6. Optimalsteuerungen

Bei der Frage der Steuerbarkeit eines Systems ging es um die Frage, ob sich ein gesteuertes System $\dot{x}(t) = f\big(t, x(t), u(t)\big)$ von einem Anfangszustand $x(t_0) = x_0$ aus durch geeignete Wahl der Steuerung $t \mapsto u(t)$ in einen beliebigen Zielzustand $x(t_1) = x_1$ überführen läßt. Wir bezeichnen mit x_u die zur Steuerung u gehörige Lösung des Anfangswertproblems

$$\dot{x}(t) \;=\; f\big(t, x(t), u(t)\big), \quad x(t_0) \;=\; x_0.$$

Typischerweise gibt es viele verschiedene Steuerungen, mit denen sich der Zielzustand ansteuern läßt, und wir können fragen, welche dieser Steuerungen in einem gewissen Sinn "optimal" ist. Wir könnten etwa daran interessiert sein, den Zielzustand in möglichst kurzer Zeit oder mit möglichst kleinem Energieverbrauch oder zu möglichst geringen Kosten zu erreichen. Die Kosten, die die Steuerung u verursacht, lassen sich typischerweise in der Form

$$J[u] \;=\; \int_{t_0}^{t_1} L\big(t, x_u(t), u(t)\big)\,\mathrm{d}t \;+\; K\big(t_1, x_u(t_1)\big)$$

schreiben, setzen sich also zusammen aus laufenden Kosten, die während des zeitlichen Entwicklung des Systems akkumuliert werden, sowie Endkosten, die beim Erreichen des Ziels fällig werden. (Bei einem Flugzeug sind etwa die Kosten für den Treibstoffverbrauch laufende Kosten, anfallende Landegebühren dagegen Endkosten.) Im allgemeinsten Fall wird ein Optimalsteuerungsproblem durch folgende Daten spezifiziert:

- eine Systemgleichung $\dot{x}(t) = f\big(t, x(t), u(t)\big)$;
- eine Anfangsbedingung $x(t_0) = x_0$;
- für den Endzustand $x(t_1) = x_1$ eine Zielbedingung $(t_1, x_1) \in S$ mit $S \subseteq [t_0, \infty) \times \mathbb{R}^n$;
- eine Klasse $\mathfrak{U}$ zulässiger Steuerungen $u : [t_0, t_1] \to \mathbb{R}^m$;
- Steuerbeschränkungen: für jede Zeit $t \in [t_0, t_1]$ ist die Menge $U_t \subseteq \mathbb{R}^m$ der Werte gegeben, die die Steuerung u zur Zeit t annehmen darf;
- Zustandsbeschränkungen: für jede Zeit $t \in [t_0, t_1]$ ist die Menge $Z_t \subseteq \mathbb{R}^n$ der Werte gegeben, die der Zustand x_u zur Zeit t annehmen darf;
- ein Kostenfunktional

$$J[u] = \int_{t_0}^{t_1} L\big(t, x_u(t), u(t)\big)\,\mathrm{d}t + K\big(t_1, x_u(t_1)\big).$$

Die Optimierungsaufgabe besteht dann darin, unter allen zulässigen Steuerungen $u \in \mathfrak{U}$ mit $u(t) \in U_t$ und $x_u(t) \in Z_t$ für alle $t \in [t_0, t_1]$, die die Zielbedingung $\big(t_1, x_u(t_1)\big) \in S$ erfüllen, eine solche Steuerung u zu finden, für die $J[u]$ minimal wird. Bevor wir daran gehen, diese Aufgabe zu lösen, wollen wir die oben genannten Daten noch etwas erläutern.

• **Zielbedingungen:** Die einfachste Bedingung ist es, sowohl t_1 als auch x_1 fest vorzugeben; man will also zu einer festen Zeit einen festen Zielzustand erreichen. In diesem Fall ist $S = \{(t_1, x_1)\}$. Es könnte aber auch sein, daß zwar der Zielzustand vorgegeben ist, daß es uns aber egal ist, wann genau wir ihn erreichen (freie Endzeit); in diesem Fall ist $S = [t_0, \infty) \times \{x_1\}$. Auch der umgekehrte Fall (vorgegebene Endzeit, beliebiger Zielzustand) ist möglich; beispielsweise, wenn ein Unternehmen über einen festen Zeithorizont hinweg den Gewinn maximieren will, daß aber die Höhe des Unternehmenskapitals am Ende dieses Zeitraums egal ist. In diesem Fall ist $S = \{t_1\} \times \mathbb{R}^n$. Es kann auch sein, daß der Zielzustand x_1 weder fest vorgeschrieben noch völlig beliebig ist. (Als Kolumbus nach Amerika segelte, wollte er nicht einen fest vorgegebenen Zielpunkt erreichen, sondern irgendeinen Punkt der amerikanischen Küste.) In diesem Fall wäre S von der Form $S = \{t_1\} \times M$ (bei fester Endzeit) bzw. $S = [t_0, \infty) \times M$ (bei freier Endzeit) mit einer vorgegebenen Zielmenge $M \subseteq \mathbb{R}^n$, typischerweise einer Mannigfaltigkeit (also etwa einer Kurve oder einer Fläche).

• **Steuerungsbeschränkungen:** Für die Menge $\mathfrak{U}$ muß man eine zulässige Funktionenklasse vorgeben, etwa meßbare, (stückweise) stetige, (stückweise) glatte, stückweise konstante Funktionen (je nachdem, welche Steuerungen praktisch realisiert werden können). Gibt es sonst keine Beschränkungen, so ist $U_t = \mathbb{R}^m$ für alle $t \in [t_0, t_1]$. Meist gibt es aber Beschränkungen, etwa weil ein Antriebsgerät (z. B. ein Elektromotor) nur eine begrenzte Leistung hat. Typisch sind Steuerungsbeschränkungen der Form $a_i \leq u_i(t) \leq b_i$ für $1 \leq i \leq m$. In diesem Fall wäre also $U_t = [a_1, b_1] \times \cdots \times [a_n, b_n]$ für alle $t \in [t_0, t_1]$. Daß U_t tatsächlich von der Zeit t abhängen darf, ist in der Praxis eher unwichtig, aber denkbar. (Mit welcher Intensität Fischfang möglich ist, hängt etwa von der Jahreszeit ab.)

