

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen
Gesamtfakultät
der
Ruprecht-Karls-Universität
Heidelberg

vorgelegt von
Diplom-Mathematiker
THORSTEN STOSSMEISTER
aus Karlsruhe
2002

Optimales Laufen

Gutachter:

Prof. Dr. Dr. h.c. Hans Georg Bock

Prof. Dr. Dr. h.c. mult. Willi Jäger

Danksagung

Diese Arbeit ist am Interdisziplinären Zentrum für Wissenschaftliches Rechnen der Universität Heidelberg (IWR) entstanden. Daher gilt meiner erster Dank der Universität und der Fakultät für Mathematik und Informatik selbst.

Mein besonderer Dank gilt natürlich meinen Betreuern Prof. Dr. Dr. h.c. Hans-Georg Bock und Dr. Johannes Schlöder, die mich zum Thema dieser Arbeit hinführten und mich durch zahlreiche Diskussionen inspirierten und leiteten. So entgegnete Prof. Bock meinen Zweifeln an der Mathematik mit seinen Zweifeln an meinem Coding und behielt schließlich Recht.

Weiter möchte ich herzlichsten Dank an Prof. Dr. Dr. h.c. mult. Willi Jäger sagen, der meine Arbeit im Rahmen des Graduiertenkollegs „Modellierung und Wissenschaftliches Rechnen in Mathematik und Naturwissenschaften“ ebenfalls betreute. Sein aufgeschlossenes Interesse und seine Freude am Thema waren von unschätzbbarer Hilfe.

Sollte ich mit dieser Arbeit weiter geblickt haben können als der wissenschaftliche Horizont es bisher zuließ, so lag dies einzig daran, daß ich auf den Schultern von Riesen stehen durfte. So danke ich Dr. Michael Winckler und Dr. Reinhold von Schwerin für viele Anregungen, Diskussionen und für eine Integrator-Bibliothek, die in ihrem klaren und verständlichen Programmierstil ein Leuchtfeuer in ansonsten trüber Nacht darstellt. Gleichmaßen geht mein Dank an Dr. Daniel Leineweber und Andreas Schäfer, deren Programm MUSCOD-II nicht nur mächtig, sondern auch vollkommen fehlerfrei ist. Dr. Katja Mombaur möchte ich ebenso danken, ihre Arbeit an verwandten Fragestellungen hat mein Blick auf die Probleme geschärft und bereichert.

Speziell möchte ich mich für die Zusammenarbeit mit Christian Kraus bedanken. Sein Ideenreichtum und seine Vorschläge haben sich an vielen Stellen in dieser Arbeit niedergeschlagen, und es waren seine Anregungen, die mich in Berührung mit der algebraischen Geometrie brachten.

Jan Simon war stets mein Retter in der Not, wenn es darum ging, Daten in absurder Zeit zu visualisieren. Evelin Hund möchte ich danken für die Heirat mit Jan Simon. Beide haben die Zumutung des Korrekturlesens auf ihre Schultern geladen und die Last auf meinen damit reichlich verringert.

Margret Rothfuß war stets auch die gute Seele der Arbeitsgruppe, Thomas Klöpfer hielt die Rechner am Laufen, beiden sei für ihren Beitrag zum Gelingen dieser Arbeit Anerkennung gezollt.

Meinem Freund Bernhard Werner danke ich für die Liebe, Geduld und Unterstützung durch die Jahre meines Studiums hindurch und während vieler Abende, die vom Computer vereinnahmt wurden.

Schließlich möchte ich mich bei meinen Eltern, Heinz und Hannelore Stoßmeister, bedanken, für ihre Liebe und Hingabe. Nur durch ihre Kraft ist all das hier möglich geworden.

MEINEN ELTERN
HEINZ UND HANNELORE STOSSMEISTER

Einleitung

Zentraler Gegenstand der vorliegenden Arbeit ist das Laufen. Das Laufen umgibt uns überall, ist Bestandteil unseres Alltags und wird dennoch, oder gerade deswegen, selten beachtet. Der Mensch bewegt sich durch den Vorgang des Laufens fort. Bei archäo-anthropologischen Forschungen ist ein Kriterium zur Unterscheidung zwischen Mensch und Menschenaffe der aufrechte Gang. Doch brauchen Menschen lange, das Laufen zu erlernen. Was so einfach scheint, stellt sich für den neugeborenen Menschen als zunächst unmöglich heraus¹. Das Erlernen dauert Jahre, und der Tag, an dem ein Mensch selbständig seine ersten Schritte macht, gehört zu den bedeutendsten Momenten im Leben (auch, wenn dieser bewusst nur von den Eltern wahrgenommen wird).

Welch ungeheure Bedeutung dem Laufen wirklich zukommt, zeigt sich dann an einzelnen Vorgängen, welche singulär aus dem Grau des Alltäglichen herausragen. So erlangt das Laufen bzw. Rennen wieder Aufmerksamkeit im Sport. Es ist eine klassische Disziplin und neue Rekorde, die vorhergesagte Grenzen des biomechanisch Machbaren überschreiten, erregen weltweites Interesse.

Doch häufig findet das Laufen erst dann wieder Beachtung, wenn es nicht oder nicht mehr funktioniert, wenn Mensch oder Tier durch Alter, Krankheit und/oder Behinderung die Fähigkeit des Laufens verlieren. Hier zeigt sich, wie sehr das tägliche Leben vom Laufen abhängig ist. Umständlich ist vor allem die Beschränkung, fahren zu müssen, wo Laufen angebracht wäre. Der Hauptgrund der Natur für die Erfindung des Laufens dürfte der chronische Mangel an natürlich gewachsenen Straßen sein. Das Fahren benötigt einen kontinuierlichen und geraden Untergrund für die Räder. Diese Situation liegt sogar in unseren Wohnungen selten vor (Treppen), erst recht nicht in freier Wildbahn. Mindestens 50% des irdischen Festlandes sind für fahrende Systeme ungeeignet². Und nicht nur auf der Erde: war der Bewegungsradius des Mars-Mobils Sojourner der Pathfinder-Mission nicht furchtbar klein, bei allem Respekt der technologischen Leistung gegenüber³? Hier darf man fragen: wäre eine solche Sonde nicht viel mobiler gewesen, wenn sie hätte laufen können? Kleine und mittlere Steine, Löcher oder Spalten, schiefe und schräge Ebenen können von laufenden Mechanismen leicht bewältigt werden, fahrende haben dabei die größten Schwierigkeiten.

Insbesondere in der medizinischen Prothetik und der Robotik werden große Anstrengungen unternommen das Laufen zu ergründen.⁴ Die Robotik ist dabei zunächst von statisch stabilen Systemen ausgegangen. Ganz nach dem Vorbild der Insekten besitzen diese mehrere Beine und haben damit die Möglichkeit, immer mindestens ein statisch stabiles

¹Man vergleiche diesbezüglich Menschen und Pferde. Pferde benötigen ca. eine halbe Stunde, Menschen ca. 1,5 Jahre, um laufen zu lernen.

²siehe auch RAIBERT[57]

³Der ausgesetzte Rover Sojourner der Mars Mission Pathfinder legte insgesamt durch 114 Bewegungskommandos 52m zurück, siehe SCIENCE MAGAZINE[62]

⁴Die ersten Untersuchungen können sogar bis zu dem berühmten russischen Mathematiker TSCHEBYSCHEFF datiert werden, siehe LUCAS[44].

Dreibein auf dem Boden zu haben.⁵ Damit kann die Maschine nicht umfallen, was angesichts der teuren Geräte durchaus verständlich ist. So ist die Maschine insbesondere zum langsamen Gehen befähigt (was ja der Startpunkt bei einer sich erst entwickelnden Antriebstechnik sein muß)⁶. Speziell beim Zwei-Beiner ist diese statische Stabilität aber nicht mehr gegeben. Hier gibt es Phasen, in denen der Schwerpunkt nicht gestützt ist. Störungen können sich fatal bemerkbar machen und die Bewegung kann auch nicht beliebig langsam ablaufen. Mit diesen Schwierigkeiten muß die medizinische Forschung z. B. Querschnittsgelähmter kämpfen: einerseits muß Stabilität gewährt werden, auch gegen Störungen und daraus folgenden Verletzungen, andererseits soll eine natürliche Fortbewegung ermöglicht werden.⁷

Doch wie sieht die natürliche Fortbewegung beim Menschen aus? Der Mensch ist durch seinen anatomischen Aufbau befähigt, fast jede seiner Bewegungen auf unterschiedliche Art durchzuführen. So kann das gleiche Ziel, das nicht immer im Laufen bestehen muß sondern durchaus auch das Fangen eines Balls o. ä. sein kann, ruckartig, umständlich, elegant, hektisch oder fließend erreicht werden. Sehr schön kann man dies beobachten, wenn ein neuer Bewegungsvorgang erlernt werden muß. Als Beispiel kann man das Schlittschuhlaufen nennen. Die Bewegungen sind anfangs ungeschickt, zackig, eckig und unbeholfen. Dies verbessert sich zusehends bis ein „runder“ Bewegungsablauf entsteht, der sich dann nicht mehr wahrnehmbar verändert.

Doch worin unterscheidet sich ein „runder“ Bewegungsablauf von einem „eckigen“?⁸ Hier kann man nun eine Vermutung äußern: natürlich wirkende, „runde“ Bewegungsabläufe sind in irgendeiner Hinsicht optimal. Wenn aus einem Bündel an Bewegungs-Möglichkeiten nur eine Bewegungsform verwirklicht wird, liegt die Vermutung nahe, daß diese sich vor den anderen dadurch auszeichnet, bezüglich eines bestimmten, jedoch nicht näher bekannten Optimalitätskriteriums besser zu sein.

Diese Vermutung kann noch durch ein Indiz gestützt werden: so gab es in den letzten Jahren deutliche Fortschritte in der Tricktechnik des Films, den Visual Special Effects⁹. Es gab einige Produktionen, die das Leben und Verhalten ausgestorbener Tiere zum Inhalt hatten, speziell der Dinosaurier. Mit großem Aufwand wurden die Bewegungen dieser Tiere simuliert und animiert. Dennoch gibt es in diesen aufwendigen Werken einige Stellen, an denen auch ein in Paläoontologie unbedarfter Zuschauer sich genötigt sieht zu sagen: „Die Bewegung sieht komisch aus.“ Obwohl kein Mensch einen lebenden Dinosaurier je in Bewegung gesehen hat, sind wir in der Lage zu sagen, daß er sich so bestimmt nicht bewegt hätte, selbst wenn die Bewegung physikalisch korrekt aussieht. Kann es sein, daß in einer solchen Situation intuitiv erkannt wird, daß der Vorgang nicht dem „üblichen“ Optimalitätskriterium genügt?¹⁰

Welches Optimalitätskriterium könnte das nun sein? Aus den vorliegenden Bewegungen läßt sich umgekehrt kein eindeutiges Kriterium bestimmen, denn es lassen sich beliebig artifizielle Kriterien aufbauen, die diese Bewegungen als Lösung haben. Hier kann

⁵Von erheiterndem Design sind beispielsweise die Entwürfe in PAI ET AL.[54], die sich in ihrem Aussehen an den platonischen Körpern orientieren.

⁶Umfangreiche Literaturverweise in das Gebiet der Robotik finden sich z. B. bei BERNS[5], beim MIT[47] oder auch bei RAIBERT UND SUTHERLAND[58].

⁷Bezüglich der Ermittlung stabiler Laufbewegungen bei nicht-rückgekoppelten Systemen gibt es umfangreiche Untersuchungen von MOMBAUR ET AL.[48]-[52].

⁸Von GAO[23] sind für ein zweidimensionales Modell mehrphasige Bewegungen gesucht worden, bei denen der Fuß stoßfrei aufsetzt, also Bewegungen, die gewisse Randbedingungen erfüllen.

⁹Einen Überblick über ältere Arbeiten findet sich bei MAGNENAT-THALMANN[45].

¹⁰Möglicherweise spielen in diesem Zusammenhang die Bewegungen der „Schwabbelmassen“ eine wichtige Rolle, wie sie z. B. von RUDER ET AL[59] untersucht worden sind und werden.

man die Vermutung äußern, daß das Optimalitätskriterium ein einfaches ist, also z. B. die energieärmste, zeitkürzeste oder beschleunigungsminimierendste Bewegung¹¹. Ausgehend von einer solchen Vermutung kann man nun versuchen, Bewegungen zu bestimmen, die optimal sind bezüglich eines solchen Kriteriums, und die Resultate mit vorliegenden Bewegungen vergleichen.

Dabei wird sich aller Voraussicht nach kein eindeutiges Optimalitätskriterium ergeben: die Bewegung eines Menschen, der durch einen Wald spaziert, dürfte wohl so etwas wie die energieärmste Bewegung sein. Aber sobald der gleiche Mensch im Wald auf ein Wildschwein trifft, wird sich seine Bewegung in eine zeitkürzeste zur Erreichung eines bestimmten Zielpunktes verwandeln. Das Kriterium hängt stark von den Umständen ab, vielleicht sogar noch von den Möglichkeiten, die aufgrund der Anatomie gegeben sind. Und über allem thront noch das Zielfunktional der Evolution, letztendlich aus den gegebenen Umständen das Beste zu machen und die Gene überleben zu lassen.

Die vorliegende Arbeit hat sich zum Ziel gesetzt, ein kleines Stück zur Beantwortung dieser Fragen beizutragen. Genauer soll untersucht werden, wie die bereits vorliegenden Techniken von Modellierung, Simulation und Optimierung eines einfachen mechanischen Systems verschmolzen werden können, um dreidimensionale optimierte Laufbewegungen zu erhalten. Dabei soll insbesondere keinerlei Vorwissen über mögliche Lösungen in die Startwerte einfließen. Laufmuster und Periodizität sollen sich durch die Optimierung ergeben. Das Optimalitätskriterium soll dabei nicht auf minimierte quadratisch-aufintegrierte Steuerungen beschränkt sein, sondern frei gewählt werden können.

Auch soll das System nicht sensorisch rückgekoppelt sein, es ist also ein „open-loop“-System. Es wird also implizit davon ausgegangen, daß die mechanischen Eigenschaften des Systems sowie der Umgebung a priori bekannt sind und perfekt modelliert werden können. Dadurch konzentriert sich die Untersuchung nicht auf Probleme der Steuerung, sondern auf die Natur der Bewegung selbst.¹²

Insbesondere soll dabei ein Modell eines laufenden Mechanismus mit nur zwei Beinen betrachtet werden, wodurch keine ständige statische Stabilität herrscht.

Dabei ergeben sich folgende Fragestellungen:

- (1) Wie lautet eine geschickte Modellierung des Systems? Dabei muß sowohl der Boden-Fuß-Kontakt modelliert werden als auch das ganze mechanische System an sich. Die Formulierung soll möglichst so sein, daß Korrekturen oder Erweiterungen nicht zu einer totalen Neuberechnung führen, sondern größtmögliche Wiederverwendbarkeit beinhalten. Am besten wäre es, wenn die Formulierung so gewählt werden kann, daß ein solches Modell auch von einem Algorithmus bzw. Computerprogramm zusammengestellt werden kann.
- (2) Was für die Modellierung am geschicktesten ist, führt bei der Simulation, d. h. der Lösung der Bewegungsgleichungen, zu numerischen Effekten, die geeignet behandelt werden müssen. So wird sich das Problem ergeben, Gleichungsnebedingungen der Bewegung konsistent erfüllen zu müssen. Insbesondere während des Optimierungsvorgangs ist diese Bedingung aber nicht erfüllt. Gleichzeitig führt die oben erwähnte sehr bequeme Modellierung dazu, daß das Modell sehr „groß“ wird (gemessen in der Anzahl der auftretenden Variablen).

¹¹Bei DENOTH[16] finden sich als Vorschläge für das Zielfunktional z. B. „Die Summe aller Muskelkräfte [im Fuß] ist minimal“ oder „Die Summe aller Muskelspannungen ist minimal“. Zu diesen Vorschlägen wurden aber keine Berechnungen durchgeführt.

¹²Zu den Vorteilen der Untersuchung eines open-loop-Systems siehe auch MOMBAUR[48]

- (3) Schließlich interessieren dann die Ergebnisse der Optimierung. Welche Bewegungen resultieren ausgehend vom Startwert einer Null-Steuerung¹³? Entsteht von selbst ein periodischer Ablauf? Was verändert sich an den Bewegungen beim Wechsel des Zielfunktional?

In Kapitel I wird versucht, die geschickteste Modellierung des Problems zu finden. Hier werden sich ganz von selbst die Vorteile einer Formulierung als differential-algebraisches System¹⁴ in *Natürlichen Koordinaten*¹⁵ zeigen. Es wird erläutert, auf welche Weise dann die Bewegungsgleichungen dieses Modells gefunden und gelöst werden können. Das übergeordnete Problem ist dabei ein sehr allgemeines nichtlineares Optimierungsproblem. Dieses wird ebenfalls formuliert und Lösungsmethoden genannt. Außerdem wird die Aufteilung des Modells auf Integration und Optimierung geklärt. So sind z. B. die mechanischen Zwangsbedingungen dem Verfahren zur numerischen Lösung mitzuteilen (z. B. um Projektionen zur Vermeidung von Drift durchführen zu können), die Konsistenzbedingungen dürfen aber nur dem Optimierungsverfahren zugeschlagen werden, da ansonsten gemeinsam mit konsistenten Anfangswerten und Stetigkeitsbedingungen Redundanz der Gleichungen auftritt.

Die oben erwähnte Formulierung der Bewegungsgleichungen des Modells als DAE führt zum Problem, daß die Gleichungsnebenbedingungen während des Optimierungsvorgangs möglicherweise verletzt werden. Kapitel II untersucht verschiedene Relaxierungsmethoden, die dieses Problem lösen sollen. Die übliche Relaxierungsmethode, die lineare Relaxierung, führt bedingt durch Geschwindigkeitsinkonsistenzen sehr schnell zum Auftreten von Singularitäten. Dies kann durch orthogonale Relaxierung vermieden werden. Es wird ebenfalls eine Behandlung von Ortsinkonsistenzen dargelegt. Als bestes, allerdings nur lokales, Verfahren wird die Methode der initialen Projektion eingeführt. Diese beseitigt sowohl Orts- als auch Geschwindigkeitsinkonsistenzen und hat zusätzlich in Bezug auf die Optimierung einen unschätzbaren Vorteil. Sie reduziert die Anzahl der zu berechnenden Richtungsableitungen für in redundanten Koordinaten formulierte Modelle auf das notwendige Minimum, welches auch in der Formulierung in Minimalkoordinaten notwendig ist, nämlich Anzahl der Freiheitsgrade. Diese initiale Projektion macht die Optimierung großer Modelle in redundanten Koordinaten erst in annehmbarer Zeit durchführbar.

In Kapitel III wird ein neuer Algorithmus vorgestellt, der sich eher als Nebenprodukt dieser Arbeit ergeben hat, sich aber inhaltlich an die Untersuchung des Auftretens von Singularitäten in Kapitel II anschließt. Der Algorithmus der *Aufblasung* ermöglicht es, unter bestimmten, häufig vorliegenden Voraussetzungen durch eine Singularität auf einer Mannigfaltigkeit (genauer gesagt: einer Varietät) hindurchzuintegrieren (sie zu durchqueren). Dabei kann der Algorithmus als eine Methode aufgefasst werden, ein singularitätsbehaftetes Modell automatisch umzuformulieren, um die Singularität lokal zu beseitigen. Die zugrunde liegende Methode, der blow-up, ist dabei der algebraischen Geometrie schon lange bekannt, aber in diesem Spezialfall der Durchquerung konnte ein sehr einfacher Beweis des Verfahrens gefunden werden. Die Fähigkeiten des Algorithmus werden in Kapitel IV demonstriert. Zunächst wird als Beispiel der entartete Kurbeltrieb singularitätsfrei umformuliert, dann wird das Beispiel der liegenden Acht vorgeführt. Hier muß sich ein Massepunkt auf einer Mannigfaltigkeit bewegen, die das Aussehen des ∞ -Zeichens hat. Dabei

¹³Im Gegensatz z. B. zu MOMBUR[48] wird in dieser Arbeit die Abfolge der Kontaktphasen nicht a priori vorgegeben.

¹⁴Engl. DAE, d. i. einfach formuliert eine Differentialgleichung mit zusätzlichen Bedingungen in Gleichungsform.

¹⁵siehe z. B. GARCÍA DE JALÓN[25]

ermöglicht die Aufblasung das problemlose Durchqueren der Singularität im Kreuzungspunkt.

In Kapitel V wird das Modell eines einfachen, dreidimensionalen zweibeinigen Mechanismus formuliert. Er besteht aus punktförmigen Massen, die als Füße, Knie, Hüfte und Schultern gedeutet werden. Auf den Mechanismus wirken als Kräfte Schwerkraft, Drehmomente als Antriebskräfte sowie die Boden-Reaktions-Kraft. Die Boden-Kräfte wirken dabei nur auf die beiden punktförmigen Füße und sind über Schaltfunktionen abgebildet (d. h. die Funktion der Kräfte ist nicht überall differenzierbar und wird beim Auftreten gewisser Ereignisse umgeschaltet, wie z. B. wenn der Fuß den Boden erreicht).

In Kapitel VI werden dann optimale Bewegungen dieses Mechanismus gesucht. Z. B. soll er sich in gegebener Zeit soweit wie möglich in eine bestimmte Richtung fortbewegen. Hier werden zwar Lösungen gefunden, sie entsprechen aber nicht den Erwartungen (keinerlei Andeutung einer sich ergebenden Periodizität, Schwierigkeiten bei der Fortsetzung über einen längeren Zeitraum). Aber nachdem eine stabilisierende Bedingung (nämlich: Oberkörper soll senkrecht bleiben) hinzugefügt wird, ergeben sich Bewegungen, die schließlich auch in zumindest periodenähnliche Abläufe münden. Diese können wiederum genommen werden, um als Starttrajektorie für periodische Bewegungen zu dienen. Damit werden dann auch periodische Bewegungen gefunden und Optimierungen mit unterschiedlichem Zielfunktional gestartet. Um noch einmal den Komfort bezüglich der Formulierung als Mehrkörpersystem in redundanten Koordinaten zu betonen, wird im letzten Abschnitt das zweibeinige Modell kurzerhand zu einem vierbeinigen gemacht und damit zwei Optimierungsläufe durchgeführt.

Schließlich hat sich zusammenfassend gezeigt, daß die vorgestellten Methoden und Vorgehensweisen in der Lage sind, in akzeptabler Rechenzeit und ohne Vorwissen über die Lösung, optimale Laufbewegungen eines dreidimensionalen Zweibeiners bei unterschiedlichen Randbedingungen, wie z. B. Periodizität, zu ermitteln. Dabei muß das Modell nicht in Minimalkoordinaten, sondern darf in natürlichen (bzw. im Falle von Punktmassen dazu äquivalenten kartesischen) Koordinaten mit Zwangsbedingungen formuliert werden, was einen extremen Komfort bei Änderungen am Modell selbst darstellt.

Die sich ergebenden Bewegungen sind zum Teil sehr überraschend. Das liegt aber weniger an den Methoden als eher an der noch nicht genügend menschenähnlichen Anatomie des Mechanismus und der intuitiv nicht erkennbaren Ausnutzung der sich damit ergebenden Freiheiten. So sind einige der Optimalitätsbedingungen nur formuliert worden, um die Ausnutzung aller Bewegungsmöglichkeiten einzuschränken, wie z. B. das Durchrotieren der Beine um das Knie.

Die Ergebnisse wurden dabei auf einem PC berechnet, der nach heutigen Maßstäben schon wieder als sehr leistungsschwach bezeichnet werden darf (PC, Pentium II, 266 MHz). Es steht mit diesen Methoden¹⁶ aber nun ein Weg offen, größere und kompliziertere Mechanismen nach dem Baukastenprinzip zusammenzubauen und sich von ihren harmonischen Laufbewegungen überraschen zu lassen.

¹⁶Insbesondere, wenn man die hier gar nicht benutzte Möglichkeit verwendet, den Integrationsprozess, der ca. 99% der Rechenzeit beansprucht, komplett zu parallelisieren, siehe GALLITZENDÖRFER[22].

Inhaltsverzeichnis

Danksagung	III
Einleitung	V
Kapitel I Problemformulierung	1
§1 Modellierung.....	1
§2 Bewegungsgleichungen und Numerische Integration.....	4
§3 Optimierung	9
§4 Aufteilung.....	12
Kapitel II Relaxierung	17
§1 Singularitäten durch Geschwindigkeiten.....	17
§1.1 Lineare Relaxierung	18
§1.2 Orthogonale Relaxierung	21
§1.3 Orthogonale Relaxierung in gemischten Koordinaten	28
§1.4 Orthogonale Relaxierung und Wronski-Matrix.....	32
§2 Inkonsistente Startpositionen	34
§2.1 Unbehebbarkeit von Rangverlusten	34
§2.2 Globaler Vollrang.....	36
§3 Initiale Projektion.....	40
§3.1 Motivation	40
§3.2 Initiale Projektion.....	42
§3.3 Variationsdifferentialgleichung	45
Kapitel III Aufblasung	47
§1 Problemstellung	47
§1.1 Rangverlust ohne Freiheitsgrad	47
§1.2 Rangverlust mit Freiheitsgrad	49
§2 Lösungsansätze	50
§3 t -Ordnung.....	51
§4 Die Kurvenaufblasung	55
§5 Die Polaraufblasung	65
§6 Der komplexe Fall.....	70
Kapitel IV Beispiele für Aufblasung	73
§1 Der Kurbeltrieb	73
§2 Und nun ... Physik	76
§3 Die liegende Acht	80
§4 Numerische Stabilität.....	81
Kapitel V Das Modell	87
§1 Die mechanische Formulierung	87
§1.1 Punktmassen.....	87

§1.2	Zwangsbedingungen	88
§1.3	Kräfte	89
§1.4	Initiale Projektion	93
§2	Die Optimierungsformulierung	93
§2.1	Mechanische Bedingungen	93
§2.2	Lösungsbedingungen	96
§2.3	Optimierungsbedingungen	97
Kapitel VI	Optimierungsdurchführung	99
§1	Vorbemerkung	99
§2	Initialwerte	101
§3	Erste Versuche	101
§3.1	$t_{end} = 0.1s$	102
§3.2	$t_{end} = 0.2s$	102
§3.3	$t_{end} = 0.3s$	103
§3.4	$t_{end} = 0.4s$	103
§3.5	$t_{end} = 0.5s$	104
§3.6	$t_{end} = 0.6s$	104
§3.7	$t_{end} = 0.7s$	104
§4	Zweite Versuche	104
§5	Stabilisierende Bedingungen	105
§6	Ein paar Schritte	106
§6.1	Ein erstes Vorwärtkommen	106
§6.2	Tiefgelegter Schwerpunkt	106
§6.3	Strenger tiefergelegter Schwerpunkt	106
§7	Periodische Bewegungen	107
§7.1	Erreichen einer periodischen Bewegung	107
§7.2	Laufen	108
§7.3	Steuerungsminimal	108
§7.4	Laufen bei freier Endzeit	108
§8	Vierbeinige Versuche	109
Kapitel VII	Zusammenfassung und Ausblick	127
§1	Zusammenfassung	127
§2	Vergleich	128
§3	Ausblick	130
Literaturverzeichnis	131	
Anhang	135	

KAPITEL I

Problemformulierung

Die mathematischen Voraussetzungen zur Bestimmung optimaler Bewegungen lassen sich grob in drei Aufgaben einteilen:

- (1) Erstellung eines mathematischen Modells des mechanischen Systems und Bestimmung der Bewegungsgleichungen dieses Modells,
- (2) Lösen der Bewegungsgleichungen (d. i. numerische Integration),
- (3) Optimierung der Steuerung in Hinsicht auf ein vorgegebenes Zielfunktional und gegebenen Randbedingungen.

Zur Bewältigung dieser Aufgaben existieren bereits verschiedene wissenschaftliche Software-Lösungen. Diese Arbeit verwendet die Programmpakete, die in der Forschungsgruppe „Simulation und Optimierung“ unter der Leitung von Prof. Dr. h. c. H.-G. BOCK am Interdisziplinären Zentrum für wissenschaftliches Rechnen der Universität Heidelberg entwickelt wurden. Sie bilden das sichere Fundament, auf dem diese Arbeit ruht, und sollen hier kurz vorgestellt werden.

Gemäß der obigen Einteilung wird dabei zunächst das Prinzip der Formulierung als Mehrkörpersystem vorgestellt. Die beiden Zugänge Minimalkoordinaten und redundante Koordinaten werden verglichen und insbesondere die natürlichen Koordinaten eingeführt. Im nächsten Abschnitt werden dann die Bewegungsgleichungen eines solchen Mehrkörpersystems hergeleitet. Speziell wird das Problem der Drift und seine Behandlung durch Projektion sowie das Auftreten von Schaltfunktionen genannt.

Dann wird das Optimierungsproblem allgemein formuliert und numerische Lösungsverfahren skizziert, insbesondere die Mehrzielmethode.

Schließlich wird an einem einfachen Beispiel gezeigt, wie diese drei Komponenten zusammenarbeiten sollen, um die Bewegung eines Mehrkörpersystems zu optimieren. Wesentliche Fragestellung ist hier, welchem Verfahren welche Informationen zur Verfügung gestellt werden muß.

§1 Modellierung

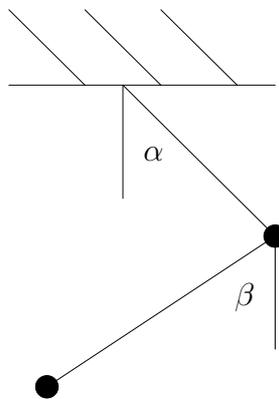
Ein lauffähiges biologisches System detailgetreu zu modellieren liegt jenseits des wissenschaftlich Realisierbaren; das Modell eines Knochens, eines Nervs, einer Sehne kann selbst bis auf chemische Ebene weiter verfeinert werden. Beim Modellieren ist es jedoch so, daß das mathematisch Machbare dazu zwingt sich zu überlegen, welches einfachst mögliche Modell die Bewegung des Laufens physikalisch leisten kann.¹

Als vielversprechende Herangehensweise bietet sich hierfür an, das System als Mehrkörpersystem (MKS oder engl. MBS) zu modellieren. Bei dieser Vorgehensweise denkt man sich den Gesamtmechanismus aus mehreren starren Teilen zusammengesetzt, wie z. B. Quader, Zylinder oder Kugeln. Verbunden werden die einzelnen Teile durch Gelenke, wie z. B. Kugel- oder Scharniergelenk. Als Aktuatoren treten Kräfte und Drehmomente auf. Mit

¹siehe z. B. METZGER[46] oder KRABBE[38] für Versuche der Modellierung auf einfacher mechanischer Ebene

dieser Herangehensweise werden u. a. verformbare Körper, komplexe Verbindungsstrukturen und komplizierte Kraftverläufe mit z. B. Hysterese ausgeschlossen. Auf diese Art und Weise ist es möglich, ein Modell nach dem Baukastenprinzip zusammenzusetzen und zu modifizieren, bis man erwarten kann, genügend Bewegungsfreiheit und Antriebsmöglichkeiten eingebaut zu haben, damit der Mechanismus in der Lage ist sich fortzubewegen. Der nächste Schritt in der Modellierung ist die wichtige Wahl der Art der Koordinaten. Hier gibt es prinzipiell zwei Möglichkeiten: Minimalkoordinaten (unabhängige) oder redundante (abhängige) Koordinaten.

- Bei Minimalkoordinaten oder unabhängigen Koordinaten gibt es so viele Variablen wie es Freiheitsgrade gibt. Die in der Physik häufig benutzten generalisierten Koordinaten des Lagrange-Formalismus sind solche Minimalkoordinaten.

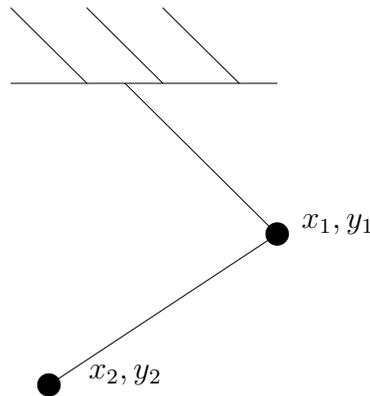


Doppelpendel mit Minimalkoordinaten

Der große Vorteil dieser Koordinaten besteht in ihrer kompakten Form: es gibt genau so viele Variablen wie eben mindestens notwendig sind und nicht mehr. Die Anzahl der Bewegungsdifferentialgleichungen wird dadurch minimal und es müssen keine sonstigen Abhängigkeiten bei der Integration berücksichtigt werden.

Der Nachteil der Minimalkoordinaten liegt in der Praxis häufig darin, daß sie meist Winkel oder winkelähnliche Größen darstellen. Demzufolge treten bei der Formulierung der Bewegungsgleichungen trigonometrische Terme auf, die selbst bei einfach strukturierten Mehrkörpersystemen wie dem Doppelpendel außerordentlich komplex sein können. So kann ihre numerische Auswertung „teuer“ sein und ihre Behandlung nur noch maschinell erfolgen. Eine eindrucksvolle Gegenüberstellung von Bewegungsgleichungen in minimalen und kartesischen Koordinaten findet sich bei SIMON[64].

- Bei redundanten oder abhängigen Koordinaten gibt es mehr Koordinaten als tatsächlich vorhandene Freiheitsgrade. Beispielsweise kann man zur Beschreibung des Doppelpendels kartesische Koordinaten verwenden, die die Lage der Pendelgewichte direkt beschreiben. Diese Möglichkeit der direkten Positionsbeschreibung führt dazu, daß die Bewegungsdifferentialgleichungen einfacher aussehen.



Doppelpendel mit kartesischen Koordinaten

Aber neben der vergrößerten Anzahl an Variablen muß als weiterer Preis das Auftreten von algebraischen Gleichungen bezahlt werden, die die Zulässigkeit der Koordinatenwerte einschränken. Beim Doppelpendel z. B. treten zwei Gleichungen hinzu, die die Länge der beiden Stangen beschreiben:

$$\begin{aligned}x_1^2 + y_1^2 - l_1^2 &= 0 \\(x_1 - x_2)^2 + (y_1 - y_2)^2 - l_2^2 &= 0.\end{aligned}$$

Mit den redundanten Koordinaten lassen sich also unzulässige Zustände beschreiben, z. B. falsche Fadenlänge, die durch neu auftretende Gleichungen („Zwangsbedingungen“) ausgeschlossen werden müssen. Die Menge der zulässigen Zustände bildet nur noch eine Untermannigfaltigkeit im Koordinatenraum, im Gegensatz zu den Minimalkoordinaten, bei denen der Koordinatenraum nur zulässige Punkte enthält (mit Einschränkungen wie z. B. Intervallbeschränkungen der Werte).

Zu den redundanten Koordinaten gehören neben den kartesischen Koordinaten die weniger bekannten natürlichen Koordinaten, die auf Arbeiten von GARCÌA DE JALÒN ET AL.[24],[25] zurückgehen². Die Lage jedes starren Körpers läßt sich durch eine Rotation des im Ursprung befindlichen Körpers und anschließender Verschiebung vom Ursprung weg beschreiben:

$$\begin{aligned}x' &= Rx + v \\R &: \text{Rotationsmatrix} \\v &: \text{Verschiebung.}\end{aligned}$$

Bei den natürlichen Koordinaten werden jedem Starrkörper zwölf Koordinaten zugeordnet, nämlich drei Koordinaten für die Ortsverschiebung v und die neun Einträge der Rotationsmatrix R . Diese Rotationsmatrix R muß aber eine Matrix sein, die die Winkel zwischen den Koordinatenachsen und ihre Längen unverändert läßt, da der Körper starr ist. Um dies zu fordern, sind sechs Gleichungsbedingungen notwendig („Starrkörperbedingungen“), so daß also jedem Körper zwölf Koordinaten und sechs Gleichungsbedingungen zugeordnet werden. Weitere Gleichungsbedingungen treten hinzu, wenn verschiedene Körper mit Gelenken verbunden werden sollen („Gelenkbedingungen“).

Der große Vorteil natürlicher Koordinaten besteht darin, daß auf jedem Körper ein lokales Koordinatensystem vorhanden ist³, was die Bestimmung der Lage jedes einzelnen Punkts

²siehe auch KRAUS[39]

³Dies ist in der Formulierung nach KRAUS[40] der Fall, in der Formulierung nach DE JALÒN[24] tritt eventuell das „point-sharing“ dazwischen.

des Starrkörpers sehr einfach macht. Ebenso lassen sich Gelenke zwischen den Körpern mit Hilfe dieser lokalen Koordinatensysteme sehr leicht formulieren. Dies ist i. allg. sogar so leicht, daß der Vorgang der Bestimmung der Modellgleichungen von einem Algorithmus durchgeführt werden kann, und das insbesondere auch in den Fällen, in denen geschlossene Schleifen vorhanden sind.

Ein Nachteil der Formulierung eines Modells in natürlichen Koordinaten ist die große Zahl an Variablen und Gleichungsbedingungen. Dieser Nachteil wird dadurch ausgeglichen, daß diese Bedingungen bzw. die aus ihnen resultierenden Matrizen sehr strukturiert sind und daß diese Strukturen zur effizienten Behandlung ausgenutzt werden können. Für Details diesbezüglich sei auf KRAUS, WINCKLER UND BOCK[40],[41] verwiesen.

Aufbauend auf diesen natürlichen Koordinaten ist von C. KRAUS das in C^{++} geschriebene Software-Paket MBSNAT entwickelt worden. Zu seinen Vorzügen zählen:

- Graphische Bedienungsfläche zur Konstruktion von Mehrkörpersystemen
- Objektorientierter Aufbau
- Lokale Koordinatensysteme sind auf jedem einzelnen Körper verfügbar
- Gelenke sind leicht modellierbar, die aus ihnen resultierenden Zwangsbedingungen sind in den meisten praktischen Fällen linear oder quadratisch
- Massenmatrix ist konstant (fließt sie in Gleichungssysteme ein, muß sie nur einmal zerlegt werden, was ein sehr großer numerischer Vorteil ist)
- Matrizen in den Bewegungsgleichungen sind dünnbesetzt und strukturiert
- Ausnutzung der Matrizenstruktur in der linearen Algebra.

§2 Bewegungsgleichungen und Numerische Integration

Gemäß den im vorigen Abschnitt geäußerten Überlegungen soll also das laufende System als Mehrkörpersystem in redundanten Koordinaten formuliert werden. Seien zu diesem Zweck⁴

$p(t)$: Vektor der Ortskoordinaten zur Zeit t , $p(t) \in \mathbb{R}^n$

$v(t)$: Vektor der Geschwindigkeitskoordinaten zur Zeit t , $v(t) \in \mathbb{R}^n$

$a(t)$: Vektor der Beschleunigungskordinaten zur Zeit t , $a(t) \in \mathbb{R}^n$

$F(t, p(t), v(t))$: Vektor der externen Kräfte auf die Körper zur Zeit t bei Ortslage $p(t)$ und Geschwindigkeitslage $v(t)$, $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$

M : Matrix der Massen der einzelnen Körper, $M \in \mathbb{R}^n \times \mathbb{R}^n$.

Diese Matrix M kann bei der Benutzung von Minimalkoordinaten durchaus eine volle und zeitabhängige Matrix sein, weil sie z. B. Zentrifugalterme berücksichtigen muß. Bei Verwendung natürlicher oder kartesischer Koordinaten ist sie konstant und blockdiagonal. Setzt man zusätzlich voraus, daß die lokalen Koordinatenachsen in Richtung der drei Trägheitsachsen zeigen und daß der Aufpunkt der Schwerpunkt des Körpers ist, so ist die Massenmatrix sogar konstant und diagonal, ein weiterer Vorteil natürlicher Koordinaten. Die Bewegungsgleichungen würden ohne Zwangsbedingungen gemäß Newtons zweitem

⁴In späteren Abschnitten wird auch gelegentlich

$$F(t, p(t), v(t), q(t))$$

geschrieben werden. Dabei ist dann $q(t)$ der Vektor der Kontrollen, die im Rahmen der Optimierung bestimmt werden sollen.

Gesetz der Mechanik als System von Differentialgleichungen erster Ordnung mit Anfangswertbedingungen

$$\begin{aligned}\dot{p}(t) &= v(t) \\ \dot{v}(t) &= a(t) \\ Ma(t) &= F(t, p(t), v(t)) \\ p(t_0) &= p_0 \\ v(t_0) &= v_0\end{aligned}$$

oder in einer etwas abkürzenden Schreibweise, in der die Zeitvariable t wo möglich unterdrückt wird,

$$\begin{aligned}\dot{p} &= v \\ \dot{v} &= a \\ Ma &= F(t, p, v) \\ p(t_0) &= p_0 \\ v(t_0) &= v_0\end{aligned}$$

lauten. Da aber Zwangsbedingungen vorliegen, die als

$$g_{pos}(p(t)) = 0$$

$$g_{pos} : \text{Ortszwangsbedingungen, } g_{pos} : \mathbb{R}^n \rightarrow \mathbb{R}^l$$

(n Anzahl der Variablen, l Anzahl der Zwangsbedingungen) oder kürzer

$$g_{pos}(p) = 0 \tag{1}$$

formuliert sind, treten auch Zwangskräfte auf.

Definition 1 (Ortskonsistenzmannigfaltigkeit). *Ein Vektor p von Ortskoordinaten, der die Gleichung (1) erfüllt, soll konsistent heißen. Die Menge aller konsistenten p heißt Ortskonsistenzmannigfaltigkeit.*

Die Zwangskräfte sorgen dafür, daß Kraftkomponenten in den Differentialgleichungen der Dynamik senkrecht zur Konsistenzmannigfaltigkeit verschwinden und das System somit konsistent bleibt. In der Euler-Lagrange-Formulierung werden diese Zwangskräfte als⁵ $-\lambda(t)G^T(p(t))$ geschrieben, wobei gilt:

$\lambda(t)$: Lagrange-Multiplikatoren zur Zeit t

$G(p(t))$: $G = g'_{pos}$ Zwangsbedingungsmatrix, Ableitung von g_{pos} nach allen

Ortsvariablen, $G : \mathbb{R}^n \rightarrow \mathbb{R}^l \times \mathbb{R}^n$.

Damit lassen sich die Bewegungsgleichungen im Euler-Lagrange-Formalismus finden:

$$\begin{aligned}\dot{p} &= v \\ \dot{v} &= a \\ Ma &= F(t, p, v) - \lambda G^T(p) \\ p(t_0) &= p_0 \\ v(t_0) &= v_0 \\ g_{pos}(p) &= 0.\end{aligned} \tag{2}$$

⁵Das Vorzeichen von λ ist reine Konvention und diese Wahl erspart später ein dauerhaftes Minuszeichen.

Dies ist die wohlbekannte Index-3-Form, da es dreimaliger Differentiation bedarf, um dieses Gleichungssystem in ein System gewöhnlicher Differentialgleichungen zu überführen.⁶ Um die Gleichung (2) numerisch zu lösen, wird nun die letzte Zeile

$$g_{pos}(p) = 0$$

zweimal nach der Zeit abgeleitet. Dabei ergeben sich nacheinander

$$\underbrace{G(p)v}_{:=g_{vel}(p,v)} = 0 \quad (3)$$

und

$$\underbrace{v^T \partial_p G(p)v}_{:= -\gamma(p,v)} + G(p)a = 0.$$

Die Gleichung (3) mit $g_{vel} : \mathbb{R}^n \rightarrow \mathbb{R}^l \times \mathbb{R}^n$ bezeichnet man auch als Geschwindigkeitszwangsbedingung. Dies führt zur

Definition 2 (Konsistenzmannigfaltigkeit). Ein Vektor $\begin{pmatrix} p \\ v \end{pmatrix}$ von Orts- und Geschwindigkeitsvariablen, der das Gleichungssystem

$$\begin{aligned} g_{pos}(p) &= 0 \\ g_{vel}(p, v) &= 0 \end{aligned}$$

erfüllt, soll konsistent heißen. Die Menge aller konsistenten $\begin{pmatrix} p \\ v \end{pmatrix}$ heißt Konsistenzmannigfaltigkeit.

Ersetzt man nun in (2) die letzte Gleichung durch ihre zweite Ableitung, so ergibt sich

$$\begin{aligned} \dot{p} &= v \\ \dot{v} &= a \\ Ma &= F(t, p, v) - \lambda G^T(p) \\ G(p)a &= \gamma(p, v) \\ p(t_0) &= p_0 \\ v(t_0) &= v_0. \end{aligned}$$

Dies läßt sich gefälliger als

$$\begin{aligned} \dot{p} &= v \\ \dot{v} &= a \\ \begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} &= \begin{pmatrix} F(t, p, v) \\ \gamma(p, v) \end{pmatrix} \\ p(t_0) &= p_0 \\ v(t_0) &= v_0 \end{aligned} \quad (4)$$

schreiben. (4) ist die Index-1-Form, da es einer weiteren Differentiation bedarf, um das System in ein System gewöhnlicher Differentialgleichungen zu überführen. (4) heißt auch *differential-algebraische Gleichung* (engl. differential algebraic equation, DAE). Dieses Gleichungssystem wird aber nicht noch einmal abgeleitet, wie es der Index-1-Begriff impliziert, sondern direkt wie eine gewöhnliche Differentialgleichung behandelt. Dies geschieht

⁶Zur Definition des differentiellen Index siehe HAIRER UND WANNER[29].

im Prinzip so: man verwendet einen Löser für gewöhnliche Differentialgleichungen und gibt vor, das Gleichungssystem würde

$$\frac{d}{dt} \begin{pmatrix} p \\ v \end{pmatrix} = f(t, p, v)$$

lauten. Jedesmal, wenn der Löser („Integrator“) eine Auswertung der rechten Seite f verlangt, wird das Gleichungssystem

$$\begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} = \begin{pmatrix} F(t, p, v) \\ \gamma(p, v) \end{pmatrix}$$

gelöst und als Ergebnis der Auswertung der rechten Seite f wird

$$\begin{pmatrix} v \\ a \end{pmatrix}$$

zurückgeliefert.

Das hier vorgestellte Verfahren funktioniert sehr gut mit zwei Einschränkungen:

- (1) Die Auswertung der rechten Seite ruht darauf, daß die Matrix

$$\begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix}$$

invertierbar ist. Im Normalfall ist M diagonal mit positiven Einträgen und G hat Vollrang. Unter diesen Umständen ist die Matrix invertierbar. Probleme mit dieser Voraussetzung werden später ausführlich behandelt.

