Universität Ulm
Fakultät für Mathematik und Wirtschaftswissenschaften

Zur Definitheit quadratischer Funktionale

Diplomarbeit zur Erlangung des akademischen Grades „Diplom-Mathematiker“

von Markus Wahrheit

1. Gutachter: Prof. Dr. W. Kratz
2. Gutachter: Prof. Dr. U. Stadtmüller

Ulm, im September 2003

Inhaltsverzeichnis

Bezeichnungen

Grundsätzlich wollen wir in dieser Arbeit Matrizen mit Großbuchstaben, Vektoren und skalare Größen mit Kleinbuchstaben bezeichnen. Außerdem sollen folgende Konventionen gelten:

Es bezeichnet

$\mathbb{N}$	die Menge der natürlichen Zahlen.
$\mathbb{R}$	die Menge der reellen Zahlen.
$\mathbb{R}^n$	die Menge der reellwertigen Vektoren der Dimension n.
$\mathbb{R}^{m\times n}$	die Menge der $m \times n$-Matrizen, deren Elemente reelle Zahlen sind.
$[a,b]$	das abgeschlossene reelle Zahlenintervall mit den Eckpunkten a und b.
$[a,b)$	das halboffene reelle Zahlenintervall mit den Eckpunkten a und b, wobei der Eckpunkt a im Gegensatz zum Eckpunkt b zum Intervall gehört.
$(a,b]$	das halboffene reelle Zahlenintervall mit den Eckpunkten a und b, wobei der Eckpunkt a im Gegensatz zum Eckpunkt b nicht zum Intervall gehört.
(a,b)	das offene reelle Zahlenintervall mit den Eckpunkten a und b.
$\mathcal{C}_s([a,b])$	die Menge der stückweise stetigen reellwertigen Funktionen auf $[a,b]$, d.h. die Menge aller reellen Funktionen auf $[a,b]$, die mit Ausnahme endlich vieler Punkte stetig sind und an ihren Unstetigkeitsstellen einseitige Grenzwerte besitzen.
$0_{n\times n}, I_{n\times n}$	die $n\times n$-Nullmatrix, bzw. die $n \times n$-Einheitsmatrix. Geht die Dimension klar aus dem Kontext hervor, so verzichten wir auf die Dimensionsangabe $n \times n$.
$\mathrm{Im} M$	das Bild der Matrix $M \in \mathbb{R}^{m\times n}$.
$\ker M$	den Kern der Matrix $M \in \mathbb{R}^{m\times n}$.
$\mathrm{rg} M$	den Rang der Matrix $M \in \mathbb{R}^{m\times n}$.
M^T	die Transponierte der Matrix $M \in \mathbb{R}^{m\times n}$.
M^{-1}	die Inverse der Matrix $M \in \mathbb{R}^{n\times n}$.
$M^\dagger$	die Moore-Penrose-Inverse von $M \in \mathbb{R}^{m\times n}$ (siehe [4], Seite 91).
$M \geq 0$	eine nichtnegativ definite Matrix $M \in \mathbb{R}^{n\times n}$, d.h. $M = M^T$ und $x^T M x \geq 0$ für alle Vektoren $x \in \mathbb{R}^n$.
$M > 0$	eine positiv definite Matrix $M \in \mathbb{R}^{n\times n}$, d.h. $M = M^T$ und $x^T M x > 0$ für alle Vektoren $x \in \mathbb{R}^n \backslash \{0\}$.
$M(t)$	eine Matrix, deren Einträge reellwertige Funktionen der Variable t sind. Entsprechend sollen Operationen wie Differenzieren oder Integrieren solcher Objekte immer komponentenweise verstanden sein, also etwa
$\dot{M}(t)$	die Matrix $(\dot{m}_{\mu\nu}(t)) = (\frac{d}{dt}\, m_{\mu\nu}(t))$ für $M(t) = (m_{\mu\nu}(t))$.
$\|x\|$	$= \sqrt{\sum_{\mu=1}^{n} x_\mu^2}$ die Euklidische Vektornorm.
$\|M\|$	$= \max_{\|x\|=1} \|Mx\|$ die Spektralnorm der Matrix $M \in \mathbb{R}^{n\times n}$.
$\oplus$	eine direkte Summe.

Vorwort

Die vorliegende Arbeit beschäftigt sich mit der Definitheit quadratischer Funktionale, die gewissen Bewegungsgleichungen und linearen homogenen Randwertbedingungen genügen. Diese Untersuchung ist von besonderem Interesse in der klassischen Variationsrechnung, in der quadratische Funktionale als zweite Variation auftreten. Vergleiche hierzu [[8], Chapter 1.8] und [[4], Chapter 2.3 und 8.2].

Die Positivität solcher Funktionale, die in der klassischen Variationsrechnung bei der hinreichenden Bedingung für (schwache) Minima eingeht, ist wohlbekannt. W.T. Reid gibt hinreichende und notwendige Bedingungen für die Positivität dieser quadratischen Funktionale in den Spezialfällen der so genannten *Dirichlet Nebenbedingungen* (siehe Seite 2 und [[6], Chapter VII, Theorem 4.5]) und bei festen Randpunkten in b mit freien Randpunkten in a der betrachteten Funktionen (vergleiche [6], Chapter VII, Theorem 6.2) an. In [[4], Theorem 2.4.2] beweist W. Kratz den allgemeinen Fall, der die Steuerbarkeit (siehe Definition 3.1.) des betrachteten Systems voraussetzt.

Die Nichtnegativität dieser Funktionale, die eine notwendige Bedingung für ein Minimum darstellt, ist zwar bereits in [[4], Remark 2.4.2] skizziert, doch erst in diesem Jahr von W. Kratz ohne die Voraussetzung der Steuerbarkeit, derer wir uns in dieser Arbeit bedienen möchten, in [5] veröffentlicht. In diesem Sinn sind die Sätze und Beweise, die wir im Fall der Nichtnegativität führen werden, „neu“ und insbesondere ist die Beweistechnik verschieden von [5].

Bevor wir jedoch die Sätze zur Definitheit jener quadratischen Funktionale im sechsten Abschnitt formulieren und beweisen, die die bekannte *(verschärfte) Jacobi Bedingung* als Spezialfall beinhaltet, siehe [[4], Theorem 8.2.6], führen wir im ersten Abschnitt sowohl das quadratische Funktional als auch die Bewegungsgleichungen und Randbedingungen, derer es obliegt, ein und formulieren die gegebene Problematik. Zudem beweisen wir bereits zu diesem Zeitpunkt eine erste notwendige Bedingung für die Nichtnegativität dieser Funktionale, die der bekannten *Legendre Bedingung* entspricht. Diese werden wir im weiteren Verlauf dieser Arbeit gelegentlich voraussetzen.

Um weitere notwendige Bedingungen für die Nichtnegativität zu ermitteln, nutzen wir die Erkenntnisse der klassischen Variationsrechnung, die auf ein selbstadjungiertes lineares Gleichungssystem, dem so genannten *Hamilton System*, führen, das die *Eulergleichungen* beinhaltet. Wir erweitern das Hamilton System, welches wir im zweiten Abschnitt betrachten möchten, wie in [[4], Remark 2.4.2] mit einem Parameter λ, um so den Fall der Nichtnegativität auf den Fall der Positivität zurückführen zu können. Diese Vorgehensweise unterscheidet sich grundsätzlich von der in [5]. Des Weiteren studieren wir besondere matrixwertige Lösungen des Hamilton Systems, die so genannten *konjugierten Basen* und ihre *fokalen Punkte*.

Im dritten Abschnitt werden wir feststellen, dass die Steuerbarkeit, die wir im Gegensatz zu [5] vereinfachend annehmen, äquivalent ist mit der Aussage, dass die fokalen Punkte jeder konjugierten Basis des Hamilton Systems isoliert sind.

Im Gegensatz zur in [[4], Remark 2.4.2] vorgeschlagenen Vorgehensweise, möchten wir jedoch nicht auf den ersten Oszillationssatz [[4], Theorem 4.2.3] zurückgreifen, sondern einen elementareren Einstieg wählen. Uns genügt nur ein Teil des Grenzwertsatzes [[4], Theorem 3.3.7], der in den Beweis des Oszillationssatzes eingeht, um die Nichtnegativität zu zeigen. Diesen betrachten wir im vierten Abschnitt.

Eine Erweiterung der Picone-Identität, die wir im fünften Abschnitt ansehen werden, liefert eine Abschätzung des betrachteten Funktionales, mit dessen Hilfe wir die hinreichenden Bedingungen sowohl für die Positivität als auch für die Nichtnegativität nachweisen werden.

In dieser Arbeit benutzen wir einige mathematische Grundlagen, wie etwa den Existenz- und Eindeutigkeitssatz für Systeme stückweise stetiger linearer Differentialgleichungen. Um das Wesentliche nicht aus den Augen zu verlieren, sind diese Hilfsmittel im Anhang aufgelistet und wir werden an den jeweiligen Stellen, an denen wir sie benötigen, auf diese verweisen.

Abschließend möchte mich ganz besonders bei Herrn Prof. Dr. Werner Kratz für die exzellente Betreuung dieser Arbeit bedanken, sowie bei Herrn Prof. Dr. Ulrich Stadtmüller, der sich freundlicherweise als Gutachter für diese Diplomarbeit zur Verfügung gestellt hat. Mein Dank gilt auch meiner Familie für ihren Zuspruch und die finanzielle Unterstützung, ohne die mein Studium nicht möglich gewesen wäre.

Ulm, im September 2003 Markus Wahrheit

1 Einführung

In dieser Arbeit werden wir stets von den folgenden Annahmen ausgehen:

$$\begin{array}{l} \mathcal{I} = [a, b] \text{ ist ein kompaktes nicht-entartetes Intervall,} \\ A(t), B(t), C(t) \text{ sind stückweise stetige reelle } n \times n\text{-matrixwertige} \\ \text{Funktionen auf } \mathcal{I}\text{, wobei } B(t) \text{ und } C(t) \text{ symmetrisch auf } \mathcal{I} \text{ sind,} \\ \text{d.h. } B(t) \equiv B^T(t), C(t) \equiv C^T(t) \text{ für alle } t \in \mathcal{I}. \end{array} \tag{1}$$

Ferner seien

$$\begin{array}{l} R_a, R_b, S_a, S_b \in \mathbb{R}^{n \times n} \text{ und} \\ S_a, S_b \text{ symmetrisch.} \end{array} \tag{2}$$

Wir betrachten das quadratische Funktional

$$\begin{array}{l} \mathcal{F}(x) = \mathcal{F}_0(x) + x^T(b) S_b x(b) - x^T(a) S_a x(a) \text{ mit} \\ \mathcal{F}_0(x) = \int_a^b \{x^T C x + u^T B u\}(t)\, dt \end{array} \tag{3}$$

für ein *zulässiges* x, d.h. x ist eine absolut stetige reelle vektorwertige Funktion (siehe Definition A.1.), die den so genannten „Bewegungsgleichungen“

$$\dot{x} = Ax + Bu \text{ auf } \mathcal{I} \text{ für eine reelle vektorwertige „Kontrollfunktion“ } u \in \mathcal{C}_s(\mathcal{I})$$

und den (separierten) Randwertbedingungen

$$x(a) \in \mathrm{Im} R_a \text{ und } x(b) \in \mathrm{Im} R_b$$

genügt.

Da wir die matrixwertigen Funktionen $A(t), B(t)$, sowie die vektorwertige Funktion $u(t)$ nicht zwingend stetig, sondern nur stückweise stetig voraussetzen, kann es vorkommen, dass die Funktion $\{Ax + Bu\}(t)$ an endlich vielen Stellen Sprünge aufweist, so dass die Ableitung von $x(t)$ an diesen Stellen nicht existiert. Wir lesen dann $\dot{x}(t)$ als Ableitung von $x(t)$ außer in endlich vielen t, in denen $\dot{x}(t)$ nicht definiert ist und sprechen davon, dass eine reelle Funktion $x(t)$ die Differentialgleichung $\dot{x} = Ax + Bu$ löst, falls $x(t)$ absolut stetig auf $\mathcal{I}$ ist, so dass gemäß Satz A.2.

$$x(t) - x(\tau) = \int_\tau^t \{A(\xi)x(\xi) + B(\xi)u(\xi)\}\, d\xi.$$

Definition 1.1.
Das in (3) definierte Funktional $\mathcal{F}$ heißt *positiv definit* (bzw. *nichtnegativ definit*), falls $\mathcal{F}(x) > 0$ (bzw. $\mathcal{F}(x) \geq 0$) für alle zulässigen x mit $x(t) \not\equiv 0$ auf $\mathcal{I}$.

Die zentrale Frage, der wir in dieser Arbeit nachgehen werden, ist also, unter welchen Bedingungen das Funktional (3) positiv oder nichtnegativ definit ist.

Bemerkung 1.2.

(i) Die Variationsrechnung beschäftigt sich mit der Aufgabe, auf einem abgeschlossenem nicht-entarteten Intervall eine oder mehrere Funktionen so zu bestimmen, dass ein gegebenes von der Wahl der Funktionen abhängiges bestimmtes Integral seinen größten oder kleinsten Wert annimmt. Ein besonderes Interesse gilt dabei den Randwertbetrachtungen, d.h. der Suche nach solchen Funktionen, die an einem oder an den beiden Rändern des Intervalls durch gegebene Punkte verlaufen. Sind die Randpunkte aller Funktionen fest gegeben, so ist man an Variationen mit $x(a) = x(b) = 0$, den so genannten *Dirichlet Nebenbedingungen*, interessiert. Vergleiche hierzu [[4], Theorem 8.2.1]. Diese Bedingung erreichen wir durch das Setzen von $R_a = R_b = 0$. Dieser Spezialfall wird ausführlich in [[6], Chapter VII, Theorem 4.5] diskutiert. Werden nur einige der Variablen des Vektors x in den Randpunkten festgehalten, so lässt sich dies durch eine entsprechende Wahl der Matrizen R_a, bzw. R_b, realisieren, siehe [[8], Section 2.8]. Durch Setzen von $R_a = I$, bzw. $R_b = I$, erreicht man, dass der Randpunkt $x(a)$, bzw. $x(b)$, 'frei' ist.

(ii) Die triviale Lösung $(x \equiv 0, u \equiv 0)$ ist stets zulässig. Ferner kann man jedoch unter der Annahme der Steuerbarkeit von (A, B), vergleiche Definition 3.1., mittels des Existenz- und Eindeutigkeitssatzes für Systeme von linearen Differentialgleichungen, Satz A.3., die Existenz eines zulässigen x zu jedem beliebig gewählten $x(a)$ und $x(b)$ gewährleisten. Siehe etwa [[4], Remark 3.5.1] und [[5], Lemma 1].

