Kapitel 1

Predictive planning and systematic action – on the control of technical processes

1.1 Executive Summary

Since the beginning of the industrial revolution control engineering has been a key technology in many technical fields. James Watt's centrifugal governor for steam engines is one of the early examples of an extremely successful controller concept, of which at the end of the 1860s approximately 75 000 devices were in use only in England [2, p. 24]. Around this time, motivated by the increasing complexity of the plants that had to be controlled, engineers started to investigate systematically the theoretical foundations of control theory. The dynamic behavior of controlled systems, however, can only be understood and advanced with the help of mathematics, or as Werner von Siemens formulated: „Without mathematics you are always in the dark."

Therefore, in control engineering the question is not – today like more than one hundred years ago – *if*, but rather *which kind of* mathematics has to be used. In fact, from algebra over geometry and the theory of dynamic systems to optimization and numerics there is hardly any mathematical field that has not contributed significantly to control engineering and its neighboring mathematical disciplines systems and control theory.

In this article we want to point out two aspects. On the one hand, by means of two examples, we show the impact mathematics had in the long history of control engineering and which factors have been fundamental for the success of control methods. On the other hand – after a brief overview of the state-of-the-art in control engineering – we introduce a modern control methodology in more detail, namely *model predictive control* (MPC). We will briefly explain underlying mathematical concepts from systems theory, numerics, and optimization, and sketch some future challenges for mathematics.

In its linear form MPC has already become a standard tool in industrial applications, in particular in process engineering. In its nonlinear form (NMPC), on which we focus here, it is generally considered as one of the most promising modern methods for the control of complex technical processes. NMPC keeps finding its way into new application areas, due to the rapid progress on the theoretical as well as on the algorithmic side. Moreover, it is a prime example for a method that has reached its present state of development only by the interdisciplinary collaboration in the fields of control engineering, mathematical systems theory, numerics, and optimization, and whose future success will depend on an interdisciplinary approach, as well. Fortunately, such approaches are more and more supported by a variety of universities in clusters of excellence, research centers, or graduate schools. These joint interdisciplinary efforts in research and teaching will be an essential factor to tackle the future challenges and exploit the full potential of the production factor mathematics in control engineering.

1.2 A long success story

Mathematical methods play an essential role in control theory since the very beginning of this research field. We revise two prominent examples that illustrate this strong impact: the stability criterion of Hurwitz and Pontryagin's maximum principle. We will also discuss which factors made the mathematical developments this successful and are in a sense prototypical for mathematical impact in control engineering.

1.2.1 The stability criterion of Hurwitz

The stability criterion of Hurwitz was developed in 1893–1894 by the mathematician Adolf Hurwitz, an employed professor at the Polytechnikum in Zurich (today ETH Zurich) [30]. The cause for the development was not so much Hurwitz' mathematical curiosity, but rather a concrete request of his Zurichian colleague Aurel Stodola, a mechanical engineer, who had been engaged in the development of regulators for hydraulic turbines. Stodola's problem is best illustrated by means of an everyday problem. Consider the control of a heater in which the temperature of the room needs to be adjusted to a desired value. An obvious strategy is to regulate the valve for the hot water flow in the heater depending on the current measured temperature. If the current temperature is lower than the desired one, the valve is opened, otherwise the valve is closed. Obviously, one does not need any mathematics to see that this approach will eventually lead to the desired temperature.

However, this only holds in the ideal case. The case gets more complicated if the influence of the control input (in the example the flow valve) on the value (i.e., the temperature) is less direct, e.g., due to delays in the system. Any user of a shower whose hot water supply reacts only delayed on the opening or closing of the valves knows this effect: instead of the desired temperature one constantly gets values alternating between „too hot“ and „too cold“. Only with much effort and fine tuning one eventually succeeds in adjusting the right temperature. This is a classical example for an unstable behavior of a control circuit. The effects in the mechanical turbine systems that Stodola considered are quite similar, caused by a multitude of mechanical couplings. Also in this case it is not an easy task to design a control that keeps the system stable on a given reference value.

Stodola knew and applied mathematical results that had been discovered approximately two decades before [62]. They allow to deduce a condition for the stability from a model of the control circuit. This condition says basically that a polynomial, i.e., a mathematical function of the form

$$P(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

has only roots (i.e., complex numbers x with $P(x) = 0$) with negative real parts. If such a polynomial has only few summands, then this condition is easy to check. However, polynomials with few summands correspond to very simple control systems for which this theory is not necessary to ensure stability – an experienced engineer would have this knowledge anyway. Relevant for practice are polynomials with many terms. The main problem is that the roots cannot simply be computed due to algebraic reasons that are known since the work of Évariste Galois in the early 19th century.

Thus, at the suggestion of Stodola, Hurwitz looked for a while into this problem and finally found a criterion that allows to determine if the roots have negative real parts, without actually having to compute them. Mathematically, his solution consisted of combining the coefficients $a_0, a_1, \ldots, a_n$ of the polynomial in n different matrices of different sizes. Then it is sufficient to compute the determinants of these matrices in order to respond to the original question of the sign of the real parts of the eigenvalues.

Stodola was enthusiastic about this solution. In a letter to Hurwitz he wrote: „I am exceedingly indebted to you for the ingenious solution of the root problem that has bothered me so much.“ [5]. But also Hurwitz mentioned in a footnote to his corresponding publication in the „Mathematische Annalen“, not entirely without pride, the successful practical application of his criterion: „Mr. Stodola uses my result in his paper [...] whose results have been used in the turbines system in the bath Davos with splendorous success.“ Remarkably in this context is the fact that the English mathematician Edward John Routh developed a similar solution already 20 years before Hurwitz, therefore today Hurwitz' criterion is mostly referred to as the method of Routh-Hurwitz. However, Hurwitz' result has found the way into applications earlier. For decades Routh's result was known only to a small academic circle, but hardly to any engineer working in practice [2, pp. 81f.].

1.2.2 Pontryagin's maximum principle

In spite of its name, referring to the Russian mathematician Lew S. Pontryagin, Pontryagin's maximum principle has several authors. On the one hand, Pontryagin published and proved the principle released in 1956 not alone, but with essential contributions of the rarely mentioned mathematicians V. G. Boltyansky and R. L. Gamkrelidze [8] (see also [51]). On the other hand, cf. [49, pp. 55f.], the principle can be found already in publications of the American mathematician M.R. Hestenes from 1950 and – even in more general but not completely proven form – of the American mathematician R. P. Isaacs from 1954/55. Important fundamental ideas for the maximum principle can be found even earlier in Carathéodory's book from 1935, as stated in [50].

The maximum principle addresses the issue of how the motion of a system – such as the trajectory of a rocket or the movement of the arm of an industrial robot – can be optimally planned. Thereby, „optimal“ is always related to a pre-defined criterion with corresponding constraints. For example: how does a rocket get with minimum energy consumption (criterion) in a given time (constraint) from Earth to a particular orbit, or how can the arm of an industrial robot be steered as fast as possible (criterion) and with constant energy expense (constraint) from a given position to another one? Here, the example of the rocket is not chosen at random, since space travel as well as military rocket technology was one of the main driving forces for the development of this kind of mathematics in the beginning of the cold war between the USSR and the USA. The improvement in efficiency and stability of industrial robots, surely more significant for the industrial development, can therefore – like the famous Teflon-pan – be seen as a by-product of space research.

Optimization problems of similar form had already been solved about 250 years before Pontryagin and the other developers of the maximum principle with the help of the so-called calculus of variations. A comprehensive description of this method surely exceeds the scope of this article, but at least a brief outline of the principle's functionality shall be given here. For a detailed description see, e.g., [49]. Starting point for the approach is a model describing a motion via mathematical equations. Usually, this is done by means of a system of differential equations. For the solution of the problem the position on the optimal path has to be computed at every time instant. The problem is that the choices of values at different time points depend on one another. For example, higher acceleration in the beginning requires stronger braking at the end and vice versa. The calculus of variations solves these complex dependencies by setting up an extended system of differential equations (the Euler-Lagrange equations), whose first part is solved „normally“ forwards, while the second part is solved backwards in time – a so-called boundary value problem. Then, for each time point the optimal path can be determined from the solution of this system.

The main problem in the application of the calculus of variations to practical problems is that here the optimal trajectories are computed, rather than the actually relevant optimal control input which has to be transmitted to the rocket engines or to the motors of the robot arms. Since these input values do not occur in the computation, it is not possible to provide physically meaningful ranges in which these control values have to lie. In other words, the calculus of variations can indeed determine optimal trajectories, but it is impossible to exclude that these trajectories require, e.g., a thrust of the engines that is way beyond the physical possibilities. This is quite obviously a deficiency which causes serious practical problems not only in space travel but in virtually every imaginable industrial application.