• **Zustandsbeschränkungen:** Es kann etwa sein, daß in einem Reaktor bestimmte Temperaturen oder Drücke nicht überschritten werden dürfen; dies wären Zustandsbeschränkungen der Form $x_i(t) \leq m_i$. Ein anderes Beispiel wäre die Vorgabe, daß beim Finden einer optimalen Flugroute bestimmte Gebiete nicht durchflogen werden dürfen oder man gegebenen Hindernissen ausweichen muß; dies wären Bedingungen der Art, daß $x(t)$ außerhalb eines (möglicherweise zeitabhängigen) verbotenen Bereichs Ω_t liegen muß, so daß dann $Z_t = \mathbb{R}^n \setminus \Omega_t$ wäre. (Zustandsbeschränkungen bedeuten in der Regel eine massive mathematische Komplikation für das Lösen des Optimierungsproblems.)

Wir formulieren nun einen allgemeinen Satz, das Pontrjaginsche Maximumprinzip, der notwendige Bedingungen angibt, die eine Steuerung $u_\star$ erfüllen muß, wenn sie optimal ist. Entsprechend der Vielzahl der Varianten, in denen Optimalsteuerungsprobleme auftreten können, gibt es auch viele Varianten des Pontrjaginschen Maximumprinzips. Dieses Prinzip, entdeckt und bewiesen in den späten 1950er Jahren von den sowjetischen Mathematikern Lev Pontrjagin (1908-1988), Vladimir Boltjanskii (*1925), Revaz Gamkrelidze (*1927) und Evgenij Mishchenko (1922-2010), gehört zu den Kronjuwelen der Mathematik des 20. Jahrhunderts. Es verbindet mathematische Tiefe mit praktischem Anwendungsbezug und stellt den Kulminationspunkt von Forschungen in den Bereichen Optimierung und Variationsrechnung dar, die sich über einen Zeitraum von rund 250 Jahren hinzogen und an denen die größten Mathematiker beteiligt waren (Newton, Jakob und Johann Bernoulli, Euler, Lagrange, Legendre, Hamilton, Jacobi, Weierstraß, Hilbert und Carathéodory, um nur die wichtigsten zu nennen). Die Herleitung dieses Prinzips (insbesondere, wenn man keine Details des Beweises unterschlagen und auch die geometrische Grundidee des Prinzips herausarbeiten will) erfordert weitaus mehr Zeit, als in dieser Vorlesung zur Verfügung steht. Wir beschränken uns daher darauf, eine einfache Variante des Prinzips zu formulieren und diese rezeptartig auf verschiedene Probleme anzuwenden. Dies ist wichtig, um ein "Gefühl" für den Inhalt des Pontrjagin-Prinzips zu entwickeln, und auch der Beweis des Prinzips (in einer Vorlesung des Master-Studiengangs) wird viel klarer, wenn man einige konkrete Beispiele vor Augen hat.

(6.1) Pontrjaginsches Maximumprinzip. *Wir betrachten das Steuerungsproblem*

$$\dot{x}(t) \;=\; f\big(t, x(t), u(t)\big) \;\in\; \mathbb{R}^n$$

mit einer C^1-Funktion f. Wir geben uns die folgenden Daten vor:

- *ein festes Zeitintervall $[t_0, t_1]$, einen festen Anfangszustand x_0 und einen festen Zielzustand x_1;*
- *einen Vektorraum $\mathfrak{U}$ meßbarer Funktionen $u : [t_0, t_1] \to \mathbb{R}^m$ mit der Eigenschaft, daß man eine Funktion $u \in \mathfrak{U}$ auf einem Teilintervall von $[t_0, t_1]$ durch eine konstante Funktion mit einem Wert in U abändern darf und die abgeänderte Funktion immer noch in $\mathfrak{U}$ liegt;*
- *eine abgeschlossene Teilmenge $U \subseteq \mathbb{R}^m$;*
- *ein Kostenfunktional, das mit einer C^1-Funktion L die folgende Form hat:*

$$J[u] \;=\; \int_{t_0}^{t_1} L\big(t, x_u(t), u(t)\big)\,\mathrm{d}t\,.$$

Die **Hamilton-Funktion** $H : [t_0, t_1] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ *werde definiert durch*

$$H(t, x, u, p, p_0) \;:=\; \langle p, f(t, x, u)\rangle + p_0 \cdot L(t, x, u).$$

Eine Steuerung $u \in \mathfrak{U}$ heiße zulässig, wenn $u(t) \in U$ für alle $t \in [t_0, t_1]$ gilt und wenn die Bedingungen $x_u(t_0) = x_0$ und $x_u(t_1) = x_1$ erfüllt sind. Unter all diesen zulässigen Steuerungen u sei $u_\star$ optimal; d.h., es gelte $J[u_\star] \le J[u]$ für alle zulässigen Steuerungen u. Es sei $x_\star = x_{u_\star}$ die aus der Steuerung $u_\star$ resultierende optimale Systemtrajektorie. Dann gibt es eine Funktion $p_\star : [t_0, t_1] \to \mathbb{R}^n$ und eine Konstante $p_{0\star} \le 0$ derart, daß die folgenden Bedingungen erfüllt sind:

- **Nichttrivialität:**

$$\big(p_{0\star}, p_\star(t)\big) \ne (0,0) \text{ für alle } t \in [t_0, t_1];$$

- **Systemgleichung:**

$$\dot{x}_\star(t) \;=\; \frac{\partial H}{\partial p}\big(t, x_\star(t), u_\star(t), p_\star(t), p_{0\star}\big);$$

- **adjungierte Gleichung:**

$$\dot{p}_\star(t) \;=\; -\frac{\partial H}{\partial x}\big(t, x_\star(t), u_\star(t), p_\star(t), p_{0\star}\big);$$

- **Maximumsbedingung:** *für jedes $t \in [t_0, t_1]$ maximiert $u^\star(t)$ die Hamiltonfunktion bezüglich u, d.h., für jdes $t \in [t_0, t_1]$ gilt*

$$\begin{aligned} & H\big(t, x_\star(t), u_\star(t), p_\star(t), p_{0\star}\big) \\ =\; & \max_{u \in U} H\big(t, x_\star(t), u, p_\star(t), p_{0\star}\big). \end{aligned}$$

Ist die Endzeit t_1 frei, so verschwindet die Hamilton-Funktion entlang der optimalen Lösung; d.h., zusätzlich gilt $H\big(t, x_\star(t), u_\star(t), p_\star(t), p_{0\star}\big) = 0$ für alle $t \in [t_0, t_1]$.