- (2) Die zweite Einschränkung ist numerischer Natur. Sie liegt darin begründet, daß beim Übergang von (2) zu (4) die Ortszwangsbedingung $g_{pos}(p) = 0$ zweimal differenziert wurde. Um die daraus resultierenden Schwierigkeiten zu erläutern, sei ein einfaches Beispiel genommen: gelöst werden soll die Gleichung

$$f(t) = 0. \tag{5}$$

Diese hat trivialerweise die Lösung $f \equiv 0$. Statt der Gleichung (5) wird nun aber deren zweite Ableitung

$$\frac{d^2}{dt^2} f(t) = 0$$

unter Berücksichtigung der Startwerte $f(0) = 0$, $\frac{d}{dt}f(0) = 0$ gelöst. Die exakte Lösung ist wieder $f \equiv 0$. Nun werden aber in der Praxis in die Lösung einer solchen Differentialgleichung numerische Fehler eingeschleppt. Nimmt man an, daß zu einem bestimmten Zeitpunkt t_1 aufgrund von angehäuften Integrationsfehlern

$$\begin{aligned} f(t_1) &= \varepsilon_1 \\ \dot{f}(t_1) &= \varepsilon_2 \end{aligned}$$

gilt, so würde sich selbst bei folgender exakter Integration

$$f(t) = \varepsilon_1 + (t - t_1)\varepsilon_2$$

ergeben. Die Lösung driftet i. allg. von der korrekten Null-Lösung immer weiter weg. Im nichtlinearen Fall kann diese Drift eine exponentiell wachsende Geschwindigkeit entwickeln und selbst auf kurzen Integrationsintervallen zu unbrauchbaren Ergebnissen führen. Katastrophal werden die Folgen, wenn aufgrund dieses Effekts die Voraussetzungen von Punkt 1 nicht mehr gültig sind. Für Einzelheiten sei auf VON SCHWERIN[66] und auf später verwiesen.

Zur Vermeidung des in Punkt 2 genannten Drift-Effekts muß man Wissen um die Lösung der Differentialgleichung verwenden. So weiß man, daß die Lösung der Ortszwangsbedingung

$$g_{pos}(p) = 0$$

genügt, denn aus dieser ist ja die Differentialgleichung durch zweimaliges Ableiten entstanden. Ebenso genügt die Lösung der ersten Ableitung dieser Bedingung, der Geschwindigkeitszwangsbedingung

$$g_{vel}(p, v) = \frac{d}{dt}g_{pos}(p) = 0.$$

Diese beiden Bedingungen wurden durch die Umformulierung zu Invarianten der Differentialgleichung, d. h. genügen die Startwerte diesen beiden Bedingungen und würde man die Differentialgleichung exakt lösen, so genügte die Lösung wieder diesen beiden Bedingungen.

Da sich aber numerische Fehler einschleichen, verletzt die Lösung möglicherweise diese Invarianten. Die verwendeten Integrator-Bibliotheken MBSSIM/MBSNAT umgehen dies während des Integrationsprozesses: der Integrator überprüft nach jedem elementaren Integrationschritt, ob die Gleichungen

$$\begin{aligned} g_{pos}(p) &= 0 \\ g_{vel}(p, v) &= 0 \end{aligned} \tag{6}$$

noch innerhalb einer gegebenen Toleranz gültig sind. Falls nein, wird ein Projektionsverfahren eingeleitet: die Werte von p und v werden auf die durch die Bedingungen (6) gegebene Konsistenzmannigfaltigkeit zurückprojiziert. Dies hält die Drift in Grenzen und verringert darüberhinaus den Gesamtfehler der Integration. Für Details sei wieder auf VON SCHWERIN[66] verwiesen.

Damit lautet das zu lösende Gesamtsystem:

$$\begin{aligned} \dot{p} &= v \\ \dot{v} &= a \\ \begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} &= \begin{pmatrix} F(t, p, v) \\ \gamma(p, v) \end{pmatrix} \\ p(t_0) &= p_0 \\ v(t_0) &= v_0 \end{aligned} \tag{7}$$

unter Berücksichtigung der Invarianten

$$\begin{aligned} g_{pos}(p) &= 0 \\ g_{vel}(p, v) &= 0. \end{aligned}$$

Dabei muß vorausgesetzt werden, daß die Initialwerte p_0, v_0 konsistent sind, d. h. es gilt

$$\begin{aligned} g_{pos}(p_0) &= 0 \\ g_{vel}(p_0, v_0) &= 0, \end{aligned}$$

ansonsten hat das System (7) keine Lösung, da die Anforderungen an die Lösung inkonsistent sind.

Ein weiterer wesentlicher Aspekt der Integration ist die Behandlung von Unstetigkeiten. Das später betrachtete Modell soll einen laufenden Mechanismus beschreiben. Dazu ist der Kontakt mit dem Boden wesentlich. Die Berührung des Bodens durch einen Fuß ist aber von qualitativ anderer Natur als der freie Flug des Fußes über dem Boden. Man könnte zwar versucht sein, diesen Übergang durch hinreichend schnell wachsende Funktionen

differenzierbar zu glätten, doch dabei gewinnt man nichts: um einigermaßen realistisch zu sein (keine Kraft auf den Fuß, solange er noch über dem Boden schwebt, eine einsetzende Kraft mit spürbar rücktreibender Wirkung, sobald er den Boden berührt) müssen derartig schnell wachsende Funktionen verwendet werden, daß das Problem steif wird. Letztendlich macht es für den Integrator keinen Unterschied, ob er eine differenzierbare, aber sehr schnell wachsende Funktion oder einen unstetigen Sprung behandeln muß.

Dieses Problem wird wie folgt gelöst: man formuliert zwei getrennte Modelle, eines mit rücktreibender Kraft, eines ohne. Das Eintreten des unstetigen Übergangs wird durch eine *Schaltfunktion*⁷ bestimmt. Das ist eine stetig differenzierbare Funktion, deren Nulldurchgang den Modellwechsel bestimmt. Im Beispiel des Fußes ist dies schlicht die z -Koordinate

$$h_{\text{Schalt}}(z) := z.$$

Detektiert der Integrator einen Nulldurchgang der Schaltfunktion, so versucht er den genauen Zeitpunkt zu bestimmen, integriert die DAE noch bis zum Schaltzeitpunkt, wechselt dann das Modell und startet den Integrationsvorgang im Schaltzeitpunkt neu. Dies ist notwendig, da die in modernen Integratoren arbeitenden Fehlerschätzer von einer hinreichend oft differenzierbaren rechten Seite der DAE ausgehen und wegen des unstetigen Übergangs ihrer notwendigen Voraussetzungen beraubt würden.

Zur Lösung von Problemen der Form (7)⁸ ist von M. WINCKLER und R. VON SCHWERIN das Software-Paket MBSSIM entwickelt worden⁹. Unter seinen Funktionalitäten befindet sich:

- MBSSIM ist eine Bibliothek für Index-3-Differential-Algebraische Gleichungen
- Reduktion auf Index-1-Formel
- Projektionen, um Drift zu vermeiden
- Adaptive Diskretisierungsmethoden
- Behandlung von Unstetigkeiten durch Schaltfunktionen
- Berechnung von Sensitivitäten mit Hilfe von Interner Numerischer Differentiation.

§3 Optimierung

Die Bestimmung einer optimalen Laufbewegung ist ein nichtlineares Optimierungsproblem mit Randbedingungen. Das mechanische System wird von Aktuatoren angetrieben und es sollen die Steuerungen der Aktuatoren ermittelt werden, die die optimale Bewegung erzeugen.

Der Zustand des mechanischen Systems sei durch $y(t)$ beschrieben, die Steuerung der Aktuatoren (die Kontrollen) durch $u(t)$.

Die Bewegung des Systems wird dann durch eine Differentialgleichung der Form

$$\dot{y}(t) = f(t, u(t), y(t)), \quad y(t_0) = y_0$$

beschrieben¹⁰.

Um eine optimale Bewegung zu bestimmen, muß ein zu minimierendes Optimalitätskriterium festgelegt werden: $\mathcal{J}(y, u)$. Z. B. kann das Optimalitätskriterium sein

⁷siehe auch WINCKLER UND VON SCHWERIN[70]

⁸siehe auch SIMEON[63]

⁹siehe auch WINCKLER[68] oder BOCK[10]

¹⁰Dies ist zwar nur eine gewöhnliche Differentialgleichung, die folgenden Ausführungen gelten aber analog für differential-algebraische Gleichungen.

- maximale zurückgelegte Entfernung zum Zeitpunkt t_e :

$$\mathcal{J}(y, u) = -y(t_e)$$

- minimale Kontrolle

$$\mathcal{J}(y, u) = \int_{t_0}^{t_e} u(t)^2 dt$$

und weitere oder Mischterme daraus.

Weiter soll die Bewegung gewissen Einschränkungen genügen, die als Gleichungs- und Ungleichungsbedingungen formuliert werden. Z. B. können dies Winkelbeschränkungen, Abstandsbeschränkungen oder Periodizitätsforderungen sein. Zusammen wird dies als

$$\begin{aligned} \min_u \quad & \mathcal{J}(y, u) \\ \dot{y}(t) = & f(t, u(t), y(t)), \quad y(t_0) = y_0 \\ r(y(t)) = & 0 \\ h(y(t), u(t)) \geq & 0 \end{aligned} \quad (8)$$

formuliert, wobei h insbesondere Beschränkungen an Zustände und Steuerungen der Art

$$\begin{aligned} u_{min} \leq u(t) \leq u_{max} \\ y_{min} \leq y(t) \leq y_{max} \end{aligned} \quad (9)$$

enthält.

Dieses Optimierungsproblem wird zunächst von der kontinuierlichen in eine diskrete Form überführt, um es zu einem endlichen Problem zu machen: dazu wird die kontinuierliche Funktion $u(t)$ diskretisiert, z. B. indem man das Zeitintervall $[t_0, t_e]$ in endliche viele Subintervalle $t_0 < t_1 < \dots < t_N = t_e$ einteilt, auf denen die Funktion u konstant sein soll:

$$u(t) \equiv u_i \text{ für } t_i \leq t < t_{i+1}.$$

Auch die Randbedingungen werden auf diese Zeitpunkte eingeschränkt:

$$\begin{aligned} r(y(t_i)) = 0 \\ h(y(t_i), u(t_i)) \geq 0 \end{aligned}$$

oder kürzer

$$\begin{aligned} r(y_i) = 0 \\ h(y_i, u_i) \geq 0. \end{aligned}$$

Des weiteren hat es sich als numerisch vorteilhaft herausgestellt¹¹, das Verfahren des „multiple shooting“ (dt. „Mehrzielmethode“) zu verwenden: es wird nicht nur versucht, die richtigen Kontrollen und eventuelle Anfangswerte zu ermitteln, sondern im Laufe der Optimierung werden auch die richtigen Zustände zu den Zwischenzeitpunkten gesucht, d. h.

$$\begin{aligned} \min_{y_i, u_i} \quad & \mathcal{J}(y, u) \\ \dot{y} = & f(t, y, u), \quad y(t_i) = y_i, \quad t \in [t_i, t_{i+1}[\\ r(y_i) = & 0 \\ h(y_i, u_i) \geq & 0. \end{aligned}$$

Die Lösung muß allerdings stetig und damit als Lösung der Differentialgleichung stetig differenzierbar sein. Also muß man zusätzlich die Bedingung stellen, daß die Lösungen auf

¹¹siehe z. B. BULIRSCH[13], BOCK[7] und [8] oder BOCK UND PLITT[9]

den einzelnen Teilintervallen aneinander anschließen müssen. Ist $Y(t, t_i, y_i)$ die Lösung der Differentialgleichung zum Zeitpunkt t bei Anfangsbedingung $y(t_i) = y_i$, so lauten diese Bedingungen

$$Y(t_{i+1}, t_i, y_i) - y_{i+1} = 0. \quad (10)$$

Diese Gleichungen werden zu den Gleichungsbedingungen r hinzugefügt.

Das so formulierte Optimierungsproblem gehört zu einer recht allgemeinen Klasse: nichtlineares Zielfunktional unter nichtlinearen Gleichungs- und Ungleichungsbedingungen. Die Auswahl an Verfahren mit einer akzeptablen Konvergenzrate ist daher nicht besonders groß. Das Newton-Verfahren bietet (lokal) quadratische Konvergenz, das Gauß-Newton- und das SQP-Verfahren bieten (lokal) superlineare Konvergenz. Alle Verfahren sind iterativ und setzen voraus, daß die vorkommenden Funktionen zweimal stetig differenzierbar sind.

Definiert man auf folgende Weise die Lagrange-Funktion

$$\mathcal{L}(y, \lambda, \mu) := \mathcal{J}(y, u) - \lambda^T r(y) - \mu^T h(y, u),$$

so ist eine notwendige (und unter zusätzlichen Annahmen hinreichende) Bedingung für das Vorliegen eines Optimums das Auffinden einer Lösung des Gleichungssystems („Karush-Kuhn-Tucker-Bedingungen“)¹²

$$\begin{aligned} \nabla_y \mathcal{L}(y, \lambda, \mu) &= 0 \\ r(y) &= 0 \\ h(y, u) &\geq 0 \\ \mu &\geq 0 \\ \mu_j h_j(y, u) &\geq 0. \end{aligned}$$

Die Ungleichungen können mit einer Active-Set-Strategie behandelt werden. Diese besteht darin, in einem iterativen Verfahren in jedem Iterationsschritt die Ungleichungen als aktiv, d. h. $h_j(y, u) = 0$, oder als inaktiv, d. h. $h_j(y, u) > 0$, zu klassifizieren. Nur die Aktiven werden dann zur Menge der Gleichungsbedingungen hinzugenommen. Somit wird dieses System (in jedem Iterationsschritt erneut) in ein Gleichungssystem verwandelt.¹³

Das Newton-Verfahren versucht dieses System auf die übliche Weise zu lösen. Da aber in der ersten Gleichung $\nabla_y \mathcal{L}$ vorkommt, kommt im Algorithmus des Newton-Verfahrens die zweite Ableitung $\nabla_y^2 \mathcal{L}$ oder Hesse-Matrix vor, deren Berechnung numerisch sehr teuer ist und bis zur Inpraktikabilität führen kann.

Das Gauß-Newton-Verfahren nähert daher die Hesse-Matrix nur an, indem mit einer Einheits-Matrix begonnen wird und in jedem Schritt Updates auf die genäherte Hesse-Matrix stattfinden. Die Konvergenzrate sinkt dadurch von (lokal) quadratisch auf (lokal) superlinear, aber die Berechnung ist nun praktisch durchführbar.

Das SQP-Verfahren geht von einer anderen Lösungsidee aus, stellt sich aber bei genauerer Analyse als mit dem Gauß-Newton-Verfahren verwandt heraus: die Zielfunktion (oder auch die Lagrange-Funktion) wird quadratisch approximiert, die Nebenbedingungen werden linear approximiert. Dies ist ein quadratisches Programm (QP), welches gelöst werden kann. Um die Lösung dieses QP-Teilproblems findet dann eine neue Näherung statt, etc.¹⁴

¹²siehe z. B. GROSSMANN UND TERNO[28]

¹³Im SQP-Verfahren (s. u.) wird zunächst das Gleichungssystem genähert und die Active-Set-Strategie dann auf die genäherten Ungleichungen angewandt.

¹⁴Eine Beschreibung des SQP-Verfahrens kann auch den Lehrbüchern von GILL ET AL.[27] und FLETCHER[20] entnommen werden.

Das Programmpaket MUSCOD-II, das von D. LEINEWEBER entwickelt wurde, löst Optimierungsprobleme mit dem SQP-Verfahren und multiple shooting. Die Hesse-Matrix wird mit einem Block-update-Verfahren genähert. Für Details sei auf LEINEWEBER[42] und [43] verwiesen.

§4 Aufteilung

Das Zusammenspiel der drei oben genannten Komponenten kann man sich wie folgt denken: nach erfolgreicher Modellierung werden Teile des Modells dem Integrator, andere Teile des Modells dem Optimierer mitgeteilt. Dann beginnt die iterative Optimierungsschleife, die selbst immer wieder den Integrator aufruft.

Damit stellt sich die Frage, welche Teile des Modells wem mitgeteilt werden müssen. Diese Aufteilung soll an einem sehr kleinen Beispiel demonstriert werden. Ein MBS soll aus zwei punktförmigen Massen M_1 und M_2 bestehen, die durch einen festen Abstand l (eine gedachte starre Stange) getrennt sind.

Beispiel 1.



Als Koordinaten seien kartesische Koordinaten in zwei Dimensionen gewählt:

$$p_{1x}, p_{1y}, p_{2x}, p_{2y},$$

analog Geschwindigkeit v und Beschleunigung a , die Massen seien m_1 und m_2 . Als externe Kräfte werden un spezifizierte $F_{1x}, F_{1y}, F_{2x}, F_{2y}$ angenommen. Die Zwangsbedingung lautet dann

$$g_{pos}(p_{1x}, p_{1y}, p_{2x}, p_{2y}) = (p_{1x} - p_{2x})^2 + (p_{1y} - p_{2y})^2 - l^2.$$

Die Ableitung nach der Zeit ergibt die Geschwindigkeitszwangsbedingung

$$g_{vel}(p_{1x}, p_{1y}, p_{2x}, p_{2y}, v_{1x}, v_{1y}, v_{2x}, v_{2y}) = 2(p_{1x} - p_{2x})(v_{1x} - v_{2x}) + 2(p_{1y} - p_{2y})(v_{1y} - v_{2y}).$$

Die Ableitungsmatrix G lautet

$$G = \begin{pmatrix} 2(p_{1x} - p_{2x}) & 2(p_{1y} - p_{2y}) & -2(p_{1x} - p_{2x}) & -2(p_{1y} - p_{2y}) \end{pmatrix}.$$

Für γ ergibt eine kurze Rechnung

$$\gamma = -2((v_{1x} - v_{2x})^2 + (v_{1y} - v_{2y})^2).$$

Damit lautet das differential-algebraische Gleichungssystem

$$\dot{p}_{1x} = v_{1x}$$

$$\dot{p}_{1y} = v_{1y}$$

$$\dot{p}_{2x} = v_{2x}$$

$$\dot{p}_{2y} = v_{2y}$$

$$\dot{v}_{1x} = a_{1x}$$

$$\dot{v}_{1y} = a_{1y}$$

$$\dot{v}_{2x} = a_{2x}$$

$$\dot{v}_{2y} = a_{2y}$$

$$\begin{pmatrix} m_1 & 0 & 0 & 0 & 2(p_{1x} - p_{2x}) \\ 0 & m_1 & 0 & 0 & 2(p_{1y} - p_{2y}) \\ 0 & 0 & m_2 & 0 & -2(p_{1x} - p_{2x}) \\ 0 & 0 & 0 & m_2 & -2(p_{1y} - p_{2y}) \\ 2(p_{1x} - p_{2x}) & 2(p_{1y} - p_{2y}) & -2(p_{1x} - p_{2x}) & -2(p_{1y} - p_{2y}) & 0 \end{pmatrix} \begin{pmatrix} a_{1x} \\ a_{1y} \\ a_{2x} \\ a_{2y} \\ \lambda \end{pmatrix} = \begin{pmatrix} F_{1x} \\ F_{1y} \\ F_{2x} \\ F_{2y} \\ -2((v_{1x} - v_{2x})^2 + (v_{1y} - v_{2y})^2) \end{pmatrix} \quad (11)$$

mit den Anfangswerten

$$\begin{aligned} p_{1x}(0) &= p_{01x}, v_{1x}(0) = v_{01x}, \\ p_{1y}(0) &= p_{01y}, v_{1y}(0) = v_{01y}, \\ p_{2x}(0) &= p_{02x}, v_{2x}(0) = v_{02x}, \\ p_{2y}(0) &= p_{02y}, v_{2y}(0) = v_{02y} \end{aligned}$$

und unter Berücksichtigung der Invarianten

$$g_{pos}(p_{1x}, \dots, p_{2y}) = (p_{1x} - p_{2x})^2 + (p_{1y} - p_{2y})^2 - l^2 \quad (12)$$

$$g_{vel}(p_{1x}, \dots, p_{2y}, v_{1x}, v_{1y}, v_{2x}, v_{2y}) = 2(p_{1x} - p_{2x})(v_{1x} - v_{2x}) + 2(p_{1y} - p_{2y})(v_{1y} - v_{2y}).$$

Für die Optimierung werde folgendes Szenario angenommen: die Stange sei zur Zeit $t = 0$ in einer festen, unveränderlichen Startposition, z. B. bei $l = 2$

$$\begin{aligned} p_{1x}(0) &= -1, v_{1x}(0) = 0, \\ p_{1y}(0) &= 0, v_{1y}(0) = 0, \\ p_{2x}(0) &= 1, v_{2x}(0) = 0, \\ p_{2y}(0) &= 0, v_{2y}(0) = 0. \end{aligned} \quad (13)$$

Es gebe eine Kontrolle u (z. B. ein auf die Stange wirkendes Drehmoment), mit deren Hilfe die Stange gedreht werden kann. Das Ziel ist es nun, die Stange bis zum Zeitpunkt $t = 1$ in die Position

$$\begin{aligned} p_{1x}(1) &= 0 \\ p_{1y}(1) &= 1 \\ p_{2x}(1) &= 0 \\ p_{2y}(1) &= -1 \end{aligned} \quad (14)$$

zu drehen und dabei so wenig Kontrolle wie möglich aufzuwenden:

$$\min_u \int_0^1 u(t)^2 dt. \quad (15)$$

Schließlich werden noch Schranken an Zustände und Kontrollen benötigt:

$$\begin{aligned} -2 &\leq p_{ix}, p_{iy} \leq 2 \\ -10 &\leq v_{ix}, v_{iy} \leq 10 \\ -100 &\leq u \leq 100. \end{aligned} \quad (16)$$

Diese Daten werden dann wie folgt aufgeteilt:

- Optimierer:

Der Optimierer erhält die Anzahl der Zustandsvariablen, nämlich 8 (4 Positionen und 4 Geschwindigkeiten), die Anzahl der Kontrollen, nämlich 1, das Ziel-funktional (15), die Schranken (16), die Startwerte (13), die Randbedingungen (14), die Unveränderlichkeit der Startposition und Startwerte („Initial guess“). Die Differentialgleichung wird dem Optimierer *nicht* mitgeteilt. Er sieht den Integrationsvorgang nur als „black box“, in die Anfangswerte und Kontrollwerte hineingegeben werden und aus der sich Endwerte ergeben.

Im Mehrzielverfahren vergrößert sich die Anzahl der Variablen um den Faktor der Anzahl von Zwischenzeitpunkten und als zusätzliche zu erfüllende Innere-Punkte-Bedingungen treten die Stetigkeitsanschlussbedingungen hinzu.

- Integrator:

Der Integrator erhält die Differentialgleichung (11) mit den Invarianten (12). Die Startwerte werden bei jedem Aufruf vom Optimierer geliefert.

- Denkbar ist auch eine gemischte Darstellung: der Integrator benutzt obige kartesische Koordinaten, während die Optimierung andere Koordinaten, z. B. Minimalkoordinaten verwendet. Der Vorteil besteht darin, daß vom Optimierer nun nur noch konsistente Startwerte geliefert werden können und daß die Anzahl der zu berechnenden Sensitivitäten minimal ist. Der Nachteil besteht in i. allg. sehr komplizierten Umrechnungen. Insbesondere bei der Rückrechnung von kartesischen in Minimalkoordinaten kann es Schwierigkeiten geben (z. B. wenn eine Kreisbewegung eine volle Umdrehung gemacht hat, muß die Winkelvariable dies durch ein Inkrement von 2π berücksichtigen, ansonsten erscheint dies als undifferenzierbarer Sprung von $2\pi - \varepsilon$ nach 0).

Insbesondere stellt sich die Frage, ob die Invarianten g_{pos} und g_{vel} dem Optimierer mitgeteilt werden müssen. Der Optimierer verfügt ja über die Werte y_i (die Zustände des Systems zu den inneren Zeitpunkten t_i) und damit wird die Integration des nächsten Zeitintervalls gestartet. Aber es gilt:

Feststellung 1. *Ist die Anfangskonfiguration konsistent, d. h.*

$$\begin{aligned} g_{pos}(p_0) &= 0 \\ g_{vel}(p_0, v_0) &= 0, \end{aligned}$$

und ist eine stetige Lösung $(p(t), v(t))$ der Deskriptorform (11) gegeben, so ist diese Lösung überall konsistent, d. h.

$$\begin{aligned} g_{pos}(p(t)) &= 0 \\ g_{vel}(p(t), v(t)) &= 0. \end{aligned}$$

BEWEIS. Der Beweis ist trivial, da g_{pos} und g_{vel} ja Invarianten jeder Lösung von (11) sind. \square

Im Vergleich dazu gilt:

Feststellung 2. *Ist eine stückweise stetige, überall konsistente Lösung von (11) gegeben, so ist diese nicht notwendigerweise stetig.*

BEWEIS. Formuliert man Zwangsbedingungen so, daß sie auch von der spiegelbildlichen Konfigurationen erfüllt werden, so kann die Lösung an einer Stelle in die spiegelbildliche Stellung springen. Sie ist dann trotzdem überall konsistent, aber nicht notwendigerweise stetig. Ist z. B. die Zwangsbedingung

$$g_{pos}(p) = p^2 - 1$$

zusammen mit $p_0 = 1$ gegeben, so ist

$$p(t) = \begin{cases} 1 & \text{für } t_0 \leq t \leq t_0 + 1 \\ -1 & \text{für } t_0 + 1 < t \leq t_0 + 2 \end{cases}$$

eine stückweise stetige, überall konsistente, jedoch nicht stetige Lösung. \square

Es liegt daher nahe, dem Optimierer nur die im Mehrzielverfahren auftretenden Stetigkeitsanschlussbedingungen mitzuteilen.¹⁵ Wegen der konsistenten und unveränderlichen Anfangswerte des Modellproblems ist dann die Lösung stetig und konsistent nach Feststellung 1. Dies führt aber wiederum dazu, daß im Laufe der Iterationen inkonsistente Startwerte in den Mehrzielknoten gebildet werden und integriert werden müssen.¹⁶ Es tritt also das Problem auf Gleichung (11) mit Startwerten zu lösen, für die

$$\begin{aligned} g_{pos}(p_0) &\neq 0 \\ g_{vel}(p_0, v_0) &\neq 0 \end{aligned}$$

gilt. Dies wird im nächsten Kapitel genauer analysiert.

¹⁵Beide Bedingungen (Stetigkeit und Konsistenz) können nicht gemeinsam gefordert werden, da dies zu Redundanzen im Optimierungsproblem führt; allerdings kann eine nicht-redundante Auswahl aus der Menge der Stetigkeits- und Konsistenzbedingungen gefordert werden, siehe SCHLÖDER[60].

¹⁶Dieses Problem tritt auch auf, wenn man die Sensitivität um konsistente Werte herum mit variierten Trajektorien berechnen möchte und dabei redundante Koordinaten verwendet.

KAPITEL II

Relaxierung

Die im vorhergehenden Kapitel vorgeschlagene Aufteilung der Problemformulierung hat die Konsequenz, daß das numerische Verfahren zur Lösung der differential-algebraischen Gleichungen häufig mit inkonsistenten Initialwerten starten muß. Um dies zu ermöglichen, ist die Relaxierung der Gleichungsnebenbedingungen notwendig. Die einfache Form der Relaxierung kann aber fatale Konsequenzen haben, wie gezeigt werden wird. Um dies zu vermeiden, wird das Verfahren der orthogonalen Relaxierung eingeführt, das zumindest die durch inkonsistente Anfangsgeschwindigkeiten verursachten Probleme beseitigt. Diese Form der Relaxierung hat zahlreiche Vorteile: kein numerischer Mehraufwand, Forminvarianz, Globalität sowie Anwendbarkeit auch in gemischten Koordinaten.

Des Weiteren wird eine Methode vorgestellt, um die durch inkonsistente Anfangspositionen verursachten Schwierigkeiten zu beseitigen. Das Verfahren beruht auf der Hinzufügung weiterer Variablen und wird lediglich theoretisch präsentiert. Es ist zwar ein globales Verfahren, aber das danach vorgestellte (lokale) Verfahren der initialen Projektion ist für die Anwendung besser geeignet. Denn es beseitigt simultan eventuell inkonsistente Positionen oder Geschwindigkeiten und hat zudem bezüglich der Optimierung den Vorteil, die Anzahl der notwendigen Richtungsableitungen auf das notwendige Minimum zu beschränken. Hierdurch wird der numerische Aufwand bei einer Formulierung in redundanten Koordinaten im wesentlichen auf den Aufwand in Minimalkoordinaten reduziert, wodurch die Verwendung redundanter Koordinaten bei großen Systemen erst sinnvoll wird.

§1 Singularitäten durch Geschwindigkeiten

Die im vorigen Abschnitt vorgestellten Verfahren arbeiten alle mit einer Näherung des Problems, insbesondere einer Linearisierung der Nebenbedingungen. Speziell die Stetigkeitsanschlussbedingungen (10) müssen auch linear genähert werden:

$$\frac{\partial}{\partial y_i} Y(t_{i+1}, t_i, y_i) \Delta y_i - \Delta y_{i+1} = 0.$$

Der Ausdruck $\frac{\partial}{\partial y_i} Y(t_{i+1}, t_i, y_i)$ ist dabei die „Sensitivität“ der Lösung der Differentialgleichung in Hinsicht auf die Variation der Startwerte und ist eine Matrix, die *Wronski-Matrix* oder *Sensitivitätsmatrix*.

In unabhängigen Koordinaten ist die Definition der Wronski-Matrix W :

$$W_{mn}(t_{i+1}, t_i, y_i) := \frac{\partial}{\partial y_{in}} Y_m(t_{i+1}, t_i, y_i) = \lim_{h \rightarrow 0} \frac{Y_m(t_{i+1}, t_i, y_i + h e_n) - Y_m(t_{i+1}, t_i, y_i)}{h},$$

wobei e_n der n -te Einheitsvektor ist. In Minimalkoordinaten ist der Ausdruck

$$Y_m(t_{i+1}, t_i, y_i + h e_n)$$

wohlbestimmt, denn der Startwert $y_i + h e_n$ ist konsistent und damit eine Lösung der Differentialgleichung i. allg. wohldefiniert.

In redundanten Koordinaten ist dem aber nicht so, der Ausdruck $Y_m(t_{i+1}, t_i, y_i + h e_n)$ ist i. allg. nicht definiert, da i. allg. $y_i + h e_n$ nicht mit den Zwangsbedingungen verträglich

ist und daher keine Lösung der DAE existiert. Um den Begriff der Sensitivitätsmatrix im Kontext der DAE zu erhalten, ist es also nötig zu definieren, was es heißen soll, Gleichung (11) mit Startwerten zu lösen, für die

$$\begin{aligned} g_{pos}(p_0) &\neq 0 \\ g_{vel}(p_0, v_0) &\neq 0 \end{aligned}$$

gilt. Das gleiche Problem ergibt sich nochmal im Rahmen des multiple shooting. Die Zustände zu den Zwischenzeiten sind im Laufe der Rechnung i. allg. inkonsistent und werden erst in der Lösung konsistent. Mit diesen inkonsistenten Zwischenzuständen soll aber der Integrator gestartet werden.

Zu diesem Zweck müssen die Invarianten bzw. Zwangsbedingungen modifiziert werden.¹

§1.1 Lineare Relaxierung. Eine Möglichkeit ist die (siehe SCHULZ, BOCK UND STEINBACH[61] oder auch VON SCHWERIN[66])²

Definition 3 (Lineare Relaxation).

$$\begin{aligned} \tilde{g}_{pos,p_0,v_0}(p) &:= g_{pos}(p) - g_{vel}(p_0, v_0)(t - t_0) - g_{pos}(p_0) \\ \tilde{g}_{vel,p_0,v_0}(p, v) &:= g_{vel}(p, v) - g_{vel}(p_0, v_0). \end{aligned}$$

Entsprechend sei die Ableitung bzgl. p \tilde{G}_{p_0,v_0} von \tilde{g}_{pos,p_0,v_0} definiert.

Der Name „Lineare Relaxation“ bezieht sich darauf, daß die ursprünglichen Bedingungen durch einen linearen Term ergänzt werden.

Es ist offensichtlich, daß für beliebige p_0, v_0 und $t = t_0$ nun für die entsprechend relaxierten Zwangsbedingungen gilt:

$$\begin{aligned} \tilde{g}_{pos,p_0,v_0}(t_0, p_0) &= 0 \\ \tilde{g}_{vel,p_0,v_0}(p_0, v_0) &= 0. \end{aligned}$$

Diese Form der Relaxierung hat den Vorteil, daß die dynamischen Gleichungen in (11) unverändert bleiben, da

$$\tilde{G}_{p_0,v_0} = G$$

gilt, weshalb hier der Index an \tilde{G} entfällt. Auch bleibt die hierarchische Struktur der Zwangsbedingungen ($g_{vel} = \frac{d}{dt}g_{pos}$) erhalten:

$$\tilde{g}_{vel,p_0,v_0} = \frac{d}{dt}g_{pos,p_0,v_0}.$$

Nach dem Begriff der Relaxation kann man nun auch den Begriff der Sensitivitätsmatrix für DAEs einführen.

Definition 4 (Wronski-Matrix für DAEs). Sei $Y_{y_0}(t, t_0, y_0)$ die Lösung der DAE zum Zeitpunkt t mit Anfangswerten $y(t_0) = y_0$ und zu Anfangswerten y_0 relaxierten Zwangsbedingungen. Dann sei als Wronski-Matrix definiert

$$W(t, t_0, y_0) := \frac{\partial}{\partial y_0} Y_{y_0}(t, t_0, y_0),$$

¹Die Sensitivität wird z. B. bei GERDTS[26] auf eine andere Weise definiert. Dort wird davon ausgegangen, daß es eine Funktion gibt, die zu den variierten Parametern konsistente Anfangswerte liefert.

² t_0 kann zwar auch als Anfangswert aufgefaßt werden, wird aber im folgenden immer als fest gedacht. Die Einbeziehung würde aber keine Schwierigkeiten bereiten, lediglich die Notation würde weniger anschaulich.

genauer

$$W_{mn}(t, t_0, y_0) := \frac{\partial}{\partial y_{0n}} (Y_{y_0})_m(t, t_0, y_0) = \lim_{h \rightarrow 0} \frac{(Y_{y_0 + h e_n})_m(t, t_0, y_0 + h e_n) - (Y_{y_0})_m(t, t_0, y_0)}{h}.$$

Aber die lineare Relaxierung hat Nachteile. Um diese Nachteile zu demonstrieren, soll ein triviales Beispiel mit 0 Freiheitsgraden eingeführt werden.

Beispiel 2.

Das System bestehe aus einem Massepunkt der Masse M , der sich auf einer eindimensionalen Mannigfaltigkeit (beschrieben durch $p \in \mathbb{R}$) befinde. Zusätzlich gelte nun noch die Zwangsbedingung

$$g_{pos}(p) := p^2 - 1.$$

Konsistente Anfangswerte wären also $p_0 = 1, v_0 = 0$ und der Massepunkt bliebe dann statisch an dieser Stelle.

Nun sollen inkonsistente Anfangsgeschwindigkeiten betrachtet werden, wobei $t_0 = 0$ und $p_0 = 1$, also konsistent, gesetzt werden.

Über die Invarianten

$$\begin{aligned} \tilde{g}_{pos, p_0, v_0} &= p^2 - 1 - 2p_0 v_0 (t - t_0) - p_0^2 + 1 \\ &= p^2 - p_0^2 - 2p_0 v_0 t \\ &= p^2 - 1 - 2v_0 t \end{aligned}$$

und

$$\tilde{g}_{pos, p_0, v_0} \equiv 0$$

erhält man somit die totale Lösungstrajektorie³

$$p(t) = \sqrt{2p_0 v_0 t + p_0^2} = \sqrt{2v_0 t + 1}.$$

An der Lösung kann man folgendes erkennen:

- (1) Die intuitiv erwartete Funktion $p(t) = p_0 + v_0 t$ ist nicht die Lösung.
- (2) Die Lösung $p(t) = \sqrt{2p_0 v_0 t + p_0^2}$ (bei angenommenem $p_0 > 0$) ist nur für $v_0 > 0$ global definiert, für $v_0 < 0$ bricht die Lösung nach endlicher Zeit zusammen (und führt zuvor zu numerischen Instabilitäten).

Der letzte Punkt soll noch einmal genauer betrachtet werden. Die Formulierung

$$g_{pos} = p^2 - 1$$

führt zur Ableitungsmatrix

$$G = (2p).$$

Die dynamischen Gleichungen (7) sind aber nur gültig, solange die Matrix

$$\begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix}$$

³Bereits hier macht sich ein Nachteil der linearen Relaxierung bemerkbar, nämlich die Abhängigkeit von der Wahl der Formulierung; statt obigem g_{pos} hätte man auch

$$g_{pos} = p - 1$$

verwenden können, das hätte aber zur totalen Lösung

$$p(t) = p_0 + v_0 t$$

geführt.

invertierbar ist, und dafür ist der Vollrang von G notwendig. Offenbar hat G Vollrang, wenn p von 0 verschieden ist und in $p = 0$ hat G Rangverlust. Nun betrachte man unter der Annahme $v_0 < 0$ die lineare Relaxierung ($t_0 = 0$)

$$\begin{aligned}\tilde{g}_{pos,p_0,v_0}(t,p) &= p^2 - p_0^2 - 2p_0v_0t \\ \tilde{g}_{vel,p_0,v_0}(p,v) &= 2pv - 2p_0v_0.\end{aligned}$$

Es ist für $t = 0$

$$\tilde{g}_{pos,p_0,v_0}(0,p_0,v_0) = 0$$

und für die Lösung $(p(t), v(t))$ des relaxierten Systems gilt für beliebige Zeiten

$$\tilde{g}_{pos,p_0,v_0}(t,p(t)) = 0.$$

Da nun aber $v_0 < 0$ ist, kann letzteres bei anwachsendem t nur erreicht werden, indem p kleiner wird. Das System wird also in den Rangverlust hineingetrieben! Dies funktioniert formal bis $p = 0$ ist. Dies ist zur Zeit

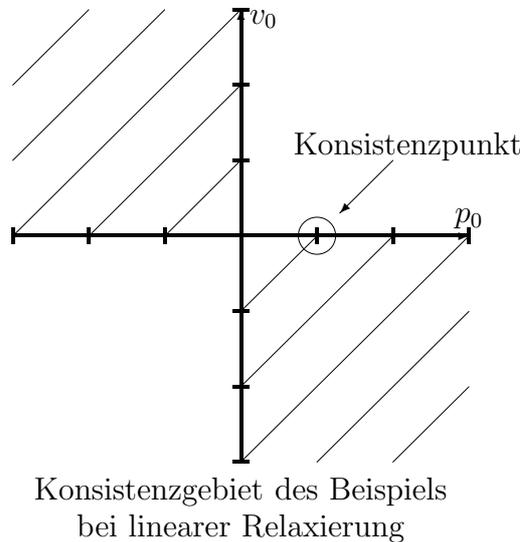
$$t = \frac{-p_0^2}{2p_0v_0} = -\frac{p_0}{2v_0}$$

der Fall. Gleichzeitig ist zu beachten, daß

$$\tilde{g}_{vel,p_0,v_0}(p(t), v(t)) = 0$$

gilt. Da aber p gegen 0 geht und letzteres $2p(t)v(t) = const \neq 0$ erzwingt, muß damit v nach Unendlich gehen. Dies zeigt, daß schon in der Nähe des Rangverlusts Probleme auftreten werden.

In diesem einfachen Beispiel wird das System in 50% aller inkonsistenten Fälle nach endlicher Zeit in eine Singularität getrieben, in der die Integration zusammenbricht.⁴



⁴Dies läßt sich allgemein so sagen: gilt

$$g_{pos}(p_0) = g_0 \in \mathbb{R}$$

und hat g_{pos} eine Singularität für einen Wert $s < g_0$ ($s > g_0$), so wird für

$$g_{vel}(p_0, v_0) < 0 \text{ (} g_{vel}(p_0, v_0) > 0 \text{)}$$

das System in die Singularität getrieben. Für $g_{pos}(p_0) \in \mathbb{R}^n$ ist die Situation komplizierter, da neben dem Rangverlust einzelner Zeilen die Möglichkeit der hinzukommenden linearen Abhängigkeit besteht.

Die schraffierte Fläche zeigt die Menge aller Initialwerte, die nach endlicher Zeit in einem Rangverlust enden. Die ungeschraffierte Menge der „problemlosen“ Startwerte ist nicht zusammenhängend, die beiden Quadranten berühren sich in einem Punkt, der ausgerechnet die Singularität selbst ist. D. h. aber, der „problemlose“ Teil ist nicht zusammenhängend und nicht konvex. Insbesondere liegt der Konsistenzpunkt $p_0 = 1, v_0 = 0$ am Rand des problemlosen Gebiets.

Anmerkung: Ähnlich bleibt die Situation bei Verwendung der Baumgarte-Relaxierung (siehe SCHULZ, BOCK UND STEINBACH[61] oder BAUMGARTE[3]). Hierbei werden die Ersetzungen

$$\begin{aligned}\tilde{g}_{pos,p_0,v_0}(t,p) &:= g_{pos}(p) - \\ &\quad - [g_{pos}(p_0) + (g_{vel}(p_0, v_0) + \alpha g_{pos}(p_0))(t - t_0) + \\ &\quad + (\alpha g_{vel}(p_0, v_0) + \frac{\alpha^2}{2} g_{pos}(p_0))(t - t_0)^2] e^{-\alpha(t-t_0)} \\ \tilde{g}_{vel,p_0,v_0}(t,p,v) &:= g_{vel}(p, v) - \\ &\quad - [g_{vel}(p_0, v_0) + \alpha g_{vel}(p_0, v_0)(t - t_0) - \\ &\quad - (\alpha^2 g_{vel}(p_0, v_0) + \frac{\alpha^3}{2} g_{pos}(p_0))(t - t_0)^2] e^{-\alpha(t-t_0)} \\ \tilde{\gamma}(t,p,v) &:= \gamma(p, v) + \\ &\quad + [-(3\alpha^2 g_{vel}(p_0, v_0) + \alpha^3 g_{pos}(p_0))(t - t_0) + (\alpha^3 g_{vel}(p_0, v_0) + \\ &\quad + \frac{\alpha^4}{2} g_{pos}(p_0))(t - t_0)^2] e^{-\alpha(t-t_0)}\end{aligned}$$

vorgenommen. Sie bewirken, daß eine auf die Konsistenzmannigfaltigkeit rücktreibende Dynamik hinzutritt. Nun ist es so: ohne diese Dynamik kommt es zum Rangverlust zur Zeit

$$t = -\frac{p_0}{2v_0}.$$

Soll also der Rangverlust verhindert werden, muß die Dynamik vorher das System wieder in Richtung Konsistenzmannigfaltigkeit schieben. Für kleines p_0 und großes v_0 muß das aber sehr schnell passieren, d. h. α muß groß gewählt und damit das System steif gemacht werden. Der Zerfall des „problemlosen“ Gebiets bleibt bestehen, allerdings liegt der Konsistenzpunkt nun nicht mehr am Rand dieses Gebiets, sondern im Inneren.

§1.2 Orthogonale Relaxierung. Diese Probleme können behoben werden, wenn man eine andere Form der Relaxierung wählt. Zunächst einmal muß der Begriff der Relaxierung weiter gefaßt werden.

Definition 5 (Relaxation). Gegeben sei eine differenzierbare Funktion

$$h : D \rightarrow \mathbb{R}^k, \quad D \subseteq \mathbb{R}^n,$$

die als

$$h(x) = 0$$

k Zwangsbedingungen formuliert.

Eine Relaxation ist eine Funktion

$$\tilde{h} : D' \times D \rightarrow \mathbb{R}^k, \quad D' \subseteq \mathbb{R}^n, D \subseteq D'$$

mit den folgenden Eigenschaften:

(1) Sei $x_0 \in \mathbb{R}^n$. Gilt

$$h(x_0) = 0,$$

so sei

$$\forall x \in \mathbb{R}^n (\tilde{h}(x_0, x) = h(x)),$$

d. h. für konsistente Werte soll die Relaxation die ursprüngliche Bedingungsfunktion sein.

(2) $\tilde{h}(x_0, x_0) = 0$ für beliebige $x_0 \in D'$.

(3) $\tilde{h}(x_0, x)$ ist differenzierbar.

Die zu den Anfangswerten x_0 relaxierte Bedingung ist dann

$$\tilde{h}(x_0, x) = 0.$$

In praktischen Anwendungen wird man verlangen, daß $\tilde{h}(x_0, x)$ hinreichend oft nach x_0 differenzierbar ist und auch in x mindestens die Differenzierbarkeitsordnung von h hat. Auch die Erhaltung von hierarchischen Strukturen innerhalb der Bedingungen (wie z. B. $g_{vel} = \frac{d}{dt}g_{pos}$) wird man fordern.

Innerhalb dieses erweiterten Relaxierungsbegriffs ist nun eine andere Art von Relaxation möglich.

Definition 6 (Orthogonale Relaxation).

$$\tilde{g}_{pos,p_0,v_0}(t, p) = g_{pos}(p - v_i(t - t_0)) - g_{pos}(p_0)$$

$$\tilde{g}_{vel,p_0,v_0}(t, p, v) = g_{vel}(p - v_i(t - t_0), v - v_i) = G(p - v_i(t - t_0))(v - v_i)$$

Dabei ist v_i („inkonsistenter Anteil der Anfangsgeschwindigkeiten v_0 “) Lösung des folgenden Gleichungssystems

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ G(p_0)v_0 \end{pmatrix}.$$

Dabei muß vorausgesetzt werden, daß $G(p_0)$ Vollrang hat und damit die Matrix des obigen Gleichungssystems invertierbar ist.

Der Name „Orthogonale Relaxation“ bezieht sich darauf, daß die inkonsistenten Anfangsgeschwindigkeiten v_0 auf die konsistente Mannigfaltigkeit (beschrieben durch $G(p_0)v = 0$) projiziert wird, und zwar orthogonal in der Energienorm⁵.

Feststellung 3. Die orthogonale Relaxation ist eine Relaxation.

BEWEIS. Als wesentliche Eigenschaft von v_i wird

$$G(p_0)v_i = G(p_0)v_0 \tag{17}$$

benötigt, welches Bestandteil der Gleichung

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ G(p_0)v_0 \end{pmatrix} \tag{18}$$

ist. Es werden nun die drei Eigenschaften einer Relaxation nachgewiesen.