Here, the maximum principle provides a remedy. Although conceptually quite similar, as also boundary value problems need to be solved, the focus is not on the computation of the optimal trajectory, but on the direct computation of the corresponding controls. This allows to directly include physical or economic restrictions on the control values in the computation. At the same time, the maximum principle yields a rather intuitive criterion, since the optimal control values in every time point are nothing else as the solutions of a new „small“ optimization problem, with parameters resulting from the boundary value problem. Since such problems can be solved numerically, the principle is not only useful for theoretical analysis, but also as a basis for powerful algorithms.

1.2.3 Conclusion

The aforementioned mathematical concepts for the analysis and solution of control problems are only two of many possible examples. Yet they have a number of properties that are more or less typical for successful mathematical concepts:

- They are applicable without knowledge of the underlying mathematical theories that were necessary for their derivation.
- They yielded considerable progress for real applications and thus enabled new industrial developments.
- They are constructive in the sense that they are easy to formulate as algorithms and therefore also easy to implement with the availability of digital computers.

Despite the evident advantages that these mathematical concepts had for industrial applications, their development was fostered by yet another factor. The control engineers were aware of the necessity of mathematical methods in the first place, and accordingly had a solid mathematical basic education. Without Stodola's knowledge of the interpretation of polynomial roots, Hurwitz would probably never have thought of developing his criterion. And without the availability of convenient models of differential equations the development of the maximum principle would have been impossible from the very beginning.

1.3 State-of-the-art and current developments: The example „model predictive control"

1.3.1 A short introduction to control engineering

The objective of control engineering is the manipulation of dynamical systems such that their behavior has one or several desired characteristics. Here the dynamic systems are influenced by the so-called *control input* or *control variables* u. In the dynamical system „car" these are, e.g., the turning angle of the steering wheel or the position of the gas pedal. Then, by suitable changing of the control variable over the time, the system dynamics can be influenced in the desired manner. The behavior of interest of the controlled system is usually summarized in the so-called *output variables* y. 3.1 shows schematically such a *controlled system*.

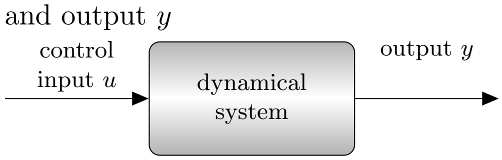
Two important control concepts are *open-loop* and *closed-loop control*. As shown in 3.2, in open-loop control the dynamic behavior is given in form of an open functional chain. The desired behavior of the output variable y is given by the *reference variable* w that provides the reference signal for the open-loop control. From this signal the control input is generated which influences the dynamical system such that the output follows the reference signal as accurately as possible.

In industrial practice open-loop controls are very important and in many cases they perform their task very well. Examples are robots, machine tools and production facilities. The basic assumption for the proper function of an open-loop control is that the real-world behavior of the dynamical system can be predicted sufficiently precise by means of the given model; if this is not the case, the actual behavior can significantly deviate from the theoretical forecast used for the computation of the control variable u.

This can happen if the dynamic system is affected by large external perturbations, as, e.g., strong side wind in automobiles. Often the behavior of the dynamic system is not precisely known, because the mathematical model represents not all aspects of the real system, hence *uncertainty* is an issue. The effects of these uncertainties are particularly dramatic when the system behavior is *unstable*. In the shower example in Section 1.2.1 instability would mean small changes in the control variable (valve) lead to large changes in the controlled variable (water temperature). In these cases *closed-loop control* has to be used in order to ensure stability of the system, i.e. in our example to ensure a temperature that stays close to the reference value.

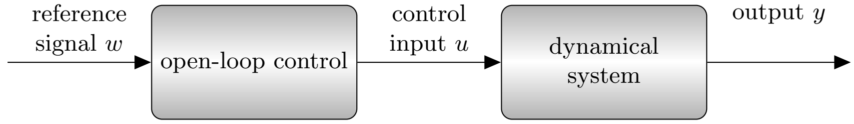
In contrast to open-loop control, a closed-loop control makes use of a feedback structure as shown in 3.3. The controlled variable is measured, fed back and compared to the desired value, i.e., the reference signal. The difference between reference and current value, also called the *control error*, is provided to the controller. Due to this closed-loop structure, the influence of external perturbations and uncertainties can be explicitly detected and adjusted before they lead to larger deviations from the reference value. The-

Abbildung 3.1: Schematic diagram of a control system with control input u and output y



refore, closed-loop controls are also applicable in the case of perturbations, uncertainties and unstable control systems.

Abbildung 3.2: The principle of open-loop control



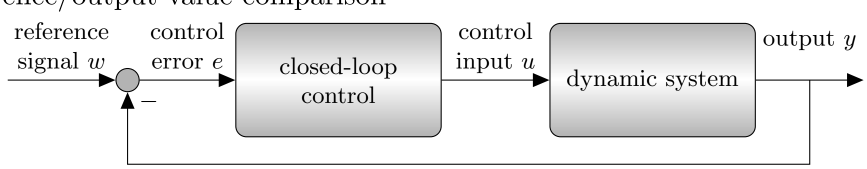
Certainly, the feedback structure alone does not automatically lead to stability of the closed-loop system: It is of crucial importance, by which rule the control input u is computed from the control error e. As already exemplified in Section 1.2, the dynamic behavior of a controlled system can be analyzed with the mathematical methods of control theory. The most important analysis methods examine stability and the so-called *robustness of stability* of the controlled system. The analysis of the robustness of stability addresses the issue if stability is preserved when the current dynamic system differs from the assumptions made in the model. This property plays a central role for the use of control methods in practical applications.

Mathematics plays a vital role in the *design* of open-loop and closed-loop controls, i.e., in the derivation of the rule that tells how the control input is computed from the reference variable or the control error. Today, modern controllers are designed *model-based*, i.e., it is assumed that a mathematical model of the system to be controlled is given, in most cases in the form of differential equations. In addition, the control objectives are formulated in mathematical terms, too. Besides the aforementioned stability these objectives often include the so-called *control performance*, typically formulated in form of an optimality criterion. Different classes of controller design methods differ in both the assumptions for the models of the control circuit and the control objectives. Regarding the model assumptions the most important distinction is between linear and nonlinear systems. Linear systems are characterized by – slightly simplified – a proportional relation between the control input u and the output y. This means that if, for instance, the control input u is doubled then the output y will be twice as large, too (mathematically one says that the system fulfills the superposition principle). Linear systems are described by linear differential equations, usually in the so-called state space form

$$\dot{x} = Ax + Bu \tag{3.1a}$$

$$y = Cx + Du. \tag{3.1b}$$

Abbildung 3.3: Principle of a closed-loop control with feedback and reference/output value comparison



The design of controllers for linear systems is well-understood and there is a large number of extremely powerful linear controller design methods available. At this point we want to mention exemplary the LQ methods (LQR, LQG, LQG/LTR, etc.) [36, 42, 23, 47], in which for linear systems of the form (3.1) a controller is designed, such that the closed-loop system has optimal behavior with regard to the minimization of a quadratic integral criterion

$$J = \int_0^\infty x^T(\tau)Qx(\tau) + u^T(\tau)Ru(\tau)\,\mathrm{d}\tau.$$

However, no external perturbations such as fluctuating side wind are considered in LQ methods. In the more recent H_∞-controller design [24, 64] controllers are designed such that external disturbances are optimally rejected, i.e., their negative effects on the control system is minimized. This minimization can be carried out by different mathematical optimization criteria, which apart from the H_∞-methods also lead to, e.g., L_1-methods [14].

Even if hardly any system occurring in industrial practice is actually linear, linear models can nevertheless often be used. The reason for this is that most systems „almost" behave like linear systems when only a small domain for the values of the controlled variable is considered: If the temperature in a room is to be increased from 20° to 22° Celsius, then the valve of the heater has to be opened approximately twice as far as if the temperature is only to be increased to 21° – the relationship between valve opening and temperature increase is linear. For larger temperatures this does not hold anymore, simply because further heating is not possible once the maximal temperature is reached, regardless of how the valve is set.

If the controlled system in the relevant domain does not exhibit linear behavior, which is frequently the case in practice, then nonlinear controller design methods have to be used. Here, the dynamic system forming the basis of the design is described by a nonlinear differential equation of the form

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x, u). \end{aligned}$$

In the last 20 years huge progress has been made in the field of nonlinear control and many current research projects deal with this subject. As an example, we mention differential geometric methods of exact linearization [33], flatness-based control [22, 58], passivity-based control [9, 55], and backstepping methods [37, 41]. In the simplest case[1] these methods yield a static

[1]Often the controller is not given by a static equation, but is itself given by a differential equation.

function

$$u = k(x), \tag{3.3}$$

i.e., a mathematical formula $k(x)$ that depends on the current system state x and determines how the control variable u has to be chosen.

In the nonlinear case, the primary focus often lies on the stabilization problem. However, certainly the control performance is of interest, too. The nonlinear equivalent to the previously mentioned LQ methods is optimal control, in which again an optimization criterion and constraints are provided – as already described in Section 1.2.2. However, in the case of the nonlinear optimal control the computation of an explicit formula (3.3) in feedback form for u is often impossible even with high-performance computers. On the other hand, the computation of nonlinear optimal open-loop controls (i.e., a control input u depending on time) with computer based, e.g., on Pontryagin's maximum principle, cf. Section 1.2.2, or the more modern direct methods, cf. Section 1.3.5, is relatively easy.