Dieser Satz sieht (jedenfalls, wenn man ihm zum ersten Mal begegnet) furchterregend und völlig unverständlich aus. Was ist etwa die Bedeutung der neben der Zustandsvariablen $x_\star$ auftretenden adjungierten Variablen $p_\star$? Was ist die Bedeutung der Hamilton-Funktion? Statt diese (völlig berechtigten) Fragen zu beantworten, wenden wir das Prinzip rezeptartig auf einige konkrete Beispiele an, um seine Nützlichkeit zu demonstrieren. (Erst, wenn man vom Sinn einer komplizierten Konstruktion überzeugt ist, ist man ja auch psychologisch bereit, sich intensiv mit dieser Konstruktion zu beschäftigen.) Zunächst machen wir aber einige Beobachtungen, die auch für das konkrete Rechnen in Beispielen wichtig sind.

(6.2) Beobachtungen. (a) Die Gleichungen

$$(\star) \qquad \dot{x} \;=\; \frac{\partial H}{\partial p}, \quad \dot{p} \;=\; -\frac{\partial H}{\partial x}$$

bezeichnet man als **Hamiltonsche Gleichungen**. Die erste dieser Gleichungen reproduziert wegen $\partial H/\partial p = f$ einfach nur die Systemgleichung $\dot{x} = f(t, x, u)$. Der Sinn der zweiten Gleichung bleibt zunächst unklar. Es stellt sich heraus, daß es sinnvoll ist, $(\star)$ nicht als System von zwei Gleichungen aufzufassen, sondern als eine einzige Differentialgleichung für die Funktion $t \mapsto \big(x(t), p(t)\big)$. Diese nimmt (in differentialgeometrischer Sprechweise) Werte im Kotangentialraum des Zustandsraums an – was immer das bedeuten mag.

(b) Die Hamiltonfunktion ist linear in (p, p_0), und keine der im Pontrjagin-Prinzip auftretenden Bedingungen wird durch eine Skalierung $(p, p_0) \mapsto (\alpha p, \alpha p_0)$ mit einem positiven Faktor $\alpha > 0$ beeinflußt. Ist also $p_0 \ne 0$, so dürfen wir o.B.d.A. $p_0 = -1$ annehmen. Wir müssen bei der Anwendung des Pontrjagin-Prinzips also letztlich nur die beiden Fälle $p_0 = -1$ und $p_0 = 0$ separat untersuchen (wobei der Fall $p_0 = 0$ ein Sonderfall ist, der in den meisten praktisch relevanten Beispielen nicht auftritt).

Wir beginnen mit einem Beispiel, das wir früher schon einmal (mit völlig anderen Methoden) behandelten.

(6.3) Insektizid-Problem. *Eine Insektenpopulation $t \mapsto x(t)$ wachse mit der natürlichen Wachstumsrate $a > 0$ und werde durch ein Insektizid bekämpft. Ist $u(t)$ die Abgaberate des Insektizids zur Zeit t, so wird die Funktion $t \mapsto x(t)$ beschrieben durch die Systemgleichung*

$$\dot{x}(t) \;=\; ax(t) - u(t).$$

Ausgehend von einer Populationsgröße $x(0) = x_0$ sei das Ziel, zu einer vorgegebenen Zeit $T > 0$ (etwa zur Zeit der nächsten Apfelblüte) alle Insekten zu töten, also die Funktion $u : [0,T] \to \mathbb{R}$ so zu wählen, daß $x_u(T) = 0$ gilt. Der angerichtete Umweltschaden sei gegeben durch das Integral $J[u] := \int_0^T u(t)^2 \mathrm{d}t$. Wie ist u zu wählen, damit einerseits das Ziel $x(T) = 0$ erreicht wird und andererseits der Ausdruck $J[u]$ möglichst klein ist?

Lösung. Wir wählen $U = \mathbb{R}$, lassen also (theoretisch) zunächst sogar negative Werte für U zu. Die Hamiltonfunktion ist

$$H(x,u,p,p_0) \;=\; p \cdot (ax - u) + p_0 \cdot u^2.$$

Wäre $p_0 = 0$, so hätten wir $H = p(ax - u)$, und diese Funktion besitzt kein Maximum bezüglich u. Wir haben also $p_0 \neq 0$ und dürfen daher $p_0 = -1$ annehmen. Dies liefert die Hamilton-Funktion

$$H(x,u,p) \;=\; p \cdot (ax - u) - u^2.$$

Als Funktion von u beschreibt dieser Ausdruck eine nach unten geöffnete Parabel, die ein eindeutig bestimmtes Maximum annimmt. Dieses finden wir durch Nullsetzen der Ableitung $\partial H / \partial u = 0$, also durch Lösen der Gleichung $-p - 2u = 0$ bzw. $u = -p/2$. Wir haben also $u_\star(t) = -p_\star(t)/2$ für alle $t \in [0,T]$. Die adjungierte Gleichung lautet

$$\dot{p} \;=\; -\frac{\partial H}{\partial x} \;=\; -ap$$

und hat die allgemeine Lösung $p(t) = -2C \cdot e^{-at}$ mit einer Konstanten C. Wegen $u = -p/2$ hat dann u die Form $u(t) = Ce^{-at}$. Setzen wir dies in die Systemgleichung ein, so ergibt sich

$$\dot{x} \;=\; ax - Ce^{-at}.$$

Die allgemeine Lösung dieser Gleichung lautet

$$x(t) \;=\; \frac{C}{2a} e^{-at} + De^{at}$$

mit einer weiteren Konstanten D. Aus den Bedingungen $x(0) = x_0$ und $x(T) = 0$ erhalten wir

$$C \;=\; \frac{2a}{1 - e^{-aT}} \quad \text{und} \quad D \;=\; \frac{-1}{e^{2aT} - 1}.$$

Wir erhalten als optimale Steuerung also

$$u(t) \;=\; \frac{2ae^{-at}}{1 - e^{-aT}}$$

(so daß qualitativ die optimale Lösung also darin besteht, die Sprührate des Insektizids als exponentiell abnehmende Funktion zu wählen) und als zugehörige Systemtrajektorie folglich

$$x(t) \;=\; \frac{e^{-at} - e^{-2aT} \cdot e^{at}}{1 - e^{-2aT}}.$$

■

(6.4) Bemerkung. Streng genommen haben wir im vorigen Beispiel nicht bewiesen, daß die gefundene Funktion u eine optimale Steuerung ist; wir haben nur bewiesen, daß außer u keine andere Funktion als Kandidat für eine Optimalsteuerung in Frage kommt. Dieses Problem ist inhärent bei Anwendung des Pontrjagin-Prinzips, denn dieses ist nur eine *notwendige* Bedingung für Optimalität: *Wenn* $u_\star$ eine optimale Steuerung ist, *dann* müssen die Bedingungen des Pontrjagin-Prinzips erfüllt sein.* Zusätzlich sind also noch Existenzsätze von Interesse, die für Optimalsteuerungsprobleme *a priori* die Existenz einer optimalen Steuerung garantieren.