Eine erste notwendige Bedingung für die Nichtnegativität quadratischer Funktionale liefert der folgende Satz, der der wohlbekannten *Legendre Bedingung*, vergleiche [[4], Theorem 8.2.3], entspricht.

Satz 1.3.
Es gelte $\mathcal{F}(x) \geq 0$ für alle zulässigen x. Dann ist $B(t) \geq 0$ für alle $t \in \mathcal{I}$.

Beweis
Wir lehnen diesen Beweis an den Beweis von [[6], Chapter VII, Theorem 4.2] an und nehmen an, dass es ein konstantes $u_0 \in \mathbb{R}^n$ mit $\|u_0\| = 1$ und $u_0^T B(t) u_0 < 0$ (beachte, dass $B(t) \in \mathcal{C}_s(\mathcal{I})$) auf einem kompakten nicht-entarteten Teilintervall $\mathcal{I}_1 \subset \mathcal{I}$ gibt. Wir konstruieren nun ein Gegenbeispiel für ein zulässiges x mit $x(a) = x(b) = 0$.
Sei $\Phi(t)$ eine beliebige Fundamentalmatrix des Systems von Differentialgleichungen

$$\dot{x} = A(t)x,\ t \in \mathcal{I}_1.$$

Dann gibt es aufgrund der Annahmen (1) und der Stetigkeit von Fundamentalmatrizen Konstanten $k_1, k_2 > 0$ mit

$$\|B(t)\| \leq k_1,\ \|C(t)\| \leq k_1,\ \|\Phi(t)\Phi(s)^{-1}\| \leq k_1,\ u_0^T B(t) u_0 \leq -k_2 \text{ für alle } s, t \in \mathcal{I}_1.$$

Es sei $[c, d] \subset \mathcal{I}_1, c < d$, mit

$$(d - c)^2 < \frac{k_2}{k_1^5}.$$

Setze

$$u(t) := \begin{cases} u_0 p(t), & t \in (c,d) \\ 0, & \text{sonst}, \end{cases}$$

wobei $p(t) = \sum_{\nu=0}^{n} \alpha_\nu t^\nu \in \mathbb{R}$ ein Polynom vom Grade $\leq n$ ist.
Betrachte die Lösung x des Systems von linearen Differentialgleichungen

$$\dot{x} = A(t)x + B(t)u(t),\ t \in \mathcal{I}_1$$

zu dem Anfangswert $x(c) = 0$.
Nach der Variation der Konstanten, Lemma A.4., besitzt x die Gestalt

$$x(t) = \int_c^t \Phi(t)\Phi(s)^{-1}B(s)u(s)\,ds.$$

Demzufolge ist $x_\mu(t) = \sum_{\nu=0}^{n} \alpha_{\mu\nu}(t)\alpha_\nu$ für

$$\alpha_{\mu\nu}(t) = \int_c^t s^\nu(\Phi(t)\Phi(s)^{-1}B(s)u_0)_\mu\,ds,\ \mu = 1,\dots,n,$$

wobei $x_\mu(t)$ die μ-te Komponente des Vektors $x(t)$ bezeichnet.
Das lineare homogene Gleichungssystem

$$x_\mu(d) = 0 \text{ für } \mu = 1,\dots,n,$$

bestehend aus n Gleichungen mit $n+1$ Unbekannten $\alpha_0,\dots,\alpha_n$, besitzt eine nichttriviale Lösung $\alpha_0,\dots,\alpha_n$ mit $\sum_{\nu=0}^{n} |\alpha_\nu|^2 = 1$.
Sei im Folgenden also $\alpha_0,\dots,\alpha_n$ so gewählt, dass $x(d) = 0$ und $\sum_{\nu=0}^{n} |\alpha_\nu|^2 = 1$ gilt. Dann ist $x(t) \not\equiv 0$ auf (c,d) und $x(t) \equiv 0$ auf $[a,c] \cup [d,b]$. Insbesondere ist $x(a) = x(b) = 0$.
Ferner gilt für $t \in [c,d]$:

$$\begin{aligned} \|x(t)\| &\leq \int_c^t \|\Phi(t)\Phi(s)^{-1}B(s)u_0 p(t)\|\,ds \\ &\leq k_1^2 \int_c^t |p(s)|\,ds \end{aligned}$$

und

$$\begin{aligned}
\mathcal{F}(x) &= \int_c^d \{x^T C x + u^T B u\}(t)\,dt \\
&\leq \int_c^d \|x(t)\|^2 \|C(t)\|\,dt + \int_c^d u_0^T B(t) u_0 |p(t)|^2\,dt \\
&\leq k_1^5 \int_c^d (\int_c^t |p(s)|\,ds)^2\,dt - k_2 \int_c^d |p(t)|^2\,dt \\
&\leq k_1^5 (d-c) (\int_c^d |p(t)|\,dt)^2 - k_2 \int_c^d |p(t)|^2\,dt.
\end{aligned}$$

Mit der Schwarzschen Ungleichung, vergleiche [[3], Seite 475], gilt nun

$$(\int_c^d |p(t)|\,dt)^2 \leq (d-c) \int_c^d |p(t)|^2\,dt.$$

Dies bedeutet

$$\mathcal{F}(x) \leq -\{k_2 - k_1^5 (d-c)^2\} \int_c^d |p(t)|^2 dt < 0$$

und damit ein Widerspruch zu unserer Voraussetzung. □

Im weiteren Verlauf werden wir bei der Suche nach weiteren notwendigen Bedingungen für die Definitheit quadratischer Funktionale gelegentlich auf diese Bedingung zurückgreifen.

2 Hamilton System und konjugierte Basen

Wir legen das folgende *Hamilton System* zu Grunde, das seinen Ursprung in der Variationsrechnung findet. Neben den Bewegungsgleichungen, die wir mit unserem quadratischen Funktional $\mathcal{F}$ verbinden, sind die so genannten *Eulergleichungen* $\dot{u} = Cx - A^T u$ eine notwendige Bedingung für die schwache Minimalität eines Extremals. Vergleiche hierzu [[4], Chapter 1.3, 2.3 und 8.2], sowie [[1], Kapitel 25].
Zunächst studieren wir die matrixwertigen Lösungen, insbesondere die *konjugierten Basen*, des Hamilton Systems. Es wird sich zeigen, dass wir die Definitheit des quadratischen Funktionales $\mathcal{F}$ mittels einer speziellen konjugierten Basis charakterisieren können. Dabei wandeln das Hamilton System in der Definition 2.16. dahingehend ab, dass wir den Parameter λ hinzufügen. Wie wir im Lemma 2.15. sehen werden, können wir damit das Problem der Nichtnegativität des Funktionales $\mathcal{F}$ auf den wohlbekannten Fall der Positivität, siehe [[4], Theorem 2.4.2], zurückführen, indem wir $C(t)$ durch die matrixwertige Funktion $C(t,\lambda) = C(t) - \lambda I$ für $t \in \mathcal{I}$ ersetzen. Dies wird unser Ansatz zur Ermittlung der notwendigen Bedingungen für die Nichtnegativität des Funktionales $\mathcal{F}$ sein.

Definition 2.1.
Das System von linearen Differentialgleichungen

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ \dot{u} &= Cx - A^T u \end{aligned} \tag{H}$$

mit matrixwertigen Funktionen $A(t), B(t), C(t), t \in \mathcal{I}$, die die Annahmen (1) erfüllen, nennt man *Hamilton System*.

Bemerkung 2.2.

(i) Ähnlich wie bei der Definition eines zulässigen x verstehen wir als Lösung des Hamilton Systems das Paar zweier absolut stetiger Funktionen (x, u), so dass gemäß Satz A.2.

$$\begin{aligned} x(t) - x(\tau) &= \int_\tau^t \{A(\xi)x(\xi) + B(\xi)u(\xi)\}\, d\xi, \\ u(t) - u(\tau) &= \int_\tau^t \{C(\xi)x(\xi) - A(\xi)^T u(\xi)\}\, d\xi. \end{aligned}$$

(ii) Das Hamilton System ist äquivalent zu dem System von linearen Differentialgleichungen

$$\begin{pmatrix} \dot{x} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} A & B \\ C & -A^T \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}.$$

Daher ist das Hamilton System zu den gegebenen Anfangswerten $x(t_0) = x_0, u(t_0) = u_0$ für ein $t_0 \in \mathcal{I}$ eindeutig lösbar nach dem

Existenz- und Eindeutigkeitssatz für Systeme linearer Differentialgleichungen, Satz A.3.

(iii) Die Nützlichkeit einer Lösung des Hamilton Systems lässt sich leicht erkennen: Ist nämlich (x, u) eine Lösung von (H), so folgern wir mittels partieller Integration

$$\begin{aligned}\mathcal{F}_0(x) &= \int_a^b \{x^T C x + u^T B u\}(t)dt \\ &= \int_a^b \{x^T C x + (\dot{x} - Ax)^T u\}(t)dt \\ &= \int_a^b \{x^T (Cx - A^T u - \dot{u})\}(t)dt + \{x^T u\}(t)|_a^b \\ &= \{x^T u\}(t)|_a^b.\end{aligned}$$

Definition 2.3.
Wir bezeichnen (X, U) als $(n \times m)$-*matrixwertige Lösung des Hamilton Systems* (kurz: matrixwertige Lösung von (H)), falls $X(t)$ und $U(t)$ absolut stetige reelle $n \times m$-matrixwertige Funktionen auf $\mathcal{I}$ sind, die das System von Differentialgleichungen $\dot{X} = AX + BU$, $\dot{U} = CX - A^T U$ lösen.

Proposition 2.4.
Für zwei matrixwertige Lösungen $(X_1, U_1), (X_2, U_2)$ von (H) ist die *Wronski-Matrix*

$$\{X_1^T U_2 - U_1^T X_2\}(t)$$

eine konstante Matrix auf $\mathcal{I}$.

Beweis
Wir verifizieren die Behauptung mittels Differentiation, wobei wir der Übersichtlichkeit halber auf die Angabe der Variablen t verzichten wollen:

$$\begin{aligned}\frac{d}{dt}\{X_1^T U_2 - U_1^T X_2\} &= \dot{X}_1^T U_2 + X_1^T \dot{U}_2 - \dot{U}_1^T X_2 - U_1^T \dot{X}_2 \\ &= (AX_1 + BU_1)^T U_2 + X_1^T (CX_2 - A^T U_2) - \\ &\quad\ (CX_1 - A^T U_1)^T X_2 - U_1^T (AX_2 + BU_2) \\ &= 0.\end{aligned}$$

□

Korollar 2.5.
Für matrixwertige Lösungen (X, U) von (H) gilt: $\{X^T U - U^T X\}(t)$ ist konstant.

Beweis
Die Behauptung folgt unmittelbar aus der Proposition 2.4. durch Setzen von $X_1 = X_2 = X, U_1 = U_2 = U$. □

Proposition 2.6.

(i) Für eine matrixwertige Lösung (X,U) von (H) gilt

$\{X^TU - U^TX\}(t_0) = 0$ für ein $t_0 \in \mathcal{I} \Leftrightarrow \{X^TU - U^TX\}(t) \equiv 0$ für alle $t \in \mathcal{I}$.

(ii) Es ist $\operatorname{rg}\begin{pmatrix} X \\ U \end{pmatrix}(t_0) = r$ für ein $t_0 \in \mathcal{I}$ und ein $r \in \mathbb{N}$ genau dann, wenn $\operatorname{rg}\begin{pmatrix} X \\ U \end{pmatrix}(t) \equiv r$ für alle $t \in \mathcal{I}$ ist.

Beweis

(i) Die Behauptung folgt aus Korollar 2.5.

(ii) Es sei $\Phi(t) \in \mathbb{R}^{2n\times 2n}$ ein Fundamentalsystem von (H). Dann ist jede matrixwertige Lösung (X,U) von (H) von der Form $\begin{pmatrix} X \\ U \end{pmatrix}(t) = \Phi(t)M$ für eine konstante Matrix $M \in \mathbb{R}^{2n\times m}$. Da nun $\operatorname{rg}\begin{pmatrix} X \\ U \end{pmatrix}(t_0) = r$ folgt $\operatorname{rg} M = r$ und damit $\operatorname{rg}\begin{pmatrix} X \\ U \end{pmatrix}(t) \equiv r$ auf $\mathcal{I}$. Die umgekehrte Richtung ist trivial.

□

Wir werden im weiteren Verlauf eine $n \times n$-matrixwertige Lösung (X,U) von (H) mit maximalem Rang so wählen, dass die resultierende Wronski-Matrix die Nullmatrix ist. Die hiermit erzielte Symmetrie von X^TU wird sich als hilfreich erweisen.

Definition 2.7.

(i) (X,U) heißt eine *konjugierte Basis von* (H), falls (X,U) eine $n \times n$-matrixwertige Lösung von (H) ist mit

$$\operatorname{rg}\begin{pmatrix} X(t) \\ U(t) \end{pmatrix} \equiv n \text{ und } \{X^TU - U^TX\}(t) \equiv 0 \text{ auf } \mathcal{I}.$$

(ii) Zwei konjugierte Basen $(X_1,U_1), (X_2,U_2)$ von (H) heißen *normalisierte konjugierte Basen von* (H), falls $\{X_1^TU_2 - U_1^TX_2\}(t) \equiv I$ auf $\mathcal{I}$ gilt.

(iii) $t_0 \in \mathcal{I}$ heißt *fokaler Punkt von* X für eine konjugierte Basis (X,U) von (H), falls $X(t_0)$ singulär ist.

Bemerkung 2.8.
Da eine konjugierte Basis (X,U) von (H) eine Lösung des Hamilton Systems mit maximalem Rang ist, kann man (X,U) als „halbe Fundamentalmatrizen“ ansehen.

Um die Existenz konjugierter Basen zu zeigen, benötigen wir zunächst das

Lemma 2.9.
Es sei $K \in \mathbb{R}^{n\times n}$. Dann gibt es ein $M \in \mathbb{R}^{n\times n}$ mit $K^T M = 0, \mathrm{rg}(K^T, M^T) = n$.