Therefore, a new class of methods for the optimal control of nonlinear systems has been developed in the last years, in which no explicit formula of the form (3.3) is computed, but the feedback law $k(x)$ is computed online in real time from open loop optimal control signals. Although such an online-computation looks more complicated at first sight, this approach has considerable advantages and is appropriate for the solution of large scale nonlinear control problems in practice. The most prominent representative of this new class of control methods is model predictive control that will be considered and discussed in more detail in the following sections.

1.3.2 The principle of model predictive control

If a controller is designed, e.g., in order to keep the temperature in a house on a constant value, then the easiest idea is to measure the current temperature and subsequently to increase or decrease the hot water flow rate appropriately – as described in Section 1.2.1 –, such that by well-chosen control parameters a stable controlled system can be reached that more or less compensates for the external variations in temperature caused by the weather. As the heater cannot supply heat arbitrary fast and also the cooling needs some time, the room temperature controlled this way will typically always slightly oscillate.

But what would happen if we used the weather forecast? Then we could slightly pre-heat the apartment if a cold front approaches, or, if warm weather is expected, the heater could be shut off already some hours in advance to have the house cool down. It is intuitively clear that the use of this „predictive“ principle allows to better maintain the control objective „as small deviation from the reference value as possible“, taking into account our limited heating or cooling capabilities.

When we are driving a car, it is utterly indispensable that we drive „anticipatory“, i.e., „predictive“. If we drove into a bend without braking in time, we would skid off the road. That this does not happen is due to the advanced control technique in our brain that has already anticipated the drive through the bend a long time before we learn about the centrifugal force in the bend the hard way, see 3.4.

Model predictive control (MPC) realizes exactly this principle for the automatic control of technical systems, cf., e.g., [11, 46]. We explain this in more detail using the example of a typical stabilization problem: the system state is to be controlled to a reference value and kept there by repeatedly measuring the state and, if necessary, adjusting detected deviations by suitable adaptation of the control input.

Based on the current state x_0 of the system to be controlled and on the current predictions of the external influences, a model predictive controller computes an (here for simplicity piecewise constant) optimal control for the near future, the „prediction horizon“ of length T. Typically, the employed optimality criterion minimizes the distance to the reference value, such that the computed optimal solution has a distance from the reference value as small as possible. Then, only the first part u_0 of the control is applied to the plant and used in order to steer the system for a short time – the sampling time δ. After this short time the new current state is determined, the prediction horizon is shifted forward by δ, and a new optimal control on this horizon is computed from which again only the first short part is used at the plant. In this way, the online computed optimal open-loop controls are „composed“ to a closed-loop control. Figure 3.5 illustrates one step in this procedure.

The two essential advantages of model predictive control methods are the ability to make use of look-ahead information and to include optimality criteria and restrictions. These advantages become apparent, e.g., in energy-intensive or time-critical applications, since goals as „with minimal energy expense“ or „within a given time“ can be directly included in the optimization criterion or in the constraint.

The two most important questions are:

- How can stability of the MPC controlled system be guaranteed?
- How can the optimal control problems be solved fast and reliably on the prediction horizon?

Both questions are of fundamentally mathematical nature. The first one belongs to mathematical systems theory, the second one to numerical mathematics. We want to discuss these questions – along with a presentation of applications of MPC – in the following sections.

1.3.3 Stability of MPC

As we want to avoid mathematical technicalities, here „stability“ will simply denote the fact that the control algorithm fulfills its goal, i.e., that the system state is steered to the reference value. Intuitively, it seems plausible that by minimizing the distance to the reference value a stable control strategy is obtained. However, this may not be the case, a we illustrate with a small example, cf. 3.6. In this figure the position of a car (system state, in the beginning at location B) shall be steered to location A (reference value), while maintaining the velocity restrictions on the road. As there are mountains in the way, the road does not lead directly from A to B, but takes a detour via location C.

If you consider the air-line distance of the car from location A as optimization criterion, then this distance has to be increased along the road from B to A before it can eventually be decreased. An MPC strategy with a prediction horizon that covers, e.g., only the part of the road up to point C in the bend will not notice this, since every motion of the car within this horizon would only increase the distance. Therefore it is optimal to simply stay in point B and consequently the MPC controlled solution will stay forever in point B and never reach point A: the control algorithm is not stable.

Of course, this problem is known for a long time and different possibilities to find a remedy have been proposed. One approach often considered in the literature is to allow only for such solutions in the optimization on the horizon T that end up in the desired reference value. This method works theoretically (what has rigorously been proven mathematically, cf. [38]) and is also occasionally employed in practice, but it has the disadvantage that a further restriction is added to the optimization problem that has to be solved and in addition in general it requires a very large horizon T to ensure that under the given restriction (i.e., for instance the speed limit in our road example) it is possible at all to reach the reference value at the time T. Either can make the solution of the optimization problem considerably more expensive. A further approach is the use of corrector terms in the optimization that compensate the effect of too short horizons T, see, e.g., [13]. The computation of these terms, however, is in general quite tedious, which is why this method is rarely applied in practice.

The approach most frequently used in practice is the most apparent: the length of the horizon T is simply increased until the algorithm becomes stable. Interestingly enough, this solution is *not* considered in the major part of theoretical MPC literature. Actually, for nonlinear MPC methods only recently it has been mathematically rigorously proven under relatively general assumptions that it indeed works [34, 26]. The disadvantage of this method is that in general it is not clear in advance how large T has to be chosen – and as T grows larger, more time is needed for the solution of the optimization problem that has to be carried out online during the runtime of the system. Current research approaches [27, 28] therefore try to estimate the required horizon T from the system properties and to use such estimates in order to determine how the optimality criterion has to be chosen in order to obtain stability with T as small as possible. At present, one can only speculate about the practicality of this approach for industrially relevant processes. For our road example, however, this approach yields a natural as well as efficient solution which perhaps already came to the mind of some of readers: if you measure the distance from location A to B not by air-line, but via the length of the road, then for arbitrary short horizons T it is always better to move towards A, since this will decrease the distance in any case. If you optimize over this distance, then the method is stable for arbitrary horizons T.

1.3.4 Application fields

Model predictive control has originally been developed in chemical engineering, where it is used, e.g., to control large and slow distillation columns – these systems are extremely energy-intensive and the slow time scale leaves much time for the computation of optimal controls. Chemical engineering is even today the main application area of MPC: Dittmar and Pfeiffer [20] have counted 9456 MPC applications in an investigation in 2005 (but estimate a considerably higher number in reality), of which more than 80 % lie in the field of chemical process control. Primarily, MPC is used in the control of continuous reactors in which the state of the process (temperature, concentration of chemicals ...) has to be kept close to a reference value for an arbitrary long time. Less often, MPC is used in so-called batch reactors in which the reaction does not occur continuously, but in a given time frame and in which the control problem consists of following a predetermined path of the process states.

Although there are examples for successful applications of nonlinear MPC (NMPC) in chemical engineering [52], here mostly linear MPC methods are applied, i.e., methods in which the underlying differential equation model is linear. This has the great advantage that the resulting optimization problems are linear, such that faster and more reliable solution algorithms are available. Although a comparative study on an industrial batch reactor [48] revealed that NMPC methods may well lead to a better control performance, so far the improvements are not as significant as to economically justify the time-consuming implementation of the method – it remains to be seen if this will change in the future with short running resources and more expensive energy. A further obstacle for the application of NMPC is often the lack of suitable nonlinear models in chemical engineering [52]. However, in this field recently considerable progress has been made with the aid of sophisticated mathematical methods, see Section 1.4.1.

Due to more efficient algorithms and theoretical advances, recently more and more application fields beyond chemical engineering have become accessible for MPC and in particular NMPC methods. Examples are the control of water gates in canal systems, the control of heating and climate systems in large buildings, or the optimal control of seasonal heat storage devices. The latter use the heat in the summer to heat up the ground under a building and at the same time to cool down the building, and reuse this heat in winter with the aid of a heat pump to save heating costs. Only predictive control allows to decide when the heat pump and additional heating can be used efficiently based on the seasonal weather forecast. Increasingly faster optimization algorithms allow even faster applications of model predictive control. Today, to-the-point control of train runs is possible and experimentally tested [25], as well as the control of the amount of fuel injection and air in automotive engine at sampling intervals of 50 milli-seconds [21]. Also, NMPC control of robot arms that need to carry out complex maneuvers as fast as possible and in constantly changing environments is within reach, see Figure 3.7. One of the most visionary possible future application of NMPC is the automatic control of flying kites that supply wind energy from high altitudes by flying on periodic paths under changing wind conditions [12, 32] in a stable way, see Figure 3.8. Another application of NMPC that is recently studied is the automatic control of insulin supply for diabetics.