(1) Sei

$$\begin{aligned} g_{pos}(p_0) &= 0 \\ g_{vel}(p_0, v_0) &= G(p_0)v_0 = 0. \end{aligned}$$

⁵siehe VON SCHWERIN[66]

Dann ist $v_i = 0$, denn offenbar ist der Nullvektor Lösung des Gleichungssystems (18). Folglich gilt

$$\begin{aligned}\tilde{g}_{pos,p_0,v_0} &= g_{pos} \\ \tilde{g}_{vel,p_0,v_0} &= g_{vel}.\end{aligned}$$

(2) Es ist

$$\begin{aligned}\tilde{g}_{pos,p_0,v_0}(t_0, p_0) &= g_{pos}(p_0 - v_i(t_0 - t_0)) - g_{pos}(p_0) \\ &= g_{pos}(p_0) - g_{pos}(p_0) \\ &= 0 \\ \tilde{g}_{vel,p_0,v_0}(t_0, p_0, v_0) &= G(p_0 - v_i(t_0 - t_0))(v_0 - v_i) \\ &= G(p_0)(v_0 - v_i) \\ &= G(p_0)v_0 - G(p_0)v_i \\ &= 0 \text{ nach (17)}.\end{aligned}$$

Außerdem bleibt die hierarchische Struktur erhalten, denn

$$\begin{aligned}\frac{d}{dt}\tilde{g}_{pos,p_0,v_0}(t, p) &= \frac{d}{dt}g_{pos}(p - v_i(t - t_0)) \\ &= G(p - v_i(t - t_0))(v - v_i) \\ &= g_{vel}(p - v_i(t - t_0), v - v_i) \\ &= \tilde{g}_{vel,p_0,v_0}(t, p, v).\end{aligned}$$

(3) Offenbar hängt die orthogonale Relaxation differenzierbar von p_0 und v_0 ab, einzig die Abhängigkeit $v_i(p_0, v_0)$ ist zu beachten; v_i ist aber Lösung eines Gleichungssystems in diesen Variablen und $G(p_0)$ und ist damit differenzierbar abhängig von diesen unter der schon getroffenen Einschränkung, daß $G(p_0)$ keinen Rangverlust hat und daß G differenzierbar ist. Ist $G(p_0)$ aber rangdefizient, so ist keine Relaxation definiert.⁶

□

Zu welcher Veränderung der dynamischen Gleichungen führt nun die orthogonale Relaxation? Es ist

$$\begin{aligned}\tilde{G}(p) &:= \frac{\partial}{\partial p}\tilde{g}_{pos,p_0,v_0}(t, p) \\ &= G(p - v_i(t - t_0)) \\ \tilde{\gamma}(p, v) &:= -\frac{d^2}{dt^2}\tilde{g}_{pos,p_0,v_0}(t, p) + \tilde{G}(p)a \\ &= -\frac{d}{dt}G(p - v_i(t - t_0))(v - v_i) + \tilde{G}(p)a \\ &= -(v - v_i)^T \frac{\partial}{\partial p}G(p - v_i(t - t_0))(v - v_i) - G(p - v_i(t - t_0))a + \\ &\quad + G(p - v_i(t - t_0))a \\ &= -(v - v_i)^T \frac{\partial}{\partial p}G(p - v_i(t - t_0))(v - v_i) \\ &= \gamma(p - v_i(t - t_0), v - v_i).\end{aligned}$$

⁶Detailliertere Angaben zur Abhängigkeit von Anfangswerten finden sich im Abschnitt §1.4.

Damit gilt also folgende Dynamik:

$$\begin{aligned}
& \dot{p} = v \\
& \dot{v} = a \\
& \begin{pmatrix} M & G^T(p - v_i(t - t_0)) \\ G(p - v_i(t - t_0)) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} = \\
& \begin{pmatrix} F(t, p, v, q) \\ \gamma(p - v_i(t - t_0), v - v_i) \end{pmatrix} \\
& \tilde{g}_{pos, p_0, v_0}(t, p) = 0 \\
& \tilde{g}_{vel, p_0, v_0}(t, p, v) = 0.
\end{aligned} \tag{19}$$

Nun wird eine Modifikation dieser Gleichung vorgenommen, die später gerechtfertigt werden soll, und zwar

$$F(t, p, v, q) \Rightarrow F(t, p - v_i(t - t_0), v - v_i, q). \tag{20}$$

Danach lauten die Gleichungen (19)

$$\begin{aligned}
& \dot{p} = v \\
& \dot{v} = a \\
& \begin{pmatrix} M & G^T(p - v_i(t - t_0)) \\ G(p - v_i(t - t_0)) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} = \\
& \begin{pmatrix} F(t, p - v_i(t - t_0), v - v_i, q) \\ \gamma(p - v_i(t - t_0), v - v_i) \end{pmatrix} \\
& g_{pos}(p - v_i(t - t_0)) - g_{pos}(p_0) = 0 \\
& g_{vel}(p - v_i(t - t_0), v - v_i) = 0.
\end{aligned} \tag{21}$$

Substituiert man nun

$$\begin{aligned}
P(t) &:= p(t) - v_i(t - t_0) \\
V(t) &:= v(t) - v_i,
\end{aligned}$$

so geht (21) über in

$$\begin{aligned}
& \dot{P} = V \\
& \dot{V} = a \\
& \begin{pmatrix} M & G^T(P) \\ G(P) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} = \begin{pmatrix} F(t, P, V, q) \\ \gamma(P, V) \end{pmatrix} \\
& g_{pos}(P) - g_{pos}(p_0) = 0 \\
& g_{vel}(P, V) = 0.
\end{aligned} \tag{22}$$

Dies ist aber nichts anderes als die Gleichung (11) mit den Anfangswerten $P(0) = p(0)$ und $V(0) = v_0 - v_i$ und einer linearen Relaxation in den eventuell inkonsistenten Positionen! Damit ergibt sich der folgende

ALGORITHMUS DER ORTHOGONALEN RELAXATION

(1) Seien p_0 und v_0 gegeben. Berechne v_i gemäß

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ G(p_0)v_0 \end{pmatrix}.$$

(2) (Substitution) Setze $v_0 := v_0 - v_i$. Danach ist v_0 konsistent.

(3) (Innere Integration) Integriere die MBS-Gleichung

$$\begin{aligned} \dot{p} &= v \\ \dot{v} &= a \\ \begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} &= \begin{pmatrix} F(t, p, v, q) \\ \gamma(p, v) \end{pmatrix} \\ g_{pos}(p) - g_{pos}(p_0) &= 0 \\ g_{vel}(p, v) &= 0 \end{aligned}$$

von t_0 bis t_{end} , wobei die Ortsbedingungen linear relaxiert wurden.

(4) Substituiere zurück:

$$\begin{aligned} p(t_{end}) &:= p(t_{end}) + v_i(t_{end} - t_0) \\ v(t_{end}) &:= v(t_{end}) + v_i. \end{aligned}$$

Offensichtlich ist der Aufwand der orthogonalen Relaxation nicht größer als der der linearen. Während der Integration ist der Aufwand sogar geringer, weil auf Positionsebene nur Konstanten von den Ortszwangsbedingungen subtrahiert werden müssen, und keine lineare Funktionen. Bei den Geschwindigkeitszwangsbedingungen muß sogar nichts verändert werden. Ein Mehraufwand entsteht lediglich zur Berechnung des v_i . Die dafür notwendige Zerlegung der MBS-Matrix $\begin{pmatrix} M & G^T \\ G & 0 \end{pmatrix}$ zur Lösung des Gleichungssystems ist aber nicht mit Zeitkosten verbunden, da sie bei Integrationsbeginn sowieso benötigt wird.

Nun muß noch die Ersetzung (20) gerechtfertigt werden. Für dieses Vorgehen gibt es vier gute Gründe:

- (1) Diese Ersetzung ist die Basis dafür, daß die orthogonale Relaxation nicht langsamer ist als die lineare. Behielte man das ursprüngliche F bei, so wären die globalen Ersetzungen $p \rightarrow P$ und $v \rightarrow V$ nicht möglich bzw. bei jeder Auswertung von F müsste man statt auf P und V auf $P + v_i(t - t_0)$ und $V + v_i$ zurückgreifen.
- (2) Durch diese Ersetzung tauchen als Argumente von F nur konsistente v auf. Sonst müsste man sich gegebenenfalls überlegen, wie man F auf inkonsistente v fortsetzt, was eventuell gar nicht möglich ist.
- (3) War p_0 konsistent ($g_{pos}(p_0) = 0$), so bleibt $p - v_i(t - t_0)$ konsistent, i. allg. aber nicht p , und man braucht F nicht auf inkonsistente p zu erweitern, was eventuell nicht möglich ist.
- (4) Sind die Anfangsgeschwindigkeiten konsistent, so findet keine Veränderung von F statt, da dann $v_i = 0$ ist. Weiter ist die Veränderung von F differenzierbar in p_0 und v_0 , da v_i dies ist (immer unter der Annahme, daß $G(p_0)$ keinen Rangverlust hat und differenzierbar ist).

Verwendet man nun für das Beispiel 2 die orthogonale Relaxation, so ergibt sich als totale Lösung

$$p(t) = p_0 + v_0 t, \quad v(t) = v_0$$

für beliebiges $p_0 \neq 0$ und beliebiges v_0 . Hier zeigt sich ein Vorteil der orthogonalen Relaxation: i. allg. ergibt sich die gleiche Lösungstrajektorie unabhängig von der Art der Formulierung der Zwangsbedingung. Dies soll genauer formuliert werden.

Feststellung 4 (Forminvarianz der orthogonalen Relaxierung). Sei $g_{pos} : \mathbb{K}^n \rightarrow \mathbb{K}^k$ und sei $h : \mathbb{K}^k \rightarrow \mathbb{K}^k$ eine differenzierbare Funktion und $h'(g_{pos}(p_0))$ sei invertierbar. Dann liefert der Algorithmus der orthogonalen Relaxierung die gleiche Lösung für das Problem (7) einmal mit den Zwangsbedingungen als

$$g_{pos}(p) = 0,$$

ein andermal mit den Zwangsbedingungen als

$$h(g_{pos}(p)) = 0$$

formuliert.

BEWEIS. Sei

$$\tilde{g}_{pos}(p) := h(g_{pos}(p)), \quad \tilde{G} := \tilde{g}'.$$

Dann gilt

$$\tilde{G}(p) = h'(g_{pos}(p))G(p)$$

und

$$\tilde{g}_{vel}(p(t)) := \frac{d}{dt} \tilde{g}_{pos}(p(t)) = \tilde{G}(p(t))v(t).$$

Weiter wird $\tilde{\gamma}$ bestimmt durch die folgende Überlegung:

$$\begin{aligned} \frac{d}{dt} \tilde{G}(p(t))v(t) &= 0 \\ \Rightarrow \tilde{G}(p(t))a(t) + \frac{d}{dt} \left(\begin{array}{c} h'(\underbrace{g_{pos}(p(t))}_{= \text{const bei orth. Rel.}})G(p(t)) \end{array} \right) v(t) &= 0 \\ \Rightarrow \tilde{G}(p(t))a(t) + h'(g_{pos}(p(t)))v^T(t)G_p(p(t))v(t) &= 0. \end{aligned}$$

Dies führt zur Definition

$$\tilde{\gamma} := -h'(g_{pos}(p(t)))v^T(t)G_p(p(t))v(t).$$

Sei nun v_i, ξ die Lösung von

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ G(p_0)v_0 \end{pmatrix}.$$

Dann ist $v_i, (h'(g_{pos}(p_0)))^T)^{-1}\xi$ Lösung von

$$\begin{pmatrix} M & \tilde{G}^T(p_0) \\ \tilde{G}(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi' \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{G}(p_0)v_0 \end{pmatrix}.$$

Damit ist die Bestimmung von v_i in beiden Fällen gleich.

Für die Integration gilt aber (wegen $g_{pos}(p(t)) - g_{pos}(p_0) \equiv 0$)

$$\begin{aligned} & \begin{pmatrix} M & \tilde{G}^T(p) \\ \tilde{G}(p) & 0 \end{pmatrix} \begin{pmatrix} \tilde{a} \\ \xi' \end{pmatrix} = \begin{pmatrix} F \\ \tilde{\gamma} \end{pmatrix} \\ \iff & \begin{pmatrix} M & G^T(p)h'(g_{pos}(p_0))^T \\ h'(g_{pos}(p_0))G(p) & 0 \end{pmatrix} \begin{pmatrix} \tilde{a} \\ \xi' \end{pmatrix} = \begin{pmatrix} F \\ h'(g_{pos}(p_0))\gamma \end{pmatrix}. \end{aligned}$$

Dies hat offensichtlich die gleiche Lösung a wie die Lösung von

$$\begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} a \\ \xi \end{pmatrix} = \begin{pmatrix} F \\ \gamma \end{pmatrix}.$$

Damit ergeben sich für beide Formulierungen das gleiche a , mithin die gleiche Lösungstrajektorie. \square

Bei linearer Relaxierung gilt Feststellung 4 nicht, wie Beispiel 2 zeigt.

Die nächste Feststellung gibt nun ein hinreichendes Kriterium, für welche Startwerte eine globale Lösung existiert.

Feststellung 5. Sei p_0 so, daß $G(p_0)$ Vollrang hat. Gilt für alle $p \in \mathbb{R}$ mit $g_{pos}(p) = g_{pos}(p_0)$ auch

$G(p)$ hat Vollrang,

so ist die MBS-Matrix

$$\begin{pmatrix} M & G^T \\ G & 0 \end{pmatrix}$$

bei Verwendung der orthogonalen Relaxation auf der ganzen Lösungstrajektorie invertierbar und die Integration läuft nicht in eine Singularität.

BEWEIS. Da

$$g_{pos}(P) - g_{pos}(p_0) = 0$$

invariant ist, ergeben sich im Laufe der Integration nur Werte $P(t)$ für die gilt

$$g_{pos}(P(t)) = g_{pos}(p_0).$$

Nach Voraussetzung hat dann aber $G(P(t))$ Vollrang, mithin bricht die Integration nicht zusammen. \square

Die Voraussetzungen der obigen Feststellung sind häufig gegeben. So gilt z. B. für die Funktion

$$g_{pos}(p) = p^2 - 1,$$

daß es nur einen Rangverlust für $p = 0$ gibt, und diese Stelle ist durch $g_{pos}(p) = -1$ charakterisiert. Hat man also ein p_0 mit $g_{pos}(p_0) \neq -1$, so sind alle p mit $g_{pos}(p) = g_{pos}(p_0)$ ebenfalls von 0 verschieden. Startet man also nicht im Rangverlust, so gelangt man auch in keinen hinein. Das Gebiet aller global integrierbaren Initialwerte besteht also aus ganz \mathbb{R}^2 ohne die Achse $\{p_0 = 0\}$.

Bemerkung: Schaltfunktionen und Ausgaberroutinen von Zwischenergebnissen können wahlweise von P, V bzw. p, v abhängig gemacht werden. Die Ausgabe des Endergebnisses muß sich natürlich auf p, v beziehen. Dabei ist es einfacher, Schaltfunktionen auf P zu beziehen (dies wird im folgenden angenommen). Ausgaberroutinen sollten sich allerdings eher auf p, v beziehen, sonst kommt es zu unerwarteten Differenzen zum Endergebnis.

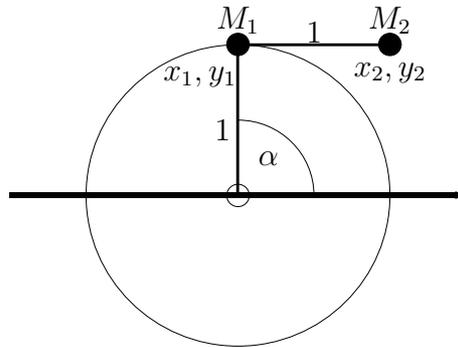
§1.3 Orthogonale Relaxierung in gemischten Koordinaten. Die Verwendung der MBS-Matrix im Schritt (1) des Algorithmus geschieht aus Aufwandsgründen; wesentlich ist nur die Gleichung

$$G(p_0)v_i = G(p_0)v_0.$$

Mit anderen Worten: es wird auf die Konsistenzmannigfaltigkeit $G(p_0)v = 0$ projiziert, aber in welcher Norm dies geschieht, steht zur freien Wahl. Die Wahl im Schritt (1) war dadurch motiviert, daß damit kein numerischer Mehraufwand entsteht. Die zu zerlegende Matrix wird für die Integration sowieso benötigt. Man könnte aber für andere Zwecke eine andere Matrix verwenden, müsste deren Zerlegung aber in Kauf nehmen (was nicht besonders tragisch ist, da es sich nur um eine Zerlegung handelt im Vergleich zur wahrscheinlich viel aufwendigeren Integration).

Ein solch anderer Zweck soll hier beispielhaft vorgeführt werden. Dabei wird sich die Flexibilität der orthogonalen Relaxation zeigen. Ausgangspunkt ist die Formulierung eines Mehrkörpermodells, das teilweise in Minimalkoordinaten und teilweise in abhängigen Koordinaten vorliegen soll.

Beispiel 3.



Ein einfacher Mechanismus mit unterschiedlichen Koordinaten beschrieben

Das angegebene Modell bestehe aus einer Masse M_1 , die in einem festen Abstand vom Ursprung diesen umlaufen kann und einer Masse M_2 , die einen festen Abstand zu M_1 habe.

Der vom üblichen abweichende Zweck bestehe nun darin, daß auf Optimiererseite andere Koordinaten als auf Integratorseite verwendet werden sollen. Der Optimierer soll ein Gemisch aus minimalen und nichtminimalen Koordinaten verwenden. So werde die Position von M_1 durch den Winkel α eindeutig festgelegt, die Position von M_2 durch die Koordinaten x_2 und y_2 . Analog werden auf Geschwindigkeitsebene die Koordinaten v_α , v_{x_2} und v_{y_2} benutzt.

Der Integrator dagegen soll die Koordinaten x_1 , y_1 und x_2 , y_2 sowie die dazugehörigen Geschwindigkeiten v_{x_1} , v_{y_1} , v_{x_2} , v_{y_2} verwenden. Der Integrator benötigt außerdem die beiden Zwangsbedingungen

$$g_{pos1}(x_1, y_1) = x_1^2 + y_1^2 - 1$$

$$g_{pos2}(x_1, y_1, x_2, y_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2 - 1$$

und deren zeitliche Ableitung g_{vel1} und g_{vel2} .

Durch die Wahl der Koordinaten auf Optimiererseite sind die vom Optimierer gelieferten

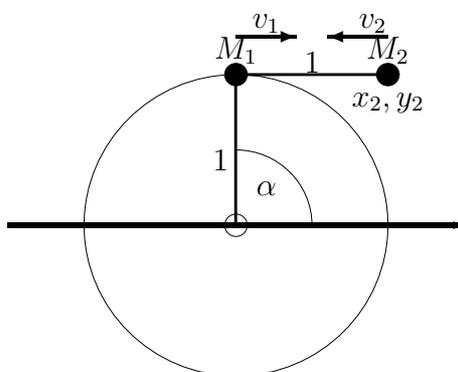
Initialwerte immer derart, daß die erste Zwangsbedingung g_{pos1} sowie g_{vel1} automatisch erfüllt ist, die zweite dagegen i. allg. nicht. Nun seien konsistente initiale p -Werte und inkonsistente initiale v -Werte gegeben

$$\begin{aligned}\alpha &= \pi/2 \\ x_2 &= 1 \\ y_2 &= 1 \\ v_\alpha &= -1 \\ v_{x_2} &= -1 \\ v_{y_2} &= 0.\end{aligned}$$

Werden diese in Koordinaten des Integrators umgerechnet, so resultiert

$$\begin{aligned}x_1 &= 0 \\ y_1 &= 1 \\ x_2 &= 1 \\ y_2 &= 1 \\ v_{x_1} &= 1 \\ v_{y_1} &= 0 \\ v_{x_2} &= -1 \\ v_{y_2} &= 0,\end{aligned}$$

so daß sich folgendes Bild ergibt:



Ausgangssituation
mit inkonsistentem v .

Wendet man hierauf die orthogonale Relaxierung an, so ergibt sich für die Berechnung des v_i :

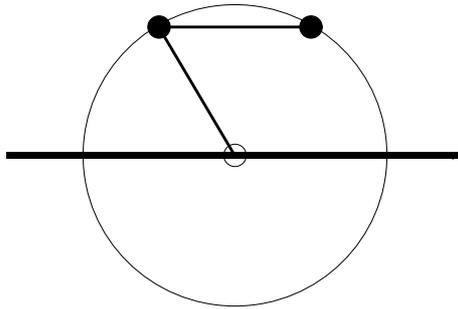
$$v_i = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

Damit wird die Integration also mit

$$v - v_i = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

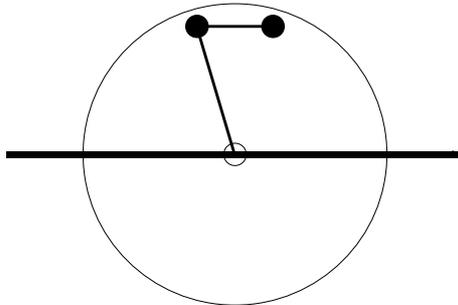
gestartet, einem konsistenten Wert.

Nach der Integration wird aber auf die Orte der Wert $v_i(t - t_0)$ und auf die Geschwindigkeiten v_i aufgeschlagen. Interessant ist, was mit den Orten passiert. Angenommen das Integrationsende wäre nach $\Delta t = 0.25$ und die Stellung sieht so aus:



Ergebnis nach der
(inneren) Integration.

Dann sieht die Stellung nach Addition von $v_i(t - t_0)$ so aus:



Endergebnis.

Insbesondere ist die Position der Masse M_1 so, daß sie nicht mehr konsistent mit g_{pos1} ist. Dies bereitet Schwierigkeiten, wenn man von (x_1, y_1) wieder auf α zurückrechnen will, um dieses Ergebnis dann dem Optimierer mitteilen zu können. Noch schwieriger ist im übrigen in solch einer Stellung die Rückrechnung von v_{x1}, v_{y1} auf v_α . Und stellt man die numerischen Werte so ein, daß sich am Ende der Rechnung für M_1 die Position $x_1 = 0, y_1 = 0$ ergibt, so ist eine Rückrechnung gänzlich unmöglich.

Dieses Problem kann man umgehen. Man erzwingt, daß zu den Koordinaten, die man nicht direkt zurückreicht, sondern umrechnen muß, die entsprechende Komponente von v_i gleich 0 ist. Denn dann wird nach der Integration an diesen Koordinaten nichts mehr

geändert und sie können problemlos transformiert werden. Im obigen Beispiel würde man also erzwingen, daß von v_i die ersten beiden Komponenten gleich 0 wären. In der praktischen Umsetzung geht man am besten wie folgt vor: dem Gleichungssystem

$$G(p_0)v_i = G(p_0)v_0 \quad (23)$$

fügt man Gleichungen hinzu, daß gewisse Werte unverändert bleiben sollen. Im obigen Beispiel 3 ergibt sich:

$$G = \begin{pmatrix} 0 & 2 & 0 & 0 \\ -2 & 0 & 2 & 0 \end{pmatrix}.$$

Möchte man v_{x1} und v_{y1} unverändert lassen, so fügt man hinzu:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} v_i = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Nun löst man das gemeinsame Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ -2 & 0 & 2 & 0 \end{pmatrix} v_i = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -4 \end{pmatrix}.$$

Dieses Gleichungssystem ist zwar rangdefizient, hat aber z. B. die Lösung⁷

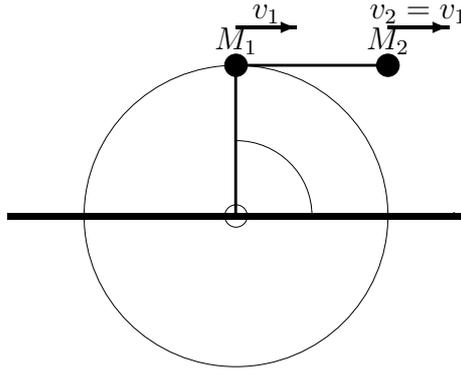
$$v_i = \begin{pmatrix} 0 \\ 0 \\ -2 \\ 0 \end{pmatrix}.$$

Subtrahiert man nun von v_0 das v_i , so wie es die orthogonale Relaxierung vorsieht, erhält man

$$\begin{aligned} v_{x1} &= 1 \\ v_{y1} &= 0 \\ v_{x2} &= 1 \\ v_{y2} &= 0, \end{aligned}$$

also konsistente Geschwindigkeiten bei unverändertem v_{x1}, v_{y1} :

⁷ v_{i4} kann beliebig gewählt werden, ohne die Konsistenz zu zerstören. Ist der Rangverlust an dieser Stelle störend, so können weitere Forderungen hinzugefügt werden, z. B. minimale Norm oder keine Veränderung in v_{y2} .



Startsituation der (inneren)
Integration bei angepasster Relaxierung.

§1.4 Orthogonale Relaxierung und Wronski-Matrix. Verwendet man orthogonale Relaxierung, so muß dies natürlich bei der Berechnung von Wronski-Matrizen, d. h. Ableitungen nach Anfangswerten und Parametern, berücksichtigt werden. Hier sollen nun die entsprechenden Formeln hergeleitet werden.

Seien dazu die folgenden Abkürzungen eingeführt: y fasse Orte, Geschwindigkeiten und Parameter zusammen:

$$y := \begin{pmatrix} p \\ v \\ q \end{pmatrix}.$$

Sei n die Dimension von p bzw. auch von v und sei E_n die $n \times n$ -Einheitsmatrix. Die Lösung der DAE zum Zeitpunkt T bei Ortsanfangswert p_0 und konsistentem Geschwindigkeitsanfangswert v_0 zum Startzeitpunkt⁸ t_0 und (konstanten) Parametern q werde mit

$$Y(T, p_0, v_0, q)$$

bezeichnet; Y ist also das Ergebnis der (inneren) Integration, angesetzt auf p_0 , v_0 und q . Das Gesamtergebnis der orthogonale Relaxierung und Integration werde mit

$$\tilde{Y}(T, p_0, v_0, q)$$

abgekürzt. $v_i(p_0, v_0)$ sei der inkonsistente Teil der Anfangsgeschwindigkeiten, der gemäß Algorithmus abgespaltet wird. Das Verfahren lässt sich dann kurz so beschreiben:

$$\tilde{Y}(T, p_0, v_0, q) = Y(T, p_0, v_0 - v_i, q) + \begin{pmatrix} v_i(T - t_0) \\ v_i \\ 0 \end{pmatrix}. \quad (24)$$

⁸Ableitungen nach t_0 sollen nicht betrachtet werden, da sie von geringem Interesse sind und die Herleitung der entsprechenden Formeln analog verläuft.

Unmittelbar aus (24) folgt als erster Zwischenschritt

$$\begin{aligned} \frac{\partial}{\partial p_0} \tilde{Y}(T, p_0, v_0, q) &= \frac{\partial}{\partial p_0} Y(T, p_0, v_0 - v_i, q) + \\ &\quad + \frac{\partial}{\partial v_0} Y(T, p_0, v_0 - v_i, q) \left(-\frac{\partial}{\partial p_0} v_i \right) + \begin{pmatrix} (T - t_0) \frac{\partial}{\partial p_0} v_i \\ \frac{\partial}{\partial p_0} v_i \\ 0 \end{pmatrix} \quad (25) \\ \frac{\partial}{\partial v_0} \tilde{Y}(T, p_0, v_0, q) &= \frac{\partial}{\partial v_0} Y(T, p_0, v_0 - v_i, q) \left(E_n - \frac{\partial}{\partial v_0} v_i \right) + \begin{pmatrix} (T - t_0) \frac{\partial}{\partial v_0} v_i \\ \frac{\partial}{\partial v_0} v_i \\ 0 \end{pmatrix} \\ \frac{\partial}{\partial q} \tilde{Y}(T, p_0, v_0, q) &= \frac{\partial}{\partial q} Y(T, p_0, v_0, q). \end{aligned}$$

Ist der Integrator in der Lage, auch Sensitivitäten gleich mitzuberechnen (z. B. über Interne Numerische Differentiation), so liefert er die Terme

$$\frac{\partial}{\partial p_0} Y(T, p_0, v_0 - v_i, q), \quad \frac{\partial}{\partial v_0} Y(T, p_0, v_0 - v_i, q), \quad \frac{\partial}{\partial q} Y(T, p_0, v_0, q)$$

gleich mit. Um daraus die letztendlich interessierenden Ableitungen von \tilde{Y} zu gewinnen, braucht man dann nur noch die Terme

$$\frac{\partial}{\partial p_0} v_i, \quad \frac{\partial}{\partial v_0} v_i.$$

Zur Berechnung dieser Matrizen geht man von der v_i bestimmenden Gleichung aus:

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ G(p_0)v_0 \end{pmatrix}. \quad (26)$$

Zunächst wird der einfachere Fall angenommen, nämlich $\frac{\partial}{\partial v_0} v_i$. Dazu leitet man die Gleichung (26) nach der j -ten Komponente v_0^j von v_0 ab:

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial v_i}{\partial v_0^j} \\ \frac{\partial \xi}{\partial v_0^j} \end{pmatrix} = \begin{pmatrix} 0 \\ G(p_0)e_j \end{pmatrix},$$

wobei e_j der j -te Einheitsvektor ist und $\frac{\partial v_i}{\partial v_0^j}$ ist die j -te Spalte von $\frac{\partial v_i}{\partial v_0}$. Dieses Gleichungssystem wird nun für $j = 1, \dots, n$ gelöst. Praktischerweise ist die Matrix auf der linken Seite gleich geblieben, so daß sie nur einmal zerlegt werden muß und dann dazu dient nacheinander alle Spalten von $\frac{\partial v_i}{\partial v_0}$ zu liefern. Der Term $\frac{\partial \xi}{\partial v_0}$ ist dabei uninteressant, seine Berechnung kann unterdrückt werden.

Nun folgt der kompliziertere Fall, $\frac{\partial}{\partial p_0} v_i$. Leitet man (26) nach der j -ten Komponente von p_0 ab, so ergibt sich

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial v_i}{\partial p_0^j} \\ \frac{\partial \xi}{\partial p_0^j} \end{pmatrix} + \begin{pmatrix} \frac{\partial}{\partial p_0^j} M & \frac{\partial}{\partial p_0^j} G^T(p_0) \\ \frac{\partial}{\partial p_0^j} G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\partial}{\partial p_0^j} G(p_0)v_0 \end{pmatrix}.$$

Diese Gleichung ermöglicht die spaltenweise Berechnung von $\frac{\partial}{\partial p_0} v_i$. Die zu zerlegende Matrix ist wieder dieselbe wie oben. Allerdings braucht man hier den Wert von ξ sowie die „Spalten“ $\frac{\partial}{\partial p_0^j} G(p_0)$ des Tensors $\frac{\partial}{\partial p} G(p)$. Diese sollten analytisch als Modellbestandteil vorliegen.

§2 Inkonsistente Startpositionen

Im vorigen Abschnitt wurde des öfteren „ $G(p_0)$ habe keinen Rangverlust“ vorausgesetzt. Auch in der Anmerkung nach Feststellung 5 wurde nur gesagt: „Startet das System nicht im Rangverlust, so läuft es auch in keinen.“ Ebenso wird das Gebiet der „problemlosen“ Startwerte für das Beispiel 2 durch die Achse $\{p_0 = 0\}$ immer noch in zwei unzusammenhängende Gebiete geteilt (wenn auch der Konsistenzpunkt nun nicht mehr am Rand dieses Gebiets liegt). Dennoch drängt sich die Frage auf:

„Gibt es eine Relaxierung, die die Integration für beliebige Startwerte erlaubt?“

In diesem Abschnitt soll auf diese Frage geantwortet werden. Generell sei in diesem Abschnitt vorausgesetzt, daß die Konsistenzmannigfaltigkeit selbst keine Stelle mit Rangverlust besitzt, d. h.

$$\forall p (g_{pos}(p) = 0 \rightarrow G(p) \text{ hat Vollrang}).$$

Im Kapitel III soll dann die Verletzung dieser Voraussetzung untersucht werden. Da im vorigen Abschnitt die Geschwindigkeitsinkonsistenz umfassend geklärt wurde, darf man im folgenden davon ausgehen, daß keine Geschwindigkeitsinkonsistenzen vorliegen.

§2.1 Unbehebbarkeit von Rangverlusten. Von Interesse ist der folgende Begriff:

Definition 7 (Globaler Vollrang). *Eine Relaxation*

$$\tilde{h} : D' \times D \rightarrow \mathbb{R}^k$$

hat globalen Vollrang, wenn gilt, daß für beliebige $(p_0, p) \in D' \times D$ mit $\tilde{h}(p_0, p) = 0$ die Ableitungsmatrix

$$\frac{\partial}{\partial p} \tilde{h}(p_0, p)$$

Vollrang hat.

Diese Definition ist strenger als eigentlich notwendig, denn ausgehend von einem bestimmten p_0 ist nicht jedes p in der Dynamik erreichbar. Dieser Gedanke ist aber in voller Allgemeinheit zu kompliziert.

Zunächst seien zwei Beispiele eingeführt.

Beispiel 4.

Das System werde durch vier Variablen a, b, c, d beschrieben und durch die Zwangsbedingung

$$g_{pos}(a, b, c, d) = (a - b)(c - d) - 1.$$

Beispiel 5.

Das System werde durch eine Variable p beschrieben und durch die Zwangsbedingung

$$g_{pos}(p) = p^2 - 1.$$

Dies entspricht dem Beispiel 2 aus dem vorherigen Abschnitt.

Von Interesse ist der Rangverlust der Ableitungsmatrix G der Beispiele. Im Beispiel 4 gilt

$$G_{Bsp4}(a, b, c, d) = \begin{pmatrix} c - d & d - c & a - b & b - a \end{pmatrix},$$

im Beispiel 5

$$G_{Bsp5}(p) = \begin{pmatrix} 2p \end{pmatrix}.$$

G_{Bsp4} verliert seinen Vollrang dann und genau dann, wenn $a = b$ und $c = d$ ist. G_{Bsp5} hat Rangverlust in $p = 0$.

Feststellung 6. Für Beispiel 4 gibt es eine Relaxation, so daß kein Rangverlust mehr möglich ist, d. h. eine Relaxation mit globalem Vollrang.

BEWEIS. Die Relaxation lautet

$$\begin{aligned} \tilde{g}_{Bsp4}(a, b, c, d) = & f_0((a - b)(c - d) - 1) + (1 - f_0)(a + b + c + d) - \\ & - f_0((a_0 - b_0)(c_0 - d_0) - 1) - (1 - f_0)(a_0 + b_0 + c_0 + d_0) \end{aligned}$$

mit

$$f_0 = (a_0 - b_0)(c_0 - d_0).$$

Die Ableitungsmatrix $\tilde{G}_{Bsp4}(a, b, c, d)$ lautet

$$\left(\begin{array}{cccc} f_0(c - d) + (1 - f_0) & f_0(d - c) + (1 - f_0) & f_0(a - b) + (1 - f_0) & f_0(b - a) + (1 - f_0) \end{array} \right).$$

\tilde{G}_{Bsp4} hat Rangverlust, wenn alle Komponenten gleich 0 sind. Sei dies der Fall für ein a, b, c, d mit $\tilde{g}_{Bsp4}(a, b, c, d) = 0$. Aus der Tatsache, daß die ersten beiden Komponenten gleich 0 sind, folgert man

$$2 - 2f_0 = 0$$

oder

$$f_0 = 1.$$

Damit geht die Bedingung, daß alle Komponenten von \tilde{G}_{Bsp4} gleich 0 sind, über in

$$\begin{aligned} a - b &= 0 \\ c - d &= 0. \end{aligned}$$

Nun gilt aber

$$\tilde{g}_{Bsp4}(a, b, c, d) = 0,$$

was sich nun schreiben lässt als

$$-1 - (a_0 - b_0)(c_0 - d_0) + 1 = 0$$

oder

$$(a_0 - b_0)(c_0 - d_0) = 0.$$

Dies ist aber nichts anderes als

$$f_0 = 0,$$

was ein Widerspruch zu obigem $f_0 = 1$ ist. □

Im Gegensatz dazu gilt

Feststellung 7. Es gibt für Beispiel 5 keine Relaxation mit globalem Vollrang, so daß das Definitionsgebiet der Relaxation die Konsistenzmenge umfasst und konvex ist, d. h. es gibt keine Relaxation von $\{p \mid \|p\| = 1\} \subseteq D$ auf

$$\tilde{g}_{Bsp5}(p_0, p) : D' \times D \rightarrow \mathbb{R}, D \subseteq D', D' \text{ konvex,}$$

die Vollrang auf D' hat.

BEWEIS. Angenommen \tilde{g}_{Bsp5} sei so eine Relaxation. Seien

$$\begin{aligned} E &= \{p \mid \|p\| = 1\}, \\ E_0 &= \{p \mid \|p\| \leq 1\}. \end{aligned}$$

Da $D' E$ umfassen und konvex sein soll, umfasst es auch E_0 . Nach der Definition einer Relaxation muß aber gelten

$$\forall p \in E \quad (\tilde{g}_{Bsp5}(p, p) = 0).$$

Nun ist E_0 kompakt, daher nimmt \tilde{g}_{Bsp5} ein Maximum und ein Minimum auf E_0 an. Ist nun der Wert im Maximum und im Minimum gleich 0, so ist \tilde{g}_{Bsp5} auf ganz E_0 identisch gleich 0, somit hat \tilde{g}_{Bsp5} auf ganz E_0 Rangverlust.

Ist dagegen der Wert im Maximum oder Minimum von 0 verschieden, so wird das Maximum bzw. Minimum nicht am Rand E angenommen (denn dort sind alle Werte gleich 0), sondern im Inneren von E an einer Stelle p_M . Dort gilt dann aber $\tilde{g}'_{Bsp5}(p_M) = 0$. Dann aber hat \tilde{g}_{Bsp5} keinen globalen Vollrang. \square

§2.2 Globaler Vollrang. Trotz der negativen Ergebnisse der Feststellung 7 gibt es eine Möglichkeit für $\mathbb{K} = \mathbb{R}$ globalen Vollrang zu erreichen. Dieses Verfahren soll hier vorgestellt werden. Im Rahmen der Definition von Relaxierung ist dies nach besagter Feststellung nicht möglich. Die Definition müsste erweitert werden, was hier aber nicht geschehen soll, da die Eingriffe in die Modellformulierung bereits keine kleinen Korrekturen mehr sind. Daher soll nicht mehr von Relaxierung, sondern von Modellformulierung gesprochen werden.

Der wesentliche Unterschied zu den Relaxierungen besteht darin, daß sich nun die Anzahl der Variablen ändert. Was bei gleichbleibender Variablenzahl nicht möglich ist, wird bei zunehmender Zahl machbar.

Feststellung 8. *Es sei ein Mehrkörpermodell in Deskriptorform (7) gegeben. Die Anzahl der Positionsvariablen p sei n . Von der gleichen Dimension sei v , insgesamt habe das System also $2n$ Variable (zzgl. eventueller Parameter). Die Anzahl der Zwangsbedingungen bzw. Invarianten auf Positionsebene sei k , d. h. g_{pos} ist eine Funktion mit Bild im \mathbb{R}^k . Dann gibt es eine Formulierung dieses Modells in $2n + 2k$ Variable (zzgl. eventueller Parameter), die globalen Vollrang besitzt.*

BEWEIS. Das Modell laute ursprünglich

$$\begin{aligned} \dot{p} &= v \\ \dot{v} &= a \\ \begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} &= \begin{pmatrix} F(t, p, v, q) \\ \gamma(p, v) \end{pmatrix} \\ g_{pos}(p) &= 0 \\ g_{vel}(p, v) &= 0. \end{aligned}$$

Es treten nun auf Ortsebene k neue Variablen \mathbf{p} und auf Geschwindigkeitsebene k neue Variablen \mathbf{v} hinzu. Als Hilfsvariable zur Formulierung taucht auch \mathbf{a} auf. Ebenfalls benötigt werden k „Pseudomassen“ \mathbf{m} , die konstant gesetzt werden können (am einfachsten gleich

1). Das System wird in einem ersten Schritt erweitert zu

$$\begin{pmatrix} \dot{p} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} v \\ \mathbf{v} \end{pmatrix}$$

$$\begin{pmatrix} \dot{v} \\ \dot{\mathbf{v}} \end{pmatrix} = \begin{pmatrix} a \\ \mathbf{a} \end{pmatrix}.$$

Die algebraische Gleichung verändert sich zu

$$\begin{pmatrix} \begin{pmatrix} M & 0 \\ 0 & \mathbf{m} \end{pmatrix} & \mathbf{G}^T(p, \mathbf{p}) \\ \mathbf{G}(p, \mathbf{p}) & 0 \end{pmatrix} \begin{pmatrix} a \\ \mathbf{a} \\ \lambda \end{pmatrix} = \begin{pmatrix} F(t, p, v, q) \\ \mathfrak{J}(t, p, v, q, \mathbf{p}, \mathbf{v}) \\ \gamma(p, v) \end{pmatrix}, \quad (27)$$

wobei \mathfrak{J} noch zu wählen ist (z. B. zu $\equiv 0$) und \mathbf{G} aus den veränderten Zwangsbedingungen resultiert. Diese werden geändert und dann relaxiert (linear/orthogonal/Baumgarte). Sei $g_{pos,i}$ die Zwangsbedingung Nr. i und G_i die i -te Zeile der Ableitungsmatrix G . Dann sei

$$\mathfrak{g}_{pos,i}(p, \mathbf{p}) := g_{pos,i}(p) + \mathbf{p}_i \sum_{j=1}^k g_{pos,j}(p)^2. \quad (28)$$

Unterteilt man die Zeile \mathbf{G}_i in einen ersten Block mit n Einträgen und einen zweiten Block mit k Einträgen, so lässt sie sich schreiben als

$$\mathbf{G}_i = \left(G_i + \mathbf{p}_i \sum_{j=1}^k 2g_{pos,j}(p)G_j(p) \quad \sum_{j=1}^k g_{pos,j}(p)^2 e_i^T \right).$$

Dabei ist e_i der i -te Einheitsbasisvektor. Führt man die Abkürzungen

$$C(p) := \sum_{j=1}^k 2g_{pos,j}(p)G_j(p)$$

$$c(p) := \sum_{j=1}^k g_{pos,j}(p)^2$$

$$E_k := k \times k \text{ Einheitsmatrix}$$

ein, so erhält man kürzer

$$\mathbf{G}(p, \mathbf{p}) = \left(G(p) + C(p)\mathbf{p} \quad c(p)E_k \right).$$

Damit nun \mathbf{G} einen Rangverlust hat, muß $c(p) = 0$ sein, denn ansonsten hat \mathbf{G} allein schon auf Grund des hinteren Teils $c(p)E_k$ Vollrang. $c(p) = 0$ impliziert aber, daß alle Zwangsbedingungen eingehalten werden:

$$\forall j = 1, \dots, k (g_{pos,j}(p) = 0).$$

Damit gilt auch $C(p) = 0$ und $\mathbf{G}(p, \mathbf{p}) = G(p)$. Nach Voraussetzung hatte aber $G(p)$ Vollrang, wenn p konsistent ist. Damit hat \mathbf{G} für beliebige Werte von p und \mathbf{p} Vollrang, also globalen Vollrang. Nun muß noch das relaxierte $\tilde{\mathbf{G}}$ betrachtet werden. Bei linearer bzw. Baumgarte-Relaxierung gilt $\tilde{\mathbf{G}} = \mathbf{G}$, damit globaler Vollrang, bei orthogonaler Relaxierung argumentiert man analog zu obiger Beweisführung mit P statt p . \square

Einige Bemerkungen dazu:

- (1) Im Modellproblem (11) waren die Anfangswerte fix. Diese können dann um

$$\begin{aligned}\mathbf{p}_0 &= 0 \\ \mathbf{v}_0 &= 0\end{aligned}$$

ergänzt werden. In der stetigen und konsistenten Lösung gilt dann

$$\begin{aligned}\mathbf{p} &\equiv 0 \\ \mathbf{v} &\equiv 0.\end{aligned}$$

- (2) Die Formulierung (27) hat noch viele Freiheiten. So darf man recht frei über \mathbf{m} verfügen. Am einfachsten ist es, man setzt alle Komponenten konstant, am besten gleich 1. Ebenso hat man die Möglichkeit in (28) die Terme $g_{pos,j}(p)^2$ mit Skalierungen zu versehen.
- (3) Eine weitere Freiheit besteht in der Wahl von \mathfrak{F} . Dieses kann beispielsweise auf 0 gesetzt werden. Dies führt dazu, daß im Falle von konsistenten p immer \mathfrak{a} gleich 0 ist und damit \mathbf{v} konstant bleibt. Geht man zusätzlich davon aus, daß in der Lösung des zugrunde liegenden Optimierungs-Modellproblems auch $\mathbf{p} \equiv 0$, $\mathbf{v} \equiv 0$ ist⁹, dann ist auch eine Formulierung der Art

$$\mathfrak{F}(p, v, q, \mathbf{p}, \mathbf{v}) = -\mathbf{p} - \mathbf{v}$$

sinnvoll, die versucht, diese Variablen wieder gegen die 0 zu treiben.

