if for every point $\mathbf{a} \in M$ there exists an open neighborhood $U \subset \mathbb{R}^{n+m+k}$ and a continuously differentiable function $\Phi : U \rightarrow \mathbb{R}^{n+m+1}$, so that the following is valid:

- (i) $M \cap U = \{z \in U \mid \Phi(z) = \mathbf{0}\}$
- (ii) $\text{rank } \Phi'(\mathbf{a}) = n + m + 1$

By limiting $\boldsymbol{\alpha}^*$ to the positive orthant one ensures that there really exists an open neighborhood U with the property (i). The other requirements follow directly from the definition of M with $\Phi = \mathbf{F}$. ■

If the requirement that the Rank Condition (5.6) must be valid for all points of M is weakened to the requirement that $\mathbf{F}'$ in **one** point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ must have the full rank, the assertion of Theorem 5.1 is nonetheless still valid in a ‘local version’.

Theorem 5.2:

Let all premises of Theorem 5.1 be fulfilled except for the requirement that the Rank Condition (5.6) is to be met by all points of M . Let furthermore a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ be given which complies with the Condition (5.6).

Then there exists an open neighborhood $U \subset \mathbb{R}^{n+m+k}$ of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, so that $M \cap U$ is a $(k-1)$ -dimensional differentiable submanifold of the $\mathbb{R}^{n+m+k}$.

Proof. The full rank of $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ implies that a $(n+m+1) \times (n+m+1)$ -submatrix $\mathbf{A}$ of $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ exists with $\det \mathbf{A} \neq 0$. Because of the continuity of $\mathbf{F}'$ also $\det \mathbf{A}$ is a continuous function and $\det \mathbf{A} \neq 0$ is valid for an entire open neighborhood U of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$. Consequently, the Rank Condition (5.6) is satisfied for all points of U . ■

5.2 Criteria for the Rank Condition

In this section we will first (Paragraph 5.2.1) elaborate in the form of Theorem 5.3 a necessary and sufficient criterion for the full rank $n + m + 1$ of the Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ in a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$. Subsequently, this criterion will be illustrated in Paragraph 5.2.2 by showing that by means of some corollaries a connection between the fulfillment of the Rank Condition and the character of the point $\mathbf{x}^*$ with respect to scalar-valued optimization can be made — remember that $\mathbf{x}^*$ is a Karush-Kuhn-Tucker point in the scalar-valued optimization problem with the objective function $g_{\boldsymbol{\alpha}^*}$ and the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. In the end of this section (Paragraph 5.2.3) we will make some observations about the connection between the character of the KKT-point $\mathbf{x}^*$ and the unrestricted variability of the weight vector $\boldsymbol{\alpha}$ in the neighborhood of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$.

5.2.1 A Necessary and Sufficient Criterion

First we have to make some preparations for Theorem 5.3.

We assume that the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ in the point $\mathbf{x}^*$ satisfy the mentioned constraint qualification, i.e. that the vectors $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$ are linearly independent. Under this condition in a neighborhood of $\mathbf{x}^*$ the m equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ define an $(n - m)$ -dimensional submanifold of the $\mathbb{R}^n$, which is also called constraint surface. Its tangent plane is an $(n - m)$ -dimensional linear subspace of the $\mathbb{R}^n$ which can be written as the orthogonal complement $S^\perp$ of the subspace $S \subset \mathbb{R}^n$ defined by $\text{span}\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_{n-m}\}$, where $\mathbf{v}_i \in \mathbb{R}^n$, be an orthonormal basis of $S^\perp$, and denote the $n \times (n - m)$ -matrix which is made up of these basis vectors by $\mathbf{V} := (\mathbf{v}_1 \dots \mathbf{v}_{n-m})$.

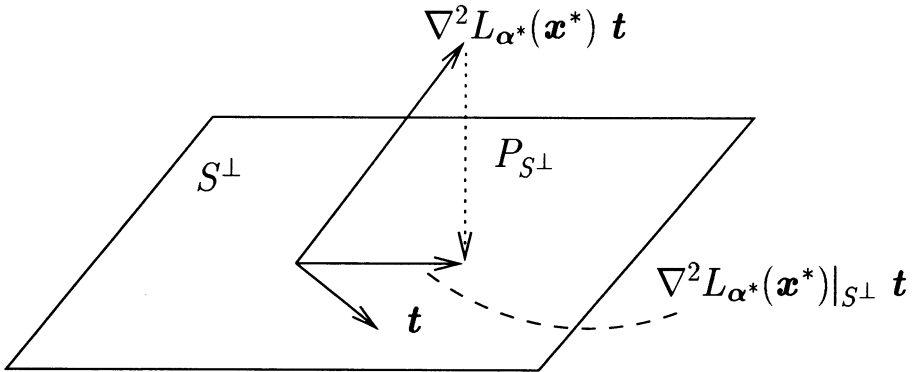


Figure 5.1: Illustration of the linear mapping $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$.

The Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, the rank of which is under investigation, has an important $(n \times n)$ -submatrix (see below, Equation (5.10)):

$\nabla^2 L_{\alpha^*}(\mathbf{x}^*) := \nabla_{\mathbf{x}}^2 (\alpha^{*T} \mathbf{f}(\mathbf{x}) + \boldsymbol{\lambda}^{*T} \mathbf{h}(\mathbf{x}))|_{\mathbf{x}=\mathbf{x}^*}$, i.e. the Hessian matrix (with regard to $\mathbf{x}$) of the Lagrangian function $L_{\alpha^*}(\mathbf{x}, \boldsymbol{\lambda})|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^*}$ to the scalar-valued objective function g_{α^*} . If one restricts the linear mapping of the $\mathbb{R}^n$ into the $\mathbb{R}^n$, which is given by this matrix, to the subspace $S^\perp$, i.e. to the tangent space of the constraint surface, one obtains the linear mapping $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ defined by

$$\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp} : \begin{cases} S^\perp & \rightarrow S^\perp \\ \mathbf{t} & \mapsto P_{S^\perp} \nabla^2 L_{\alpha^*}(\mathbf{x}^*) \mathbf{t} \end{cases}, \quad (5.7)$$

where $P_{S^\perp}$ denotes the projection mapping onto the subspace $S^\perp$ (see Figure 5.1). The matrix representation of this linear mapping $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ with regard to the basis $\{\mathbf{v}_1, \dots, \mathbf{v}_{n-m}\}$ of the subspace $S^\perp$ is $\mathbf{V}^T \nabla^2 L_{\alpha^*}(\mathbf{x}^*) \mathbf{V}$. For this matrix representation we have

$$(\mathbf{V}^T \nabla^2 L_{\alpha^*}(\mathbf{x}^*) \mathbf{V})^T = \mathbf{V}^T (\nabla^2 L_{\alpha^*}(\mathbf{x}^*))^T \mathbf{V} = \mathbf{V}^T \nabla^2 L_{\alpha^*}(\mathbf{x}^*) \mathbf{V}, \quad (5.8)$$

because $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)$, as a Hessian matrix, is symmetrical. Therefore, $\mathbf{V}^T \nabla^2 L_{\alpha^*}(\mathbf{x}^*) \mathbf{V}$ is a symmetrical matrix. Consequently $S^\perp$ can be spanned by an orthonormal basis consisting of eigenvectors of $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$, and all the eigenvalues $\nu_1, \dots, \nu_{n-m}$ are real numbers.

Now we are in the position to state a necessary and sufficient condition³ for the fulfillment of the Rank Condition (5.6) in the form of the following theorem:

Theorem 5.3:

Consider a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \alpha^*) \in M$ (for the definition of M see Theorem 5.1), i.e. let $\mathbf{x}^*$ be a Karush-Kuhn-Tucker point to the scalar-valued optimization problem with the objective function g_{α^*} and the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Let the constraint qualification be fulfilled in $\mathbf{x}^*$, i.e. the vectors $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$ are linearly independent.

The subspace $\text{span}\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\} \subset \mathbb{R}^n$ is denoted by S and its orthogonal complement by $S^\perp$.

Then the following equivalence holds:

$$\text{rank } \mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \alpha^*) = n + m + 1 \iff$$

The set of the vectors $\mathbf{u} \in S^\perp \subseteq \mathbb{R}^n$ with $\mathbf{u} \neq \mathbf{0}$, for which the following is true:

- $\mathbf{u} \in \{\text{eigenraum of } \nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp} \text{ associated with the eigenvalue } 0\}$

and

- $\mathbf{u} \in \text{kernel } \mathbf{f}'(\mathbf{x}^*)$

³ At this point the author would like to thank Prof. Dr. Klaus Ritter for his suggestion of extending the claim of Theorem 5.3 to the generality presented here (see [RITTER, 1998]).

is the void set.

[or equivalent to this: the intersection (which shall be denoted by E) of the linear spaces $\text{kernel } \mathbf{f}'(\mathbf{x}^*) \subseteq \mathbb{R}^n$ and $\{\text{eigenraum of } \nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp} \text{ associated with the eigenvalue } 0\} \subseteq S^\perp \subseteq \mathbb{R}^n$ consists of the zero vector.]

Proof. First, we note that $\text{rank } \mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) = n + m + 1$ is equivalent to saying that the linear equation

$$\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)^T \mathbf{z} = \mathbf{0} \quad (5.9)$$

has only the trivial solution $\mathbf{z}_0 = \mathbf{0} \in \mathbb{R}^{n+m+1}$.

Proof of $\Leftarrow$:

We start by a closer look at (5.9). Differentiation of the function $\mathbf{F}$ results in

$$\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)^T = \left(\begin{array}{c|ccc|c} \nabla^2 L_{\alpha^*}(\mathbf{x}^*) & \nabla h_1(\mathbf{x}^*) & \dots & \nabla h_m(\mathbf{x}^*) & \mathbf{0} \\ \hline \nabla h_1(\mathbf{x}^*)^T & & & & \\ \vdots & & \mathbf{0} & & \mathbf{0} \\ \nabla h_m(\mathbf{x}^*)^T & & & & \\ \hline \nabla f_1(\mathbf{x}^*)^T & & & & 1 \\ \vdots & & \mathbf{0} & & \vdots \\ \nabla f_k(\mathbf{x}^*)^T & & & & 1 \end{array} \right). \quad (5.10)$$

Writing a solution vector $\mathbf{z} \in \mathbb{R}^{n+m+1}$ as $\mathbf{z} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \\ c \end{pmatrix}$ with $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $c \in \mathbb{R}$, we get a linear system of equations which is equivalent to (5.9):

$$\nabla^2 L_{\alpha^*}(\mathbf{x}^*) \mathbf{a} = -\mathbf{H} \mathbf{b}, \quad (5.11)$$

$$\text{where } \mathbf{H} := (\nabla h_1(\mathbf{x}^*) \dots \nabla h_m(\mathbf{x}^*)) \in \mathbb{R}^{n \times m}$$

$$\nabla h_1(\mathbf{x}^*)^T \mathbf{a} = 0 \quad (5.12)$$

$$\vdots$$

$$\nabla h_m(\mathbf{x}^*)^T \mathbf{a} = 0$$

$$\nabla f_1(\mathbf{x}^*)^T \mathbf{a} = -c \quad (5.13)$$

$$\vdots$$

$$\nabla f_k(\mathbf{x}^*)^T \mathbf{a} = -c.$$

The Equations (5.12) are equivalent to the statement

$$\mathbf{a} \in S^\perp, \quad (5.14)$$

and the Equations (5.13) lead to the conclusion

$$\nabla g_{\alpha^*}(\mathbf{x}^*)^T \mathbf{a} \equiv \sum_{i=1}^k \alpha_i^* \nabla f_i(\mathbf{x}^*)^T \mathbf{a} = -c \left(\sum_{i=1}^k \alpha_i^* \right) = -c. \quad (5.15)$$

Since $\mathbf{F}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) = \mathbf{0}$, we have $\nabla g_{\boldsymbol{\alpha}^*}(\mathbf{x}^*) \in S$. On the other hand, $\mathbf{a} \in S^\perp$, so we conclude

$$c = 0 . \quad (5.16)$$

Thus, the Equations (5.13) imply that $\mathbf{f}'(\mathbf{x}^*) \mathbf{a} = \mathbf{0}$ or, equivalently,

$$\mathbf{a} \in \text{kernel } \mathbf{f}'(\mathbf{x}^*) . \quad (5.17)$$

The columns of $\mathbf{H}$ span the subspace $\text{image } \mathbf{H} \subset \mathbb{R}^n$. On the other hand, they form a basis of S thanks to the constraint qualification. Thus, we conclude $\text{image } \mathbf{H} = S$. Application of the projector $P_{S^\perp}$ to the equation (5.11) therefore results in

$$P_{S^\perp} \nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*) \mathbf{a} = -P_{S^\perp} \mathbf{H} \mathbf{b} = \mathbf{0} . \quad (5.18)$$

Since $\mathbf{a}$ is a vector of $S^\perp$, this is equivalent to

$$\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp} \mathbf{a} = \mathbf{0} , \quad (5.19)$$

i.e. $\mathbf{a}$ is an eigenvector belonging to the eigenvalue 0 of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$.

Having provided the material needed, we now assume that the intersection E as defined in the proposition is empty and that the equation (5.9) has a non-trivial solution $\mathbf{z} \neq \mathbf{0}$. The latter assumption implies $\mathbf{a} \neq \mathbf{0}$, since $\mathbf{a} = \mathbf{0}$ would lead to $\mathbf{b} = \mathbf{0}$ [according to (5.11) the vector $-\mathbf{b}$ can be considered as the vector of coefficients of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*) \mathbf{a} \in S$ with respect to the basis $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$ of S], and since $c = \mathbf{0}$ holds anyway. Thus, $\mathbf{a} \in S^\perp$ is an element of E conflicting with the assumption that E is empty. Hence, the $\Leftarrow$ -direction is proven.

Proof of $\Rightarrow$:

We suppose that (5.9) has only the trivial solution $\mathbf{z}_0 = \mathbf{0}$ and that there is a vector $\mathbf{u} \in E$, $\mathbf{u} \neq \mathbf{0}$. Consider the vector $\mathbf{z} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ 0 \end{pmatrix} \in \mathbb{R}^{n+m+1}$, where $-\mathbf{v}$ is the (uniquely determined) vector of coefficients of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*) \mathbf{u} \in S$ with respect to the basis $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$ of S . By construction, the triple $\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ 0 \end{pmatrix}$ solves the system of equations (5.11), (5.12), and (5.13). Therefore, $\mathbf{z}$ is a non-trivial solution of (5.9). Due to this contradiction, the assumption that there is a non-trivial vector $\mathbf{u} \in E$ cannot be true. ■

5.2.2 Interpretation in View of Optimization

At first sight the criterion for the fulfillment of the Rank Condition (5.6), contained in Theorem 5.3, looks rather abstract. We shall now fill it with life, deriving, by means of a few corollaries, the fulfillment of the Rank Condition (5.6) from the character of the KKT-point $\mathbf{x}^*$ of the objective function $g_{\boldsymbol{\alpha}^*}$. Let us first state the following corollary:

Corollary 5.4:

Let a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ be given, i.e. let $\mathbf{x}^*$ be a Karush-Kuhn-Tucker point to the scalar-valued optimization problem with the objective function $g_{\boldsymbol{\alpha}^*}$ and the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Furthermore, assume the constraint qualification to be fulfilled in $\mathbf{x}^*$, i.e. the vectors $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$ to be linearly independent.

We conclude:

If the linear mapping $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ is regular, $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ has the full rank $n + m + 1$.

Proof. Since the linear mapping $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ is diagonalizable, its regularity is equivalent to the statement that for all eigenvalues $\nu_1, \dots, \nu_{n-m}$ we have:

$$\nu_i \neq 0 \quad \forall i = 1, \dots, n - m. \quad (5.20)$$

Consequently, there exists no non-trivial eigenvector to the eigenvalue 0. Therefore, according to Theorem 5.3, $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ has the full rank. ■

The sufficient criterion (for the rank condition) given in Corollary 5.4 obtains an immediate meaning, once one adopts the view of the scalar-valued optimization.

First let us point out that the affiliation of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ to the manifold M means that $\mathbf{x}^*$ meets the necessary condition of **first order** for a local extremal point of the objective function $g_{\boldsymbol{\alpha}^*}$ under the constraint $\mathbf{h}(\mathbf{x}) = \mathbf{0}$.

When classifying stationary points $\mathbf{x}^*$ of $g_{\boldsymbol{\alpha}^*}$ by means of information of **second order**, $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ (i.e. the Hessian matrix of the Lagrangian function, restricted to the tangent plane of the constraint surface) plays the same part as the Hessian matrix of the objective function in the unconstrained case (see e.g. [LUENBERGER, 1984]). The regularity of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$, which is according to Corollary 5.4 a sufficient criterion for the fulfillment of the Rank Condition (5.6), implies in the context of optimization that an analysis of the eigenvalues of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ (and hence an analysis of the information of second order) allows us to determine⁴ whether the Karush-Kuhn-Tucker point $\mathbf{x}^*$ is a local minimizer, a saddle point or a local maximizer of the objective function $g_{\boldsymbol{\alpha}^*}$ (subject to the constraint $\mathbf{h}(\mathbf{x}) = \mathbf{0}$). The distinction can be made by means of the sign of the eigenvalues:

- a) $\nu_i > 0 \quad \forall i = 1, \dots, n - m \iff \nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ is positively definite. In accordance with the sufficient optimality condition of second order $\mathbf{x}^*$ is an isolated minimizer of $g_{\boldsymbol{\alpha}^*}$ [subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$].
- b) $\exists (i, j) \in \{1, \dots, n - m\} \times \{1, \dots, n - m\} : \nu_i > 0, \nu_j < 0$. In this case $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ is indefinite (and regular), and the point $\mathbf{x}^*$ is a saddle point of $g_{\boldsymbol{\alpha}^*}$ [subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$].

⁴ This question can be decided from the eigenvalues of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ only if none of these eigenvalues is equal to zero. Otherwise, for the determination of the character of the stationary point $\mathbf{x}^*$ information of *higher order* is required.

- c) $\nu_i < 0 \forall i = 1, \dots, n - m$. Since this is equivalent to stating that $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ is negatively definite, $\mathbf{x}^*$ is an isolated maximizer of g_{α^*} [under $\mathbf{h}(\mathbf{x}) = \mathbf{0}$].

While case c) does not occur when one searches for efficient points — such points are situated on the border of $\mathbf{f}(R)$ which is opposite to the efficient set —, cases a) and b) lead to the following important corollary to Theorem 5.3:

Corollary 5.5:

Consider a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$, i.e. $\mathbf{x}^*$ is a Karush-Kuhn-Tucker point to the scalar-valued optimization problem with the objective function g_{α^*} and the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Furthermore, let the mentioned constraint qualification be satisfied in $\mathbf{x}^*$.

Let $\mathbf{x}^*$ further be either

- *a local minimizer of g_{α^*} meeting the sufficient optimality condition of second order, i.e. $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ is positively definite,*

or

- *a saddle point of g_{α^*} , such that $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ is regular and indefinite.*

Then $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ has the full rank $n + m + 1$. ■

Now one could ask in a colloquial way what happens during a transition from a minimizer to a saddle point on the set M . Let us examine the following scenario, which is shown schematically in Figure 5.2, in order to examine this question more closely.

Consider two points of M , $(\mathbf{x}^1, \boldsymbol{\lambda}^1, \boldsymbol{\alpha}^1)$ and $(\mathbf{x}^2, \boldsymbol{\lambda}^2, \boldsymbol{\alpha}^2)$. Let $\mathbf{x}^1$ be a local minimizer of the function g_{α^1} [subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$], which meets the sufficient optimality condition of second order, and let $\mathbf{x}^2$ be a saddle point of g_{α^2} [subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$] with the regular Hessian matrix $\nabla^2 L_{\alpha^2}(\mathbf{x}^2)|_{\text{span}\{\nabla h_1(\mathbf{x}^2), \dots, \nabla h_m(\mathbf{x}^2)\}^\perp}$. Let us assume moreover that it is possible to connect both points by a continuous curve

$$\boldsymbol{\tau} : \begin{cases} \mathbb{R} \supset [0, 1] & \rightarrow M \\ t & \mapsto (\mathbf{x}(t), \boldsymbol{\lambda}(t), \boldsymbol{\alpha}(t)) \\ & \text{with } \boldsymbol{\tau}(0) = (\mathbf{x}^1, \boldsymbol{\lambda}^1, \boldsymbol{\alpha}^1) \text{ and } \boldsymbol{\tau}(1) = (\mathbf{x}^2, \boldsymbol{\lambda}^2, \boldsymbol{\alpha}^2), \end{cases} \quad (5.21)$$

where for all $\mathbf{x}(t)$, $t \in [0, 1]$ the constraint qualification is valid. Stating to each curve point $\boldsymbol{\tau}(t)$ the $(n - m)$ -tuple formed by the eigenvalues of the linear mapping $\nabla^2 L_{\alpha(t)}(\mathbf{x}(t))|_{\text{span}\{\nabla h_1(\mathbf{x}(t)), \dots, \nabla h_m(\mathbf{x}(t))\}^\perp}$, one obtains a continuous curve $\tilde{\boldsymbol{\tau}} : t \mapsto (\nu_1(t), \dots, \nu_{n-m}(t))^T$ which corresponds to $\boldsymbol{\tau}$. By assumption one has: $\nu_i(0) > 0 \forall i \in \{1, \dots, n - m\}$ and $\exists j \in \{1, \dots, n - m\} : \nu_j(1) < 0$. Because of

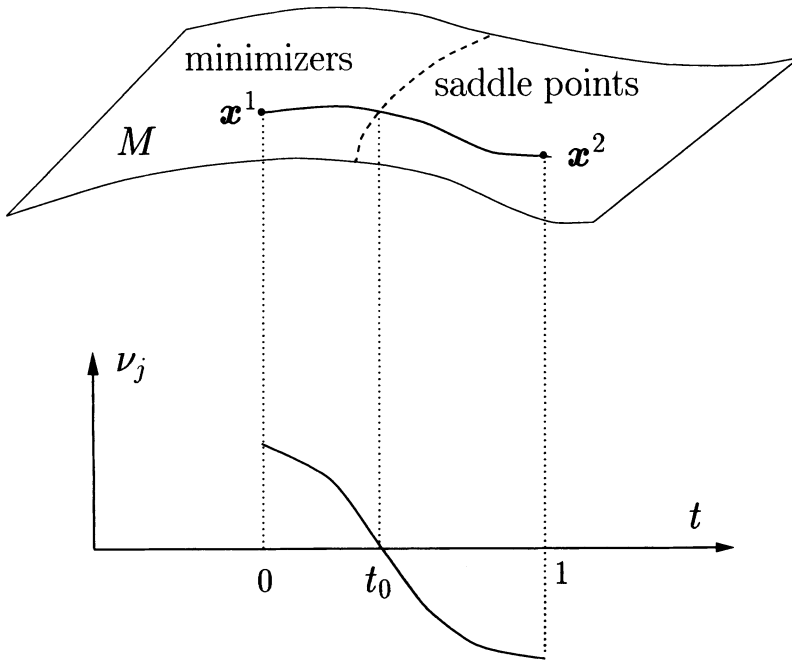


Figure 5.2 During the transition from a region of minimizers of g_α to a region of saddle points of g_α at least one eigenvalue ν_j of $\nabla^2 L|_{S^\perp}$ has to be zero.

the continuity of $\tilde{\tau}$ there must exist a curve parameter $t_0 \in [0, 1]$ with $\nu_j(t_0) = 0$. The point $\tau(t_0) \in M$ meets neither of the two conditions which, according to Corollary 5.5, are sufficient for the Rank Condition (5.6).

The points $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ of M , in which $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ has the eigenvalue 0 and which therefore do not meet either of the two sufficient conditions of Corollary 5.5, can be discussed from two different points of view:

- (A) From the point of view of **scalar-valued optimization** an eigenvalue 0 of the linear mapping $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ signifies that the local model of second (i.e. quadratic) order of the Lagrangian function $L_{\boldsymbol{\alpha}^*}(\mathbf{x}, \boldsymbol{\lambda})|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^*}$ (briefly denoted by $L_{\boldsymbol{\alpha}^*}$) in the direction of the corresponding eigenvector is flat and thus only models of higher order can give information about the behavior (increase or decrease) of $L_{\boldsymbol{\alpha}^*}$ along this direction.
For example, — like in the scenario described — when passing through the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ an eigenvalue of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ can pass through zero, which signifies that the curvature of the Lagrangian function with regard to this eigenvector changes its sign. If e.g. all eigenvalues (i.e. curvatures) were positive before, such a zero passing indicates that the character of the stationary point of g_α has changed: A minimum point has become a saddle point.
- (B) The second aspect is the question, whether in the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ the

Rank Condition (5.6) is met all the same. This viewpoint is significant in the context of this book for the following reason.

Our aim is to move around on the set M of candidates for efficient points by making use of a homotopy method. The method becomes particularly valuable, once it enables us to get from local minima to saddle points (and vice versa) in the sense of the above scenario. To this end, the feature (which for local minima and ‘regular’ saddle points is guaranteed by Corollary 5.5) of the zero manifold M of being — at least locally — a $(k - 1)$ -**dimensional differentiable manifold** must be valid also for those points $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, in which an eigenvalue of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ passes through zero (or several eigenvalues pass through zero simultaneously).

In the light of the above viewpoint now a further aspect of Theorem 5.3 discloses itself: According to the necessary and sufficient criterion given there, the Rank Condition (5.6) is fulfilled in points, in which the eigenraum belonging to the eigenvalue 0 of the linear mapping $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ does not disappear, if and only if none of the (non-trivial) vectors of this eigenraum is contained in the kernel of the mapping $\mathbf{f}'(\mathbf{x}^*)$.

If the eigenraum (associated with the eigenvalue 0) is one-dimensional, i.e. if only a single eigenvalue of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ in the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ passes through zero, this criterion can easily be verified: If the normalized eigenvector is denoted by $\mathbf{u}$, there has to exist at least one individual objective function f_i with $\nabla f_i(\mathbf{x}^*)^T \cdot \mathbf{u} \neq 0$.