1.3.5 Direct optimal control methods

In this section we want to give some insight into the current state-of-the-art of solution methods for the computationally most expensive step in model predictive control, the repeated solution of optimal control problems on finite time horizons. In what follows we concentrate on the more challenging case of nonlinear systems.

In Section 1.2.2, the maximum principle was addressed. So-called *indirect* methods based on this mathematical theorem are one possibility to compute optimal solutions for control problems. Even today, these methods are indeed applied. In particular, this holds for aerospace technology, see for example [10], but also for the analysis of problems in chemical engineering, see [59]. The maximum principle still plays an important role in the analysis of properties of optimal control problems. But, for the construction of algorithms that can be transferred to a computer these indirect methods have some disadvantages:

- the significant analytical effort even for small changes in the model or in the parameters,
- the complicated treatment of path and control restrictions which lead to state-dependent jumps in the differential variables,
- the boundary value problems with extremely small domains of convergence which are often very difficult to solve numerically, frequently only by the use of homotopies.

Because of these disadvantages the *direct* methods that are also known under the key word „first discretize, then optimize“ have established themselves for the computation of control strategies in practice, see also the discussion in [4].

A further important question is the evolution of the computed optimal control functions over time. In the maximum principle the optimal control is a continuous function, structural restrictions on its shape cannot be made. From a practical point of view this does not necessarily make sense. For instance, you may think of a temperature controller that due to technical reasons cannot be adjusted continuously but only in fixed intervals from one constant value to another. Such a controller cannot exactly follow any arbitrary control function, but only approximate it. Such an approximation is the main principle of direct methods in the first place.

In these methods the control functions are approximated by functions that are determined by a finite number of control variables. In the simplest case these are piecewise constant functions, i.e., they take a constant value q_i on a given time interval $[t_i, t_{i+1}]$. In 3.9 this is illustrated.

This transformation to a mathematical optimization problem that now only depends on a finite number of degrees of freedom (namely exactly the values q_i) allows the use of sophisticated methods of nonlinear optimization (nonlinear programming). Here, mainly two types of algorithms compete: interior point and active set-based methods. While in the latter, the set of active inequalities (denoted as the active set) is carried along and modified from iteration to iteration, the interior point methods move within the admissible domain towards the solution on the so-called central path with the aid of logarithmic barrier functions.

Both types of algorithms have certain assets and drawbacks which lead to the fact that they determine the optimal solution faster for certain classes of optimization problems. As a rule of thumb, interior point methods often perform better for problems with many inequality constraints (since they come from the inside and do not „notice“ the many intersections of the inequality constraints that for active set-based methods have to be „visited“ sequentially), while active set-based methods allow more efficient warm starts in the solution of related optimization problems. This is due to the fact that the sets of active inequalities often differ only marginally, e.g., when a slightly altered optimization problem results from newly measured data. Then, starting in the already computed solution of the original problem often leads to the result in a few iterations.

In the discretization of optimal control problems with direct methods commonly three types of algorithms are distinguished, see [4]: direct *single shooting*, direct *multiple shooting* and direct *collocation. Single shooting* was developed in the 1970's, see for example [29] and [54]. The basic idea is to consider the differential variables as dependent quantities of the independent control variables and to solve an initial value problem in each iteration of the optimization algorithm in order to compute the objective function and the constraints as well as the derivatives with respect to the control variables. In other words: we have an outer optimization algorithm in which our degrees of freedom are optimized, and an inner one, in which we „shoot“ for fixed solutions, i.e., we solve the differential equation by integration in order to compute the value of the states at all time points. 3.10 visualizes this basic idea.

In optimization problems often the state at the end of the time horizon is of particular interest. If you think of determining the controls to fly a plane, then it becomes clear how important the fine tuning of all control variables is, if a reference in a long distance is to be reached exactly. To distribute these difficulties along the time horizon, the concept of single shooting has been extended to the direct *multiple shooting*. This method has been published in 1981 for the first time [6]. Since then, the method has constantly been improved or newly implemented, see for example [43].

As in the case of single shooting, the controls are discretized on a time grid. But, in order to better coordinate the flight of the plane over a long distance, checkpoints in form of additional variables are incorporated. Now, the plane only has to be „flown“ to the next checkpoint, similar as in a human chain for extinguishing a fire. Since nonlinearities and possibly instabilities of the problem have smaller effects on a short time horizon than on a long one, it is now considerably easier to reach the focused target. This approach is a good example for the strategy to separate a hard problem into several small ones. This principle *divide et impera*, divide and conquer, can be found in many mathematical algorithms.

As new „checkpoints“, multiple shooting variables $s_i^x \in \mathbb{R}^{n_x}$ with $0 \leq i < n_{\text{ms}}$ are introduced for the states on a given time grid $t_0 \leq t_1 \leq \cdots \leq t_{n_{\text{ms}}} = t_f$ that serve as initial values in the decoupled integration on the time intervals $[t_i, t_{i+1}]$. As shown in Figure 3.11, such a composed trajectory is only continuous on the whole time horizon if the conditions

$$s_{i+1}^x = x(t_{i+1}; s_i^x, q_i, p) \tag{3.4}$$

hold. The left-hand side of equation (3.4) corresponds to the initial value on the time interval $[t_{i+1}, t_{i+2}]$, the right-hand side to the integrated value of the final state on the previous time interval $[t_i, t_{i+1}]$, depending on the initial value s_i^x and the control variables q_i and p. These equations are added to the optimization problem and have to be fulfilled by an optimal solution. For this reason, the direct multiple shooting method is often denoted as *all-at-once* approach, since simulation and optimization are carried out at the same time in contrast to the sequential approach of single shooting.

Graphically spoken, each checkpoint has to make sure that his flight reaches the respective successor so that the plane is eventually controlled from the start to the end point. The decisions that need to be made between every two checkpoints, our controls q_i, of course also have to be made in the single shooting. But now, at first everyone takes care prior about his own time period and about how the plane can be brought to the neighbor, and in the same time nevertheless behaves optimally. This „minding one's own business“ is one of the biggest strengths of this approach. Besides the gain in stability, it leads to unfolding of structures in the optimization problem that can algorithmically be used. The matrix that results from the analysis of the necessary optimality conditions has block form, whereby the blocks include all variables that belong to a time interval.

Because of the parameterizations of the state space, *direct collocation* and direct multiple shooting are quite similar. Collocation goes back to [61] and has been expanded amongst others in [7], [3] and [56]. Here, the basic idea is not to use any independent integrators for the solution of the differential equation, but to discretize all equations and incorporate them in the optimization problem. Collocation and multiple shooting share a number of advantages in comparison to single shooting. For instance

- it is possible to use previous knowledge about the process (that is given, e.g., in form of measured data) for the initialization of the state variables,
- it is guaranteed that the arising systems of differential equations are actually solvable also for nonlinear systems, while single shooting can run into a singularity for badly chosen controls,
- also unstable systems can be solved, if they are well-posed, since perturbations do not propagate over the complete time horizon, but are damped by the tolerances in the parameterization variables,
- a considerably better convergence behavior in higher dimensional space is observed in practice.

The resulting optimization problems are larger than in the case of single shooting. But, this is compensated by specific structure exploiting algorithms. In the case of collocation these are solution algorithms that use the particular band structure of the sparse matrices. For direct multiple shooting condensation algorithms are used to reduce the size of the quadratic programs that need to be solved, while for the approximation of the Hessian so-called high-rank updates are used. These accommodate the special block structure given by the shooting-intervals and thus accumulate more curvature information in each iteration than general quasi-Newton methods.

In summary, from our point of view direct multiple shooting and collocation are the methods of choice to solve nonlinear optimal control problems efficiently. Furthermore, these methods can be adapted efficiently to the special structure of NMPC applications: in particular, the fact that here iteratively a series of similar optimal control problems is solved has been utilized in recent years to develop variants of these methods that are particularly adapted to NMPC [44, 18, 57] and for which, also with inaccurate numerical solutions, the stability of the controlled system can be verified [17, 19, 63].

1.4 Challenges

We hope that so far we have convinced the reader of the significant role of mathematics as a key factor for technical control processes and that we have given some – of course subjective – insight into the current state-of-the-art and the variety of applications. In this chapter, we want to address some of the tasks and challenges that lie ahead. For this purpose in Section 1.4.1 we consider aspects that arise in the modeling of the considered processes and see to what extent mathematics can contribute significantly also at this level. In Section 1.4.2, we discuss the robustness of MPC controls and in the concluding Section 1.4.3 we highlight some further current topics.

1.4.1 Modeling

The methods for the model-based control of processes described in the preceding sections have an essential assumption: *for the considered process an adequate mathematical description in form of a system of differential equations can be found.*

The keywords in this sentence suggest where the difficulties lie.

- „for the considered process“ indicates that it is a matter of problem-specific modeling. From one plant to another several components may slightly differ – but for this reason the new process may require a completely different mathematical model.
- „adequate“ implicates at least three facts. First, that the model is correct, i.e., the essential properties that are of interest in the particular context are captured. Second, that the model only includes what is absolutely necessary in order to avoid excessive computing times. And

third, that the resulting model is suited for numerical simulation and optimization.