Beweis
Im Fall $K = 0$ erfüllt $M = I$ unsere Behauptung.
Sei nun $r := \mathrm{rg}K > 0$, so gibt es mit den elementaren Zeilenoperationen ein reguläres $T \in \mathbb{R}^{n\times n}$, so dass $K^T = T \begin{pmatrix} K_1 \\ 0 \end{pmatrix}$ mit $K_1 \in \mathbb{R}^{r\times n}$ und $\mathrm{rg}K_1 = r$.
Sei $m_1, \ldots, m_{n-r}$ eine reelle Basis des Kerns von $\begin{pmatrix} K_1 \\ 0 \end{pmatrix}$. Wir definieren $M := \begin{pmatrix} 0 & \widetilde{M} \end{pmatrix} T^T$ mit $\widetilde{M} = (m_1, \ldots, m_{n-r})$.
Damit ist $K^T M = 0$ und es gilt

$$\mathrm{rg}(K^T, M^T) = \mathrm{rg}T \begin{pmatrix} K_1 & 0 \\ 0 & \widetilde{M}^T \end{pmatrix} = \mathrm{rg} \begin{pmatrix} K_1 & 0 \\ 0 & \widetilde{M}^T \end{pmatrix}.$$

Da $\mathrm{rg}K_1 = r$ und $\mathrm{rg}\widetilde{M} = n - r$ ist, ergibt sich die Behauptung. □

Proposition 2.10.
Zu jedem gegebenen $t_0 \in \mathcal{I}$ und $X_0 \in \mathbb{R}^{n\times n}$ gibt es eine (nicht notwendigerweise eindeutige) konjugierte Basis (X, U) von (H) mit $X(t_0) = X_0$.

Beweis
Nach dem obigen Lemma gibt es ein $U_0 \in \mathbb{R}^{n\times n}$ mit $X_0^T U_0 = 0$ und $\mathrm{rg}(X_0^T, U_0^T) = n$. Betrachte nun die $n \times n$-matrixwertige Lösung (X, U) von (H) zu den Anfangswerten $X(t_0) = X_0, U(t_0) = U_0$, die es nach dem Existenz- und Eindeutigkeitssatz für Systeme linearer Differentialgleichungen, Satz A.3., gibt. Dann ist (X, U) nach Proposition 2.6. eine konjugierte Basis von (H). □

Im Folgenden werden wir zeigen, dass sich zu jeder konjugierten Basis (X_1, U_1) von (H) eine konjugierte Basis (X_2, U_2) von (H) finden lässt, so dass $(X_1, U_1), (X_2, U_2)$ normalisierte konjugierte Basen von (H) bilden. Überdies können wir sogar die Regularität von $X_2(t_0)$ in einem vorgegebenen Punkt $t_0 \in \mathcal{I}$ sichern.

Proposition 2.11.
Es seien $(X_1, U_1), (X_2, U_2)$ normalisierte konjugierte Basen von (H). Dann gelten die folgenden Gleichungen für alle $t \in \mathcal{I}$, wobei wir auf die Angabe der Variablen t verzichten.

$$X_\nu^T U_\nu = U_\nu^T X_\nu \text{ für } \nu = 1, 2, \quad X_1 X_2^T = X_2 X_1^T, \quad U_1 U_2^T = U_2 U_1^T$$
$$X_1^T U_2 - U_1^T X_2 = X_1 U_2^T - X_2 U_1^T = I.$$

Beweis
Mit der Definition 2.7. rechnen wir nach, dass

$$\{\begin{pmatrix} U_2^T & -X_2^T \\ -U_1^T & X_1^T \end{pmatrix}\begin{pmatrix} X_1 & X_2 \\ U_1 & U_2 \end{pmatrix}\}(t) = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

ist, und da eine Matrix mit ihrer Inversen kommutiert, folgt die Behauptung. □

Satz 2.12.
Gegeben seien reelle $n \times n$-Matrizen Q_1, Q_2 so, dass $Q_1 Q_2^T$ symmetrisch ist. Ferner sei $\text{rg}(Q_1, Q_2) = n$.
Dann gilt:

(i) $\text{rg}(Q_1^T, Q_2^T) = n$ und $\ker Q_1 Q_2^T = \ker Q_1^T \oplus \ker Q_2^T$.

(ii) $Q_1 + t_0 Q_2$ ist regulär, falls $t_0 > 0$ und

$$t_0 \|Q_2^T\| < \min\{|\mu| : \mu \text{ ist negativer Eigenwert von } Q_1 Q_2^T\} := c_0.$$

Beweis

(i) Die Matrix $D := (Q_1, Q_2)(Q_1, Q_2)^T = Q_1 Q_1^T + Q_2 Q_2^T \in \mathbb{R}^{n \times n}$ ist positiv definit aufgrund der Rangbedingung $\text{rg}(Q_1, Q_2) = n$.
Wir zeigen (i) für die Matrizen $\widetilde{Q}_\nu = D^{-\frac{1}{2}} Q_\nu$, $\nu = 1, 2$. Mit der Regularität von $D^{-\frac{1}{2}}$ ergibt sich dann die Behauptung.
Mit der Symmetrie von $Q_1 Q_2^T$ ist

$$\begin{pmatrix} \widetilde{Q}_1 & \widetilde{Q}_2 \\ -\widetilde{Q}_2 & \widetilde{Q}_1 \end{pmatrix}\begin{pmatrix} \widetilde{Q}_1^T & -\widetilde{Q}_2^T \\ \widetilde{Q}_2^T & \widetilde{Q}_1^T \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

und da eine Matrix mit ihrer Inversen kommutiert, folgt $\widetilde{Q}_1^T \widetilde{Q}_1 + \widetilde{Q}_2^T \widetilde{Q}_2 = I$ und damit $\text{rg}(Q_1^T, Q_2^T) = n$, sowie die Symmetrie von $\widetilde{Q}_1^T \widetilde{Q}_2$.
Gelte $\widetilde{Q}_1 \widetilde{Q}_2^T d = 0$ für ein $d \in \mathbb{R}^n$. Dann ist

$$\widetilde{Q}_1^T d = \widetilde{Q}_1^T (\widetilde{Q}_1 \widetilde{Q}_1^T + \widetilde{Q}_2 \widetilde{Q}_2^T) d = \widetilde{Q}_1^T \widetilde{Q}_1 \widetilde{Q}_1^T d,$$

also

$$d - \widetilde{Q}_1 \widetilde{Q}_1^T d \in \ker \widetilde{Q}_1^T$$

und

$$\widetilde{Q}_2^T \widetilde{Q}_1 \widetilde{Q}_1^T d = \widetilde{Q}_1^T \widetilde{Q}_2 \widetilde{Q}_1^T d = \widetilde{Q}_1^T \widetilde{Q}_1 \widetilde{Q}_2^T d = 0,$$

also

$$\widetilde{Q}_1 \widetilde{Q}_1^T d \in \ker \widetilde{Q}_2^T.$$

Demzufolge ist $\ker(\widetilde{Q}_1 \widetilde{Q}_2^T) \subset \ker \widetilde{Q}_1^T + \ker \widetilde{Q}_2^T$. Die umgekehrte Inklusion folgt aufgrund der Symmetrie von $\widetilde{Q}_1 \widetilde{Q}_2^T$, denn sei $d = d_1 + d_2$ mit $d_1 \in \ker \widetilde{Q}_1^T$ und $d_2 \in \ker \widetilde{Q}_2^T$, so gilt $\widetilde{Q}_1 \widetilde{Q}_2^T d = \widetilde{Q}_2 \widetilde{Q}_1^T d_1 + \widetilde{Q}_1 \widetilde{Q}_2^T d_2 = 0$. Diese Summe ist direkt, da $\text{rg}(\widetilde{Q}_1, \widetilde{Q}_2)^T = n$.

(ii) Gelte nun nach (i) $\mathbb{R}^n = P \oplus N \oplus \ker Q_1^T \oplus \ker Q_2^T$, mit $Q_1 Q_2^T > 0$ auf dem Untervektorraum P und < 0 auf dem Untervektorraum N.
Es sei $Q := Q_2(Q_1^T + t_0 Q_2^T)$. Dann gilt $Q = 0$ auf $\ker Q_2^T$, da mit $Q_1 Q_2^T$ ebenfalls Q symmetrisch ist, und $Q > 0$ auf $P \oplus \ker Q_1^T$ nach (i).
Ferner gilt $Q < 0$ auf N, denn es gilt

$$d^T Q d \leq -c_0 \|d\|^2 + |t_0| \|Q_2^T\| \|d\|^2 < 0 \text{ für alle } d \in N, d \neq 0.$$

Also ist $\ker Q = \ker Q_2^T$.
Gelte nun $(Q_1 + t_0 Q_2)^T d = 0$, so ist $Qd = 0$ und demzufolge $Q_2^T d = 0 = Q_1^T d$.
Mit (i) folgt nun $d = 0$ und damit ist $Q_1 + t_0 Q_2$ regulär.

□

Proposition 2.13.
Es sei (X_1, U_1) eine konjugierte Basis von (H) und $t_0 \in \mathcal{I}$. Dann existiert eine konjugierte Basis (X_2, U_2) des Hamilton Systems so, dass $(X_1, U_1), (X_2, U_2)$ normalisierte konjugierte Basen von (H) sind und dass $X_2(t_0)$ regulär ist.

Beweis
Die Matrix $K(t) := \{X_1^T X_1 + U_1^T U_1\}(t)$ ist regulär aufgrund von $\operatorname{rg}(X_1^T(t), U_1^T(t)) = n$ nach der Definition einer konjugierten Basis. Definiere für $\eta \in \mathbb{R}$

$$\begin{aligned} \widetilde{X}(t,\eta) &:= -U_1(t)K^{-1}(t) + \eta X_1(t) \\ \widetilde{U}(t,\eta) &:= X_1(t)K^{-1}(t) + \eta U_1(t). \end{aligned}$$

Dann gilt für alle $\eta \in \mathbb{R}$

$$\begin{pmatrix} \widetilde{U}^T(t,\eta) & -\widetilde{X}^T(t,\eta) \\ -U_1^T(t) & X_1^T(t) \end{pmatrix} \begin{pmatrix} X_1(t) & \widetilde{X}(t,\eta) \\ U_1(t) & \widetilde{U}(t,\eta) \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \qquad (*)$$

Da also die obigen Matrizen regulär sind, gilt $\operatorname{rg}(X_1(t), \widetilde{X}(t,\eta)) = n$ für jedes $\eta \in \mathbb{R}$. Zudem folgt die Symmetrie von $\widetilde{X}(t,\eta) X_1^T(t)$, da eine Matrix mit ihrer Inversen kommutiert.
Wir wenden nun den Satz 2.12.(ii) mit $Q_1(t) = \widetilde{X}(t,0)$ und $Q_2(t) = X_1(t)$ an und können folgern, dass $\widetilde{X}(t,\eta_0) = \widetilde{X}(t,0) + \eta_0 X_1(t)$ regulär für hinreichend kleines $\eta_0 > 0$ ist.
Sei nun (X_2, U_2) eine $n \times n$-matrixwertige Lösung von (H) zu den Anfangswerten

$$\begin{aligned} X_2(t_0) &:= \widetilde{X}(t_0,\eta_0) = -U_1(t_0)K^{-1}(t_0) + \eta_0 X_1(t_0) \\ U_2(t_0) &:= \widetilde{U}(t_0,\eta_0) = X_1(t_0)K^{-1}(t_0) + \eta_0 U_1(t_0). \end{aligned}$$

Mit (*) folgt nun $\{X_1^T U_2 - U_1^T X_2\}(t_0) = I$ und demzufolge mit Proposition 2.4.

$$\{\begin{pmatrix} X_1 \\ U_1 \end{pmatrix}^T \begin{pmatrix} U_2 \\ -X_2 \end{pmatrix}\}(t) = \{X_1^T U_2 - U_1^T X_2\}(t) \equiv I \text{ auf } \mathcal{I}$$

und $\operatorname{rg}\begin{pmatrix} X_2 \\ U_2 \end{pmatrix}(t) \equiv n$ auf $\mathcal{I}$.

Ferner gilt nach (*) $\{X_2^T U_2 - U_2^T X_2\}(t_0) = 0$ und mit Proposition 2.6.(i) $\{X_2^T U_2 - U_2^T X_2\}(t) \equiv 0$ auf $\mathcal{I}$.
Damit haben wir gezeigt, dass $(X_1, U_1), (X_2, U_2)$ normalisierte konjugierte Basen von (H) sind, für die $X_2(t_0)$ regulär ist. □

Aus der nachstehenden Definition und dem folgenden Lemma ist ersichtlich, dass wir unter Zuhilfenahme des Parameters λ den Fall der Nichtnegativität auf den Fall der Positivität zurückführen können.

Definition 2.14.
Für ein zulässiges x und $\lambda \in \mathbb{R}$ definieren wir

$$\mathcal{F}(x, \lambda) := \mathcal{F}(x) - \lambda \int_a^b \|x(t)\|^2 \, dt.$$

Lemma 2.15.
Es sei $x \not\equiv 0$ auf $\mathcal{I}$ und zulässig. Dann ist $\mathcal{F}(x) \geq 0$ genau dann, wenn $\mathcal{F}(x, \lambda) > 0$ für alle $\lambda < 0$.

Beweis
Mit $-\lambda \int_a^b \|x(t)\|^2 \, dt > 0$ für alle $\lambda < 0$ und der Stetigkeit in λ folgt unmittelbar die Behauptung. □

Da $\mathcal{F}(x, \lambda) = \int_a^b \{x^T(C - \lambda I)x - u^T B u\}(t) + x^T(b) S_b x(b) - x^T(a) S_a x(a)$ können wir also die Nichtnegativität auf die Positivität zurückführen, indem wir statt der matrixwertigen Funktion $C(t)$ die Funktion $C(t, \lambda) := C(t) - \lambda I$ für $\lambda < 0$ betrachten. Da auch die Funktion $C(t) - \lambda I$ für alle λ stückweise stetig und symmetrisch ist, bleiben die bisherigen Betrachtungen für $C(t, \lambda)$ für alle festen $\lambda \in \mathbb{R}$ erhalten. Im weiteren Verlauf studieren wir das Verhalten des Parameters λ auf entsprechende konjugierte Basen.

Definition 2.16.
Zu einem gegebenen $\lambda \in \mathbb{R}$ bezeichnen wir das System von linearen Differentialgleichungen

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ \dot{u} &= (C - \lambda I)x - A^T u \end{aligned} \tag{H_λ}$$

mit matrixwertigen Funktionen $A(t), B(t), C(t) - \lambda I, t \in \mathcal{I}$, die die Annahmen (1) erfüllen, als *Hamilton System* (H_λ).

Lemma 2.17.
Es sei (X, U) eine konjugierte Basis von (H_λ). Ferner seien $X(a,\lambda)$ und $U(a,\lambda)$ konstant für $\lambda \in \mathbb{R}$. Dann sind $X(t,\lambda)$ und $U(t,\lambda)$ differenzierbar (sogar analytisch) in λ bei festem $t \in \mathcal{I}$.