Im nächsten Beispiel (einem “Klassiker” in der Theorie der Optimalsteuerungen) betrachten wir ein Problem, in dem Steuerungsbeschränkungen auftreten.

(6.5) Raketenwagen. *Wir betrachten ein Fahrzeug der Masse m, das sich nur in einer Richtung bewegen kann; wir bezeichnen mit $x(t)$ seine Position zur Zeit t. Ist $F(t)$ die Antriebskraft, die wir auf das Fahrzeug ausüben, so gilt das Newtonsche Grundgesetz $m\ddot{x}(t) = F(t)$. Wir setzen $u(t) := F(t)/m$ und haben dann die Systemgleichung*

$$\ddot{x} = u \qquad \textit{bzw.} \qquad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ u \end{bmatrix}$$

mit der Position $x_1 := x$ und der Geschwindigkeit $x_2 := \dot{x}_1$. Beginnend mit einem beliebigen Anfangszustand $\big(x_1(0), x_2(0)\big) = (x_0, v_0)$ wollen wir das Fahrzeug in möglichst kurzer Zeit T in den Zielzustand $\big(x_1(T), x_2(T)\big) = (0,0)$

---

* Eine ähnliche Situation ist uns aus der Analysis I bekannt: *Wenn* eine Funktion $f : \mathbb{R} \to \mathbb{R}$ an einer Stelle x_0 ein Maximum oder Minimum annimmt, *dann* muß $f'(x_0) = 0$ gelten. Aus der Bedingung $f'(x_0) = 0$ kann aber *nicht* geschlossen werden, daß bei x_0 ein Extremum auftritt, selbst wenn die Gleichung $f'(x_0) = 0$ eine eindeutige Lösung hat. (Man denke an $f(x) = x^3$ und $x_0 = 0$.)

steuern, wobei wir eine Antriebsbeschränkung $|u(t)| \le u_{\max}$ *voraussetzen. Der Einfachheit halber seien die Einheiten so gewählt, daß* $u_{\max} = 1$ *gilt. Wie ist die optimale Steuerung zu wählen?*

Lösung. Zu minimieren ist das Integral $\int_0^T 1\,dt$ bei freier Endzeit T. Wir haben also $L \equiv 1$, und die Hamiltonfunktion lautet

$$H(x_1, x_2, u, p_1, p_2, p_0) \;=\; p_1 x_2 + p_2 u + p_0.$$

Die adjungierte Gleichung lautet

$$\begin{bmatrix} \dot p_1 \\ \dot p_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -p_1 \end{bmatrix}$$

und hat zur Folge, daß p_1 konstant und p_2 linear ist; genauer gilt $p_2(t) = -p_1 t + C$ mit einer Konstanten C. Die Maximumsbedingung erzwingt $u(t) = 1$, falls $p_2(t) > 0$ gilt, und $u(t) = -1$, falls $p_2(t) < 0$ gilt. Zu einer Zeit t, zu der $p_2(t) = 0$ gilt, kann der Wert $u(t)$ nicht aus der Maximumsbedingung hergeleitet werden, aber da p_2 linear ist und daher höchstens einen Vorzeichenwechsel hat, kann dies nur zu einem einzigen isolierten Zeitpunkt passieren (und insbesondere nicht auf einem ganzen Zeitintervall). Wir haben also das folgende Ergebnis: Die gesuchte Steuerung kann nur die Werte ± 1 annehmen und höchstens einmal ihr Vorzeichen ändern. Diese Information reicht aus, um für einen gegebenen Anfangszustand (x_0, v_0) die zugehörige Optimalsteuerung zu finden. Auf einem Intervall, auf dem $u \equiv 1$ gilt, haben wir nämlich

$$\begin{bmatrix} \dot x_1 \\ \dot x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ 1 \end{bmatrix} \text{ und damit } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (t^2/2) + at + b \\ t + a \end{bmatrix}$$

mit geeigneten Konstanten a und b und folglich

$$x_1 \;=\; \frac{x_2^2}{2} + A$$

mit einer geeigneten Konstanten A $\big($nämlich $A = b - (a^2/2)\big)$. Auf einem Intervall, auf dem $u \equiv -1$ ist, haben wir dagegen

$$\begin{bmatrix} \dot x_1 \\ \dot x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -1 \end{bmatrix} \text{ und damit } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -(t^2/2) + at + b \\ -t + a \end{bmatrix}$$

mit geeigneten Konstanten a und b und folglich

$$x_1 \;=\; \frac{-x_2^2}{2} + B$$

mit einer geeigneten Konstanten B $\big($nämlich $B = b + (a^2/2)\big)$. Die optimalen Zustandstrajektorien sind also Parablen der Form $x_1 = \pm(x_2^2/2) + C$. Die folgende Abbildung zeigt einige typische Kurven (wobei die blauen Parabelbögen zur Steuerung $u \equiv 1$ und die roten Parabelbögen zur Steuerung $u \equiv -1$ gehören).

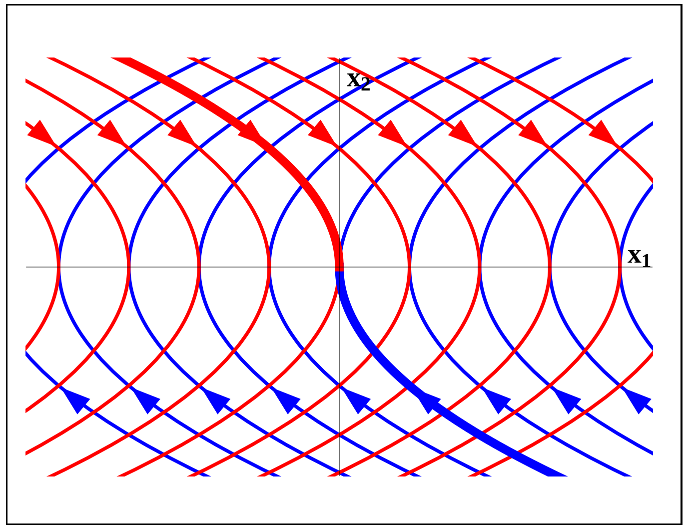

Abb. 6.5: Optimale Trajektorien für den Raketenwagen.

Die fett markierte Kurve mit der Gleichung $x = -\operatorname{sign}(y) \cdot y^2/2$ bzw. $y = -\operatorname{sign}(x) \cdot \sqrt{2|x|}$ wird als **Umschaltkurve** bezeichnet. Liegt der Anfangszustand genau auf dieser Kurve, so wird der Zielzustand $(0,0)$ ohne Umschalten erreicht; die optimale Steuerung ist $u \equiv -1$ für $y_0 > 0$ und $u \equiv 1$ für $y_0 < 0$.