- „mathematical description“ refers to the underlying principles as well as to the estimation of the specific problem parameters, both based on the comparison of simulated with measured data.

- the phrase „can be found“ suggests a certain randomness. The question arises, how such „finding processes“ can be organized systematically.

Mathematical modeling addresses the problem of finding a mathematical model for a real process. This is based on knowledge from the respective field of application. It is known at least since Newton that the acceleration is the second derivative of the (time dependent) position, which can therefore in turn be computed from the fixed initial position and velocity and given acceleration. Another example are conservation laws which often form the basis of physical or chemical models. All such models are of course only approximations of reality and simplify matters here and there. From the consideration of friction forces to the interaction on atomic level such a model can be refined arbitrarily. An important task of the modeler is to find the right compromise between simplicity of the mathematical model and a sufficiently close description of reality.

Unfortunately, for many problems in practice the considered processes are often still too complex, or not sufficiently understood. This is particularly true for systems biology, which has started to systematically make use of scientific computing methods only very recently. A different situation prevails in robotics, mechanics, or chemical engineering, where already for some decades methods of modeling, simulation and also optimization are applied. Here, the basic principles are often so-called *first principles*, i.e., basic and well-understood conservation laws of natural sciences. Nevertheless, a successful modeling – in particular in the field of nonlinear models – can only be achieved by very labor-intensive efforts based on all available expertise.

Therefore, for many relevant technical processes the assumptions mentioned in the beginning of this section are in fact a challenge for the future. Yet, there are encouraging results towards an algorithmic modeling that we want to highlight in the following – not without emphasizing that also this shifting away from the tedious trial-and-error based expert modeling towards standardized and verifiable methods would not be possible without underlying mathematical methods.

Parameter estimation Even though the basic principles are known, often certain model parameters are unknown. One may think for example of the masses of components in multi-body systems, of activation energies, or of reaction velocities. For the determination of these parameters measurements are required, in which frequently the parameters cannot be directly measured; instead only states of the system, or functions of these states are available. If you want to know certain physical properties of your heating, you will most probably need to measure the temperature in the room under different conditions and determine the intrinsic heating parameters from this experimental data and your mathematical model. The parameter estimation aims to minimize the distance between a model response and collected measured data by *fitting* of the unknown parameters. The method of least squares has been developed by Carl-Friedrich Gauss and Adrien-Marie Legendre already in the beginning of the 19th century. However, it experienced a significant advancement in the last decades, in particular concerning efficient numerics, restrictions on the parameter estimation problem, and robust parameter estimation by use of objective functionals that allow for outliers, see e.g., [40].

Optimum experimental design The parameter estimation yields parameter values for which a simulation of the system of differential equations has a minimal distance to measured values. In addition, a statistical analysis allows for the computation of the covariance matrix, which is a measure for the accuracy with which the parameters have been determined. In simplified terms, this is analogous to the error bars which are frequently plotted in diagrams for measured data.

As a matter of fact, different experimental setups lead to different accuracies. The goal of optimum experimental design is to determine experimental setups that lead to confidence regions (error bars) as small as possible. Experimental setup means everything that can be influenced in carrying out an experiment. This includes control inputs as well as the decision at which time instances (or at which positions in spatially distributed processes, respectively) measurements are to be made.

While optimum experimental design has been examined in statistics for many years, there was hardly any transfer to technical applications. This is due to the fact that optimum experimental design has been examined for a long time as a theoretical construct, and concrete computations have been restricted to academic example problems – a comparable situation to the one of Hurwitz mentioned earlier. Only in recent years researchers started working on efficient numerical implementations and on a generalization to the nonlinear, dynamic case; pushed, of course, by a concrete demand from engineers in academia and industry. The achieved breakthroughs in the advancement of statistical analysis now allow the application even within industrial practice, where nonlinear dynamic systems with up to hundred differential states, and more and more also spatially distributed processes are considered.

In particular, for nonlinear processes the computed optimal experimental designs are often anything but intuitive. Yet they can yield significant improvements and insight. For example, methods developed at the Interdisciplinary Center for Scientific Computing in Heidelberg are by now regularly applied at the industrial partner BASF in Ludwigshafen: empirical data from more than two dozen industrial projects has shown that compared to traditional experimental setups designed by experts the number of experiments could be reduced by more that 80 % while maintaining or even improving statistical quality [1]. In [39] an optimal experimental design is described with only 2 experiments that determines the parameters that have to be estimated accurately to 1 %, while a test plan with 15 experiments designed by an expert for comparison determines a parameter merely up to plus/minus 30 %. In the highly competitive field of chemical engineering this paves the way for potential savings in terms of costs, environment pollution, and time-to-market.

Further important research fields are *model discrimination* which focuses on the automatic design of experiment designs in order to set a boundary as significant as possible between competing models, as well as *model reduction* that deals with the simplification of an existing model without losing the relevant properties. One approach for this is the spectral analysis of the modes that separates the slow components from the fast ones, which often play only a subordinate role for the control.

1.4.2 Robustness of solutions, uncertainties

While (N)MPC is a concept for the control of processes that implicitly allows for uncertainties, in particular in the methods for offline-optimization introduced in Section 1.3.5, it is assumed that the process to be optimized behaves deterministically according to the mathematical model. In practice, this will not be the case as different kinds of uncertainties occur:

- *Model uncertainties.* Despite the efforts to obtain an adequate model for a considered process presented in Section 1.4.1, there will always be model errors that lead to (hopefully only marginal) deviations of the simulation results from the real process.

- *External perturbations.* Random perturbations (noise) cause influences on the system that are unknown at the time of optimization. Examples are the side wind for automobiles, rainfall fluctuations in sewerage plants, consumer behavior in economy, or external temperature variations.

- *Different scenarios.* This kind of uncertainty arises due to the fact that certain parameters of the examined process are open a priori. Turbines blades or aircrafts, for example, should have good aerodynamic properties for a whole bandwidth of upstream flow angles. Frequently, one further distinguishes between a discrete number of scenarios, a probability distribution in a continuous domain, and a worst-case scenario.

All three types of uncertainties pose challenges both to the the user as well as to the mathematician.

Optimal solutions often have a property that limits their applicability in practice: they are very sensitive against uncertainties. This is due to the fact that the constraints and nonlinear effects are exploited as much as possible. Typically, there are one or more restrictions that are active in an optimal solution, i.e., they will be violated whenever a little something changes. If you want to drive in optimal time from A to B, then at some places you will drive exactly at the maximally allowed velocity. If now the model is inaccurate, or external perturbations occur (think of head or tail wind), then these restrictions will possibly be violated. In case of a speed limit small violations may be acceptable, but this is not the case in so-called *runaway* processes that cannot be reversed once a certain threshold is crossed. Who wants to be reliable for achieving maximal energy efficiency in a nuclear power plant at the expense of constantly being on the brink of an irreversible chain reaction?

From a practical point of view a solution that is *almost* as good as the optimal, but less sensitive against perturbations would be preferred. Approaches in this direction work with (a) feedback, in the same way as MPC does by including measured data in a realtime-context, or (b) with safety margins that are added to the restrictions, or (c) with weighted sums, multi-criteria optimization or game-theoretical approaches (here, there are first approaches in the context of NMPC [45] that are however algorithmically considerably more complex as the „normal“ NMPC), or as well (d) by the integration of higher derivatives in the optimization routine. The sensitivity of the solution against uncertain parameters can be formally computed as a derivative. Modern approaches in robust optimization therefore often extend the mathematical model by terms with higher derivatives of the objective function or the restrictions to these parameters, respectively. Despite encouraging advances in this field, uncertainties and the determination of robust solutions will remain a big challenge further on.

1.4.3 Further challenges

While theory and algorithmics of optimal control have long since been developed for finite dimensional dynamic systems, the door for the treatment of *spatial effects* has opened due to algorithmic advances and the availability of faster hardware obeying Moore's Law. For instance, this means that the temperature in a room should not be controlled by only using a representative sample value at a specific position, but rather by taking into account the temperature at each position in space.

While this will be hardly ever necessary for simple heaters in a room, there are many processes where the spatial component is important. In the production of steel beams, for example, you need to make sure the temperature variations within the beam do not become too large in the cooling phase to avoid subsequent cracks, [60]. In aerodynamics the fluid behavior plays a decisive role. The spatial positions of air turbulence and shock fronts directly influences the flight properties, [35].

The optimization of processes described by partial differential equations has gained more and more attention in recent years, for example due to a Priority Research Program of the Deutsche Forschungsgemeinschaft [15]. Nevertheless, it appears that the gap between mathematical fundamental research and practical application is still rather large at this point, not only in the field of algorithms for the solution of optimal control problems, but also in the system-theoretical fundamentals of MPC methodology.