We can summarize: When considering the zero transition of an eigenvalue of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ from the point of view of scalar-valued optimization, the relevant information is of higher than quadratic order. In contrast, the question, whether M in a neighborhood of $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is a differentiable manifold of well-defined dimension $k - 1$, is decided according to whether there is an individual objective function f_i , the **gradient** of which has a component along the eigenvector associated with the eigenvalue 0 of the **Hessian matrix** [of the Lagrangian function].

5.2.3 Variability of the Weight Vector

In the following paragraph we assume that there is an open neighborhood $U \subset \mathbb{R}^{n+m+k}$ of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$, so that $M \cap U$ is a $(k - 1)$ -dimensional differentiable submanifold of the $\mathbb{R}^{n+m+k}$. Now we will discuss the question, under which premises this manifold is parametrizable in a (possibly more restricted) neighborhood of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ by the components of the weight vectors $\boldsymbol{\alpha}$. In other words: Under which premises can $\boldsymbol{\alpha}$ be locally varied without restrictions?

The condition $\sum_{i=1}^k \alpha_i = 1$, which $\boldsymbol{\alpha}$ has to fulfill and which can be written in an equivalent way as $\boldsymbol{\alpha}^T \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 1$, determines a plane in the $\mathbb{R}^k$. This plane

— or the part of it situated in $\overset{\circ}{\mathbb{R}}_+^k$ — can be parametrized by an arbitrary choice of $(k-1)$ components of α . The above question has to be asked more precisely: under which premises can $k-1$ (arbitrarily chosen) components α_i of α be varied freely in a neighborhood of α^* without leaving the manifold M ?

The following theorem gives an answer to this:

Theorem 5.6:

Consider a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$, i.e. $\mathbf{x}^*$ is a Karush-Kuhn-Tucker point for the scalar-valued optimization problem with the objective function g_{α^*} and the equality constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Let the constraint qualification be fulfilled in $\mathbf{x}^*$.

Let furthermore $U \in \mathbb{R}^{n+m+k}$ be an open neighborhood of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ such that $M \cap U$ is a $(k-1)$ -dimensional differentiable submanifold of the $\mathbb{R}^{n+m+k}$.

Then we have:

$M \cap U$ is parametrizable in an appropriate neighborhood $\tilde{U}$ of $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ by $k-1$ arbitrarily chosen components α_i of α , if and only if the linear mapping $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ is regular.

Proof. First we want to bring to mind that the j -th column of the Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is formed by $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})_j}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, i.e. the derivative of $\mathbf{F}$ with respect to the j -th component of the extended variable vector $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})$. According to the implicit-function theorem $M \cap U$ can be parametrized locally by the $k-1$ components α_i , if and only if the derivative with respect to one of these components α_i is not necessary for the completion of the rank of the Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$. For the full rank of the Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ $n+m+1$ linearly independent columns are required. It is therefore *necessary* for the above-mentioned local parametrizability of $M \cap U$, that the submatrix $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ of $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, formed out of the first $n+m$ columns (i.e. the derivatives with respect to $\mathbf{x}$ and $\boldsymbol{\lambda}$), has the full rank $n+m$. On the other hand, as one learns from the explicit form of $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)^T$ in Equation (5.10), all columns of $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ have a 0 as $(n+m+1)$ -th element, all columns $\frac{\partial \mathbf{F}}{\partial \alpha_i}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, however, a 1. Therefore, the rank of the submatrix $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is increased by 1 when we add an arbitrary column of $\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$. Hence, the full rank of $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is also *sufficient* for the aforesaid local parametrizability of $M \cap U$.

It remains to be shown that $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ has the full rank $n+m$, if and only if $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ is regular.

This proof can be executed to a large extent in analogy to the proof of Theorem 5.3 and will be stated in the following very briefly. In advance let us note that the $(n+m+1)$ -th row of $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is the zero vector, so that the rank of $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is identical with the rank of the symmetri-

cal $(n + m) \times (n + m)$ -submatrix $\frac{\partial \mathbf{F}_{1 \dots n+m}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, which one obtains by eliminating this last row.

Proof of ' $\Leftarrow$ ':

It suffices to show that from the existence of a non-trivial solution $\mathbf{z} \in \mathbb{R}^{n+m}$, $\mathbf{z} \neq \mathbf{0}$, of the equation

$$\frac{\partial \mathbf{F}_{1 \dots n+m}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \mathbf{z} = \mathbf{0} \quad (5.22)$$

the singularity of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ follows.

By introducing the notation $\mathbf{z} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$ with $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$ Equation (5.22) is equivalent to the system of equations built by the Equations (5.11) and (5.12). As in the proof of Theorem 5.3 one can therefore conclude for a solution of this system of equations that $\mathbf{a}$ is an element of the eigenraum of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ associated with the eigenvalue 0. Since $\mathbf{z} \neq \mathbf{0}$ implies $\mathbf{a} \neq \mathbf{0}$, from the existence of a non-trivial solution $\mathbf{z} \neq \mathbf{0}$ of Equation (5.22) one can conclude that $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ has a non-vanishing eigenraum associated with the eigenvalue 0 and is therefore singular.

Proof of ' $\Rightarrow$ ':

This assertion shall be proved by contradiction. Let Equation (5.22) have only the trivial solution $\mathbf{z} = \mathbf{0}$ and let us assume that $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ is singular, i.e. has a non-vanishing eigenraum belonging to the eigenvalue 0. Let us now choose a vector $\mathbf{u} \neq \mathbf{0}$ of this eigenraum and examine (as in the proof of Theorem 5.3) the vector $\mathbf{z} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \in \mathbb{R}^{n+m}$, where $-\mathbf{v}$ is the well-determined coefficient vector of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*) \mathbf{u} \in S$ with regard to the basis $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$ of the subspace S . By construction, $\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$ solves the system of Equations (5.11) and (5.12). This is a contradiction to the assumption that Equation (5.22) has only the trivial solution. ■

The above proof shows that when $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ [and therefore also $\frac{\partial \mathbf{F}}{\partial (\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$] is regular, an arbitrary column $\frac{\partial \mathbf{F}}{\partial \alpha_j}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ of the submatrix $\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ can be utilized to complete the rank of $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$. The chosen component α_j is not available for the local parametrization of $M \cap U$, according to the implicit-function theorem. By choosing a component α_j [or, equivalently, by choosing the other $k - 1$ $\boldsymbol{\alpha}$ -components] one determines simultaneously, which $k - 1$ $\boldsymbol{\alpha}$ -components shall parametrize the plane $\boldsymbol{\alpha}^T \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 1$ in the $\mathbb{R}^k$.

If the linear mapping $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ is singular, in accordance with Theorem 5.6 $k - 1$ (arbitrarily chosen) components of the weight vector $\boldsymbol{\alpha}$ are no longer freely variable. In order to examine this limitation of the variability of $\boldsymbol{\alpha}$ more closely, let us assume that the eigenraum of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$ associated with the eigenvalue 0 is one-dimensional and is spanned by the vector $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0}$. As one can infer from the proof of Theorem 5.3, we then have for the kernel of the

mapping $\frac{\partial \mathbf{F}_{1\dots n+m}}{\partial(\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$:

$$\text{kernel } \frac{\partial \mathbf{F}_{1\dots n+m}}{\partial(\mathbf{x}, \boldsymbol{\lambda})}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) = \text{span} \left\{ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \right\}, \quad (5.23)$$

where $-\mathbf{v} \in \mathbb{R}^m$ is the well-determined coefficient vector of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*) \mathbf{u} \in S$ with regard to the basis $\{\nabla h_1(\mathbf{x}^*), \dots, \nabla h_m(\mathbf{x}^*)\}$ of the subspace S . Consequently, the first $n+m$ columns of the Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ generate the $(n+m-1)$ -dimensional subspace $T := \text{span} \left\{ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \right\}^\perp \times \{0\}$ of the $\mathbb{R}^{n+m} \times \mathbb{R}$. We require (at least) two columns of the submatrix $\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ to complete the dimension of the span of the columns to $n+m+1$. The i -th column of $\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is $(\nabla f_i(\mathbf{x}^*), \mathbf{0}, 1)^T$, where $\mathbf{0} \in \mathbb{R}^m$. If one picks out two columns (the i -th and the j -th), they span the subspace $\text{span} \{(\nabla(f_i - f_j)(\mathbf{x}^*), \mathbf{0}, 0)^T, (\nabla(f_i + f_j)(\mathbf{x}^*), \mathbf{0}, 2)^T\}$. As in the second basis vector $(\nabla(f_i + f_j)(\mathbf{x}^*), \mathbf{0}, 2)^T$ of this subspace the $(n+m+1)$ -th component does not vanish, it is automatically not included in the sum space $(T + \text{span}\{(\nabla(f_i - f_j)(\mathbf{x}^*), \mathbf{0}, 0)^T\})$. The first basis vector $(\nabla(f_i - f_j)(\mathbf{x}^*), \mathbf{0}, 0)^T$ has a non-vanishing component in $T^\perp$ and is therefore not included in T , if and only if:

$$\begin{pmatrix} \nabla f_i(\mathbf{x}^*) - \nabla f_j(\mathbf{x}^*) \\ \mathbf{0} \\ 0 \end{pmatrix}^T \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ 0 \end{pmatrix} \equiv \nabla f_i(\mathbf{x}^*)^T \mathbf{u} - \nabla f_j(\mathbf{x}^*)^T \mathbf{u} \neq 0. \quad (5.24)$$

Hence, two columns i and j of $\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ [which contain the derivative of the function $\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})$ with respect to the components α_i or α_j of the weight vector $\boldsymbol{\alpha}$] can complete the rank of the Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, if and only if the gradients of the associated individual objectives f_i and f_j have different components in the direction of the eigenvector $\mathbf{u}$.

If, on the other hand, one tries to answer the question, *whether* and, if so, *which* $k-2$ components of $\boldsymbol{\alpha}$ are freely variable (under the above assumption of a one-dimensional eigenraum belonging to the eigenvalue 0 of $\nabla^2 L_{\boldsymbol{\alpha}^*}(\mathbf{x}^*)|_{S^\perp}$), one obtains: A choice of $k-2$ $\boldsymbol{\alpha}$ -components can be varied (locally) freely, if and only if the gradients of both individual objectives, which correspond to the other two $\boldsymbol{\alpha}$ -components, have different components in the direction of the eigenvector $\mathbf{u}$.

The observations just made can be applied analogously to scenarios, in which the eigenraum associated with the eigenvalue 0 has a dimension larger than 1.

It was the aim of the considerations of this paragraph to show in which way limitations of the local variability of the weight vector $\boldsymbol{\alpha}$, which follow from certain curvature properties of the border of the image set $\mathbf{f}(R)$ (see Section 4.4, last paragraph) and which can also be observed numerically (see Section 7.2, Figure 7.9), are connected to the structure of the Jacobian matrix $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ and the rank properties of its submatrices. Both phenomena, i.e. the limitation

of the variability of α — which is induced by a change of curvature of the border of the image set $f(R)$ during the transition from minima to saddle points region (see Section 4.4) — as well as the collapse of the rank of the submatrix $\frac{\partial F}{\partial(\mathbf{x}, \lambda)}(\mathbf{x}^*, \lambda^*, \alpha^*)$, which according to the implicit-function theorem must accompany this limitation of the variability, have a common cause: the transition of an eigenvalue of the Hessian matrix $\nabla^2 L_{\alpha^*}(\mathbf{x}^*)|_{S^\perp}$ through zero.

From the above considerations a second important conclusion can be drawn. A comparison of Theorems 5.3 and 5.6 shows that the local parametrizability of M by α (or $k - 1$ of its components) is based on substantially stricter premises than the property of $M \cap U$ of being a $(k - 1)$ -dimensional differentiable manifold. When drawing up a homotopy method for vector optimization we are therefore not going to use α (or $k - 1$ of its components) directly as a $[(k - 1)$ -dimensional] homotopy parameter, but will develop a generalized method which is based on a parametrization that is realizable under the weakest possible assumption, namely the property of $M \cap U$ of being a $(k - 1)$ -dimensional differentiable manifold. The discussion of the basic principle of this method will be the subject of the following section.

5.3 A Special Class of Local Charts

Given a point $(\mathbf{x}^*, \lambda^*, \alpha^*)$ of the set M of candidates for Pareto optimal points which meets the Rank Condition (5.6), we want to investigate the neighborhood of this point on M . That is, we want to find other points of $M \cap U$, where $U \subset \mathbb{R}^{n+m+k}$ is an open neighborhood of $(\mathbf{x}^*, \lambda^*, \alpha^*)$.

In accordance with Theorem 5.2, U can be chosen in such a way that $M \cap U$ is a $(k - 1)$ -dimensional differentiable submanifold of $\mathbb{R}^{n+m+k}$. This property guarantees for $M \cap U$ the existence of a local chart. A local chart φ of $M \cap U$ is defined as a C^1 -homeomorphism $\varphi : T \rightarrow V$ which maps an open subset $T \subset \mathbb{R}^{k-1}$ onto an open neighborhood $V \subset (M \cap U) \subset \mathbb{R}^{n+m+k}$ of the point $(\mathbf{x}^*, \lambda^*, \alpha^*)$ and which meets the rank condition $\text{rank } \varphi'(\xi) = k - 1 \forall \xi \in T$ (see Section 4.2).

The basic idea of our approach is to construct an appropriate local chart φ of $M \cap U$ and to generate points of $M \cap U$ by varying the chart parameters ξ . Figure 5.3 schematically illustrates this plan. Let $\xi_0 := \varphi^{-1}(\mathbf{x}^*, \lambda^*, \alpha^*)$ denote the inverse image of $(\mathbf{x}^*, \lambda^*, \alpha^*)$ under the mapping φ . According to our plan, we generate a set of chart parameter points in the neighborhood of ξ_0 . In Figure 5.3 these points are denoted by $\{\xi_{(1)}, \xi_{(2)}, \xi_{(3)}, \xi_{(4)}\}$. The numerical evaluation of the mapping φ for these points will yield the new points $\{\varphi(\xi_{(1)}), \varphi(\xi_{(2)}), \varphi(\xi_{(3)}), \varphi(\xi_{(4)})\}$ on $M \cap U$.

The explicit numerical construction of an appropriate local chart φ will be the subject of Chapter 6. By scrutinizing the aim of exploring the local neighborhood of the point $(\mathbf{x}^*, \lambda^*, \alpha^*)$ on $M \cap U$, general guidelines for the construction of φ can be gained. These shall be discussed now.

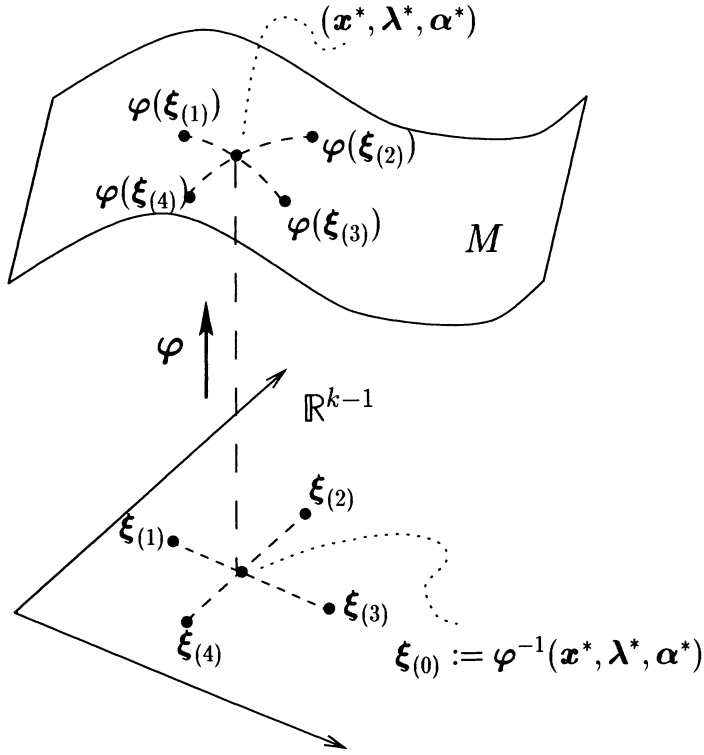


Figure 5.3 The basic idea of generating new points of $M \cap U$ by numerical evaluations of an appropriate local chart φ (see text).

- (i) The image set of φ has to be a neighborhood of the point $(x^*, \lambda^*, \alpha^*)$. Therefore, it is natural to demand that $(x^*, \lambda^*, \alpha^*)$ be the image of the parameter origin, i.e. that we have

$$\varphi(0) = (x^*, \lambda^*, \alpha^*). \quad (5.25)$$

Any arbitrary chart $\tilde{\varphi}$ can be brought into this form by translation.

- (ii) The chart φ has to be evaluated numerically. The following method for constructing φ permits us to apply the tools of numerical linear algebra effectively:

The space $\mathbb{R}^{n+m+k}$ is decomposed into a $(k-1)$ -dimensional linear subspace L and the orthogonal complement $L^\perp$ associated to it. Let us assume $\{q_1, \dots, q_{n+m+k}\}$ to be an orthonormal basis of the $\mathbb{R}^{n+m+k}$ such that $\text{span}\{q_1, \dots, q_{k-1}\} = L$ and $\text{span}\{q_k, \dots, q_{n+m+k}\} = L^\perp$. The chart φ now describes a point $(x, \lambda, \alpha) \in M \cap U$ as a function of its projection onto the subspace L which has been attached to the point $(x^*, \lambda^*, \alpha^*)$. Chart parameters are the coordinates ξ of the vector thus projected with regard to

the basis $\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$. Such a chart φ has the form

$$\varphi : \xi \mapsto (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q} \begin{pmatrix} \xi \\ \eta(\xi) \end{pmatrix}, \quad (5.26)$$

where $\mathbf{Q} := (\mathbf{q}_1 \dots \mathbf{q}_{n+m+k})$ is the orthogonal matrix constructed out of the basis vectors and η denotes a continuously differentiable mapping $\eta : \mathbb{R}^{k-1} \supseteq T \rightarrow \mathbb{R}^{n+m+1}$, with $\eta(\mathbf{0}) = \mathbf{0}$.

- (iii) The neighborhood V of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ on the manifold $M \cap U$ should be accessible to our exploration along all ‘directions’ without leaving $\varphi(T)$. The heuristic notion of a direction on V can be formalized naturally by means of a generalized local coordinate curve $\tau_t : [0, a) \rightarrow V$, $\gamma \mapsto \varphi(\gamma \cdot \mathbf{t})$, where $\mathbf{t} \in \mathbb{R}^{k-1}$, $\|\mathbf{t}\| = 1$, and $a \cdot \mathbf{t} \in \partial T$ (the boundary of T). Therefore, we are led to require that the infimum of the set of distances $\{\|\mathbf{p}\| \mid \mathbf{p} \in \partial T \subset \mathbb{R}^{k-1}\}$ between the origin $\mathbf{0} \in T$ and boundary points of T should be as large as possible.

In order to illustrate requirement (iii) we take as an example the one-dimensional manifold S^1 , i.e. the unit circle in the $\mathbb{R}^2$ centered at the origin, as shown in Figure 5.4.

Let us have a closer look at the point $(x^*, y^*) = (0, 1)^T$ and search for a parametrization of S^1 in the neighborhood of this point which satisfies the requirements (i) to (iii). A chart which clearly meets the requirements (i) and (ii) is given by

$$\varphi_{S^1} : \begin{cases} (-1, +1) \rightarrow \mathbb{R}^2 \\ x \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ \sqrt{1-x^2} - 1 \end{pmatrix} = \begin{pmatrix} x \\ \sqrt{1-x^2} \end{pmatrix} \end{cases}. \quad (5.27)$$

In this case the x -coordinate is the chart parameter, the vectors $\mathbf{q}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{q}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ constitute the orthonormal basis, the matrix $\mathbf{Q}$ is the identity matrix, and the function η is defined as $\eta(x) = \sqrt{1-x^2} - 1$.

In order to verify whether φ_{S^1} meets also requirement (iii), one has to take into consideration the borders of the domain of definition T of this chart. These borders are characterized by the divergence of the derivative $\frac{d}{dx}\eta(x) = \frac{-x}{\sqrt{1-x^2}}$ in the points $x = -1$ and $x = +1$ (see also Figure 5.4). Requirement (iii) is indeed met, as both borders of the domain of definition are equally distant⁵ from the parameter origin $x = 0$.

If one asks for the reason of this property, one realizes the following particularity of

⁵ Because of the constant curvature of the circle — a special property of this example — the parameter interval T for all charts, which have the form (5.26), is of equal total length, namely 2. Therefore the verification of (iii) in this special case is identical with the verification of the symmetrical position of T with regard to the origin.

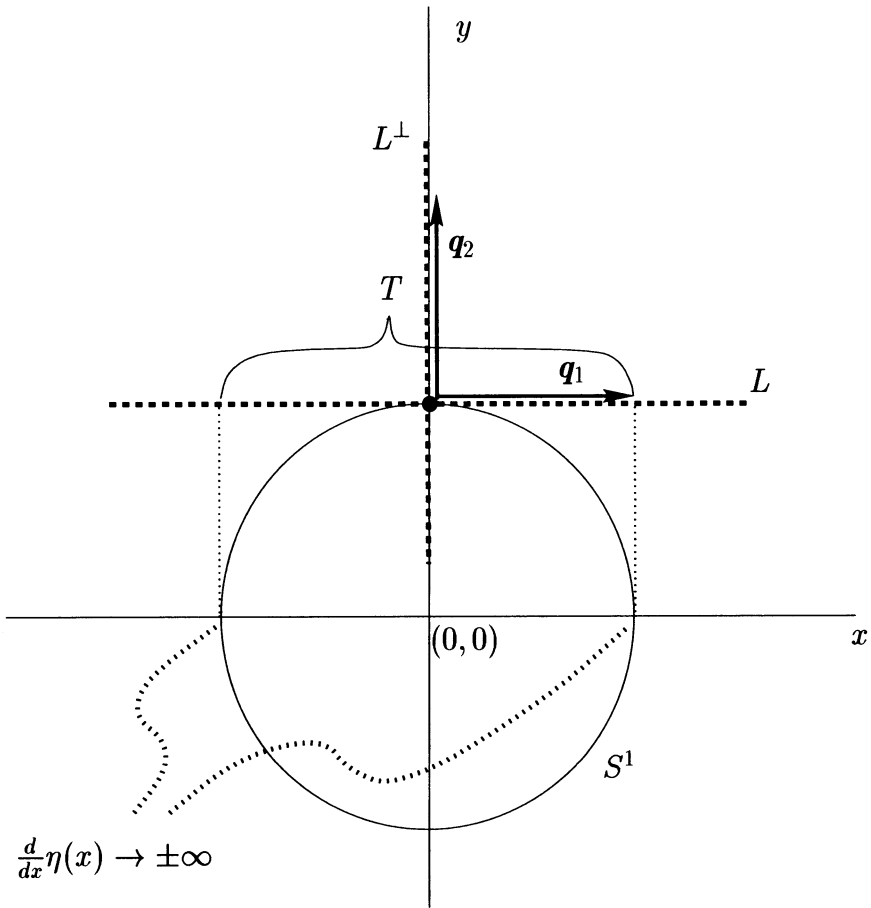


Figure 5.4 A local chart of the unit circle S^1 . The domain of definition T is limited by divergencies of $\frac{d}{dx}\eta(x)$.

the chart φ_{S^1} : The derivative $\frac{d}{dx}\eta(x) = \frac{-x}{\sqrt{1-x^2}}$, which diverges in the border points, has the value zero in the parameter origin. If one enlarges the notion of distance intuitively to $\mathbb{R} \cup \{+\infty, -\infty\}$, in the origin the derivative $\frac{d}{dx}\eta(x)$ therefore has a ‘maximum distance’ from $+\infty$ and $-\infty$, the values, to which it tends at the borders. If — as is the case in our problem — one has no knowledge of the curvature properties of the manifold $M \cap U$, this is the best measure one can take to fulfill requirement (iii).

When we apply the result of the above discussion to the case of a general chart φ of the form (5.26), a consequence of requirement (iii) is the additional constraint

$$\frac{\partial \eta}{\partial \xi}(0) = 0 \quad (5.28)$$

on the Jacobian matrix of η .

Before we shall prove that a chart with the properties (5.26) and (5.28) really exists, we go into an important implication of requirement (5.28). This constraint determines the subspace L (see point (ii)), which underlies the construction of the chart φ . To see that, let us have a look at the columns $\frac{\partial \varphi}{\partial \xi_1}(\mathbf{0}), \dots, \frac{\partial \varphi}{\partial \xi_{k-1}}(\mathbf{0})$ of the Jacobian matrix $\varphi'(\mathbf{0}) \equiv \frac{\partial \varphi}{\partial \xi}(\mathbf{0})$. These vectors form a basis of the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}(M \cap U)$ to the manifold $M \cap U$ in the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ (see Section 4.2, point (c)). If, on the other hand, one calculates $\frac{\partial \varphi}{\partial \xi_i}(\mathbf{0})$ by making use of the Equations (5.26) and (5.28), one gets

$$\frac{\partial \varphi}{\partial \xi_i}(\mathbf{0}) = \mathbf{Q} \frac{\partial}{\partial \xi_i} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta}(\boldsymbol{\xi}) \end{pmatrix}(\mathbf{0}) = \mathbf{Q} \mathbf{e}_i, \quad (5.29)$$

where the vector $\mathbf{e}_i \in \mathbb{R}^{n+m+k}$ at the i -th position has 1 and otherwise only zeros. Therefore $\mathbf{Q} \mathbf{e}_i$ is the i -th column of the matrix $\mathbf{Q}$. Since, by construction, the i -th column of $\mathbf{Q}$ is the vector $\mathbf{q}_i$, which lies in the subspace L and belongs to the orthonormal basis we use, one can conclude:

As a consequence of Equation (5.28) the basis $\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$ of the subspace L is at the same time also a basis of the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}(M \cap U)$, and the *span* of this basis, i.e. the subspace L , is identical with the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}(M \cap U)$. The chart parameters $\boldsymbol{\xi}$ of a point $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) \in (M \cap U)$ are hence the coordinates of the vector, which is generated by projecting $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})$ onto the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}(M \cap U)$, with regard to an orthonormal basis of this tangent plane [which has been attached to $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$]. Thus, the local chart φ is based on a coordinate system which is adapted to the local geometry of the manifold $(M \cap U)$. Figure 5.5 illustrates this crucial feature of the chart φ .