- Liegt der Anfangszustand (x_0, y_0) unterhalb der Umschaltkurve, so muß man zunächst maximal beschleunigen ($u \equiv 1$), bis man die Umschaltkurve erreicht hat, und dann maximal abbremsen ($u \equiv -1$), bis man den Zielzustand $(0,0)$ erreicht hat.

- Liegt der Anfangszustand (x_0, y_0) oberhalb der Umschaltkurve, so muß man zunächst maximal abbremsen ($u \equiv -1$), bis man die Umschaltkurve erreicht hat, und dann maximal beschleunigen ($u \equiv 1$), bis man den Zielzustand $(0,0)$ erreicht hat.

Übungsaufgabe: Berechne für einen beliebigen Anfangszustand $\big(x_1(0), x_2(0)\big) = (x_0, v_0)$ die Zeit T bis zum Erreichen des Zielzustands sowie die Zeit θ bis zum Umschalten zwischen voller Beschleunigung und vollem Abbremsen.

(6.6) Beobachtungen. (a) In diesem Beispiel liefert das Pontrjagin-Prinzip die generelle Form der optimalen Lösungen, aber man muß dann noch aus den möglichen optimalen Trajektorien die für gegebene Randbedingungen richtige Trajektorie "zusammensetzen". Diese Vorgehensweise kommt häufig vor und wird als **Kontrollsynthese** bezeichnet.

(b) Beim Raketenwagen sind *a priori* alle Beschleunigungswerte im Intervall $[-1, 1]$ möglich, aber es stellt sich heraus, daß eine optimale Steuerung nur die extremen Werte 1 und -1 annimmt. Eine solche Steuerung wird als "Bang-Bang-Steuerung" bezeichnet. (Sie entspricht einem Fahrstil, den man sich vielleicht bei einer Rallye, aber nicht im Straßenverkehr angewöhnen sollte.) Es stellt sich heraus, daß solche Bang-Bang-Steuerungen für lineare Systeme mit konstanten Koeffizienten typische zeitoptimale Lösungen sind.

(6.7) Bang-Bang-Prinzip. *Wir betrachten ein lineares gesteuertes System $\dot{x}(t) = Ax(t) + Bu(t)$ mit $x \in \mathbb{R}^n$ und $u \in \mathbb{R}^m$. Dieses System soll in minimaler Zeit T aus einem Anfangs- in einen Endzustand überführt werden, wobei die Steuerung u Einschränkungen der Form $|u_i(t)| \leq a_i$ für $1 \leq i \leq m$ unterliege; d.h., es sind nur Steuerungen $u : [0,T] \to U$ mit $U := [-a_1, a_1] \times \cdots \times [-a_m, a_m]$ zulässig. Es seien $b_1, \ldots, b_m$ die Spalten von B, und für jeden Index $1 \leq i \leq m$ sei das System $\dot{x}(t) = Ax(t) + b_i u_i(t)$ vollständig steuerbar. (Im Jargon der Regelungstechniker sagt man, das ursprüngliche System sei über jeden einzelnen Steuerungskanal schon vollständig steuerbar.) Dann kann eine optimale Steuerung nur die Werte $(\pm a_1, \pm a_2, \ldots, \pm a_m)$ annehmen.*

Beweis. Hier ist $\int_0^T 1\,\mathrm{d}t$ zu minimieren; d.h., wir haben $L \equiv 1$. Die Hamiltonfunktion ist

$$\begin{aligned} H(x,u,p,p_0) &= \langle p, Ax + Bu\rangle + p_0 \\ &= \langle A^T p, x\rangle + \langle p, Bu\rangle + p_0. \end{aligned}$$

Es seien $u = u_\star$ eine optimale Steuerung, $x = x_\star$ die zugehörige Systemtrajektorie und $p = p_\star$ die zugehörige adjungierte Variable. Die adjungierte Gleichung lautet $\dot{p}(t) = -A^T p(t)$ und hat die allgemeine Lösung $p(t) = e^{-tA^T} p(0)$. Gälte $p(t) = 0$ auch nur für einen einzigen Zeitpunkt t, so wäre $p \equiv 0$, aufgrund der Nichttrivialitätsbedingung also $p_0 \neq 0$, was aber nicht sein kann, weil dann einerseits $H \equiv p_0$ wäre, andererseits aber $H \equiv 0$, weil die Endzeit frei ist. Wir haben also $p(t) \neq 0$ für alle $t \in [0,T]$. Die Maximumsbedingung erzwingt, daß für jede Zeit t der Wert $u_\star(t)$ den Ausdruck

$$\langle p, Bu\rangle = \langle p, \sum_{i=1}^m b_i u_i\rangle = \sum_{i=1}^m u_i \langle p, b_i\rangle$$

maximiert. Da die einzelnen Steuerkomponenten u_i voneinander unabhängig sind, bedeutet dies, daß für jeden einzelnen Index i der Ausdruck $u_i\langle p, b_i\rangle$ maximiert wird, und dies bedeutet

$$u_i(t) = \begin{cases} a_i, & \text{falls } \langle p(t), b_i\rangle > 0; \\ -a_i, & \text{falls } \langle p(t), b_i\rangle < 0. \end{cases}$$

Falls $\langle p(t), b_i\rangle = 0$ gilt, so kann $u_i(t)$ nicht aus der Maximumsbedingung bestimmt werden. Wir sind fertig, wenn wir zeigen können, daß dies nur zu einzelnen isolierten Zeitpunkten passieren kann, daß also nicht auf einem ganzen Zeitintervall $I \subseteq [0,T]$ die Bedingung $\langle p(t), b_i\rangle \equiv 0$ gelten kann. Wäre dies der Fall, so müßte auf dem Zeitintervall I auch

$$\begin{aligned} 0 &= (\mathrm{d}/\mathrm{d}t)\langle p(t), b_i\rangle = \langle \dot{p}(t), b_i\rangle \\ &= \langle -A^T p(t), b_i\rangle = -\langle p(t), Ab_i\rangle \end{aligned}$$

gelten, folglich auch

$$\begin{aligned} 0 &= (\mathrm{d}/\mathrm{d}t)\langle p(t), Ab_i\rangle = \langle \dot{p}(t), Ab_i\rangle \\ &= \langle -A^T p(t), Ab_i\rangle = -\langle p(t), A^2 b_i\rangle \end{aligned}$$

und so weiter, insgesamt also $\langle p(t), A^k b_i\rangle = 0$ für alle $k \in \mathbb{N}_0$. Dann würde $p(t)$ aber senkrecht auf allen Spalten der Kalman-Matrix $(b_i \mid Ab_i \mid \cdots \mid A^{n-1}b_i)$ stehen, damit aber wegen der vorausgesetzten Steuerbarkeitsbedingung auf ganz $\mathbb{R}^n$, was nur für $p(t) = 0$ möglich ist. Dieser Widerspruch zeigt, daß die Bedingung $\langle p(t), b_i\rangle = 0$ nicht auf einem ganzen Zeitintervall $I \subseteq [0,T]$ gelten kann. Damit ist die Behauptung bewiesen. ■

Als nächstes Beispiel behandeln wir wieder ein Problem, bei dem eine zeitoptimale Lösung gesucht ist, dieses Mal aber mit einer nichtlinearen Systemgleichung.