A challenge of quite different nature is the combination of continuous and discrete events. Here, continuous means a connected domain from which a variable or a control can take a value. You can think for example of the velocity of a vehicle or of the position of the gas pedal. Discrete means a non-connected domain that is mathematically composed of a finite set of possible values. Here, the gearshift is a good example for a discrete control, while „the walking robot has contact with the ground or not“ is a state in which either one applies. In the same way, the digital couplings in linked systems that are composed of a variety of smaller physical subsystems lead to discrete state components. Systems that combine both types of states and/or controls are denoted as *hybrid systems*. Frequently, hybrid systems arise due to a multi-scale modeling. Here, fast transient transitions are assumed to be instantaneous. A good example are valves that are assumed to be either open or closed in the model, while in reality they have to be transferred from one state to the other.

Mathematically, the integral or discrete variables that arise in the modeling pose a substantial challenge for the optimization. This may surprise on first sight, as only a restricted number of possibilities are available that theoretically can all be enumerated. However, the number of possibilities grows rapidly if the number of variables is increased. This is the case for hybrid systems, in particular to decide when and how often should be switched. A treatment of these variables by trial-and-error is impractical because of the immense number of possibilities. An overview of possible methods for discrete control functions and further literature on hybrid systems can be found for example in [53].

Also this research field turns out to be extremely active, and is for example supported by the European research network HYCON [31] and – for the special case of digital linked systems – by a DFG Priority Research Program [16]. In particular, the development of reliable and efficient algorithms and a stability analysis in the sense of Section 1.3.3 will require much attention in the next years to be able to exploit the full optimization potential in hybrid systems.

1.5 Visions and recommendations

In our opinion, MPC and in particular NMPC belong to the most promising methods for the solution of complex nonlinear control problems, which is why they will gain further importance in the future. Their ability not „only“ to solve control problems but also to include technical and economical restrictions via optimality criteria and constraints yields – in particular in times of short running resources – distinct advantages over other methods. Moreover, the rapid advance in the field of optimization algorithms allows continuously new applications with increasingly faster dynamics.

Just like the success factors of the classical methods of Hurwitz and Pontryagin, MPC has the distinct advantage to be intuitively easy to understand and therefore to be applicable without deeper understanding of the underlying mathematics – be it the systems theoretical analysis of the control behavior or the basics of the used algorithms stemming from numerics and optimization.

Therefore, we want to restrict the following comments to this control method, although surely analogous statements are also valid for other modern methods in control engineering.

1.5.1 Theory and praxis

From a theoretical point of view, the biggest conceptual advantage of (N)MPC – namely the explicit inclusion of a mathematical model for the prediction of the system behavior – is at the same time one of the essential disadvantages in practice. MPC inevitably requires a sufficiently exact mathematical model for the complex process to be controlled, which is not always available. As described in Section 1.4.1, recently promising mathematical modeling methods have become available also for the nonlinear case. Nevertheless, the design of suitable models will most likely remain a major challenge. The necessary investments may only be made by industry if the achievable advantages for the optimization of production processes are supported by concrete examples. Therefore, many of the emerging applications sketched in Section 1.3.4 play a pioneering role and their success or failure (and naturally also the communication of corresponding achievements) will decide over the acceptance of NMPC in industrial application in the long run.

But also on the theoretical side there are a number of open questions. In particular, the assumptions in the major part of theoretical MPC literature differ distinctively from the methods used in practice, as presented in Section 1.3.3. Here, it seems desirable to close this gap between theory and praxis. Especially, theoretical work that aims at concrete improvements of MPC methods – such as improved stability, faster algorithms, better control quality, higher robustness – should try to achieve their results and improvements under realistic assumptions.

Finally, on the algorithmic side further advance is expected, mainly by intensified cooperations of system and control theorists on the one and optimizers and numerical analysists on the other hand. As an example, only a systems theoretical stability analysis of the overall control algorithm can give information in what sense the integration of higher derivatives in the optimization routine explained in Section 1.4.2 really increases the robustness – and how the optimality criterion should look like. On the other hand, e.g., the systems theoretical analysis of MPC algorithms should always consider realistic properties of the optimization routines (including the manifold „tricks“ used in practice) in order to be able to make realistic and reliable statements about the efficiency of control. Thus, in both examples the competence from both fields is necessary.

1.5.2 Interdisciplinary in education

From our point of view, all these goals are only to be achieved if – just as in the cooperation of Hurwitz and Stodola – already in the education the co-operation between mathematicians and engineers and between the different fields of applied mathematics is emphasized. Only this way mathematicians can discover the methodical challenges of practical problems, only in this way engineers can profit from the newest mathematical methods, and only in this way methods and techniques from different mathematical areas can complement each other in a meaningful way.

Fortunately, there are a quite a number of very promising current interdisciplinary teaching initiatives, e.g.,

- the International Doctorate Program „Identification, Optimization and Control with Applications in Modern Technologies“ of the Universities Bayreuth, Erlangen-Nürnberg and Würzburg within the Elite Network Bayern
- the Graduate School „Mathematical and Computational Methods for the Sciences“ of the University Heidelberg
- the Center of Excellence „Optimization in Engineering“ (OPTEC) at the University Leuven
- the Cluster of Excellence „Simulation Technology“ (SimTech) of the University Stuttgart with postgraduate school as well as bachelor and master degree programs

to mention only those activities in which the authors of this article are involved. The formal framework for further advances is currently quite good, especially in the field of control engineering. It remains to hope that these efforts are continued such that many of the developed mathematical methods find their way into industrial applications and that the opportunities of the production factor mathematics are also taken in the future.

Abbildung 3.4: Anticipatory driving – as it is supposed to be . . .

Abbildung 3.5: Schematic diagram of model predictive control

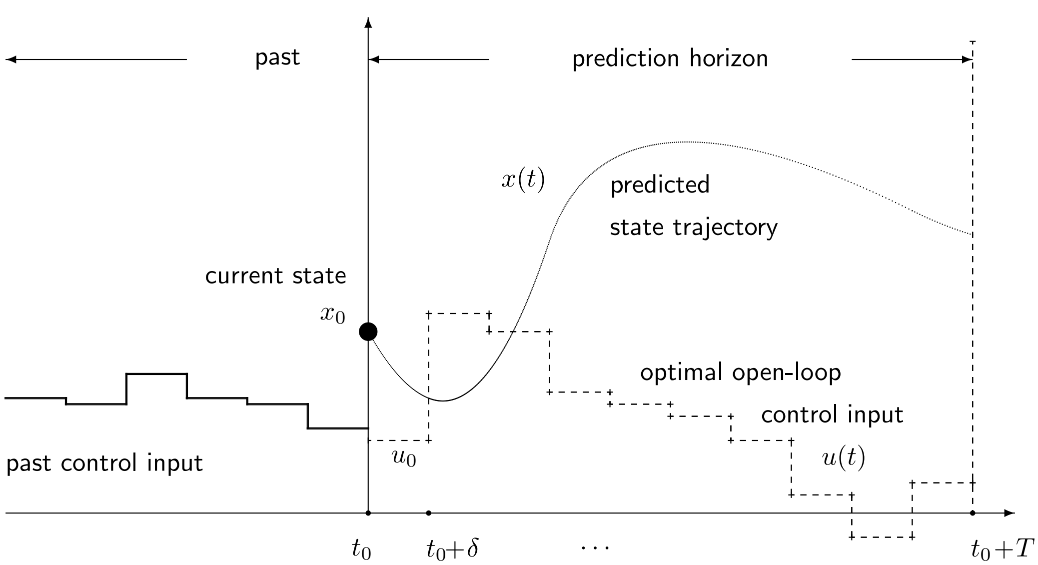


Abbildung 3.6: Example of a control problem

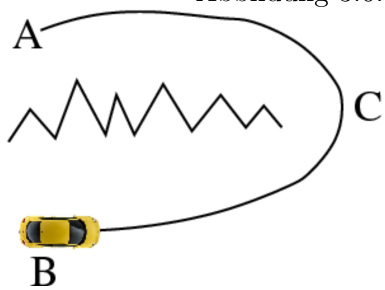


Abbildung 3.7: Robots at the university of Leuven currently running experiments with MPC in milli-second intervals.

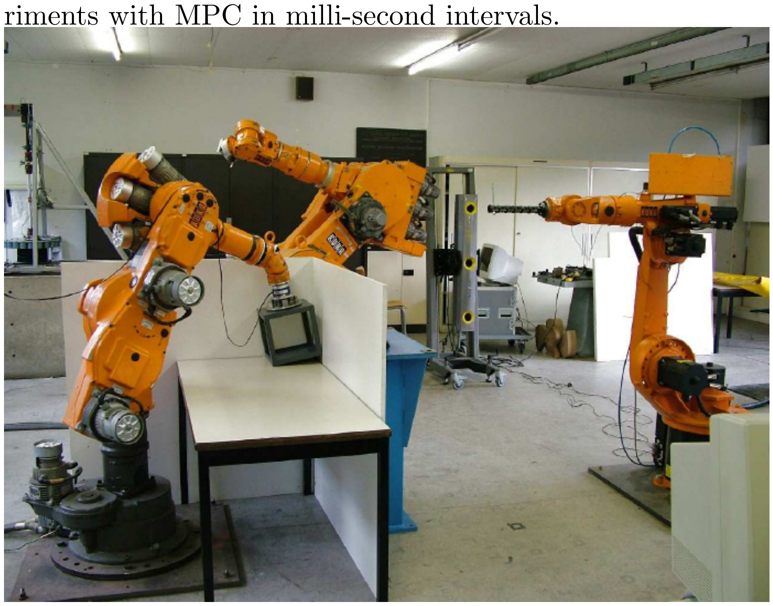


Abbildung 3.8: Automatically controlled kite for energy production [12]

Abbildung 3.9: Top left: control computed with indirect method. The other diagrams show approximations with the direct method on different discretization grids.