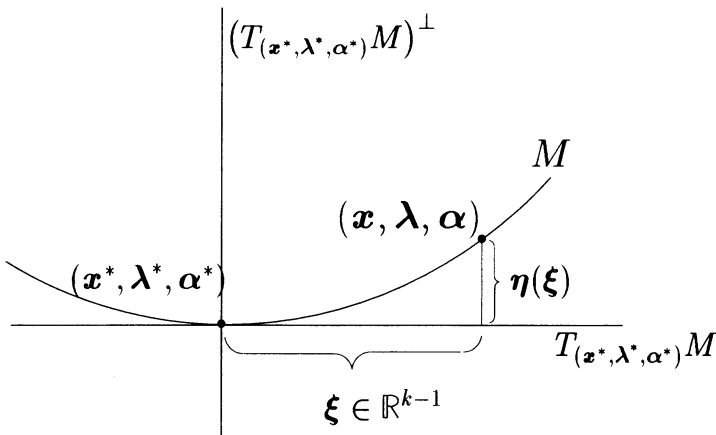


Figure 5.5 The decomposition of the $\mathbb{R}^{n+m+k}$ into the tangent space $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}(M \cap U)$ and its orthogonal complement enables the construction of a chart φ which is adapted to the local geometry of the manifold $M \cap U$.

The following theorem ensures the existence of such a local chart φ .

Theorem 5.7:

Consider a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ and assume that there exists an open neighborhood $U \subset \mathbb{R}^{n+m+k}$ of $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ such that $M \cap U$ is a $(k-1)$ -dimensional C^1 -submanifold of $\mathbb{R}^{n+m+k}$.

Let furthermore $\{\mathbf{q}_1, \dots, \mathbf{q}_{n+m+k}\}$ be an orthonormal basis of the $\mathbb{R}^{n+m+k}$ such that $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\} = T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}(M \cap U)$ [tangent plane to $M \cap U$ in the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$]. Let $\mathbf{Q} := (\mathbf{q}_1 \dots \mathbf{q}_{n+m+k})$ denote the orthogonal matrix formed by the basis vectors $\mathbf{q}_i$.

Then there exists an open neighborhood $T \subset \mathbb{R}^{k-1}$ of the origin $\mathbf{0} \in \mathbb{R}^{k-1}$, an open neighborhood V [relative to $(M \cap U)$] of the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ and a local chart of the form

$$\varphi: \begin{cases} T \rightarrow V \subset (M \cap U) \\ \boldsymbol{\xi} \mapsto (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta}(\boldsymbol{\xi}) \end{pmatrix} \end{cases}, \quad (5.30)$$

where:

$$\boldsymbol{\eta}(\mathbf{0}) = \mathbf{0} \quad \text{and} \quad \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\xi}}(\mathbf{0}) = \mathbf{0}. \quad (5.31)$$

Proof. First let us state that the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}(M \cap U)$ has the dimension $k-1$ and, therefore, a basis $\{\mathbf{q}_1, \dots, \mathbf{q}_{n+m+k}\}$ really exists, which has the properties of the assumption of the theorem.

Let $\bar{U} \subseteq U$ be a neighborhood of $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ such that $\bar{\alpha}_i > 0$ for all $i \in \{1, \dots, k\}$ and $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\alpha}}) \in \bar{U}$. The manifold $M \cap U$ to be parametrized is defined as the intersection of the zero manifold

$$\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = \mathbf{0} \quad (5.32)$$

with $\bar{U}$. Let $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})$ be an arbitrary point of $M \cap \bar{U}$ and let us denote the coordinates of $((\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) - (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*))$ with respect to the basis $\{\mathbf{q}_1, \dots, \mathbf{q}_{n+m+k}\}$ by $(\boldsymbol{\xi}, \boldsymbol{\mu})^T$, $\boldsymbol{\xi} \in \mathbb{R}^{k-1}$, $\boldsymbol{\mu} \in \mathbb{R}^{n+m+1}$, i.e.

$$(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\mu} \end{pmatrix}. \quad (5.33)$$

The inverse image of the neighborhood $\bar{U}$ with respect to this coordinate transformation is an open neighborhood $\tilde{U}$ of the origin in the space of the $(\boldsymbol{\xi}, \boldsymbol{\mu})$ -coordinates. A point of $\mathbb{R}^{n+m+k}$ solves the equation $\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = \mathbf{0}$, if and only if its $(\boldsymbol{\xi}, \boldsymbol{\mu})$ -coordinates solve the following equation:

$$\tilde{\mathbf{F}}(\boldsymbol{\xi}, \boldsymbol{\mu}) := \mathbf{F}(\mathbf{x}(\boldsymbol{\xi}, \boldsymbol{\mu}), \boldsymbol{\lambda}(\boldsymbol{\xi}, \boldsymbol{\mu}), \boldsymbol{\alpha}(\boldsymbol{\xi}, \boldsymbol{\mu})) = \mathbf{F}\left((\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\mu} \end{pmatrix}\right) = \mathbf{0}. \quad (5.34)$$

Describing the set of solutions of (5.34) by $\tilde{M} := \{(\xi, \mu) \in \mathbb{R}^{n+m+k} \mid \tilde{F}(\xi, \mu) = \mathbf{0}\}$, we can conclude that the coordinate transformation (5.33) establishes a diffeomorphism between the C^1 -manifolds $\tilde{M} \cap \tilde{U}$ and $M \cap \tilde{U}$.

Our next step is to construct a local chart of $\tilde{M} \cap \tilde{U}$.

The Jacobian matrix of $\tilde{F}$, evaluated at the point $(\xi, \mu) = \mathbf{0}$, is given by

$$\tilde{F}'|_{(\xi, \mu)=\mathbf{0}} \equiv \left. \frac{\partial \tilde{F}}{\partial (\xi, \mu)} \right|_{(\xi, \mu)=\mathbf{0}} = \mathbf{F}'(\mathbf{x}^*, \lambda^*, \alpha^*) \mathbf{Q}. \quad (5.35)$$

Let us examine the matrix $\mathbf{F}'(\mathbf{x}^*, \lambda^*, \alpha^*) = \begin{pmatrix} \nabla_{(\mathbf{x}, \lambda, \alpha)} F_1(\mathbf{x}^*, \lambda^*, \alpha^*)^T \\ \vdots \\ \nabla_{(\mathbf{x}, \lambda, \alpha)} F_{n+m+1}(\mathbf{x}^*, \lambda^*, \alpha^*)^T \end{pmatrix}$,

where $\nabla_{(\mathbf{x}, \lambda, \alpha)} \equiv (\frac{\partial}{\partial \mathbf{x}}, \frac{\partial}{\partial \lambda}, \frac{\partial}{\partial \alpha})$. Its rows form a base of the subspace $(T_{(\mathbf{x}^*, \lambda^*, \alpha^*)}(M \cap U))^\perp$. For any $l \in \{k, \dots, n+m+k\}$, the l -th column of the matrix $\mathbf{F}'(\mathbf{x}^*, \lambda^*, \alpha^*) \mathbf{Q}$ can be interpreted as the tuple of coefficients of the vector $\mathbf{q}_l \in [T_{(\mathbf{x}^*, \lambda^*, \alpha^*)}(M \cap U)]^\perp$ with regard to that basis $\{\nabla_{(\mathbf{x}, \lambda, \alpha)} F_1, \dots, \nabla_{(\mathbf{x}, \lambda, \alpha)} F_{n+m+1}\}$. As the linear independence of the vectors $\{\mathbf{q}_k, \dots, \mathbf{q}_{n+m+k}\}$ is preserved during this change of basis, the last $n+m+1$ columns of the matrix $\tilde{F}'|_{(\xi, \mu)=\mathbf{0}} = \mathbf{F}'(\mathbf{x}^*, \lambda^*, \alpha^*) \mathbf{Q}$ are linearly independent vectors, and we obtain

$$\text{rank} \left. \frac{\partial \tilde{F}}{\partial \mu} \right|_{(\xi, \mu)=\mathbf{0}} = n+m+1. \quad (5.36)$$

Therefore, according to the implicit-function theorem, there exist an open neighborhood $T \subset \mathbb{R}^{k-1}$ of the origin $\mathbf{0} \in \mathbb{R}^{k-1}$, an open neighborhood $W \subset \mathbb{R}^{n+m+1}$ of the origin $\mathbf{0} \in \mathbb{R}^{n+m+1}$, and a continuously differentiable function $\eta : T \rightarrow W$ such that the equation (5.34) has exactly one solution $(\xi, \mu) = (\xi, \eta(\xi))$ for each $\xi \in T$. Since the point $(\mathbf{0}, \mathbf{0})$ solves the system of Equations (5.34), we have $\eta(\mathbf{0}) = \mathbf{0}$.

The set $\tilde{V} := \tilde{M} \cap (T \times W)$ is an open neighborhood [relatively to $\tilde{M} \cap \tilde{U}$] of the origin $\mathbf{0} \in (\tilde{M} \cap \tilde{U}) \subset \mathbb{R}^{n+m+k}$. We choose the neighborhoods T and W small enough to ensure $T \times W \subset \tilde{U}$. As a consequence, $\tilde{V} \subset (\tilde{M} \cap \tilde{U})$, and the mapping

$$\tilde{\varphi} : \begin{cases} T & \rightarrow \tilde{V} \\ \xi & \mapsto \begin{pmatrix} \xi \\ \eta(\xi) \end{pmatrix} \end{cases} \quad (5.37)$$

is a local chart of the C^1 -manifold $\tilde{M} \cap \tilde{U}$. Composing $\tilde{\varphi}$ with the coordinate transformation (5.33) and defining V as the image of $\tilde{V}$ under this coordinate transformation, we obtain a mapping φ of the form (5.30) as a chart of the C^1 -manifold $M \cap U$.

In order to verify the second equation of (5.31), we write the formula for the Jacobian matrix of $\eta(\xi)$ at the point $\xi = \mathbf{0}$, which is supplied by the implicit-function theorem:

$$\frac{\partial \eta}{\partial \xi}(\mathbf{0}) = - \left(\frac{\partial \tilde{F}}{\partial \eta} \Big|_{(\xi, \eta(\xi))=\mathbf{0}} \right)^{-1} \frac{\partial \tilde{F}}{\partial \xi} \Big|_{(\xi, \eta(\xi))=\mathbf{0}} . \quad (5.38)$$

As a result of (5.35), the matrix $\frac{\partial \tilde{F}}{\partial \xi} \Big|_{(\xi, \eta(\xi))=\mathbf{0}}$ consists of the first $(k-1)$ columns of the matrix $F'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \mathbf{Q}$. By construction of the basis vectors $\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$ we have

$$\nabla_{(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})} F_i(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)^T \mathbf{q}_j = \mathbf{0} \quad \forall i \in \{1, \dots, n+m+1\}, j \in \{1, \dots, k-1\} , \quad (5.39)$$

i.e. these columns are all null vectors. Thus, Property (5.31) is proven. ■

Before we start to present the homotopy strategy in the following chapter, let us make a further remark concerning the feature $\frac{\partial \eta}{\partial \xi}(\mathbf{0}) = \mathbf{0}$ of our local chart φ . $\frac{\partial \eta}{\partial \xi}(\mathbf{0}) = \mathbf{0}$ is the decisive property of φ on which the homogeneous discretization of the Pareto set is based (see Paragraph 6.1.3 below).

Homotopy Strategies

In the present chapter we will develop a numerical method which enables us to generate neighboring points on the manifold M , starting from a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$, and thus to explore, step by step, the set of candidates for Pareto optimal points.

Section 5.3 has already outlined the strategy: The manifold M is parametrized locally by a chart φ . By specific variation of the chart parameters one determines, in which direction the exploration is to proceed on M [*procedural step 1*]. Subsequently, the function value of the chart φ , evaluated in the chosen parameter point, is determined numerically by a Newton method [*procedural step 2*]. This value of the function φ is nothing else than the wanted neighboring point of M . From the point of view of numerical mathematics this way of acting is a homotopy (or continuation) method generalized to a multidimensional homotopy parameter: procedural step 1 corresponds to the predictor, procedural step 2 to the corrector of the homotopy method.

From the viewpoint of the decision-maker there are two important application scenarios for this kind of homotopy method.

In scenario I a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ of the candidate set M is given and the decision-maker would like to get to know better a neighborhood ($\subset M$) of this point in all directions, in order to obtain a local overall picture of efficient solution alternatives.

Scenario II also starts from a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ on M . The weight vector $\boldsymbol{\alpha}^*$ gives information about the relative weight of the individual objectives, which is associated with the (candidate for a) Pareto optimal point $\mathbf{x}^*$. The decision-maker in scenario II now wants to know to where efficient solutions move, when the weight shifts in a definite direction characterised by a vector $\delta \bar{\boldsymbol{\alpha}} \in \mathbb{R}^k$.

The above homotopy concept is indeed usable in both application scenarios. In the two following sections we will develop made-to-order methods for scenarios I or II and cast each into a numerical algorithm.

6.1 Method I: Local Exploration of M

6.1.1 Method Principle

Let a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ be given, in which the Rank Condition (5.6) is fulfilled. According to the strategy outlined above, the set M in a neighborhood of $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is to be explored by choosing a set of points $\boldsymbol{\xi}_{(i)}$ out of the domain of definition $T \subset \mathbb{R}^{k-1}$ of the chart $\boldsymbol{\varphi}$ and by evaluating $\boldsymbol{\varphi}$ numerically at these points. The following two steps result in an evaluation $\boldsymbol{\varphi}(\boldsymbol{\xi}_{(i)})$ of the chart $\boldsymbol{\varphi}$.

- a) In the first step we determine the projection $\boldsymbol{\varphi}^P(\boldsymbol{\xi}_{(i)})$ of $\boldsymbol{\varphi}(\boldsymbol{\xi}_{(i)})$ onto the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}M$.

The chart $\boldsymbol{\varphi}$ is constructed in such a way that the chart parameter $\boldsymbol{\xi}$ of a point $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) \in M$ is formed by the coordinates of the vector which results from projecting $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})$ onto the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}M$ [attached at the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$] with regard to the basis $\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$ of this tangent plane. Therefore one can write immediately

$$\boldsymbol{\varphi}^P(\boldsymbol{\xi}_{(i)}) = (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + (\mathbf{q}_1 \dots \mathbf{q}_{k-1}) \boldsymbol{\xi}_{(i)}. \quad (6.1)$$

- b) Step 2 has to lead us directly to the manifold M , starting from the point $\boldsymbol{\varphi}^P(\boldsymbol{\xi}_{(i)})$ on the tangent plane to M . Because of $\boldsymbol{\varphi}(\boldsymbol{\xi}_{(i)}) \in M$, $\boldsymbol{\varphi}(\boldsymbol{\xi}_{(i)})$ solves the system of Equations (5.5), i.e. one has

$$\mathbf{F}(\boldsymbol{\varphi}(\boldsymbol{\xi}_{(i)})) = \mathbf{F}\left((\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q}\left(\begin{array}{c} \boldsymbol{\xi}_{(i)} \\ \boldsymbol{\eta}_{(i)} \end{array}\right)\right) = \mathbf{0}. \quad (6.2)$$

(6.2) is a system of $n + m + 1$ equations for the $n + m + 1$ unknown quantities $\boldsymbol{\eta}(\boldsymbol{\xi}_{(i)}) =: \boldsymbol{\eta}_{(i)}$ and has a solution due to the premise $\boldsymbol{\xi}_{(i)} \in T$ [remember: T denotes the domain of definition of the chart $\boldsymbol{\varphi}$]. To calculate this numerically, we make use of a Newton method (see e.g. [HÄMMERLIN & HOFFMANN, 1989] or [SCHWARZ, 1996]). The starting point is the value $\mathbf{0}$ of the $\boldsymbol{\eta}$ -coordinate of the predictor $\boldsymbol{\varphi}^P(\boldsymbol{\xi}_{(i)})$, i.e. $\boldsymbol{\eta}_{(i)}^{[0]} = \mathbf{0}$. The Newton method generates approximate solutions in an iterative way, which converge towards the wanted zero – if the starting point lies in the range of convergence of the method. The transition from the l -th approximate solution $\boldsymbol{\eta}_{(i)}^{[l]}$ to the $(l+1)$ -st approximate solution $\boldsymbol{\eta}_{(i)}^{[l+1]}$ is based on a linearization of the function $\tilde{\mathbf{F}}(\boldsymbol{\eta}_{(i)})$, which is defined as

$$\tilde{\mathbf{F}} : \begin{cases} \mathbb{R}^{n+m+1} & \rightarrow \mathbb{R}^{n+m+1} \\ \boldsymbol{\eta}_{(i)} & \mapsto \mathbf{F}\left((\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q}\left(\begin{array}{c} \boldsymbol{\xi}_{(i)} \\ \boldsymbol{\eta}_{(i)} \end{array}\right)\right) \end{cases}. \quad (6.3)$$

To this end, one develops $\tilde{\mathbf{F}}$ in a Taylor series around the point $\boldsymbol{\eta}_{(i)}^{[l]}$,

$$\tilde{\mathbf{F}}(\boldsymbol{\eta}_{(i)}) = \tilde{\mathbf{F}}(\boldsymbol{\eta}_{(i)}^{[l]}) + \tilde{\mathbf{F}}'(\boldsymbol{\eta}_{(i)}^{[l]}) \cdot (\boldsymbol{\eta}_{(i)} - \boldsymbol{\eta}_{(i)}^{[l]}) + o(\|\boldsymbol{\eta}_{(i)} - \boldsymbol{\eta}_{(i)}^{[l]}\|), \quad (6.4)$$

and breaks off the Taylor expansion after the linear terms. The zero of this linear approximation for $\tilde{\mathbf{F}}(\boldsymbol{\eta}_{(i)})$ is taken as the $(l+1)$ -th approximate solution $\boldsymbol{\eta}_{(i)}^{[l+1]}$. $\boldsymbol{\eta}_{(i)}^{[l+1]}$ is therefore determined by demanding that $\mathbf{v} := (\boldsymbol{\eta}_{(i)}^{[l+1]} - \boldsymbol{\eta}_{(i)}^{[l]})$ solves the linear system of equations in $\mathbf{z}$

$$\tilde{\mathbf{F}}'(\boldsymbol{\eta}_{(i)}^{[l]}) \cdot \mathbf{z} = -\tilde{\mathbf{F}}(\boldsymbol{\eta}_{(i)}^{[l]}). \quad (6.5)$$

By explicitly calculating the Jacobian matrix $\tilde{\mathbf{F}}'(\boldsymbol{\eta}_{(i)}^{[l]})$, one transforms Equation (6.5) into the equivalent system of equations

$$\mathbf{F}'\left((\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q}\left(\begin{array}{c} \boldsymbol{\xi}_{(i)} \\ \boldsymbol{\eta}_{(i)}^{[l]} \end{array}\right)\right) \cdot \left(\mathbf{q}_k \dots \mathbf{q}_{n+m+k}\right) \mathbf{z} = -\tilde{\mathbf{F}}(\boldsymbol{\eta}_{(i)}^{[l]}). \quad (6.6)$$

6.1.2 Comparison with the Classical Homotopy Method

We will now compare the procedure steps presented above with a classical homotopy method (see e.g. [SCHWETLICK, 1979], [GARCIA & ZANGWILL, 1981] and [ALLGOWER & GEORG, 1990]).

The classical homotopy method is an approach to the solution of systems of nonlinear equations. It is based on the idea of forging a link between the system of equations, the solution of which one searches for, and a system of equations, the solution of which one has on hand. Let $\mathbf{H}_0(\mathbf{y}) = \mathbf{0}$ [with $\mathbf{y} \in \mathbb{R}^l$ and $\mathbf{H}_0: \mathbb{R}^l \rightarrow \mathbb{R}^l$] be the system of equations with a known solution and $\mathbf{G}(\mathbf{y}) = \mathbf{0}$ [with $\mathbf{G}: \mathbb{R}^l \rightarrow \mathbb{R}^l$] the system of equations to be solved. A link is established by embedding¹ both systems of equations in a family $\mathbf{H}(\mathbf{y}, t) = \mathbf{0}$ of systems of equations, parametrized by the homotopy parameter $t \in \mathbb{R}$.

Assume that such an embedding has already been found in the form of a continuously differentiable function $\mathbf{H}: \mathbb{R}^{l+1} \rightarrow \mathbb{R}^l$ with the property $\mathbf{H}(\mathbf{y}, 0) = \mathbf{H}_0(\mathbf{y})$ and $\mathbf{H}(\mathbf{y}, 1) = \mathbf{G}(\mathbf{y})$. Once a point $(\mathbf{y}^*, t^*)$ is known which solves the embedding system of equations and for which the $l \times l$ -matrix $\frac{\partial \mathbf{H}}{\partial \mathbf{y}}(\mathbf{y}^*, t^*)$ is regular, the solutions of the system of equations in a neighborhood of $(\mathbf{y}^*, t^*)$ generate, according to the implicit-function theorem, a space curve $\boldsymbol{\tau}$ in the $\mathbb{R}^l \times \mathbb{R}$, $\boldsymbol{\tau}: t \mapsto (\mathbf{y}(t), t)$, which is parametrizable by t (and continuously differentiable).

¹ One possible form of embedding is a linear combination of the two functions $\mathbf{H}_0$ and $\mathbf{G}$: $\mathbf{H}(\mathbf{y}, t) := t \cdot \mathbf{G}(\mathbf{y}) + (1 - t) \cdot \mathbf{H}_0(\mathbf{y})$.

The classical homotopy methods start in the well-known solution point $(\mathbf{y}^*, t^* = 0)$ and construct numerically — for parameter values $t_{(i)} \equiv t^* + i \cdot \delta t$, successively augmented by $\delta t > 0$ — the points $\tau(t_{(i)})$ on this solution curve, in order to get to the curve point for $t = 1$, $\tau(1) = (\mathbf{y}(1), 1)$. Its $\mathbf{y}$ -component $\mathbf{y}(1)$ is the desired solution of the problem $\mathbf{G}(\mathbf{y}) = \mathbf{0}$. Let $\tau(t_{(i)})$ be the curve point calculated last, then the calculation of $\tau(t_{(i+1)})$ [i.e. the calculation of $\mathbf{y}(t_{(i+1)})$] is carried out in two stages (see Figure 6.1).

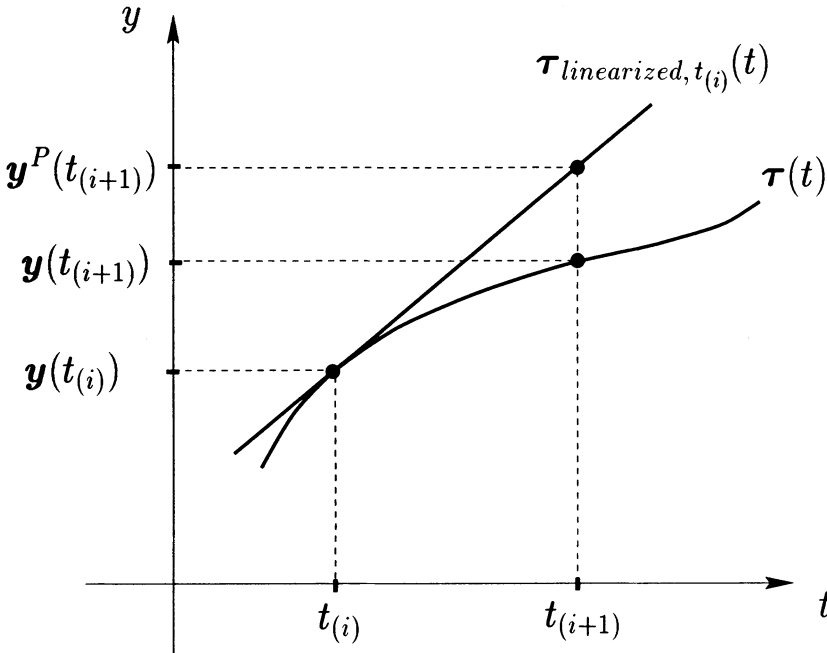


Figure 6.1 Prediction of the curve point $\tau(t_{(i+1)}) = (\mathbf{y}(t_{(i+1)}), t_{(i+1)})$ by linearizing the curve τ [i.e. by linearizing the implicitly defined mapping $\mathbf{y}(t)$] at the point $t = t_{(i)}$: $\mathbf{y}^P(t_{(i+1)}) = \mathbf{y}(t_{(i)}) + (t_{(i+1)} - t_{(i)}) \cdot \mathbf{y}'(t_{(i)})$, where according to the implicit-function theorem $\mathbf{y}'(t_{(i)})$ is given by $\mathbf{y}'(t_{(i)}) = - \left(\frac{\partial \mathbf{H}}{\partial \mathbf{y}}(\mathbf{y}(t_{(i)}), t_{(i)}) \right)^{-1} \frac{\partial \mathbf{H}}{\partial t}(\mathbf{y}(t_{(i)}), t_{(i)})$.

First, one calculates the tangent vector $\tau'(t_{(i)}) \equiv \frac{d\tau}{dt}|_{t=t_{(i)}}$ to the curve τ in the point $\tau(t_{(i)})$ and makes the straight line $\tau(t_{(i)}) + \beta \cdot \tau'(t_{(i)})$, $\beta \in \mathbb{R}$, intersect the plane defined by $t = t_{(i+1)}$ in the $\mathbb{R}^{l+1}$. The result is denoted by $\tau^P(t_{(i+1)}) = (\mathbf{y}^P(t_{(i+1)}), t_{(i+1)})$ [see Figure 6.1].