- (4) Die Dynamik der Variablen \mathbf{p} und \mathbf{v} wird weniger vom Term \mathfrak{F} als vielmehr von \mathfrak{G} bestimmt. Sind die p konsistent, dann ist $\mathfrak{G}(p, \mathbf{p}) = G(p)$ und die Gleichungen für \mathfrak{a} entkoppeln von den übrigen Gleichungen in (27). Dann bestimmt \mathfrak{F} allein die Dynamik von \mathbf{p} , \mathbf{v} . Sind die p aber nicht konsistent, so wird die Dynamik von \mathbf{p} , \mathbf{v} durch Einhaltung der Zwangsbedingung (28) bestimmt. Die Stärke der Inkonsistenz der p bestimmt den Grad der Ankopplung der \mathfrak{a} -Gleichungen an die a -Gleichungen.
- (5) Gewisse Vereinfachungen sind möglich. So ist es denkbar, daß die Zeilen von G in gewisse unabhängige Teile zerfallen, und zwar derart, daß der Rangverlust eines Teils nicht abhängig ist von den Zwangsbedingungen, die den Zeilen anderer Teile entsprechen. In diesem Fall kann man den Term $\sum_{j=1}^k g_{pos,j}(p)^2$ durch eine Summe nur über die Zeilen des eigenen Teils ersetzen. Ist z. B. Rangverlust nur durch Nullwerdung einer Zeile (nicht aber durch lineare Abhängigkeit von anderen Zeilen) möglich, so kann man diesen Term durch $\mathbf{p}_i g_{pos,i}(p)^2$ ersetzen. Nach Zusatzüberlegungen ist es auch möglich dann zu $\mathbf{p}_i g_{pos,i}(p)$ überzugehen. Die entsprechende Zwangsbedingung lautet dann

$$\begin{aligned}\mathfrak{g}_{pos,i}(p, \mathbf{p}) &= g_{pos,i}(p) + \mathbf{p}_i g_{pos,i}(p) \\ &= (1 + \mathbf{p}_i) g_{pos,i}(p),\end{aligned}$$

und deren Ableitung

$$\mathfrak{G}_i(p, \mathbf{p}) = \begin{pmatrix} (1 + \mathbf{p}_i) G_i(p) & g_{pos,i}(p) \end{pmatrix}.$$

Ist die Optimierungsumgebung nun so, daß die Integration nie mit $\mathbf{p}_i = -1$ und immer mit $\mathbf{v}_i = 0$ gestartet wird (z. B. weil die Schranken in (11) für \mathbf{p}_i und \mathbf{v}_i entsprechend gewählt wurden), so hat \mathfrak{G}_i immer Vollrang. Denn war

$$g_{pos,i}(p_0) = 0,$$

⁹Z. B. indem man die fixen Anfangswerte von \mathbf{p} und \mathbf{v} auf 0 setzt.

so entkoppeln die Gleichungen für \mathbf{p}_i und \mathbf{v}_i von den übrigen, \mathbf{p}_i bleibt konstant, $\mathbf{g}_{pos,i}(p, \mathbf{p}) = 0$ impliziert $g_{pos,i}(p) = 0$, p bleibt also konsistent, G_i behält Vollrang. War $g_{pos,i}(p) \neq 0$, so bleibt es wegen $\mathbf{g}_{pos,i}(p, \mathbf{p}) = const \neq 0$ auch ungleich 0, mithin hat \mathbf{G}_i Vollrang.

- (6) Die Anzahl der Variablen \mathbf{p} , die hinzugefügt werden müssen, korrespondiert zur Größe des Rangverlusts, der auftreten kann. Muß man mit einem totalen Rangverlust rechnen, müssen k Variablen hinzugefügt werden. Kann der Rangdefekt höchstens 1 betragen, so reicht die Hinzufügung einer Variablen.
- (7) Interessant ist die Anwendung der Baumgarte-Relaxierung auf die modifizierte Formulierung von Beispiel 5, gestartet im Wert $p = 0$ und $\mathbf{p} = 0$. Die Dynamik der Baumgarte-Relaxierung versucht das System auf die Konsistenzmannigfaltigkeit zurückzutreiben. Hier besteht die Konsistenzmannigfaltigkeit aber aus den Punkten $p = -1$ und $p = 1$ und der Startwert liegt genau symmetrisch dazwischen. Aus Symmetriegründen dürfte keiner der beiden Werte bevorzugt werden und dies geschieht auch: statt p ändert sich der Wert von \mathbf{p} , um die Zwangsbedingung einzuhalten. Für $p = \mathbf{p} = 0$ und dem Baumgarte-Ansatz

$$g_{pos}(p) = \Phi(t)$$

ergibt sich (vereinfachte Gleichung dadurch, daß $p = 0$ bleibt)

$$\begin{pmatrix} a \\ \mathbf{a} \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 2v^2 + 2v^2\mathbf{p} + \ddot{\Phi}(t) \\ -\mathbf{f} + \mathbf{m}(2v^2 + 2v^2\mathbf{p} + \ddot{\Phi}(t)) \end{pmatrix}.$$

Dabei wird die Dynamik von \mathbf{p} vollständig von dieser Zwangsbedingung diktiert. Die Wahl von \mathbf{f} hat in dieser Situation nur Einfluss auf λ . Wählt man aber $\mathbf{f} = -\mathbf{p} - \mathbf{v}$ und weicht p von 0 nur ein wenig ab, so wird \mathbf{p} schließlich wieder Richtung 0 und p Richtung Mannigfaltigkeitspunkt geschoben.

- (8) Die modifizierte Modellformulierung (27) führt weiterhin zu entsprechend oft differenzierbaren Lösungen. I. a. stehen nun aber innerhalb des F inkonsistente p und v . Daher muß man F auf solche Werte fortsetzen. Dabei muß man Entartungen berücksichtigen, wenn man die Differenzierbarkeit erhalten möchte. So könnten z. B. zwei Variablen x_1, x_2 eine Einheitsachse $\frac{x_1 - x_2}{\|x_1 - x_2\|}$ festgelegt haben, um die z. B. ein Drehmoment wirkt. „Wandern“ nun x_1 und x_2 aufeinander zu und durcheinander durch, so springt die Achse in die entgegengesetzte Richtung und dies führt zum Verschwinden der Differenzierbarkeit. Dies kann aber behoben werden, indem man das wirkende Drehmoment mit einem Wert multipliziert, der 1 ist, wenn x_1, x_2 konsistent sind, und der 0 wird, wenn sie sich durchdringen. Lautet die ursprüngliche Zwangsbedingung

$$g_{pos}(x_1, x_2) = (x_1 - x_2)^2 - 1,$$

so kann als besagter Term

$$(x_1 - x_2)^2$$

gewählt werden. Der Sprung der Achse wird also mit der Entartung geglättet.

- (9) Wählt man \mathbf{f} wie oben geschildert $\equiv 0$ oder $= -\mathbf{p} - \mathbf{v}$, dann bleiben \mathbf{p}, \mathbf{v} linear bzw. konstant, falls ihre Gleichungen von den übrigen entkoppeln, also falls die p konsistent sind. Sind die p nicht konsistent, ist mit starken Veränderungen von \mathbf{p}, \mathbf{v} zu rechnen. In diesem Sinne darf man die neuen Variablen \mathbf{p} und \mathbf{v} „Seismographen der Inkonsistenz“ nennen.

- (10) Entfernt sich das System zu weit von der Konsistenzmannigfaltigkeit, so werden die Trajekturen von \mathbf{p} und \mathbf{v} immer chaotischer. I. a. werden sie durch heftige Schwingungen gekennzeichnet, die dazu führen, daß der Optimierer Probleme bei der Bestimmung einer Abstiegsrichtung bekommt.
- (11) Auch wenn starke Abweichungen von der Mannigfaltigkeit zu einem schnellen Ende der Optimierung führen können, gibt es Bereiche, in denen sie angewandt werden kann, so bei der Bestimmung der initialen Optimierungswerte im Mehrzielverfahren. Es gibt die Situation, in der die Anfangstrajektorie der Optimierung unbekannt ist und durch eine initiale Integration aus Anfangswert und Initialsteuerung gewonnen wird. Verlässt die Trajektorie dabei den durch (11) definierten Bereich, so setzt der Optimierer sie im nachfolgenden Knoten auf den nächsten nach (11) zulässigen Punkt. Dabei kann es durchaus passieren, daß mehrere Massen (die z. B. allesamt den Boden durchschlagen haben) auf den gleichen Punkt gesetzt werden, mithin also in einer sehr großen Entartung gestartet wird. Bricht hier die Integration zusammen, so kann auch der Optimierer ohne die nötigen Anfangswerte nicht starten. Trotz z. T. heftiger Schwingungen liefert aber die Formulierung (28) Initialwerte. Überwiegen diese nicht, so wird im Laufe weniger Optimierungsschritte die Inkonsistenz immer schwächer und sinkt dann auf das normale Maß. Nachdem diese Phase erreicht wird, ist es überlegenswert, wieder mit diesen Zwischenresultaten zum Modell (27) zurückzukehren. Denn i. allg. sind keine weiteren starken Abweichungen von der Konsistenzmannigfaltigkeit zu erwarten und das Modell mit den zusätzlichen Variablen ist natürlich in der Integration langsamer als das ohne.

§3 Initiale Projektion

Ein Nachteil der redundanten Koordinaten besteht in der großen Zahl an Variablen. Besonders schwerwiegend macht sich dies bei Berechnung der Wronski-Matrix gemäß Definition 4 bemerkbar. Es müssen offensichtlich so viele Richtungsableitungen bestimmt werden wie es Variablen gibt und hinter jeder Richtungsableitung steht i. allg. eine vollständige Integration einer im Anfangswert variierten DAE. Besonders gravierend wirkt sich das bei den natürlichen Koordinaten aus, bei denen jeder Starrkörper zwölf Orts- und zwölf Geschwindigkeitsvariablen zugeordnet bekommt, im Gegensatz zu drei Orts- und drei Geschwindigkeitsvariablen in der Formulierung mit Minimalkoordinaten. Der Unterschied von 24 Integrationen zu 6 Integrationen spräche hier eindeutig gegen die Anwendung von natürlichen Koordinaten, gäbe es nicht eine Möglichkeit, die Anzahl der notwendigen Richtungsableitungen auf die Anzahl zu reduzieren, die auch für Minimalkoordinaten notwendig sind.

Diese Möglichkeit gibt es und sie soll hier vorgestellt werden. Sie reduziert nicht nur den Integrationsaufwand auf das notwendige Minimum, sondern löst auch die in den beiden vorherigen Abschnitten besprochenen Probleme der Singularitäten durch inkonsistente Startwerte. Ihr Nachteil ist, daß die vorgestellte Methode nur lokal funktioniert, d. h. nur durchgeführt werden kann, wenn die inkonsistenten Startwerte nicht allzu stark von der Konsistenzmannigfaltigkeit entfernt sind (dies ist in der Praxis i. allg. der Fall) und wenn wie schon öfters der Vollrang von G vorausgesetzt wird.

§3.1 Motivation. Die Methode beruht auf Ideen von C. KRAUS, welcher wiederum Arbeiten von J. SCHLÖDER¹⁰ auf die Mechanik angewandt hat.

¹⁰siehe SCHLÖDER[60]

Zur Motivation sei noch einmal an den Algorithmus der orthogonalen Relaxation in Abschnitt §1.2 erinnert. Im letzten Schritt des Algorithmus findet die Rücksubstitution statt:

$$\begin{aligned} p(t_{end}) &:= p(t_{end}) + v_i(t_{end} - t_0) \\ v(t_{end}) &:= v(t_{end}) + v_i. \end{aligned}$$

Welche Konsequenzen hat es, wenn die Rücksubstitution unterlassen wird?

- (1) Die „Lösung“ zum Zeitpunkt t_0 erfüllt nicht mehr die Anfangsbedingungen

$$p(t_0) = p_0, v(t_0) = v_0,$$

sondern nur noch

$$p(t_0) = p_0, v(t_0) = v_0 - v_i.$$

- (2) Die Wronski-Matrix ist rangdefizitär. Dies sieht man am deutlichsten, wenn man die Wronski-Matrix zum Zeitpunkt t_0 ermittelt. Die Wronski-Matrix ist nicht mehr die Einheitsmatrix, sondern nur noch der zur Ortsvariablen gehörige Teil ist eine Einheitsmatrix. Der zu Geschwindigkeitsvariablen gehörige Teil ist i. allg. nicht mehr die Einheitsmatrix. Der Einfachheit halber darf man sich im folgenden die Massenmatrix M als Einheitsmatrix E vorstellen. Die Matrix $G(p_0)$ bestimmt im Punkt p_0 einen Tangentenraum (Richtungen r mit $G(p_0)r = 0$) und einen Normalenraum (Linearkombinationen der Zeilen von $G(p_0)$). In Richtung der Normalen ist die Sensitivität 0. Dies sieht man folgendermaßen: sei v_i, ξ Lösung des Gleichungssystems (1)

$$\begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ G(p_0)v_0 \end{pmatrix}.$$

Nun werde v_i variiert in eine Richtung λ , die sich als Linearkombination von Zeilen von $G(p_0)$ schreiben lässt:

$$v'_0 := v_0 + \lambda, \lambda = M^{-1}G^T\mu.$$

Dann folgt aus

$$\begin{aligned} \begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i + \lambda \\ \xi - \mu \end{pmatrix} &= \begin{pmatrix} M & G^T(p_0) \\ G(p_0) & 0 \end{pmatrix} \begin{pmatrix} v_i + M^{-1}G^T(p_0)\mu \\ \xi - \mu \end{pmatrix} \\ &= \begin{pmatrix} Mv_i + G^T(p_0)\mu + G^T(p_0)\xi - G^T(p_0)\mu \\ G(p_0)v_i + G(p_0)M^{-1}G^T(p_0)\mu \end{pmatrix} \\ &= \begin{pmatrix} Mv_i + G^T(p_0)\xi \\ G(p_0)(v_i + \lambda) \end{pmatrix} \\ &= \begin{pmatrix} Mv_i + G^T(p_0)\xi \\ G(p_0)(v_0 + \lambda) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ G(p_0)(v_0 + \lambda) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ G(p_0)v'_0 \end{pmatrix}. \end{aligned}$$

Damit wird im Rahmen der Relaxierung von v'_0 der Term $v_i + \lambda$ subtrahiert, also ergibt sich

$$V(0) := v'_0 - v_i - \lambda = v_0 + \lambda - v_i - \lambda = v_0 - v_i.$$

Dies ist aber der gleiche Startwert für V wie bei v_0 ! Die Variation $v'_0 := v_0 + \lambda$ wurde durch das Subtrahieren des inkonsistenten Anteils völlig kompensiert (was insofern klar ist, als daß um eine Größe variiert wurde, die vollständig senkrecht zur Konsistenzmannigfaltigkeit ist). Damit ergibt sich, daß die Sensitivität in Richtungen des Normalenraums gleich 0 ist. Folglich ist die Wronski-Matrix rangdefizitär, sie hat „in diese Richtungen“ Nullspalten.

- (3) Die Berechnung der Wronski-Matrix hat sich reduziert. Führt man eine Basis-Transformation durch auf eine neue Basis, die aus dem Normalenraum und dem Tangentenraum von $G(p_0)$ besteht, so ist das Ergebnis des Normalenraums a priori bekannt: es besteht aus Nullspalten. Nur noch die Sensitivität im Tangentenraum muß berechnet werden. Da der Normalenraum von den Zeilen von $G(p_0)$ aufgespannt wird, bedeutet dies, daß man sich die Berechnung von so vielen variierten Trajektorien erspart wie es Zwangsbedingungen gibt. Das heißt aber, daß nur noch so viele Richtungsableitungen gebraucht werden wie es tatsächlich Freiheitsgrade gibt. Damit ist man bei der minimal notwendigen Größe, die auch bei Minimalkoordinaten benötigt wird (dies allerdings hier immer nur bezogen auf den Raum der Geschwindigkeitskoordinaten, noch nicht der Ortskoordinaten).
- (4) Der Rangverlust der Wronski-Matrix ist im multiple-shooting-Kontext ungefährlich. Im multiple-shooting werden die Stetigkeitsanschlußbedingungen (10)

$$Y(t_{i+1}, t_i, y_i) - y_{i+1} = 0$$

gefordert, deren Linearisierung lautet

$$\frac{\partial}{\partial y_i} Y(t_{i+1}, t_i, y_i) \Delta y_i - \Delta y_{i+1} = 0$$

bzw. in einer Matrix-Block Schreibweise bei geeigneter Sortierung der Variablen

$$\left(\begin{array}{cc} \frac{\partial}{\partial y_i} Y(t_{i+1}, t_i, y_i) & -E \end{array} \right) \begin{pmatrix} \Delta y_i \\ \Delta y_{i+1} \end{pmatrix} = 0, \quad (29)$$

wobei E die Einheitsmatrix ist. Der Rangverlust des Matrixblocks $\frac{\partial}{\partial y_i} Y(t_{i+1}, t_i, y_i)$ ist ungefährlich, da der Matrixblock E für den Vollrang des Systems (29) sorgt.

Insgesamt bietet sich hier also die Möglichkeit, den Aufwand zur Berechnung der Wronski-Matrix auf das notwendige Minimum zu reduzieren und redundante Koordinaten damit wieder konkurrenzfähig zu machen. Allerdings muß die zu lösende DAE umformuliert werden, da das Verfahren wegen Punkt 1 (Seite 41) die Anfangswertbedingung nicht erfüllt. Auch sollen die Vorteile eines solchen Vorgehens ebenfalls auf die Ortskoordinaten ausgedehnt werden. Die Vorstellung, daß eine Variation in eine inkonsistente Richtung einfach wegprojiziert wird, ist dabei die zentrale Idee.

§3.2 Initiale Projektion. Die differential-algebraische Gleichung (7) wird umformuliert zu

$$\begin{aligned} \dot{p} &= v \\ \dot{v} &= a \\ \begin{pmatrix} M & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} a \\ \lambda \end{pmatrix} &= \begin{pmatrix} F(t, p, v) \\ \gamma(p, v) \end{pmatrix} \\ \begin{pmatrix} p(t_0) \\ v(t_0) \end{pmatrix} &= \Pi \left(\begin{pmatrix} p_0 \\ v_0 \end{pmatrix} \right) \end{aligned} \quad (30)$$

unter Berücksichtigung der Invarianten

$$\begin{aligned} g_{pos}(p) &= 0 \\ g_{vel}(p, v) &= 0. \end{aligned}$$

Dabei ist $\Pi : \mathbb{K}^{2n} \rightarrow \mathbb{K}^{2n}$ ein Projektor, der die folgenden Eigenschaften hat:

- (1) Π projiziert die Werte p_0 und v_0 auf die Konsistenzmannigfaltigkeit, d. h. es gilt

$$\begin{aligned} g_{pos} \left(\left[\Pi \left(\begin{pmatrix} p_0 \\ v_0 \end{pmatrix} \right) \right]_1 \right) &= 0 \\ g_{vel} \left(\left[\Pi \left(\begin{pmatrix} p_0 \\ v_0 \end{pmatrix} \right) \right]_1, \left[\Pi \left(\begin{pmatrix} p_0 \\ v_0 \end{pmatrix} \right) \right]_2 \right) &= 0, \end{aligned}$$

wobei $[\cdot]_i$ die i -te Komponente des Vektors bedeutet.

- (2) Π ist i. allg. nur lokal definiert, d. h. nur für Werte p_0, v_0 , die hinreichend nahe an der Konsistenzmannigfaltigkeit liegen. Dies macht das gesamte Verfahren zu einem lokalen Verfahren.
- (3) Π ist die identische Abbildung auf der Konsistenzmannigfaltigkeit selbst, d. h. gilt $g_{pos}(p_0) = 0$ und $g_{vel}(p_0, v_0) = 0$, so ist

$$\Pi \left(\begin{pmatrix} p_0 \\ v_0 \end{pmatrix} \right) = \begin{pmatrix} p_0 \\ v_0 \end{pmatrix}.$$

Zusammenfassend läßt sich also sagen: die Einführung eines solchen Projektors Π macht das Ganze im Gegensatz zu den fast global anwendbaren Relaxierungen zu einem nur noch lokal anwendbaren Verfahren. Dieses erspart aber die Relaxierungen selbst, da nach der Projektion die Anfangswerte konsistent sind, und nimmt auf der Konsistenzmannigfaltigkeit selbst keinerlei Veränderungen vor, da dort die Projektion die Identität ist.

Nun geht es darum, einen geschickten Projektor Π zu wählen, damit das Verfahren die im Abschnitt §3.1 versprochenen Eigenschaften besitzt.

Definition 8 (Initiale Projektion). *Das Verfahren der initialen Projektion besteht darin, das differential-algebraische System (7) durch die Formulierung (30) zu ersetzen, dabei sei*

$$\Pi \left(\begin{pmatrix} p_0 \\ v_0 \end{pmatrix} \right) := \begin{pmatrix} p \\ v \end{pmatrix}$$

und p sei Lösung des folgenden nichtlinearen Gleichungssystems

$$\begin{aligned} A(p - p_0) + G^T(p)\xi &= 0 \\ g_{pos}(p) &= 0 \end{aligned} \tag{31}$$

und v sei Lösung des nachgeordneten linearen Gleichungssystems

$$\begin{pmatrix} A & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} v_i \\ \xi' \end{pmatrix} = \begin{pmatrix} 0 \\ G(p)v_0 \end{pmatrix} \tag{32}$$

$$v = v_0 - v_i.$$

Die Matrix A ist dabei wählbar, i. allg. wird $A = M$ (Massenmatrix) oder $A = E$ (Einheitsmatrix) sein. Voraussetzung ist, daß $G(p)$ Vollrang hat.

Feststellung 9 (Irrelevante Richtungen). *Die Spalten von $A^{-1}G^T(p)$ stellen irrelevante Richtungen dar, d. h. werden die Anfangswerte p_0 und v_0 in Richtung einer Linearkombination aus Spalten von $A^{-1}G^T(p)$ (=Normalraum) variiert, so wird die Variation wegprojiziert und es ergibt sich das gleiche p und v wie im unvariierten Fall.*

BEWEIS. Seien p , v und ξ die Lösungen des obigen Systems bei Anfangswerten p_0 und v_0 und sei $p'_0 := p_0 + A^{-1}G^T(p)\mu$. Dann gilt:

$$\begin{aligned} A(p - p_0 - A^{-1}G^T(p)\mu) + G^T(p)(\xi + \mu) &= A(p - p_0) - G^T(p)\mu + G^T(p)\xi + G^T(p)\mu \\ &= A(p - p_0) + G^T(p)\xi \\ &= 0 \end{aligned}$$

sowie

$$g_{pos}(p) = 0.$$

Also ist p , $\xi + \mu$ Lösung des Systems (31) für das variierte p_0 , mithin ergibt sich das gleiche p wie im unvariierten Fall.

Damit wird in das Gleichungssystem (32) ein unverändertes p eingesetzt. Der Beweis, daß dann auch die Variation des v_0 wegprojiziert wird, wurde schon in Punkt 2 von §3.1 erbracht. \square

Zur Berechnung der Wronski-Matrix mit Hilfe variiertes Trajektorien ergibt sich folgender

ALGORITHMUS DER INITIALEN PROJEKTION

- (1) Projiziere die gegebenen Anfangswerte p_0 und v_0 nach p und v durch Lösen der Gleichungssysteme (31) und (32).
- (2) Erstelle eine neue Basis

$$B = \begin{pmatrix} B_N & B_T \end{pmatrix},$$

deren erster Teil B_N aus den Spalten von $A^{-1}G^T(p)$ gebildet ist und deren zweiter Teil B_T den Orthogonalraum zu B_N aufspannt. (Hier kann wahlweise eine Orthogonalisierung oder eine Orthonormalisierung vorgenommen werden.)

- (3) Setze in der zu berechnenden Wronski-Matrix die ersten Spalten auf 0, die zu B_N korrespondieren. Dies sind jeweils so viele Spalten wie es Zwangsbedingungen gibt (Anzahl Zeilen von g_{pos} plus Anzahl Zeilen von g_{vel}).
- (4) Setze Variationen in Richtungen der Spalten von B_T . Die Anzahl der Variationen entspricht der Anzahl der tatsächlichen Freiheitsgrade.
- (5) Projiziere die variierten Anfangswerte wieder auf die Konsistenzmannigfaltigkeit durch Lösen der Gleichungssysteme (31) und (32).
- (6) Integriere alle Trajektorien, die unveränderte Haupttrajektorie ebenso wie die variierten.
- (7) Bestimme aus dem Ergebnis den Teil der Wronski-Matrix, der zu B_T korrespondiert. Nun liegt die Wronski-Matrix bzgl. der Basis B vor.
- (8) Führe eine Rücktransformation von B auf das ursprüngliche Koordinatensystem durch. Dabei wird die Wronski-Matrix bzgl. B in die zu berechnende Wronski-Matrix überführt.

Das Gleichungssystem (31) löst man am besten mit dem Newton-Verfahren. Der Iterationsschritt lautet dann

$$\begin{pmatrix} A + \partial_p G^T(p)\xi & G^T(p) \\ G(p) & 0 \end{pmatrix} \begin{pmatrix} \Delta p \\ \Delta \xi \end{pmatrix} = \begin{pmatrix} -A(p - p_0) - G^T(p)\xi \\ -g_{pos}(p) \end{pmatrix}.$$

Dabei tritt der Ableitungstensor $\partial_p G$ auf, der als Bestandteil des Modells analytisch vorliegen sollte.

Der Aufwand zur Durchführung dieses Iterationsverfahrens ist i. allg. vernachlässigbar klein im Vergleich damit, daß eine größere Anzahl an Integrationen damit gespart wird. Ansonsten besteht noch die Möglichkeit, hier Maßnahmen zur Beschleunigung des Verfahrens zu ergreifen, als da wären

- feste Anzahl an Iterationsschritten (z. B. nur ein Schritt) oder
- Vereinfachung der Matrix durch Weglassen des Tensorterms $\partial_p G$.

Hier wäre aber eine genaue Fehleranalyse nötig, denn eine ungenaue Bestimmung von p führt zu einer ungenauen Bestimmung der Richtungen in B und damit zu Fehlern in der zu berechnenden Wronski-Matrix.

Ein Verfahren wie die initiale Projektion findet sich bereits bei GERDTS[26]. Dort werden zu i. allg. inkonsistenten Startwerten x_0 konsistente Werte x'_0 gesucht mit

$$\begin{aligned} \min_{x'_0} \quad & \|x'_0 - x_0\|^2 \\ & g(x'_0) = 0. \end{aligned}$$

Dies entspricht in obiger Darstellung einer Projektion mit A als Einheitsmatrix. Allerdings wird diese Projektion bei GERDTS nicht dazu benutzt, die Anzahl der für die Richtungsableitungen notwendigen Integration zu minimieren.

§3.3 Variationsdifferentialgleichung. Für viele theoretische Untersuchungen ist es vorteilhafter, statt der Differentialgleichung

$$\dot{y}(t) = f(t, y(t)), \quad y(t_0) = y_0$$

die zugehörige Integralgleichung zu betrachten, denn sie beinhaltet in einer Gleichung sowohl den dynamischen Verlauf als auch die Anfangswerte:

$$y(t, y_0) = y_0 + \int_{t_0}^t f(\tau, y(\tau, y_0)) d\tau.$$

Daher soll an dieser Stelle die Veränderung der Differentialgleichung auch in der Integraldarstellung genannt werden. Ein Übergang der Art, wie er von (7) nach (30) vollzogen wird, würde in diesem Zusammenhang als

$$y(t, y_0) = \Pi(y_0) + \int_{t_0}^t f(\tau, y(\tau, y_0)) d\tau$$

formuliert werden. Die Veränderung besteht also darin, daß statt des Initialwertes y_0 der Wert $\Pi(y_0)$ verwendet wird.

Von Interesse z. B. für die Interne Numerische Differentiation ist die Variationsdifferentialgleichung, d. h. die Differentialgleichung für die Wronski-Matrix

$$W(t, t_0, y_0) := \frac{\partial}{\partial y_0} y(t, t_0, y_0).$$

Die zugehörige Integralgleichung lautet in einer halbformalen Schreibweise

$$\frac{\partial y(t, y_0)}{\partial y_0} = \frac{\partial \Pi(y_0)}{\partial y_0} + \int_{t_0}^t \frac{\partial}{\partial y} f(\tau, y(\tau, y_0)) \frac{\partial}{\partial y_0} y(\tau, y_0) d\tau.$$

Für die Differentialgleichung der Wronski-Matrix $W(t, t_0, y_0)$ ergibt sich damit

$$\frac{d}{dt} W(t, t_0, y_0) = \frac{\partial}{\partial y} f(t, y(t, y_0)) W(t, t_0, y_0), \quad W(t_0, t_0, y_0) = \frac{\partial \Pi(y_0)}{\partial y_0}.$$

(Hier sieht man sehr schön, wie der Anfangswert der Wronski-Matrix die Sensitivität des Projektors ist, und dieser hat „entlang den Projektionsachsen“ die Sensitivität 0).

Dies ist die Form für ein System der Art (30), während für ein System der Art (7) die Variationsdifferentialgleichung

$$\frac{d}{dt} W(t, t_0, y_0) = \frac{\partial}{\partial y} f(t, y(t, y_0)) W(t, t_0, y_0), \quad W(t_0, t_0, y_0) = E$$

lautet.

KAPITEL III

Aufblasung

Das vorherige Kapitel hat Methoden für den Fall entwickelt, daß Singularitäten durch Abweichung von der Orts- oder Geschwindigkeitskonsistenzmannigfaltigkeit auftreten. Es ist naheliegend zu fragen, wie der Fall behandelt werden soll, bei dem Singularitäten auf der Mannigfaltigkeit selbst auftreten¹. Die Behandlung dieser Frage ist ein völlig anderes Problem, das in das Gebiet der algebraischen Geometrie führt. Dort wurden tiefgreifende Resultate zur Lösung dieses Problems erarbeitet, die jedoch in ihrem Beweis zum kompliziertesten zählen, was die Mathematik zu bieten hat (Satz von HIRONAKA). In der algebraischen Geometrie lautet das Problem, zu einer gegebenen Mannigfaltigkeit eine äquivalente ohne Singularität zu finden. Schwächt man diese Formulierung dahingehend ab, eine Singularität auf einer Mannigfaltigkeit numerisch durchqueren zu können, ohne daß die Integration zusammenbricht, so läßt sich ein einfaches Verfahren angeben. Dessen Beweis ist zudem noch sehr elementar. Dieses Verfahren (die Aufblasung) wird nun vorgestellt und der sehr einfache Beweis der Endlichkeit geführt.²

§1 Problemstellung

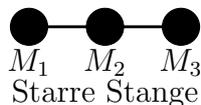
Es wird nun der Fall untersucht, daß $G(p)$ einen Rangverlust besitzt, obwohl $g_{pos}(p) = 0$ ist, also p konsistent ist.

Zunächst seien wieder Beispiele eingeführt, die die Problematik veranschaulichen sollen.

§1.1 Rangverlust ohne Freiheitsgrad.

Beispiel 6.

Der zugrundeliegende Raum sei der \mathbb{R}^2 . Drei Massen M_1, M_2, M_3 , jeweils beschrieben durch zwei Koordinaten x_i, y_i , sollen in einer Linie in festen Abständen liegen. M_1 soll zu M_3 den euklidischen Abstand 2 haben, M_2 soll jeweils die Entfernung 1 von M_1 und von M_3 haben.



Dies werde als Zwangsbedingung wie folgt formuliert: der Abstand von M_1 zu M_3 sei 2.

$$g_{Bsp6,1}(x_1, y_1, x_2, y_2, x_3, y_3) = (x_1 - x_3)^2 + (y_1 - y_3)^2 - 4.$$

Nun ist noch keine Bedingung an M_2 gestellt worden, also: der Abstand von M_2 zu M_1 sei 1.

$$g_{Bsp6,2}(x_1, y_1, x_2, y_2, x_3, y_3) = (x_1 - x_2)^2 + (y_1 - y_2)^2 - 1.$$

¹Eine verwandte Problematik, nämlich das Vorliegen eines Rangdefekts der Größe 1 bei linear-impliziten DAEs vom Index 1, sogenannte Impasse-Punkte, wird z. B. bei RABIER UND RHEINBOLDT[55] und [56], BAUER[1] und [2] behandelt.

²Dieses Verfahren hat sich als Nebenprodukt der Untersuchung von Singularitäten ergeben. Es wird später bei der Bestimmung optimaler Laufbewegungen selbst nicht eingesetzt.

Nun könnte aber M_2 noch in einem Kreis von Radius 1 um M_1 herum rotieren, also muß noch die Bedingung hinzugefügt werden, daß M_2 von M_3 den Abstand 1 hat.

$$g_{Bsp6,3}(x_1, y_1, x_2, y_2, x_3, y_3) = (x_2 - x_3)^2 + (y_2 - y_3)^2 - 1.$$

Nun erst ist die gewünschte Konfiguration erreicht, aber erstaunlicherweise liegt nun auch Rangverlust von G_{Bsp6} vor. Denn es ist

$$G_{Bsp6}(x_1, y_1, x_2, y_2, x_3, y_3) = 2 \begin{pmatrix} x_1 - x_3 & y_1 - y_3 & 0 & 0 & -(x_1 - x_3) & -(y_1 - y_3) \\ x_1 - x_2 & y_1 - y_2 & -(x_1 - x_2) & -(y_1 - y_2) & 0 & 0 \\ 0 & 0 & x_2 - x_3 & y_2 - y_3 & -(x_2 - x_3) & -(y_2 - y_3) \end{pmatrix}.$$

Setzt man hierin die konsistenten Werte

$$x_1 = 0, y_1 = 0, x_2 = 1, y_2 = 0, x_3 = 2, y_3 = 0$$

ein, so erhält man

$$G_{Bsp6}(0, 0, 1, 0, 2, 0) = \begin{pmatrix} -4 & 0 & 0 & 0 & 4 & 0 \\ -2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 2 & 0 \end{pmatrix},$$

und dieses hat nicht Vollrang. Dies ist überraschend, denn der Rangverlust tritt offensichtlich erst nach Hinzufügung der dritten Zeile auf, und diese hat doch eine vorhandene Freiheit (Rotation von M_2 um M_1 eliminiert). Auch ist kein neuer Freiheitsgrad aufgetreten, denn das System hat drei Freiheitsgrade (Translation von M_1 liefert zwei Freiheitsgrade, dann Rotation von M_3 um M_1 liefert einen weiteren Freiheitsgrad. Damit ist das System vollständig beschrieben), dies entspricht sechs Variablen weniger drei Zwangsbedingungen. Dies sei festgehalten in der folgenden

Feststellung 10. *Es gibt Rangverluste von G , ohne daß ein (ihn verursachender) Freiheitsgrad neu aufgetreten wäre.*

Allerdings gibt es von der durch Beispiel 6 beschriebenen Konsistenzmannigfaltigkeit eine Formulierung, die keinen Rangverlust besitzt, nämlich „ M_2 sitzt im Mittelpunkt von M_1 und M_3 “

$$\begin{aligned} g_{pos,1}(x_1, y_1, x_2, y_2, x_3, y_3) &= (x_1 - x_3)^2 + (y_1 - y_3)^2 - 4 \\ g_{pos,2}(x_1, y_1, x_2, y_2, x_3, y_3) &= x_2 - \frac{1}{2}(x_1 + x_3) \\ g_{pos,3}(x_1, y_1, x_2, y_2, x_3, y_3) &= y_2 - \frac{1}{2}(y_1 + y_3). \end{aligned}$$

Hier lautet die Ableitungsmatrix

$$G(x_1, y_1, x_2, y_2, x_3, y_3) = \begin{pmatrix} 2(x_1 - x_3) & 2(y_1 - y_3) & 0 & 0 & -2(x_1 - x_3) & -2(y_1 - y_3) \\ -\frac{1}{2} & 0 & 1 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 1 & 0 & -\frac{1}{2} \end{pmatrix}$$

und setzt man obige Werte hier ein, ergibt sich

$$G(0, 0, 1, 0, 2, 0) = \begin{pmatrix} -2 & 0 & 0 & 0 & 2 & 0 \\ -\frac{1}{2} & 0 & 1 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 1 & 0 & -\frac{1}{2} \end{pmatrix},$$

was keinen Rangverlust besitzt. Damit stellt sich die Frage

„Gibt es einen Algorithmus, der ausgehend von der Formulierung mit Rangverlust die Formulierung ohne Rangverlust findet?“

§1.2 Rangverlust mit Freiheitsgrad.

Beispiel 7.

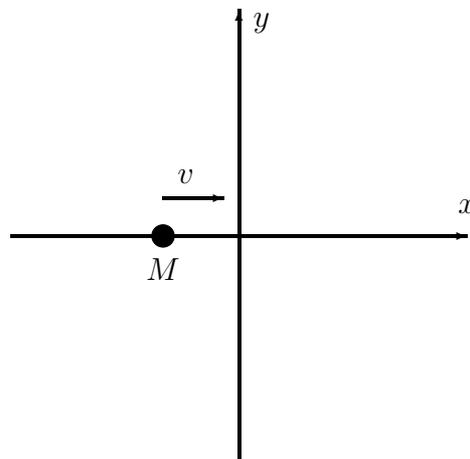
Betrachtet wird die Mannigfaltigkeit des Koordinatenkreuzes im \mathbb{R}^2 , beschrieben durch

$$g_{Bsp7}(x, y) = xy.$$

Die Ableitungsmatrix lautet hier

$$G_{Bsp7} = \begin{pmatrix} y & x \end{pmatrix}.$$

Offensichtlich hat G_{Bsp7} einen Rangverlust in $x = y = 0$, also im Ursprung des Koordinatenkreuzes. Dies ist auch leicht verständlich, treffen sich doch im Ursprung zwei eindimensionale Mannigfaltigkeiten. Betrachtet wird nun die Bewegung eines einzelnen Massepunktes M , der in $x = -1, y = 0$ mit $v_x = 1, v_y = 0$ gestartet wird bei externen Kräften gleich 0.



Die Mannigfaltigkeit des
Koordinatenkreuzes

Die Lösung der Bewegungsgleichung ist offensichtlich

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} -1 + t \\ 0 \end{pmatrix}.$$

Der Massepunkt fliegt die x -Achse entlang mit konstantem v . Ein Integrator, der auf dem Index-1-System (7) arbeitet, bricht allerdings im Ursprung bzw. kurz davor zusammen. Der Rangverlust ist ein unüberwindbares Hindernis, obwohl die Bewegung (physikalisch) eindeutig vorgegeben ist, da der im Ursprung spontan auftretende zweite Freiheitsgrad bei $v > 0$ nicht erreichbar ist (v müsste einen Sprung machen).³ Es ist so, daß allein die Geschwindigkeit schon entscheidet, daß der Freiheitsgrad $\{x = 0\}$ nicht erreicht werden

³Unter Umständen kann ein Integrator den Rangverlust allein aufgrund numerischer Ungenauigkeiten überwinden, z. B. durch eine sehr grobe Schrittweite, oder auch, weil die Zwangsbedingung nur als

$$g_{pos}(x, y) = xy - \varepsilon, \varepsilon \neq 0$$

mit kleinem $|\varepsilon|$ eingehalten wird. Zu welchem erstaunlichen Verhalten letzteres führt, kann man sich vorstellen, wenn man sich die Mannigfaltigkeiten $xy = 0$ und $xy = \varepsilon$ vor Augen führt.

kann. Dennoch destabilisiert dieser Freiheitsgrad die Integration und führt zum Abbruch. Somit stellt sich hier die Frage:

„Gibt es einen Algorithmus, der die Mannigfaltigkeit so umformuliert, so daß die physikalisch sinnvolle Lösung gefunden werden kann?“

§2 Lösungsansätze

Die in den beiden Abschnitten §1.1 und §1.2 gestellten Fragen können hier nicht vollständig beantwortet werden, da es sich hierbei um Problemstellungen handelt, die tief in das Gebiet der algebraischen Geometrie hineinreichen.

Dabei muß allerdings angemerkt werden, daß bisher die Funktion g_{pos} eine beliebige stetig differenzierbare Funktion sein durfte. Der zentrale Untersuchungsgegenstand der algebraischen Geometrie sind allerdings *Polynomgleichungen*. Die Beschränkung auf polynomielle Zwangsbedingungen stellt aber keine große Einschränkung dar:

- Findet die Modellierung in natürlichen Koordinaten statt, so sind alle praktisch vorkommenden Bedingungen Polynomgleichungen zweiten oder sogar ersten Grades.
- Nach dem Satz von Stone-Weierstraß liegen bei kompaktem Definitionsbereich die Polynome dicht in der Menge der stetigen Funktionen, letztere lassen sich also beliebig gut durch Polynome approximieren.
- Häufig werden gar keine anderen Funktionen außer Polynomen zur Berechnung zugelassen. Selbst wenn in einer Gleichung z. B. eine trigonometrische Funktion vorkommt, wird diese in der Auswertung durch den Computer häufig durch einen hinreichend langen Anfangsteil einer polynomiellen Entwicklung (Taylor- oder besser Tschebyscheff-Polynome) ersetzt. Numerisch gesehen entsteht bei der Ersetzung dieser Funktion durch ein geeignetes Polynom kein relevanter Fehler.

Daher seien in diesem Kapitel nur noch polynomielle Gleichungen betrachtet.

Im Lichte der algebraischen Geometrie lassen sich die beiden obigen Fragen so zusammenfassen:⁴

„Gibt es zu jeder Varietät eine birational äquivalente Varietät ohne Singularitäten?“

Die Antwort lautet „Ja“ und wird mit dem berühmten Satz von HEISUKE HIRONAKA gelöst⁵, einem der schönsten Sätze der Mathematik mit einem der kompliziertesten Beweise. Der Satz von Hironaka zeigt, daß es eine Sequenz von „blow-ups“⁶ gibt, die in einer singularitätsfreien Varietät endet. Inzwischen sind u. a. von O. VILLAMAYOR Fortschritte dahingehend erzielt worden, den nicht-konstruktiven Beweis von Hironaka in einen effektiven Algorithmus zu verwandeln⁷. Beispielsweise ist ein solcher Algorithmus als Paket DESING für das Computer-Algebra-Programm XMAPLEV erhältlich. Diese Algorithmen, die global Singularitäten eliminieren, sind allerdings sehr aufwendig.

Zu den beiden obigen Teilfragen kann man allerdings mehr sagen als einen bloßen Beweis auf den globalen Satz von Hironaka zu geben. So sind die Probleme des Abschnitts §1.1 damit verknüpft, daß das eine Varietät (oder Mannigfaltigkeit) definierende Gleichungssystem nicht das einfachst mögliche ist. Im Falle eines algebraisch abgeschlossenen

⁴Varietät ist in der algebraischen Geometrie das Pendant zur Mannigfaltigkeit, nämlich die Nullstellenmenge eines Systems von Polynomgleichungen, ohne daß eine weitere Differenzierbarkeitsstruktur wie Karten oder Atlanten eingeführt werden.

⁵siehe HARTSHORNE[34]

⁶Der Begriff des blow-ups oder der Aufblasung wird im nächsten Abschnitt erklärt.

⁷siehe z. B. VILLAMAYOR[65] oder ENCINAS[18]

Grundkörpers wie z. B. \mathbb{C} liefert der Hilbertsche Nullstellensatz die Aussage, daß das eigentliche beschreibende Gleichungssystem durch das Radikalideal definiert wird⁸. Algorithmen zur Berechnung des Radikalideals bzw. einer Basis des Radikalideals existieren⁹. Wird ein Gleichungssystem iterativ aufgebaut, so sollte man zu verhindern suchen, daß redundante Gleichungen zum Gleichungssystem hinzugefügt werden, denn sie verursachen einen globalen Rangverlust. In einem solchen iterativen Verfahren sollte vor dem Hinzufügen einer weiteren Gleichung geprüft werden, ob die Gleichung nicht schon eine Konsequenz des existierenden Systems ist („Ideal-Membership-Problem“). Hier führt der Weg zwangsläufig zu den Gröbner-Basen. Ist ein Gleichungssystem in Form einer Gröbner-Basis dargestellt, so läßt sich aufs einfachste überprüfen, ob eine Gleichung eine Konsequenz daraus ist oder nicht (es muß ein einfacher Divisionsalgorithmus angewandt werden). Auch verträgt sich die Berechnung einer Gröbner-Basis (Buchberger-Algorithmus) mit dem iterativen Charakter eines solchen Verfahrens. Eine vorhandene Basis kann verwendet werden, um nach dem Hinzufügen einer neuen Gleichung die Basis eines neuen Systems zu berechnen¹⁰. Allerdings kann ein doppelt-exponentielles Wachstum in der Anzahl der Gleichungen nicht ausgeschlossen werden. Die Gröbner-Basis selbst dient zwar zur Lösung des Ideal-Membership-Problems, ist selbst aber als beschreibendes Gleichungssystem i. allg. untauglich, da sie zu viele Gleichungen enthält, die selbst schon Konsequenz anderer Gleichungen sind. Sie wird i. allg. also als Gleichungssystem redundant sein. Sie könnte aber parallel zum eigentlichen Gleichungssystem mitgeführt werden. Die zweite Frage kann nun teilweise auch einfacher gelöst werden als durch die Verwendung des Hironaka-Satzes. In dem hier interessierenden Fall, daß es eine Kurve gibt, die sich einer isolierten Singularität von außen nähert, kann man mit wesentlich einfacheren Mitteln zum Ziel kommen. Der folgende Abschnitt soll sie vorführen.

§3 *t*-Ordnung

Generalannahme: Im folgenden wird angenommen, daß es in einem Gleichungssystem mindestens so viele Variablen wie Gleichungen gibt. Weiter wird angenommen, daß der zugrunde liegende Körper \mathbb{K} der Körper der reellen Zahlen \mathbb{R} oder der Körper der komplexen Zahlen \mathbb{C} ist.

Definition 9 (Singularität). *Seien*

$$g_1 = 0, \dots, g_m = 0$$

eine endliche Anzahl von polynomiellen Gleichungen in mehreren Unbekannten x_0, \dots, x_n , in vektorieller Schreibweise

$$g(x) = 0.$$

Sei weiter $G = g'$ die Ableitung von g nach allen Variablen. x heißt Singularität des Gleichungssystems $g = 0$, wenn gilt

$$|G(x)G^T(x)| = \det(G(x)G^T(x)) = 0.$$

Diese Definition der Singularität unterscheidet sich von der Definition der Singularität in der algebraischen Geometrie. Dort wird eine Singularität über die Varietät („Lösungsmenge“) einer Konsequenzenmenge eines Gleichungssystems („Ideal“) definiert¹¹. Die Definition 9 bezieht sich auf das Gleichungssystem selbst. Der Unterschied zeigt sich an einem

⁸siehe z. B. COX, LITTLE UND O'SHEA[15], Kapitel 4, §1.

⁹Eine Beschreibung findet sich in BECKER UND WEISPFENNIG[4]

¹⁰siehe BUCHBERGER[12]

¹¹siehe z. B. COX, LITTLE UND O'SHEA[15], Kapitel 9, §6, Definition 7.

einfachen Beispiel: die Gleichung

$$x^2 = 0$$

hat nach Definition 9 eine singuläre Lösung ($x = 0$). Gemäß der algebraischen Geometrie ist die nur aus der 0 bestehende Punktmenge aber singularitätsfrei.

Definition 10 (t-Ord). Sei $x :] - 1, 1[\rightarrow \mathbb{K}$ eine analytische Funktion, d. h. in eine Potenzreihe entwickelbar und auf $] - 1, 1[$ mit ihr übereinstimmend:

$$\forall t \in] - 1, 1[\quad (x(t) = x_n t^n + x_{n+1} t^{n+1} + \dots).$$

Dann sei das kleinste n , für das $x_n \neq 0$ gilt, die t -Ordnung $t\text{-Ord}$ von $x(t)$:

$$t\text{-Ord}(x) := n.$$

Gilt $x \equiv 0$, so sei die t -Ordnung auf ∞ gesetzt:

$$t\text{-Ord}(x) := \infty, \text{ für } x \equiv 0.$$

Das Auftreten einer Singularität wird durch den Rangverlust der Matrix G charakterisiert. Der Rangverlust von G ist wiederum mit $|GG^T| = 0$ äquivalent. Für die weiteren Betrachtungen erweist sich aber der Ausdruck $|GG^T|$ als ungünstig, an seiner statt soll ein anderer verwendet werden, der auf die t -Ordnungen der Matrixelemente zurückgreift.