(6.8) Zermelo-Problem. *In einem unendlich ausgedehnten Gewässer sei die Strömung durch ein Vektorfeld $(x,y) \mapsto \big(f(x,y), g(x,y)\big)$ gegeben. Ein Schwimmer kann sich mit konstanter Geschwindigkeit (deren Betrag wir o.B.d.A. als 1 annehmen können) relativ zum Wasser bewegen. Er will in möglichst kurzer Zeit T von einem gegebenen Ausgangspunkt (x_0, y_0) zu einem gegebenen Zielpunkt (x_1, y_1) gelangen. Entlang welcher Bahnkurve muß er schwimmen? Bezeichnen wir mit $\varphi(t)$ den Winkel, den die Schwimmrichtung zur Zeit t mit der x-Achse bildet, so lautet die Aufgabe folgendermaßen: Für welche Steuerungsfunktion $t \mapsto \varphi(t)$ erfüllt die Lösung des Anfangswertproblems*

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} f(x,y) + \cos\varphi \\ g(x,y) + \sin\varphi \end{bmatrix}, \qquad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

für einen möglichst kleinen Wert $T > 0$ die Bedingung $\big(x(T), y(T)\big) = (x_1, y_1)$?

Hier lösen wir das Problem nicht vollständig, sondern leiten eine Differentialgleichung her, die die optimale Steuerung erfüllen muß. In einem konkret gegebenen Beispiel ist dann ein Randwertproblem zu lösen, um die optimale Steuerung zu finden.

(6.9)) Satz. *Ist im Zermelo-Problem die Steuerung $t \mapsto \varphi(t)$ optimal gewählt und ist $t \mapsto \big(x(t), y(t)\big)$ die aus dieser Steuerung resultierende Bahnkurve, so gilt*

$$\begin{aligned} \dot{x} &= f(x,y) + \cos\varphi, \\ \dot{y} &= g(x,y) + \sin\varphi, \\ \dot{\varphi} &= g_x(x,y)\sin^2\varphi - f_y(x,y)\cos^2\varphi \\ &\qquad + \big(f_x(x,y) - g_y(x,y)\big)\sin\varphi\cos\varphi. \end{aligned}$$

Beweis. Das Kontrollsystem lautet

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} f(x,y) + \cos\varphi \\ g(x,y) + \sin\varphi \end{bmatrix};$$

zu minimieren ist die Zeit $T = \int_0^T 1\,\mathrm{d}t$. Die Hamiltonfunktion lautet daher

$$H = 1 + \lambda\big(f(x,y) + \cos\varphi\big) + \mu\big(g(x,y) + \sin\varphi\big).$$

Die adjungierten Gleichungen lauten

$$\begin{aligned}\dot\lambda &= -\frac{\partial H}{\partial x} = -\lambda\, f_x - \mu\, g_x,\\ \dot\mu &= -\frac{\partial H}{\partial y} = -\lambda\, f_y - \mu g_y,\end{aligned}$$

und die Maximumsbedingung liefert

$$0 = \frac{\partial H}{\partial \varphi} = -\lambda \sin\varphi + \mu\cos\varphi.$$

Die letzte Bedingung drückt aus, daß $(\mu,-\lambda)^T$ senkrecht auf dem Vektor $(\cos\varphi,\sin\varphi)^T$ steht und daher $(\lambda,\mu)^T$ ein Vielfaches dieses Vektors sein muß. Es gibt also eine Funktion $t\mapsto w(t)$ mit

$$\begin{bmatrix}\lambda\\ \mu\end{bmatrix} = w\begin{bmatrix}\cos\varphi\\ \sin\varphi\end{bmatrix}.$$

Einsetzen in die adjungierten Gleichungen ergibt

$$\begin{aligned}\dot w\cos\varphi - w\dot\varphi\sin\varphi &= -wf_x\cos\varphi - wg_x\sin\varphi,\\ \dot w\sin\varphi + w\dot\varphi\cos\varphi &= -wf_y\cos\varphi - wg_y\sin\varphi.\end{aligned}$$

Diese Gleichungen lassen sich zusammenfassen zu der Vektorgleichung

$$\dot w\begin{bmatrix}\cos\varphi\\ \sin\varphi\end{bmatrix} + w\dot\varphi\begin{bmatrix}-\sin\varphi\\ \cos\varphi\end{bmatrix} = -w\begin{bmatrix}f_x\cos\varphi + g_x\sin\varphi\\ f_y\cos\varphi + g_y\sin\varphi\end{bmatrix}.$$

Bilden wir auf beiden Seiten das Skalarprodukt mit $(-\sin\varphi,\cos\varphi)^T$, so ergibt sich die Gleichung $w\dot\varphi = w\big(f_x\sin\varphi\cos\varphi + g_x\sin^2\varphi - f_y\cos^2\varphi - g_y\sin\varphi\cos\varphi$, nach Division durch w also

$$\dot\varphi = g_x\sin^2\varphi + (f_x-g_y)\sin\varphi\cos\varphi - f_y\cos^2\varphi.$$

Ist also $t\mapsto\varphi(t)$ eine optimale Steuerung und ist $t\mapsto \big(x(t),y(t)\big)$ die zugehörige optimale Bahnkurve, so haben wir

$$\begin{aligned}\dot x &= f(x,y)+\cos\varphi,\\ \dot y &= g(x,y)+\sin\varphi,\\ \dot\varphi &= g_x\sin^2\varphi + (f_x-g_y)\sin\varphi\cos\varphi - f_y\cos^2\varphi.\end{aligned}$$

■

Zum Abschluß betrachten wir das Dubinssche Problem. In einer Ebene seien eine Anfangsposition A, eine Endposition B, eine Anfangsrichtung v und eine Endrichtung w gegeben. Ein Fahrzeug soll entlang einer möglichst kurzen Kurve so von A nach B bewegt werden, und zwar so, daß v die Fahrtrichtung am Anfang und w die Fahrtrichtung am Ende ist. Die Kurve muß dabei so sein, daß der Krümmungsradius der Kurve nicht kleiner wird als ein vorgegebener Wert R (der Wenderadius des Fahrzeugs). Wir können die Längeneinheit so festlegen, daß $R=1$ gilt. Gibt der Winkel $\varphi(t)$ die Fahrtrichtung zur Zeit t an, so muß also $|\dot\varphi(t)|\le 1$ für alle t gelten. Die Aufgabenstellung lautet in kontrolltheoretischer Terminologie also folgendermaßen.