If one interprets the $\mathbf{y}$ -component $\mathbf{y}(t)$ of the space curve $\tau(t)$ as a solution of the differential equation in t (see [SCHWETLICK, 1979])

$$\dot{\mathbf{y}} = - \left(\frac{\partial \mathbf{H}}{\partial \mathbf{y}} \right)^{-1} \frac{\partial \mathbf{H}}{\partial t}, \quad (6.7)$$

then the stage of the method outlined above corresponds to a step of the Euler method for the numerical integration of this differential equation, where the steplength is chosen to be δt . Because the geometry of the solution curve $\tau(t)$ for values $t > t_{(i)}$ is, so to speak, predicted from the derivative information in the point $\tau(t_{(i)})$, this homotopy step is also called **predictor** step.

After that the error produced by the predictor step (i.e. the deviation of the point $\tau^P(t_{(i+1)})$ from the graph of the curve τ) has to be corrected in the so-called **corrector** step. One achieves this by making the result $\tau^P(t_{(i+1)})$ of the predictor step the start value of a Newton method. Since we have $l + 1$ unknown quantities, one has to add another equation to the system of equations $\mathbf{H}(\mathbf{y}, t) = \mathbf{0}$. In the classical homotopy method the additional equation $t = t_{(i+1)}$ is taken and an actual calculation of the curve point $\tau(t_{(i+1)})$ is carried out (see Figure 6.1). If one refrains from evaluating the curve τ exactly at the point $t_{(i+1)}$ of the homotopy parameter t , one can, for instance, add alternatively the equation $(\mathbf{y}, t) \cdot \tau'(t_{(i)}) = \tau^P(t_{(i+1)})^T \cdot \tau'(t_{(i)})$. By this one achieves that all iterations of the Newton method lie in the plane which passes through the predictor and which is orthogonal to the curve tangent $\tau'(t_{(i)})$. Thus, the corrector step (viewed as a step in the space) is orthogonal to the previous predictor step and tends to the graph of the curve τ .

Our way of looking at the problem of exploring the manifold M locally differs from the problem the classical homotopy methods start from, predominantly as regards the dimension of the respective zero manifolds. M has the dimension $k - 1$, whereas the homotopy curve τ can be interpreted as a one-dimensional manifold.

If in our way of looking at the problem we consider the special case $k - 1 = 1$ (i.e. the case of a bicriterial optimization problem), M becomes a curve τ , and the tangent plane at M becomes the span of the tangent vector at the curve τ . In this special case the procedural step a) described in Paragraph 6.1.1 corresponds to the predictor step of the classical homotopy method.

An important difference, however, consists in the way of parametrization.

The classical homotopy method described above has the aim of solving the system of equations for a given value of the homotopy parameter t determined in advance, therefore parametrizes the curve τ by t and, consequently, has to start from the rigorous assumption $\text{rank} \left(\frac{\partial \mathbf{H}}{\partial \mathbf{y}} \right) = l$. However, there may exist curve points, where the complete matrix $\frac{\partial \mathbf{H}}{\partial (\mathbf{y}, t)}$ has the full rank, i.e. where the zero manifold of the system of equations (in the following in a casual way also denoted by the 'curve τ ') is a locally differentiable one-dimensional manifold, but where the tangent vector at τ is orthogonal to the vector $(0, \dots, 0, 1)$ [i.e. to the t -axis]. As an example, Figure 6.2 shows a cuspidal point of the curve τ which has this property. In such points the submatrix $\frac{\partial \mathbf{H}}{\partial \mathbf{y}}$ is necessarily singular. Hence, a reparametrization is necessary and is indeed carried out within a strategy for cuspidal points in classical homotopy methods (see e.g. [SCHWETLICK, 1979]).

An example of such a change of parametrization is the exchange of the column $\frac{\partial \mathbf{H}}{\partial y_i}$, $i \in \{1, \dots, l\}$ in the submatrix $\frac{\partial \mathbf{H}}{\partial \mathbf{y}}$ against the column $\frac{\partial \mathbf{H}}{\partial t}$ of the complete Jacobian matrix $\frac{\partial \mathbf{H}}{\partial (\mathbf{y}, t)}$. If the submatrix thus generated is regular, y_i can be used as a ‘new’ local parameter of the curve τ .

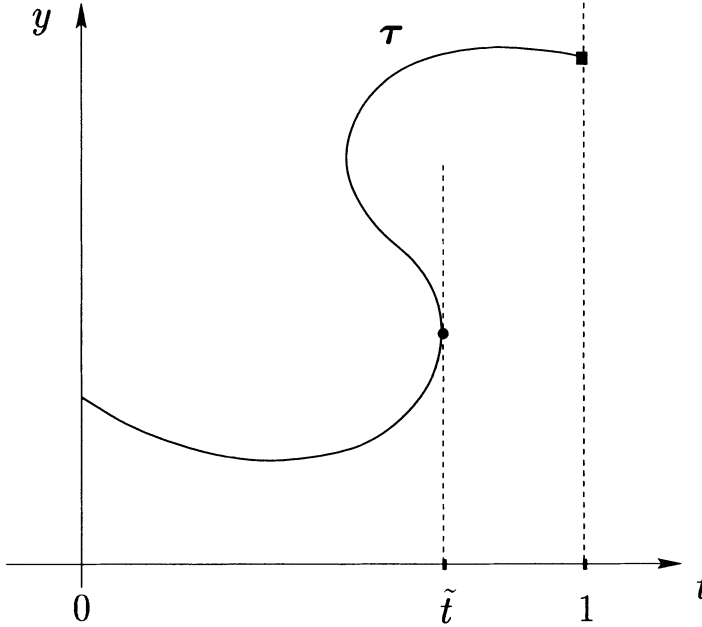


Figure 6.2 In the cuspidal point (marked by $\bullet$) the curve τ cannot be locally parametrized by t . Nevertheless, a strategy of reparametrization allows reaching finally the desired point (marked by $\blacksquare$) by homotopy.

Our method — considered in the special case $k = 2$ — makes a change of chart in every newly generated curve point and fits the parametrization (chart) constantly to the curve geometry. We demonstrated in Theorem 5.7 that such a parametrization requires only the weakest possible assumption $\text{rank} \left(\frac{\partial \mathbf{H}}{\partial (\mathbf{y}, t)} \right) = l$, which in any case is necessary for the local character of the zero manifold τ to be a differentiable one-dimensional manifold. We discussed in Section 5.3 that this choice of a chart is at the same time the best measure based on linear information about the geometry of τ to push away the borders of the domain of definition of the chart as far as possible from the current parameter point (which, put as an argument into the chart, produces the relevant curve point) and thus to obtain ‘maximal freedom of action’ for the next homotopy step. Let us note here that the choice of the chart in our method is related to the parametrization of the curve τ by its arc length, an approach well-known in literature, by means of which the calculation of points of the homotopy curve can be put down to the solution of the initial-value problem of an ordinary differential equation (see

e.g. [RAKOWSKA ET AL., 1991]).

The procedural step b) corresponds in the special case $k = 2$ to the corrector step of the classical homotopy method, if one adds the equation $(\mathbf{y}, t) \cdot \boldsymbol{\tau}'(t_{(i)}) = \boldsymbol{\tau}^P(t_{(i+1)})^T \cdot \boldsymbol{\tau}'(t_{(i)})$ (see above) to the system of equations $\mathbf{H}(\mathbf{y}, t) = \mathbf{0}$. However, the Newton method functioning as a corrector in the classical method acts in the $\mathbb{R}^{l+1}$, while in procedural step b) the Newton method acts — because of the constructed orthonormal basis — in the $\mathbb{R}^l$ (namely $\text{span}\{\text{tangent vector to } \boldsymbol{\tau}\}^\perp$).

Summarizing the result of the comparison just made, one can state the following. The construction of the orthonormal basis $\{\mathbf{q}_1, \dots, \mathbf{q}_{n+m+k}\}$ and the subsequent method steps a) and b) can be interpreted as a generalization of the classical predictor-corrector homotopy method. This generalization allows an application of that method to systems of equations which depend upon several parameters (so-called homotopy parameters). Looking at the zero manifold M from a differential-topological point of view, as discussed in chapter 5, one obtains almost automatically a parametrization of the points of M , which are to be generated by homotopy: Instead of being parametrized by the original homotopy parameters $\boldsymbol{\alpha}$, these points are parametrized by $k - 1$ coordinates with regard to a coordinate system fitted to the local geometry of the manifold M . The corresponding $k - 1$ coordinate axes span the tangent plane to M in the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, the neighborhood of which shall be explored.

6.1.3 Homogeneous Discretization of the Efficient Set

The user (decision-maker), who wants to obtain a survey of the set of efficient points, wants to have sufficient information about the mutual competition (i.e. the ‘trade-off’) of the individual objectives in all regions of interest. To get to this with the least effort possible, a method of vector optimization should be able to generate a homogeneous distribution of efficient points (in the objective space) or — in the ideal case — should be able to control this distribution in a simple way. For all parametric methods of multiobjective optimization this ability depends, of course, on the respective parametrization. For example, one does not succeed in generating a homogeneous distribution of efficient points by applying the weighting method, which parametrizes the efficient points by the corresponding weight vectors $\boldsymbol{\alpha}$ (see [DAS & DENNIS, 1996A]).

On the contrary, the parametrization in our method enables us in a simple way to control the local density of discretization of the efficient set. We will demonstrate this in the following.

A measure of the density of discretization is the distance (in the objective space) between two neighbored efficient points calculated by the method. Let us consider a situation, where the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ is already known and φ is a chart for a neighborhood of this point, constructed according to the rule of The-

orem 5.7. We choose a chart parameter vector $\xi_{(i)} := \delta_{(i)} \cdot \mathbf{e}_i$, where $\delta_{(i)} \in \mathbb{R}$ and $|\delta_{(i)}| \ll 1$ and where $\mathbf{e}_i$ denotes the i th unit vector in $\mathbb{R}^{k-1}$, and calculate the Euclidean distance ρ between the image points $\mathbf{f}(\mathbf{x}^*)$ and $\mathbf{f}(P_{\mathbf{x}} \varphi(\xi_{(i)}))$ in the objective space. Here, $P_{\mathbf{x}}$ denotes the projector onto the $\mathbf{x}$ -space, i.e. $P_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = \mathbf{x}$. The distance ρ can be made a function of $\delta_{(i)}$ by defining

$$\rho(\delta_{(i)}) := \|\mathbf{f}(\mathbf{x}^*) - \mathbf{f}(P_{\mathbf{x}} \varphi(\delta_{(i)} \cdot \mathbf{e}_i))\| \equiv \|\mathbf{f}(\mathbf{x}^*) - \tilde{\mathbf{f}}(\delta_{(i)})\|, \quad (6.8)$$

where the function $\tilde{\mathbf{f}}$ is defined as $\tilde{\mathbf{f}} : \mathbb{R} \rightarrow \mathbb{R}^k$, $\delta_{(i)} \mapsto \mathbf{f}(P_{\mathbf{x}} \varphi(\delta_{(i)} \cdot \mathbf{e}_i))$. On our way to computing the Taylor series expansion of $\rho(\delta_{(i)})$ near the point $\delta_{(i)}^0 = 0$, we first develop $\tilde{\mathbf{f}}(\delta_{(i)})$ around the point $\delta_{(i)}^0 = 0$ in a Taylor series:

$$\tilde{\mathbf{f}}(\delta_{(i)}) = \mathbf{f}(\mathbf{x}^*) + \delta_{(i)} \cdot \left(\begin{array}{c} \frac{d\tilde{\mathbf{f}}_1}{d\delta_{(i)}} \\ \vdots \\ \frac{d\tilde{\mathbf{f}}_k}{d\delta_{(i)}} \end{array} \right) \bigg|_{\delta_{(i)}=0} + \mathbf{o}(\delta_{(i)}) \quad (6.9)$$

$$\frac{d\tilde{\mathbf{f}}_j}{d\delta_{(i)}} \bigg|_{\delta_{(i)}=0} = \nabla f_j(\mathbf{x}^*)^T \frac{d\mathbf{x}}{d\delta_{(i)}} \bigg|_{\delta_{(i)}=0}, \quad j = 1, \dots, k \quad (6.10)$$

$$\frac{d\mathbf{x}}{d\delta_{(i)}} \bigg|_{\delta_{(i)}=0} = \bar{\mathbf{Q}} \left(\frac{d(\delta_{(i)} \cdot \mathbf{e}_i)}{d\delta_{(i)}} \right) \bigg|_{\delta_{(i)}=0} = \bar{\mathbf{Q}} \left(\begin{array}{c} \mathbf{e}_i \\ 0 \end{array} \right) = \bar{\mathbf{q}}_i, \quad (6.11)$$

where $\bar{\mathbf{Q}}$ is the submatrix of $\mathbf{Q}$ formed by the first n rows, $\bar{\mathbf{q}}_i \in \mathbb{R}^n$ is the vector formed by the first n elements of the basis vector $\mathbf{q}_i$ (see Paragraph 5.3), and $\mathbf{o}(\delta_{(i)})$ denotes a mapping $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^k$ with $\mathbf{g}(0) = \mathbf{0}$ and $\lim_{\delta_{(i)} \rightarrow 0, \delta_{(i)} \neq 0} \frac{g_j(\delta_{(i)})}{\delta_{(i)}} = 0 \quad \forall j = 1, \dots, k$. It should be emphasized that the second identity in (6.11) is a consequence of $\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\xi}}(\mathbf{0}) = \mathbf{0}$.

Inserting (6.11) into (6.10) and the resulting equation into (6.9) gives

$$\tilde{\mathbf{f}}(\delta_{(i)}) = \mathbf{f}(\mathbf{x}^*) + \mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i \cdot \delta_{(i)} + \mathbf{o}(\delta_{(i)}), \quad (6.12)$$

where $\mathbf{f}'$ denotes the Jacobian matrix of $\mathbf{f}$. By means of the argumentation given in the footnote² we obtain

$$\rho(\delta_{(i)}) = \|\mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i\| \cdot |\delta_{(i)}| + \mathbf{o}(\delta_{(i)}). \quad (6.17)$$

² In order to obtain equation (6.17) we first state that

$$\|\mathbf{o}(\delta_{(i)})\| = \mathbf{o}(\delta_{(i)}). \quad (6.13)$$

This is true since (6.13) is another way of writing $\lim_{\delta_{(i)} \rightarrow 0, \delta_{(i)} \neq 0} \frac{(\sum_{j=1}^k g_j(\delta_{(i)})^2)^{\frac{1}{2}}}{\delta_{(i)}} = 0$

and since $\lim_{\delta_{(i)} \rightarrow 0, \delta_{(i)} \neq 0} \frac{(\sum_{j=1}^k g_j(\delta_{(i)})^2)^{\frac{1}{2}}}{\delta_{(i)}} = \lim_{\delta_{(i)} \rightarrow 0, \delta_{(i)} \neq 0} \left(\sum_{j=1}^k \left(\frac{g_j(\delta_{(i)})}{\delta_{(i)}} \right)^2 \right)^{\frac{1}{2}}$ and

$$\lim_{\delta_{(i)} \rightarrow 0, \delta_{(i)} \neq 0} \frac{g_j(\delta_{(i)})}{\delta_{(i)}} = 0 \quad \forall j = 1, \dots, k.$$

Now we are prepared to put our intention to produce a uniform spread of Pareto points in concrete terms. Assume, again, that a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ is given and that the homotopy algorithm is to compute further points $(\mathbf{x}_{(i)}, \boldsymbol{\lambda}_{(i)}, \boldsymbol{\alpha}_{(i)}) \in M$ in the neighborhood of $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$. In order to obtain a uniform spread in the objective space, the user of the algorithm should be able to predetermine the Euclidean distance $c \in \mathbb{R}^+$ between $\mathbf{f}(\mathbf{x}_{(i)})$ and $\mathbf{f}(\mathbf{x}^*)$, i.e. $\|\mathbf{f}(\mathbf{x}_{(i)}) - \mathbf{f}(\mathbf{x}^*)\| = c$. In the framework of a linear approximation, which is close to reality for small stepsizes $|\delta_{(i)}| \ll 1$, this requirement can be fulfilled due to (6.17) by choosing the chart parameter vectors $\boldsymbol{\xi}_{(i)}$ as $\boldsymbol{\xi}_{(i)} = \delta_{(i)} \cdot \mathbf{e}_i$, $i = 1, \dots, k-1$ with

$$\delta_{(i)} = \frac{c}{\|\mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i\|} . \quad (6.18)$$

Thus, the discretization of the Pareto surface in objective space can be well controlled by an appropriate **rescaling** of the coordinate axes in the space of the $\boldsymbol{\xi}$ -parameters.

Let us emphasize once again that the special property (5.28) $[\frac{\partial \eta}{\partial \boldsymbol{\xi}}(\mathbf{0}) = \mathbf{0}]$ of the constructed chart $\boldsymbol{\varphi}$ is the decisive reason for this simple controllability of the discretization density [see last identity in Equation (6.11)].

6.1.4 Numerical Algorithm

Now we shall put the method outlined above in the form of an algorithm describing the numerical computation of a set of candidates for Pareto optimal solutions. Each homotopy step comprises the following ten partial steps.

- (1) The **starting point** for a homotopy step is a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$.

When starting the method, i.e. when no homotopy step has been carried out yet, one obtains $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ by starting with a weight vector $\boldsymbol{\alpha}^* \in \mathbb{R}_+^k$ [with $\sum_{i=1}^k \alpha_i^* = 1$] and by solving the scalar-valued optimization problem ‘Minimize $g_{\boldsymbol{\alpha}^*}(\mathbf{x}) \equiv \boldsymbol{\alpha}^{*T} \mathbf{f}(\mathbf{x})$ under the constraint $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ ’ with a common optimization method. To this aim, one has at one’s disposal e.g. the method of the ‘Sequential Quadratic Programming’ (see e.g. [LUENBERGER, 1984],

Now we insert (6.12) into (6.8) and get

$$\rho(\delta_{(i)}) = \|\mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i \cdot \delta_{(i)} + \mathbf{o}(\delta_{(i)})\| . \quad (6.14)$$

The triangle axiom allows the conclusion

$$-\|\mathbf{o}(\delta_{(i)})\| \leq \|\mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i \cdot \delta_{(i)} + \mathbf{o}(\delta_{(i)})\| - \|\mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i \cdot \delta_{(i)}\| \leq \|\mathbf{o}(\delta_{(i)})\| \quad (6.15)$$

$$\Leftrightarrow \left| \|\mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i \cdot \delta_{(i)} + \mathbf{o}(\delta_{(i)})\| - \|\mathbf{f}'(\mathbf{x}^*) \bar{\mathbf{q}}_i \cdot \delta_{(i)}\| \right| \leq \|\mathbf{o}(\delta_{(i)})\| , \quad (6.16)$$

from which (6.17) follows immediately with the help of (6.13).

[FLETCHER, 1987], [GROSSMANN & TERNO, 1993]) or the (Best/Braeuninger/Ritter/Robinson)-method (see [BEST ET AL., 1981]). Once the homotopy method has been started, we can choose arbitrary³ points out of M , generated by homotopy, as new starting points $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$.

- (2) Calculate the **Jacobian matrix** $\mathbf{F}'$ of $\mathbf{F}$ in the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$.

$$\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) = \left(\begin{array}{c|cc} \nabla^2(\boldsymbol{\alpha}^{*T}\mathbf{f}(\mathbf{x}^*) + \boldsymbol{\lambda}^{*T}\mathbf{h}(\mathbf{x}^*)) & \nabla h_1 & \dots & \nabla h_m & \nabla f_1 & \dots & \nabla f_k \\ \hline \nabla h_1(\mathbf{x}^*)^T & & & & & & \\ \vdots & & & & & & \\ \nabla h_m(\mathbf{x}^*)^T & & & & & & \\ \hline \mathbf{0} & \mathbf{0} & & & 1 & \dots & 1 \end{array} \right) \quad (6.19)$$

The information of first and of second order (i.e. ∇ and ∇^2) of the functions $\mathbf{f}$ and $\mathbf{h}$ in the point $\mathbf{x}^*$, which is required for the calculation of $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$, can be gained either by symbolic differentiation — which yields an exact result, but is not always practicable —, by automatic differentiation (see [FISCHER, 1988] and [FISCHER, 1996]) or by numerical differentiation (i.e. by means of approximation of partial derivatives by difference quotients).

- (3) Generate a **QR-factorization** of the matrix $(\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*))^T$ by Householder reflections (see e.g. [WERNER, 1992]).

From this factorization, which does not make demands on the rank of $(\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*))^T$, an orthogonal matrix $\tilde{\mathbf{Q}} \in \mathbb{R}^{(n+m+k) \times (n+m+k)}$ and a matrix $\mathbf{R} \equiv \begin{pmatrix} \mathbf{R}_I \\ \mathbf{0} \end{pmatrix} \in \mathbb{R}^{(n+m+k) \times (n+m+1)}$ result (see Figure 6.3), where $\mathbf{R}_I \in \mathbb{R}^{(n+m+1) \times (n+m+1)}$ is an upper triangular matrix such that:

$$(\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*))^T = \tilde{\mathbf{Q}} \mathbf{R} \quad (6.20)$$

- (4) The triangular matrix $\mathbf{R}_I$ contains the information about whether $\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ has the **full rank**.

To understand this, let us examine the j -th column of $(\mathbf{F}'(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*))^T$. Because of Equation (6.20) and of the triangular shape of $\mathbf{R}_I$, it is a linear combination of the first j columns of $\tilde{\mathbf{Q}}$, where the linear coefficients are in the j -th column of $\mathbf{R}_I$. If and only if $(\mathbf{R}_I)_{jj} = 0$ [or, from a numerical viewpoint, if $|(\mathbf{R}_I)_{jj}| < s$, with the numerical bound $s \in \mathbb{R}^+$], the j -th column

³ The test, whether the Rank Condition (5.6) in the point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ is fulfilled, is carried out only in step (4).

$$\begin{array}{c} n+m+1 \\ \hline (F')^T \end{array} = \begin{array}{c} n+m+k \\ \hline \tilde{Q} \end{array} \cdot \begin{array}{c} n+m+1 \\ \hline \begin{array}{c} R_1 \\ 0 \end{array} \end{array}$$

$(q_k | \dots | q_{n+m+k}) \quad (q_1 | \dots | q_{k-1})$

Figure 6.3 Structure of the matrices resulting from a QR -factorization of $(F'(\mathbf{x}^*, \lambda^*, \alpha^*))^T$.

of $(F'(\mathbf{x}^*, \lambda^*, \alpha^*))^T$ is situated in the span of the first $j-1$ columns of $\tilde{Q}$ and hence also in the span of the first $j-1$ columns of $(F'(\mathbf{x}^*, \lambda^*, \alpha^*))^T$.

$(F'(\mathbf{x}^*, \lambda^*, \alpha^*))^T$ has therefore full rank $n+m+1$, if and only if all diagonal elements of R_1 are unequal to zero. If this is not the case, the Rank Condition (5.6) in the point $(\mathbf{x}^*, \lambda^*, \alpha^*)$ is not fulfilled. Consequently, the point cannot be a starting point of a homotopy step and we have to go back to step (1).

- (5) By reordering of the columns of the matrix $\tilde{Q}$ one can get a matrix Q , the columns of which form the **orthonormal basis** of the $\mathbb{R}^{n+m+k}$ which is required for the local chart φ (see Equation (5.30)).

According to the aforementioned the span of the columns of $(F'(\mathbf{x}^*, \lambda^*, \alpha^*))^T$ is identical with the span of the first $n+m+1$ columns of $\tilde{Q}$ [the fulfillment of the Rank Condition (5.6) has been checked in the last step and from now on shall be taken for granted.]. Since we have $\text{span}\{\text{columns of } (F'(\mathbf{x}^*, \lambda^*, \alpha^*))^T\} = (T_{(\mathbf{x}^*, \lambda^*, \alpha^*)} M)^\perp$, and since $\tilde{Q}$ is orthogonal, we can conclude immediately: The columns of $\tilde{Q}$ are an orthonormal basis of the $\mathbb{R}^{n+m+k}$, such that $\text{span}\{\tilde{q}_1, \dots, \tilde{q}_{n+m+1}\} = (T_{(\mathbf{x}^*, \lambda^*, \alpha^*)} M)^\perp$ and $\text{span}\{\tilde{q}_{n+m+2}, \dots, \tilde{q}_{n+m+k}\} = T_{(\mathbf{x}^*, \lambda^*, \alpha^*)} M$, where the j -th column vector of $\tilde{Q}$ is denoted by $\tilde{q}_j$.

Let us write $\tilde{Q}$ as $\tilde{Q} = (\tilde{Q}_1 | \tilde{Q}_2)$, where $\tilde{Q}_1 \in \mathbb{R}^{(n+m+k) \times (n+m+1)}$ and $\tilde{Q}_2 \in \mathbb{R}^{(n+m+k) \times (k-1)}$. The orthogonal matrix Q required for our chart φ is simply obtained (cf. Figure 6.3) by exchanging the order of the two sub-

matrices $\tilde{Q}_1$ and $\tilde{Q}_2$, i.e.

$$Q = \left(\tilde{Q}_2 | \tilde{Q}_1 \right) . \quad (6.21)$$

- (6) Generate a **set of chart parameter vectors** $\{\xi_{(i)}\}$, labeled by an index set I .

Let us assume that each point $\mathbf{x}_{(i)}$ computed by the current step of the homotopy algorithm has a given distance $c \in \mathbb{R}^+$ in the objective space from the starting point $\mathbf{x}^*$, i.e. $\|\mathbf{f}(\mathbf{x}_{(i)}) - \mathbf{f}(\mathbf{x}^*)\| \approx c$. This requirement can be met by the following rule:

Choose the index set $I = \{1, \dots, k-1\}$ and the i -th chart parameter as

$$\xi_{(i)} = \frac{c}{\|\mathbf{f}'(\mathbf{x}^*)\bar{\mathbf{q}}_i\|} \mathbf{e}_i , \quad (6.22)$$

where $\mathbf{e}_i$ is the i -th unit vector in the $\mathbb{R}^{k-1}$ and $\bar{\mathbf{q}}_i \in \mathbb{R}^n$ is the vector constructed out of the first n elements of the basis vectors $\mathbf{q}_i$.

- (7) Carry out **steps (8) to (10) for all indices** $i \in I$.