Definition 11 (t-Matrix). Eine Matrix, deren Elemente analytische Funktionen im Sinne der Definition 10 sind, heiße t -Matrix.

Definition 12 (t-Ord einer Rechtecksmatrix, Q-Symbol). Sei A eine rechteckige $m \times n$ -Matrix, $m \leq n$ und sei A auch eine t -Matrix. Weiter sei Q_A die Menge aller quadratischen $m \times m$ -Submatrizen von A :

$$Q_A := \{B \mid B \text{ } m \times m\text{-Submatrix von } A, \\ \text{d. h. } B \text{ entsteht aus } A \text{ durch Weglassen von } n - m \text{ Spalten.}\}$$

Da auch jedes Element von Q_A eine t -Matrix ist, sei weiter definiert:

$$t\text{-Ord}(A) := \min_{B \in Q_A} t\text{-Ord}(\det(B)).$$

Es gilt mit obiger Notation folgende einfache

Feststellung 11. Sei A eine t -Matrix. Rangverlust von A für $t = 0 \iff t\text{-Ord}(A) > 0$.

BEWEIS. Hat A für $t = 0$ einen Rangverlust, so gilt für jede quadratische Submatrix B

$$|B(t)| = 0 \text{ für } t = 0.$$

Damit gilt $t\text{-Ord}(B) > 0$, und da B beliebige Submatrix war, gilt auch $t\text{-Ord}(A) > 0$. Hat A keinen Rangverlust für $t = 0$, so gibt es eine quadratische Submatrix B mit

$$|B(t)| \neq 0 \text{ für } t = 0,$$

damit ist dann $t\text{-Ord}(B) = 0$ und folglich $t\text{-Ord}(A) = 0$. □

Definition 13 (Durchquerbarkeit). Seien $g_1 = 0, \dots, g_m = 0$ eine endliche Anzahl von polynomiellen Gleichungen in mehreren Unbekannten x_0, \dots, x_n , in vektorieller Schreibweise $g(x) = 0$, und sei a eine Singularität des Gleichungssystems. Die Ableitungsmatrix von g werde wieder mit G bezeichnet. Dann heißt a durchquerbare Singularität, wenn folgendes gilt:

(1) Es gibt eine analytische Funktion $X = (x_1, \dots, x_n)^T :] - 1, 1[\rightarrow \mathbb{K}$ mit

$$\forall t \in] - 1, 1[\quad (g(x(t)) = 0).$$

(2) Es ist $x(0) = a$.

(3) Es ist $t\text{-Ord}(G \circ x) < \infty$. Dies bedeutet, daß es eine quadratische Submatrix gibt, deren Determinante für $t = 0$ eine isolierte Nullstelle hat.

$x(t)$ heißt durchquerende Funktion oder durchquerende Kurve.

Die Bedingungen in der Definition einer durchquerbaren Singularität besagen, daß es eine analytische Kurve gibt, die in der Lösungsmenge des Gleichungssystems verläuft, die Singularität selbst auch als Wert hat (sie im topologischen Sinne durchquert) und daß das singuläre Verhalten selbst in gewissem Sinne isolierbar ist.

Beispiel 8.

(1) Sei das Gleichungssystem $g(x, y) = 0$ mit

$$g(x, y) := xy$$

betrachtet (die Lösung dieses Systems bildet das Koordinatenkreuz). Dieses Gleichungssystem hat für $x = 0, y = 0$ eine Singularität. Diese ist aber durchquerbar, denn setzt man

$$x(t) := t$$

$$y(t) := 0,$$

so sind die Bedingungen der Definition erfüllt.

(2) Sei das Gleichungssystem $g(x) = 0$ mit

$$g(x) := x^2$$

betrachtet (die Lösung dieses Systems besteht nur aus dem Punkt $x = 0$). Die Singularität ist nicht durchquerbar. Denn ist $x(t)$ eine durchquerende Funktion, so folgt aus der ersten Bedingung in der Definition 13

$$\forall t \in] - 1, 1[\quad (x(t)^2 = 0)$$

die Feststellung

$$\forall t \in] - 1, 1[\quad (x(t) = 0),$$

damit ist die dritte Bedingung der Definition nicht erfüllt.

(3) Ebenfalls nicht durchquerbar ist die Singularität des Gleichungssystems

$$x = 0$$

$$xy = 0.$$

Denn die zweite Gleichung ist automatisch erfüllt, wenn die erste Gleichung gilt, daher gibt es eine globale Redundanz der Gleichungen, die die Undurchquerbarkeit zur Folge hat.

Bemerkung: Die Eigenschaft der Durchquerbarkeit ist eine lokale Eigenschaft. Es muß zwar in der Umgebung des durchquerten Punktes eine Kurve geben, die in der Lösungsmenge des Gleichungssystems liegt, aber es reicht, daß diese Kurve in einer hinreichend kleinen Umgebung existiert. Existiert z. B. eine solche Kurve nur auf $] - \varepsilon, \varepsilon[$, so wird durch $\tilde{f}(t) := f(\varepsilon t)$ eine ebensolche auf $] - 1, 1[$ definiert.

Die Kurve, die die Singularität durchquert, wird die Rolle eines Lotsen durch die Singularität spielen.

Definition 14 (Kurvenaufblasung). Sei $g(x) = 0$ ein Gleichungssystem polynomieller Gleichungen in mehreren Unbekannten und sei 0 eine Lösung des Gleichungssystems, d. h. $g(0) = 0$. Unter der Aufblasung von g bzgl. x_0 versteht man das sich wie folgt ergebende Gleichungssystem g^* :

- (1) g^* besteht zunächst aus den Gleichungen (Transformationsgleichungen)

$$\begin{aligned}x_1 - \lambda_1 x_0 &= 0 \\x_2 - \lambda_2 x_0 &= 0 \\&\vdots = \vdots \\x_n - \lambda_n x_0 &= 0.\end{aligned}$$

- (2) Dann wird in allen Gleichungen von $g(x) = 0$ jedes Vorkommen von x_i ersetzt durch $\lambda_i x_0$, dies ergebe das Gleichungssystem $\tilde{g}(x_0, \lambda) = 0$ (Substitution).
 (3) Anschließend wird aus den so modifizierten Gleichungen die höchste Potenz von x_0 abdividiert, die sich noch ohne Entstehung von Quotienten abdividieren lässt. Die so entstandenen Gleichungen bilden \bar{g} (Abdivision).
 (4) Die Gleichungen von \bar{g} werden zu g^* hinzugefügt.

Der Name „Kurvenaufblasung“ soll darauf anspielen, daß später die die Singularität durchquerende Kurve entscheiden wird, bezüglich welcher der Variablen aufgeblasen werden wird, d. h. welche der Komponenten die Funktion der Variable x_0 in obiger Definition übernehmen wird.

Definition 15 (Polaraufblasung). Sei $g(x) = 0$ ein Gleichungssystem polynomieller Gleichungen in mehreren Unbekannten und sei 0 eine Lösung des Gleichungssystems, d. h. $g(0) = 0$. Unter der Polaraufblasung von g versteht man das sich wie folgt ergebende Gleichungssystem g^* :

- (1) g^* besteht zunächst aus den Gleichungen (Transformationsgleichungen)

$$\begin{aligned}x_0 - r e_0 &= 0 \\x_1 - r e_1 &= 0 \\x_2 - r e_2 &= 0 \\&\vdots = \vdots \\x_n - r e_n &= 0 \\e_0^2 + e_1^2 + \cdots + e_n^2 - 1 &= 0 \text{ (Normzeile)}.\end{aligned}$$

- (2) Dann wird in allen Gleichungen von $g(x) = 0$ jedes Vorkommen von x_i ersetzt durch $r e_i$, dies ergebe das Gleichungssystem $\tilde{g}(r, e) = 0$ (Substitution).
 (3) Anschließend wird aus den so modifizierten Gleichungen die höchste Potenz von r abdividiert, die sich noch ohne Entstehung von Quotienten abdividieren lässt. Die so entstandenen Gleichungen bilden \bar{g} (Abdivision).
 (4) Die Gleichungen von \bar{g} werden zu g^* hinzugefügt.

Der Name „Polaraufblasung“ soll an den Übergang von kartesischen Koordinaten zu Polarkoordinaten erinnern, was im wesentlichen (bis auf die Normierungsbedingung) der gleiche Vorgang ist. Die Polaraufblasung ist der in der algebraischen Geometrie wohlbekannte „blow-up“¹².

¹²siehe z. B. COX, LITTLE UND O'SHEA[15], Kapitel 9, §7

Beispiel 9. *Das Gleichungssystem*

$$g(x_0, x_1) := x_1^2 - x_0^3 = 0$$

hat 0 als Lösung. Die Kurvenaufblasung bzgl. x_0 lautet

$$\begin{aligned} x_1 - \lambda_1 x_0 &= 0 \\ \lambda_1^2 - x_0 &= 0. \end{aligned}$$

Die Polaraufblasung lautet

$$\begin{aligned} x_0 - r e_0 &= 0 \\ x_1 - r e_1 &= 0 \\ e_1^2 - r e_0^3 &= 0 \\ e_0^2 + e_1^2 - 1 &= 0. \end{aligned}$$

§4 Die Kurvenaufblasung

Sei die Kurvenaufblasung bzgl. x_0 noch einmal formal hingeschrieben: seien dazu

$$g_i(x_1, \dots, x_n) = 0$$

die Gleichungen des Gleichungssystems $g(x) = 0$, 0 sei eine Lösung des Gleichungssystems und durchquerbare Singularität, $x(t)$ sei eine durchquerende Kurve und $x_0(t)$ eine der Komponenten von $x(t)$ mit niedrigster t -Ordnung. Dann sei weiter

$$\tilde{g}_i(x_0, \lambda_1, \dots, \lambda_n) := g_i(x_0, \lambda_1 x_0, \dots, \lambda_n x_0)$$

und

$$\bar{g}_i(x_0, \lambda_1, \dots, \lambda_n) := \frac{\tilde{g}_i(x_0, \lambda_1, \dots, \lambda_n)}{x_0^{m_i}},$$

wobei m_i der niedrigste in g_i vorkommende Monomgrad ist.

Von Bedeutung sind die Ableitungsmatrizen

$$G_i := g'_i = \left(\frac{\partial g_i}{\partial x_0} \dots \frac{\partial g_i}{\partial x_n} \right),$$

analog G^* , \tilde{G}_i und \bar{G}_i (diese aber bzgl. der in ihnen vorkommenden Variablen). Dann ist

$$\frac{\partial \tilde{g}_i}{\partial \lambda_j} = x_0 \frac{\partial g_i}{\partial x_j} \tag{33}$$

$$\frac{\partial \tilde{g}_i}{\partial x_0} = \frac{\partial g_i}{\partial x_0} + \sum_{j=1}^n \lambda_j \frac{\partial g_i}{\partial x_j}. \tag{34}$$

Weiter ist $\tilde{g}_i = x_0^{m_i} \bar{g}_i$ und somit

$$\frac{\partial \tilde{g}_i}{\partial \lambda_j} = x_0^{m_i} \frac{\partial \bar{g}_i}{\partial \lambda_j},$$

hat man eine durchquerende Kurve gegeben, so gilt auf dieser¹³

$$\frac{\partial \bar{g}_i}{\partial \lambda_j} = \frac{1}{x_0^{m_i}} \frac{\partial \tilde{g}_i}{\partial \lambda_j}. \tag{35}$$

¹³Diese Gleichung gilt nur lokal, d. h. in einer hinreichend kleinen Umgebung von $t = 0$ und ist darüber hinaus im Sinne einer Gleichung zwischen Elementen eines Körpers meromorpher Funktionen zu verstehen. So ist der Quotient auf der rechten Seite der Gleichung eigentlich für $t = 0$ nicht definiert, es gibt aber eine eindeutige analytische Fortsetzung in $t = 0$. Siehe dazu auch FREITAG UND BUSAM[21], KAPITEL III.

Ebenso gilt

$$\frac{\partial \tilde{g}_i}{\partial x_0} = x_0^{m_i} \frac{\partial \bar{g}_i}{\partial x_0} + \underbrace{m_i x_0^{m_i-1} \bar{g}_i}_{=0 \text{ auf durchquerender Kurve}},$$

woraus sich ergibt

$$\frac{\partial \bar{g}_i}{\partial x_0} = \frac{1}{x_0^{m_i}} \frac{\partial \tilde{g}_i}{\partial x_0}. \quad (36)$$

Daraus folgt

$$\bar{G}_i = \frac{1}{x_0^{m_i}} \tilde{G}_i$$

und für die komplette Ableitungsmatrix auf einer durchquerenden Kurve

$$\bar{G} = \underbrace{\begin{pmatrix} \frac{1}{x_0^{m_1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{x_0^{m_n}} \end{pmatrix}}_{:=D} \tilde{G}. \quad (37)$$

Teilt man G_i auf gemäß

$$G_i = (G_{i0} \quad G_{i1\dots n}),$$

so kann man schreiben

$$\tilde{G}_i = (G_{i0} + G_{i1\dots n} \lambda \quad x_0 G_{i1\dots n}),$$

als Gesamtmatrix

$$\begin{aligned} \tilde{G} &= (G_0 + G_{1\dots n} \lambda \quad x_0 G_{1\dots n}) \\ &= G \underbrace{\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \lambda_1 & x_0 & & \vdots \\ \vdots & 0 & \ddots & 0 \\ \lambda_n & 0 & \cdots & x_0 \end{pmatrix}}_{:=M}. \end{aligned} \quad (38)$$

Sei T die Ableitungsmatrix der Transformationsgleichungen

$$\begin{aligned} x_1 - \lambda_1 x_0 &= 0 \\ x_2 - \lambda_2 x_0 &= 0 \\ &\vdots \\ x_n - \lambda_n x_0 &= 0, \end{aligned}$$

also (mit der Anordnung der Variablen als $x_1, \dots, x_n, x_0, \lambda_1, \dots, \lambda_n$)

$$T = \begin{pmatrix} 1 & & -\lambda_1 & -x_0 & & \\ & \ddots & \vdots & & \ddots & \\ & & 1 & -\lambda_n & & -x_0 \end{pmatrix}.$$

Spaltet man T auf in Einheitsmatrix und Rest

$$T =: (E|R),$$

dann gilt schließlich

$$G^* = \begin{pmatrix} T \\ 0|D\tilde{G} \end{pmatrix} = \begin{pmatrix} T \\ 0|DGM \end{pmatrix} = \begin{pmatrix} E & R \\ 0 & DGM \end{pmatrix}. \quad (39)$$

Mit obiger Notation gilt folgende

Feststellung 12. *Seien die Matrizen G, D, M und G^* wie oben definiert und sei 0 Lösung und durchquerbare Singularität des Gleichungssystems $g(x) = 0$ mit durchquerender Kurve $x(t)$ und entsprechend definiertem $\lambda(t)$. Dann gilt:*

$$t\text{-Ord}(G^*(x(t), \lambda(t))) \leq t\text{-Ord}(DGM(x(t), \lambda(t))).$$

BEWEIS. Sei $X \in Q_{DGM}$ und sei R' die analoge Auswahl an Spalten aus R . Dann gilt

$$\begin{pmatrix} E & R' \\ 0 & X \end{pmatrix} \in Q_{G^*}.$$

Eine solche Matrix hat aber als Determinante $|X|$. Daraus folgt die Behauptung. \square

Um das folgende Lemma beweisen zu können, sind genauere Untersuchungen zur Berechnung der t -Ordnung einer Matrix nötig. Die folgenden sehr elementaren Feststellungen liefern das Nötigste dazu.

Die erste Feststellung betrifft die Multiplikation einer Zeile mit einer Funktion mit t -Ordnung 0 .

Feststellung 13. *Die t -Matrizen (in Zeilenschreibweise)*

$$A = \begin{pmatrix} \dots & \vdots & \dots \\ \dots & z^T & \dots \\ \dots & \vdots & \dots \end{pmatrix}$$

und

$$B = \begin{pmatrix} \dots & \vdots & \dots \\ \dots & ez^T & \dots \\ \dots & \vdots & \dots \end{pmatrix}$$

mit einer Funktion $e(t)$, für die gilt: $t\text{-Ord}(e) = 0$, haben die gleiche t -Ordnung:

$$t\text{-Ord}(A) = t\text{-Ord}(B).$$

BEWEIS. Zu jedem $M_1 \in Q_A$ gibt es ein $M_2 \in Q_B$ mit

$$|M_1| = e|M_2|$$

und daher

$$t\text{-Ord}(M_1) = t\text{-Ord}(M_2).$$

Umgekehrt gibt es zu jedem $M_2 \in Q_B$ ein $M_1 \in Q_A$ mit dieser Eigenschaft. Daraus folgt die Behauptung. \square

Analoges gilt für die Spalten.

Feststellung 14. *Die t -Matrizen (in Spaltenschreibweise)*

$$A = \begin{pmatrix} \dots & \vdots & \dots \\ \dots & z & \dots \\ \dots & \vdots & \dots \end{pmatrix}$$

und

$$B = \begin{pmatrix} \dots & \vdots & \dots \\ \dots & ez & \dots \\ \dots & \vdots & \dots \end{pmatrix}$$

mit einer Funktion $e(t)$, für die gilt: $\text{t-Ord}(e) = 0$, haben die gleiche t -Ordnung:

$$\text{t-Ord}(A) = \text{t-Ord}(B).$$

BEWEIS. Sei M_1 eine quadratische Submatrix von B . Enthält M_1 die mit e multiplizierte Spalte nicht, so ist M_1 auch eine quadratische Submatrix von A . Enthält M_1 die Spalte, so gibt es eine quadratische Submatrix M_2 von A mit

$$|M_1| = e|M_2|,$$

dann ist wieder

$$\text{t-Ord}(M_1) = \text{t-Ord}(M_2).$$

Umgekehrt gibt es zu jeder quadratischen Submatrix M_2 von A ein analoges M_1 von B mit der gleichen Eigenschaft. Daraus ergibt sich wieder die Behauptung. \square

Feststellung 15. *Die t -Rechtecksmatrizen (in Spaltenschreibweise)*

$$A := \begin{pmatrix} s_0 + \alpha(t)s_1 & s_1 & s_2 & \dots & s_n \end{pmatrix}$$

und

$$B := \begin{pmatrix} s_0 & s_1 & s_2 & \dots & s_n \end{pmatrix}$$

haben die gleiche t -Ordnung für eine analytische Funktion $\alpha(t)$.

BEWEIS. Sind die beiden Matrizen quadratisch, so ist die Behauptung klar. Im echten Rechtecksfall sei P_A die Menge aller quadratischen Submatrizen von A , die die erste Spalte nicht enthalten, S_A die Menge der Matrizen, die die ersten beiden Spalten enthalten und T_A die sonstigen, analog mit B :

$$P_A := \{M \in Q_A \mid M = \begin{pmatrix} s_i & \dots & \dots \end{pmatrix}, i > 0\}$$

$$S_A := \{M \in Q_A \mid M = \begin{pmatrix} s_0 + \alpha(t)s_1 & s_1 & \dots \end{pmatrix}\}$$

$$T_A := Q_A \setminus (P_A \cup S_A)$$

$$P_B := \{M \in Q_B \mid M = \begin{pmatrix} s_i & \dots & \dots \end{pmatrix}, i > 0\}$$

$$S_B := \{M \in Q_B \mid M = \begin{pmatrix} s_0 & s_1 & \dots \end{pmatrix}\}$$

$$T_B := Q_B \setminus (P_B \cup S_B).$$

(1) Da $P_A = P_B$ gilt, folgt

$$\text{t-Ord}(P_A) = \text{t-Ord}(P_B).$$

- (2) Zu jedem $M \in S_A$ gibt es ein analoges $M' \in S_B$ (die gleiche Auswahl an Spalten) mit $|M| = |M'|$ und umgekehrt, daher gilt

$$\text{t-Ord}(S_A) = \text{t-Ord}(S_B).$$

- (3) Sei nun $M \in T_A$. Es wird gezeigt, daß zu M eine Matrix $M' \in T_B$ mit gleicher t -Ordnung existiert oder daß es eine Matrix $N \in P_A = P_B$ gibt mit nicht höherer t -Ordnung. Umgekehrt findet sich auch zu einem $M' \in T_B$ ein solches $M \in T_A$ oder wieder ein entsprechendes N . Daraus folgt dann die Gesamtbehauptung. Sei also

$$M = \left(\begin{array}{ccc} s_0 + \alpha(t)s_1 & s_j & \dots \end{array} \right), j > 1.$$

Es gilt dann

$$|M| = \left| \underbrace{\left(\begin{array}{ccc} s_0 & s_j & \dots \end{array} \right)}_{:=M'} + \alpha(t) \underbrace{\left(\begin{array}{ccc} s_1 & s_j & \dots \end{array} \right)}_{:=N} \right|,$$

woraus folgt

$$\text{t-Ord}(M) \geq \min\{\text{t-Ord}(M'), \text{t-Ord}(\alpha) + \text{t-Ord}(N)\}.$$

Gilt nun $\text{t-Ord}(M') < \text{t-Ord}(N)$, so ist $\text{t-Ord}(M) = \text{t-Ord}(M')$ (da αN die niedrigste t -Potenz in $|M'|$ nicht annihilieren kann). Gilt dagegen $\text{t-Ord}(M') \geq \text{t-Ord}(N)$, so ist die t -Ordnung von N nicht höher als die von M und es gilt $N \in P_A = P_B$.

Ist umgekehrt ein solches $M' \in T_B$ gegeben, so gilt wieder mit analogem N : falls $\text{t-Ord}(M') < \text{t-Ord}(N)$, dann gilt $\text{t-Ord}(M) = \text{t-Ord}(M')$ und das gewünschte M ist gefunden. Andernfalls gibt es ein $N \in P_A = P_B$ mit nicht höherer Ordnung als die von M' .

□

Von entscheidender Bedeutung für Algorithmus ist das folgende

Lemma 1 (Abstieg in der t -Ord bei Kurvenaufblasung). *Sei*

$$g_0(x) = 0, \dots, g_l(x) = 0$$

ein Gleichungssystem polynomieller Gleichungen in mehreren Unbekannten x_0, \dots, x_n ($l \leq n$ nach Generalannahme) und sei 0 durchquerbare Singularität des Gleichungssystems mit durchquerender Kurve $x(t)$. Der niedrigste Monomgrad der Zeile g_i sei m_i und es existiere mindestens ein $m_i \geq 2$. Sei weiter $x_j(t)$ die Komponente mit niedrigster t -Ordnung und $g^ = 0$ die Kurvenaufblasung von $g = 0$ bezüglich x_j . Dann gilt:*

$$\text{t-Ord}(G^*(x(t), \lambda(t))) < \text{t-Ord}(G(x(t))).$$

Insbesondere gilt auch

$$\text{t-Ord}(DGM(x(t), \lambda(t))) < \text{t-Ord}(G(x(t))),$$

wobei DGM das aufgeblasene Gleichungssystem ohne Transformationsgleichungen darstellt.

BEWEIS. O. B. d. A. sei $j = 0$.

Da 0 Lösung des Gleichungssystems ist, gilt $\forall i(m_i > 0)$.

Auch gilt, da $x(t)$ die Singularität 0 durchquert $\forall i(\text{t-Ord}(x_i) > 0)$.

Nach Feststellung 12 gilt

$$\text{t-Ord}(G^*) \leq \text{t-Ord}(DGM).$$

Es ist daher nur noch

$$\text{t-Ord}(DGM) < \text{t-Ord}(G)$$

zu zeigen. Sei dazu $A \in Q_{DG}$. Dann gibt es ein $B \in Q_G$ mit $A = DB$. Umgekehrt gibt es zu $B \in Q_G$ ein entsprechendes $A = DB \in Q_{DG}$. Für diese A und B gilt:

$$\begin{aligned} |A| &= |DB| \\ &= x_0^{-\sum_{i=0}^l m_i} |B|. \end{aligned}$$

Damit ergibt sich

$$\text{t-Ord}(A) = \text{t-Ord}(B) - \text{t-Ord}(x_0) \sum_{i=0}^l m_i,$$

woraus folgt

$$\text{t-Ord}(DG) = \text{t-Ord}(G) - \text{t-Ord}(x_0) \sum_{i=0}^l m_i. \quad (40)$$

Für den nächsten Beweisschritt sei eine Matrix Z mit ZM verglichen. Sei dazu Z spaltenweise niedergeschrieben:

$$Z = \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ z_0 & z_1 & \dots & z_n \\ \vdots & \vdots & \dots & \vdots \end{pmatrix}.$$

ZM lässt sich dann schreiben als

$$ZM = \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ z_0 + \sum_{i=1}^n \lambda_i z_i & x_0 z_1 & \dots & x_0 z_n \\ \vdots & \vdots & \dots & \vdots \end{pmatrix}.$$

Sei weiter definiert

$$Z' := \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ z_0 + \sum_{i=1}^n \lambda_i z_i & z_1 & \dots & z_n \\ \vdots & \vdots & \dots & \vdots \end{pmatrix}.$$

Es ist

$$\text{t-Ord}(Z') = \text{t-Ord}(Z)$$

nach Feststellung 15. Ist nun eine quadratische Submatrix $A \in Q_{ZM}$ gewählt, so gibt es ein $B \in Q_{Z'}$ mit

$$|A| = |B|x_0^e,$$

und umgekehrt gibt es zu jedem $B \in Q_{Z'}$ ein $A = BM \in Q_{ZM}$ ebenfalls mit

$$|A| = |B|x_0^e,$$

wobei $e = l + 1$ gilt, wenn die erste Spalte aus ZM nicht in der Auswahl war, ansonsten gilt $e = l$, insgesamt aber $e \leq l + 1$. Damit folgt

$$\text{t-Ord}(A) = \text{t-Ord}(B) + e \text{t-Ord}(x_0)$$

und weiter

$$\begin{aligned} \text{t-Ord}(ZM) &\leq \text{t-Ord}(Z') + (l+1) \text{t-Ord}(x_0) \\ &= \text{t-Ord}(Z) + (l+1) \text{t-Ord}(x_0). \end{aligned}$$

Dies wird speziell zu

$$\text{t-Ord}(DGM) \leq \text{t-Ord}(DG) + (l+1) \text{t-Ord}(x_0). \quad (41)$$

(40) und (41) ergeben zusammen

$$\begin{aligned} \text{t-Ord}(DGM) &\leq \text{t-Ord}(DG) + (l+1) \text{t-Ord}(x_0) \\ &= \text{t-Ord}(G) - \text{t-Ord}(x_0) \sum_{i=0}^l m_i + (l+1) \text{t-Ord}(x_0) \\ &= \text{t-Ord}(G) - \text{t-Ord}(x_0) \left(\sum_{i=0}^l m_i - l - 1 \right). \end{aligned}$$

Es ist aber $\sum_{i=0}^l m_i > \sum_{i=0}^l 1 = (l+1)$, da alle m_i größer als 0 sind und mindestens ein m_i größer als 1 ist. Damit folgt weiter

$$\sum_{i=0}^l m_i - l - 1 > 0$$

und wegen $\text{t-Ord}(x_0) > 0$ ergibt sich damit

$$\text{t-Ord}(DGM) < \text{t-Ord}(G),$$

was zu zeigen war. □

Lemma 2 (Erhaltung der Durchquerbarkeit bei Kurvenaufblasung). *Sei*

$$g(x) = 0$$

ein polynomielles Gleichungssystem in den Unbekannten x_0, \dots, x_n , wobei eine Gleichung mindestens kleinsten Monomgrad 2 hat. $x = 0$ sei Lösung des Gleichungssystems und durchquerbare Singularität, weiter $x(t)$ eine durchquerende Kurve und $x_0(t)$ eine Komponente niedrigster t -Ordnung. Seien weiter

$$\lambda_i(t) := \frac{x_i(t)}{x_0(t)}, \quad \lambda_i := \lambda_i(0)$$

(dies ist in einer hinreichend kleinen Umgebung von $t = 0$ wohldefiniert und wieder im Sinne einer Gleichung im Körper meromorpher Funktionen gemeint) und sei g^ die Kurvenaufblasung von g bzgl. x_0 .*

Ist dann $(0, \lambda)$ eine Singularität von g^ , so ist sie wieder durchquerbar.*

BEWEIS. Es werden die drei Bedingungen der Definition 13 nachgewiesen.

- (1) Die $\lambda_i(t)$ sind analytische Funktionen, durch Übergang von $\lambda_i(t)$ zu $\lambda_i(\alpha t)$ kann angenommen werden auf ganz $] - 1, 1[$.

Weiter gilt

$$x_i(t) - \lambda_i(t)x_0(t) = 0$$

und

$$\begin{aligned}\bar{g}_i(x_0(t), \lambda_1(t), \dots, \lambda_n(t)) &= \frac{\tilde{g}_i(x_0(t), \lambda_1(t), \dots, \lambda_n(t))}{x_0(t)^{m_i}} \\ &= \frac{g_i(x_0(t), \lambda_1(t)x_0(t), \dots, \lambda_n(t)x_0(t))}{x_0(t)^{m_i}} \\ &= \frac{g_i(x_0(t), x_1(t), \dots, x_n(t))}{x_0(t)^{m_i}} \\ &= 0, \text{ denn im Zähler steht eine Funktion } \equiv 0.\end{aligned}$$

(2) $(x(0), \lambda(0)) = (0, \lambda_1, \dots, \lambda_n)$.

(3) Es ist $t\text{-Ord}(G^*) < t\text{-Ord}(G) < \infty$ nach Lemma 1.

□

Allerdings sichert dieses Lemma nur die Durchquerbarkeit des Punktes, der durch die Kurve bestimmt ist. Bei der Aufblasung wird die Singularität aber eventuell in mehrere Punkte vergrößert, von denen sich andere als undurchquerbar herausstellen können, z. B.

$$g(x, y) = (y + x)(y - x)^2.$$

Die Ableitung G ist

$$G = \left(\begin{array}{cc} (y - x)^2 - 2(y - x)(y + x) & (y - x)^2 + 2(y - x)(y + x) \end{array} \right).$$

Sei

$$\left(\begin{array}{c} x(t) \\ y(t) \end{array} \right) := \left(\begin{array}{c} t \\ -t \end{array} \right)$$

eine durchquerende Kurve, auf dieser Kurve gilt

$$G(x(t), y(t)) = \left(\begin{array}{cc} (2t)^2 & (2t)^2 \end{array} \right).$$

Nach der Aufblasung $y = \lambda x$ lautet die Gleichung

$$\tilde{g} = (\lambda + 1)x(\lambda - 1)^2x^2$$

und weiter

$$\bar{g} = (\lambda + 1)(\lambda - 1)^2.$$

Die letzte Gleichung hat die beiden Lösungen $\lambda = -1$ (der Grenzwert der obigen durchquerenden Kurve) und $\lambda = 1$. Letzteres ist aber undurchquerbar (als einzige durchquerende Kurve kommt $\lambda(t) \equiv 1$ in Frage, die immer in der Singularität verharrt). Die Lösung $\lambda = 1$ entspricht dem Limes von

$$\left(\begin{array}{c} x(t) \\ y(t) \end{array} \right) := \left(\begin{array}{c} t \\ t \end{array} \right),$$

auf dieser Kurve ist aber

$$G(x(t), y(t)) = \left(\begin{array}{cc} 0 & 0 \end{array} \right).$$

Eine immer wieder auftretende Maßnahme ist die Umentwicklung: eine Variable x_i nimmt nicht den Wert 0 an, sondern einen Wert $c_i \neq 0$, daher wird eine neue Variable per

$$x_i = \xi_i + c_i$$

eingeführt. Die neue Variable ξ_i nimmt nun den Wert 0 an. Die folgende Feststellung besagt, daß sich dabei die Ableitungsmatrix nicht ändert.

Feststellung 16. Sei $g(x_0, \dots, x_n) = 0$ ein polynomielles Gleichungssystem und sei definiert

$$\hat{g}(x_0, \dots, x_{i-1}, \xi_i, x_{i+1}, \dots, x_n) := g(x_0, \dots, x_{i-1}, \xi_i + c_i, x_{i+1}, \dots, x_n).$$

Sei weiter $x(t)$ eine reell-analytische Kurve mit $x_i(0) = c_i$ und sei

$$\hat{x}_j(t) := \begin{cases} x_j(t) & \text{für } j \neq i \\ x_i(t) - c_i & \text{für } j = i \end{cases}.$$

Dann gilt

$$\partial_j \hat{g} = \partial_j g,$$

also

$$\text{t-Ord}(\hat{G} \circ \hat{x}) = \text{t-Ord}(G \circ x).$$

BEWEIS. Der Beweis ist offensichtlich. □

Eine zweite elementare Maßnahme ist die Bildung der Linearkombination zweier Gleichungen. Dies ist notwendig, wenn zwei Gleichungen denselben linearen Anteil besitzen. In diesem Fall muß von einer Gleichung der kleinste Monomgrad, nämlich der zum linearen Teil gehörige, beseitigt werden.

Feststellung 17. Sei $g(x_0, \dots, x_n) = 0$ ein polynomielles Gleichungssystem und es gehe \hat{g} daraus hervor, indem eine Gleichung $g_i = 0$ durch $g_i + \alpha g_j = 0$, $i \neq j$ ersetzt werde. Sei weiter $x(t)$ eine reell-analytische Kurve. Dann gilt

$$\text{t-Ord}(\hat{G} \circ x) = \text{t-Ord}(G \circ x).$$

BEWEIS. Der Beweis ist trivial. □

Satz 1 (Algorithmus Kurvenaufblasung). Sei $g(x) = 0$ ein Gleichungssystem polynomieller Gleichungen in mehreren Unbekannten und sei 0 eine durchquerbare Singularität des Gleichungssystems mit durchquerender Kurve $x(t)$. Dann endet der folgende Algorithmus nach endlich vielen Schritten in einem Gleichungssystem ohne Rangverlust.

ALGORITHMUS (KURVENAUFBLASUNG)

- (1) Analog zum Gaußschen Eliminationsverfahren kann man durch Linearkombination der Gleichungen erreichen, daß die linearen Anteile eine rechte obere Dreiecksmatrix bilden. Dadurch wird $t\text{-Ord}(G)$ nach Feststellung 17 nicht verändert.
Entweder besitzen danach alle Gleichungen noch einen linearen Anteil (dann hat G Vollrang und der Algorithmus terminiert) oder mindestens eine Gleichung besitzt keinen linearen Anteil. Damit sind die Voraussetzungen von Lemma 1 gegeben.
- (2) Bezüglich der Komponente von $x(t)$ mit der niedrigsten t -Ordnung wird eine Kurvenaufblasung durchgeführt. Die durchquerende Kurve $x(t)$ wird analog Lemma 2 um $\lambda(t)$ erweitert. Nach Lemma 1 sinkt die t -Ordnung.
- (3) Gilt $t\text{-Ord}(G^*) = 0$, hat G^* keinen Rangverlust mehr, der Algorithmus terminiert mit dem Ergebnis g^* .
- (4) Die Singularität wird durch eine Umentwicklung wieder in den Ursprung geschoben. Existieren Komponenten der durchquerenden Kurve, die für $t = 0$ nicht den Wert 0 annehmen, so kann nach einer Umentwicklung angenommen werden, daß dies der Fall ist. Die Umentwicklungsgleichungen werden nicht zu den anderen Gleichungen hinzugefügt (sie stören aber lediglich die Aussage, daß alle Komponenten der durchquerenden Kurve den Wert 0 annehmen).
- (5) Es wird nun $g := g^*$ ohne Transformations- und Umentwicklungsgleichungen, $x(t) := (x_0(t), \lambda(t))$ gesetzt und wieder bei Punkt 1 begonnen.

Das Weglassen der Transformations- und Umentwicklungsgleichungen für die nachfolgenden Aufblasungen wird ermöglicht, weil nach Lemma 1 nicht nur $t\text{-Ord}(G^*)$ sondern auch $t\text{-Ord}(DGM)$ kleiner ist als $t\text{-Ord}(G)$.

Abgesehen davon, daß der Algorithmus nur lokal funktioniert, d. h. in einer hinreichend kleinen Umgebung der Singularität, kann bei einer Kurvenaufblasung ein Teil der Lösungsmenge des Gleichungssystems verloren gehen. Sei z. B. $g(x, y) = 0$ mit

$$g(x, y) := xy.$$

Die Lösungsmenge besteht aus dem Koordinatenkreuz. Sei weiter

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} t \\ 0 \end{pmatrix}$$

eine durchquerende Kurve. Die Aufblasung lautet dann

$$y = \eta x,$$

somit $\bar{g} = 0$ mit

$$\bar{g} := \eta.$$

Damit „fehlen“ aber alle Lösungen mit $y \neq 0$. Diese würden geliefert werden, wäre die durchquerende Kurve

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 \\ t \end{pmatrix}$$

gewesen, aber dann fehlten alle Lösungen mit $x \neq 0$. Der obige Algorithmus zur Kurvenaufblasung liefert also eventuell nur einen Teil der ursprünglichen Varietät. Es ist allerdings immer der Teil der Varietät vorhanden, in der die durchquerende Kurve verläuft.

Hierauf beruht der Wert des Algorithmus. Es geht zwar möglicherweise bei der Umformulierung lokal ein Teil der Mannigfaltigkeit verloren, die durchquerende Kurve ist aber nicht betroffen und auf ihr kann die Singularität durchquert werden.

Diese Eigenschaft der Nichtäquivalenz der Umformulierung besitzt die Polaraufblasung nicht, sie ist im Gegensatz zur Kurvenaufblasung „uniform“. Daher soll hier der Zusammenhang zwischen Kurvenaufblasung und Polaraufblasung genauer beleuchtet werden.

§5 Die Polaraufblasung

Die Polaraufblasung ist mit der Kurvenaufblasung sehr ähnlich, jedoch gilt es feine Unterschiede zu beachten. Zunächst einmal gilt es, einen Zusammenhang zwischen den beiden Verfahren herzustellen.

Lemma 3 (Zusammenhang Kurven- und Polaraufblasung). *Sei $g = 0$ eine einzelne Gleichung in den Unbekannten x_0, \dots, x_n , 0 sei durchquerbare Singularität und $x(t)$ sei durchquerende Kurve und $x_0(t)$ eine der Komponenten mit niedrigster t -Ordnung. Sei \bar{g} das Ergebnis der Kurvenaufblasung bzgl. x_0 und sei \hat{g} das Ergebnis der Polaraufblasung der Zeile $g = 0$. Dann gilt für ein $k > 0$ in einer hinreichend kleinen Umgebung von $t = 0$:*

$$\begin{aligned}\frac{\partial \hat{g}}{\partial e_i} &= \frac{\partial \bar{g}}{\partial \lambda_i} e_0^{k-1} \text{ für } i > 0 \\ \frac{\partial \hat{g}}{\partial r} &= \frac{\partial \bar{g}}{\partial x_0} e_0^{k+1} \\ \frac{\partial \hat{g}}{\partial e_0} &= r \frac{\partial \bar{g}}{\partial x_0} e_0^k - \sum_{i=1}^n e_i \frac{\partial \bar{g}}{\partial \lambda_i} e_0^{k-2}.\end{aligned}$$

BEWEIS. Sei

$$x_i(t) = \lambda_i(t)x_0(t)$$

und

$$x_i(t) = r(t)e_i(t).$$

Dann ist in einer hinreichend kleinen Umgebung von $t = 0$

$$\lambda_i(t) = \frac{e_i(t)}{e_0(t)}.$$

Da mindestens ein $e_i(0)$ von Null verschieden ist (wg. $\sum_i e_i^2 = 1$), schließt man hieraus

$$e_0(t) \neq 0, \text{ in hinreichend kleiner Umgebung von } t = 0.$$

Weiter ist

$$\begin{aligned}r^k \hat{g}(e_0, \dots, e_n, r) &= g(re_0, \dots, re_n) \\ &= g(x_0, \dots, x_n) \\ &= g(x_0, \lambda_1 x_0, \dots, \lambda_n x_0) \\ &= x_0^k \bar{g}(x_0, \lambda_1, \dots, \lambda_n).\end{aligned}$$

Daraus ergibt sich

$$\begin{aligned}\hat{g}(e_0, \dots, e_n, r) &= \frac{x_0^k}{r^k} \bar{g}(x_0, \lambda_1, \dots, \lambda_n) \\ &= e_0^k \bar{g}\left(x_0, \frac{e_1}{e_0}, \dots, \frac{e_n}{e_0}\right),\end{aligned}$$

also

$$\hat{g}(e_0, e_1, \dots, e_n, r) = \bar{g}\left(re_0, \frac{e_1}{e_0}, \dots, \frac{e_n}{e_0}\right)e_0^k, \quad (42)$$

wobei k der niedrigste vorkommende Monomgrad ist.

Aus (42) folgt unmittelbar durch differenzieren nach e_i

$$\frac{\partial \hat{g}}{\partial e_i} = \frac{\partial \bar{g}}{\partial \lambda_i} e_0^{k-1} \text{ für } i > 0.$$

Durch differenzieren von (42) nach r ergibt sich

$$\frac{\partial \hat{g}}{\partial r} = \frac{\partial \bar{g}}{\partial x_0} e_0^{k+1}.$$

Für die Differenzierung nach e_0 folgt

$$\frac{\partial \hat{g}}{\partial e_0} = r \frac{\partial \bar{g}}{\partial x_0} e_0^k - \sum_{i=1}^n e_i \frac{\partial \bar{g}}{\partial \lambda_i} \frac{e_0^k}{e_0^2} + \underbrace{ke_0^{k-1} \bar{g}}_{=0 \text{ wg. } \bar{g}=0}.$$

□

Das obige Lemma lässt sich für mehrere Gleichungen noch einmal so zusammenfassen: ist \bar{g} das Ergebnis der Kurvenaufblasung bzgl. x_0 und \bar{G} die entsprechende Ableitungsmatrix mit den Spalten d_i ,

$$\bar{G} = \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ d_0 & d_1 & \dots & d_n \\ \vdots & \vdots & \dots & \vdots \end{pmatrix}, \quad (43)$$

so lässt sich die Ableitungsmatrix \hat{G} von \hat{g} mit der Variablenreihenfolge r, e_0, e_1, \dots, e_n so schreiben:

$$\hat{G} = \begin{pmatrix} e_0^{k_0-2} & & \\ & \ddots & \\ & & e_0^{k_l-2} \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots & \dots & \vdots \\ e_0^3 d_0 & re_0^2 d_0 - \sum_{i=1}^n e_i d_i & e_0 d_1 & \dots & e_0 d_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{pmatrix}.$$

Nimmt man nun zu den Gleichungen auch noch die Normzeile $\sum_{i=0}^n e_i^2 = 1$ hinzu, so lautet die gesamte Matrix

$$G^* = \begin{pmatrix} 2 & & \\ & e_0^{k_0-2} & \\ & & \ddots & \\ & & & e_0^{k_l-2} \end{pmatrix} \begin{pmatrix} 0 & e_0 & e_1 & \dots & e_n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ e_0^3 d_0 & re_0^2 d_0 - \sum_{i=1}^n e_i d_i & e_0 d_1 & \dots & e_0 d_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{pmatrix}. \quad (44)$$

(Da die $e_0 \neq 0$ in einer hinreichend kleinen Umgebung von $t = 0$ sind, stören eventuelle negative Exponenten nicht.) Damit lässt sich die t -Ordnung der Matrix (44) wie folgt

berechnen:

$$\text{t-Ord}(G^*) = \text{t-Ord} \left(\left(\begin{array}{cccccc} 0 & e_0 & e_1 & \dots & e_n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ e_0^3 d_0 & r e_0^2 d_0 - \sum_{i=1}^n e_i d_i & e_0 d_1 & \dots & e_0 d_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{array} \right) \right) \quad (\text{Feststellung 13})$$

$$= \text{t-Ord} \left(\left(\begin{array}{cccccc} 0 & e_0 & \frac{e_1}{e_0} & \dots & \frac{e_n}{e_0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ d_0 & r e_0^2 d_0 - \sum_{i=1}^n e_i d_i & d_1 & \dots & d_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{array} \right) \right) \quad (\text{Feststellung 14})$$

$$= \text{t-Ord} \left(\left(\begin{array}{cccccc} 0 & e_0 + \frac{e_1^2}{e_0} + \dots + \frac{e_n^2}{e_0} & \frac{e_1}{e_0} & \dots & \frac{e_n}{e_0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ d_0 & r e_0^2 d_0 & d_1 & \dots & d_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{array} \right) \right) \quad (\text{Feststellung 15}]$$

$$= \text{t-Ord} \left(\left(\begin{array}{cccccc} 0 & \frac{1}{e_0} & \frac{e_1}{e_0} & \dots & \frac{e_n}{e_0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ d_0 & 0 & d_1 & \dots & d_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{array} \right) \right) \quad (\text{Feststellung 15})$$

$$= \text{t-Ord} \left(\left(\begin{array}{cccccc} 0 & 1 & \frac{e_1}{e_0} & \dots & \frac{e_n}{e_0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ d_0 & 0 & d_1 & \dots & d_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{array} \right) \right) \quad (\text{Feststellung 14].}$$

Aus letzterem und mit der Bezeichnung aus Gleichung (43) folgt weiter

$$\text{t-Ord}(G^*) \leq \text{t-Ord}(\bar{G}), \quad (45)$$

denn jede quadratische Submatrix Q von \bar{G} kann per

$$\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$$

zu einer quadratischen Submatrix von G^* gemacht werden.

Das folgende Lemma benötigt im Gegensatz zu den bisherigen Folgerungen die Voraussetzung

$$\mathbb{K} = \mathbb{R}.$$

Lemma 4 (Erhaltung der Durchquerbarkeit bei Polaraufblasung).

Sei der zu Grunde liegende Körper der Körper der reellen Zahlen. Sei weiter

$$g(x_0, \dots, x_n) = 0$$

ein polynomielles Gleichungssystem in den Unbekannten x_0, \dots, x_n . $x = 0$ sei Lösung des Gleichungssystems und durchquerbare Singularität und $x(t)$ eine durchquerende Kurve, sowie $x_0(t)$ eine Komponente niedrigster t -Ordnung. Seien weiter¹⁴

$$r(t) := x_0(t) \sqrt{1 + \sum_{i=1}^n \frac{x_i(t)^2}{x_0(t)^2}}, \quad e_i(t) := \frac{x_i(t)}{r(t)}, \quad e_i := e_i(0),$$

und sei g^* die Polaraufblasung von g .