Beweis
Siehe [[6], Chapter I, Section 10, Theorem 10.1]. □

Definition 2.18.
Eine symmetrische und reelle $n \times n$-matrixwertige Funktion $M(t), t \in \mathcal{I}$, heißt monoton wachsend (bzw. monoton fallend) auf $\mathcal{I}$, falls ihre quadratische Form $c^T M(t) c$ für alle $c \in \mathbb{R}^n$ monoton wachsend (bzw. monoton fallend) in $\mathcal{I}$ ist. Wächst (bzw. fällt) ihre quadratische Form streng für alle $c \neq 0$, so sagt man, $M(t)$ ist streng monoton wachsend (bzw. streng monoton fallend) auf $\mathcal{I}$.

Proposition 2.19.
Es seien $(X_1, U_1), (X_2, U_2)$ normalisierte konjugierte Basen von (H_λ) mit Anfangsbedingungen, die für den Randwert a in λ fixiert sind, d.h. es gilt $\frac{\partial}{\partial\lambda} X_\nu(a,\lambda) \equiv \frac{\partial}{\partial\lambda} U_\nu(a,\lambda) \equiv 0, \lambda \in \mathbb{R}, \nu = 1, 2$.
Dann sind für jedes feste $t \in \mathcal{I}$ die matrixwertigen Funktionen

(i) $-\{X_1^{-1} X_2\}(t,\lambda)$ und

(ii) $\{U_1 X_1^{-1}\}(t,\lambda)$

monoton fallend in λ auf den Intervallen, auf denen $X_1(t,\lambda)$ regulär ist.

Beweis
Betrachte die Matrizen

$$X_*(t,\lambda) := \begin{pmatrix} 0 & I \\ X_1 & X_2 \end{pmatrix}(t,\lambda),\ U_*(t,\lambda) := \begin{pmatrix} I & 0 \\ U_1 & U_2 \end{pmatrix}(t,\lambda).$$

Wir zeigen, dass die matrixwertige Funktion

$$\begin{aligned} M(t,\lambda) &:= \{U_* X_*^{-1}\}(t,\lambda) = \{\begin{pmatrix} I & 0 \\ U_1 & U_2 \end{pmatrix} \begin{pmatrix} -X_1^{-1}X_2 & X_1^{-1} \\ I & 0 \end{pmatrix}\}(t,\lambda) \\ &= \begin{pmatrix} -X_1^{-1}X_2 & X_1^{-1} \\ (X_1^{-1})^T & U_1 X_1^{-1} \end{pmatrix}(t,\lambda) \end{aligned}$$

für jedes feste $t \in \mathcal{I}$ monoton fallend in λ ist.
Dann folgt mit

$$\begin{aligned} \{-X_1^{-1}X_2\}(t,\lambda) &= \begin{pmatrix} I_{n\times n} & 0 \end{pmatrix} M(t,\lambda) \begin{pmatrix} I_{n\times n} \\ 0 \end{pmatrix} \quad \text{und} \\ \{U_1 X_1^{-1}\}(t,\lambda) &= \begin{pmatrix} 0 & I_{n\times n} \end{pmatrix} M(t,\lambda) \begin{pmatrix} 0 \\ I_{n\times n} \end{pmatrix} \end{aligned}$$

die Behauptung.

Zunächst beobachten wir, dass $\begin{pmatrix} X_* \\ U_* \end{pmatrix}$ das Hamilton System

$$\frac{\partial}{\partial t}\begin{pmatrix} X_* \\ U_* \end{pmatrix} = (\mathcal{H} - \lambda\mathcal{H}_0)\begin{pmatrix} X_* \\ U_* \end{pmatrix}$$

mit

$$\mathcal{H} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A & 0 & B \\ 0 & 0 & 0 & 0 \\ 0 & C & 0 & -A^T \end{pmatrix} \text{ und } \mathcal{H}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 \end{pmatrix}$$

löst, wobei wir mit 0 die $n \times n$-dimensionale Nullmatrix bezeichnen. Ferner folgt mit Proposition 2.11. für alle $\lambda \in \mathbb{R}$ und alle $t \in \mathcal{I}$, dass $\{X_*^T U_* - U_*^T X_*\}(t,\lambda) \equiv 0$, d.h. es gilt $\{X_*^T U_*\}(t,\lambda) = \{U_*^T X_*\}(t,\lambda)$.

Mit $\mathcal{J} := \begin{pmatrix} 0 & I_{2n\times 2n} \\ -I_{2n\times 2n} & 0 \end{pmatrix} = -\mathcal{J}^T$ und Lemma 2.17. evaluieren wir die Gleichung

$$\begin{aligned} & \tfrac{\partial}{\partial t}\{X_*^T \tfrac{\partial}{\partial\lambda} U_* - U_*^T \tfrac{\partial}{\partial\lambda} X_*\} = \tfrac{\partial}{\partial t}\{\begin{pmatrix} X_* \\ U_* \end{pmatrix}^T \mathcal{J} \tfrac{\partial}{\partial\lambda}\begin{pmatrix} X_* \\ U_* \end{pmatrix}\} \\ = & \begin{pmatrix} X_* \\ U_* \end{pmatrix}^T (\mathcal{H} - \lambda\mathcal{H}_0)^T \mathcal{J} \tfrac{\partial}{\partial\lambda}\begin{pmatrix} X_* \\ U_* \end{pmatrix} + \begin{pmatrix} X_* \\ U_* \end{pmatrix}^T \mathcal{J}\left[(\mathcal{H} - \lambda\mathcal{H}_0)\tfrac{\partial}{\partial\lambda}\begin{pmatrix} X* \\ U_* \end{pmatrix} - \mathcal{H}_0\begin{pmatrix} X_* \\ U_* \end{pmatrix}\right] \\ = & -\begin{pmatrix} X_* \\ U_* \end{pmatrix}^T \mathcal{J}\mathcal{H}_0\begin{pmatrix} X_* \\ U_* \end{pmatrix} = \begin{pmatrix} X_* \\ U_* \end{pmatrix}^T \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -I_{n\times n} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} X_* \\ U_* \end{pmatrix} \le 0. \end{aligned}$$

Die dritte Gleichung folgt mittels Symmetrie von

$$\mathcal{J}(\mathcal{H} - \lambda\mathcal{H}_0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C - \lambda I & 0 & -A^T \\ 0 & 0 & 0 & 0 \\ 0 & -A & 0 & -B \end{pmatrix},$$

denn damit gilt:

$$\mathcal{J}(\mathcal{H} - \lambda\mathcal{H}_0) = (\mathcal{H} - \lambda\mathcal{H}_0)^T\mathcal{J}^T = -(\mathcal{H} - \lambda\mathcal{H}_0)^T\mathcal{J}.$$

Mit dem obigen Hilfsresultat können wir die Aussage bezüglich der Monotonie von $M(t,\lambda)$ in λ beweisen. Es gilt mit der Symmetrie von $X_*^T U_*$ und der elementaren Kenntnis aus der linearen Algebra $(X_*^{-1})^T = (X_*^T)^{-1}$:

$$\begin{aligned} \tfrac{\partial}{\partial\lambda} M(t,\lambda) &= (X_*^T)^{-1} X_*^T (\tfrac{\partial}{\partial\lambda} U_* X_*^{-1}) X_* X_*^{-1} \\ &= (X_*^T)^{-1} X_*^T ((\tfrac{\partial}{\partial\lambda} U_*) X_*^{-1} - U_* X_*^{-1} (\tfrac{\partial}{\partial\lambda} X_*) X_*^{-1}) X_* X_*^{-1} \\ &= (X_*^{-1})^T (X_*^T \tfrac{\partial}{\partial\lambda} U_* - U_*^T \tfrac{\partial}{\partial\lambda} X_*) X_*^{-1} \le 0. \end{aligned}$$

Der letzte Ausdruck ist nichtpositiv, da nach Voraussetzung $\frac{\partial}{\partial\lambda} X_*(a,\lambda) \equiv \frac{\partial}{\partial\lambda} U_*(a,\lambda) \equiv 0$ und damit $\{X_*^T \frac{\partial}{\partial\lambda} U_* - U_*^T \frac{\partial}{\partial\lambda} X_*\}(t,\lambda) \le 0$, denn dieser Ausdruck entspricht für $t = a$ der Nullmatrix und ist monoton fallend in $\mathcal{I}$. $\square$

3 Steuerbarkeit und fokale Punkte

In diesem Abschnitt möchten wir das Konzept der Steuerbarkeit vorstellen, welche wir im Unterschied zu [5] vereinfachend annehmen werden. Das zentrale Resultat dieses Abschnitts ist Satz 3.4., der die Äquivalenz der Steuerbarkeit mit der Isoliertheit fokaler Punkte des Hamilton Systems aufzeigt.

Definition 3.1.
Das Paar (A, B) heißt *steuerbar auf* $\mathcal{I}$, falls $\dot{v} = -A^T(t)v, B^T(t)v(t) \equiv 0$ auf einem nicht-entarteten Intervall $[t_1, t_2] \subset \mathcal{I}$ stets $v(t) \equiv 0$ auf $[t_1, t_2]$ impliziert.

Proposition 3.2.
Es seien $(X_1, U_1), (X_2, U_2)$ normalisierte konjugierte Basen des Hamilton Systems und $X_1(t)$ regulär auf einem nicht-entarteten Teilintervall $\mathcal{I}_1 \subset \mathcal{I}$. Ferner sei (A, B) steuerbar auf $\mathcal{I}_1$ und $B(t) \geq 0$ auf $\mathcal{I}_1$. Dann ist

$$X_1^{-1}(t)X_2(t) \text{ ist streng monoton wachsend auf } \mathcal{I}_1.$$

Beweis
Wir verifizieren die Behauptung durch Differentiation, wobei wir die Variable t weglassen, und mit Proposition 2.11., sowie Lemma A.5.

$$\begin{aligned}
\frac{d}{dt}X_1^{-1}X_2 &= -X_1^{-1}(AX_1 + BU_1)X_1^{-1}X_2 + X_1^{-1}(AX_2 + BU_2) \\
&= -X_1^{-1}AX_2 - X_1^{-1}BU_1X_1^{-1}X_2 + X_1^{-1}AX_2 + X_1^{-1}BU_2 \\
&= X_1^{-1}B(X_1^{-1})^T\{X_1^TU_2 - X_1^TU_1X_1^{-1}X_2\} \\
&= X_1^{-1}B(X_1^{-1})^T\{X_1^TU_2 - U_1^TX_2\} \\
&= X_1^{-1}B(X_1^{-1})^T \\
&\geq 0 \text{ für alle } t \in \mathcal{I}.
\end{aligned}$$

Damit ist $\{X_1^{-1}X_2\}(t)$ monoton wachsend auf $\mathcal{I}_1$. Um die strenge Monotonie zu zeigen, seien $t_1, t_2 \in \mathcal{I}_1$ mit $t_1 < t_2$ und $d \in \mathbb{R}^n$:

$$d^T\{X_1^{-1}(t_2)X_2(t_2) - X_1^{-1}(t_1)X_2(t_1)\}d = 0.$$

Dann gilt $X_1^{-1}(t)X_2(t)d = X_1^{-1}(t_1)X_2(t_1)d$ für alle $t \in [t_1, t_2]$.
Betrachte nun die Lösung des Hamilton Systems

$$\begin{aligned}
x(t) &= \{X_2(t) - X_1(t)X_1^{-1}(t_1)X_2(t_1)\}d \\
u(t) &= \{U_2(t) - U_1(t)X_1^{-1}(t_1)X_2(t_1)\}d,
\end{aligned}$$

wobei $X_1^{-1}(t_1)X_2(t_1)$ nach Proposition 2.11. symmetrisch ist.
Es ist $x(t) \equiv 0$ auf $[t_1, t_2]$ und damit $\dot{u} = -A^T(t)u, B(t)u(t) \equiv 0$ auf $[t_1, t_2]$. Mit der Steuerbarkeit von (A, B) auf $\mathcal{I}$ folgt $u(t) \equiv 0$ auf $[t_1, t_2]$ und damit $x(t) \equiv u(t) \equiv 0$ auf $\mathcal{I}$ aufgrund des Existenz- und Eindeutigkeitssatzes für Systeme von linearen Differentialgleichungen.

Nach der Definition einer konjugierten Basis und Proposition 2.11. ist

$$\begin{pmatrix} U_2(t) - U_1(t)X_1^{-1}(t_1)X_2(t_1) \\ -\{X_2(t) - X_1(t)X_1^{-1}(t_1)X_2(t_1)\} \end{pmatrix}^T \begin{pmatrix} X_1(t) \\ U_1(t) \end{pmatrix} = I,$$

und damit $\ker \begin{pmatrix} X_2(t) - X_1(t)X_1^{-1}(t_1)X_2(t_1) \\ U_2(t) - U_1(t)X_1^{-1}(t_1)X_2(t_1) \end{pmatrix} = \{0\}$, d.h. $d = 0$. Damit ist $\{X_1^{-1}X_2\}(t)$ streng monoton wachsend auf $\mathcal{I}_1$. □

Proposition 3.3.
Es sei $Q(t)$ eine stetige reelle symmetrische und streng monoton wachsende $n \times n$-matrixwertige Funktion auf einem nicht-entarteten Teilintervall $\mathcal{I}_1 \subset \mathcal{I}$ mit Eigenwerten $\mu_\nu(t)$, $\nu = 1, \ldots, n$, die der Größe nach ansteigend geordnet sein sollen, d.h. $\mu_1(t) \leq \ldots \leq \mu_n(t)$ für $t \in \mathcal{I}_1$.
Dann sind die Eigenwerte $\mu_\nu(t), \nu = 1, \ldots, n$, stetig und streng monoton wachsend auf $\mathcal{I}_1$.