- (8) **Predictor step:**

$$\varphi^P(\xi_{(i)}) = (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + (\mathbf{q}_1 \dots \mathbf{q}_{k-1}) \xi_{(i)} \quad (6.23)$$

- (9) **Corrector step by the Newton method:**

Generate a sequence of points $\boldsymbol{\eta}_{(i)}^{[l]} \in \mathbb{R}^{n+m+1}$, $l = 0, 1, \dots$, with the starting point $\boldsymbol{\eta}_{(i)}^{[0]} = \mathbf{0}$ and the iteration rule $\boldsymbol{\eta}_{(i)}^{[l+1]} = \mathbf{z} + \boldsymbol{\eta}_{(i)}^{[l]}$, where $\mathbf{z}$ is the solution of the linear system of equations

$$\mathbf{F}'(\varphi^P(\xi_{(i)})) \cdot (\mathbf{q}_k \dots \mathbf{q}_{n+m+k}) \mathbf{z} = -\mathbf{F}\left((\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + Q \begin{pmatrix} \xi_{(i)} \\ \boldsymbol{\eta}_{(i)}^{[l]} \end{pmatrix}\right). \quad (6.24)$$

This is a modification of the Newton method, the so-called simplified Newton method (see e.g. [WERNER, 1992]). Since only the right-hand side of (6.24) changes during the Newton iterations, this modification has the advantage that the Jacobian matrix $\mathbf{F}'$, which contains second derivatives of $\mathbf{f}$ and $\mathbf{h}$ and is therefore costly to be calculated, has to be determined only once every corrector step. For the solution of the linear system of Equations (6.24) one calculates once, at the beginning of the Newton iterations, a LR -decomposition of the matrix $[\mathbf{F}'(\varphi^P(\xi_{(i)})) \cdot (\mathbf{q}_k \dots \mathbf{q}_{n+m+k})] \in \mathbb{R}^{(n+m+1) \times (n+m+1)}$ by means of the Gaussian elimination with columns pivot search, and then obtains in every iteration step the solution of (6.24) by descending and ascending substitution.

If after a given maximum number of iterations l_{max} the norm of the residual vector $\mathbf{F} \left((\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) + \mathbf{Q} \begin{pmatrix} \boldsymbol{\xi}_{(i)} \\ \boldsymbol{\eta}_{(i)}^{[l_{max}]} \end{pmatrix} \right)$ is not yet smaller than a given error bound, this is regarded as an indication for the Newton method not to converge. The failure to converge again suggests that the distance of the predictor point $\boldsymbol{\varphi}^P(\boldsymbol{\xi}_{(i)})$ to the manifold M is too large, i.e. the predictor step was too large in view of the local curvature of the manifold M .

A remedy for this is reducing the chart parameter step $\boldsymbol{\xi}_{(i)}$ to half and repeating the predictor step (8) with the chart parameter value $\frac{1}{2}\boldsymbol{\xi}_{(i)}$. The behavior of convergence of the Newton method is therefore a sensor for the steplength control of the single homotopy steps.

- (10) The last step is the **check**, whether the generated point $\boldsymbol{\varphi}(\boldsymbol{\xi}_{(i)})$ is really in the manifold M .

For this purpose one still has to check, whether $P_{\boldsymbol{\alpha}} \boldsymbol{\varphi}(\boldsymbol{\xi}_{(i)}) \in \overset{\circ}{\mathbb{R}}_+^k$ is true, where $P_{\boldsymbol{\alpha}}$ is the projection onto the $\boldsymbol{\alpha}$ -space [i.e. $P_{\boldsymbol{\alpha}}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = \boldsymbol{\alpha}$]. If this is not the case, one has to go back again to the predictor step (8), using the halved chart parameter value $\frac{1}{2}\boldsymbol{\xi}_{(i)}$.

6.2 Method II: Purposeful Change of the Weights

6.2.1 Significance of the Weight Vector for the User

The application scenario II (see introduction of this chapter) assumes that the decision-maker wants to vary the weight vector $\boldsymbol{\alpha}$ in a purposeful way. To underline the relevance of this scenario we will discuss briefly, what significance the knowledge of $\boldsymbol{\alpha}$ has for the user.

Let a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$ be given. What information does the weight vector $\boldsymbol{\alpha}^*$ have for the decision-maker?

In case $\mathbf{x}^*$ is a *minimum* of the scalar-valued function $g_{\boldsymbol{\alpha}^*}$, the information is clear. The components of $\boldsymbol{\alpha}^*$ indicate the relative weights of the individual objectives within the total objective function $g_{\boldsymbol{\alpha}^*}$, which is a linear combination of the k individual objectives and is minimized by $\mathbf{x}^*$.

If $\mathbf{x}^*$ is a *saddle point* of $g_{\boldsymbol{\alpha}^*}$, the interpretation of $\boldsymbol{\alpha}^*$ cannot be gathered so clearly. In this case reverting to the geometric significance of the weight vector, as demonstrated in Chapter 4, is helpful: $\boldsymbol{\alpha}$ is orthogonal to the tangent plane to the ‘hypersurface of efficient points’ in the objective space. If one portrays an efficient point $\mathbf{y}$ which is a neighboring point to $\mathbf{f}(\mathbf{x}^*)$ by the difference vector $\delta\mathbf{y}$, i.e. $\mathbf{y} = \mathbf{f}(\mathbf{x}^*) + \delta\mathbf{y}$, and ignores the component of $\delta\mathbf{y}$ which projects out of the tangent plane one has accordingly

$$\boldsymbol{\alpha}^{*T} \cdot \delta\mathbf{y} = 0 . \quad (6.25)$$

One can gather from this that α^* indicates the **relative weight of the individual objectives in the point $\mathbf{x}^*$** also in such cases, when $\mathbf{x}^*$ is a saddle point of g_{α^*} . For example, we will pick out the first two objectives, i.e. we are only interested in such efficient neighboring points of $\mathbf{f}(\mathbf{x}^*)$ which differ from $\mathbf{f}(\mathbf{x}^*)$ in the first two objectives. For the difference vectors $\delta \mathbf{y} = (\delta y_1, \delta y_2, 0, \dots, 0)$ associated with these neighboring points we can conclude from Equation (6.25)

$$\delta y_2 = -\frac{\alpha_1^*}{\alpha_2^*} \delta y_1 . \quad (6.26)$$

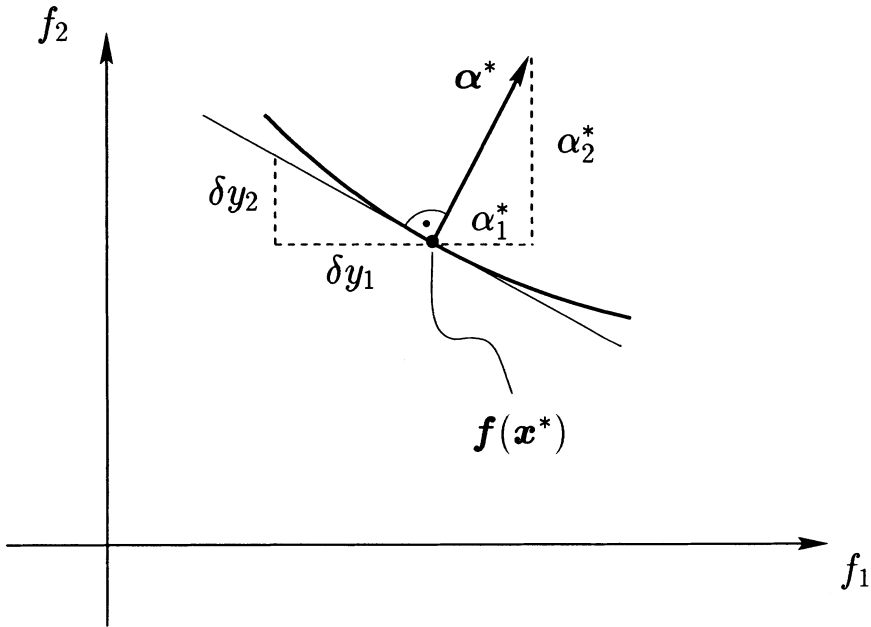


Figure 6.4 The significance of the weight vector α^* as the 'exchange rate' [valid in the point $\mathbf{f}(\mathbf{x}^*)$] between the individual objectives.

If one intends, starting from $\mathbf{f}(\mathbf{x}^*)$, to improve the first objective by a certain amount, one has to pay with a deterioration of the second objective by the $\frac{\alpha_1^*}{\alpha_2^*}$ -fold of this amount (see Figure 6.4). One can compare the situation to the foreign exchange market: In order to get something in the 'currency' of the objective 1 one has to pay the corresponding amount in the currency of the objective 2. In this metaphor the components of α^* determine the 'exchange factors' or 'exchange rates', which are valid in the point $\mathbf{x}^*$ [or $\mathbf{f}(\mathbf{x}^*)$] between the different objectives. The user can interpret α^* in any case directly as the relative weight of the individual objectives in the point $\mathbf{x}^*$.

6.2.2 Principle of the Procedure

In the scenario of application II the user has a point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*) \in M$, in which the Rank Condition (5.6) is fulfilled, and would like to know which efficient solutions he obtains when shifting the weight vector from the present value $\boldsymbol{\alpha}^*$ iteratively by a small vector $\delta\boldsymbol{\alpha}$. In the following we will develop a variant of the homotopy method described in the last section, which calculates the corresponding wandering of the efficient points.

The strategy of this method variant II is to insert in the method version I the requirement that every **predictor step** observes the given change⁴ $\delta\bar{\boldsymbol{\alpha}}$ of the weight vector in the best possible way. If we introduce the projection operator $P_{\boldsymbol{\alpha}}$ onto the $\boldsymbol{\alpha}$ -space by means of the definition

$$P_{\boldsymbol{\alpha}} : \begin{cases} \mathbb{R}^{n+m+k} & \rightarrow \mathbb{R}^k \\ (\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) & \mapsto \boldsymbol{\alpha} \end{cases}, \quad (6.28)$$

this requirement reads as follows:

$$P_{\boldsymbol{\alpha}} [\boldsymbol{\varphi}^P(\boldsymbol{\xi}) - (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)] \equiv (\bar{\mathbf{b}}_1 \dots \bar{\mathbf{b}}_{k-1}) \boldsymbol{\xi} = \delta\bar{\boldsymbol{\alpha}}, \quad (6.29)$$

where the vectors $\bar{\mathbf{b}}_i$ denote the partial vectors constructed out of the last k elements of the vectors $\mathbf{q}_i$ (of the tangent plane basis). This requirement at first looks like an overdetermined system of equations, since it has k equations for the $k-1$ components of $\boldsymbol{\xi}$. The last equation in (6.29) is $(\bar{\mathbf{b}}_1)_k, \dots, (\bar{\mathbf{b}}_{k-1})_k \cdot \boldsymbol{\xi} = (\delta\bar{\boldsymbol{\alpha}})_k$. Now every vector $\mathbf{q}_i, i = 1, \dots, k-1$, as a vector of the tangent plane, is orthogonal to the row vector $\frac{\partial F_{n+m+1}}{\partial(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})} \equiv (0, \dots, 0, 0, \dots, 0, 1, \dots, 1)$ of the Jacobian matrix $\mathbf{F}'$ [see equation (6.19)]. It follows that $\sum_{l=1}^k (\bar{\mathbf{b}}_i)_l = 0$ or $(\bar{\mathbf{b}}_i)_k = -\sum_{l=1}^{k-1} (\bar{\mathbf{b}}_i)_l$ for all $i = 1, \dots, k-1$. As in the form of Equation (6.27) the analogue is equally valid for the right-hand side of Equation (6.29), the k -th equation in (6.29) is the negative sum of the first $k-1$ equations and therefore redundant. With the denotations $\mathbf{b}_i := ((\bar{\mathbf{b}}_i)_1, \dots, (\bar{\mathbf{b}}_i)_{k-1})^T$, $\delta\boldsymbol{\alpha} := ((\delta\bar{\boldsymbol{\alpha}})_1, \dots, (\delta\bar{\boldsymbol{\alpha}})_{k-1})^T$ and $\mathbf{B} := (\mathbf{b}_1 \dots \mathbf{b}_{k-1}) \in \mathbb{R}^{(k-1) \times (k-1)}$ Equation (6.29) is therefore equivalent to

$$(\mathbf{b}_1 \dots \mathbf{b}_{k-1}) \boldsymbol{\xi} \equiv \mathbf{B} \boldsymbol{\xi} = \delta\boldsymbol{\alpha}. \quad (6.30)$$

The solubility of this linear system of equations is determined by the rank of the matrix $\mathbf{B}$. Two cases can be distinguished and shall be discussed in the following.

⁴ Since $\boldsymbol{\alpha}^* + \delta\bar{\boldsymbol{\alpha}}$ must also be a valid weight vector, and therefore $\sum_{l=1}^k (\boldsymbol{\alpha}^* + \delta\bar{\boldsymbol{\alpha}})_l \equiv \sum_{l=1}^k \boldsymbol{\alpha}_l^* + \sum_{l=1}^k (\delta\bar{\boldsymbol{\alpha}})_l = 1$ must be true, we assume that $\delta\bar{\boldsymbol{\alpha}}$ satisfies the equation

$$(\delta\bar{\boldsymbol{\alpha}})_k = -\sum_{l=1}^{k-1} (\delta\bar{\boldsymbol{\alpha}})_l. \quad (6.27)$$

- (A) $\text{rank } \mathbf{B} = k - 1 \iff \mathbf{B}$ is regular.

In this case Equation (6.30) has always exactly one solution $\xi_{\delta\alpha}$, and this is true for arbitrarily chosen $\delta\alpha$. This freedom in choosing the weight shift $\delta\alpha$ shows that the manifold M could be locally parametrized directly by the weight shift $\delta\alpha$ (see Paragraph 5.2.3). Equation (6.30) is in this case the conversion formula of a parametrization by $\delta\alpha$ (see footnote⁵) into a parametrization by means of the coordinates ξ with regard to an orthonormal basis of the tangent plane.

- (B) $\text{rank } \mathbf{B} =: l < k - 1 \iff \mathbf{B}$ is not regular.

In this case Equation (6.30) is not soluble for arbitrary $\delta\alpha$, and therefore the manifold M cannot be parametrized in its full dimensionality by (arbitrary) $\delta\alpha$. Depending on the relation between the matrix $\mathbf{B}$ and the value of $\delta\alpha$, which is given by the user, two cases can again be discriminated:

- (a) $\delta\alpha \in \text{image } \mathbf{B} [= \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}]$

Under this assumption Equation (6.30) is soluble, but the solution is not determined in a unique way. Since no further requirements are made for the next point on M to be computed, it is sufficient to calculate **one** solution of (6.30). A procedure to find such a solution is the following:

Take l linearly independent vectors out of the set $\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}$. Since these vectors are a basis of the subspace $\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\} \subset \mathbb{R}^{k-1}$, $\delta\alpha$ can be represented in a unique way as a linear combination of these basis vectors. The following vector $\xi_{\delta\alpha}$ now solves Equation (6.30): For all indices $i \in \{1, \dots, k-1\}$, for which $\mathbf{b}_i$ is part of the chosen basis, ξ_i is equal to the coefficient of $\delta\alpha$ with regard to $\mathbf{b}_i$ in the mentioned linear combination. For all indices $j \in \{1, \dots, k-1\}$, for which $\mathbf{b}_j$ is not part of the basis, we set $\xi_j = 0$.

- (b) $\delta\alpha \notin \text{image } \mathbf{B}$

In this case Equation (6.30) has no solution, i.e. the variation of the weight vector wanted by the user cannot be fully carried out. However, continuing the iterations of the homotopy method can make sense, in order to find eventually a situation in later homotopy steps, in which the wanted variation $\delta\alpha$ is feasible.

To define a predictor step which complies with the $\delta\alpha$ desired by the user as well as possible, we make Equation (6.30) less rigorous by requiring that ξ should be chosen such that the α -component of the

⁵ Let us note here that a strict parametrization by $\delta\alpha$ requires – unlike our method – that not only the predictor step, but also the corrector step should observe exactly the given $\delta\alpha$. This requirement leads back to a one-dimensional homotopy method, in which instead of the full manifold M only a curve on M is parametrized by $t \cdot \delta\alpha$, $t \in \mathbb{R}$, $\delta\alpha$ fixed. One-dimensional homotopy methods are described, e.g., in [SCHWETLICK, 1979] and [SCHWETLICK & KRETZSCHMAR, 1991].

predictor **minimizes** the **distance to** $\delta\alpha$. This is equivalent to the equation

$$(\bar{b}_1 \dots \bar{b}_{k-1}) \xi = P_{\text{span}\{\bar{b}_1, \dots, \bar{b}_{k-1}\}} \delta\alpha, \quad (6.31)$$

where $P_{\text{span}\{\bar{b}_1, \dots, \bar{b}_{k-1}\}}$ denotes the projection of the $\mathbb{R}^{k-1}$ onto the span of the vectors $\{\bar{b}_1, \dots, \bar{b}_{k-1}\}$. If one interprets $P_{\text{span}\{\bar{b}_1, \dots, \bar{b}_{k-1}\}} \delta\alpha$ as a ‘new $\delta\alpha$ ’, one finds the initial situation of case (a) and can borrow the solution method⁶ given there.

To complete the picture let us add that in both cases (A) and (B)(a) we have $P_{\text{span}\{\bar{b}_1, \dots, \bar{b}_{k-1}\}} \delta\alpha = \delta\alpha$. That is why in these cases the modified Equation (6.31) is equivalent to Equation (6.30). Equation (6.30) can therefore be substituted formally in all cases by Equation (6.31).

Also in method variant II we will stick to the principle of determining points of M by numerical evaluations of the local chart φ , because this principle has the mentioned advantages of taking into consideration the full dimensionality of M in a natural way and of adapting the calculation procedure to the local geometry of M . The principle implies, however, that the given $\delta\bar{\alpha}$ can only be taken into consideration in the predictor step. The corrector step is always carried out orthogonally to the tangent plane $T_{(x^*, \lambda^*, \alpha^*)}M$ and is completely determined by Equation (6.2). Therefore one can ask immediately to what extent the **corrector step is compatible with the given $\delta\bar{\alpha}$** .

In order to settle this question, the α -component of the point $\varphi(\xi) \in M$, which has been constructed by method variant II, has to be compared to the wanted new weight vector $(\alpha^* + \delta\bar{\alpha})$ or, equivalently hereto, $P_\alpha [\varphi(\xi) - (x^*, \lambda^*, \alpha^*)]$ has to be compared to $\delta\bar{\alpha}$. To this end, we examine the Taylor expansion of $P_\alpha [\varphi(\xi) - (x^*, \lambda^*, \alpha^*)]$, interpreted as a function of ξ , around the point $\xi_0 = 0$. Let the partial vectors constructed out of the last k elements of the vectors q_i (of the tangent plane basis) be denoted again by $\bar{b}_i$.

$$P_\alpha [\varphi(\xi) - (x^*, \lambda^*, \alpha^*)] = \left[(\bar{b}_1 \dots \bar{b}_{n+m+k}) \frac{\partial \begin{pmatrix} \xi \\ \eta(\xi) \end{pmatrix}}{\partial \xi}(0) \right] \cdot \xi + o(\|\xi\|) \quad (6.32)$$

$$= \left[(\bar{b}_1 \dots \bar{b}_{k-1}) \frac{\partial \xi}{\partial \xi}(0) + (\bar{b}_k \dots \bar{b}_{n+m+k}) \frac{\partial \eta(\xi)}{\partial \xi}(0) \right] \cdot \xi + o(\|\xi\|) \quad (6.33)$$

$$= (\bar{b}_1 \dots \bar{b}_{k-1}) \xi + o(\|\xi\|) \quad (6.34)$$

The last equality sign is a consequence of the property $\frac{\partial \eta(\xi)}{\partial \xi}(0) = 0$, which is the main construction feature of our local chart φ [see Equation (5.28)]. Because

⁶ For a detailed discussion of the solution method see Paragraph 6.2.3.

of this property the corrector step, i.e. the determination of $\boldsymbol{\eta}(\boldsymbol{\xi})$, does not have an influence in linear order on the $\boldsymbol{\alpha}$ -component of the newly determined point. On the contrary, $P_{\boldsymbol{\alpha}} [\boldsymbol{\varphi}(\boldsymbol{\xi}) - (\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)]$ is, in linear order, only given by the predictor step $(\bar{\boldsymbol{b}}_1 \dots \bar{\boldsymbol{b}}_{k-1}) \boldsymbol{\xi}$ which, according to requirement (6.31), takes into account the given $\delta\bar{\boldsymbol{\alpha}}$ in the best possible way.

Summarizing one can state: Deviations from the given $\delta\bar{\boldsymbol{\alpha}}$ which originate from the corrector step do not appear in linear, but only in quadratic order with regard to the steplength $\|\boldsymbol{\xi}\|$ of the homotopy step. This is a consequence of the special construction of the chart $\boldsymbol{\varphi}$, which parametrizes the manifold M locally.

6.2.3 Numerical Algorithm

A homotopy step of the method variant II consists of nine partial steps. The partial steps (1) to (5) are identical with the corresponding steps of the method variant I. Instead of the generation of a *set* of chart parameter vectors [partial step (6) of variant I] in variant II only *one* chart parameter vector $\boldsymbol{\xi}$ is now determined according to the given $\delta\bar{\boldsymbol{\alpha}}$ [partial step (6)]. The subsequent partial steps (7) to (9) are identical with the steps (8) to (10) of variant I.

After each homotopy step the calculated point $\boldsymbol{\varphi}(\boldsymbol{\xi})$ is taken as the starting point $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ of a new homotopy step.

The only partial step which is changed compared to variant I is step (6). The aim of this step is finding a solution $\boldsymbol{\xi}$ of Equation (6.31). The numerical method leading to this aim can be formulated in a way which is valid for all three cases [(A), (B)(a) and (B)(b)] discussed earlier in Paragraph 6.2.2. The partial steps to be executed are listed below.

- (6.a) Determine the matrix $\boldsymbol{B} \in \mathbb{R}^{(k-1) \times (k-1)}$ out of the matrix $\boldsymbol{Q} \in \mathbb{R}^{(n+m+k) \times (n+m+k)}$, calculated in step (5), by eliminating the first $(n+m)$ rows and the $(n+m+k)$ -th row of the submatrix which is formed by the first $k-1$ columns of $\boldsymbol{Q}$.
- (6.b) Initialize a matrix $\tilde{\boldsymbol{B}}$ to $\tilde{\boldsymbol{B}} = \boldsymbol{B}$. In the following $\tilde{\boldsymbol{B}}$, by eliminating linearly dependent columns, will be reduced to a matrix, the columns of which form a basis of the subspace $\text{span}\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_{k-1}\} \subseteq \mathbb{R}^{k-1}$.
- (6.c) Initialize a binary vector $\boldsymbol{a} \in \{0, 1\}^{k-1}$ with the value $\boldsymbol{a} = (1, \dots, 1)$. The i -th element (= the i -th bit) of this vector is planned to contain the information, whether the i -th column of $\boldsymbol{B}$ is part of the chosen basis of the subspace $\text{span}\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_{k-1}\} \subseteq \mathbb{R}^{k-1}$ [then: $a_i = 1$] or not [then: $a_i = 0$]. Initialize a pointer such that it points to the first element of $\boldsymbol{a}$.
- (6.d) Generate a $\boldsymbol{QR}$ -factorization of the matrix $\tilde{\boldsymbol{B}}$ by (at most) $k-1$ Householder reflections according to the following rule:

Rename the matrix $\tilde{\mathbf{B}}$, transformed by $s-1$ Householder reflections, $\tilde{\mathbf{B}}^{(s-1)}$. Check **before** the s -th Householder reflection, whether the vector $\mathbf{c}_s := (\tilde{\mathbf{B}}_{ss}^{(s-1)}, \tilde{\mathbf{B}}_{(s+1)s}^{(s-1)}, \dots, \tilde{\mathbf{B}}_{(k-1)s}^{(s-1)})^T$ (which is transformed next) fulfills the condition

$$\|\mathbf{c}_s\| > \epsilon, \quad (6.35)$$

where $\epsilon \ll 1$ is a numerical bound set in advance.

If it does, execute the s -th Householder reflection as described in the textbooks (see e.g. [WERNER, 1992]) and advance the pointer by one position within the binary vector $\mathbf{a}$.

If not, replace the s -th Householder reflection with the following two actions without increasing the index s by 1:

- Eliminate the column vector $\tilde{\mathbf{b}}_s$ of the matrix $\tilde{\mathbf{B}}$, since it is a linear combination of the vectors $\{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_{s-1}\}$. Eliminate the s -th column in $\tilde{\mathbf{B}}^{(s-1)}$ correspondingly.
- In the vector $\mathbf{a}$ (indicating the choice of the basis vectors) the value, at which the pointer currently points, is changed to 0. Advance the pointer subsequently by one position (within $\mathbf{a}$).

At the end of step (6.d) one has the following results:

- A matrix $\tilde{\mathbf{B}}$, the columns of which span the subspace $\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\} \subseteq \mathbb{R}^{k-1}$ and represent a selection of the vectors $\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}$.
- The number l of columns of $\tilde{\mathbf{B}}$. We know that $l = \text{rank } \mathbf{B}$.
- The ‘basis-choice-vector’ $\mathbf{a}$, by means of which the original position of the columns $\tilde{\mathbf{b}}_i$ of $\tilde{\mathbf{B}}$ (i.e. the chosen basis vectors) in the initial matrix $\mathbf{B}$ can be reconstructed.
- A $\mathbf{QR}$ -decomposition of $\tilde{\mathbf{B}}$,

$$\tilde{\mathbf{B}} = \hat{\mathbf{Q}}\hat{\mathbf{R}}, \quad (6.36)$$

where $\hat{\mathbf{Q}} \in \mathbb{R}^{(k-1) \times (k-1)}$ is an orthogonal matrix, the first l columns $\{\hat{\mathbf{q}}_1, \dots, \hat{\mathbf{q}}_l\}$ of which form an orthonormal basis of the subspace $\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}$. The first l rows of $\hat{\mathbf{R}} \in \mathbb{R}^{(k-1) \times l}$ constitute an $l \times l$ -upper triangular matrix $\hat{\hat{\mathbf{R}}}$, the last $((k-1) - l)$ rows contain only zeros.