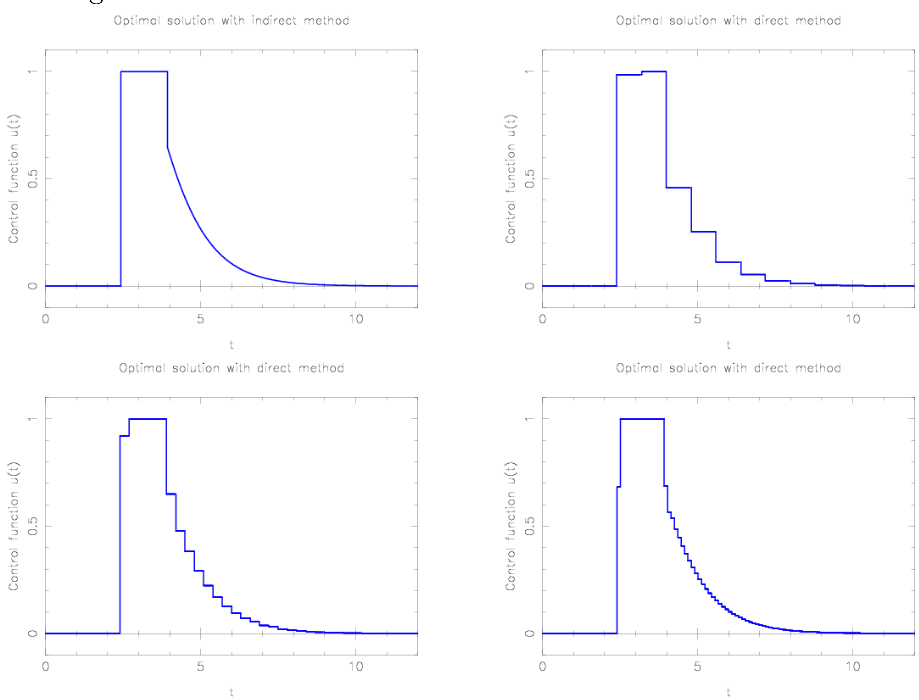


Abbildung 3.10: Schematic illustration of the direct single shooting method. The controls are given by piecewise constant functions q_i, the corresponding states are determined by integration. The lengths of the intervals do not necessarily have to be equidistant, but might also be longer on the last intervals, as indicated.

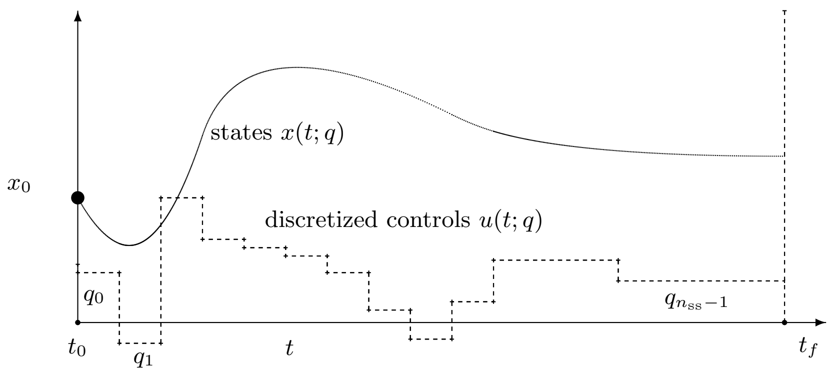
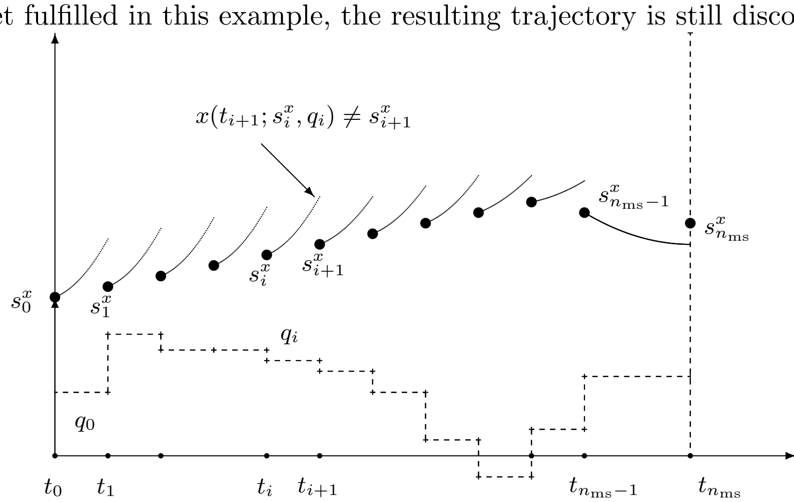


Abbildung 3.11: Schematic illustration of direct multiple shooting. The controls are given by piecewise constant functions q_i, the corresponding states are determined by piecewise integration. The connection conditions are not yet fulfilled in this example, the resulting trajectory is still discontinuous.



Literaturverzeichnis

[1] BASF SE. BASF und Universität Heidelberg entwickeln gemeinsam neue Mathematik-Software für die Forschung. Press release P-08-308, June, 16th 2008.

[2] S. Bennett. *A history of control engineering 1800–1930*. Peter Peregrinus Ltd., London, 1979. Paperback reprint 1986.

[3] L.T. Biegler. Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Computers and Chemical Engineering*, 8:243–248, 1984.

[4] T. Binder, L. Blank, H.G. Bock, R. Bulirsch, W. Dahmen, M. Diehl, T. Kronseder, W. Marquardt, J.P. Schlöder, and O.v. Stryk. Introduction to model based optimization of chemical processes on moving horizons. In M. Grötschel, S.O. Krumke, and J. Rambau, editors, *Online Optimization of Large Scale Systems: State of the Art*, pages 295–340. Springer, 2001.

[5] C. C. Bissell. Stodola, Hurwitz and the genesis of the stability criterion. *Int. J. Control*, 50:2313–2332, 1989.

[6] H.G. Bock. Numerical treatment of inverse problems in chemical reaction kinetics. In K.H. Ebert, P. Deuflhard, and W. Jäger, editors, *Modelling of Chemical Reaction Systems*, volume 18 of *Springer Series in Chemical Physics*, pages 102–125. Springer, Heidelberg, 1981.

[7] H.G. Bock. Recent advances in parameter identification techniques for ODE. In P. Deuflhard and E. Hairer, editors, *Numerical Treatment of Inverse Problems in Differential and Integral Equations*, pages 95–121. Birkhäuser, Boston, 1983.

[8] V. G. Boltyanski, R. V. Gamkrelidze, and L. S. Pontryagin. On the theory of optimal processes (russian). *Dokl. Akad. Nauk SSSR*, 110:7–10, 1956.

[9] C.I. Byrnes, A. Isidori, and J.C. Willems. Passivity, feedback equivalence, and the global stabilization of minimum phase nonlinear systems. *IEEE Transactions on Automatic Control*, 36(11):1228–1240, 1991.

[10] J.-B. Caillau, J. Gergaud, T. Haberkorn, P. Martinon, and J. Noailles. Numerical optimal control and orbital transfers. In *Proceedings of the Workshop Optimal Control, Sonderforschungsbreich 255: Transatmosphärische Flugsysteme, Heronymus Munchen, ISBN 3-8979-316-X*, pages 39–49, Greifswald, Germany, 2002.

[11] E. F. Camacho and C. Bordons. *Model predictive control.* Springer–Verlag, London, 2nd edition, 2004.

[12] M. Canale, L. Fagiano, M. Ippolito, and M. Milanese. Control of tethered airfoils for a new class of wind energy generators. In *Proceedings of the 45th IEEE Conference on Decision and Control, San Diego, California*, pages 4020–4026, 2006.

[13] H. Chen and F. Allgöwer. A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability. *Automatica*, 34(10):1205–1217, 1998.

[14] M.A. Dahleh and I.J. Diaz-Bobillo. *Control of Uncertain Systems : A Linear Programming Approach.* Prentice Hall, Englewood Cliffs, N.J, 1995.

[15] DFG Schwerpunktprogramm 1253, Optimization with Partial Differential Equations. http://www.am.uni-erlangen.de/home/spp1253/.