Beweis
Es seien $a \leq t_0 < t_1 \leq b$ und s_k Eigenvektoren von $Q(t_1)$ zum Eigenwert $\mu_k(t_1)$, $k = 1, \ldots, n$. Sei S_ν der von $s_1, \ldots, s_\nu, \nu = 1, \ldots, n$, aufgespannte Untervektorraum von $\mathbb{R}^n$. Mit $R(t, d) = \frac{d^T Q(t) d}{\|d\|^2}$ und Satz A.6. folgt:

$$\begin{aligned} -\mu_\nu(t_1) &= \min_{d \in S_\nu \setminus \{0\}} -R(t_1, d) \\ &= \min_{d \in S_\nu \setminus \{0\}} [R(t_0, d) - R(t_1, d) - R(t_0, d)] \\ &\leq \max_{d \in \mathbb{R}^n \setminus \{0\}} [R(t_0, d) - R(t_1, d)] + \min_{d \in S_\nu \setminus \{0\}} -R(t_0, d) \\ &< -\mu_\nu(t_0) \end{aligned}$$

und damit die strenge Monotonie der Eigenwerte. Die gleiche Argumentation liefert

$$\begin{aligned} \mu_\nu(t_1) &= \max_{d \in S_\nu \setminus \{0\}} R(t_1, d) = \max_{d \in S_\nu \setminus \{0\}} [R(t_1, d) - R(t_0, d) + R(t_0, d)] \\ &\leq \max_{d \in \mathbb{R}^n \setminus \{0\}} [R(t_1, d) - R(t_0, d)] + \max_{d \in S_\nu \setminus \{0\}} R(t_0, d) = \gamma + \mu_\nu(t_0), \end{aligned}$$

wobei γ der größte der Eigenwert der Matrix $Q(t_1) - Q(t_0)$ ist. Mit dem Korollar A.7. folgt nun für alle $\nu = 1, \ldots, n$

$$\mu_\nu(t_1) - \mu_\nu(t_0) \leq \gamma \leq \|Q(t_1) - Q(t_0)\|$$

und aufgrund der Stetigkeit von $Q(t)$ die Stetigkeit der $\mu_\nu(t), \nu = 1, \ldots, n$, auf $\mathcal{I}_1$. □

Satz 3.4.
Es sei $B(t) \geq 0$ für $t \in \mathcal{I}$. Dann ist (A, B) genau dann steuerbar auf $\mathcal{I}$, wenn die fokalen Punkte von $X(t)$ in $\mathcal{I}$ isoliert sind für jede konjugierte Basis (X, U) von (H).

Beweis
Es sei zunächst angenommen, dass (A, B) nicht steuerbar auf $\mathcal{I}$ ist. Dies bedeutet, dass es ein $u(t), t \in \mathcal{I}$, gibt mit $\dot{u} = -A^T(t)u, B(t)u(t) \equiv 0$ auf einem nichtentarteten Intervall $[t_1, t_2] \subset \mathcal{I}$ und $u(t_1) \neq 0$. Mit einem solchen u ist $(x(t) \equiv 0, u(t))$ eine Lösung des Hamilton Systems auf $[t_1, t_2]$. Es sei (X, U) eine konjugierte Basis zu dem Anfangswert $X(t_1) = 0$, die es nach Proposition 2.10. gibt. Dann ist $\mathrm{rg} U(t_1) = n$ nach der Definition einer konjugierten Basis und demzufolge bilden die Spaltenvektoren von $U(t_1)$ eine Basis des $\mathbb{R}^n$. Wir ersetzen nun die erste Spalte von $X(t)$ mit $x(t) \equiv 0$, sowie die erste Spalte von $U(t)$ mit $u(t)$ für $t \in [t_1, t_2]$ und bezeichnen die entstehenden Matrizen mit $\widetilde{X}(t)$ und $\widetilde{U}(t)$, wobei nach dem Austauschsatz von Steinitz die Spaltenvektoren von $\widetilde{U}(t)$ ebenfalls eine Basis des $\mathbb{R}^n$ bilden sollen, ansonsten betrachte eine geeignete Permutation der Spaltenvektoren von $\begin{pmatrix} X \\ U \end{pmatrix}(t)$. Dann ist $(\widetilde{X}, \widetilde{U})$ eine $n \times n$-matrixwertige Lösung von (H), für die $\mathrm{rg}(\widetilde{X}^T(t_1), \widetilde{U}^T(t_1)) = n$ und $\{\widetilde{X}^T\widetilde{U} - \widetilde{U}^T\widetilde{X}\}(t_1) = 0$ gilt. Nach Proposition 2.6. ist $(\widetilde{X}, \widetilde{U})$ damit eine konjugierte Basis, für die jedes $t \in [t_1, t_2]$ ein fokaler Punkt von $\widetilde{X}(t)$ ist. Ergo sind die fokalen Punkte der konjugierten Basis $(\widetilde{X}, \widetilde{U})$ nicht isoliert.
Es sei nun (A, B) steuerbar auf $\mathcal{I}$, (X, U) eine konjugierte Basis von (H) und es sei t_1 ein innerer Punkt von $\mathcal{I}$. Nach der Proposition 2.13. und Lemma A.8. gibt es nun eine zweite konjugierte Basis (X_1, U_1) derart, dass $(X_1, U_1), (X_2 = X, U_2 = U)$ eine normalisierte konjugierte Basis bilden und dass $X_1(t)$ regulär auf $\mathcal{I}_1 = (t_1 - \varepsilon, t_1 + \varepsilon) \subset \mathcal{I}$ für ein $\varepsilon > 0$ ist. Mit Proposition 3.2. folgt, dass die stetige matrixwertige Funktion $H(t) = X_1^{-1}(t)X_2(t) = X_1^{-1}(t)X(t)$ streng monoton wachsend auf $\mathcal{I}_1$ ist. Damit sind auch die Eigenwerte $\mu_1(t), \ldots, \mu_n(t)$ von $H(t)$ stetig und streng monoton wachsend nach Proposition 3.3. Daher gibt es ein $0 < \delta < \varepsilon$ mit $\mu_\nu(t) \neq 0$ für alle $t \in (t_1 - \delta, t_1 + \delta) \backslash \{t_1\}$ und allen $\nu = 1, \ldots, n$. Also ist $X(t)$ regulär für alle $t \in (t_1 - \delta, t_1 + \delta) \backslash \{t_1\}$. Daher sind alle fokalen Punkte in $\mathcal{I}$ isoliert. $\square$

4 Der Grenzwertsatz

Im zweiten Abschnitt haben wir das wohlbekannte Hamilton System um den Parameter λ erweitert, um so den Fall der Nichtnegativität auf den der Positivität zurückführen zu können. Wir werden später sehen, dass gerade der Grenzwert $\lambda \to 0-$ besonders interessant ist. Diesen untersuchen wir in dem folgenden Grenzwertsatz, den wir so allgemein fassen, dass wir auch Untersuchungen der Variable t darauf anwenden können.

Satz 4.1. (Grenzwertsatz)
Es seien $X(t)$ und $U(t)$ $n \times n$-matrixwertige Funktionen für $t \in (t_0, t_0 + \varepsilon]$, $t_0 \in \mathbb{R}, \varepsilon > 0$, für die gilt:

$$\begin{aligned}
&X^T(t)U(t) = U^T(t)X(t) \text{ für } t \in (t_0, t_0 + \varepsilon],\\
&X(t) \to X, U(t) \to U (t \to t_0+),\\
&\mathrm{rg}(X^T, U^T) = n \text{ und } X^T U = U^T X,\\
&X(t) \text{ ist regulär für } t \in (t_0, t_0 + \varepsilon] \text{ und}\\
&Q(t) := U(t)X^{-1}(t) \text{ ist monoton fallend auf } (t_0, t_0 + \varepsilon].
\end{aligned}$$

Dann gilt:

$$\lim_{t \to t_0+} d^T Q(t) d = \infty \text{ für alle } d \notin \mathrm{Im} X.$$

Beweis
Wir führen den Beweis analog zu [[4], Proposition 3.3.10]. Es bezeichne $r := \mathrm{rg} X$. Da $Q(t)$ reell und symmetrisch ist, ist $Q(t)$ diagonalisierbar, d.h. es gibt eine orthogonale Matrix $T(t)$, also eine Matrix $T(t) \in \mathbb{R}^{n \times n}$ mit $T(t)T^T(t) = I$, mit

$$T^T(t)Q(t)T(t) = \begin{pmatrix} D_1(t) & 0 \\ 0 & D_2(t) \end{pmatrix},$$

wobei

$$D_1(t) = \begin{pmatrix} \mu_1(t) & & 0 \\ & \ddots & \\ 0 & & \mu_r(t) \end{pmatrix}, \quad D_2(t) = \begin{pmatrix} \mu_{r+1}(t) & & 0 \\ & \ddots & \\ 0 & & \mu_n(t) \end{pmatrix}$$

Diagonalmatrizen und die $\mu_\nu(t), \nu = 1, \ldots, n$, die Eigenwerte der Matrix $Q(t)$ sind, die betragsmäßig der Größe nach ansteigend geordnet sein sollen, d.h. $|\mu_1(t)| \leq \ldots \leq |\mu_n(t)|$.
Wir untersuchen nun das Verhalten der Eigenwerte für $t \to t_0+$. Dazu benutzen wir den Satz A.6. mit $V_1(t) = \{X(t)c | c \in \ker X\}$, so dass $\dim V_1(t) = n - r$. Da nach Satz 2.12.(i) $Ud \neq \{0\}$ für alle $d \in \ker X \backslash \{0\}$ gibt es ein $\delta > 0$ und ein $\eta > 0$ mit

$$\begin{aligned}
|\mu_\nu(t)| \quad &\geq \quad |\mu_{r+1}(t)| = \max_{\dim V = n-r} \min\{\frac{\|Q(t)d\|}{\|d\|} | d \in V \backslash \{0\}\} \geq \inf_{d \in V_1(t) \backslash \{0\}} \frac{\|Q(t)d\|}{\|d\|}\\
&= \inf_{c \in \ker X \backslash \{0\}} \frac{\|U(t)c\|}{\|(X(t) - X)c\|} \geq \frac{\eta}{\gamma(t)} \text{ für } t \in (t_0, t_0 + \delta] \text{ und } \nu = r+1, \ldots, n,
\end{aligned}$$

wobei $\gamma(t) = \|X(t) - X\| \to 0$ für $t \to t_0+$.
Sei nun $V_2(t) = \{X(t)c | c \in \mathrm{Im}X^T\}$, so ist $\dim V_2(t) = r = \mathrm{rg}X$. Dann gibt es nach unseren Voraussetzungen und Lemma A.9. ein $\alpha > 0$ mit $\|XX^Tc\| \geq \alpha\|X^Tc\|$ für alle $c \in \mathbb{R}^n$, ein $\beta \geq 0$ mit $\|U(t)\| \leq \beta$ für alle $t \in \mathcal{I}$, sowie ein $\delta > 0$ mit $\|X(t) - X\| \leq \frac{\alpha}{2}$ für alle $t \in (t_0, t_0 + \delta]$, so dass

$$\begin{aligned}
|\mu_\nu(t)| &\leq |\mu_r(t)| = \min_{\dim V = r} \max\{\frac{\|Q(t)d\|}{\|d\|} | d \in V \backslash \{0\}\} \\
&\leq \sup_{d \in V_2(t) \backslash \{0\}} \frac{\|Q(t)d\|}{\|d\|} = \sup_{c \in \mathbb{R}^n \backslash \ker X^T} \frac{\|U(t)X^Tc\|}{\|X(t)X^Tc\|} \\
&\leq \sup_{c \in \mathbb{R}^n \backslash \ker X^T} \frac{\|U(t)X^Tc\|}{\|XX^Tc\| - \|(X(t) - X)X^Tc\|} \\
&\leq \sup_{c \in \mathbb{R}^n \backslash \ker X^T} \frac{\beta\|X^Tc\|}{\alpha\|X^Tc\| - \alpha/2\|X^Tc\|} = 2\frac{\beta}{\alpha}
\end{aligned}$$

für $t \in (t_0, t_0 + \delta]$ und $\nu = 1, \ldots, r$.
Die Eigenwerte $\mu_1(t), \ldots, \mu_r(t)$ sind also beschränkt für $t \to t_0+$, und die Eigenwerte $\mu_{r+1}(t), \ldots, \mu_n(t)$ konvergieren gegen ∞ für $t \to t_0+$, da $Q(t)$ monoton fallend auf $(t_0, t_0 + \varepsilon]$ ist.
Nehmen wir nun an, (d_k) sei eine konvergente Folge in $\mathbb{R}^n$ mit Grenzwert d, (t_k) sei eine konvergente Folge in $\mathbb{R}$ mit Grenzwert t_0 und $t_0 < t_k$ für alle $k \in \mathbb{N}$, und $d_k^T Q(t_k) d_k$ sei beschränkt für $k \to \infty$. Aufgrund der Kompaktheit dürfen wir ohne Beschränkung der Allgemeinheit annehmen, dass $T(t_k) \to T(k \to \infty)$ konvergiert, wobei T orthogonal ist, ansonsten betrachte eine entsprechende Teilfolge, die nach dem Auswahlprinzip von Bolzano-Weierstraß, siehe [[3], Seite 156], konvergent ist. Demnach konvergiert die Folge $c_k := T^T(t_k)d_k \to T^Td(k \to \infty)$.
Sofern

$$d_k^T Q(t_k) d_k = c_k^T \begin{pmatrix} D_1(t_k) & 0 \\ 0 & D_2(t_k) \end{pmatrix} c_k$$

beschränkt ist, besitzt d die Gestalt $d = T \begin{pmatrix} \xi \\ 0 \end{pmatrix}$ mit einem $\xi \in \mathbb{R}^r$. Betrachte nun die konvergente Folge $\widetilde{d}_k = X(t_k)d$ mit Grenzwert Xd für $t_k \to t_0+$. Da für diese $\widetilde{d}_k^T Q(t_k)\widetilde{d}_k = d^T X^T(t_k)Q(t_k)X(t_k)d \to d^T X^T U d(t_k \to t_0+)$ beschränkt ist, ist nach obiger Überlegung ist also $\mathrm{Im}X \subset \mathrm{Im}T \begin{pmatrix} I_{r \times r} & 0 \\ 0 & 0 \end{pmatrix}$. Ferner folgt aufgrund der Rangbedingung die Gleichheit. Folglich impliziert jede konvergente Folge $d_k \to d(k \to \infty)$ mit beschränktem $d_k^T Q(t_k) d_k$ für eine von rechts gegen t_0 konvergente Folge, dass $d \in \mathrm{Im}T \begin{pmatrix} I_{r \times r} & 0 \\ 0 & 0 \end{pmatrix} = \mathrm{Im}X$. Hieraus ergibt sich die Behauptung. $\square$

Bemerkung 4.2.
Es ergibt sich ein entsprechendes Resultat, wenn man den linksseitigen Grenzwert betrachtet, d.h. unter Voraussetzungen analog zu denen des vorigen Satzes, wobei wir das halboffene Intervall $[t_0 - \varepsilon, t_0)$ und entsprechend den Grenzwert $t \to t_0-$ betrachten, gilt

$$\lim_{t \to t_0-} d^T Q(t) d = -\infty \text{ für alle } d \notin \mathrm{Im} X.$$

Korollar 4.3.
Es sei (A, B) steuerbar und $(X_1, U_1), (X_2, U_2)$ normalisierte konjugierte Basen von (H), so dass $X_2(t_0)$ regulär für ein gegebenes $t_0 \in \mathcal{I}$ ist. Dann gilt

$$\lim_{t \to t_0+} d^T \{-X_1^{-1} X_2\}(t) d = \infty \text{ für alle } d \notin \mathrm{Im} X_1^T(t_0).$$

Beweis
Wir benutzen den Grenzwertsatz, Satz 4.1., mit $X(t) = -X_1^T(t)$ und $U(t) = X_2^T(t)$. Die Voraussetzungen des Grenzwertsatzes sind erfüllt, denn es gilt:

- $-\{X_1 X_2^T\}(t) = -\{X_2 X_1^T\}(t)$ für alle $t \in \mathcal{I}$, insbesondere auch für t_0, nach Proposition 2.11.
- $-X_1^T(t) \to -X_1^T(t_0), X_2^T(t) \to X_2^T(t_0) (t \to t_0+)$ aufgrund der Stetigkeit in t.
- $\mathrm{rg}(X_1(t_0), X_2(t_0)) = n$, da $X_2(t_0)$ regulär ist.
- Es gibt ein $\varepsilon > 0$ so, dass $X_1(t)$ regulär ist für $t \in (t_0, t_0 + \varepsilon]$. Ist nämlich $X_1(t_0)$ singulär, so ist der fokale Punkt von $X_1(t_0)$ nach Satz 3.4. isoliert. Andernfalls sagt dieses Korollar nichts aus.