(6.e) Now is the time to make a distinction of cases.

(case 1) $l > 0$.

First calculate an auxiliary variable $\tilde{\boldsymbol{\xi}} \in \mathbb{R}^l$ which solves the equation

$$\tilde{\mathbf{B}}\tilde{\boldsymbol{\xi}} = P_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}} \delta \boldsymbol{\alpha}. \quad (6.37)$$

This equation can be transformed in such a way that the solution can be found easily:

Since from the property of $\{\hat{\mathbf{q}}_1, \dots, \hat{\mathbf{q}}_l\}$ of being an orthonormal basis of $\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}$ follows the equation

$$P_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}} = \begin{pmatrix} \hat{\mathbf{q}}_1 & \dots & \hat{\mathbf{q}}_l \end{pmatrix} \cdot \begin{pmatrix} \hat{\mathbf{q}}_1^T \\ \vdots \\ \hat{\mathbf{q}}_l^T \end{pmatrix}, \quad \text{and because of}$$

$\hat{\mathbf{Q}}\hat{\mathbf{R}} = \begin{pmatrix} \hat{\mathbf{q}}_1 & \dots & \hat{\mathbf{q}}_l \end{pmatrix} \hat{\mathbf{R}}$, Equation (6.37) can also be written in the form

$$\begin{pmatrix} \hat{\mathbf{q}}_1 & \dots & \hat{\mathbf{q}}_l \end{pmatrix} \hat{\mathbf{R}} \tilde{\boldsymbol{\xi}} = \begin{pmatrix} \hat{\mathbf{q}}_1 & \dots & \hat{\mathbf{q}}_l \end{pmatrix} \begin{pmatrix} \hat{\mathbf{q}}_1^T \\ \vdots \\ \hat{\mathbf{q}}_l^T \end{pmatrix} \delta \boldsymbol{\alpha}. \quad (6.38)$$

Multiplying this equation from the left with $\begin{pmatrix} \hat{\mathbf{q}}_1^T \\ \vdots \\ \hat{\mathbf{q}}_l^T \end{pmatrix}$ and making use

of $\begin{pmatrix} \hat{\mathbf{q}}_1^T \\ \vdots \\ \hat{\mathbf{q}}_l^T \end{pmatrix} \begin{pmatrix} \hat{\mathbf{q}}_1 & \dots & \hat{\mathbf{q}}_l \end{pmatrix} = \text{Identity} \in \mathbb{R}^{l \times l}$ finally results in

$$\hat{\mathbf{R}} \tilde{\boldsymbol{\xi}} = \begin{pmatrix} \hat{\mathbf{q}}_1^T \\ \vdots \\ \hat{\mathbf{q}}_l^T \end{pmatrix} \delta \boldsymbol{\alpha}. \quad (6.39)$$

From this form of Equation (6.37) the solution $\tilde{\boldsymbol{\xi}}$ can be calculated directly by ascending substitution.

Now we obtain a solution $\boldsymbol{\xi} \in \mathbb{R}^{k-1}$ of Equation (6.31) from the auxiliary variable $\tilde{\boldsymbol{\xi}} \in \mathbb{R}^l$ by the following procedure:

Copy for $i = 1, \dots, l$ the elements $\tilde{\xi}_i$ of the vector $\tilde{\boldsymbol{\xi}}$ one by one to those positions in the vector $\boldsymbol{\xi}$, in which there is a 1 in the basis-choice-vector $\mathbf{a}$. Fill all other $((k-1)-l)$ positions in $\boldsymbol{\xi}$ with zeros.

(Case 2) $l = 0$.

In this case the subspace $\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_{k-1}\}$ has the dimension 0, and the Equation (6.31) we want to solve has the trivial form $\mathbf{0} = \mathbf{0}$. For the determination of $\boldsymbol{\xi}$ we therefore need a different criterion. A reasonable requirement for $\boldsymbol{\xi}$ is that the current homotopy step should not lead back to that position on the manifold M , from where one has just arrived. This is guaranteed by the following procedure:

- Determine $(\Delta \mathbf{x}, \Delta \boldsymbol{\lambda}, \Delta \boldsymbol{\alpha})$, i.e. the difference vector between the starting point $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)$ of the current homotopy step and the starting point of the last homotopy step.

- Calculate $(\Delta \mathbf{x}, \Delta \boldsymbol{\lambda}, \Delta \boldsymbol{\alpha})^T \cdot \mathbf{q}_1$ [remember: $\mathbf{q}_1$ is the first vector of an orthonormal basis of the tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*)}M$].
- If $(\Delta \mathbf{x}, \Delta \boldsymbol{\lambda}, \Delta \boldsymbol{\alpha})^T \cdot \mathbf{q}_1 \geq 0$ is true, the last homotopy step and the basis vector $\mathbf{q}_1$ include an acute angle. In this case set $\boldsymbol{\xi} = (1, 0, 0, \dots)$.
If $(\Delta \mathbf{x}, \Delta \boldsymbol{\lambda}, \Delta \boldsymbol{\alpha})^T \cdot \mathbf{q}_1$ is negative, $(\Delta \mathbf{x}, \Delta \boldsymbol{\lambda}, \Delta \boldsymbol{\alpha})^T \cdot (-\mathbf{q}_1) \geq 0$ follows. Therefore set $\boldsymbol{\xi} = (-1, 0, 0, \dots)$.

Chapter 7

Numerical Results

The aim of the present chapter is, on the one hand, to check by numerical tests the correctness of the method developed. For this purpose, in Section 7.1 an academic example of a vector optimization problem shall be solved numerically. The result of this problem can also be determined in an alternative way, thus enabling a comparison with the result of the developed homotopy method. For the sake of a meaningful graphic illustration we have chosen an example of a bicriterial problem.

On the other hand, the chapter shall demonstrate that the method makes the solution of real application problems possible. Actually, the developed homotopy method is already in use in the industrial sector of the SIEMENS company. In particular, one manages with its aid to solve numerically the two problems discussed in Chapter 2, the design optimization of a combined-cycle power plant and the optimization of the operating point of a recovery-boiler. The Sections 7.2 and 7.3 present the results of these calculations.

7.1 Example 1 (academic)

We are searching for the set of efficient solutions of the following bicriterial objective function¹:

$$f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^2 \\ \mathbf{x} & \mapsto \begin{pmatrix} \cos(a(\mathbf{x})) \cdot b(\mathbf{x}) \\ \sin(a(\mathbf{x})) \cdot b(\mathbf{x}) \end{pmatrix}, \text{ with} \end{cases} \quad (7.1)$$

$$a(\mathbf{x}) := \frac{2\pi}{360} [a_c + a_1 \cdot \sin(2\pi x_1) + a_2 \cdot \sin(2\pi x_2)] \quad (7.2)$$

$$b(\mathbf{x}) := 1 + d \cdot \cos(2\pi x_1) \quad (7.3)$$

In the computed example the following values were assigned to the constants a_c, a_1, a_2 and d : $a_c = 45$, $a_1 = 40$, $a_2 = 25$ and $d = 0.5$. The variable space is not

¹ The author would like to thank Dr. mult. Reinhart Schultz [SCHULTZ, 1998] for having communicated this optimization problem to him.

limited by any constraints.

As both variables x_1 and x_2 enter the objective function f only as arguments (angles in radian measure) of trigonometric functions, f is periodical with period 1 with regard to both variables. The search space can therefore be limited, without loss of generality, to the square $[0, 1] \times [0, 1] \subset \mathbb{R}^2$. In particular, one obtains a fairly precise representation of the image set $f(\mathbb{R}^2) \equiv f([0, 1] \times [0, 1])$, if one covers this square with a fine grid and plots the images of the grid points under the mapping f . Figure 7.1 shows the resulting image set of f .

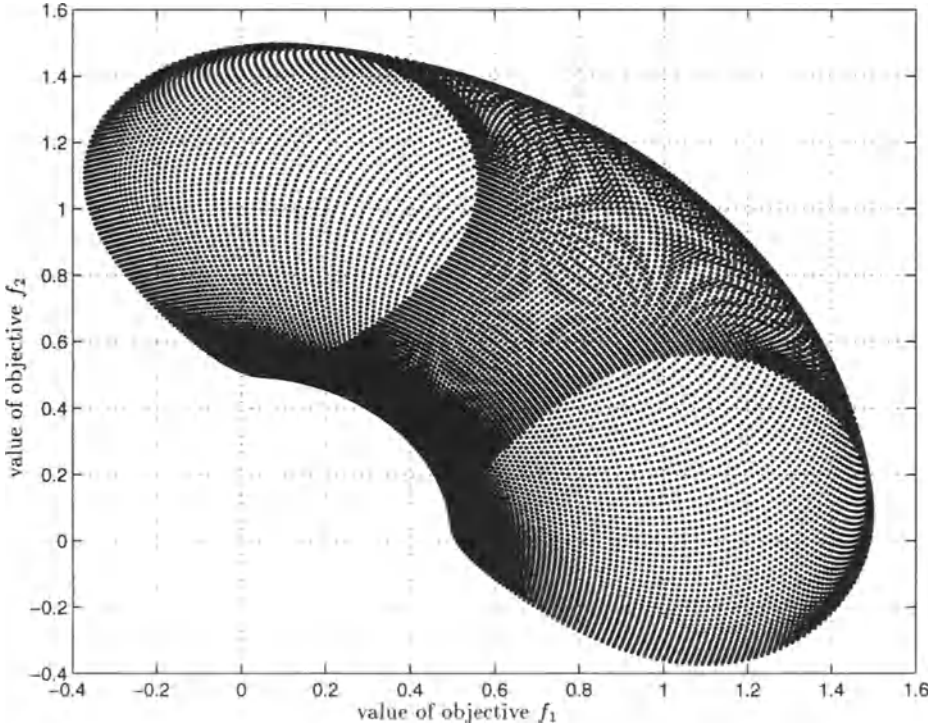


Figure 7.1: Image set $f(\mathbb{R}^2)$ of the example function 1

In order to compute by homotopy the ‘efficient curve’ (the set of efficient points), which can be gathered from Figure 7.1, a starting point $(\mathbf{x}^*, \boldsymbol{\alpha}^*) \in M$ is required. We choose it such that $\mathbf{x}^*$ is a stationary point of the convex combination $g_{\boldsymbol{\alpha}^*}$ to the weight vector $\boldsymbol{\alpha}^* = (0.5, 0.5)^T$. The search for a minimizer of $g_{\boldsymbol{\alpha}^*}$ by a (damped and regularized) Newton method, with the starting point $\mathbf{x}_0 = (0, 0)^T$, leads to the efficient point [in the objective space], which is marked by a ‘+’ in Figure 7.2, partial figure on the upper left.

Now the basis vector $\mathbf{q}$ of the one-dimensional tangent plane $T_{(\mathbf{x}^*, \boldsymbol{\alpha}^*)}M$ [i.e. of a straight line] is determined, a fixed steplength $\xi_0 = 0.06$ is chosen and a sequence

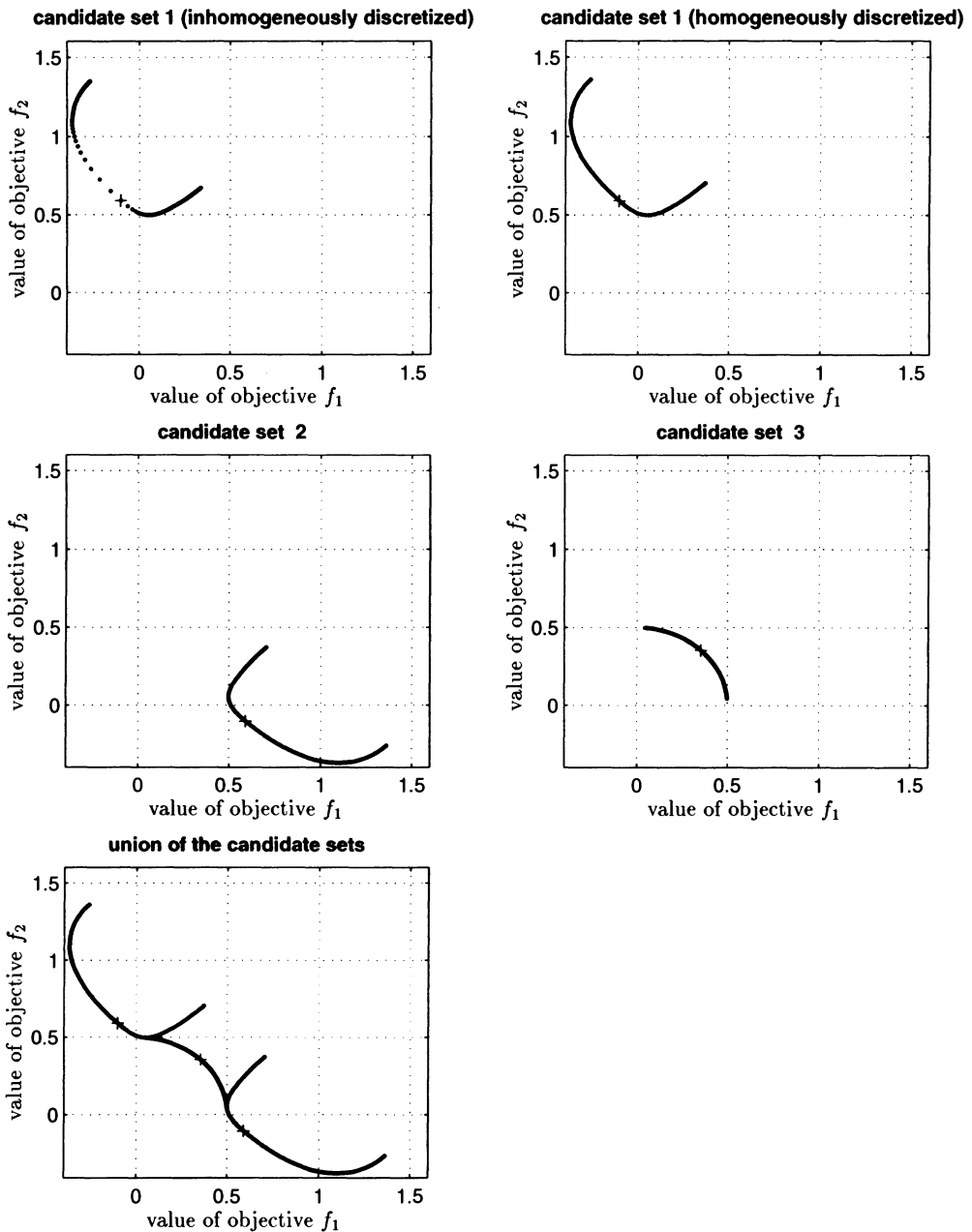


Figure 7.2 Candidates for efficient points in the objective space. The entire candidate set is composed of three partial zero manifolds, which are denoted as candidate set 1, 2 and 3. Candidate set 1 is determined once without re-scaling of the chart parameters ξ (upper left), once with re-scaling (upper right).

of homotopy steps according to the algorithm described in Paragraph 6.1.4 is carried out. In each step one avoids going back on the efficient curve: Let l be the index of the current homotopy step. If $[(\mathbf{x}, \boldsymbol{\alpha})_{(l)} - (\mathbf{x}, \boldsymbol{\alpha})_{(l-1)}]^T \cdot \mathbf{q} < 0$, the chart parameter $\xi_{l+1} = -\xi_0$ is chosen instead of the chart parameter $\xi_{l+1} = \xi_0$. In this way, by starting from $(\mathbf{x}^*, \boldsymbol{\alpha}^*)$, 300 homotopy steps are made in both directions². The result [in the objective space] is the ‘candidate set 1’ (partial figure on the upper left in Figure 7.2).

Two things are striking:

- (i) Obviously, the candidate set 1 is only a subset of the efficient curve.

[On the other hand, not all points of the candidate set 1 are efficient. The main reason for this is that negative $\boldsymbol{\alpha}$ -components were also admitted to indicate the further course of the candidate curve 1 (as a part of the entire zero manifold). In addition to that, the candidate set 1 also contains some points that are locally efficient (being minimizers of a convex combination $g_{\boldsymbol{\alpha}}$), but not globally efficient (being situated in the ordering cone of points from the candidate set 3).]

- (ii) The discretization of the efficient set is inhomogeneous.

In order to remedy defect (ii), we replace the fixed steplength ξ_0 with a steplength control according to the re-scaling rule of Paragraph 6.1.3. The result plotted in the upper right partial figure shows that a homogeneous discretization of the efficient curve can actually be obtained in this manner. Also, the number of homotopy steps required for an adequate resolution of the efficient curve is substantially reduced (100 steps instead of 300).

In order to obtain the remainder of the efficient curve (see point (i)), we repeat the same method steps, starting this time from the point $(\mathbf{x}^*, \boldsymbol{\alpha}^*) = (0.75, 0.6, 0.5, 0.5) \in M$. As a result we get the candidate set 2 (central left partial figure in Figure 7.2). The image $\mathbf{f}(\mathbf{x}^*)$ of the starting point is again marked by a ‘+’.

The candidate sets 1 and 2 are both bent ‘inwards’ [i.e. they each form the boundary of a convex subset of the image set $\mathbf{f}(\mathbb{R}^2)$] and consequently consist of (local) *minima* of linear combinations $g_{\boldsymbol{\alpha}}$, according to the argumentation of Section 4.4. The still missing subset of the efficient curve consists, to judge by its curvature (see Figure 7.1), of *saddle points* of corresponding linear combinations $g_{\boldsymbol{\alpha}}$. To compute this subset we carry out the above method steps a third time. Starting point $(\mathbf{x}^*, \boldsymbol{\alpha}^*) \in M$ is now the saddle point $\mathbf{x}^* = (0.5, 0.5)^T$ of $g_{\boldsymbol{\alpha}^*} \equiv g_{(0.5, 0.5)}$. The central right partial figure of Figure 7.2 shows the result, the candidate set 3. In order to confirm that candidate set 3 consists indeed of saddle points, in the central partial figure of Figure 7.3 the eigenvalues of the Hessian matrix $\nabla^2 g_{\boldsymbol{\alpha}}(\mathbf{x})$ were plotted against the iteration index of the homotopy steps³. Since the Hessian

² In contrast to the description of the algorithm in Paragraph 6.1.4, negative $\boldsymbol{\alpha}$ -components (which correspond to an inversion of the sign of individual objectives) are also admitted.

matrix is evidently indefinite along the entire candidate set [besides two points, which we will speak later of], the saddle point property has been proven.

The union of the three candidate sets for efficient points which have now been determined by homotopy is shown in the lower left partial figure of Figure 7.2. A comparison with the image set $\mathbf{f}(\mathbb{R}^2)$ shows that this sum-set includes the entire efficient curve (in objective space). As the above discussion makes clear, the example presented is already a non-trivial case of a vector optimization problem: The set of efficient points is composed of several (namely three) one-dimensional candidate manifolds (more precisely: connection components).

It can be gathered from the plot of the sum-set that both the candidate sets 1 and 3 and the candidate sets 2 and 3 have each one point [in the objective space] in common. An examination of the three candidate sets in the $(\mathbf{x}, \boldsymbol{\alpha})$ -space reveals these as three one-dimensional manifolds (curves), which intersect also in the inverse images of the common objective-points. In both points of intersection the zero manifold M cannot have locally the character of a one-dimensional differentiable manifold, as there exists no unambiguous local parametrization (chart) of M here. Consequently, in these points of intersection the Rank Condition (5.6) must be violated, i.e. the Jacobian matrix $\mathbf{F}'(\mathbf{x}, \boldsymbol{\alpha})$ must have a rank smaller than (the full rank) 3. An important question is now, whether our numerical homotopy method indicates clearly such a change of the dimension of the candidate manifold M , which opens up the possibility of a bifurcation.

To answer this question, in Figure 7.3 (upper partial figure) the minimum of $|(\mathbf{R}_1)_{jj}|$, $j \in \{1, 2, 3\}$, where $\mathbf{R}_1$ denotes the triangular matrix resulting from the $\mathbf{QR}$ -factorization of $(\mathbf{F}'(\mathbf{x}, \boldsymbol{\alpha}))^T$ [see step (3) of the algorithm in Paragraph 6.1.4], is plotted against the homotopy steps carried out to determine the candidate set 3. In fact, this minimum is zero in two points. The comparison with the lower partial figure (of Figure 7.3), which plots the corresponding α_1 -values, shows that both these points are exactly the points of intersection of the candidate set 3 with the two other candidate sets.

In order to round off the discussion of example 1, let us point out three things:

- The central partial figure of Figure 7.3 reveals why the rank of the Jacobian matrix $\mathbf{F}'(\mathbf{x}, \boldsymbol{\alpha})$ breaks down in both points of intersection of the candidate set: An eigenvalue of the Hessian matrix $\nabla^2 g_{\boldsymbol{\alpha}}(\mathbf{x})$ equals zero in both these points, and no gradient of an individual objective function is there to compensate for this rank deficit. As the discussion in Section 5.2 shows, this is a non-generic behavior. The numerical example 2 demonstrates the (generic) case, when the zero transition of an eigenvalue of the Hessian matrix is not connected with a jump in the dimension of the manifold M .
- Obviously the homotopy method has no difficulty in skipping both bifurcation (or intersection) points and in proceeding on the relevant partial

³ The iteration index 0 corresponds to the starting point $(\mathbf{x}^*, \boldsymbol{\alpha}^*)$, the index sign indicates the direction of the progression on the one-dimensional manifold (curve) M .

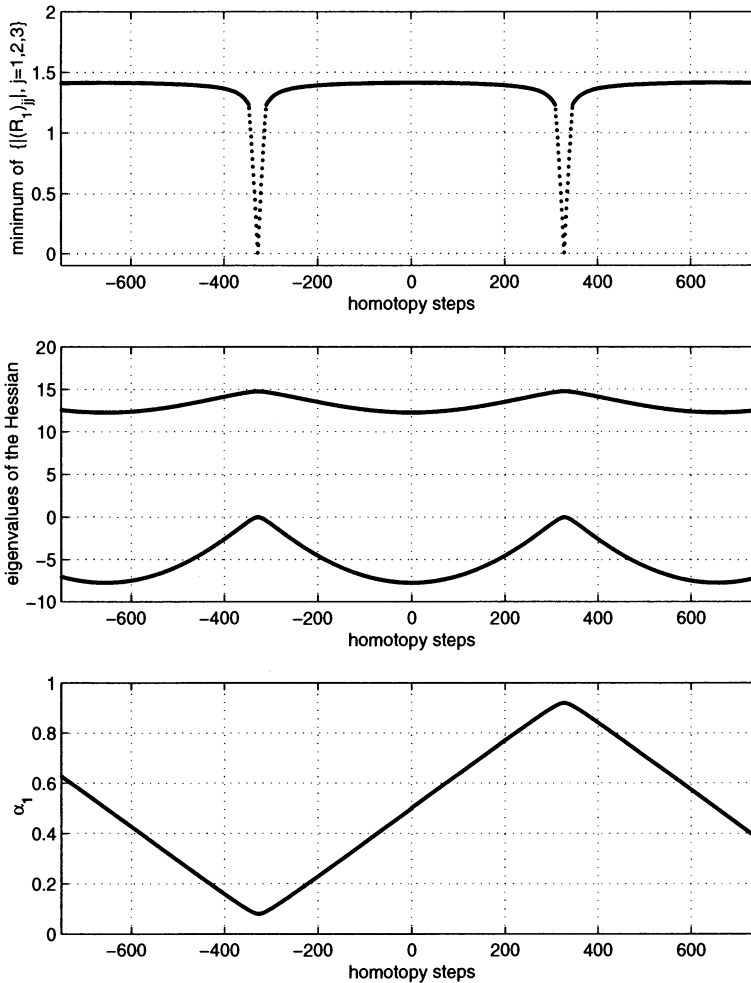


Figure 7.3 Record of three quantities during the generation of the candidate set 3 by means of the homotopy method. The sign of the iteration indices (denoted as homotopy steps) indicates the direction of the progression on the curve M , the index 0 corresponds to the starting point $(\mathbf{x}^*, \alpha^*)$.

curves of the manifold M . This behavior is in accord with arguments of [KELLER, 1977] and [ALLGOWER & GEORG, 1990]. There it is demonstrated that one-dimensional homotopy methods of the Euler-Newton-type can skip simple bifurcation points without difficulty, as for sufficiently small steplengths the predictor step leads into the ‘attraction cone’ of the Newton-corrector.

- The candidate set 3 has two ends in the objective space. The question arises whether the corresponding inverse-image set, more precisely: the corresponding subset of the zero manifold M in the $(\mathbf{x}, \alpha)$ -space, has this

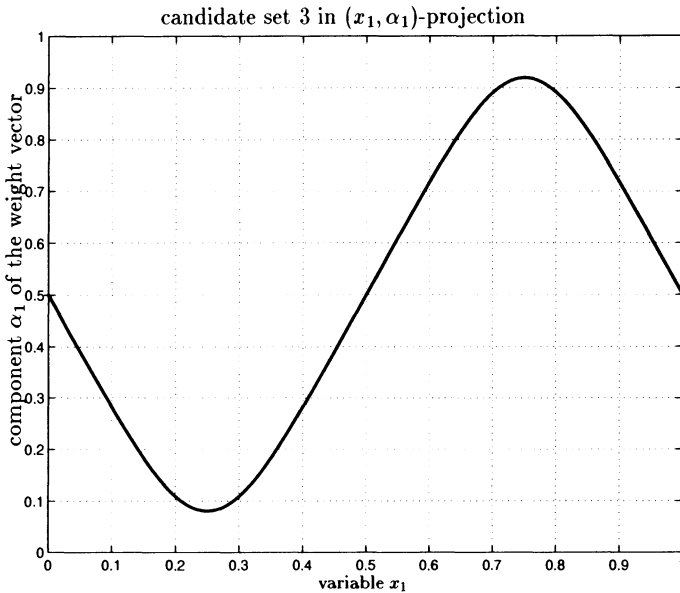


Figure 7.4 Projection of the candidate set 3 (in the extended variable space) onto the plane spanned by the x_1 - and the α_1 -axis.

property, too.