(6.10) Dubins-Problem. *Finde eine Steuerungsfunktion u mit $|u|\le 1$ derart, daß die Lösung des Anfangswertproblems*

$$\begin{bmatrix}\dot x\\ \dot y\\ \dot\varphi\end{bmatrix} = \begin{bmatrix}\cos\varphi\\ \sin\varphi\\ u\end{bmatrix},\qquad \begin{bmatrix}x(0)\\ y(0)\\ \varphi(0)\end{bmatrix} = \begin{bmatrix}x_0\\ y_0\\ \varphi_0\end{bmatrix}$$

nach möglichst kurzer Zeit T den vorgegebenen Zielzustand $\big(x(T),y(T),\varphi(T)\big) = (x_1,y_1,\varphi_1)$ erreicht. Zu minimieren ist also das Kostenfunktional $J[u]=\int_0^T 1\,\mathrm{d}t$ mit freier Endzeit T.

(6.11) Satz. *Eine optimale Lösung des Dubins-Problems ist ein Teilstück einer Kurve, die entweder aus drei aneinandergesetzten Kreisbögen vom Radius R oder aber einem Kreisbogen vom Radius R, einem Geradenstück und einem weiteren Kreisbogen vom Radius R zusammengesetzt ist.*

Beweis. Wir bezeichnen die zu (x,y,φ) gehörigen adjungierten Variablen mit (λ,μ,ρ). Die Hamilton-Funktion ist gegeben durch

$$H = -p_0 + \lambda\cos\varphi + \mu\sin\varphi + \rho u$$

mit einer Konstanten $p_0\ge 0$. Da die Endzeit frei ist, gilt $H\equiv 0$ entlang einer optimalen Trajektorie. Die adjungierten Gleichungen lauten

$$\dot\lambda = 0,\quad \dot\mu = 0,\quad \dot\rho = \lambda\sin\varphi - \mu\cos\varphi\,.$$

Wir sehen, daß λ und μ konstant sind. Schreiben wir $\lambda = \alpha\cos\beta$ und $\mu = -\alpha\sin\beta$ (mit $\alpha := \sqrt{\lambda^2+\mu^2}\ge 0$), so geht die letzte der adjungierten Gleichungen über in

$$\dot\rho(t) = \alpha\sin\big(\varphi(t)+\beta\big),$$

und die Hamilton-Funktion nimmt die Form

$$H = -p_0 + \alpha\cos\big(\varphi(t)+\beta\big) + \rho(t)\,u(t)$$

an. Wir unterscheiden nun zwei Fälle.

• **Erster Fall:** $\alpha = 0$, folglich $\lambda\equiv\mu\equiv 0$. In diesem Fall haben wir $\dot\rho = 0$, so daß $\rho(t)\equiv\rho_0$ konstant ist. (Notwendigerweise gilt $\rho_0\ne 0$, denn sonst wären alle adjungierten Variablen identisch Null.) Die Hamilton-Funktion lautet dann $H = -p_0 + \rho_0\,u$. Die Maximumbedingung liefert dann $u(t) = \mathrm{sign}(\rho_0) = \pm 1$. Diese Steuerung liefert als Fahrtkurve einen Kreisbogen (der je nach Vorzeichen von u gegen den Uhrzeigersinn oder im Uhrzeigersinn durchfahren wird).

• **Zweiter Fall:** $\alpha > 0$. Da die adjungierten Gleichungen homogen sind, können wir λ,μ,ρ durch $\lambda/\alpha,\mu/\alpha,\rho/\alpha$

ersetzen und erhalten die Gleichung $\dot\rho(t) = \sin\bigl(\varphi(t)+\beta\bigr)$ und die Hamiltonfunktion $H = -(p_0/\alpha) + \cos\bigl(\varphi(t)+\beta\bigr) + \rho(t)u(t)$. Da die Hamilton-Funktion entlang einer optimalen Lösung identisch verschwindet, gilt daher

$$\cos\bigl(\varphi(t)+\beta\bigr) + \rho(t)u(t) \;=\; \text{const} \;\geq\; 0.$$

Wir betrachten nun die Menge $J := \{t \in [0,T] \mid \rho(t) \neq 0\}$ und unterscheiden (innerhalb des gerade betrachteten zweiten Falls) zwei Unterfälle.

Erster Unterfall: J enthält ein (maximal gewähltes) offenes Intervall (τ_1,τ_2). Wir haben dann $\rho(\tau_1) = \rho(\tau_2) = 0$, aber $\rho(t) \neq 0$ für $\tau_1 < t < \tau_2$. Wegen $\cos\bigl(\varphi(\tau_1)+\beta\bigr) \geq 0$ liegt der Winkel $\theta := \varphi(\tau_1)+\beta$ in $[0,\pi/2] \cup [3\pi/2, 2\pi)$. Wir nehmen zunächst $\theta \in [0,\pi/2]$ an. Dann gilt $\dot\rho(\tau_1) = \sin\bigl(\varphi(\tau_1)+\beta\bigr) = \sin\theta \geq 0$ und daher $u(t) = +1$ bzw. $\dot\varphi(t) = 1$ für $\tau_1 < t < \tau_2$. (Der Fall $\theta \in [3\pi/2, 2\pi)$ wird vollkommen analog behandelt und entspricht der Bedingung $u \equiv -1$.) Hieraus folgt $\varphi(t) = t + \varphi(\tau_1) - \tau_1$ auf dem Intervall $[\tau_1,\tau_2]$. Wir rechnen

$$\begin{aligned} 0 = \rho(\tau_2) - \rho(\tau_1) &= \int_{\tau_1}^{\tau_2} \dot\rho(t)\,\mathrm{d}t = \int_{\tau_1}^{\tau_2} \sin\bigl(\varphi(t)+\beta\bigr)\,\mathrm{d}t \\ &= \int_{\tau_1}^{\tau_2} \sin(t+\theta-\tau_1)\,\mathrm{d}t = -\cos(\tau_2-\tau_1+\theta) + \cos(\theta) \end{aligned}$$

und sehen, daß $\cos(\theta) = \cos(\tau_2-\tau_1+\theta)$ und damit $\tau_2 - \tau_1 = 2(\pi-\theta)$ gilt. Aus dieser Gleichung ziehen wir mehrere Folgerungen.