Ist dann $(0, e)$ eine Singularität von g^* , so ist sie durchquerbar.

BEWEIS. Es werden die drei Bedingungen der Definition 13 nachgewiesen.

- (1) Zunächst einmal ist $r(t)$ eine analytische Funktion, denn $\sqrt{1+x}$ ist lokal eine analytische Funktion, darüberhinaus ist der Term unter der Wurzel

$$\sum_{i=1}^n \frac{x_i(t)^2}{x_0(t)^2}$$

wegen der Annahme $\mathbb{K} = \mathbb{R}$ nicht-negativ. Durch Übergang von $r(t)$ zu $r(\alpha t)$ kann angenommen werden, daß $r(t)$ auf ganz $] -1, 1[$ analytisch ist.

Weiter gilt

$$x_i(t) - r e_i(t) = 0$$

und

$$\begin{aligned} \bar{g}_i(r(t), e_0(t), \dots, e_n(t)) &= \frac{\tilde{g}_i(r(t), e_0(t), \dots, e_n(t))}{r(t)^{m_i}} \\ &= \frac{g_i(r(t)e_0(t), \dots, r(t)e_n(t))}{r(t)^{m_i}} \\ &= \frac{g_i(x_0(t), x_1(t), \dots, x_n(t))}{r(t)^{m_i}} \\ &= 0, \text{ denn im Zähler steht eine Funktion } \equiv 0. \end{aligned}$$

¹⁴ $r(t)$ kann entweder auf diese Weise oder per

$$r(t) := \sqrt{\sum_{i=0}^n x_i(t)^2}$$

(falls die niedrigste t -Ordnung gerade ist) bzw.

$$r(t) := \begin{cases} \sqrt{\sum_{i=0}^n x_i(t)^2} & \text{für } t \geq 0 \\ -\sqrt{\sum_{i=0}^n x_i(t)^2} & \text{für } t < 0 \end{cases}$$

(falls die niedrigste t -Ordnung ungerade ist) geschehen. Diese etwas umständlich anmutenden Definitionen rühren daher, daß die Komponenten $e_i(t)$, die zu den Komponenten $x_i(t)$ mit der niedrigsten t -Ordnung gehören, nach der Aufblasung Werte $\neq 0$ annehmen, und beim Durchgang durch die Singularität keinen Sprung erleiden dürfen, wie das z. B. bei $\frac{t}{\sqrt{t^2}} = \frac{t}{|t|}$ der Fall wäre. Somit ist eine Wahl zu treffen zwischen

$$r(t) := \sqrt{\sum_{i=0}^n x_i(t)^2}$$

und

$$r(t) := \operatorname{sgn}(t) \sqrt{\sum_{i=0}^n x_i(t)^2}.$$

$$(2) (r(0), e(0)) = (0, e_0, \dots, e_n).$$

(3) Es sei \bar{g} das Ergebnis der analogen Kurvenaufblasung. Nach Lemma 1 und (45) gilt

$$\text{t-Ord}(G^*) \leq \text{t-Ord}(\bar{G}) < \text{t-Ord}(G) < \infty.$$

□

Allerdings gilt zu diesem Lemma analog zu der nach Lemma 2 gemachten Bemerkung die Einschränkung, daß lediglich die Erhaltung der Durchquerbarkeit des durch die durchquerende Kurve bestimmten Punktes garantiert wird. Jedoch kann die Singularität in mehrere Punkte aufgeblasen werden, von denen sich einige als undurchquerbar herausstellen können. Als Beispiel hierfür diene die Gleichung

$$xy^2 = 0.$$

Die Singularität $x = 0, y = 0$ ist durchquerbar mittels der Kurve $x(t) = 0, y(t) = t$. Nach der Aufblasung lauten die Gleichungen

$$\begin{aligned} e_x e_y^2 &= 0 \\ e_x^2 + e_y^2 - 1 &= 0. \end{aligned}$$

Der durch die Kurve bestimmte Punkt ist $r = 0, e_x = 0, e_y = 1$ und nicht mehr singulär. Aber bei der Aufblasung von $x = 0, y = 0$ ist auch $r = 0, e_x = 1, e_y = 0$ entstanden, eine undurchquerbare Singularität (sie ist analog zu $x^2 = 0$ isoliert und dadurch undurchquerbar).

Zusammenfassend gilt daher der folgende reelle

Satz 2 (Algorithmus Polaraufblasung). *Es liege der Körper der reellen Zahlen zugrunde. Sei $g(x) = 0$ ein Gleichungssystem polynomieller Gleichungen in mehreren Unbekannten und sei 0 eine durchquerbare Singularität des Gleichungssystems mit durchquerender Kurve $x(t)$. Dann endet der folgende Algorithmus nach endlich vielen Schritten in einem lokal äquivalenten Gleichungssystem ohne Rangverlust.*

ALGORITHMUS (POLARAUFBLASUNG)

- (1) Analog zum Gaußschen Eliminationsverfahren kann man durch Linearkombination der Gleichungen erreichen, daß die linearen Anteile eine rechte obere Dreiecksmatrix bilden. Dadurch wird $t\text{-Ord}(G)$ nach Feststellung 17 nicht verändert.
Entweder besitzen danach alle Gleichungen noch einen linearen Anteil (dann hat G Vollrang und der Algorithmus terminiert) oder mindestens eine Gleichung besitzt keinen linearen Anteil. Damit sind die Voraussetzungen von Lemma 1 gegeben.
- (2) Es wird eine Polaraufblasung durchgeführt. Die durchquerende Kurve $x(t)$ wird analog Lemma 4 um $r(t), e_i(t)$ erweitert. Nach (45) und Lemma 1 sinkt die t -Ordnung.
- (3) Gilt $t\text{-Ord}(G^*) = 0$, hat G^* keinen Rangverlust mehr, der Algorithmus terminiert mit dem Ergebnis g^* .
- (4) Die Singularität wird durch eine Umentwicklung wieder in den Ursprung geschoben. Existieren Komponenten der durchquerenden Kurve, die für $t = 0$ nicht den Wert 0 annehmen, so kann nach einer Umentwicklung angenommen werden, daß dies der Fall ist. Die Umentwicklungsgleichungen werden nicht zu den anderen Gleichungen hinzugefügt (sie stören aber lediglich die Aussage, daß alle Komponenten der durchquerenden Kurve den Wert 0 annehmen).
- (5) Es wird nun $g := g^*$ ohne Transformations- und Umentwicklungsgleichungen, $x(t) := (r(t), e_i(t))$ gesetzt und wieder bei Punkt 1 begonnen.

§6 Der komplexe Fall

Das Lemma 4 war die einzige Stelle, an der die Voraussetzung $\mathbb{K} = \mathbb{R}$ benötigt wurde. Wird der Algorithmus der Aufblasung allerdings einer kleinen Modifikation unterzogen, so ist diese Voraussetzung nicht mehr nötig.

Sei im folgenden also $\mathbb{K} = \mathbb{C}$. Schwierigkeiten bereitet der Term

$$1 + \sum_{i=1}^n \frac{x_i(t)^2}{x_0(t)^2}.$$

Im Reellen ist er stets ≥ 1 , im Komplexen dagegen kann er beliebige Werte annehmen. Als einfaches Beispiel braucht man nur die durch

$$x^2 + y^2 = 0$$

definierte Varietät zu betrachten, die im Reellen undurchquerbar ist, im Komplexen dagegen durch die Kurve

$$x(t) := t, y(t) := it$$

durchquert wird. Für diese Kurve gilt

$$1 + \frac{y(t)^2}{x(t)^2} = 1 + \frac{-t^2}{t^2} = 0.$$

Für den Wert 0 oder sogar negative Werte verliert die Definition von $r(t)$ in Lemma 4 aber ihre Analytizität.

Dies kann auf folgende Weise umgangen werden: es findet vor der Polaraufblasung eine

lineare Transformation der Variablen statt, die dafür sorgt, daß anschließend in der Kurve nur noch eine einzige Komponente niedrigster t -Ordnung existiert. Dann nämlich gilt $\sum_{i=1}^n \frac{x_i(0)^2}{x_0(0)^2} = 0$ und die lineare Transformation hat nach Feststellung 15 die gleiche t -Ordnung.

Definition 16 (Vorschalttransformation für \mathbb{C}). Sei $x(t)$ eine durchquerende Kurve und $x_0(t)$ eine (von möglicherweise mehreren) Komponenten niedrigster t -Ordnung. Seien $x_1(t), \dots, x_j(t)$ die anderen Komponenten gleicher niedrigster t -Ordnung. Dann sei für $i = 1, \dots, j$

$$\alpha_i := \lim_{t \rightarrow 0} \frac{x_i(t)}{x_0(t)}.$$

Damit werde definiert:

$$\begin{aligned}\tilde{x}_0(t) &:= x_0(t) \\ \tilde{x}_i(t) &:= x_i(t) - \alpha_i x_0(t).\end{aligned}$$

Nach Anwendung der Vorschalttransformation gibt es nur noch eine Komponente niedrigster t -Ordnung und der Beweis von Lemma 4 kann auch im Komplexen angewandt werden.

Damit ergibt sich der folgende komplexe

Satz 3 (Algorithmus Polaraufblasung). Es liege der Körper der komplexen Zahlen zugrunde. Sei $g(x) = 0$ ein Gleichungssystem polynomieller Gleichungen in mehreren Unbekannten und sei 0 eine durchquerbare Singularität des Gleichungssystems mit durchquerender Kurve $x(t)$. Dann endet der folgende Algorithmus nach endlich vielen Schritten in einem lokal äquivalenten Gleichungssystem ohne Rangverlust.

ALGORITHMUS (KOMPLEXE POLARAUFBLASUNG)

- (1) Es werde die Vorschalttransformation angewandt. Dadurch wird nach Feststellung 15 die t -Ordnung nicht verändert und es gibt nur noch eine Komponente niedrigster t -Ordnung.
- (2) Analog zum Gaußschen Eliminationsverfahren kann man durch Linearkombination der Gleichungen erreichen, daß die linearen Anteile eine rechte obere Dreiecksmatrix bilden. Dadurch wird $t\text{-Ord}(G)$ nach Feststellung 17 nicht verändert.
Entweder besitzen danach alle Gleichungen noch einen linearen Anteil (dann hat G Vollrang und der Algorithmus terminiert) oder mindestens eine Gleichung besitzt keinen linearen Anteil. Damit sind die Voraussetzungen von Lemma 1 gegeben.
- (3) Es wird eine Polaraufblasung durchgeführt. Die durchquerende Kurve $x(t)$ wird analog zu Lemma 4 um $r(t), e_i(t)$ erweitert. Nach (45) und Lemma 1 sinkt die t -Ordnung.
- (4) Gilt $t\text{-Ord}(G^*) = 0$, hat G^* keinen Rangverlust mehr, der Algorithmus terminiert mit dem Ergebnis g^* .
- (5) Die Singularität wird durch eine Umentwicklung wieder in den Ursprung geschoben. Existieren Komponenten der durchquerenden Kurve, die für $t = 0$ nicht den Wert 0 annehmen, so kann nach einer Umentwicklung angenommen werden, daß dies der Fall ist. Die Umentwicklungsgleichungen werden nicht zu den anderen Gleichungen hinzugefügt.
- (6) Es wird nun $g := g^*$ ohne Transformations- und Umentwicklungsgleichungen, $x(t) := (r(t), e_i(t))$ gesetzt und wieder bei Punkt 1 begonnen.

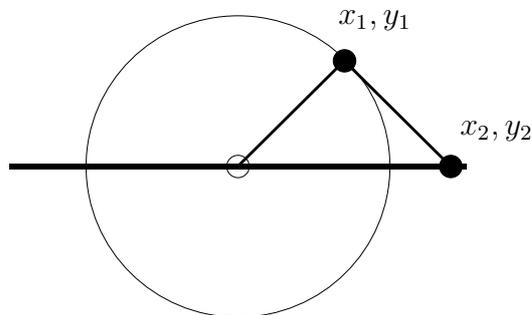
KAPITEL IV

Beispiele für Aufblasung

Hier soll nun das im vorigen Kapitel entwickelte Verfahren der Kurven- und Polaraufblasung auf einige Beispiele angewandt werden. Am entarteten Kurbeltrieb wird die Funktionsweise noch einmal ausgeführt, am Beispiel der liegenden Acht wird auch numerisch die Durchquerung der Singularität demonstriert. Schließlich wird noch kurz die numerische Stabilität des Verfahrens angedeutet.

§1 Der Kurbeltrieb

Das Verfahren soll nun am Beispiel des (entarteten) Kurbeltriebs vorgeführt werden. Der Kurbeltrieb bestehe aus zwei gleich langen Stangen (die gleiche Länge ist für die Entartung verantwortlich), von denen eine im Ursprung befestigt ist und um diesen kreisen kann. Die zweite soll an die erste anschließen, ihr Ende soll aber nur auf der x -Achse rutschen können. Ein solcher Mechanismus dient i. allg. dazu, eine kreisförmige Bewegung in eine horizontale „Pump“-Bewegung umzuwandeln.



Kurbeltrieb

Das Ende der ersten Stange werde durch die Koordinaten x_1, y_1 beschrieben, das Ende der zweiten durch x_2, y_2 . Die Längen der Stangen seien 1. Aus der Beschreibung des Mechanismus ergeben sich ganz zwanglos die drei Zwangsbedingungen¹

$$\begin{aligned}x_1^2 + y_1^2 - 1 &= 0 \\(x_1 - x_2)^2 + (y_1 - y_2)^2 - 1 &= 0 \\y_2 &= 0.\end{aligned}$$

¹Diese sich ergebenden Zwangsbedingungen werden allerdings mit einer Singularität behaftet sein. Es gibt eine bessere Formulierung des „gleichen“ Mechanismus ohne diese Singularität; der Algorithmus wird sie finden.

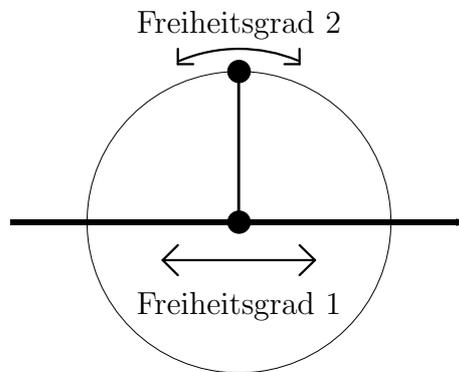
Die Ableitungsmatrix lautet

$$G = \begin{pmatrix} 2x_1 & 2y_1 & 0 & 0 \\ 2(x_1 - x_2) & 2(y_1 - y_2) & -2(x_1 - x_2) & -2(y_1 - y_2) \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die Singularität liegt im Punkt $(0, 1, 0, 0)$ vor:

$$G(0, 1, 0, 0) = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die Ursache für das Auftreten einer Singularität an dieser Stelle ist das Zusammentreffen zweier Untermannigfaltigkeiten. Der Mechanismus kann sich auf zwei Arten aus dieser Stellung herausbewegen: entweder rutscht das Ende der zweiten Stange x_2, y_2 auf der x -Achse weiter oder das Ende der zweiten Stange verharrt im Ursprung und das Ende der ersten Stange rotiert kreisförmig um den Ursprung.



Zwei Freiheitsgrade

Eine durchquerende Kurve ist gegeben durch

$$p(t) = \begin{pmatrix} \frac{t}{2} \\ \sqrt{1 - \frac{t^2}{4}} \\ t \\ 0 \end{pmatrix}.$$

Zunächst einmal werde die Singularität in den Nullpunkt transformiert, dazu seien neue Variablen q_i eingeführt:

$$\begin{aligned} q_1 &= x_1 \\ q_2 &= y_1 - 1 \\ q_3 &= x_2 \\ q_4 &= y_2. \end{aligned}$$

In diesen neuen Koordinaten lauten die Zwangsbedingungen nun

$$\begin{aligned} q_1^2 + q_2^2 + 2q_2 &= 0 \\ q_1^2 - 2q_1q_3 + q_3^2 + q_2^2 + 2q_2 - 2q_2q_4 - 2q_4 + q_4^2 &= 0 \\ q_4 &= 0 \end{aligned} \tag{46}$$

und die durchquernde Kurve lautet

$$q(t) = \begin{pmatrix} \frac{t}{2} \\ \sqrt{1 - \frac{t^2}{4}} - 1 \\ t \\ 0 \end{pmatrix}.$$

Die linearen Anteile der Gleichungen werden auf rechte obere Dreiecksgestalt gebracht (der Einfachheit halber in der Reihenfolge q_2, q_4, \dots):

$$\begin{aligned} q_1^2 + q_2^2 + 2q_2 &= 0 \\ q_4 &= 0 \\ -2q_1q_3 + q_3^2 - 2q_2q_4 + q_4^2 &= 0. \end{aligned}$$

Nun wird aufgrund der t -Ordnungen von $q(t)$ bzgl. q_1 aufgeblasen:

$$\begin{aligned} q_2 &= \lambda_2 q_1 \\ q_3 &= \lambda_3 q_1 \\ q_4 &= \lambda_4 q_1. \end{aligned}$$

Diese Transformation führt zu

$$\begin{aligned} q_1^2 + \lambda_2^2 q_1^2 + 2\lambda_2 q_1 &= 0 \\ \lambda_4 q_1 &= 0 \\ -2\lambda_3 q_1^2 + \lambda_3^2 q_1^2 - 2\lambda_2 \lambda_4 q_1^2 + \lambda_4^2 q_1^2 &= 0. \end{aligned}$$

Die nachfolgende Abdivision führt zu

$$\begin{aligned} q_1 + \lambda_2^2 q_1 + 2\lambda_2 &= 0 \\ \lambda_4 &= 0 \\ -2\lambda_3 + \lambda_3^2 - 2\lambda_2 \lambda_4 + \lambda_4^2 &= 0. \end{aligned}$$

Das ganze System hat nun keinen Rangverlust mehr. Dies ist auch die versprochene einfachere Formulierung des Systems: die letzte Zeile des Gleichungssystems ist unter Berücksichtigung von $\lambda_4 = 0$ äquivalent zu

$$\lambda_3(\lambda_3 - 2) = 0.$$

Dies hat zwei Lösungszweige, die den beiden sich treffenden Mannigfaltigkeiten entsprechen, nämlich $\lambda_3 = 0$ (der zweite Körper verharrt immer im Ursprung) oder $\lambda_3 = 2$ (die x -Koordinate des zweiten Körpers ist immer doppelt so groß wie die x -Koordinate des ersten Körpers). Wählte man nun nur einen dieser beiden Zweige als intendierte Modellierung aus, so wäre das Gleichungssystem von Anfang an ohne Rangverlust gewesen. In der obigen Formulierung sind beide Möglichkeiten immer noch enthalten, aber separiert worden ($\lambda_3 \equiv 0$ oder $\lambda_3 \equiv 2$).

Die Polaraufblasung verfährt gleich bis zum Zwischenstand

$$\begin{aligned} q_1^2 + q_2^2 + 2q_2 &= 0 \\ q_4 &= 0 \\ -2q_1q_3 + q_3^2 - 2q_2q_4 + q_4^2 &= 0. \end{aligned}$$

Hier nun werden die Transformationsgleichungen

$$\begin{aligned}
q_1 - re_{q_1} &= 0 \\
q_2 - re_{q_2} &= 0 \\
q_3 - re_{q_3} &= 0 \\
q_4 - re_{q_4} &= 0 \\
e_{q_1}^2 + e_{q_2}^2 + e_{q_3}^2 + e_{q_4}^2 - 1 &= 0
\end{aligned}$$

angewandt:

$$\begin{aligned}
r^2 e_{q_1}^2 + r^2 e_{q_2}^2 + 2re_{q_2} &= 0 \\
re_{q_4} &= 0 \\
-2r^2 e_{q_1} e_{q_3} + r^2 e_{q_3}^2 - 2r^2 e_{q_2} e_{q_4} + r^2 e_{q_4}^2 &= 0 \\
e_{q_1}^2 + e_{q_2}^2 + e_{q_3}^2 + e_{q_4}^2 - 1 &= 0.
\end{aligned}$$

Die Abdivision ergibt hier

$$\begin{aligned}
re_{q_1}^2 + re_{q_2}^2 + 2e_{q_2} &= 0 \\
e_{q_4} &= 0 \\
-2e_{q_1} e_{q_3} + e_{q_3}^2 - 2e_{q_2} e_{q_4} + e_{q_4}^2 &= 0 \\
e_{q_1}^2 + e_{q_2}^2 + e_{q_3}^2 + e_{q_4}^2 - 1 &= 0.
\end{aligned}$$

Um den Vollrang zu erkennen, muß man hier in der dritten und der vierten Zeile nur die Ableitungen nach e_{q_1} und e_{q_3} betrachten. Indem man die durchquerenden Kurven ermittelt, erkennt man $e_{q_1}(0) \neq 0$, $e_{q_3}(0) \neq 0$ und $e_{q_3}(t) = 2e_{q_1}(t)$.

Somit erhält man für die Ableitungsmatrix der einzigen beiden kritischen Gleichungen bzgl. e_{q_1} und e_{q_3} das Ergebnis

$$\begin{pmatrix} -4e_{q_1} & 2e_{q_1} \\ 2e_{q_1} & 4e_{q_1} \end{pmatrix}.$$

Damit liegt Vollrang vor. Wieder kann man an der dritten Gleichung die beiden sich treffenden Zweige $e_{q_3} = 0$ und $e_{q_3} = 2e_{q_1}$ erkennen.

§2 Und nun ... Physik

Trotz der Erfolge, mathematisch die Formulierung der Zwangsmannigfaltigkeit so umzugestalten, daß diese keinen Rangverlust mehr hat, kommt es zu einem anderen Problem. Die Beseitigung des Rangverlusts der Mannigfaltigkeit funktioniert mathematisch immer, auch dann, wenn es physikalisch nicht funktionieren darf. Dies sei an einem einfachen Beispiel illustriert:

$$g(x, y) := y^2 - x^3.$$

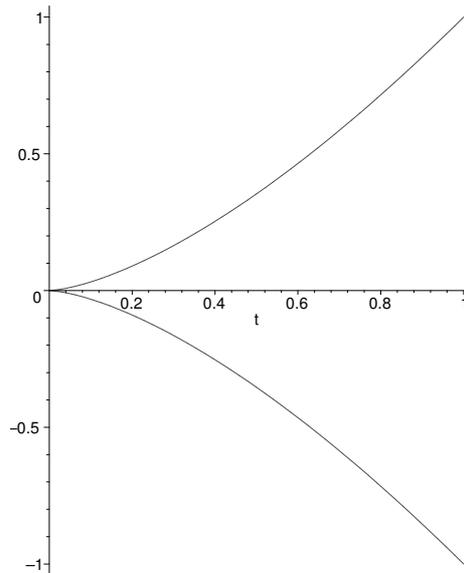
Ausgangspunkt sei folgendes Gedankenexperiment: ein Massepunkt der Masse M werde an einem Punkt der Kurve mit $y > 0$ plaziert und falle unter Einfluß von Schwerkraft nach unten.

Nach endlicher Zeit t_1 erreicht die Masse mit einer nichtverschwindenden Geschwindigkeit

$$v_x(t_1) < 0$$

den Ursprung. Was passiert nun aber mit den Geschwindigkeit v_x ? Damit der Punkt weiter auf der Kurve nach unten gleiten könnte, müssten sie einen Sprung erleiden (etwa im Sinne eines inelastischen Stoßes) zu

$$v_x(t_1 + \varepsilon) \geq 0 \text{ für alle hinreichend kleinen } \varepsilon > 0.$$

ABBILDUNG IV.1. Die Kurve $y^2 = x^3$

Aber wenn die Geschwindigkeit v_x zum Zeitpunkt t_1 springt, gibt es keine stetig differenzierbare Lösung des differential-algebraischen Gleichungssystems. Jedoch haben die Methoden des vorherigen Abschnitts Möglichkeiten geliefert, die Singularität der Kurve zu beseitigen. Und dennoch soll es keine glatte Lösung geben?

Die Antwort auf diesen scheinbaren Widerspruch liegt in der Natur der hinzugefügten Variablen. Die Kurve

$$y^2 - x^3 = 0$$

wird durch Einführung eines zusätzlichen Parameters s (nicht zu Verwechseln mit der Zeit oder der Bogenlänge) beseitigt:

$$\begin{aligned} y - s^3 &= 0 \\ x - s^2 &= 0. \end{aligned}$$

Diese zusätzliche Variable s , die „projektive Variable“, taucht nun in der DAE des physikalischen Systems auf, und somit stellt sich die Frage, welche „Masse“ dieser Variablen in der Massenmatrix zuzuordnen ist. Beschreibt man das physikalische System durch den Lagrange-Formalismus, so kommt die Variable s in den Lagrange-Gleichungen nicht vor, denn sowohl Potential als auch kinetische Energie sind nur von y bzw. x und y abhängig. Bezüglich der Variable s ergibt sich eine Nullgleichung ($0 = 0$). Insbesondere die Tatsache, daß die kinetische Energie nicht von s abhängig ist, legt nahe, der Variablen s eine Nullmasse zuzordnen:

$$m_s := 0.$$

Dies hat nun die Auswirkung, daß im Gleichungssystem 4 zwar die Ableitungsmatrix G immer Vollrang hat, dafür aber die Massenmatrix M nicht mehr. Durch diese Hintertür kehren die Nicht-Differenzierbarkeiten in den Lösungen der DAE wieder zurück.

Trotzdem sind die vorgestellten Methoden zur Singularitätsbeseitigung damit nicht sinnlos:

- (1) Wie das Beispiel

$$g(x, y) = xy$$

zeigt, gibt es Fälle, in denen, trotz der Nullzeilen der Massenmatrix, das System als Ganzes nun immer Vollrang erhält. Dieses Beispiel soll etwas weiter ausgeführt werden. Mit der Zwangsbedingung $g(x, y) = xy$ und den Massen m_x und m_y lautet die Gesamtmatrix der Gleichung 4

$$\begin{pmatrix} M & G^T \\ G & 0 \end{pmatrix} = \begin{pmatrix} m_x & 0 & y \\ 0 & m_y & x \\ y & x & 0 \end{pmatrix},$$

welches in $x = 0, y = 0$ Rangverlust hat.

Nach der Aufblasung lauten die Zwangsbedingungen

$$\begin{aligned} x - r e_x &= 0 \\ y - r e_y &= 0 \\ e_x e_y &= 0 \\ e_x^2 + e_y^2 - 1 &= 0. \end{aligned}$$

Damit ergibt sich als Gesamtmatrix (mit Nullmassen der neuen Variablen und Variablenreihenfolge x, y, r, e_x, e_y):

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -e_x & -e_y & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -r & 0 & e_y & 2e_x \\ 0 & 0 & 0 & 0 & 0 & 0 & -r & e_x & 2e_y \\ 1 & 0 & -e_x & -r & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -e_y & 0 & -r & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & e_y & e_x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2e_x & 2e_y & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Singularität $x = 0, y = 0$ wurde durch die Aufblasung in vier Punkte getrennt: $r = 0, |e_x| = 1, e_y = 0$ und $r = 0, e_x = 0, |e_y| = 1$. Im ersten Fall lautet die Gesamtmatrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

oder

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

im zweiten Fall lautet die Matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \end{pmatrix}$$

oder

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In jedem dieser vier Fälle hat die Matrix Vollrang.

- (2) Das mathematische Verfahren der Singularitätsbeseitigung funktioniert immer, sogar in solchen Fällen, in denen aufgrund physikalischer Überlegungen keine vollständig glatte Beschreibung des Systems existieren kann, wie dies z. B. bei

$$g(x, y) = y^2 - x^3$$

der Fall ist. Hier kann nicht einerseits durch einen mathematischen Trick eine glatte Lösung erzwungen werden und gleichzeitig die mathematische Lösung eine physikalisch sinnvolle sein, wenn die Physik die Existenz einer solchen Lösung ausschließt.

Zusammen legt dies folgende Vermutung nahe: immer wenn eine physikalisch sinnvolle Lösung existiert, liefert das mathematische Verfahren der Aufblasung eine Umformulierung der Gleichung (4), die sogar unter Annahme von Nullmassen für die neu hinzugekommenen Variablen Vollrang besitzt.

Ein weiteres Beispiel funktionierender Aufblasung und Integration wird im folgenden Abschnitt vorgestellt.

§3 Die liegende Acht

Ein Beispiel für das Gelingen der Aufblasung und numerischer Integration ist die Zwangsbedingung, daß ein Körper nur entlang einer liegenden Acht gleiten darf. Die entsprechende Zwangsbedingungsgleichung lautet

$$y^2 - x^2(1 - x^2) = 0.$$

Das Aussehen dieser Zwangsbedingung kann Abbildung IV.2 entnommen werden. Die

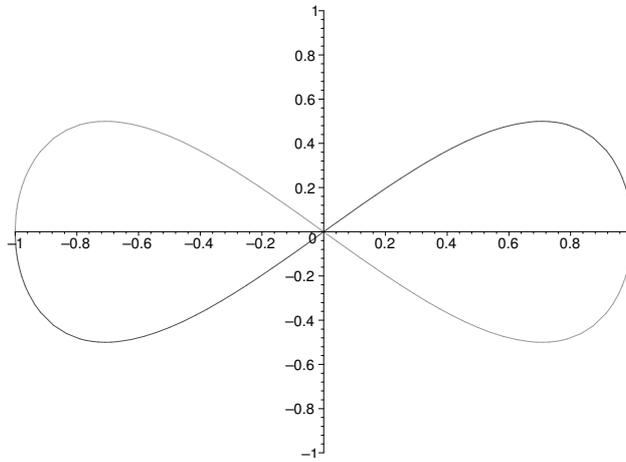


ABBILDUNG IV.2. Die Kurve $y^2 - x^2(1 - x^2) = 0$

ursprüngliche Gleichung

$$x^4 - x^2 + y^2 = 0$$

geht durch die Aufblasung über in

$$\begin{aligned} x - re_x &= 0 \\ y - re_y &= 0 \\ e_x^2 + e_y^2 - 1 &= 0 \\ r^2 e_x^4 - e_x^2 + e_y^2 &= 0. \end{aligned}$$

Dieses Gleichungssystem ist auf der Mannigfaltigkeit ohne Rangverlust. Mit den Werten

$$\begin{aligned} m_x &= 1, F_x = 0 \\ m_y &= 1, F_y = -g = -9.81 \\ m_r &= 0, F_r = 0 \\ m_{e_x} &= 0, F_{e_x} = 0 \\ m_{e_y} &= 0, F_{e_y} = 0 \end{aligned}$$

sowie den Anfangswerten

$$\begin{aligned}
x(0) &= 1, \dot{x}(0) = 0 \\
y(0) &= 0, \dot{y}(0) = 4 \\
r(0) &= 1, \dot{r}(0) = 0 \\
e_x(0) &= 1, \dot{e}_x(0) = 0 \\
e_y(0) &= 0, \dot{e}_y(0) = 4
\end{aligned}$$

wird die DAE (7) für den Zeitraum von 0 bis 2 sec integriert. Der Integrator ist dabei der Dormand-Prince85-Integrator von MBSSIM. Er benötigt 0.08 sec auf einem Pentium-II-266 MHz-Rechner und macht in der Integration 85 Schritte bzw. 1190 Funktionsauswertungen. Das Ergebnis ist in den Abbildungen IV.3 bis IV.8 wiedergegeben.

§4 Numerische Stabilität

In diesem Abschnitt soll, wenn auch nur angedeutet, untersucht werden, wie sich der Aufblasungsalgorithmus gegenüber numerischen Störungen verhält. Dies soll am Beispiel des Koordinatenkreuzes $g(x, y) := xy = 0$ verdeutlicht werden. Dazu wird die Gleichung des Koordinatenkreuzes

$$xy = 0 \tag{47}$$

einer gestörten Gleichung

$$xy = \varepsilon \tag{48}$$

gegenübergestellt.

Der Aufblasungsalgorithmus ist ein exakter Koordinaten-Transformations-Algorithmus. Er führt die qualitativ unterschiedlichen Varietäten, beschrieben durch (47) und (48), über in andere, äquivalente Darstellungen. Daher zeigt der Algorithmus beim Übergang von (47) zu (48) ein sprunghaftes Verhalten, ist also in diesem Sinne nicht numerisch stabil.

Was bedeutet es in diesem Zusammenhang, numerische Stabilität zu fordern? Betrachtet man (48) als Störung von (47), so soll die Aufblasung von (48) ein qualitativ ähnliches Verhalten zeigen wie die von (47). Dies bedeutet anschaulich folgendes: angenommen ein Massepunkt würde auf einem Hyperbelzweig von (48) sich dem Ursprung nähern. Er sollte dann in Ursprungsnähe nicht „um die Kurve biegen“, sondern auf den anderen Hyperbelast springen, so daß es bei kleinem ε und „aus größerer Entfernung“ wirkt, als ob er sich der Achse entlang bewegt hätte.

Dieses Verhalten lässt sich durch eine Modifikation des Aufblasungsalgorithmus erreichen. Während der exakte Aufblasungsalgorithmus in (48) das ε berücksichtigt, besteht die Modifikation darin, daß der Algorithmus konstante (und lineare) Terme als 0 betrachtet, wenn sie kleiner sind als eine vorgegebene Toleranzgrenze.

Am Beispiel (48) bedeutet dies, daß die Gleichung

$$xy = \varepsilon$$

aufgeblasen wird zu

$$\begin{aligned}
x - re_x &= 0 \\
y - re_y &= 0 \\
e_x^2 + e_y^2 - 1 &= 0 \\
e_x e_y &= \varepsilon.
\end{aligned}$$

Die möglichen Lösungen (x, y) bilden hier ein leicht gedrehtes Koordinatenkreuz (Abbildung IV.9). In der praktischen Durchführung kann man sich dies wie folgt vorstellen: der Massepunkt nähert sich in der ursprünglichen Formulierung dem Nullpunkt. Die Determinante von $G^T G$ sinkt schließlich unter einen kritischen Wert, wodurch zur aufgeblasenen Formulierung übergegangen wird. In dieser Formulierung durchquert der Massepunkt den Ursprung. Nachdem die Determinante der ursprünglichen Formulierung die kritische Grenze wieder übersteigt, kann zu ihr zurückgekehrt werden. Aus der „Ferne“ betrachtet wirkt dies, als hätte der Massepunkt die y -Achse im Ursprung durchquert.

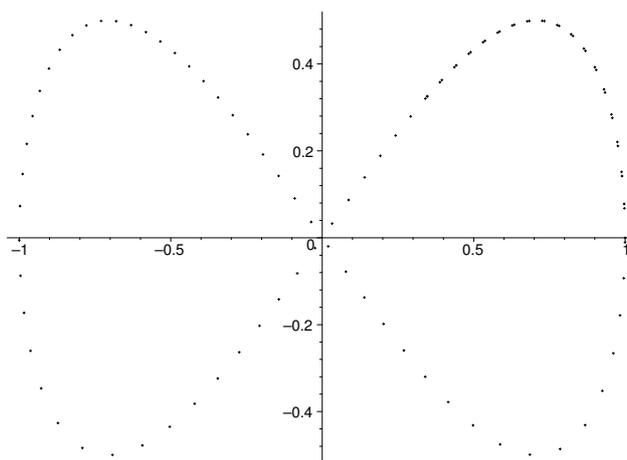


ABBILDUNG IV.3. Position in der $x - y$ -Ebene

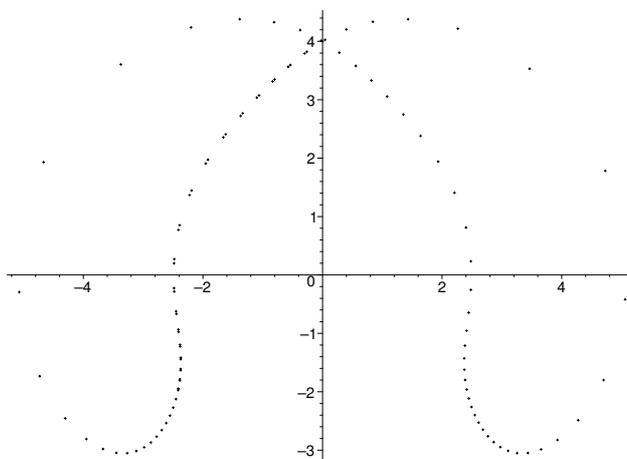


ABBILDUNG IV.4. Geschwindigkeit in der $x - y$ -Ebene

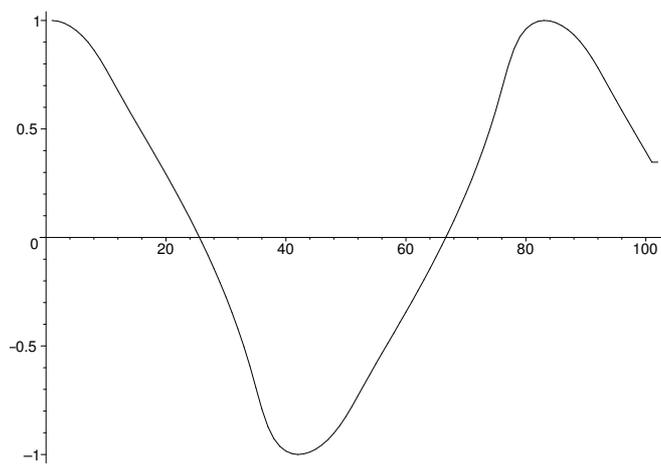
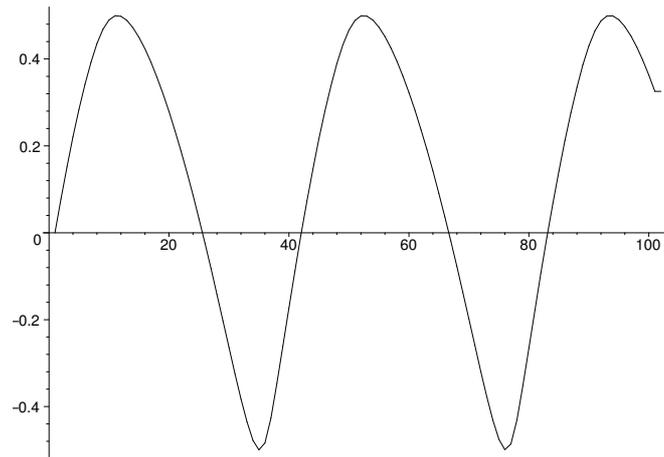
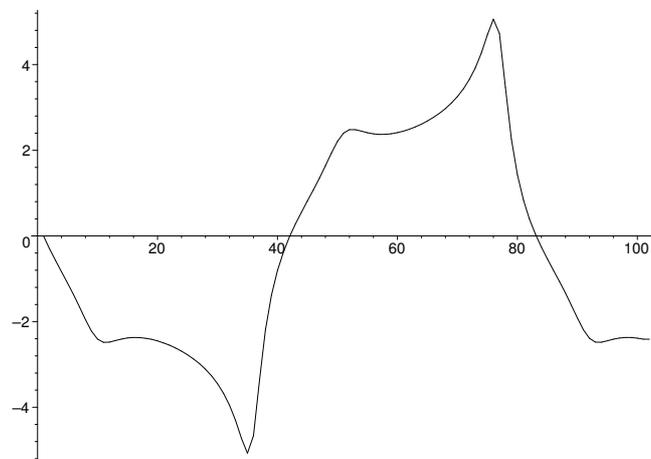
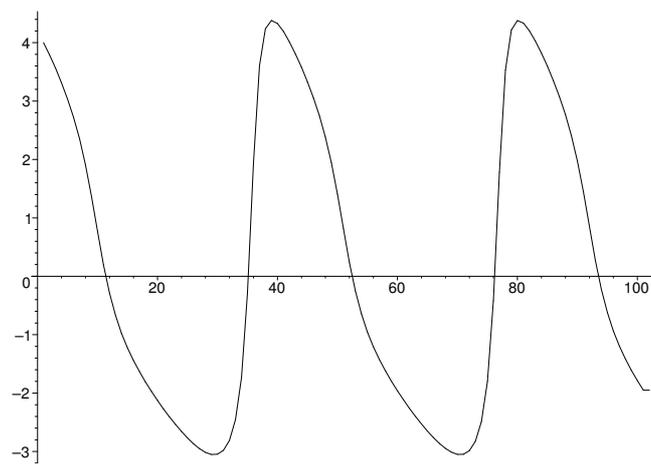
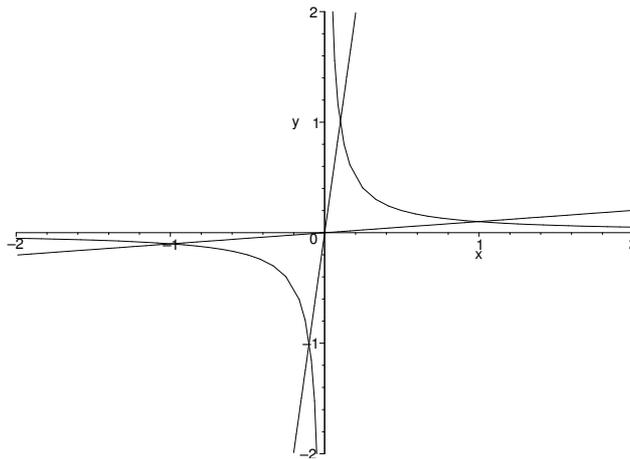


ABBILDUNG IV.5. x -Position über der Zeit

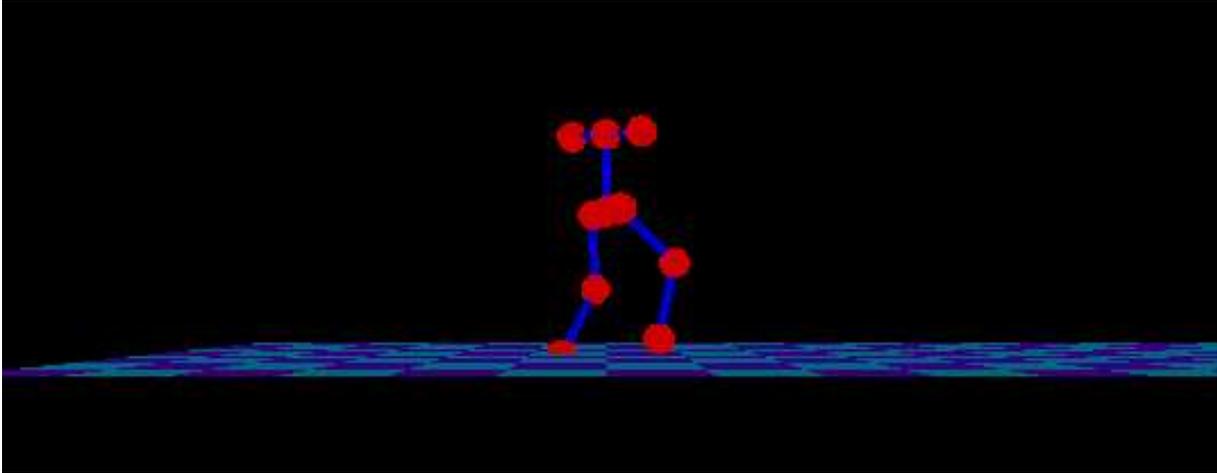
IV. BEISPIELE FÜR AUFBLASUNG

ABBILDUNG IV.6. y -Position über der ZeitABBILDUNG IV.7. x -Geschwindigkeit über der ZeitABBILDUNG IV.8. y -Geschwindigkeit über der Zeit

ABBILDUNG IV.9. Modifizierte Aufblasung für $xy = 0.1$

KAPITEL V

Das Modell



Hier wird ein Modell für ein Mehrkörpersystem aufgestellt, für welches dann anschließend optimale Laufbewegungen bestimmt werden sollen. Dabei ergeben sich aus dem Entwurf die mechanischen Zwangsbedingungen sowie wirkende Kräfte und Drehmomente. Bestandteil des Modells sind die Bodenreaktionskräfte (und sich dadurch ergebende Schaltfunktionen), ohne die keine Laufbewegung möglich wäre.

Die möglichen Bewegungen werden dann durch weitere Bedingungen auf Optimierungsseite eingeschränkt (z. B. sollen die Knie höher als die Füße bleiben).

Des Weiteren werden die Gleichungen genannt, die hinzugefügt werden müssen, wenn Periodizität verlangt werden sollte.

§1 Die mechanische Formulierung

§1.1 Punktmassen. Das Modell besteht aus 10 punktförmigen Massen (im Bild¹ als rote Kugeln dargestellt), denen jeweils die Masse 1 zugeordnet wird. Ihre Numerierung ist in Abbildung V.1 gezeigt. Wegen der Punktförmigkeit der Massen und der Massenlosigkeit aller weiteren Elemente (Gelenke, Aktuatoren) reduzieren sich die natürlichen auf gewöhnliche kartesische Koordinaten. Aus den 10 Massepunkten ergeben sich in drei Dimensionen dreißig Orts- und dreißig Geschwindigkeitskoordinaten. Die Koordinaten werden auf die dreißig Ortsvariablen wie folgt abgebildet:

$$x\text{-Koordinate von Massepunkt Nr. } i \rightarrow 3(i - 1) + 1$$

$$y\text{-Koordinate von Massepunkt Nr. } i \rightarrow 3(i - 1) + 2$$

$$z\text{-Koordinate von Massepunkt Nr. } i \rightarrow 3(i - 1) + 3.$$

¹Die folgenden Visualisierungen wurden mit Hilfe von JAFV erstellt, einem auf OpenGL basierenden Graphikprogramm von MICHAEL WINCKLER und JÖRG HUBER[69].

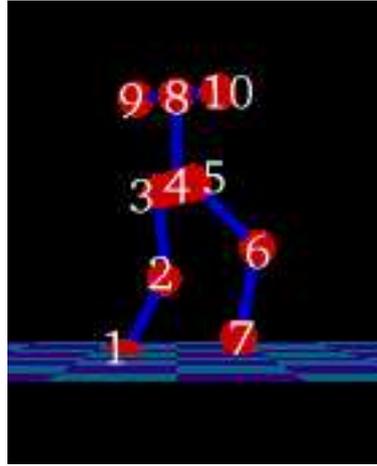


ABBILDUNG V.1. Numerierung der Massepunkte

§1.2 Zwangsbedingungen. Die Bewegungsmöglichkeiten der einzelnen Massepunkte werden durch 19 mechanische Zwangsbedingungen eingeschränkt. Die ersten 13 sind gemäß Abbildung V.2 einfach zu erkennen.