Mit dem Grenzwertsatz folgt nun unmittelbar die Behauptung, da $\{UX^{-1}\}(t) = \{(X^{-1})^T U^T\}(t)$, sofern $\{X^T U\}(t)$ symmetrisch ist. □

5 Die Picone-Identität

Um eine hinreichende Bedingung für die Positivität und die Nichtnegativität des quadratischen Funktionales $\mathcal{F}$ ermitteln, werden wir in diesem Abschnitt das Funktional $\mathcal{F}$ nach unten abschätzen. Betrachtet man den Integranden von $\mathcal{F}_0(x)$, so sieht man unmittelbar den Zusammenhang zur nachstehenden Picone-Identität. Es gilt nämlich

$$\mathcal{F}_0(x) \geq F(t,x(t))|_a^b + \int_a^b \{z^T Bz\}(t)\,dt$$

mit den Bezeichnungen des nachstehenden Satzes, wie sich im Beweis der Sätze 6.2. und 6.3. ergeben wird. Auch dieser Satz findet seinen Ursprung in der Variationsrechnung, wie man in [[4], Chapter 1.3] nachlesen kann.

Satz 5.1. (Picone-Identität)
Es sei x zulässig und $(X_1,U_1),(X_2,U_2)$ normalisierte konjugierte Basen von (H), für die X_1 regulär auf einem nicht-entarteten Teilintervall $\mathcal{I}_1 \subset \mathcal{I}$ ist. Dann gilt für $t \in \mathcal{I}_1$, für die $x(t)$ differenzierbar ist, und $\alpha \in \mathbb{R}^n$:

$$\frac{d}{dt}F(t,x(t)) = \{x^T Cx + u^T Bu - z^T Bz\}(t),$$

wobei die Funktionen $z(t)$ und $F(t,x)$ gegeben sind durch

$$\begin{aligned} z(t) &= u(t) - U_1(t)X_1^{-1}(t)x(t) - (X_1^{-1})^T(t)\alpha \\ F(t,x) &= x^T U_1(t)X_1^{-1}(t)x + 2\alpha^T X_1^{-1}(t)x - \alpha^T X_1^{-1}(t)X_2(t)\alpha. \end{aligned}$$

Beweis
Wir folgen dem Beweis von [[4], Theorem 1.2.1] und evaluieren unter Weglassen der Variablen t und mit Proposition 2.11. und Lemma A.5.

$$\begin{aligned} \tfrac{d}{dt}\{x^T U_1 X_1^{-1}x\} &= (Ax+Bu)^T U_1 X_1^{-1}x + x^T(CX_1 - A^T U_1)X_1^{-1}x \\ &\quad -x^T U_1 X_1^{-1}(AX_1 + BU_1)X_1^{-1}x + x^T U_1 X_1^{-1}(Ax+Bu) \\ &= x^T(C - U_1X_1^{-1}BU_1X_1^{-1})x + u^T BU_1X_1^{-1}x + x^T U_1 X_1^{-1}Bu \\ &= x^T Cx + u^T Bu - (u - U_1X_1^{-1}x)^T B(u - U_1X_1^{-1}x) \end{aligned}$$

$$\begin{aligned} &\tfrac{d}{dt}\{2\alpha^T X_1^{-1}x - \alpha^T X_1^{-1}X_2\alpha\} = -2\alpha^T X_1^{-1}(AX_1+BU_1)X_1^{-1}x + 2\alpha^T X_1^{-1}(Ax+Bu) \\ &\quad +\alpha^T X_1^{-1}(AX_1+BU_1)X_1^{-1}X_2\alpha - \alpha^T X_1^{-1}(AX_2+BU_2)\alpha \\ &= -2\alpha^T X_1^{-1}BU_1X_1^{-1}x + 2\alpha^T X_1^{-1}Bu + \alpha^T X_1^{-1}BU_1X_1^{-1}X_2\alpha - \alpha^T X_1^{-1}BU_2\alpha \\ &= 2\alpha^T X_1^{-1}B(u - U_1X_1^{-1}x) + \alpha^T X_1^{-1}B(U_1X_1^{-1}X_2 - U_2)\alpha \\ &= 2\alpha^T X_1^{-1}B(u - U_1X_1^{-1}x) + \alpha^T X_1^{-1}B(X_1^{-1})^T(X_1^T U_1 X_1^{-1}X_2 - X_1^T U_2)\alpha \\ &= 2\alpha^T X_1^{-1}B(u - U_1X_1^{-1}x) + \alpha^T X_1^{-1}B(X_1^{-1})^T(U_1^T X_2 - X_1^T U_2)\alpha \\ &= 2\alpha^T X_1^{-1}B(u - U_1X_1^{-1}x) - \alpha^T X_1^{-1}B(X_1^{-1})^T\alpha \end{aligned}$$

Die Addition der beiden obigen Terme ergibt die Behauptung. □

In den Vorbetrachtungen zu diesem Abschnitt haben wir festgestellt, dass wir das Funktional $\mathcal{F}(x)$ mittels der Picone-Identität nach unten abschätzen können, sofern die matrixwertige Funktion $X_1(t)$ der normalisierten konjugierten Basen (X_1, U_1), (X_2, U_2) von (H) regulär ist. Wir werden aber sehen, dass gerade der Fall interessant ist, wenn $X_1(t)$ am Rand des betrachteten Intervalls $\mathcal{I}$ singulär ist. Das folgende Korollar sichert aber zumindest die Existenz des Grenzwertes der Funktion F aus dem Satz der Picone-Identität mit von t abhängigem α an den Stellen, an denen $X_1(t)$ singulär ist.

Korollar 5.2.
Es sei (A, B) steuerbar auf $\mathcal{I}$, $(X_1, U_1), (X_2, U_2)$ seien normalisierte konjugierte Basen von (H) und $x(t)$ eine auf $\mathcal{I}$ stetige Funktion mit $x(t) \in \mathrm{Im}X_1(t)$ für $t \in \mathcal{I}$. Ferner sei $t_0 \in \mathcal{I}$. Dann gilt mit

$$\begin{aligned} F(t, x(t)) &= \{x^T U_1 X_1^\dagger x + 2\alpha^T X_1^\dagger x - \alpha^T X_1^\dagger X_2 \alpha\}(t) \text{ und} \\ \alpha(t) &= X_1^T(t) U_1(t_0) X_1^\dagger(t_0) x(t_0) - U_1^T(t) x(t), \end{aligned}$$

dass

$$\lim_{t \to t_0} F(t, x(t)) = x^T(t_0) U_1(t_0) X_1^\dagger(t_0) x(t_0) \;\text{ und }\; \lim_{t \to t_0} \alpha(t) = 0.$$

Beweis
Aufgrund der Voraussetzung, dass (A, B) steuerbar auf $\mathcal{I}$ ist, sind nach Satz 3.4. die fokalen Punkte von $X(t)$ in $\mathcal{I}$ isoliert. Da ferner die Ableitung von $x(t)$ bis auf endlich viele Punkte in $\mathcal{I}$ existiert, gibt es eine punktierte Umgebung um t_0, d.h. ohne t_0, auf der $X_1(t)$ regulär und $x(t)$ differenzierbar ist. Also ist Satz 5.1. anwendbar. Die Funktionen $X_1(t), U_1(t), X_2(t), U_2(t)$ und $x(t)$ sind absolut stetig auf $\mathcal{I}$. Allerdings ist $X_1^\dagger(t)$ unstetig in allen $t \in \mathcal{I}$, in denen $X_1(t)$ singulär ist. Durch die Wahl von $\alpha(t)$ werden wir aber genau solche Terme eliminieren.
Wir verifizieren die Behauptung unter Anwendung von Proposition 2.11. und der Voraussetzung, dass $x(t) \in \mathrm{Im}X(t)$ für $t \in \mathcal{I}$, wobei wir wie im vorangegangenen Satz den Parameter t weglassen wollen:

$$\begin{aligned} F(\cdot, x(\cdot)) = & \quad x^T U_1 X_1^\dagger x + 2x^T(t_0)(X_1^\dagger)^T(t_0) U_1^T(t_0) X_1 X_1^\dagger x - 2x^T U_1 X_1^\dagger x \\ & - \; x^T(t_0)(X_1^\dagger)^T(t_0) U_1^T(t_0) X_1 X_1^\dagger X_2 X_1^T U_1(t_0) X_1^\dagger(t_0) x(t_0) \\ & - \; x^T U_1 X_1^\dagger X_2 U_1^T x + 2x^T U_1 X_1^\dagger X_2 X_1^T U_1(t_0) X_1^\dagger(t_0) x(t_0) \\ = & - \; x^T U_1 X_1^\dagger X_1 U_2^T x + 2x^T(t_0)(X_1^\dagger)^T(t_0) U_1^T(t_0) x \\ & - \; x^T(t_0)(X_1^\dagger)^T(t_0) U_1^T(t_0) X_1 X_2^T U_1(t_0) X_1^\dagger(t_0) x(t_0) \\ & + \; 2x^T U_1 X_1^\dagger X_1 X_2^T U_1(t_0) X_1^\dagger(t_0) x(t_0) \\ = & - \; x^T U_1 U_2^T x + 2x^T(t_0)(X_1^\dagger)^T(t_0) U_1^T(t_0) x \\ & - \; x^T(t_0)(X_1^\dagger)^T(t_0) U_1^T(t_0) X_1 X_2^T U_1(t_0) X_1^\dagger(t_0) x(t_0) \\ & + \; 2x^T U_1 X_2^T U_1(t_0) X_1^\dagger(t_0) x(t_0). \end{aligned}$$

Nach Voraussetzung ist $x(t)$ stetig an t_0 und damit ist der letzte Ausdruck ist stetig an t_0 und es gilt bei Weglassen des gemeinsamen Parameters t_0

$$\begin{aligned} \lim_{t \to t_0} F(t, x(t)) &= x^T(-U_1 U_2^T + 2U_1 X_1^\dagger - U_1 X_2^T U_1 X_1^\dagger + 2U_1 X_2^T U_1 X_1^\dagger)x \\ &= x^T U_1 X_1^\dagger x. \end{aligned}$$

Die Behauptung für den Grenzwert von $\alpha(t)$ folgt unmittelbar mit der Definition einer konjugierten Basis. □

6 Hauptsätze zur Definitheit quadratischer Funktionale

Es gelten in diesem zentralen Abschnitt wie bisher (1) und (2). Wir betrachten nun eine spezielle konjugierte Basis (X_a, U_a) des Hamilton Systems, für die gilt

$$X_a(a) = R_a,\ R_a^T U_a(a) = -R_a^T S_a R_a. \tag{4}$$

Proposition 6.1.
Es gibt eine konjugierte Basis (X_a, U_a) von (H), die den Anfangsbedingungen (4) genügt.

Beweis
Es sei (X, U) eine $n \times n$-matrixwertige Lösung des Hamilton Systems (H) zu den Anfangswerten $X_a(a) = R_a, U_a(a) = -S_a R_a + \widetilde{S}$, wobei

$$R_a^T \widetilde{S} = 0,\ \mathrm{rg}(R_a^T, \widetilde{S}^T) = n.$$

Ein solches $\widetilde{S}$ existiert nach Lemma 2.9.
Dann gilt

$$\mathrm{rg}(X_a^T(a), U_a^T(a)) = \mathrm{rg}(R_a^T, -R_a^T S_a^T + \widetilde{S}^T) = \mathrm{rg}(R_a^T, \widetilde{S}^T) = n$$

und mit der Symmetrie von S_a

$$X_a^T(a)U_a(a) - U_a^T(a)X_a(a) = -R_a^T S_a R_a + R_a^T \widetilde{S} + R_a^T S_a^T R_a - \widetilde{S}^T R_a = 0.$$

Nach Proposition 2.6. ist ein solches (X, U) eine konjugierte Basis. □

Wir haben nun die nötigen Vorbereitungen getroffen um die Sätze zur Charakterisierung der Positivität und Nichtnegativität des quadratischen Funktionales $\mathcal{F}$ zu beweisen. Da der Beweis des hinreichenden Teils bei beiden Sätzen sehr ähnlich verläuft, werden wir diesen Teil der beiden Beweise zusammenfassen. Anschließend werden wir jedoch zunächst die notwendigen Bedingungen der Positivität getrennt beweisen, da wir im Fall der Nichtnegativität auf diesen zurückgreifen möchten.

Satz 6.2. (Positivität)
Es sei (A, B) steuerbar auf $\mathcal{I}$ und (X_a, U_a) sei eine spezielle konjugierte Basis des Hamilton Systems (H), die den Anfangsbedingungen (4) genügt. Dann ist $\mathcal{F}$ genau dann positiv definit, wenn die folgenden Bedingungen gelten:

(i) $B(t) \geq 0$ auf $\mathcal{I}$.

(ii) $X_a(t)$ ist regulär für alle $t \in (a, b]$.

(iii) $R_b^T\{S_b + U_a(b)X_a^{-1}(b)\}R_b$ ist positiv definit auf $\mathrm{Im}R_b^T$.

Satz 6.3. (Nichtnegativität)
Es sei (A, B) steuerbar auf $\mathcal{I}$ und (X_a, U_a) sei eine spezielle konjugierte Basis des Hamilton Systems (H), die den Anfangsbedingungen (4) genügt. Dann ist $\mathcal{F}$ genau dann nichtnegativ definit, wenn die folgenden Bedingungen gelten:

(i*) $B(t) \geq 0$ auf $\mathcal{I}$.

(ii*) $X_a(t)$ ist regulär für alle $t \in (a, b)$.

(iii*) $R_b^T\{S_b + U_a(b)X_a^\dagger(b)\}R_b$ ist nichtnegativ definit.