[16] DFG Schwerpunktprogramm 1305, Regelungstheorie digital vernetzter dynamischer Systeme. http://spp-1305.atp.rub.de/.

[17] M. Diehl, H.G. Bock, and J.P. Schlöder. A real-time iteration scheme for nonlinear optimization in optimal feedback control. *SIAM Journal on Control and Optimization*, 43(5):1714–1736, 2005.

[18] M. Diehl, H.G. Bock, J.P. Schlöder, R. Findeisen, Z. Nagy, and F. Allgöwer. Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *J. Proc. Contr.*, 12(4):577–585, 2002.

[19] M. Diehl, R. Findeisen, F. Allgöwer, H.G. Bock, and J.P. Schlöder. Nominal stability of the real-time iteration scheme for nonlinear model predictive control. *IEE Proc.-Control Theory Appl.*, 152(3):296–308, 2005.

[20] R. Dittmar and B.-M. Pfeiffer. Modellbasierte prädiktive Regelung in der industriellen Praxis. *at – Automatisierungstechnik*, 54:590–601, 2006.

[21] H.J. Ferreau, P. Ortner, P. Langthaler, L. del Re, and M. Diehl. Predictive control of a real-world diesel engine using an extended online active set strategy. *Annual Reviews in Control*, 31(2):293–301, 2007.

[22] M. Fliess, J. Lévine, P. Martin, and P. Rouchon. Flatness and defect of non-linear systems: introductory theory and examples. *International Journal of Control*, 61(6):1327–1361, 1995.

[23] O. Föllinger. *Optimierung dynamischer Systeme: Eine Einführung für Ingenieure*. Oldenbourg, München, 2. edition, 1988.

[24] B. Francis, J. Helton, and G. Zames. H_∞-optimal feedback controllers for linear multivariable systems. *IEEE Transactions on Automatic Control*, 29(10):888–900, 1984.

[25] R. Franke, M. Meyer, and P. Terwiesch. Optimal control of the driving of trains. *at – Automatisierungstechnik*, 50(12):606–614, 2002.

[26] G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel. Model predictive control: for want of a local control Lyapunov function, all is not lost. *IEEE Transactions on Automatic Control*, 50(5):546–558, 2005.

[27] L. Grüne and A. Rantzer. On the infinite horizon performance of receding horizon controllers. *IEEE Transactions on Automatic Control*, 53(9):2100–2111, 2008.

[28] L. Grüne. Analysis and design of unconstrained nonlinear MPC schemes for finite and infinite dimensional systems. *SIAM Journal on Control and Optimization* 48(2):1206–1228, 2009.

[29] G.A. Hicks and W.H. Ray. Approximation methods for optimal control systems. *Can. J. Chem. Engng.*, 49:522–528, 1971.

[30] A. Hurwitz. Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt. *Math. Annalen*, 46:273–284, 1895. Nachgedruckt in: R. Jeltsch (ed.) et al., Stability theory. Birkhäuser, Basel, 1996, 239–249.

[31] HYCON, European Network of Excellence on Hybrid Control. http://www.ist-hycon.org/.

[32] A. Ilzhoefer, B. Houska, and M. Diehl. Nonlinear MPC of kites under varying wind conditions for a new class of large scale wind power generators. *International Journal of Robust and Nonlinear Control*, 17(17):1590–1599, 2007.

[33] A. Isidori. *Nonlinear control systems*, volume 1. Springer, Berlin, 3. edition, 2002.

[34] A. Jadbabaie and J. Hauser. On the stability of receding horizon control with a general terminal cost. *IEEE Transactions on Automatic Control*, 50(5):674–678, 2005.

[35] A. Jameson. Aerodynamics. In E. Stein, R. De Borst, and T.J.R. Hughes, editors, *Encyclopedia of Computational Mechanics*, volume 3, pages 325–406. Wiley, 2004.

[36] R. E. Kalman. When is a linear control system optimal? *Trans. ASME, Series D, Journal of Basic Engn.*, 86:51–60, 1964.

[37] I. Kanellakopoulos, P.V. Kokotovic, and A.S. Morse. Systematic design of adaptive controllers for feedback linearizable systems. *IEEE Transactions on Automatic Control*, 36(11):1241–1253, 1991.

[38] S. S. Keerthy and E. G. Gilbert. Optimal infinite horizon feedback laws for a general class of constrained discrete-time systems: stability and moving horizon approximations. *J. Optimiz. Theory Appl.*, 57:265–293, 1988.

[39] S. Körkel. *Numerische Methoden für Optimale Versuchsplanungsprobleme bei nichtlinearen DAE-Modellen.* PhD thesis, Universität Heidelberg, Heidelberg, 2002.

[40] S. Körkel, E. Kostina, H.G. Bock, and J.P. Schlöder. Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optimization Methods and Software*, 19:327–338, 2004.

[41] M. Krstic, I. Kanellakopoulos, and P.V. Kokotovic. *Nonlinear and Adaptive Control Design.* Wiley, New York, 1995.

[42] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems*. Wiley, New York, 1972.

[43] D.B. Leineweber, I. Bauer, A.A.S. Schäfer, H.G. Bock, and J.P. Schlöder. An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization (Parts I and II). *Computers and Chemical Engineering*, 27:157–174, 2003.

[44] W.C. Li, L.T. Biegler, C.G. Economou, and M. Morari. A constrained pseudo-Newton control strategy for nonlinear systems. *Computers and Chemical Engineering*, 14(4/5)(451-468), 1990.

[45] D. Limon, T. Alamo, F. Salas, and E. F. Camacho. Input to state stability of min-max MPC controllers for nonlinear systems with bounded uncertainties. *Automatica*, 42(5):797–803, 2006.

[46] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: stability and optimality. *Automatica*, 36:789–814, 2000.

[47] V. L. Mehrmann. *The autonomous linear quadratic control problem. Theory and numerical solution*, volume 163 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1991.

[48] Z. Nagy, B Mahn, R. Franke, and F Allgöwer. Evaluation study of an efficient output feedback nonlinear model predictive control for temperature tracking in an industrial batch reactor. *Control Engineering Practice*, 15:839–850, 2007.

[49] H. J. Pesch. *Schlüsseltechnologie Mathematik*. Teubner, Stuttgart, 2002.

[50] H.J. Pesch and R. Bulirsch. The maximum principle, Bellman's equation and Caratheodory's work. *Journal of Optimization Theory and Applications*, 80(2):203–229, 1994.

[51] L.S. Pontryagin, V.G. Boltyanski, R.V. Gamkrelidze, and E.F. Miscenko. *The Mathematical Theory of Optimal Processes*. Wiley, Chichester, 1962.

[52] S. J. Qin and T. A. Badgwell. An overview of nonlinear model predictive control applications. In J. C. Cantor, C. E. Garcia, and B. Carnahan, editors, *Nonlinear model predictive control*. Birkhäuser, Basel, 2000.

[53] S. Sager, G. Reinelt, and H.G. Bock. Direct methods with maximal lower bound for mixed-integer optimal control problems. *Mathematical Programming*, 118(1)(109-149), 2009.

[54] R.W.H. Sargent and G.R. Sullivan. The development of an efficient optimal control package. In J. Stoer, editor, *Proceedings of the 8th IFIP Conference on Optimization Techniques (1977), Part 2*, Heidelberg, 1978. Springer.

[55] A.J. van der Schaft. *L_2-gain and Passivity Techniques in Nonlinear Control*. Springer, London, 2. edition, 2000.

[56] V.H. Schulz. Solving discretized optimization problems by partially reduced SQP methods. *Computing and Visualization in Science*, 1:83–96, 1998.

[57] Y. Shimizu, T. Ohtsuka, and M. Diehl. A real-time algorithm for nonlinear receding horizon control using multiple shooting and continuation/Krylov method. *International Journal of Robust and Nonlinear Control*, 2009. In print.

[58] H. Sira-Ramírez and S.K. Agrawal. *Differentially Flat Systems*. Marcel Dekker, New York, 2004.

[59] B. Srinivasan, S. Palanki, and D. Bonvin. Dynamic Optimization of Batch Processes: I. Characterization of the nominal solution. *Computers and Chemical Engineering*, 27:1–26, 2003.

[60] F. Tröltzsch and A. Unger. Fast solution of optimal control problems in the selective cooling of steel. *Zeitschrift für Angewandte Mathematik und Mechanik*, pages 447–456, 2001.

[61] T.H. Tsang, D.M. Himmelblau, and T.F. Edgar. Optimal control via collocation and non-linear programming. *International Journal on Control*, 21:763–768, 1975.

[62] J. Wischnegradski. Sur la théorie générale des régulateurs. *Comptes Rendus de L'Académie des Sciences de Paris*, 83:318–321, 1876.

[63] V. M. Zavala and L.T. Biegler. The advanced step NMPC controller: Optimality, stability and robustness. *Automatica*, 2008. (accepted for publication).

[64] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, Upper Saddle River, NJ, 1996.