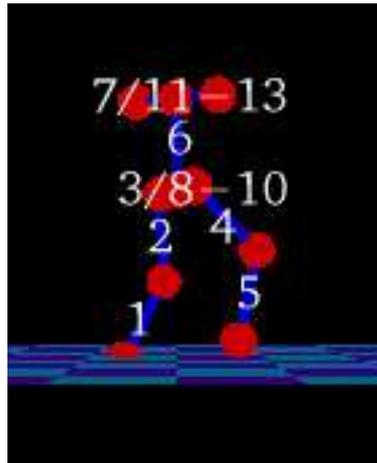


ABBILDUNG V.2. Die ersten 13 Zwangsbedingungen

Die Bedingungen 1 bis 7 beschreiben, daß sich die Massepunkte in einem fixen Abstand voneinander befinden:

$$\begin{aligned}
 gpos_1 &= (p_1 - p_4)^2 + (p_2 - p_5)^2 + (p_3 - p_6)^2 - 1 \\
 gpos_2 &= (p_4 - p_7)^2 + (p_5 - p_8)^2 + (p_6 - p_9)^2 - 1 \\
 gpos_3 &= (p_7 - p_{13})^2 + (p_8 - p_{14})^2 + (p_9 - p_{15})^2 - 1 \\
 gpos_4 &= (p_{13} - p_{16})^2 + (p_{14} - p_{17})^2 + (p_{15} - p_{18})^2 - 1 \\
 gpos_5 &= (p_{16} - p_{19})^2 + (p_{17} - p_{20})^2 + (p_{18} - p_{21})^2 - 1 \\
 gpos_6 &= (p_{10} - p_{22})^2 + (p_{11} - p_{23})^2 + (p_{12} - p_{24})^2 - 1 \\
 gpos_7 &= (p_{25} - p_{28})^2 + (p_{26} - p_{29})^2 + (p_{27} - p_{30})^2 - 1.
 \end{aligned}$$

Die Bedingungen 8 bis 13 beschreiben, daß die jeweiligen Massen sich immer in der Mitte zweier anderer Massen befinden müssen:

$$\begin{aligned} gpos_8 &= p_7 + p_{13} - 2p_{10} \\ gpos_9 &= p_8 + p_{14} - 2p_{11} \\ gpos_{10} &= p_9 + p_{15} - 2p_{12} \\ gpos_{11} &= p_{25} + p_{28} - 2p_{22} \\ gpos_{12} &= p_{26} + p_{29} - 2p_{23} \\ gpos_{13} &= p_{27} + p_{30} - 2p_{24}. \end{aligned}$$

Zwangsbedingung 14 besagt, daß die Achse zwischen Körper 1 und 2 senkrecht auf der Achse zwischen Körper 3 und 5 sein muß; damit wird ein seitliches Wegknicken des Beins verboten. Die Zwangsbedingungen 15 bis 17 fordern dies analog für die Achsen zwischen 2 und 3, zwischen 5 und 6 und zwischen 6 und 7.

$$\begin{aligned} gpos_{14} &= (p_1 - p_4)(p_7 - p_{13}) + (p_2 - p_5)(p_8 - p_{14}) + (p_3 - p_6)(p_9 - p_{15}) \\ gpos_{15} &= (p_4 - p_7)(p_7 - p_{13}) + (p_5 - p_8)(p_8 - p_{14}) + (p_6 - p_9)(p_9 - p_{15}) \\ gpos_{16} &= (p_{19} - p_{16})(p_7 - p_{13}) + (p_{20} - p_{17})(p_8 - p_{14}) + (p_{21} - p_{18})(p_9 - p_{15}) \\ gpos_{17} &= (p_{16} - p_{13})(p_7 - p_{13}) + (p_{17} - p_{14})(p_8 - p_{14}) + (p_{18} - p_{15})(p_9 - p_{15}). \end{aligned}$$

Die Bedingung 18 fordert, daß die Achse zwischen Körper 4 und 8 ebenfalls senkrecht auf der Achse zwischen Körper 3 und 5 steht („Wirbelsäule“ steht senkrecht auf „Hüfte“):

$$gpos_{18} = (p_{10} - p_{22})(p_7 - p_{13}) + (p_{11} - p_{23})(p_8 - p_{14}) + (p_{12} - p_{24})(p_9 - p_{15}).$$

Bedingung 19 nun fordert, daß die Achse zwischen den Körpern 9 und 10 wiederum senkrecht auf der Achse zwischen 4 und 8 steht („Schulter“ steht senkrecht auf „Wirbelsäule“):

$$gpos_{19} = (p_{25} - p_{28})(p_{10} - p_{22}) + (p_{26} - p_{29})(p_{11} - p_{23}) + (p_{27} - p_{30})(p_{12} - p_{24}).$$

§1.3 Kräfte. Auf das Modell wirken nun äußere Kräfte (Gravitation, Reibung mit dem Boden) sowie innere Kräfte (Drehmomente durch Aktuatoren). Durch letztere wird die Bewegung gesteuert. Seien zunächst die wirkenden Drehmomente beschrieben. Als lokales Koordinatensystem dient dabei eine Basis, die aus der „Hüfte“ (Achse von Körper 3 nach Körper 5), der „Wirbelsäule“ (Achse von Körper 4 nach 8) und einer Basisergänzung gebildet wird, wie in Abbildung V.3 gezeigt. Der Aktuator für das Drehmoment q_1 sitzt

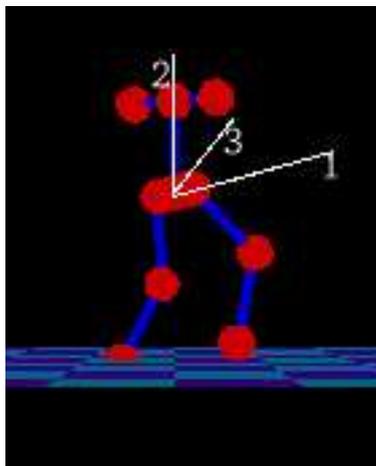


ABBILDUNG V.3. Lokale Basis der Drehmomente

sozusagen in Körper 2 und bewirkt ein Drehmoment auf die Achse zwischen Körper 1 und 2 (auf der Abbildung V.4 der gelbe, nach rechts zeigende Pfeil) und ein Gegendrehmoment auf die Achse zwischen Körper 2 und 3 (der grüne, nach links zeigende Pfeil). Ein analog wirkendes Drehmoment q_2 sitzt in Körper 6 und wirkt mit einem Moment auf

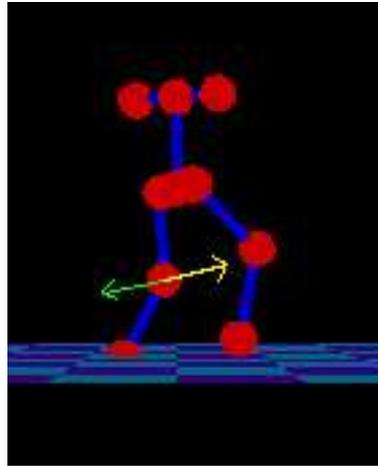


ABBILDUNG V.4. Das Drehmoment q_1

die Achse zwischen Körper 6 und 7 und mit einem Gegenmoment auf die Achse zwischen Körper 6 und 5. Die beiden Drehmomente q_1 und q_2 könnte man als Kniebeuger und -strecker bezeichnen.

Ein weiteres Drehmoment q_3 sitzt in Körper 3 und wirkt auf die Achse zwischen Körper 2 und 3 (in Abbildung V.5 der gelbe, nach rechts zeigende Pfeil) mit einem Gegenmoment auf die Achse zwischen Körper 5 und 6 (der grüne, nach links zeigende Pfeil). Mit seiner

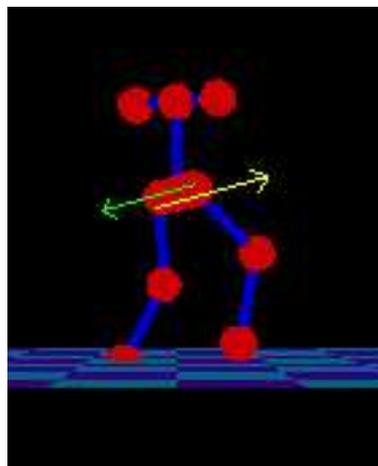


ABBILDUNG V.5. Das Drehmoment q_3

Hilfe können die beiden „Oberschenkel“ (Achse Körper 2 und 3 und Achse Körper 5 und 6) bewegt werden und zwar in entgegengesetzten Richtungen (aufeinander zu oder voneinander weg, nicht aber gemeinsam in eine Richtung).

Das Drehmoment q_4 vermag nun, beide „Oberschenkel“ in die gleiche Richtung zu bewegen mit einer Gegenbewegung der „Wirbelsäule“. Es wirkt auf die Achse zwischen Körper 2 und 3 sowie in gleichem Sinne auf die Achse zwischen Körper 5 und 6 (in Abbildung V.6

die beiden grünen, nach rechts zeigenden Pfeile) und doppelt so stark in entgegengesetzter Richtung auf die Achse zwischen Körper 4 und 8 (der gelbe, nach links zeigende Pfeil). Mit

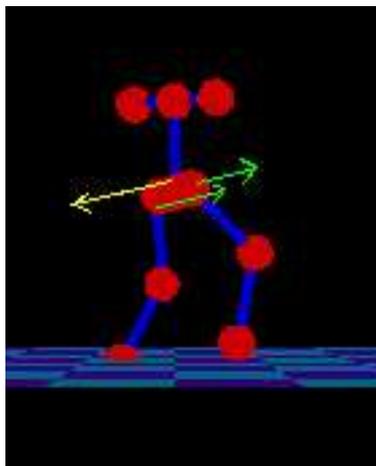


ABBILDUNG V.6. Das Drehmoment q_4

Hilfe dieses Drehmoments ist das Modell in der Lage, eine Art „Klappmesser-Bewegung“ durchzuführen.

Das letzte Drehmoment, q_5 , erlaubt nun eine „Schulterdrehung“. Es bewirkt ein Drehmoment auf die Achse zwischen Körper 9 und 10 (in Abbildung V.7 der gelbe, nach oben zeigende Pfeil) und ein Gegenmoment auf die Achse zwischen Körper 3 und 5 (der grüne, nach unten zeigende Pfeil).

Nach diesen inneren Kräften sind nun die auf das Modell wirkenden äußeren Kräfte zu

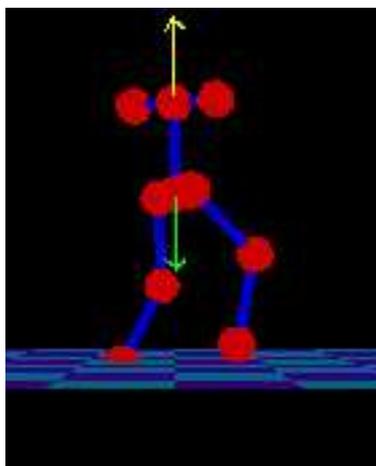


ABBILDUNG V.7. Das Drehmoment q_5

beschreiben. Zunächst einmal wirkt die Gravitation auf alle Körper gleichmäßig. Die Formulierung dieser Kraft ist offensichtlich.

Des weiteren wirken auf die Körper 1 und 7 (und nur auf diese) Bodenreaktionskräfte. Dringt einer dieser Körper in den als weich gedachten Boden ein, so erfährt er eine rücktreibende Kraft. Seien mit p_x, p_y, p_z und v_x, v_y, v_z die Orts- bzw. Geschwindigkeitskoordinaten

des eindringenden Körpers bezeichnet. Dann wird als rücktreibende Kraft gesetzt:

$$\begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{cases} \begin{pmatrix} C_{x\text{-Reibung}} v_x p_z^3 \\ C_{y\text{-Reibung}} v_y p_z^3 \\ C_{z\text{-Reibung}} v_z p_z^3 - D p_z^3 \end{pmatrix} & \text{für } p_z < 0 \\ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} & \text{sonst.} \end{cases}$$

In x - und y -Richtung wirken Reibungskräfte, die proportional zur Geschwindigkeit in dieser Richtung sind und dabei mit der dritten Potenz der Eindringtiefe multipliziert werden. In z -Richtung wirkt ebenfalls ein solcher Reibungsterm, zusätzlich aber noch eine rücktreibende Kraft, die proportional zu Eindringtiefe hoch drei gesetzt wird.

Die dritte Potenz der Eindringtiefe ergibt sich aus der Tatsache, daß es für den Optimierungsvorgang eine notwendige Voraussetzung ist, daß das Zielfunktional (über die Trajektorien) zweimal stetig differenzierbar von den Eingangsparametern abhängt. Beim Übergang von „Körper hat z -Koordinate > 0 “ zu „Körper hat z -Koordinate ≤ 0 “ findet aber mit dem Auftreten von Rücktriebs- und Reibungskräften ein Sprung in den Kräften statt. Durch die Glättung dieses Sprungs mit der dritten Potenz der Eindringtiefe geht die zweimalige Differenzierbarkeit nicht verloren.

In den numerischen Berechnungen werden die Werte

$$\begin{aligned} C_{x\text{-Reibung}} &= C_{y\text{-Reibung}} = 10^6 \\ C_{z\text{-Reibung}} &= 8 \cdot 10^4 \\ D &= 100 \end{aligned}$$

angenommen. In ihren Größenordnungen sind die Werte so gewählt, daß der Aufprall eines Körpers auf dem Boden der Schwingung eines harmonischen Oszillators im aperiodischen Grenzfall nahekommt, damit die auftretenden Schwingungen möglichst schnell abklingen. Exakt erreichbar ist dieser Wunschfall nicht, da ja z. B. nicht klar ist, ob das Modell mit einem oder zwei Füßen aufprallen wird. Ansonsten sind die gewählten Werte willkürliche Eigenschaften des Bodens.

Auch ist aus der obigen Kraftdefinition ersichtlich, daß die Funktion p_z (z -Koordinate) des Körpers als Schaltfunktion für einen Modellwechsel dient. Da die beiden Körper 1 und 7 als Füße dienen sollen, gibt es damit zwei Schaltfunktionen.

Im Prinzip wäre es möglich, auch sämtliche andere Körper des Modells mit eventuellen Boden-Reaktions-Kräften zu versehen, aber es gibt drei Gründe, die dagegen sprechen:

- (1) In der erwarteten Lösung, einer Laufbewegung, werden wohl nur die beiden Füße mit dem Boden kollidieren. Auf diese Art werden allerdings eventuelle optimale Laufbewegungen, in denen z. B. als wesentliches Bewegungselement der „Kopf“ mit dem Boden kollidiert, von vornherein ausgeschlossen und nicht gefunden. Sollte die Lösung aber wider Erwarten z. B. einen Bewegungsanteil haben, bei dem der „Kopf“ in den Boden eindringt (ohne Impuls übertragen zu können, da keine Reaktionskräfte auf ihn wirken), so müßte die Lösung verworfen und das Modell mit einer zusätzlichen Boden-Reaktionskraft auf den „Kopf“ versehen werden.
- (2) Jede zusätzliche Schaltfunktion verlangsamt die Integration der DAE spürbar. In den gerechneten Intervallen einer Simulationsdauer von $\delta t = 0.025s$ mit einem

Schaltereignis verbringt der Integrator mehr als 50% der Zeit mit der genauen Lokalisierung der Schaltstelle.

- (3) Mit der Anzahl der Schaltfunktionen steigt das Risiko, daß Schaltereignisse nicht gefunden werden können. Die Wanderung von Schaltzeitpunkten durch einander hindurch und daraus eventuell resultierende doppelte Ereignisse stellen für den Integrator eine sehr große Schwierigkeit dar (siehe VON SCHWERIN[66]).

§1.4 Initiale Projektion. Für die initiale Projektion wird noch das Produkt zwischen dem Ableitungstensor G_p und ξ benötigt. Als Ergebnis $G_p\xi$ ergibt sich folgende Matrix, die aus typographischen Gründen in drei Matrizen zu jeweils zehn Zeilen aufgespalten wird, wobei die 30 Einträge jeder Zeile in 3-er-Block-Zeilen dargestellt werden. Die Abbildungen V.8 bis V.10 zeigen diese drei Teile.

§2 Die Optimierungsformulierung

Zu den Bedingungen auf mechanischer Seite treten nun auch noch Bedingungen auf der Optimierungsseite hinzu. Dies sind z. B. Bedingungen, die sich mechanisch nur sehr schwer formulieren lassen (wie z. B. Bedingungen, die das Durchdringen zweier Körper ausschließen sollen) oder Bedingungen, die an die gewünschte Lösung gestellt werden (wie z. B. keine volle Umdrehung der Beine) oder die reiner Optimierungsart sind (wie z. B. Periodizitätsforderungen).

§2.1 Mechanische Bedingungen. Als mechanische Bedingungen können z. B. die folgenden Ungleichungsbedingungen aufgefasst werden, die besagen, daß die „Knie“ nur in einer Richtung geknickt werden dürfen. Dies als Bedingung echt auf der Mechanik-Seite einzuführen hieße, weitere Schaltfunktionen und Reaktionskräfte einzuführen. Es ist einfacher, sie auf die Optimierungsseite zu schieben und vom Optimierer zu verlangen, daß die Lösung diese Ungleichungen erfüllen soll.

$$\begin{aligned} b_1 &= p_{13} - p_7 \\ b_2 &= p_{14} - p_8 \\ b_3 &= p_{15} - p_9 \\ c_1 &= (p_{22} - p_{10})/2 \\ c_2 &= (p_{23} - p_{11})/2 \\ c_3 &= (p_{24} - p_{12})/2 \\ d_1 &= b_2c_3 - b_3c_2 \\ d_2 &= b_3c_1 - b_1c_3 \\ d_3 &= b_1c_2 - b_2c_1 \end{aligned}$$

Damit bilden b, c, d ein lokales Koordinatensystem.

$$\begin{aligned} x_1 &= (p_7 - p_4)c_1 + (p_8 - p_5)c_2 + (p_9 - p_6)c_3 \\ y_1 &= (p_7 - p_4)d_1 + (p_8 - p_5)d_2 + (p_9 - p_6)d_3 \\ x_2 &= (p_4 - p_1)c_1 + (p_5 - p_2)c_2 + (p_6 - p_3)c_3 \\ y_2 &= (p_4 - p_1)d_1 + (p_5 - p_2)d_2 + (p_6 - p_3)d_3 \\ res_1 &= (y_1x_2 - x_1y_2) \end{aligned}$$

ABBILDUNG V.8. $G_p \xi$ Zeile 1-10

$$G_p \xi = \begin{pmatrix} 2\xi_1 & 0 & 0 & -2\xi_1 & 0 & 0 & \xi_{14} & 0 & 0 & 0 \\ 0 & 0 & -\xi_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 2\xi_1 & 0 & 0 & -2\xi_1 & 0 & 0 & \xi_{14} & 0 & 0 \\ 0 & 0 & 0 & -\xi_{14} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2\xi_1 & 0 & 0 & -2\xi_1 & 0 & 0 & \xi_{14} & 0 \\ 0 & 0 & 0 & 0 & -\xi_{14} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline -2\xi_1 & 0 & 0 & 2\xi_1 + 2\xi_2 & 0 & 0 & -2\xi_2 - \xi_{14} + \xi_{15} & 0 & 0 & 0 \\ 0 & 0 & \xi_{14} - \xi_{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & -2\xi_1 & 0 & 0 & 2\xi_1 + 2\xi_2 & 0 & 0 & -2\xi_2 - \xi_{14} + \xi_{15} & 0 & 0 \\ 0 & 0 & 0 & \xi_{14} - \xi_{15} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -2\xi_1 & 0 & 0 & 2\xi_1 + 2\xi_2 & 0 & 0 & -2\xi_2 - \xi_{14} + \xi_{15} & 0 \\ 0 & 0 & 0 & 0 & \xi_{14} - \xi_{15} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \xi_{14} & 0 & 0 & -2\xi_2 - \xi_{14} + \xi_{15} & 0 & 0 & 2\xi_2 + 2\xi_3 - 2\xi_{15} & 0 & 0 & \xi_{18} \\ 0 & 0 & -2\xi_3 + \xi_{15} - \xi_{17} & 0 & 0 & -\xi_{16} + \xi_{17} & 0 & 0 & \xi_{16} & 0 \\ 0 & -\xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \xi_{14} & 0 & 0 & -2\xi_2 - \xi_{14} + \xi_{15} & 0 & 0 & 2\xi_2 + 2\xi_3 - 2\xi_{15} & 0 & 0 \\ \xi_{18} & 0 & 0 & -2\xi_3 + \xi_{15} - \xi_{17} & 0 & 0 & -\xi_{16} + \xi_{17} & 0 & 0 & \xi_{16} \\ 0 & 0 & -\xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \xi_{14} & 0 & 0 & -2\xi_2 - \xi_{14} + \xi_{15} & 0 & 0 & 2\xi_2 + 2\xi_3 - 2\xi_{15} & 0 \\ 0 & \xi_{18} & 0 & 0 & -2\xi_3 + \xi_{15} - \xi_{17} & 0 & 0 & -\xi_{16} + \xi_{17} & 0 & 0 \\ \xi_{16} & 0 & 0 & -\xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \xi_{18} & 0 & 0 & 2\xi_6 \\ 0 & 0 & -\xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2\xi_6 & 0 & 0 & \xi_{19} & 0 & 0 & -\xi_{19} & 0 & 0 \end{pmatrix}$$

ABBILDUNG V.9. $G_p \xi$ Zeile 11-20

$$G_p \xi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \xi_{18} & 0 & 0 \\ 2\xi_6 & 0 & 0 & -\xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2\xi_6 & 0 & 0 & \xi_{19} & 0 & 0 & -\xi_{19} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \xi_{18} & 0 \\ 0 & 2\xi_6 & 0 & 0 & -\xi_{18} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\xi_6 & 0 & 0 & \xi_{19} & 0 & 0 & -\xi_{19} \\ \hline -\xi_{14} & 0 & 0 & \xi_{14} - \xi_{15} & 0 & 0 & -2\xi_3 + \xi_{15} - \xi_{17} & 0 & 0 & -\xi_{18} \\ 0 & 0 & 2\xi_3 + 2\xi_4 + 2\xi_{17} & 0 & 0 & -2\xi_4 + \xi_{16} - \xi_{17} & 0 & 0 & -\xi_{16} & 0 \\ 0 & \xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & -\xi_{14} & 0 & 0 & \xi_{14} - \xi_{15} & 0 & 0 & -2\xi_3 + \xi_{15} - \xi_{17} & 0 & 0 \\ -\xi_{18} & 0 & 0 & 2\xi_3 + 2\xi_4 + 2\xi_{17} & 0 & 0 & -2\xi_4 + \xi_{16} - \xi_{17} & 0 & 0 & -\xi_{16} \\ 0 & 0 & \xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -\xi_{14} & 0 & 0 & \xi_{14} - \xi_{15} & 0 & 0 & -2\xi_3 + \xi_{15} - \xi_{17} & 0 \\ 0 & -\xi_{18} & 0 & 0 & 2\xi_3 + 2\xi_4 + 2\xi_{17} & 0 & 0 & -2\xi_4 + \xi_{16} - \xi_{17} & 0 & 0 \\ -\xi_{16} & 0 & 0 & \xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & -\xi_{16} + \xi_{17} & 0 & 0 & 0 \\ 0 & 0 & -2\xi_4 + \xi_{16} - \xi_{17} & 0 & 0 & 2\xi_4 + 2\xi_5 & 0 & 0 & -2\xi_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\xi_{16} + \xi_{17} & 0 & 0 \\ 0 & 0 & 0 & -2\xi_4 + \xi_{16} - \xi_{17} & 0 & 0 & 2\xi_4 + 2\xi_5 & 0 & 0 & -2\xi_5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\xi_{16} + \xi_{17} & 0 \\ 0 & 0 & 0 & 0 & -2\xi_4 + \xi_{16} - \xi_{17} & 0 & 0 & 2\xi_4 + 2\xi_5 & 0 & 0 \\ -2\xi_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \xi_{16} & 0 & 0 & 0 \\ 0 & 0 & -\xi_{16} & 0 & 0 & -2\xi_5 & 0 & 0 & 2\xi_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & \xi_{16} & 0 & 0 \\ 0 & 0 & 0 & -\xi_{16} & 0 & 0 & -2\xi_5 & 0 & 0 & 2\xi_5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$G_{p\xi} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \xi_{16} & 0 \\ 0 & 0 & 0 & 0 & -\xi_{16} & 0 & 0 & -2\xi_5 & 0 & 0 \\ 2\xi_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & -\xi_{18} & 0 & 0 & -2\xi_6 \\ 0 & 0 & \xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\xi_6 & 0 & 0 & -\xi_{19} & 0 & 0 & \xi_{19} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\xi_{18} & 0 & 0 \\ -2\xi_6 & 0 & 0 & \xi_{18} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\xi_6 & 0 & 0 & -\xi_{19} & 0 & 0 & \xi_{19} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\xi_{18} & 0 \\ 0 & -2\xi_6 & 0 & 0 & \xi_{18} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\xi_6 & 0 & 0 & -\xi_{19} & 0 & 0 & \xi_{19} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \xi_{19} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\xi_{19} & 0 & 0 & 2\xi_7 & 0 & 0 & -2\xi_7 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \xi_{19} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\xi_{19} & 0 & 0 & 2\xi_7 & 0 & 0 & -2\xi_7 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \xi_{19} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\xi_{19} & 0 & 0 & 2\xi_7 & 0 & 0 & -2\xi_7 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\xi_{19} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \xi_{19} & 0 & 0 & -2\xi_7 & 0 & 0 & 2\xi_7 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\xi_{19} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \xi_{19} & 0 & 0 & -2\xi_7 & 0 & 0 & 2\xi_7 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\xi_{19} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \xi_{19} & 0 & 0 & -2\xi_7 & 0 & 0 & 2\xi_7 \end{pmatrix}$$

ABBILDUNG V.10. $G_{p\xi}$ Zeile 21-30

$$\begin{aligned} x_1 &= (p_{13} - p_{16})c_1 + (p_{14} - p_{17})c_2 + (p_{15} - p_{18})c_3 \\ y_1 &= (p_{13} - p_{16})d_1 + (p_{14} - p_{17})d_2 + (p_{15} - p_{18})d_3 \\ x_2 &= (p_{16} - p_{19})c_1 + (p_{17} - p_{20})c_2 + (p_{18} - p_{21})c_3 \\ y_2 &= (p_{16} - p_{19})d_1 + (p_{17} - p_{20})d_2 + (p_{18} - p_{21})d_3 \\ res_2 &= (y_1x_2 - x_1y_2). \end{aligned}$$

Gefordert wird:

$$\begin{aligned} res_1 &\geq 0 \\ res_2 &\geq 0. \end{aligned}$$

§2.2 Lösungsbedingungen. Die nachfolgenden Bedingungen entspringen dem gewünschten Aussehen der Lösung. Sie schließen z. B. halbe Rotationen des Fußes um das Knie herum oder das Eindringen der Knie in den Boden aus.

Diese beiden Bedingungen besagen „Hüfte bleibt höher als die Knie“:

$$\begin{aligned} p_9 - p_6 &\geq 0 \\ p_{15} - p_{18} &\geq 0. \end{aligned}$$

„Das Knie höher als der dazugehörige Fuß“:

$$\begin{aligned} p_6 - p_3 &\geq 0 \\ p_{18} - p_{21} &\geq 0. \end{aligned}$$

„Die Knie über dem Boden“:

$$\begin{aligned} p_6 &\geq 0 \\ p_{18} &\geq 0. \end{aligned}$$

§2.3 Optimierungsbedingungen. Die Optimierungsbedingungen sind stark vom ausgewählten Optimierungsziel abhängig. Lautet dieses in etwa „Finde die Steuerungen, die das Modell in fester Zeit aus einer festen Position heraus möglichst weit in x -Richtung bringen.“, so braucht keine weitere Optimierungsbedingung gestellt zu werden. Gibt es aber z. B. eine Periodizitätsforderung, so treten folgende weitere Bedingungen hinzu: die Periodizitätsforderung selbst (N ist Anzahl der Körper)

$$\begin{aligned} p_{3i+1}(t_{end}) - p_{3i+1}(t_0) - p_1(t_{end}) + p_1(t_0) &= 0 & i = 1, \dots, N-1 \\ p_{3i+2}(t_{end}) - p_{3i+2}(t_0) &= 0 & i = 0, \dots, N-1 \\ p_{3i+3}(t_{end}) - p_{3i}(t_0) &= 0 & i = 0, \dots, N-1 \end{aligned}$$

und Konsistenzbedingungen der Startposition. Diese setzen sich zusammen aus den Ortskonsistenzbedingungen

$$\begin{aligned} gpos_1 &= (p_1 - p_4)^2 + (p_2 - p_5)^2 + (p_3 - p_6)^2 - 1 \\ gpos_2 &= (p_4 - p_7)^2 + (p_5 - p_8)^2 + (p_6 - p_9)^2 - 1 \\ gpos_3 &= (p_7 - p_{13})^2 + (p_8 - p_{14})^2 + (p_9 - p_{15})^2 - 1 \\ gpos_4 &= (p_{13} - p_{16})^2 + (p_{14} - p_{17})^2 + (p_{15} - p_{18})^2 - 1 \\ gpos_5 &= (p_{16} - p_{19})^2 + (p_{17} - p_{20})^2 + (p_{18} - p_{21})^2 - 1 \\ gpos_6 &= (p_{10} - p_{22})^2 + (p_{11} - p_{23})^2 + (p_{12} - p_{24})^2 - 1 \\ gpos_7 &= (p_{25} - p_{28})^2 + (p_{26} - p_{29})^2 + (p_{27} - p_{30})^2 - 1 \\ gpos_8 &= p_7 + p_{13} - 2p_{10} \\ gpos_9 &= p_8 + p_{14} - 2p_{11} \\ gpos_{10} &= p_9 + p_{15} - 2p_{12} \\ gpos_{11} &= p_{25} + p_{28} - 2p_{22} \\ gpos_{12} &= p_{26} + p_{29} - 2p_{23} \\ gpos_{13} &= p_{27} + p_{30} - 2p_{24} \\ gpos_{14} &= (p_1 - p_4)(p_7 - p_{13}) + (p_2 - p_5)(p_8 - p_{14}) + (p_3 - p_6)(p_9 - p_{15}) \\ gpos_{15} &= (p_4 - p_7)(p_7 - p_{13}) + (p_5 - p_8)(p_8 - p_{14}) + (p_6 - p_9)(p_9 - p_{15}) \\ gpos_{16} &= (p_{19} - p_{16})(p_7 - p_{13}) + (p_{20} - p_{17})(p_8 - p_{14}) + (p_{21} - p_{18})(p_9 - p_{15}) \\ gpos_{17} &= (p_{16} - p_{13})(p_7 - p_{13}) + (p_{17} - p_{14})(p_8 - p_{14}) + (p_{18} - p_{15})(p_9 - p_{15}) \\ gpos_{18} &= (p_{10} - p_{22})(p_7 - p_{13}) + (p_{11} - p_{23})(p_8 - p_{14}) + (p_{12} - p_{24})(p_9 - p_{15}) \\ gpos_{19} &= (p_{25} - p_{28})(p_{10} - p_{22}) + (p_{26} - p_{29})(p_{11} - p_{23}) + (p_{27} - p_{30})(p_{12} - p_{24}) \end{aligned}$$

und den Geschwindigkeitskonsistenzbedingungen

$$\begin{aligned}
gvel_1 &= 2(v_1 - v_4)(p_1 - p_4) + 2(v_2 - v_5)(p_2 - p_5) + 2(v_3 - v_6)(p_3 - p_6) \\
gvel_2 &= 2(v_4 - v_7)(p_4 - p_7) + 2(v_5 - v_8)(p_5 - p_8) + 2(v_6 - v_9)(p_6 - p_9) \\
gvel_3 &= 2(v_7 - v_{13})(p_7 - p_{13}) + 2(v_8 - v_{14})(p_8 - p_{14}) + 2(v_9 - v_{15})(p_9 - p_{15}) \\
gvel_4 &= 2(v_{13} - v_{16})(p_{13} - p_{16}) + 2(v_{14} - v_{17})(p_{14} - p_{17}) + 2(v_{15} - v_{18})(p_{15} - p_{18}) \\
gvel_5 &= 2(v_{16} - v_{19})(p_{16} - p_{19}) + 2(v_{17} - v_{20})(p_{17} - p_{20}) + 2(v_{18} - v_{21})(p_{18} - p_{21}) \\
gvel_6 &= 2(v_{10} - v_{22})(p_{10} - p_{22}) + 2(v_{11} - v_{23})(p_{11} - p_{23}) + 2(v_{12} - v_{24})(p_{12} - p_{24}) \\
gvel_7 &= 2(v_{25} - v_{28})(p_{25} - p_{28}) + 2(v_{26} - v_{29})(p_{26} - p_{29}) + 2(v_{27} - v_{30})(p_{27} - p_{30}) \\
gvel_8 &= v_7 + v_{13} - 2v_{10} \\
gvel_9 &= v_8 + v_{14} - 2v_{11} \\
gvel_{10} &= v_9 + v_{15} - 2v_{12} \\
gvel_{11} &= v_{25} + v_{28} - 2v_{22} \\
gvel_{12} &= v_{26} + v_{29} - 2v_{23} \\
gvel_{13} &= v_{27} + v_{30} - 2v_{24} \\
gvel_{14} &= (v_1 - v_4)(p_7 - p_{13}) + (p_1 - p_4)(v_7 - v_{13}) + (v_2 - v_5)(p_8 - p_{14}) + \\
&\quad + (p_2 - p_5)(v_8 - v_{14}) + (v_3 - v_6)(p_9 - p_{15}) + (p_3 - p_6)(v_9 - v_{15}) \\
gvel_{15} &= (v_4 - v_7)(p_7 - p_{13}) + (p_4 - p_7)(v_7 - v_{13}) + (v_5 - v_8)(p_8 - p_{14}) + \\
&\quad + (p_5 - p_8)(v_8 - v_{14}) + (v_6 - v_9)(p_9 - p_{15}) + (p_6 - p_9)(v_9 - v_{15}) \\
gvel_{16} &= (v_{19} - v_{16})(p_7 - p_{13}) + (p_{19} - p_{16})(v_7 - v_{13}) + (v_{20} - v_{17})(p_8 - p_{14}) + \\
&\quad + (p_{20} - p_{17})(v_8 - v_{14}) + (v_{21} - v_{18})(p_9 - p_{15}) + (p_{21} - p_{18})(v_9 - v_{15}) \\
gvel_{17} &= (v_{16} - v_{13})(p_7 - p_{13}) + (p_{16} - p_{13})(v_7 - v_{13}) + (v_{17} - v_{14})(p_8 - p_{14}) + \\
&\quad + (p_{17} - p_{14})(v_8 - v_{14}) + (v_{18} - v_{15})(p_9 - p_{15}) + (p_{18} - p_{15})(v_9 - v_{15}) \\
gvel_{18} &= (v_{10} - v_{22})(p_7 - p_{13}) + (p_{10} - p_{22})(v_7 - v_{13}) + (v_{11} - v_{23})(p_8 - p_{14}) + \\
&\quad + (p_{11} - p_{23})(v_8 - v_{14}) + (v_{12} - v_{24})(p_9 - p_{15}) + (p_{12} - p_{24})(v_9 - v_{15}) \\
gvel_{19} &= (v_{25} - v_{28})(p_{10} - p_{22}) + (p_{25} - p_{28})(v_{10} - v_{22}) + (v_{26} - v_{29})(p_{11} - p_{23}) + \\
&\quad + (p_{26} - p_{29})(v_{11} - v_{23}) + (v_{27} - v_{30})(p_{12} - p_{24}) + (p_{27} - p_{30})(v_{12} - v_{24}).
\end{aligned}$$

KAPITEL VI

Optimierungsdurchführung

In diesem Kapitel werden nun, aufbauend auf den bisherigen Vorarbeiten, optimale Laufbewegungen bestimmt. Zunächst ergeben sich einige „Fehlversuche“, die weniger im Mißlingen der Optimierung begründet sind als vielmehr damit, daß die Optimierung den begrenzten Zeitraum nutzt und dabei optimale Bewegungen produziert, die entweder den Erwartungen ganz und gar nicht entsprechen oder die sich z. B. nicht auf längere Zeitintervalle forsetzen lassen.

Durch eine stabilisierende Bedingung kann dann eine erwartungsgemäße Laufbewegung erreicht werden, die sogar so nah an Periodizität heranzuführt, daß ausgehend davon periodische optimale dreidimensionale Laufbewegungen gefunden werden können. Die Schrittfolge ist dabei nicht vorgegeben, sondern ergibt sich im Laufe der Optimierung. Diese gefundenen Resultate können dann wieder als Ausgangspunkt für z. B. andere Zielfunktionale dienen.

Dabei ist zu bemerken, daß die gefundenen Optima mit aller Wahrscheinlichkeit lediglich lokale Optima sind, da die benutzten Verfahren i. allg. keine globalen Optima finden.

Darüberhinaus zeigt sich ein zu erwartendes Resultat, nämlich die enge Verflechtung von „Anatomie“ und optimalem Laufergebnis. Die Bewegung nutzt alle ihr gegebenen Möglichkeiten. So kann es z. B. zu einer grotesken Voll-Rotation der Schulter kommen, da diese in den mechanischen und in den Optimierungsbedingungen nicht ausgeschlossen wird. Daher zeigt manchmal erst das Ergebnis, welche weiteren Einschränkungen (oder Möglichkeiten) dem Modell hinzugefügt werden müssen. Die Erfahrung lehrt aber, daß bei solchen Veränderungen am Modell (insbesondere bei hinzugefügten Einschränkungen) das bisherige Ergebnis als Startwert unbrauchbar ist und die Homotopiemethode (v. a. in der Zeit und Anzahl der Knoten) neu gestartet werden muß.

§1 Vorbemerkung

Zur Durchführung der Optimierung sind nun vier weitere Angaben notwendig: Zielfunktional, Anfangswerte, stabilisierende Bedingungen sowie die Beschränkungen. In ihrer geschickten Wahl besteht die Kunst der Optimierung. Das Ziel der Optimierung soll die Bestimmung einer optimalen dreidimensionalen Bewegung sein, bei der das Modell vorwärts läuft. Von Interesse ist das sich dabei ergebende Bewegungsmuster. Welche Schrittfolge entsteht etc.?

Es ist nun eine wesentliche Eigenschaft des der Optimierung innewohnenden Newton-Verfahrens, daß es nur ein *lokales* Verfahren ist. Dies führt zu folgender Erwartung:

Sind die Startwerte nicht nah genug am Optimum, so ist die Konvergenz zum Optimum wenig wahrscheinlich.

Zugleich ist aber auch klar:

Zwei unterschiedliche Bewegungsmuster liegen i. allg. nicht nahe beieinander.

Damit sieht es so aus, als ob man als Startwert der Optimierung bereits das Bewegungsmuster des Ergebnisses kennen müsste. In solch einem Fall stellt sich aber die Frage,

welchen Erkenntniswert die Durchführung einer solchen Optimierung überhaupt bringt. Des Weiteren stellen sich folgende Probleme in der Praxis ein:

- (1) Ist das Bewegungsmuster des Optimums ungefähr bekannt, ist es dennoch als Startwert unbrauchbar. Die nur ungenauen Zustandswerte zu den Zwischenzeitpunkten der multiple-shooting-Knoten zusammen mit unbekanntem Steuerungen führen i. allg. zum Abbruch der Optimierung. Die Zwischenzustände liegen für den Optimierer so weit „auseinander“, daß es (im Rahmen der stattfindenden Linearisierung) so wirkt, als ob sie gar nicht stetig aneinander angeschlossen werden können. Der Vorgang wird abgebrochen als wäre er physikalisch gar nicht lösbar.
- (2) Der Versuch, ein ungefähres Startbewegungsmuster zu wählen oder durch Interpolation zwischen Start- und gewünschtem Endzustand berechnen zu lassen, ist noch weniger geeignet. Solch eine Herangehensweise hat zur Folge, daß tatsächlich physikalisch unmögliche Bewegungen als initial guess gewählt werden. Im Dreidimensionalen sind fast alle Bewegungen mit Rotationen und Gegenrotationen verbunden. Im Falle einer Interpolation bzw. einer Modellierung per Hand kommen diese nicht korrekt vor. Dies wirkt, als müsste z. B. das System einen Schritt vorwärts machen, ohne sich dabei gleichzeitig ein wenig um den stehenden Fuß zu drehen. Dies ist nicht möglich und im Rahmen der Linearisierung wird hier auch keine Lösung erkannt. Dies führt entweder dazu,
 - (a) daß keine zulässigen Lösungen des linearisierten Optimierungsproblems gefunden werden, nach einigen Iterationen bricht die Optimierung ab, oder
 - (b) daß nach wenigen Iterationen der stetige Anschluß der ersten Mehrzielknoten hergestellt wird, die weiteren Mehrzielknoten ebenfalls stetig angeschlossen sind, aber im Grunde durch Weiterlösen der DAE aus den ersten Knoten gewonnen werden und die Startinformation der hinteren Knoten gänzlich verloren gegangen ist.

Diese Schwierigkeiten werden überwunden, wenn der initial guess der Zustände des Multiple-Shooting-Verfahrens durch Integration aus dem ersten Knoten gewonnen wird. Im Idealfall, wenn die Kontrollen der optimalen Lösung bekannt sind, ist dieser Startzustand dann auch gleich das Optimum. Im allgemeinen ist aber von den optimalen Kontrollen gar nichts bekannt.

Eine Möglichkeit besteht nun darin, ausgehend von völlig unbekanntem Kontrollen (von denen man angemessenerweise als initial guess annimmt, sie seien alle 0) durch manuelles Probieren Kontrollen herauszufinden, die eine unbeholfene Vorwärtsbewegung erzeugen, um diese dann als Startwerte zu nehmen. Auch dies stellt sich in der Praxis als außerordentlich schwierig heraus. Das Modell reagiert im Gegensatz zur Puppe eines Marionettenspielers hochgradig nichtlinear. Es gibt keine offensichtlichen Wirkungen und die nichtoffensichtlichen bewirken das ständige Zusammenfallen des Modells.

Zusammenfassend sei daher formuliert:

FAUSTREGEL:

- (1) Als Startwerte können nur stetige Trajektorien (oder sehr nah daran gelegene) mit passenden Kontrollen verwendet werden.
- (2) Die passenden Kontrollen müssen vom Optimierungsprozeß selbst ermittelt werden.

Eine Methode, die diese beiden Bedingungen erfüllt, ist die *Homotopiemethode*¹. Diese besteht darin, Bedingungen und insbesondere das Zielfunktional Schritt für Schritt zu verändern. So wird mit einfachen Bedingungen und einem einfachen Zielfunktional gestartet, damit gute Aussichten bestehen, ein Optimum zu erreichen. Dieses wird dann als Startwert für ein Problem mit leicht geänderten Bedingungen und Zielfunktional genommen. Auf diese Weise versucht man bis zum ursprünglichen Problem zu gelangen. Sollte diese Methode von Erfolg gekrönt sein, so ist mit ihr auch ein nichttrivialer Erkenntniswert gewonnen. Das sich ergebende Bewegungsmuster des gefundenen Optimums ist nicht bereits in den Startwerten vorgegeben und wiedergefunden worden, sondern hat sich ergeben.

§2 Initialwerte

Ausgehend von der Stellung

$$\begin{array}{ll}
 p_1 = -0.25844 & p_{16} = 0.5 \\
 p_2 = 0.5 & p_{17} = -0.5 \\
 p_3 = 0.001 & p_{18} = 1.101 \\
 p_4 = 0 & p_{19} = 0.5 \\
 p_5 = 0.5 & p_{20} = -0.5 \\
 p_6 = 0.9670254 & p_{21} = 0.101 \\
 p_7 = 0 & p_{22} = 0 \\
 p_8 = 0.5 & p_{23} = 0 \\
 p_9 = 1.9670254 & p_{24} = 2.9670254 \\
 p_{10} = 0 & p_{25} = 0 \\
 p_{11} = 0 & p_{26} = 0.5 \\
 p_{12} = 1.9670254 & p_{27} = 2.9670254 \\
 p_{13} = 0 & p_{28} = 0 \\
 p_{14} = -0.5 & p_{29} = -0.5 \\
 p_{15} = 1.9670254 & p_{30} = 2.9670254 \\
 v_i = 0 \text{ für } i = 1, \dots, 30 &
 \end{array}$$

erhält man mit allen Kontrollen auf 0 gesetzt für einen simulierten Zeitraum von 0.85s den Bewegungsablauf, der in Abbildung VI.1 wiedergegeben ist.

Während die beiden Füße auf dem Boden ruhen, sinkt der Körper langsam in sich zusammen und in den Boden hinein (auf die anderen Massen außer den Füßen wirkt ja keine rücktreibende Kraft).

§3 Erste Versuche

Die obige Starttrajektorie in ihrer vollen Länge ist als initial guess unbrauchbar. Der Zusammenbruch des Modells verletzt die in Abschnitt §2.2 genannten Lösungsbedingungen. Innerhalb der Linearisierung erkennt der Optimierer nicht, wie er diese Bedingungen erfüllen könnte und bricht daher bereits nach einem SQP-Schritt ab.

¹siehe z. B. BOCK[6]

§3.1 $t_{end} = 0.1s$. Dagegen kann von obiger Starttrajektorie ein Anfangsstück genommen werden, d. h. die Zeitdauer wird zunächst einmal als kürzer angenommen. Mit der obigen Starttrajektorie wird in einem ersten Versuch formuliert: in der festen Zeit von

$$t_{end} := 0.1s$$

soll das System das Zielfunktional

$$\mathcal{J}(p, v, q) := -p_{10}$$

minimieren (d. h. die x -Koordinate des Körpers Nr. 4 maximieren), also möglichst weit in x -Richtung kommen. Die Ausgangsposition, d. h. der erste Knoten, ist dabei fix. Die Zeit wird in vier multiple-shooting-Knoten unterteilt, an die Zustände werden die Schranken

$$\begin{aligned} -1.0 &\leq p_i \leq 5.0 \\ -50.0 &\leq v_i \leq 50.0 \end{aligned}$$

gestellt. Sie werden so gewählt, daß sie eigentlich nicht erreicht werden können. Der Zustand des ersten Knotens wird fix gehalten und an die stückweise linearen Kontrollen die Beschränkung

$$-1000.0 \leq q_i \leq 1000.0$$

gestellt. Das Ergebnis wird in Abbildung VI.2 gezeigt. Die Rechenzeit beträgt dabei auf einem 233 MHz PentiumII PC gerundet 3 Stunden 45 Minuten, benötigt werden 1671 SQP-Schritte. Als Abbruchgenauigkeit ist dabei die Erfüllung der KKT-Bedingungen genauer als 10^{-6} gewählt, das Zielfunktional erreicht den Wert (gerundet)

$$-0.9386.$$

Offenbar stößt sich das Modell mit einem Bein schwungvoll ab, um mit einer „Klappmesserbewegung“ die „Hüfte“ möglichst weit in x -Richtung zu schleudern. Ein wenig erinnert dies an gewisse Strategien im Sport.