(iv*) $\mathrm{Im}R_b \subset \mathrm{Im}X_a(b)$.

Beweis der Sätze 6.2. und 6.3.
Zunächst werden wir zeigen, dass die Bedingungen (i*)-(iv*) hinreichend für die Nichtnegativität und die Bedingungen (i)-(iii) hinreichend für die Positivität eines quadratischen Funktionales sind.
Es gelten die Bedingungen (i*)-(iv*) und x sei zulässig mit

$$x(a) = R_a c_a,\ x(b) = R_b c_b$$

für gewisse $c_a, c_b \in \mathbb{R}^n$. Wir benutzen Korollar 5.2. für $X_1 = X_a, U_1 = U_a$ und einer konjugierten Basis (X_2, U_2) von (H), so dass $(X_a, U_a), (X_2, U_2)$ normalisierte konjugierte Basen von (H) bilden,

$$\begin{aligned}
\alpha_a(\varepsilon) &= X_a^T(a+\varepsilon)U_a(a)X_a^\dagger(a)x(a) - U_a^T(a+\varepsilon)x(a+\varepsilon),\\
\alpha_b(\varepsilon) &= X_a^T(b-\varepsilon)U_a(b)X_a^\dagger(b)x(b) - U_a^T(b-\varepsilon)x(b-\varepsilon)
\end{aligned}$$

und einem hinreichend kleinem $\varepsilon > 0$, wobei wegen (ii*), (iv*) und (4) $x(t) \in \mathrm{Im}X_a(t)$ für alle $t \in \mathcal{I} = [a, b]$ gilt. Ferner benutzen wir die Picone-Identität, Satz 5.1., mit dem in diesem Satz definierten $F(t, x(t))$, obigen normalisierten konjugierten Basen und

$$\begin{aligned}
z_{\varepsilon,\nu}(t) &:= u(t) - U_a(t)X_a^{-1}(t)x(t) - (X_a^{-1})^T(t)\alpha_\nu(\varepsilon)\\
\to \quad z(t) &:= u(t) - U_a(t)X_a^{-1}(t)x(t) \qquad\qquad \text{für } \varepsilon \to 0+, \nu \in \{a, b\}
\end{aligned}$$

Dann gilt mit beliebigem $\tau \in (a, b)$ und für alle hinreichend kleinen $\delta > 0$ nach oben zitiertem Satz, Korollar 5.2. und wegen (i*):

$$\begin{aligned}
\mathcal{F}_0(x) &= \int_a^b \{x^T Cx + u^T Bu\}(t)\,dt\\
&= \lim_{\varepsilon\to 0+} \Big[\int_{a+\varepsilon}^{\tau} \{x^T Cx + u^T Bu\}(t)\,dt + \int_{\tau}^{b-\varepsilon} \{x^T Cx + u^T Bu\}(t)\,dt\Big]\\
&\geq \lim_{\varepsilon\to 0+} \Big[\int_{a+\delta}^{\tau} \{z_{\varepsilon,a}^T B z_{\varepsilon,a}\}(t)\,dt + F(t, x(t))|_{a+\varepsilon}^{\tau} +\\
&\qquad \int_{\tau}^{b-\delta} \{z_{\varepsilon,b}^T B z_{\varepsilon,b}\}(t)\,dt + F(t, x(t))|_{\tau}^{b-\varepsilon}\Big]\\
&= \int_{a+\delta}^{b-\delta} \{z^T Bz\}(t)\,dt + \{x^T U_a X_a^\dagger x\}(b) - \{x^T U_a X_a^\dagger x\}(a)\\
&\geq \{x^T U_a X_a^\dagger x\}(b) - \{x^T U_a X_a^\dagger x\}(a).
\end{aligned}$$

Damit folgt

$$\begin{aligned}
\mathcal{F}(x) &= \mathcal{F}_0(x) + x^T(b)S_b x(b) - x^T(a)S_a x(a) \\
&\geq \{x^T U_a X_a^\dagger x\}(b) - \{x^T U_a X_a^\dagger x\}(a) + x^T(b)S_b x(b) - x^T(a)S_a x(a) \\
&= x^T(b)\{S_b + U_a(b)X_a^\dagger(b)\}x(b) - c_a^T\{X_a^T(a)U_a(a) + R_a^T S_a R_a\}c_a \\
&\geq 0 \text{ nach (iii*) und (4).}
\end{aligned}$$

Damit haben wir nun also die hinreichende Bedingung für die Nichtnegativität eines quadratischen Funktionales gezeigt.
Gelten nun die Bedingungen (i)-(iii), so folgt aus diesen unmittelbar die Bedingungen (i*)-(iv*) und damit $\mathcal{F}(x) \geq 0$, wobei wir die Folgerung (iii) $\Rightarrow$ (iii*) kurz erläutern möchten. Es ist

$$\begin{aligned}
&c^T R_b^T\{S_b + U_a(b)X_a^{-1}(b)\}R_b c > 0 \text{ für alle } c \in \mathrm{Im} R_b^T \backslash \{0\} \\
\Leftrightarrow\ &c^T R_b^T\{S_b + U_a(b)X_a^{-1}(b)\}R_b c > 0 \text{ für alle } c \in \mathbb{R}^n \text{ mit } R_b c \neq 0,
\end{aligned}$$

denn mit $\mathbb{R}^n = \mathrm{Im} R_b^T \oplus \ker R_b$ gibt es für alle $c \in \mathbb{R}^n$ Vektoren $d, \widetilde{d} \in \mathbb{R}^n$ mit $c = R_b^T d + \widetilde{d}$ und $R_b \widetilde{d} = 0$. Dann ist $R_b c \neq 0 \Leftrightarrow R_b R_b^T d \neq 0 \Leftrightarrow R_b^T d \neq 0$. Damit folgt obige Äquivalenz und auch (iii*).
Ist $\mathcal{F}(x) = 0$, so folgt aus der Abschätzung oben, dass $x^T(b)\{S_b + U_a(b)X_a^{-1}(b)\}x(b)$ $= 0$ und damit wegen der Bedingung (iii) $x(b) = 0$. Da die obige Abschätzung für alle hinreichend kleinen $\delta > 0$ gilt, ist überdies $Bz \equiv 0$ und damit $Bu \equiv Bv$ für $v = U_a X_a^{-1} x$, wobei (x, v) das Hamilton System mit $0 = x(b) = v(b)$ löst. Mit dem Existenz- und Eindeutigkeitssatz, Satz A.3., folgt somit $x(t) \equiv 0$ auf $\mathcal{I}$. Dies ergibt die hinreichende Bedingung für die Positivität eines quadratischen Funktionales.

Die Notwendigkeit der Bedingung (i), bzw. (i*), ergibt sich unmittelbar aus Satz 1.3. Da wir im weiteren Verlauf dieses Beweises die Nichtnegativität auf den Fall der Positivität zurückführen werden, wollen wir zunächst den notwendigen Teil von Satz 6.2. getrennt beweisen. Dazu bemerken wir zunächst, dass es nach Satz 3.4. ein $\varepsilon > 0$ gibt, für das

$$X_a(t) \text{ regulär für alle } t \in (a, a+\varepsilon) \text{ ist.} \tag{*}$$

Zu (ii):
Es sei angenommen, es gäbe ein $t_0 \in (a, b]$ mit $X_a(t_0)c = 0$ für ein $c \in \mathbb{R}^n \backslash \{0\}$.
Setze nun

$$\begin{aligned}
x(t) &:= \begin{cases} X_a(t)c, & t \in [a, t_0] \\ 0, & t \in (t_0, b], \end{cases} \\
u(t) &:= \begin{cases} U_a(t)c, & t \in [a, t_0] \\ 0, & t \in (t_0, b]. \end{cases}
\end{aligned}$$

Dann ist x zulässig und mit (*) gilt $x(t) \neq 0$ für $t \in (a, a+\varepsilon)$ wegen der Regularität von $X_a(t)$, so dass $x(t) \not\equiv 0$ auf $\mathcal{I}$.

Mittels partieller Integration folgt wie in Bemerkung 2.2.(ii)

$$\begin{aligned}\mathcal{F}(x) &= x^T u|_a^{t_0} - x^T(a)S_a x(a) \\ &= -x^T(a)u(a) - x^T(a)S_a x(a) \\ &= -c^T X_a^T(a)U_a(a)c - c^T R_a^T S_a R_a c \\ &= 0,\end{aligned}$$

ein Widerspruch zu unserer Voraussetzung.

Zu (iii):
Sei nun also $X_a(t)$ regulär auf $(a,b]$, aber angenommen $R_b^T\{S_b + U_a(b)X_a^{-1}(b)\}R_b$ ist nicht positiv definit auf $\mathrm{Im}R_b^T$, d.h. es gibt ein $c \in \mathrm{Im}R_b^T$ mit

$$c^T R_b^T\{S_b + U_a(b)X_a^{-1}(b)\}R_b c \leq 0 \text{ und } d := R_b c \neq 0.$$

Betrachte die Lösung

$$\begin{pmatrix} x \\ u \end{pmatrix}(t) \quad := \quad \begin{pmatrix} X_a \\ U_a \end{pmatrix}(t)X_a^{-1}(b)d.$$

Dann gilt ähnlich wie oben:

$$\begin{aligned}\mathcal{F}(x) &= x^T u|_a^b + x^T(b)S_b x(b) - x^T(a)S_a x(a) \\ &= x^T(b)u(b) - x^T(a)u(a) + x^T(b)S_b x(b) - x^T(a)S_a x(a) \\ &= x^T(b)u(b) + x^T(b)S_b x(b) \\ &= c^T R_b^T\{S_b + U_a(b)X_a^{-1}(b)\}R_b c \leq 0,\end{aligned}$$

ein Widerspruch. Demnach haben wir die Notwendigkeit der Bedingungen (i) - (iii) gezeigt.

Wir fahren nun mit dem Beweis der Notwendigkeit der Bedingungen (ii*)-(iv*) fort. Dabei betrachten wir eine konjugierte Basis (X_a, U_a) von (H_λ), für die gilt

$$X_a(a,\lambda) \equiv X_a(a) \text{ und } U_a(a,\lambda) \equiv U_a(a) \text{ für alle } \lambda \in \mathbb{R}.$$

Zu (ii*):
Es sei angenommen, es gäbe ein $t_0 \in (a,b)$, so dass $X_a(t_0) = X_a(t_0, 0)$ singulär ist. Wegen $\mathcal{F}(x) \geq 0$ ist $\mathcal{F}(x,\lambda) > 0$ für alle $\lambda < 0$ nach Lemma 2.15. Daher ist $X_a(t,\lambda)$ regulär für alle $\lambda < 0$ und $t \in (a,b]$ nach Satz 6.2.
Wir betrachten

$$H(t,\lambda) := -X_a^{-1}(t,\lambda)X_*(t,\lambda)$$

mit einer konjugierten Basis (X_*, U_*) von (H_λ) so, dass $(X_a, U_a), (X_*, U_*)$ normalisierte konjugierte Basen von (H_λ) bilden, und dass $X_*(t_0, 0)$ regulär ist. Die Existenz einer solchen konjugierten Basis ist nach Proposition 2.13. gewährleistet.
Setze nun $\lambda := -1$ und

$$c := d^T H(t_0, -1)d \text{ mit einem } d \notin \mathrm{Im}X_a^T(t_0, 0).$$

Da (A,B) steuerbar auf $\mathcal{I}$ und damit nach Satz 3.4. die fokalen Punkte von $X_a(t,0)$ in t isoliert sind, gibt es ein $\varepsilon > 0$ so, dass $t_0 + \varepsilon \le b$ und $X_a(t,0)$ regulär ist für alle $t \in (t_0, t_0+\varepsilon]$, d.h. H existiert auf

$$\mathcal{J} := ([t_0, t_0+\varepsilon] \times [-1,0]) \backslash \{(t_0, 0)\}.$$

Nach den Propositionen 2.19. und 3.2. ist $H(t,\lambda)$ monoton fallend in den beiden Parametern t und λ und damit

$$d^T H(t,\lambda) d \le c \text{ für alle } (t,\lambda) \in \mathcal{J}.$$

Da nach dem Korollar zum Grenzwertsatz, Korollar 4.3., jedoch

$$\lim_{t \to t_0+} d^T H(t,0) d = +\infty,$$

folgt aufgrund der Stetigkeit von H auf $\mathcal{J}$ ein Widerspruch, denn dann müsste es ein $\widetilde{t} \in (t_0, t_0+\varepsilon]$ geben mit $d^T H(\widetilde{t},0) d \ge 2c$.

Zu (iv*):
Es sei angenommen es gäbe ein $d \in \mathrm{Im} R_b$ so, dass $d \notin \mathrm{Im} X_a(b,0)$ ist. Mit der Bemerkung zum Grenzwertsatz, Bemerkung 4.2., folgt nun

$$d^T \{S_b + U_a(b,\lambda) X_a^{-1}(b,\lambda)\} d \to -\infty (\lambda \to 0-),$$

wobei der Grenzwertsatz anwendbar ist, da nach Proposition 2.19. $\{U_a X_a^{-1}\}(t,\lambda)$ monoton fallend in λ ist und aufgrund der der Stetigkeit in λ nach Lemma 2.17. Wegen letztere gibt es ein $\varepsilon > 0$, so dass $\mathcal{F}(x,\lambda) \not> 0$ für $-\varepsilon \le \lambda < 0$. Dies ist jedoch ein Widerspruch dazu, dass aus $\mathcal{F}(x) \ge 0$ die Positivität von $\mathcal{F}(x,\lambda)$ folgt, Lemma 2.15.

Zu (iii*):
Es sei angenommen, es gäbe ein $d = X_a(b)c \in \mathrm{Im} R_b$ mit $d^T \{S_b + U_a(b) X_a^\dagger(b)\} d < 0$. Setze $(x(t), u(t)) = (X_a(t), U_a(t))c$.
Dann ist x zulässig und es gilt

$$\begin{aligned}\mathcal{F}(x) &= x^T u|_a^b + x^T(b) S_b x(b) - x^T(a) S_a x(a) \\ &= d^T \{S_b + U_a(b) X_a^\dagger(b)\} d < 0,\end{aligned}$$

ein Widerspruch. □

Bemerkung 6.4.

(i) Ist die matrixwertige Funktion $B(t)$ positiv definit auf $\mathcal{I}$, so tritt das quadratische Funktional $\mathcal{F}$ mit der Kontrollfunktion $u = B^{-1}(\dot{x} - Ax)$ in der klassischen Variationsrechnung auf. In diesem Fall sagt man, dass das zu untersuchende Extremal die *verschärfte Legendre Bedingung* erfüllt. Siehe hierzu [[4] Lemma 8.2.4].