Figure 7.4 shows the projection of this subset onto the (x_1, α_1) -plane. Both extrema of α_1 correspond to the bifurcation points of M (compare the lower partial figure of Figure 7.3). If we start from the point $(x_1^*, \alpha_1^*) = (0.5, 0.5)^T$, generate the curve section situated between the extrema of α_1 by homotopy and map it into the objective space, we obtain the complete candidate set 3. If one proceeds (by homotopy) beyond one of the extrema of α_1 , one moves [in the objective space] from one end of the candidate set 3 back to the center and finds the same situation as at the starting point (x_1^*, α_1^*) because of the periodicity of the objective function f . The periodical objective function thus maps a zero manifold, which is unbounded (in the x_1 -values), into the candidate set 3 with its two ends.

7.2 Example 2: Design of a Combined-Cycle Power Plant

In Section 2.1 the pinch-point design of a combined-cycle heat recovery boiler was presented as an industrial application problem of vector optimization. This problem will now be solved by homotopy.

The three design variables ‘high pressure(hp)-pinch-point’, ‘medium pressure(mp)-pinch-point’ and ‘low pressure(lp)-pinch-point’ are the variables to be optimized. Let the triple of these variables be denoted briefly by $\mathbf{x} \in \mathbb{R}^3$.

The negative (thermodynamical) efficiency and the investment costs implied by these design variables are criteria for the assessment of a value of $\mathbf{x}$. Both objectives have to be minimized, i.e. the set of efficient points with regard to the vector-valued objective function

$$\mathbf{f}_{\text{power plant}}(\mathbf{x}) \begin{pmatrix} \text{investment costs}(\mathbf{x}) \\ -\text{efficiency}(\mathbf{x}) \end{pmatrix} \quad (7.4)$$

is what we are looking for.

Since both individual objectives are *contrary* to one another in the sense that the investment costs criterion prefers large pinch-points and the (thermodynamical) efficiency criterion small ones (see Section 2.1), the efficient points of this objective function will lie in the range of technically reasonable pinch-points and, therefore, the set of feasible variable values must not be limited by constraints.

As a first step the functions $\text{efficiency}(\mathbf{x})$ and $\text{investment costs}(\mathbf{x})$ have to be provided in the form of a model which reflects sufficiently precisely both the physical correlations within the power plant and the price structure for the production of the relevant components. For computing the thermodynamic equilibrium — which is the solution of a non-linear system of equations — and the geometry of the heat exchanger surfaces corresponding to this equilibrium one can fall back upon a simulator program, which the manufacturer of power stations SIEMENS KWU uses for checking its power plant designs. The resulting mappings $\text{efficiency}(\mathbf{x})$ and $\text{investment costs}(\mathbf{x})$, which exist in the form of a computer program, cause, however, long computing times for each function evaluation (simulator program!) and have some small discontinuities (coming to light with high numerical resolution), which originate, for example, from breaking off iteration loops in the simulator.

For acceleration and smoothing both components of the objective function are therefore approximated by an (at least) twice continuously differentiable model. To this purpose, a grid with a fine mesh covers the space of technically meaningful pinch-points (a compact subspace of the $\mathbb{R}^3$), and for each of these variable points the exact (simulation) model is evaluated. The data pairs $(\mathbf{x}, \mathbf{f}_{\text{power plant}}(\mathbf{x}))$ obtained by that procedure serve to train (i.e. to fit by means of regression) a 3-layer-perceptron [with a tanh-activating function of the central neuron layer]. The resulting mapping $\mathbf{f}_{\text{approx}}$ is a neuronal approximation model of the function

$\mathbf{f}_{power\ plant}$ and is defined as

$$(f_{approx})_j = \mathbf{g}_3 \circ \mathbf{g}_2 \circ \mathbf{g}_1, \text{ for } j = 1, 2, \text{ where} \quad (7.5)$$

$$\mathbf{g}_1 : \begin{cases} \mathbb{R}^3 \rightarrow \mathbb{R}^{20} \\ \mathbf{x} \mapsto \mathbf{A}_j \cdot \mathbf{x} + \mathbf{b}_j \end{cases}$$

$$\mathbf{g}_2 : \begin{cases} \mathbb{R}^{20} \rightarrow \mathbb{R}^{20} \\ (\dots, x_i, \dots)^T \mapsto (\dots, \tanh(x_i), \dots)^T \end{cases}$$

$$\mathbf{g}_3 : \begin{cases} \mathbb{R}^{20} \rightarrow \mathbb{R} \\ \mathbf{x} \mapsto \mathbf{c}_j^T \cdot \mathbf{x} + d_j \end{cases}$$

Here, the matrix $\mathbf{A}_j \in \mathbb{R}^{20 \times 3}$, the vectors $\mathbf{b}_j, \mathbf{c}_j \in \mathbb{R}^{20}$ and the real number d_j represent the parameters (neural weights) obtained by the training procedure.

In the range of relevant variable values $\mathbf{f}_{approx}$ reproduces the mapping $\mathbf{f}_{power\ plant}$ (as determined by simulation) in an excellent way and, therefore, is used instead of $\mathbf{f}_{power\ plant}$ for the following calculations.

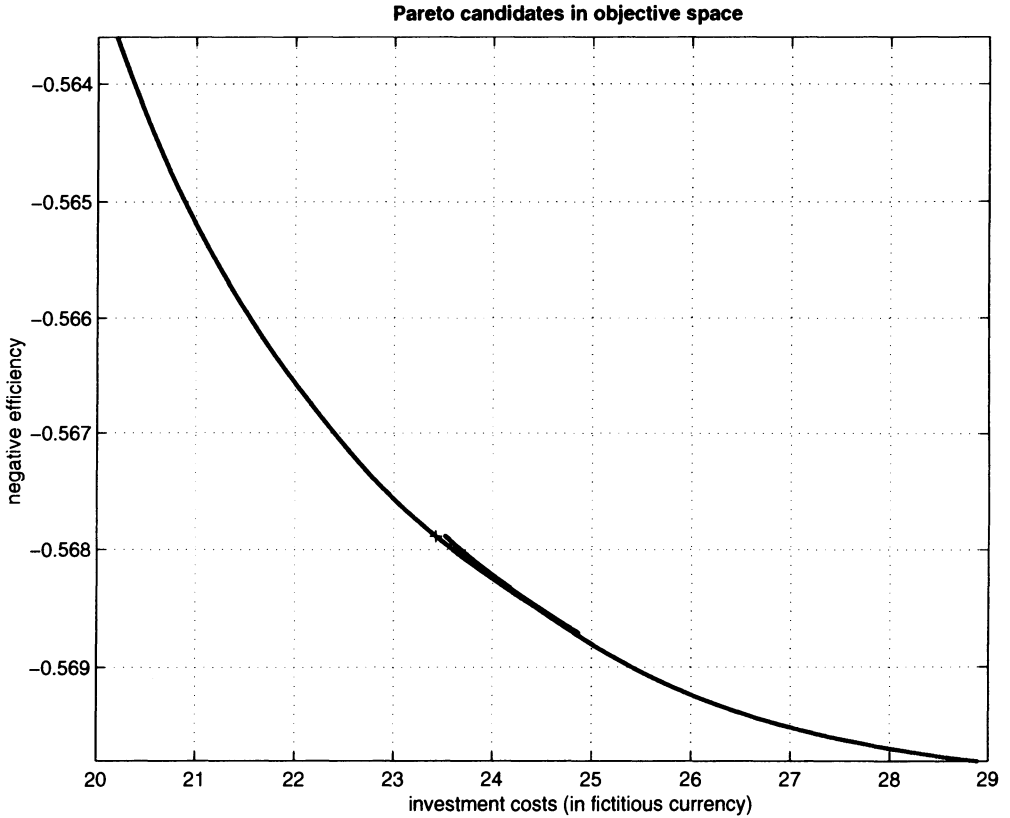


Figure 7.5: Candidates for efficient points in the objective space (power plant design).

Since the example problem 2 (as already example 1) is bicriterial, we can expect that the efficient solutions will form an efficient curve. To obtain it we

start from a minimizer of the convex combination g_{α^*} to the weight vector $\alpha^* = (0.5, 0.5)^T$ and carry out the same algorithmic steps as in example 1. The result, a curve of candidates for efficient pinch-point designs [in the objective space], is plotted in Figure 7.5.

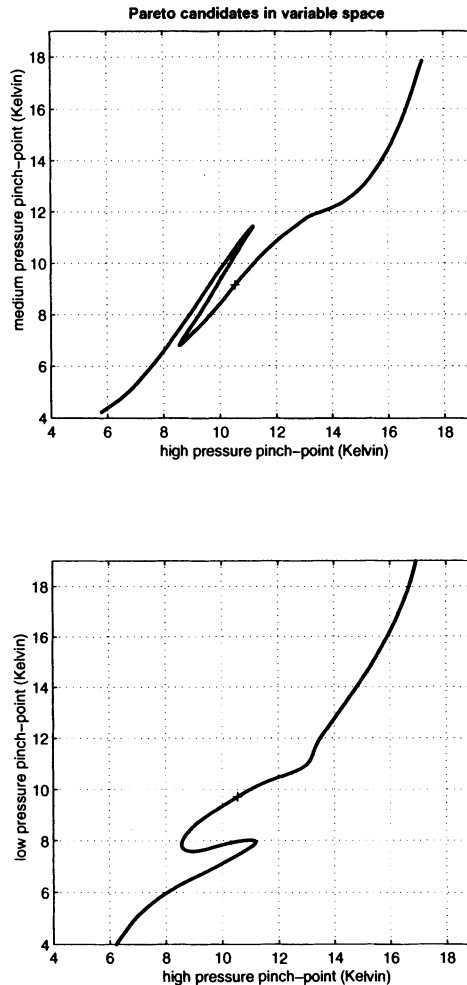


Figure 7.6 Candidates for Pareto optimal points in the variable space (power plant design). The upper figure shows the projection of these candidate points onto the plane, which is spanned by the axis of the high pressure pinch-points and the axis of the medium pressure pinch-points. Analogously, the lower figure shows the projection onto the plane spanned by the axis of the high pressure pinch-points and the axis of the low pressure pinch-points.

The knowledge of this efficient curve is profitable for the power plant manufacturer in many respects. First it enables him to determine the design optima (with regard to the electricity production costs) in the sense of a parametric analysis for

all scenarios of future annuities, fuel prices and marketable electricity quantities [see Equation (2.1) on page 11]. Furthermore, the technical sales department can discuss directly with the customer (and future operator of the projected power plant) about design variants with reference to this efficient curve, and the customer can determine the design according to his own priorities. In the third place, one can learn from the form of the efficient curve in Figure 7.5 that there is a limited range of economically reasonable efficient solutions. If one intends to force down the investment costs (for the heat recovery boiler and parts of the cooling system) substantially below 21 units (of some fictitious currency), one has to accept a great loss of thermodynamical efficiency. On the other hand, for efficiency increases beyond the mark of 56.95% (assuming certain basic conditions in the power plant) one has to pay with a disproportionate increase of the investment costs.

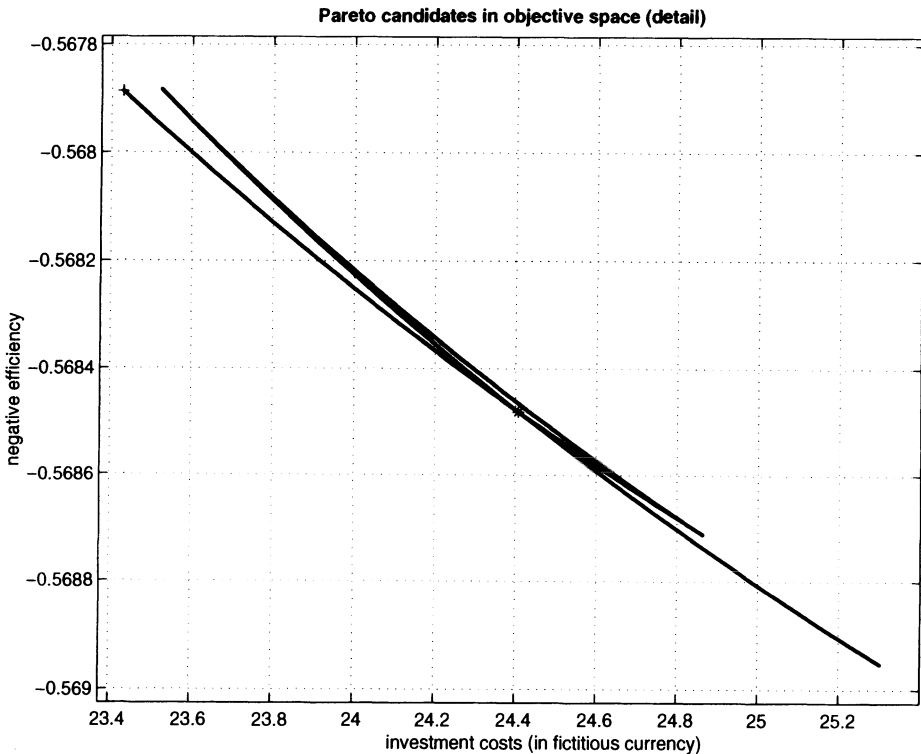


Figure 7.7 Detail of the candidate set for efficient points, which has been plotted in Figure 7.5, in the objective space (power plant design).

Except for a conspicuousness in the investment costs interval between 23 and 25 (fictitious) currency units, the appearance of the efficient curve (Figure 7.5) tempts one to conjecture that the efficient design points follow a very simple rule of formation. The inverse image of this efficient curve (more precisely: the

candidate set for Pareto optimality) in the variable space, which is plotted in Figure 7.6 in two different projections, does not at all, however, have a trivial shape.

Now we will examine more closely the mentioned conspicuousness in the investment costs interval between 23 and 25 currency units. For this purpose, we generate again the candidate curve (for efficient solutions) for this part of the objective space and increase the density of the discretization of the curve by reducing the predictor steplength. The selective enlargements of Figures 7.5 and 7.6 thus obtained are plotted in Figures 7.7 and 7.8. One can see⁴, that the candidate curve [in the objective space] consists of three arcs.

The examination of the candidate set in the objective space (Figure 7.7) reveals the following behavior: If one starts from the point marked by '+' [the minimizer to the convex combination $g_{(0.5,0.5)}$] and moves along the zero manifold M (by homotopy), one passes through a curve section (which has the form of a compressed ∇ -operator), the points of which do not represent globally efficient pinch-point designs. The point, in which the globally efficient and the not globally efficient part of the candidate set meet, is marked in Figure 7.7 by '*'. Since this point in the objective space is the point of intersection of the candidate curve with itself, it has two inverse-images in the variable space. These are also marked by '*' in Figure 7.8.

In the variable space the situation is as follows: If one starts the homotopy method from the point '+' in the direction of increasing thermodynamical efficiencies, one obtains globally efficient solutions until one gets to the first design point⁵ marked by a '*'. After this point has been passed, the homotopy method supplies solutions which are not globally efficient until the second (*)-point has been reached. After having passed this point the generated candidate points are globally efficient again.

We will now shed light upon the character of the curve section between both (*)-points.

In Figure 7.9 a number of quantities are plotted, which were recorded during the generation of the candidate set shown in Figures 7.7 and 7.8. If we examine the eigenvalues of the Hessian matrix, we notice two zero transitions of an eigenvalue. Up to the first zero transition all eigenvalues are positive, the Hessian matrix is thus positively definite and the generated candidate points are *minima* of a convex combination g_α . Up to that point, which corresponds to the first S-bend in both projections of Figure 7.8 (or the right upper corner of the ' ∇ ' in

⁴ The character of the candidate curve discloses itself best when one looks at a small angle from right below to the diagonal of the figure and screws up one eye.

⁵ This pinch-point design leads to the same result with regard to both objectives as the second design point marked by a '*'. For a power plant the technical significance is that the two pinch-point combinations 1 (low hp- and mp-pinch-points, but high lp-pinch-point) and 2 (high hp- and mp-pinch-points with a low lp-pinch-point) imply an identical thermodynamical efficiency and identical expenditure for heat exchanger surfaces.

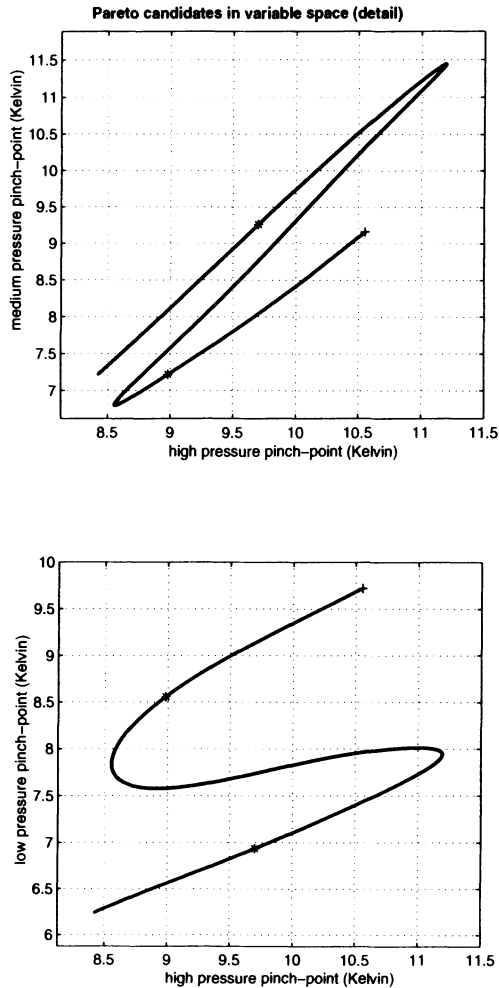


Figure 7.8 Detail of the set of candidates for Pareto optimality, which has been plotted in Figure 7.6, in the variable space (power plant design).

Figure 7.7), all candidate points are at least locally efficient. Between both zero transitions the Hessian matrix is indefinite (and regular), the generated points are hence *saddle points* of a convex combination g_α . It is interesting to note that the curve formed by these saddle points in the objective space (see Figure 7.7) has a positive curvature. According to Theorem 4.6 this fact is incompatible with the local efficiency of these points. The generated saddle points are therefore candidate points, which are *not* locally efficient. From the second zero transition (the second S-bend or the left upper corner of the ‘ ∇ ’) up, all eigenvalues are again positive, the generated candidate points are therefore at least locally (and from the (*)-point up also globally) efficient.

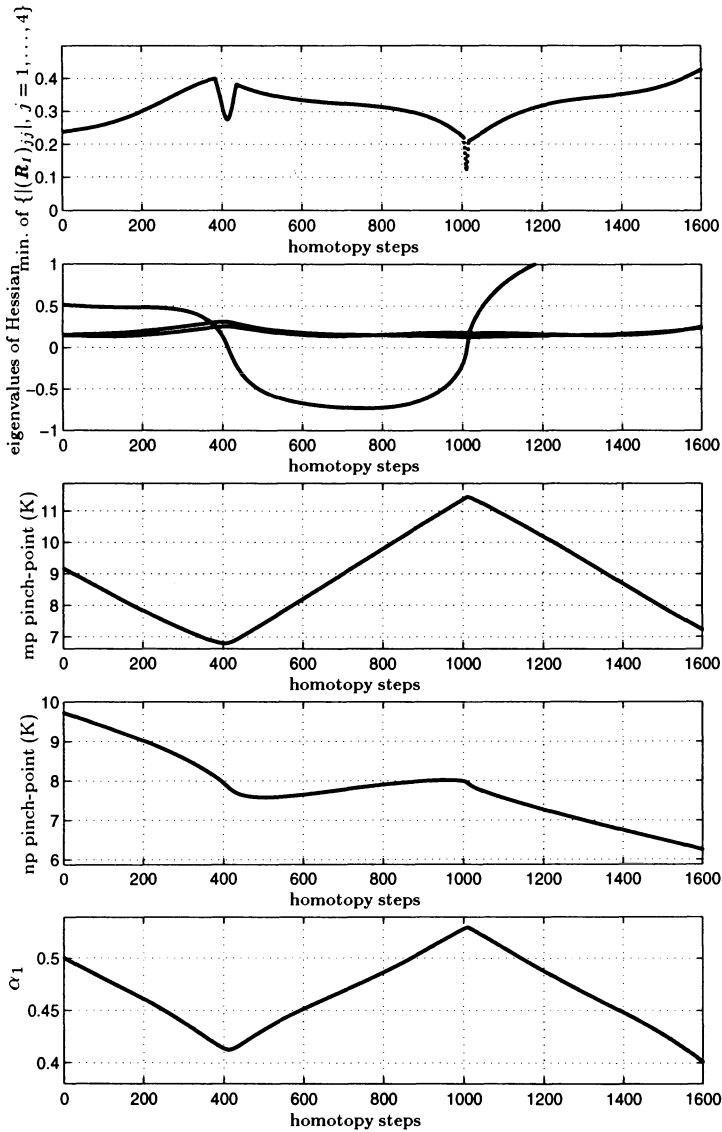


Figure 7.9 Record of five quantities during the generation (by means of the homotopy method) of the candidate set plotted in Figures 7.7 and 7.8.

In view of two zero transitions of eigenvalues of the Hessian matrix it is, of course, an interesting question, whether the dimension of the manifold M is conserved at these points. As is indicated by the minima of $\{(|\mathbf{R}_t|_{jj}|, j \in \{1, 2, 3, 4\})$ (upper partial figure of Figure 7.9), which are clearly off zero everywhere, this is the case. The problem of the pinch-point design supplies a numerical instance

of the assertion, which was proved in Section 5.2, that a zero transition of an eigenvalue of the Hessian matrix does not, in general, lead to a change (or jump) of the dimensionality of M .

As one can gather from the lower partial figure of Figure 7.9, every zero transition of an eigenvalue is connected to an inversion of the trend of the α_1 -values, when moving along the manifold M . This originates from the fact that in these transition points the two last columns of the Jacobian matrix $\mathbf{F}'(\mathbf{x}, \boldsymbol{\alpha})$, which form the submatrix $\frac{\partial \mathbf{F}}{\partial \boldsymbol{\alpha}}$, are required to fill up the rank of $\mathbf{F}'(\mathbf{x}, \boldsymbol{\alpha})$. According to the implicit-function theorem M can no longer be parametrized locally by α_1 , i.e. α_1 loses its character as a potentially free (homotopy) parameter and becomes a dependent variable. Therefore it is in particular no longer guaranteed that the user of the homotopy method can continue to reduce or increase α_1 beyond these transition points.

7.3 Example 3: The Optimal Operating Point of a Recovery-Boiler

Taking up the description of the problem in Section 2.2 we will now put in concrete terms and solve numerically the multiobjective optimization problem resulting from the operating point optimization of a recovery-boiler.

The three supplied streams of air (primary air, secondary air, tertiary air), which are henceforth combined to the variable vector $\mathbf{x} \in \mathbb{R}^3$, serve as optimization variables.

Indicators of the plant state are the four physical quantities o_2 (O_2 -concentration), so_2 (SO_2 -concentration), *steam* (mass flow of the generated steam) and *temp* (temperature of the char bed). All these four quantities are functions of the air stream vector $\mathbf{x}$. However, this functional dependence is not available as a physical model (see Section 2.2). As the functional relation is required only for a local range of the $\mathbf{x}$ -space, in which there are meaningful operating points and for which we have a great number of measurement data, a data based model formation is adequate (cf. [STURM & SCHÄFFLER, 1998]). Each of the four physical quantities is formulated as a quadratic function of $\mathbf{x}$, for instance,

$$o_2(\mathbf{x}) = a + \mathbf{b}^T \cdot \mathbf{x} + \mathbf{x}^T \cdot \mathbf{C} \cdot \mathbf{x} , \quad (7.6)$$

and the parameters of the linear regression model, combined in the scalar quantity a , the vector $\mathbf{b}$ and the symmetric matrix $\mathbf{C}$, are determined by the least squares method out of the available measurement data. In order to validate the regression approach, the absolute error of the regression function with regard to the data is computed (see [STURM & SCHÄFFLER, 1998]). It turns out that the four physical quantities can be described very well by their relative regression models within the range of meaningful operating points.

The ideal target values the plant operator has in mind for the four physical quantities — referred to the quantity of black liquor which has to be burnt in the concrete example — are:

$$o_{2,ideal} = 6.5 \text{ (\%)} \quad so_{2,ideal} = 1.2 \text{ (g/m}^3\text{)} \quad (7.7)$$

$$steam_{ideal} = 105 \text{ (t/h)} \quad temp_{ideal} = 995 \text{ (}^\circ\text{C)} \quad (7.8)$$

The vector-valued objective function consists of the quadratic deviations of the four quantities, resulting from the regression model, from these ideal values:

$$\mathbf{f}_{boiler}(\mathbf{x}) = \begin{pmatrix} (o_2(\mathbf{x}) - o_{2,ideal})^2 \\ (so_2(\mathbf{x}) - so_{2,ideal})^2 \\ (steam(\mathbf{x}) - steam_{ideal})^2 \\ (temp(\mathbf{x}) - temp_{ideal})^2 \end{pmatrix} \quad (7.9)$$

The space of feasible variable values $\mathbf{x}$ does not have to be limited by explicit constraints, as the efficient solutions will lie in a range of reasonable operating points because of the (so to speak ‘attractive’) character of the objective function.

In order to find a first efficient point of the objective function $\mathbf{f}_{boiler}$, all deviations, i.e. all components f_i , are first of all weighted equally [i.e. we set $\boldsymbol{\alpha}^* = (0.25 \ 0.25 \ 0.25 \ 0.25)^T$] and a minimizer $\mathbf{x}^*$ of the convex combination $g_{\boldsymbol{\alpha}^*}(\mathbf{x}) = \boldsymbol{\alpha}^{*T} \cdot \mathbf{f}_{boiler}(\mathbf{x})$ is determined (e.g. by the Newton method).

A set of further efficient solutions in the neighborhood of this operating point has to be computed next and the result has to be presented to the plant operator in an illustrative form. One obtains this by the following way of acting:

- Determine the (three-dimensional) tangent plane to the zero manifold M in the point $(\mathbf{x}^*, \boldsymbol{\alpha}^*)$ by computing numerically the three basis vectors $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ of this tangent plane according to the steps (2) to (5) in Paragraph 6.1.4.
- Carry out for all three basis vectors (subsequently called homotopy directions) the following steps (a) to (c):

(a) Fix a basic steplength ξ_0 (in the present example: $\xi_0 = 0.0025$).