- Es kann nicht $\theta = 0$ gelten, denn andernfalls hätten wir $u \equiv 1$ auf einem Intervall der Länge 2π und wüden daher einen vollen Kreisbogen durchfahren, was nicht optimal sei kann.
- Es gilt $\tau_2 - \tau_1 \in [\pi, 2\pi)$. Mit anderen Worten, wir durchfahren einen Kreisbogen, der mindestens ein Halbkreis ist. Die Länge dieses Bogens $\bigl($nämlich $2(\pi-\theta)\bigr)$ ist vollständig durch $\dot\rho(\tau_1) = \sin\theta > 0$ bestimmt. Dies impliziert, daß u sein Vorzeichen zur Zeit τ_1 von -1 zu $+1$ wechselt.
- Es gilt $\dot\rho(\tau_2) = \sin\bigl(\varphi(tau_2)+\beta\bigr) = \sin(2\pi-\theta) = -\sin\theta < 0$. Dies bedeutet, daß u zur Zeit τ_2 von $+1$ nach -1 wechselt.
- Die beiden letzten Bedingungen zeigen, daß in aufeinanderfolgenden Intervallen der betrachteten Art zwei kongruente Kreisbögen durchlaufen werden, der eine im Uhrzeigersinn, der andere gegen den Uhrzeigersinn.

Die folgende Abbildung zeigt, daß es nicht mehr als drei solcher Bögen geben kann und daß im Fall dreier Bögen die beiden äußeren nicht zugleich mindestens Halbkreise sein können. (Beachte, daß diese Einschränkung *nicht* aus dem Pontrjagin-Prinzip folgt, sondern aus dem expliziten Überprüfen, daß manche Trajektorien, die die Bedingungen des Pontrjagin-Prinzips erfüllen, nicht optimal sind. Das Pontrjagin-Prinzips ist eine notwendige, aber keine hinreichende Bedingung für Optimalität.)

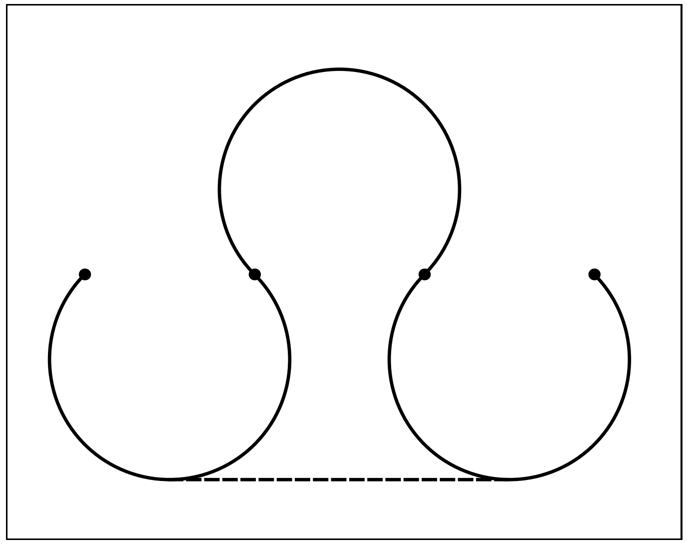

Abb. 6.11a: Nichtoptimalität mancher nach dem Pontrjagin-Prinzip möglichen Trajektorien.

Zweiter Untervall: Die Menge J habe die Form $[0,\tau_1) \cup (\tau_2, T]$ (wobei eines der Teilintervalle oder sogar beide fehlen lönnen). Gilt $\tau_1 = \tau_2 =: \tau$, so hat ρ eine einzige Nullstelle zur Zeit τ, und die Steuerung u ist gegeben durch $u(t) = \operatorname{sign}\bigl(\rho(t)\bigr)$. In diesem Fall durchfahren wir also zwei Kreisbögen, den einen im Uhrzeigersinn, den andern gegen den Uhrzeigersinn. Gilt dagegen $\tau_1 < \tau_2$, so haben wir $\rho(t) \neq 0$ für $t \in [0,\tau_1) \cup (\tau_2, T]$ und $\rho(t) = 0$ für $t \in [\tau_1,\tau_2]$. In diesem Fall können wir die Werte der Steuerung u auf dem Intervall $[\tau_1,\tau_2]$ nicht aus der Maximumsbedingung erschließen. Wir können aber folgendermaßen argumentieren: Wegen $H \equiv 0$ auf $[\tau_1,\tau_2]$ ist die Funktion $t \mapsto \cos\bigl(\varphi(t)+\beta\bigr)$ und damit auch die Funktion $t \mapsto \varphi(t)$ auf dem Intervall $[\tau_1,\tau_2]$ konstant. Hieraus folgt $u = \dot\varphi \equiv 0$ auf $[\tau_1,\tau_2]$. Während des Zeitintervalls $[\tau_1,\tau_2]$ durchfahren wir also ein Geradenstück. Die optimale Lösung in einem solchen Fall wird durch das folgende Diagramm beschrieben.

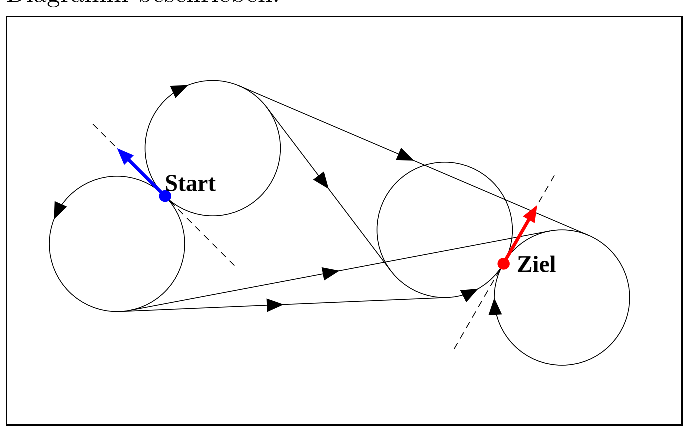

Abb. 6.11b: Optimale Trajektorien mit einem Geradenstück in der Mitte.

Damit ist alles gezeigt. Welche Lösungskurve bei einer konkret gegebenen Anfangs- und Endkonfiguration optimal ist, wird durch Durchprobieren der (endlich vielen!) Möglichkeiten ermittelt, die nach Anwendung des gerade bewiesenen Satzes verbleiben. ■