(ii) Im Beweis der Positivität haben wir die Steuerbarkeit von (A, B) auf $\mathcal{I}$ lediglich dahingehend verwendet, um zu zeigen, dass $X(t)$ regulär auf einem Bereich $(a, a+\varepsilon]$ mit einem $\varepsilon > 0$ ist (beachte, dass wir in diesem Beweis die geforderte Steuerbarkeit des Korollars 5.2. ebenfalls lediglich dahingehend verwenden). Ist $R_a = X(a)$ regulär, so benötigen wir nach Lemma A.8. die Steuerbarkeit für die Positivität nicht. In dem Fall der Nichtnegativität wird jedoch in unserem Beweis die Steuerbarkeit benötigt.

(iii) Ist $R_b = 0$, d.h. $x(b) = 0$, so sind die Bedingungen (iii), (iii*) und (iv*) stets erfüllt. In diesem Fall vereinfachen sich also die Sätze zur Positivität und Nichtnegativität. Ist zusätzlich $R_a = I$, so benötigen wir nach (ii) die Steuerbarkeit nicht und Satz 6.2. stimmt mit [[6], Chapter VII, Theorem 6.2] überein.

(iv) Das Lemma 2.9. und Proposition 6.1. geben eine Konstruktion der speziellen konjugierten Basis (X_a, U_a) an, so dass die Sätze zur Definitheit des quadratischen Funktionales $\mathcal{F}$ nicht nur theoretischen, sondern auch praktischen Nutzen haben.

A Hilfsmittel

Definition A.1.
Eine auf einem Intervall $\mathcal{I} = [a,b]$ definierte reelle Funktion f ist *absolut stetig auf* $\mathcal{I}$, wenn es zu jedem $\varepsilon > 0$ eine Zahl $\delta > 0$ existiert, so dass

$$\sum_{\nu=1}^{n} |f(\beta_\nu) - f(\alpha_\nu)| < \varepsilon$$

für jedes n und jedes disjunkte System von Segmenten $(\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n)$ aus $\mathcal{I}$ gilt, deren Längen der Ungleichung

$$\sum_{\nu=1}^{n} (\beta_\nu - \alpha_\nu) < \delta$$

genügen.

Satz A.2.
Wenn f eine reelle auf $\mathcal{I} = [a,b]$ absolut stetige Funktion ist, dann ist f in fast allen Punkten von $\mathcal{I}$ differenzierbar und

$$f(t) - f(a) = \int_a^t f'(\xi) d\xi, \quad a \leq t \leq b.$$

Beweis
Siehe [[7], Satz 7.20]. □

Satz A.3. (Existenz- und Eindeutigkeit von Systemen linearer Differentialgleichungen)
Es seien $A(t) \in \mathbb{R}^{n \times n}$ und $b(t) \in \mathbb{R}^n$ stückweise stetige Funktionen auf $\mathcal{I} = [a,b]$. Dann besitzt das System linearer Differentialgleichungen

$$\dot{x} = A(t)x + b(t), \; x(t_0) = x_0 \text{ für ein } t_0 \in \mathcal{I}$$

eine eindeutige absolut stetige Lösung $x(t) \in \mathbb{R}^n$ auf $\mathcal{I}$.

Beweis
Betrachte zunächst den Fall, dass $A(t)$ und $b(t)$ stetige Funktionen sind. Dieser Fall ist wohlbekannt und daher möchten wir auf den Beweis in [[6], Chapter I, Theorem 5.1] verweisen. Die dort benötigte Lipschitz Bedingung ist aufgrund der Linearität erfüllt, welche wir mit den Bezeichnungen des Buches [6] kurz nachrechnen möchten:

$$\|A(t)(z-y)\| \leq \|A\|\|z-y\| \text{ für alle } t \in \mathcal{I} \text{ und allen stetigen Funktionen } y, z.$$

Da eine stückweise stetige Funktion auf einem kompakten Definitionsbereich beschränkt ist, folgt die Endlichkeit von $\|A\|$ und damit die Behauptung für den stetigen Fall.
Seien nun $A(t)$ und $b(t)$ stückweise stetig für $t \in \mathcal{I}$, so gibt es eine Partition $a = t_0 < t_1 < \ldots < t_m = b, m \in \mathbb{N}$, so dass $A(t)$ und $b(t)$ stetige Funktionen auf dem offenen Intervall $(t_{\nu-1}, t_\nu), \nu = 1, \ldots, m$ sind und daher zu stetigen Funktionen auf dem kompakten Intervall $[t_{\nu-1}, t_\nu]$ ergänzt werden können. Nach dem obigen stetigem Fall gibt es also in jedem Teilintervall $[t_{\nu-1}, t_\nu], \nu = 1, \ldots, m$, eine eindeutige stetige Funktion $x(t)$ mit

$$\dot{x} = A(t)x + b(t), x(t_{\nu-1}) = x_{\nu-1},$$

wobei wir $x_\nu := x(t_\nu)$ setzen. Die so gewonnene Funktion erfüllt die Behauptung. □

Lemma A.4. (Variation der Konstanten)
Es sei $\mathcal{I} = [a, b]$ ein nicht-entartetes Intervall, $A(t) \in C_s(\mathcal{I}), x$ eine absolut stetige Funktion auf $\mathcal{I}$ und Φ eine beliebige Fundamentalmatrix von $\dot{x} = A(t)x,\ t \in \mathcal{I}$. Dann ist die Lösung der Differentialgleichung

$$\dot{x} = A(t)x + b(t), x(t_0) = x_0,\ t \in \mathcal{I}$$

für ein $t_0 \in \mathcal{I}$ gegeben durch

$$x(t) = \Phi(t)\Phi(t_0)^{-1}x_0 + \int_{t_0}^{t} \Phi(t)\Phi(s)^{-1}b(s)\, ds,\ t \in \mathbb{R}.$$

Beweis
Wir verifizieren die Behauptung mittels Differentiation:

$$\begin{aligned}
\dot{x}(t) &= \dot{\Phi}(t)\Phi(t_0)^{-1}x_0 + \dot{\Phi}(t)\int_{t_0}^{t} \Phi(s)^{-1}b(s)\, ds + \Phi(t)\Phi(t)^{-1}b(t) \\
&= A(t)\Phi(t)\Phi(t_0)^{-1}x_0 + A(t)\Phi(t)\int_{t_0}^{t} \Phi(s)^{-1}b(s)\, ds + b(t) \\
&= A(t)x(t) + b(t).
\end{aligned}$$

Ferner gilt $x(t_0) = x_0$. Damit ergibt sich aus Satz A.3. die Behauptung. □

Lemma A.5.
Es sei $X(t) \in \mathbb{R}^{n\times n}$ regulär und differenzierbar für $t \in \mathcal{I}$. Dann ist $X^{-1}(t)$ für $t \in \mathcal{I}$ differenzierbar und es gilt:

$$\frac{d}{dt}X^{-1}(t) = -X^{-1}(t)\dot{X}(t)X^{-1}(t).$$

Beweis
Nach dem Laplaceschen Entwicklungssatz gilt

$$X^{-1}(t) = \frac{1}{\det X(t)} \left((-1)^{\mu+\nu} \det \left(\begin{array}{cc|c} x_{11}(t) & \dots & x_{1n}(t) \\ \vdots & & \vdots \\ \hline x_{n1}(t) & \dots & x_{nn}(t) \end{array} \right) \right)$$

und damit ist X^{-1} differenzierbar für $t \in \mathcal{I}$ nach den Differentiationsregeln.
Ferner gilt:

$$0 = \frac{d}{dt} I = \frac{d}{dt} X(t) X^{-1}(t) = \dot{X}(t) X^{-1}(t) + X(t) \frac{d}{dt} X^{-1}(t).$$

Hieraus ergibt sich die Behauptung. □

Satz A.6.
Gegeben sei eine reelle und symmetrische $n \times n$-Matrix Q mit (reellwertigen) Eigenwerten $\mu_1, \dots, \mu_n$ und zugehörigen Eigenvektoren $s_1, \dots, s_n$. Sei $\nu \in \{1, \dots, n\}$ und S der von $s_1, \dots, s_\nu$ aufgespannte Untervektorraum von $\mathbb{R}^n$. Ferner bezeichne $R(d) = \frac{d^T Q d}{\|d\|^2}$ für $d \in \mathbb{R}^n \backslash \{0\}$ der *Rayleigh-Quotient* von Q und V ein Untervektorraum des $\mathbb{R}^n$ der angegebenen Dimension ν.
Dann gilt:

(i) $\mu_\nu = \max\{R(d) | d \in S \backslash \{0\}\}$, falls $\mu_1 \leq \ldots \leq \mu_n$.

(ii) $\mu_\nu = \min\limits_{\dim V = \nu} \max\{R(d) | d \in V \backslash \{0\}\}$, falls $\mu_1 \leq \ldots \leq \mu_n$.

(iii) $|\mu_\nu| = \max\limits_{\dim V = n-\nu+1} \min\{\frac{\|Qd\|}{\|d\|} | d \in V \backslash \{0\}\}$, falls $|\mu_1| \leq \ldots \leq |\mu_n|$.

(iv) $|\mu_\nu| = \min\limits_{\dim V = \nu} \max\{\frac{\|Qd\|}{\|d\|} | d \in V \backslash \{0\}\}$, falls $|\mu_1| \leq \ldots \leq |\mu_n|$.

Beweis
Die Teile (i),(ii),(iv) stehen direkt in [[4], Proposition 3.2.1] (siehe auch [2] und [9]), und auch der Teil (iii) ergibt sich, da dieser äquivalent zu $|\mu_\nu| = \max\limits_{\dim V = \nu} \min\{\frac{\|Qd\|}{\|d\|} | d \in V \backslash \{0\}\}$, falls $|\mu_1| \geq \ldots \geq |\mu_n|$ ist. □

Korollar A.7.
Es sei Q eine reelle und symmetrische $n \times n$-Matrix. Dann ist $|\mu| \leq \|Q\|$ für alle Eigenwerte μ von Q.

Beweis
Nach Satz A.6. gilt für alle Eigenwerte μ von Q:

$$|\mu| \leq \max_{d \in \mathbb{R}^n \backslash \{0\}} \left| \frac{d^T Q d}{\|d\|^2} \right| \leq \max_{d \in \mathbb{R}^n \backslash \{0\}} \frac{\|Q\| \|d\|^2}{\|d\|^2} = \|Q\|.$$

□

Lemma A.8.
Ist eine auf einem Intervall $\mathcal{I}$ stetige $n \times n$-matrixwertige Funktion $M(t)$ regulär in einem Punkt $t_0 \in \mathcal{I}$, so gibt es ein $\varepsilon > 0$, so dass $M(t)$ regulär auf dem Bereich $(t_0 - \varepsilon, t_0 + \varepsilon)$ ist.

Beweis
Da $M(t)$ stetig auf $\mathcal{I}$ ist, ist auch ihre Determinante auf $\mathcal{I}$ stetig. Mit der Charakterisierung der Regularität von $M(t)$ über das Nicht-Verschwinden der Determinante $\det M(t) \neq 0$ folgt nun unmittelbar die Behauptung. $\square$

Lemma A.9.
Zu jeder gegebenen quadratischen Matrix $M \in \mathbb{R}^{n\times n}$ gibt es ein reelles $\alpha > 0$ mit $\|M^T Mc\| \geq \alpha \|Mc\|$ für alle $c \in \mathbb{R}^n$.

Beweis
Die Matrix $M^T M$ ist reell und symmetrisch und demzufolge sind ihre Eigenwerte reell und es gibt zugehörige Eigenvektoren, die reell, linear unabhängig und orthogonal zueinander sind. Da $M^T M$ zudem nichtnegativ definit ist, sind die Eigenwerte $\mu_1, \dots, \mu_n$ nichtnegativ, wobei wir mit $\mu_\nu > 0$ für $\nu = 1, \dots, r = \mathrm{rg} M^T M$ die positiven und mit $\mu_\nu = 0$ für $\nu = r+1, \dots, n$ die Eigenwerte bezeichnen, die Null sind. Mit $a := \min\{\mu_1, \dots, \mu_r\}$ bezeichnen wir den kleinsten positiven Eigenwert von $M^T M$. Es seien c_ν Eigenvektoren zum Eigenwert μ_ν, d.h. $M^T M c_\nu = \mu_\nu c_\nu, \nu = 1, \dots, n$, mit $\|c_1\| = \ldots = \|c_n\| = 1$, so dass $c_1, \dots, c_n$ eine Orthonormalbasis des $\mathbb{R}^n$ bilden. Jeder Vektor $c \in \mathbb{R}$ findet damit eine Darstellung $c = \sum\limits_{\nu=1}^{n} \beta_\nu c_\nu$. Betrachten wir den Vektor $\beta = (\beta_1, \dots, \beta_n)^T$, so stellen wir aufgrund der Orthonormalität der $c_\nu, \nu = 1, \dots, n$, fest, dass $\|c\| = \|\beta\|$ und

$$\|M^T Mc\| = \|\sum_{\nu=1}^{r} \mu_\nu \beta_\nu c_\nu\| = |\sum_{\nu=1}^{n} \mu_\nu^2 \beta_\nu^2|^{\frac{1}{2}} \geq a\|\beta\| = a\|c\|.$$

Wir folgern unmittelbar

$$\|Mc\| \leq \|M\|\|c\| \leq \frac{\|M\|}{a} \|M^T Mc\|$$

und mit $\alpha = \frac{a}{\|M\|}$ die Behauptung. $\square$

Literatur

[1] Baur, G.: „Variationsrechnung und Kontrolltheorie“, Vorlesungsmanuskript, Ulm 2000.

[2] Courant, R., und Hilbert, D.: „Methods of Mathematical Physics I“, Wiley-Interscience, New York 1953.

[3] Heuser, H.: „Lehrbuch der Analysis Teil 1“, 12. Auflage, B.G. Teubner, Stuttgart 1998.

[4] Kratz, W.: „Quadratic Functionals in Variational Analysis and Control Theory“, Akademie Verlag, Berlin 1995.

[5] Kratz, W.: „Definiteness of Quadratic Functionals“, Analysis 23, 163-183, 2003.

[6] Reid, W.T.: „Ordinary Differential Equations“, Wiley, New York 1971.

[7] Rudin, W.: „Reelle und Komplexe Analysis“, Oldenbourg Verlag, München, Wien, 1999.

[8] Sagan, H.: „Introduction to the calculus of variations“, McGraw-Hill, New York 1969.

[9] Wilkinson, J.H.: „The algebraic eigenvalue problem“, Clarendon Press, Oxford 1965.

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig angefertigt habe. Es wurden keine außer den angegebenen Quellen verwendet und die aus diesen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Ulm, im September 2003 Markus Wahrheit