(b) Choose a number N of homotopy steps (both for the forward and the backward direction). In the sense of the algorithmic step (6) in Paragraph 6.1.4, for each homotopy direction the following set of chart parameters $\{\boldsymbol{\xi}_{(i)}\}$ is provided [stated here for the homotopy direction $\mathbf{q}_1$]:

$$\{\boldsymbol{\xi}_{(i)}\} = \{(-N \cdot \xi_0, 0, 0)^T, (-(N-1) \cdot \xi_0, 0, 0)^T, \dots, (0, 0, 0)^T, \dots, (+(N-1) \cdot \xi_0, 0, 0)^T, (+N \cdot \xi_0, 0, 0)^T\} \quad (7.10)$$

Note: The point on M which corresponds to the chart parameter $\boldsymbol{\xi}_{(i)} = (0, 0, 0)^T$ is the starting point $(\mathbf{x}^*, \boldsymbol{\alpha}^*)$ (for each of the three homotopy directions).

- (c) Determine for every value $\xi_{(i)}$ of the chart parameter the associated point $\varphi(\xi_{(i)})$ of the manifold M according to the algorithmic steps (7) through (10) in Paragraph 6.1.4.

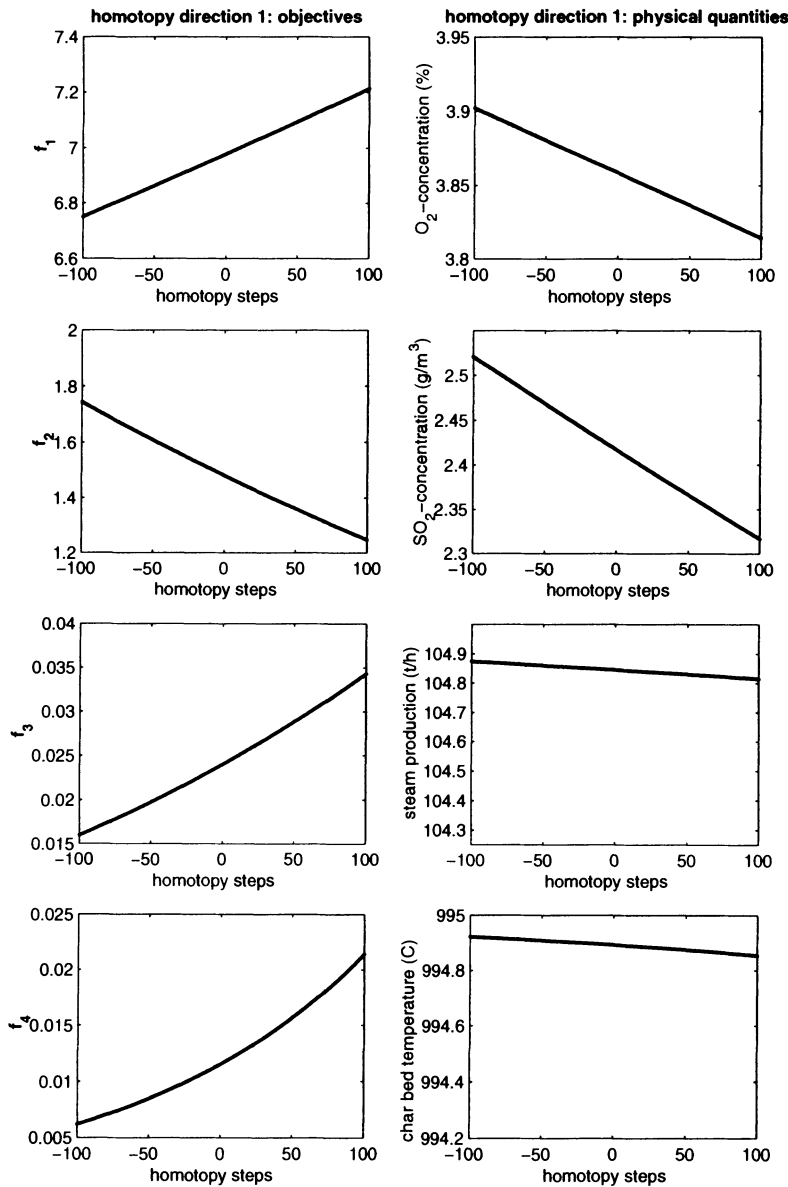


Figure 7.10 Variation of the objective function values (figures of left column) and the associated physical quantities (figures of right column) along the coordinate axis 1 of the tangent plane to the efficient manifold.

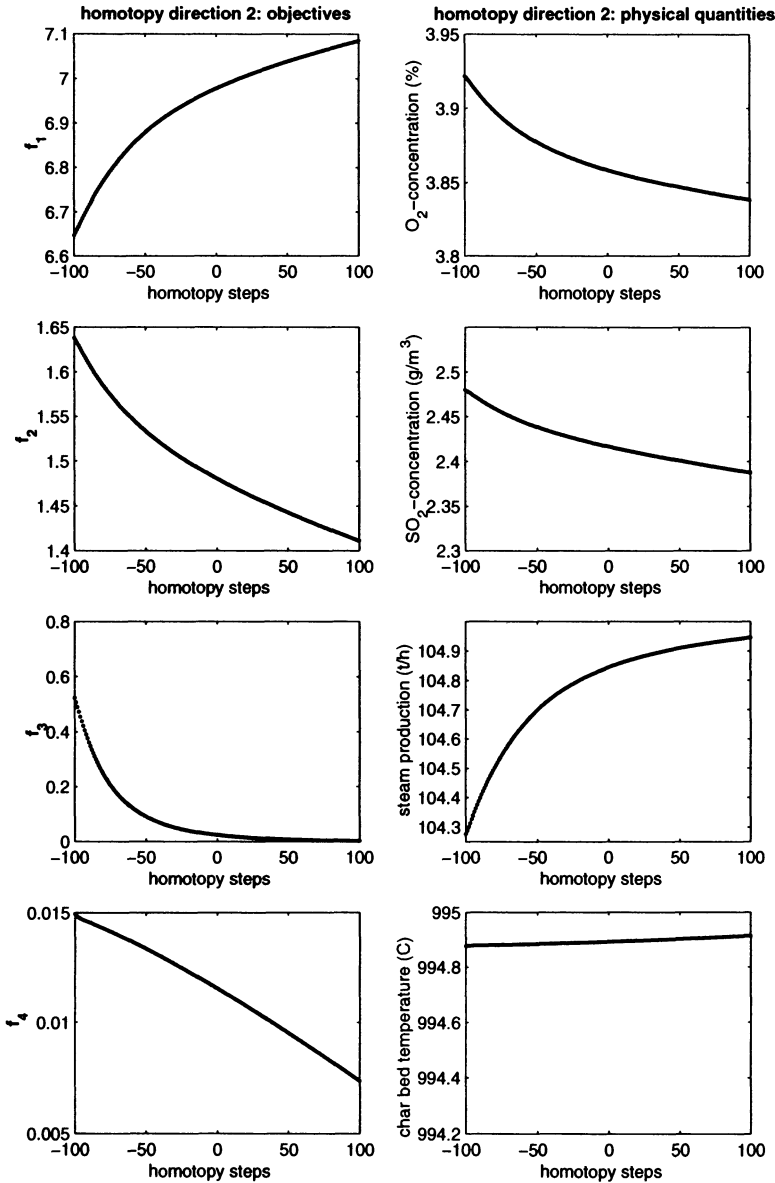


Figure 7.11 Variation of the objective function values (figures of left column) and the associated physical quantities (figures of right column) along the coordinate axis 2 of the tangent plane to the efficient manifold.

Figures 7.10 to 7.12 show the result of this procedure — one figure for each homotopy direction. On the abscissa of each partial figure the indices of the homotopy steps are indicated. The value 0 corresponds to the unchanged starting point (x^*, α^*) . The partial figures of the left column represent the values of

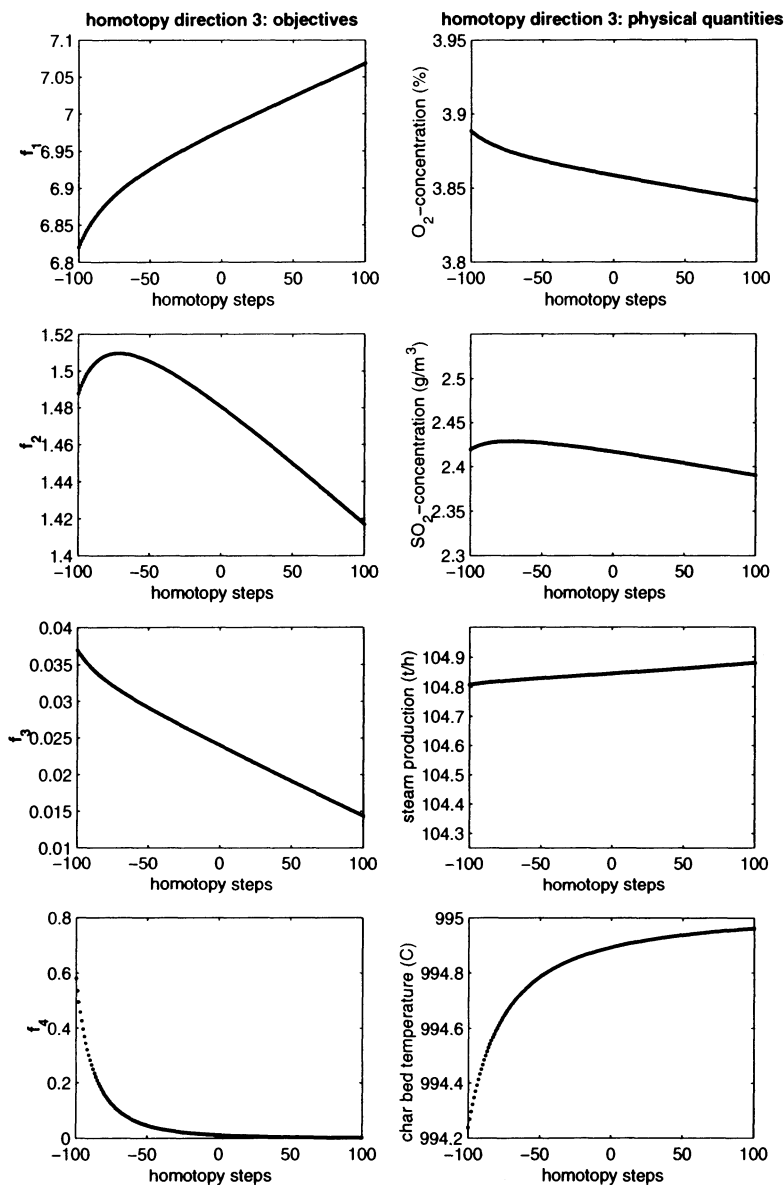


Figure 7.12 Variation of the objective function values (figures of left column) and the associated physical quantities (figures of right column) along the coordinate axis 3 of the tangent plane to the efficient manifold.

the four individual objectives which result when moving along the relative homotopy direction. On the right the values of the associated physical quantities (O_2 -concentration to temperature of the char bed) are shown.

From the course of these curves (for the physical quantities) the plant operator

can gain valuable information which backs up his decision in favor of an operating point (out of the neighborhood of $\mathbf{x}^*$).

To understand the significance of these curves compare⁶ the homotopy directions 1, 2 and 3, each gone along in negative direction. By proceeding in the (negative) homotopy direction 1, the O_2 -component of the operating point approaches (a bit) the desired value 6.5, but one has to buy this at the price of a deviation of the SO_2 -component from the desired value 1.2. The steam production and the temperature of the char bed remain more or less unchanged.

By proceeding in the (negative) homotopy direction 2, O_2 and SO_2 reveal a similar movement in opposite senses, but the deterioration with regard to the SO_2 -concentration turns out to be smaller. In return, also the steam production falls off in an unwelcome manner.

If one tries to improve the O_2 -value by proceeding in the (negative) homotopy direction 3, the SO_2 -concentration and the steam production remain approximately constant, but now the temperature of the char bed slightly decreases.

By virtue of this information the plant operator can now make an overall assessment based on his experience and his knowledge of the current urgency of the individual objectives and can make a sound decision in favor of the most appropriate operating point.

⁶ To facilitate comparison the ordinates of those partial figures, in which the values of the physical quantities are plotted, have the same scale for all three homotopy directions.

Bibliography

- [ALLGOWER & GEORG, 1990] Allgower, E. and Georg, K. (1990). *Numerical Continuation Methods*. Springer Verlag, Berlin-Heidelberg-New York.
- [ASH, 2000] Ash, R. (2000). *Probability and Measure Theory*. Harcourt/Academic Press, Burlington, Massachusetts.
- [BAUER, 1991] Bauer, H. (1991). *Wahrscheinlichkeitstheorie*. de Gruyter, Berlin-New York.
- [BEST ET AL., 1981] Best, Bräuninger, Ritter, and Robinson (1981). A globally and quadratically convergent algorithm for general nonlinear programming problems. *Computing*, 26, pages 141–153.
- [BÖWE & FURUMOTO, 1992] Böwe, J. and Furumoto, F. (1992). Laugenverbrennung und Chemikalienrückgewinnung beim Sulfatverfahren. *Technischer Bericht, Siemens AG, Unternehmensbereich ANL A221*.
- [CARMO, 1976] Carmo, M. d. (1976). *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Englewood Cliffs, New Jersey.
- [DAS, 1997] Das, I. (1997). *Nonlinear Multicriteria Optimization and Robust Optimality*. Dissertation, Rice University, Houston, Texas.
- [DAS & DENNIS, 1996A] Das, I. and Dennis, J. (1996a). A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. pages 1–12.
- [DAS & DENNIS, 1996B] Das, I. and Dennis, J. (1996b). Normal-boundary intersection: A new method for generating Pareto optimal points in multicriteria optimization problems. *Technical Report 96-11, Dept. of Computational and Applied Mathematics, Rice University, Houston, Texas*.
- [EDGEWORTH, 1881] Edgeworth, F. (1881). *Mathematical Psychics*. C. Kegan Paul & Co., London, England.
- [FISCHER, 1988] Fischer, H. (1988). Some aspects of automatic differentiation. In *Numerical Methods and Approximation Theory III, Proceedings of the Third International Conference on Numerical Methods and Approximation Theory*, pages 199–208. Nis, Yugoslavia.

- [FISCHER, 1996] Fischer, H. (1996). Automatic Differentiation: The Key Idea and an Illustrative Example. In Fischer, H., Riedmüller, B., and Schäffler, S., editors, *Applied Mathematics and Parallel Computing: Festschrift for Klaus Ritter*, pages 121–139. Physica-Verlag, Heidelberg.
- [FLETCHER, 1987] Fletcher, R. (1987). *Practical Methods of Optimization*. John Wiley, Chichester-New York.
- [FONSECA & FLEMING, 1995] Fonseca, C. and Fleming, P. (1995). An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary Computation*, 3(1), pages 1–16.
- [FORSTER, 1984] Forster, O. (1984). *Analysis 3*. Vieweg Verlag, Wiesbaden.
- [GARCIA & ZANGWILL, 1981] Garcia, C. and Zangwill, W. (1981). *Pathways to Solutions, Fixed Points, and Equilibria*. Prentice Hall, Englewood Cliffs.
- [GOLDBERG, 1989] Goldberg, D. (1989). *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley Publishing Company, Reading, Massachusetts, USA.
- [GÖPFERT & NEHSE, 1990] Göpfert, A. and Nehse, R. (1990). *Vektoroptimierung*. BSB Teubner Verlagsgesellschaft, Leipzig.
- [GROSSMANN & TERNO, 1993] Großmann, C. and Terno, J. (1993). *Numerik der Optimierung*. Teubner Verlag, Stuttgart.
- [HAIMES, 1973] Haimes, Y. (1973). Integrated system identification and optimization. *Control and Dynamic Systems: Advances in Theory and Applications*, 10, pages 435–518.
- [HÄMMERLIN & HOFFMANN, 1989] Hammerlin, G. and Hoffmann, K.-H. (1989). *Numerische Mathematik*. Springer Verlag, Berlin-Heidelberg-New York.
- [HASMINSKIJ, 1980] Hasminskij, R. (1980). *Stochastic Stability of Differential Equations*. Sijthoff and Noordhoff International Publishers, Alphen aan den Rijn.
- [JAHN, 1986] Jahn, J. (1986). *Mathematical Vector Optimization in Partially Ordered Linear Spaces*. Lang, Frankfurt.
- [JAHN, 1999] Jahn, J. (1999). *Introduction to the Theory of Nonlinear Optimization*. Springer Verlag, Berlin-Heidelberg-New York.
- [JÄNICH, 1992] Jänich, K. (1992). *Vektoranalysis*. Springer Verlag, Berlin-Heidelberg-New York.

- [KARUSH, 1939] Karush, W. (1939). *Minima of functions of several variables with inequalities as side conditions*. Master's Dissertation, University of Chicago.
- [KELLER, 1977] Keller, H. (1977). Numerical solution of bifurcation and nonlinear eigenvalue problems. In Rabinowitz, P., editor, *Application of bifurcation theory*, pages 359–384. Academic Press, New York, London.
- [KUHN & TUCKER, 1951] Kuhn, H. and Tucker, A. (1951). Nonlinear programming. In Neyman, J., editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, Berkeley.
- [LIN, 1976] Lin, J. (1976). Multiple objective problems: Pareto-optimal solutions by method of proper equality constraints. *IEEE Transactions on Automatic Control*, 21, pages 641–650.
- [LUENBERGER, 1984] Luenberger, D. (1984). *Linear and Nonlinear Programming*. Addison-Wesley Publishing Company, Reading, Massachusetts, USA.
- [MARGLIN, 1967] Marglin, S. (1967). *Public Investment Criteria*. MIT Press, Cambridge, Massachusetts.
- [PARETO, 1906] Pareto, V. (1906). *Manuale di Economica Politica*. Societa Editrice Libraria, Milano, Italy.
- [PARETO, 1971] Pareto, V. (1971). *Manual of Political Economy (English translation of 'Manuale di Economica Politica')*. MacMillan Company, New York.
- [PROTTER, 1990] Protter, P. (1990). *Stochastic Integration and Differential Equations*. Springer Verlag, Berlin-Heidelberg-New York.
- [RAKOWSKA ET AL., 1991] Rakowska, J., Haftka, R., and Watson, L. (1991). Tracing the efficient curve for multiobjective control-structure optimization. *Computing Systems in Engineering*, 2(6), pages 461–471.
- [RAO & PAPALAMBROS, 1989] Rao, J. and Papalambros, P. (1989). A nonlinear programming continuation strategy for one parameter design optimization problems. In *Proceedings of ASME Design Automation Conference*, pages 77–89. Montreal, Quebec, Canada.
- [RHEINBOLDT, 1986] Rheinboldt, W. (1986). *Numerical analysis of parametrized nonlinear equations*. John Wiley, Chichester-New York.
- [RITTER, 1998] Ritter, K. (1998). Private communication.
- [SAWARAGI ET AL., 1985] Sawaragi, Y., Nakayama, H., and Tanino, T. (1985). *Theory of Multiobjective Optimization*. Academic Press, Orlando, Florida, USA.

- [SCHÄFFLER, 1995] Schäffler, S. (1995). *Global Optimization Using Stochastic Integration*. S. Roderer Verlag, Regensburg.
- [SCHÄFFLER ET AL., 1999] Schäffler, S., Schultz, R., and Weinzierl, K. (1999). A stochastic method for the solution of unconstrained vector optimization problems. *Submitted to Journal of Optimization Theory and Application (JOTA)*.
- [SCHULTZ, 1998] Schultz, R. (1998). Private communication.
- [SCHWARZ, 1996] Schwarz, H. R. (1996). *Numerische Mathematik*. Teubner Verlag, Stuttgart.
- [SCHWETLICK, 1979] Schwetlick, H. (1979). *Numerische Lösung nichtlinearer Gleichungen*. Oldenbourg Verlag, München-Wien.
- [SCHWETLICK & KRETZSCHMAR, 1991] Schwetlick, H. and Kretzschmar, H. (1991). *Numerische Verfahren für Naturwissenschaftler und Ingenieure*. Fachbuch Verlag Leipzig.
- [STADLER, 1987] Stadler, W. (1987). Initiators of Multicriteria Optimization. In Jahn, J. and Krabs, W., editors, *Recent Advances and Historical Development of Vector Optimization*, pages 3–47. Springer Verlag, Berlin-Heidelberg-New York.
- [STADLER, 1988] Stadler, W., editor (1988). *Multicriteria Optimization in Engineering and in the Sciences*. Plenum Press, New York.
- [STRAUSS, 1994] Strauß, K. (1994). *Kraftwerkstechnik*. Springer Verlag, Berlin-Heidelberg-New York.
- [STURM & SCHÄFFLER, 1998] Sturm, T. and Schäffler, S. (1998). Datengetriebene Modellierung und Online-Optimierung eines Recovery-Boilers. *Technischer Bericht, Siemens AG, Zentralabteilung Technik, ZT PP 2*.
- [TIMMEL, 1980] Timmel, G. (1980). Ein stochastisches Suchverfahren zur Bestimmung der optimalen Kompromißlösungen bei statischen polykriteriellen Optimierungsaufgaben. *Wiss. Z. TH Ilmenau*, 6, pages 159–174.
- [WERNER, 1992] Werner, J. (1992). *Numerische Mathematik 1*. Vieweg Verlag, Braunschweig-Wiesbaden.
- [ZADEH, 1963] Zadeh, L. (1963). Optimality and non-scalar-valued performance criteria. *IEEE Transactions on Automatic Control*, 8.

Index

A

Active-set strategy, 65

B

Bicriterial optimization problem, 11, 26, 53

Black liquor, 12, 124

Bordered manifold, 50, 55, 58

Brownian motion, 38

C

Char bed, 12, 123

Chart, 47

change of $\sim$ s, 47

parameter, 49

CHIM-simplex, 25

Combined-cycle power plant, 10, 115

Constraint

equality $\sim$ s, 45, 65

inequality $\sim$ s, 45, 65

qualification, 46

surface, 68

Continuation method, *see* Homotopy method

Convex

analysis, 32

combination of the objectives, 4, 47

Corrector step, 87, 91, 98

Curvature

normal $\sim$, 51, 61

principal $\sim$, 53, 60

Curve

of dominated points, 31

of steepest descent, 31

D

Differentiable manifold, 47
bordered $\sim$, 50

Differential, 52

topology, 6, 47

Dominating point, 17

E

Edgeworth-Pareto optimality, *see*
Pareto optimality

Efficiency

plant $\sim$, 10

thermodynamical $\sim$, 11, 116

Efficient

curve, 53, 110

globally $\sim$, 17

locally $\sim$, 18

point, *see* Efficient, solution
set, 18

solution, 4, 18

Embedding, 89

ϵ -constraint method, 23

Euler method, 43

Evolutionary algorithm, 29

F

Feasible set, 15, 45

Fitness, 29

sharing, 30

G

Genetic operator, 29

H

Heat

exchanger surface, 11

recovery boiler, 11, 115

Hessian matrix, 69

Homogeneous discretization of the ef-
ficient set, 7, 93

Homotopy

approach to multiobjective opti-
mization, 27

method, 4
 classical $\sim$, 89
 parameter, 4
 Householder reflection, 96
 Hypersurface, 51

I

Implicit-function theorem, 76, 77, 85, 89
 Initial value problem, 31
 Investment costs, 11, 116

J

Jacobian matrix, 67

K

Karush-Kuhn-Tucker
 conditions, 46
 point, 5, 46

L

Lagrange multiplier, 65
 Lagrangian function, 4, 69, 72
 Lipschitz condition
 global $\sim$, 40
 local $\sim$, 32

M

Manifold, *see* Differentiable manifold
 bordered $\sim$, *see* Bordered manifold
 fold
 zero $\sim$, *see* Zero manifold
 Method
 of equality constraints, 24
 Mutation, 29

N

Neuronal approximation model, 116
 Newton method, 88
 simplified $\sim$, 98
 Normal field, 52
 Normal-Boundary Intersection, 25
 Numerical function, 37

O

Objective

function
 scalar-valued $\sim$, 4
 vector-valued $\sim$, 15
 space, 6

Optimality condition

of first order, *see* Karush-Kuhn-Tucker conditions
 of second order, 5, 47, 72

Optimization

multicriteria $\sim$, *see* Optimization, multiobjective
 multiobjective $\sim$, 3
 of models, 9
 of operating points, 9
 of system designs, 9
 parametric $\sim$, 4
 quadratic $\sim$ problem, 31
 scalar-valued $\sim$, 20, 46
 vector $\sim$, *see* Optimization, multiobjective

Order

partial $\sim$, 17
 conic $\sim$, 17
 relation, 16
 total $\sim$, 17

P

Paper production, 12
 Pareto
 optimal
 globally $\sim$, 17
 locally $\sim$, 18
 point, 17
 set, 18

Path of a stochastic process, 37

Perceptron, 116

Pinch-point, 11, 115

Population, 29

Power plant, 10

Predictor step, 87, 91, 98

QR-factorization, 96

R

Random time, 38

Recombination, 29
Recovery-boiler, 12, 123

S

Saddle point, 6, 60, 72
Scalarization, 20
Second fundamental form, 53
Selection, 29
Sequential quadratic programming,
 95
Slack variable, 65
Stochastic
 algorithm for vector optimization, 42
 differential equation, 31
 process, 37
 path of a $\sim$, 37
Submanifold, 48

T

Tangent
 plane, 49
 space, 49
 outward directed $\sim$, 50
 vector, 49

U

Unit circle, 81

V

Vector optimization, *see* Optimization, multiobjective

W

Weight
 neural $\sim$ s, 117
 vector, 5, 20, 53, 99
Weighted L_p -metric method, 21
Weighting method, 4, 20
Weingarten-mapping, 51
Wiener measure, 37

Z

Zero manifold, 6, 48