§3.2 $t_{end} = 0.2s$. In einem zweiten Versuch wird bei gleicher Starttrajektorie und 8 Mehrzielknoten die Endzeit auf

$$t_{end} := 0.2s$$

gesetzt (und ansonsten die gleiche Wahl der Schranken wie beim ersten Versuch verwendet). Hier ergibt sich nach 745 SQP-Schritten (also deutlich weniger Schritte als bei kürzerer Zeit!) und etwa 3 Stunden 25 Minuten auf dem gleichen System die Lösungsstrajektorie, die in Abbildung VI.3 gezeigt wird. Das Zielfunktional im Optimum beträgt

$$-1.592.$$

Hier wird nun in einer ersten Phase durch eine „Hüpfbewegung“ eine kleine Strecke horizontal zurückgelegt und anschließend die erfolgreiche Schleuderbewegung durchgeführt. Massiv tritt hier eine Durchrotation der Schulter um die Hüfte herum hinzu, die durch keine der Optimierungsbedingungen verboten wird (denn eine solche Bewegung wurde nicht erwartet). Hier ergibt sich die Fragestellung, die früher schon einmal angedeutet wurde: läßt man diese „Durchrotation“ zu? Denn schließlich wäre es eine zulässige Lösung. Oder verwirft man diese Lösung, fügt zu den Optimierungsbedingungen weitere hinzu, die eine solche Lösung ausschließen und startet den Optimierungsvorgang erneut? Diese Fragen brauchen aber an dieser Stelle noch nicht beantwortet zu werden, im weiteren Verlauf wird eine solche Bewegung als zulässig erlaubt.

§3.3 $t_{end} = 0.3s$. Im dritten Versuch wird bei gleicher Starttrajektorie und 12 Knoten die Endzeit auf

$$t_{end} := 0.3s$$

gesetzt. Hier werden 1360 SQP-Schritte und ca. 12 Stunden 33 Minuten Rechenzeit benötigt, im Ergebnis (Abbildung VI.4) beträgt das Zielfunktional

$$-2.680.$$

Nach einer recht erfolgversprechenden Anfangsphase der Bewegung setzt wieder der artistische Teil ein. Um am gesetzten zeitlichen Ende die Hüfte möglichst weit in x -Richtung gebracht zu haben, wird zunächst noch während des Bodenkontakts Schwung in diese Richtung geholt, und am Ende die Hüfte in x -Richtung geschleudert, wobei der restliche Körper eine entsprechende Gegenbewegung macht. Die Schulter rotiert wieder um die Hüfte herum. So erfolgreich der Beginn der Bewegung ist (hier wird schon durch Benutzung beider Beine und eines sukzessiven Bodenkontakts eine Bewegung in x -Richtung erzeugt), so unerwartet ist immer noch der Schluß. Der Endzeitpunkt zusammen mit dem Zielfunktional bewirkt, daß das Ende der Bewegung völlig auf das bestmögliche Resultat bzgl. der x -Koordinate des Körpers 4 ausgerichtet ist, selbst wenn nach dem Endzeitpunkt die Bewegung möglicherweise gar nicht mehr fortgesetzt werden könnte (z. B. wegen der Einhaltung gewisser Bedingungen).

§3.4 $t_{end} = 0.4s$. Im vierten Versuch wird bei gleicher Starttrajektorie und 16 Knoten die Endzeit auf

$$t_{end} := 0.4s$$

gesetzt. Die Optimierung zeigt hier bereits sehr starke Durchführungsschwierigkeiten. Der Optimierungsprozess bricht immer wieder ab, erlaubt aber einen „Kaltstart“ mit initialer Hesse-Matrix-Schätzung. Nach vier Tagen Rechenzeit ist aber noch immer kein Optimum gefunden. Die Karush-Kuhn-Tucker-Bedingungen werden zwar gut erfüllt (ca. 10^{-5} bis 10^{-6}), jedoch ändert sich in fast jedem Schritt die Menge der aktiven Ungleichungen. Nach ca. 3200 SQP-Schritten wird daher der Optimierungslauf abgebrochen und mit dem Zwischenergebnis ein neuer Lauf gestartet, allerdings mit einer Homotopie in der Abbruchgenauigkeit. So wird zunächst ein (lokales) Optimum mit Abbruchgenauigkeit 10^{-2} gesucht, dann 10^{-3} usw. bis 10^{-7} . Die sich ergebenden Optimalwerte sind

KKT-Bedingung genauer als : Zielfunktional im Optimum

$$10^{-2} : -1.4821473215910805354$$

$$10^{-3} : -1.4825016535114083904$$

$$10^{-4} : -1.4830743158709716045$$

$$10^{-5} : -3.5384704371819468705$$

$$10^{-6} : -3.7168608947115790997$$

$$10^{-7} : -3.7468882828827116960.$$

Da der Start der ersten Optimierung mit einem Zwischenergebnis erfolgt, bei dem das Zielfunktional schon im Bereich -3.7 liegt, sieht man sehr deutlich, daß große Teile des initial guess wohl verworfen werden und die Optimierung im Prinzip neu startet. Allerdings ist der Startunterschied groß genug, um dieses Mal zur Konvergenz zu führen und das nach nur 473 SQP-Schritten und ca. 10 Stunden Rechenzeit. Das Ergebnis zeigt Abbildung VI.5. Hier wird eine initialer Hüpfen vollführt gefolgt von einer wieder recht vielversprechenden Schrittbewegung, die aber erneut in die nun hinlänglich bekannte Flugphase verfällt. Positiv zu bemerken ist neben dieser halben Schrittbewegung die Impulsübertragung auf den

Boden, die offensichtlich stattfindet. Unerwartet ist die massive Schulterdurchrotation, ein Problem ist weiterhin die Flugphase.

§3.5 $t_{end} = 0.5s$. Im fünften Versuch wurde bei gleicher Starttrajektorie und 20 Knoten die Endzeit auf

$$t_{end} := 0.5s$$

gesetzt. Hier wird von vornherein auch eine Homotopie in der Abbruchgenauigkeit durchgeführt. Nach ca. 15h 51 Minuten Rechenzeit und 581 SQP-Schritten ergibt sich ein Optimum mit dem Wert -5.0. Im Optimum werden die Karush-Kuhn-Tucker-Bedingungen auf $1.9 \cdot 10^{-9}$ genau erfüllt. Das Ergebnis der Optimierung zeigt Abbildung VI.6. Auch hier wird nach einem einleitenden Hüpfen die Bewegung durch die Rotation der Schulter dominiert. Die Fallbewegung des Schwerpunktes wird durch das Absenken der relativ massereichen Schulter aufgefangen, so daß die Hüfte möglichst lang in ausgestreckter Position bleiben kann. Die Bewegung als Ganzes macht einen zunehmend chaotischen Eindruck. Besondere die Bewegung der Schulter spielt nun eine Hauptrolle, während die Beine, zumindest im zweiten Teil der Trajektorie, nur noch von wenig belang scheinen. Ihre Bewegungen sind nur noch sehr schwer nachzuvollziehen. So läßt sich sagen, daß die Optimierung zumindest unerwartete Ergebnisse liefert, die auch nicht in den Anfangswerten oder dem initial guess versteckt vorgegeben sind.

§3.6 $t_{end} = 0.6s$. Im sechsten Versuch wird bei gleicher Starttrajektorie und 24 Knoten die Endzeit auf

$$t_{end} := 0.6s$$

gesetzt. Auch hier wird wieder eine Homotopie in der Abbruchgenauigkeit verwendet. Nach ca. 5h Rechenzeit und 265 SQP-Schritten ergibt sich nun ein qualitativ anderes Bild (siehe Abbildung VI.7). Das Zielfunktional im Optimum beträgt -5.0.

Hier geht der anfängliche Hüpfen in ein gewagtes gesprungenes Spagat über, gefolgt von der bereits vertrauten Schulterrotation. Die Schulter durchschlägt dabei den Boden, was nicht weiter verwunderlich ist, da auf die Schultermassen keinerlei rücktreibende Kräfte wirken und für sie der Boden nicht existent ist. Hier ist eine Erwartung falsifiziert worden: es wird erwartet, es werde sich keine optimale Bewegung ergeben, bei der andere Massen als die der Füße den Boden durchdringen.

§3.7 $t_{end} = 0.7s$. Im siebten Versuch wird bei gleicher Starttrajektorie und 28 Knoten die Endzeit auf

$$t_{end} := 0.7s$$

gesetzt. Bei Homotopie in der Abbruchgenauigkeit und nach ca. 10 Stunden 30 Minuten Rechenzeit sowie 340 SQP-Schritten ergibt sich ein neues Bild (siehe Abbildung VI.8). Das Zielfunktional beträgt im Optimum -5.0.

Einer anfänglichen Sprungbewegung, die die Startposition verbessert, folgt ein Schritt zumindest eines Fußes. Doch dann verschwindet die Schulter wieder unter dem Boden. Die Endstellung darf als problematisch bezeichnet werden.

§4 Zweite Versuche

In einem zweiten Versuch wird noch einmal das gleiche Vorgehen wie in Abschnitt §3 probiert, diesmal aber mit einem anderen Zielfunktional. Statt die Maximierung von p_{10} zu verlangen, wird die Maximierung der x -Koordinate des Schwerpunkts verlangt, also

$$\mathcal{J}(p, v, q) := -\frac{1}{10}(p_1 + p_4 + p_7 + p_{10} + p_{13} + p_{16} + p_{19} + p_{22} + p_{25} + p_{28}).$$

Hier benötigt die Homotopiemethode in der Anzahl der Knoten² für bis zu 22 Knoten ca. 10 Tage Rechenzeit und resultiert in dem Ergebnis, das Abbildung VI.9 zeigt. Auch hier ist das Resultat (das Zielfunktional beträgt im Optimum übrigens -2.8697) von bemerkenswertem Aussehen, es zeichnet sich jedoch keinerlei Laufstruktur ab. Zusätzlich endet die Bewegung in einer Art und Weise, die eine Fortsetzung um weitere Knoten zunehmend schwieriger macht: das Anhängen eines Knotens bewirkt zunächst einmal nur die Fortsetzung der Bewegung (Verlängerung der Simulation um den Zeitraum eines Knotens). Der Aufprall verletzt bei Verlängerung des Simulationszeitraums aber einige der Optimierungsnebenbedingungen (z. B. „Knie höher als der Fuß“). Der Optimierer sucht nun nach Kontrollen, die es der Bewegung ermöglichen, die Nebenbedingungen zu erfüllen. Nachdem der Körper aber so seltsam aufkommt, ist dies im Rahmen der physikalischen Möglichkeiten nicht mehr erreichbar. Es können zwar alle Bedingungen erfüllt werden, indem die Bewegung anders aussieht, aber im Rahmen der Linearisierung „sieht“ der Optimierer die (weit entfernt gelegenen) Möglichkeiten nicht mehr. Dadurch kommt es zum Abbruch des Optimierungsvorgangs. Damit ist auch dieser Ansatz nicht gelungen.

§5 Stabilisierende Bedingungen

Die ermittelten, sicherlich nur lokalen Optima sind zwar sehr beeindruckend in ihrer Bewegung, nicht aber der gewünschte Bewegungsverlauf, insbesondere der Schultereinsatz und das Bewegungsende. Auch ist es im Sinne des Homotopieverfahrens sehr störend, daß das Bewegungsende in einer Art verläuft, die meist über den Endzeitpunkt nicht mit den Optimierungsbedingungen verträglich fortgesetzt werden kann. Dies macht eine Homotopie in der Anzahl der Knoten und der Dauer der simulierten Zeit unmöglich.

Um also einerseits die Bewegung in erwünschtere Bahnen zu lenken, andererseits eine Weiterführbarkeit der Bewegung zu sichern und ebenfalls um zu verhindern, daß das System in lokalen Minima stecken bleibt (wie z. B. dem Umfallen), werden an das System zusätzliche *stabilisierende* Bedingungen gestellt. Diese stabilisierenden Bedingungen werden hier zu den mechanischen Bedingungen des Abschnitts wie folgt hinzugefügt: die erste Bedingung besagt, daß das System kein „Spagat“ machen sollen, der Winkel zwischen den beiden Oberschenkeln darf höchstens $\pi/2$ betragen:

$$\begin{aligned}x_1 &= (p_7 - p_4)c_1 + (p_8 - p_5)c_2 + (p_9 - p_6)c_3 \\y_1 &= (p_7 - p_4)d_1 + (p_8 - p_5)d_2 + (p_9 - p_6)d_3 \\x_2 &= (p_{13} - p_{16})c_1 + (p_{14} - p_{17})c_2 + (p_{15} - p_{18})c_3 \\y_2 &= (p_{13} - p_{16})d_1 + (p_{14} - p_{17})d_2 + (p_{15} - p_{18})d_3 \\x_1x_2 + y_1y_2 &\geq 0.\end{aligned}$$

Die nächste Bedingung besagt, daß die Hüfte (mit einer gewissen Toleranz) „zwischen“ den Knien gehalten werden soll (im Sinne von: die beiden Knie sollen sich nicht gleichzeitig beide hinter der Hüfte befinden oder beide zugleich vor der Hüfte):

$$\begin{aligned}y_1 &= (p_7 - p_4)d_1 + (p_8 - p_5)d_2 + (p_9 - p_6)d_3 \\y_2 &= (p_{13} - p_{16})d_1 + (p_{14} - p_{17})d_2 + (p_{15} - p_{18})d_3 \\-y_1y_2 + 0.01 &\geq 0.\end{aligned}$$

²D. h. es wird ein Optimum für einen Knoten gesucht, mit diesem Ergebnis dann das Optimum für zwei Knoten usw.

Die bedeutendste Auswirkung hat aber folgende Beschränkung, die verlangt, daß der Oberkörper senkrecht bleibt:

$$\begin{aligned} p_{22} - p_{10} &= 0 \\ p_{23} - p_{11} &= 0. \end{aligned}$$

§6 Ein paar Schritte

§6.1 Ein erstes Vorwärtskommen. Mit Hilfe der stabilisierenden Bedingungen kann nun ein Homotopieverfahren in der Anzahl der Knoten und der Länge der Zeitdauer durchgeführt werden. Wieder ausgehend vom Startzustand (erster Knoten), wie in den obigen Beispielen, wird ein Optimum für einen Knoten (plus Anfangsknoten) und der Zeitdauer von $t_{end} = 0.025s$ gesucht. Das Ergebnis wird als initial guess für die gleiche Problemformulierung aber mit zwei Knoten und der Zeit $t_{end} = 0.05s$ benutzt. Liegt jeweils ein Optimum vor, so wird dieses als Startwert benutzt, um die Problemformulierung um einen Knoten und die Zeitspanne von $0.025s$ zu verlängern. Das Ergebnis zeigt Abbildung VI.10.

Das Zielfunktional ist dabei weiterhin

$$\mathcal{J}(p, v, q) := -p_{10},$$

der Wert beträgt im Optimum -1.801 .

Die Bewegung mündet in einen Ablauf, der eher an ein „Krabbeln auf Knien“ erinnert. Es sei aber daran erinnert, daß in der Modellformulierung auf die Knie keine Kräfte wirken, daher ist es nicht wirklich eine Krabbelbewegung.

§6.2 Tiefergelegter Schwerpunkt. Mit obiger Bewegung als Starttrajektorie wird dann (bei gleichbleibender Knotenzahl) das Optimum der Bewegung gesucht, in der das Zielfunktional

$$\mathcal{J}(p, v, q) := -p_{10}$$

verändert wird zu

$$\mathcal{J}(p, v, q) := -p_{10} + \int_{t_0}^{t_{end}} (z_{Schwerpunkt}(\tau) - C)^2 d\tau,$$

wobei C eine Konstante ist, die ausgehend von 1.13 (d. i. ca. $2/3$ der Schwerpunkthöhe des Modells mit ausgestreckten Beinen) dann Werte von $1.0, 0.9, \dots, 0.1, 0.0$ annimmt (Homotopie im Wert von C). Damit soll der Schwerpunkt des Systems tief gehalten werden, um der Bewegung zusätzliche Stabilität zu verleihen. Die Gesamtrechendauer der Optimierung bis $C = 0.0$ beträgt ca. 10 Tage.

Das Ergebnis für $C = 0.0$ der Koordinate p_{10} im Optimum ist -1.8166 und unterscheidet sich unwesentlich vom Wert des Optimums ohne Modifikation des Zielfunktional, der Wert des Zielfunktional selbst beträgt -0.3998 . Die Bewegung unterscheidet sich auch optisch kaum von der Startbewegung (Abbildung VI.11).

§6.3 Strenger tiefergelegter Schwerpunkt. Bei diesem Versuch wird wieder das Zielfunktional verändert, aber zu

$$\mathcal{J}(p, v, q) := -p_{10} + 100 \int_{t_0}^{t_{end}} (z_{Schwerpunkt}(\tau) - C)^2 d\tau,$$

wobei C den Wert $C = 1.13$ annimmt. Als Starttrajektorie wird das Ergebnis aus Abschnitt §6.1 verwendet. Diese Optimierung misslingt, das Verfahren konvergiert nicht. Vermutlich liegen Startwert und nächstes lokales Optimum zu weit auseinander.

Erfolgreich ist dagegen die Strategie, mit nur vier Knoten zu beginnen und ausgehend von jedem gefundenen Optimum ein neues mit angehängtem Knoten zu suchen. Bereits nach nur vier Tagen Rechenzeit zeigt sich bei 20 Knoten ein Einmünden in eine periodische Bewegung (Abbildung VI.12).

Bezüglich einer ausführlichen Darstellung des Optimierungsergebnisses sei auf den Anhang verwiesen.

§7 Periodische Bewegungen

§7.1 Erreichen einer periodischen Bewegung. In Abschnitt §6.3 wird eine Bewegung erzielt, in der ein Einmünden in einen periodischen Ablauf (zumindest näherungsweise) sichtbar wird. Aus dieser Bewegung wird nun dieser fast-periodische Teil ausgeschnitten (Knoten 13 bis 18 bei einer mit 0 beginnenden Zählung), um ihn als Startwert zur Erreichung einer periodischen Bewegung zu verwenden. Diese 6 Knoten werden auf 24 Knoten verfeinert, um eine hinreichende Anzahl von Kontrollen für die Erreichung einer periodischen Bewegung zu erhalten.

Dazu wird zunächst als Zielfunktional die Minimierung der Abweichung von einer periodischen Bewegung gewählt, wobei jeweils nach Erreichung eines Optimums eine weitere „harte“ Periodizitätsbedingung an die Bewegung gestellt wird:

$$\begin{aligned}
\delta_x &:= p_1(t_{end}) - p_1(t_{start}) & , & \quad per_{10} = v_0(t_{end}) & , \\
per_0 &= p_4(t_{end}) - \delta_x & , & \quad per_{11} = v_4(t_{end}) & , \\
per_1 &= p_7(t_{end}) - \delta_x & , & \quad per_{12} = v_7(t_{end}) & , \\
per_2 &= p_9(t_{end}) & , & \quad per_{13} = v_9(t_{end}) & , \\
per_3 &= p_{10}(t_{end}) - \delta_x & , & \quad per_{14} = v_{10}(t_{end}) & , \\
per_4 &= p_{11}(t_{end}) & , & \quad per_{15} = v_{11}(t_{end}) & , \\
per_5 &= p_{12}(t_{end}) & , & \quad per_{16} = v_{12}(t_{end}) & , \\
per_6 &= p_{16}(t_{end}) - \delta_x & , & \quad per_{17} = v_{16}(t_{end}) & , \\
per_7 &= p_{19}(t_{end}) - \delta_x & , & \quad per_{18} = v_{19}(t_{end}) & , \\
per_8 &= p_{22}(t_{end}) - \delta_x & , & \quad per_{19} = v_{22}(t_{end}) & , \\
per_9 &= p_{25}(t_{end}) - \delta_x & , & \quad per_{20} = v_{25}(t_{end}). &
\end{aligned}$$

Diese 21 Bedingungen entsprechen 11 Bedingungen an eine periodische Geschwindigkeit sowie 10 Bedingungen an periodische Orte (gemäß 11 Freiheitsgraden jeweils für Orte und Geschwindigkeiten, wobei das Vorwärtskommen in Richtung p_0 keiner einschränkenden Bedingung unterworfen wird).

Mit der Einfügung der letzten Periodizitätsbedingung wird dann das Zielfunktional geändert. Es soll nun der Ortsschwerpunkt soweit wie möglich in x -Richtung gebracht werden:

$$\begin{aligned}
\mathcal{J} = & - (p_1(t_{End}) + p_4(t_{End}) + p_7(t_{End}) + p_{10}(t_{End}) + p_{13}(t_{End}) + \\
& + p_{16}(t_{End}) + p_{19}(t_{End}) + p_{22}(t_{End}) + p_{25}(t_{End}) + p_{28}(t_{End}))/10.0.
\end{aligned}$$

Da $p_1(t_{start})$ festgehalten wird, ist es nur noch notwendig die initiale Geschwindigkeit zu beschränken (ansonsten würde sich zwar eine periodische Geschwindigkeit ergeben, allerdings im Laufe des Optimierungsprozesses mit immer größerer Startgeschwindigkeit, bis die initialen v -Werte an die Box-Schranken für v_{start} stoßen und sich dann aber eher eine Flugbewegung ergibt). Da der v -Schwerpunkt in x -Richtung im Startknoten einen Wert von ca. 2.3 hat, wird eine etwas willkürlich gewählte Schranke mit rund dem doppelten

Wert gewählt:

$$(v_1(t_{Start}) + v_4(t_{Start}) + v_7(t_{Start}) + v_{10}(t_{Start}) + v_{13}(t_{Start}) + v_{16}(t_{Start}) + v_{19}(t_{Start}) + v_{22}(t_{Start}) + v_{25}(t_{Start}) + v_{28}(t_{Start}))/10.0 \leq 5.0.$$

Nach ca. 1 Tag Rechenzeit ergibt sich dann ein Optimum. Das Resultat ist veranschaulicht in Abbildung VI.13 (ausführlicher siehe Anhang).

Die sich ergebende Bewegung ist weniger ein Laufen als vielmehr ein Hüpfen auf einem Bein. Man beachte aber, daß der Mechanismus recht flott vorwärts kommt, denn die Zeitdauer beträgt nur 0.15s.

§7.2 Laufen. Mit dem Ergebnis der Optimierung des vorherigen Abschnitts wird ein neuer Optimierungslauf gestartet mit gleichem Zielfunktional, aber ohne die Bedingung, daß der „Oberkörper“ senkrecht zu bleiben hat. Hier ergibt sich ein neues Optimum nach ca. 2,5 Tagen Rechenzeit. Es ist in Abbildung VI.14 dargestellt (ausführlicher siehe Anhang).

§7.3 Steuerungsminimal. Das Ergebnis des vorigen Abschnitts wird wieder als Starttrajektorie für eine neue Optimierungssuche benutzt, diesmal mit anderem Zielfunktional. Es wird unter Hinzufügung der Randbedingung

$$(p_1(t_{end}) + p_4(t_{end}) + p_7(t_{end}) + p_{10}(t_{end}) + p_{13}(t_{end}) + p_{16}(t_{end}) + p_{19}(t_{end}) + p_{22}(t_{end}) + p_{25}(t_{end}) + p_{28}(t_{end}))/10 \geq 2$$

(zurückgelegte Mindestentfernung) verlangt, daß

$$\mathcal{J} = \int_{t_0}^{t_{end}} \sum_{i=1}^5 \frac{u_i(\tau)^2}{10^4} d\tau$$

minimiert werden soll. Die Skalierung mit 10^{-4} ist notwendig, da sich ansonsten keine Konvergenz ergibt.

Nach nur 24 Stunden Rechenzeit wird ein Optimum gefunden, der Wert von \mathcal{J} hat dabei von 56.27 auf 13.37 abgenommen. Die Abbildung VI.15 zeigt das Ergebnis (ausführlicher siehe Anhang).

§7.4 Laufen bei freier Endzeit. Hier soll eine periodische Bewegung bei freier Endzeit gesucht werden. D. h. die obere Schranke für die Gesamtzeitdauer wird so hoch angesetzt, daß sie voraussichtlich nicht erreicht wird. Eine Verlängerung der Zeit bedeutet, daß die Kräfte länger wirken dürfen (damit das Modell wahrscheinlich weiter vorankommt). Gleichzeitig ist aber das Diskretisierungsschema fest, d. h. die Kontrollen bleiben stückweise konstant auf einer festen Anzahl von Intervallen. Damit sind bei zu langer Zeitdauer die Kontrollen möglicherweise nicht mehr fein genug, um alle Nebenbedingungen der Bewegung einzuhalten. Insofern gibt es also zwei widerstrebende Kräfte, die Zeit zu verlängern bzw. zu verkürzen.

Ausgehend von dem in Abschnitt §7.2 gefunden Optimum ergibt sich nach ca. 2 Tagen Rechenzeit ein neues Optimum, das Zielfunktional hat sich dabei von -2.3055 und einer Zeitdauer von $0.15s$ auf -4.6167 und einer Zeitdauer von 0.6407 vergrößert. Das Ergebnis ist in Abbildung VI.16 gezeigt (ausführlicher siehe Anhang).

§8 Vierbeinige Versuche

Der große Vorteil der Formulierung des Mechanismus als Mehrkörpersystem ist die Modularität. Durch ein einfaches Erweitern der Anzahl der Massen, Kräfte und Zwangsbedingungen kann das zweibeinige Modell aus Kapitel V zu einem vierbeinigen Modell erweitert werden. In dieser Variante werden lediglich an die vorhandenen Schultern des ursprünglichen Modells zwei Beine angehängt. Als weitere Kräfte treten vier Drehmomente hinzu, die die Analoga der Drehmomente q_1 bis q_4 darstellen. Auf eine ausführlichere Darstellung des vierbeinigen Modells soll hier verzichtet werden.

Dabei müssen die bereits ermittelten Formulierungen von g_{pos} , g_{vel} , γ und F nicht verändert werden, sondern es werden nur weitere Gleichungen hinzugefügt.³

Für dieses Modell werden analog zu Abschnitt §3 optimale Bewegungen berechnet: ausgehend von einer festen Ausgangsstellung soll in fester Zeit der Schwerpunkt möglichst weit in x -Richtung gebracht werden. Wieder wird ein Homotopieverfahren in der Anzahl der Knoten benutzt. Nach ca. 3 Tagen ergibt sich das in Abbildung VI.18 gezeigte Ergebnis bestehend aus fünf Knoten. Die Optimierung kann an dieser Stelle nicht fortgesetzt werden, da die um einen Knoten verlängerte Bewegung zu keinem erfolgreichen Optimierungslauf mehr führt.

Auch hier hilft die Idee der stabilisierenden Bedingung. Diese besteht in diesem Fall daraus, den „Rücken“ waagrecht zu halten, d. h. die Mitte zwischen den beiden „Schultern“ sollen die gleiche z -Koordinate haben. Mit dieser Einschränkung verläuft das Verfahren wesentlich erfolgreicher, hier ergibt sich nach ca. 8 Tagen Rechenzeit das Ergebnis in Abbildung VI.19 bestehend aus 9 Knoten. Das Zielfunktional verbessert sich in Zeitraum von 0.2s von -0.579 auf -1.74.

³Eine Ausnahme bildet der Tensor $G_p\xi$, der neu berechnet werden muß. Er wird aber aus den weiterhin gültigen g_{pos} plus neue Gleichungen berechnet.

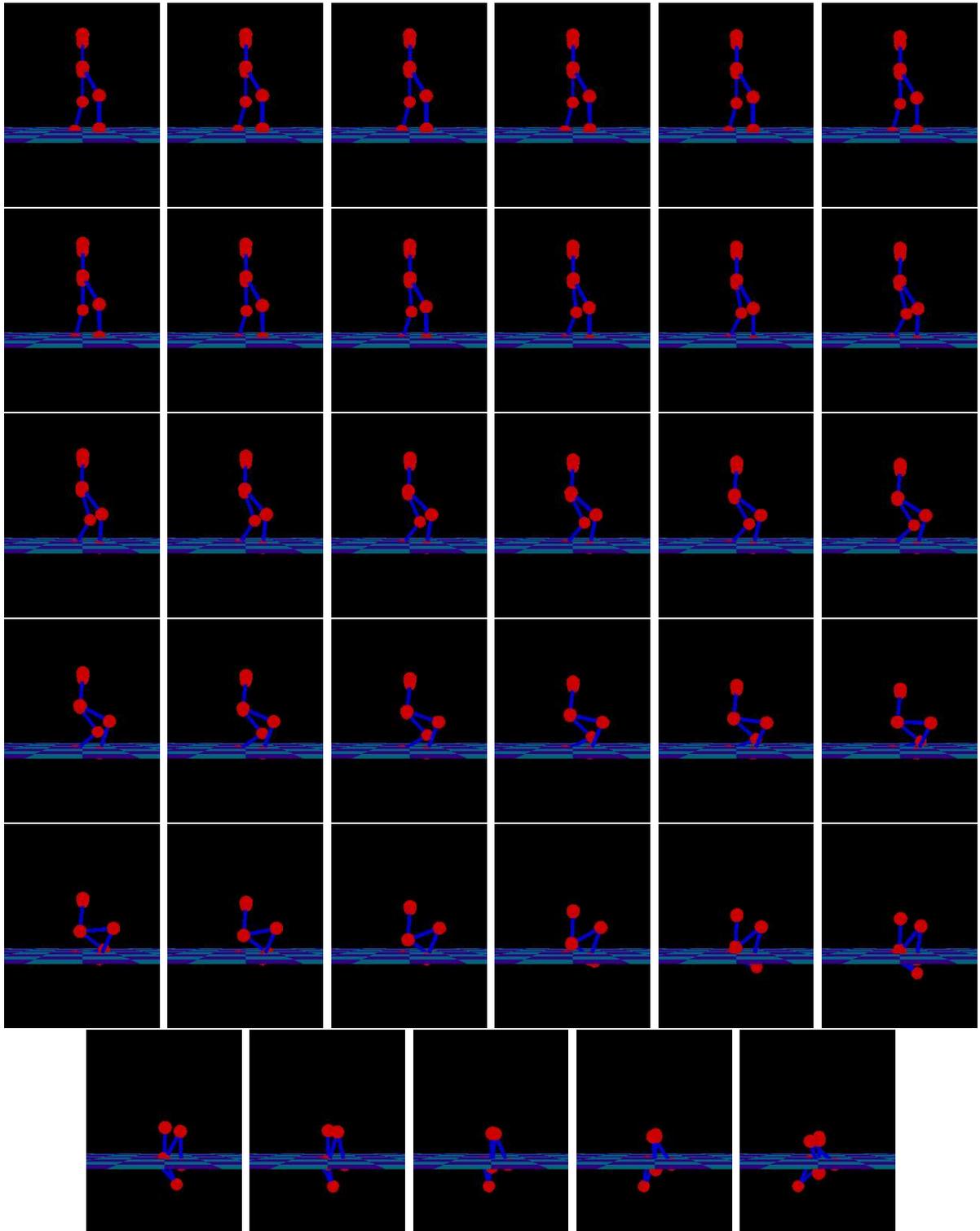


ABBILDUNG VI.1. Die Ausgangstrajektorie der Optimierung

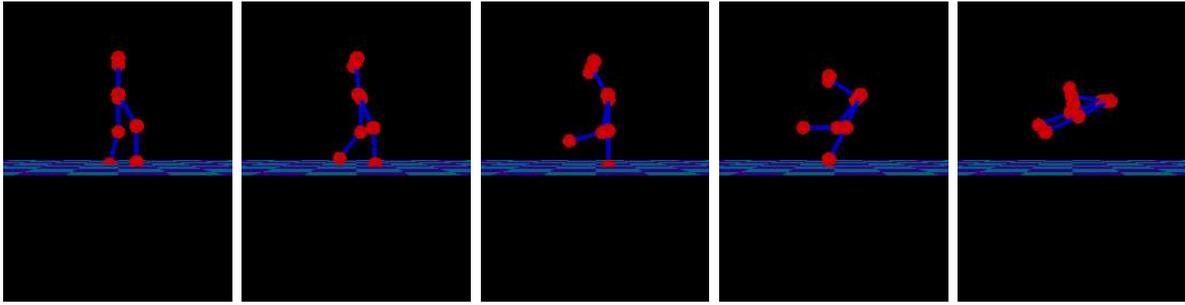


ABBILDUNG VI.2. Masse 4 möglichst weit nach rechts (Maximiere x -Koordinate des 4. Körpers)

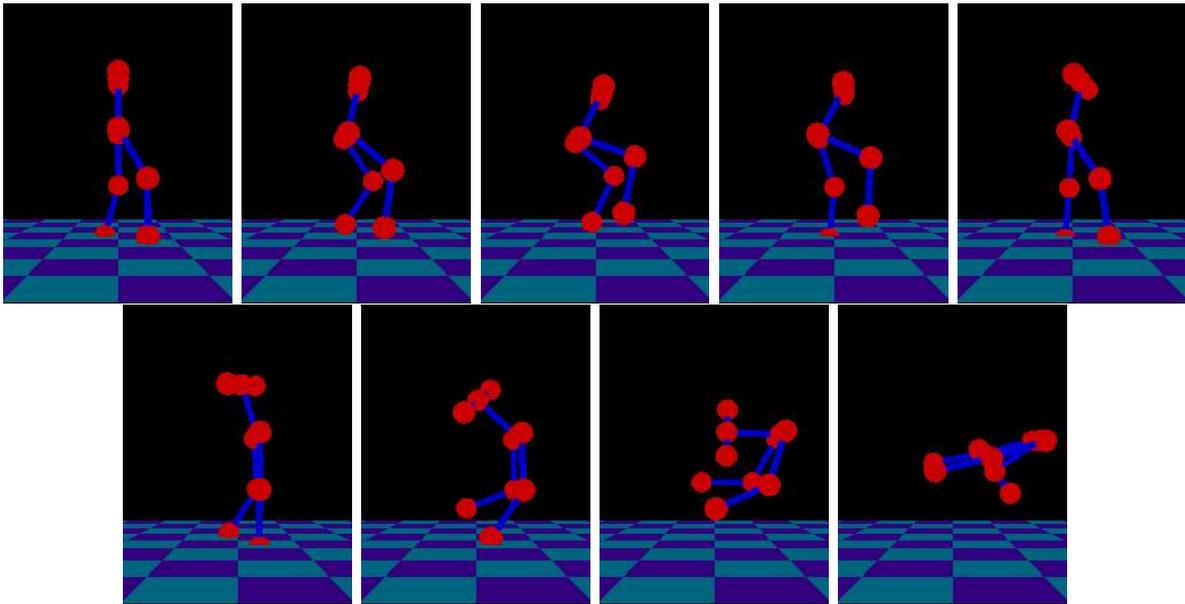


ABBILDUNG VI.3. Masse 4 möglichst weit nach rechts, nun $t_{end} = 0.2s$

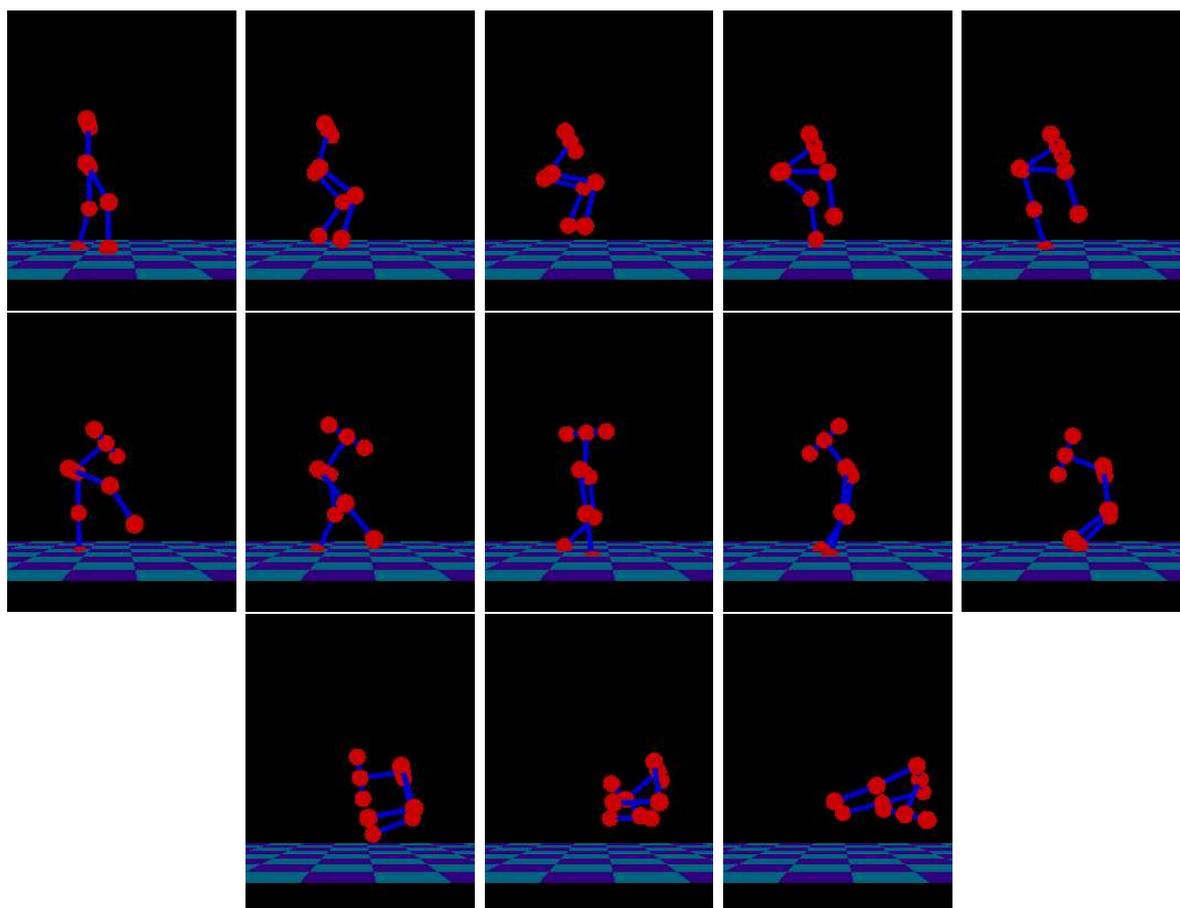


ABBILDUNG VI.4. Masse 4 möglichst weit nach rechts, nun $t_{end} = 0.3s$

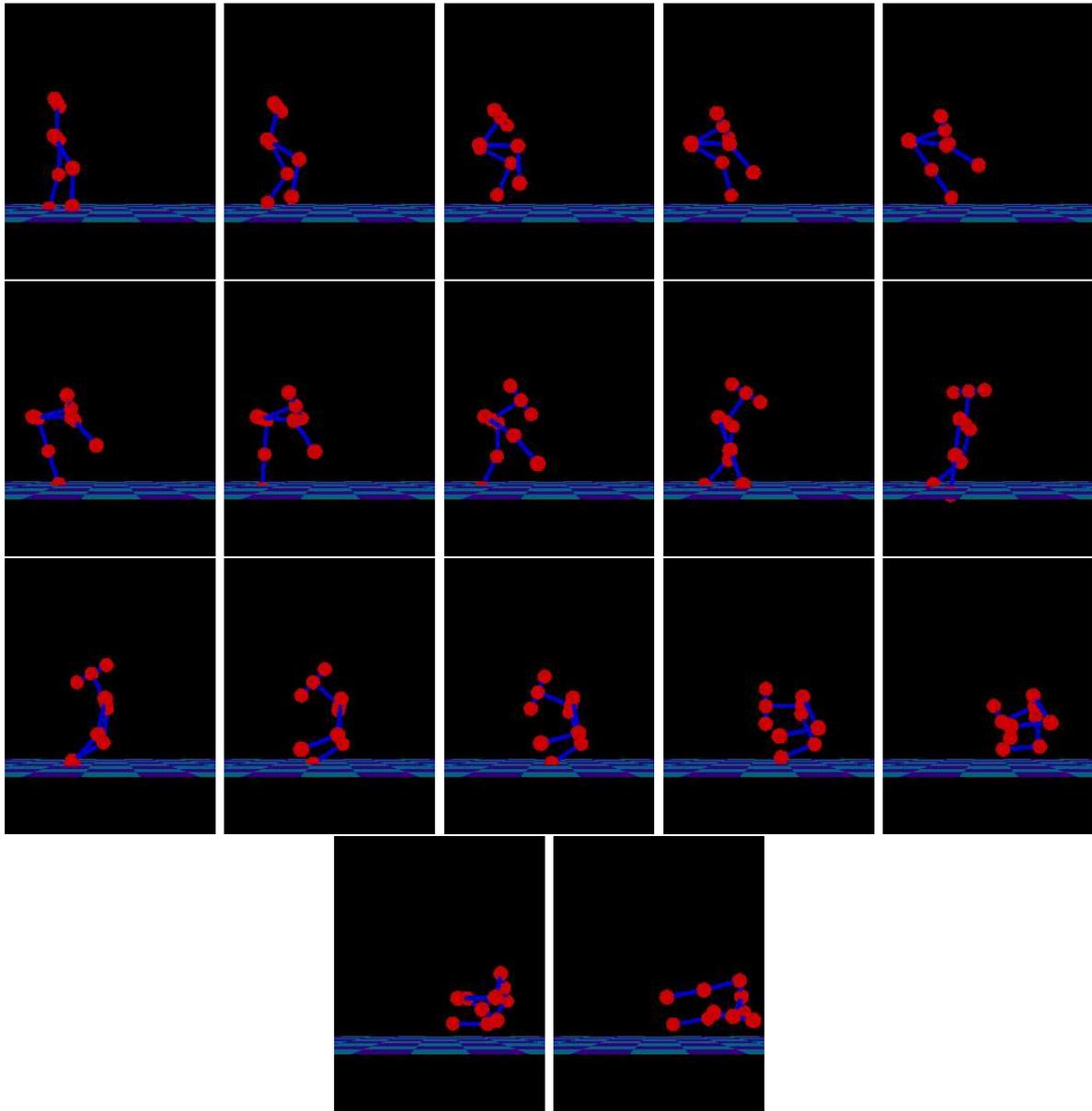


ABBILDUNG VI.5. Masse 4 möglichst weit nach rechts, nun $t_{end} = 0.4s$

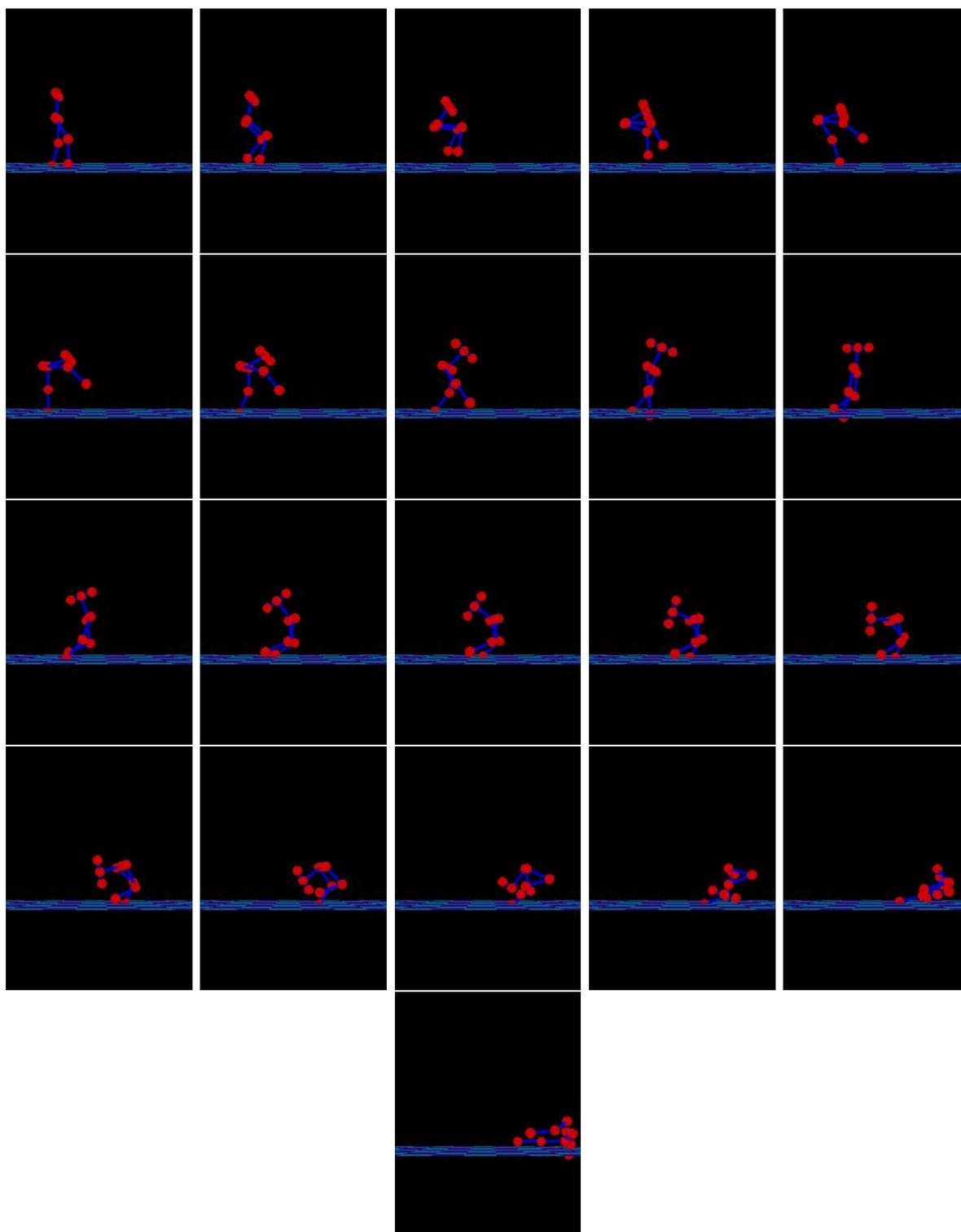


ABBILDUNG VI.6. Masse 4 möglichst weit nach rechts, nun $t_{end} = 0.5s$

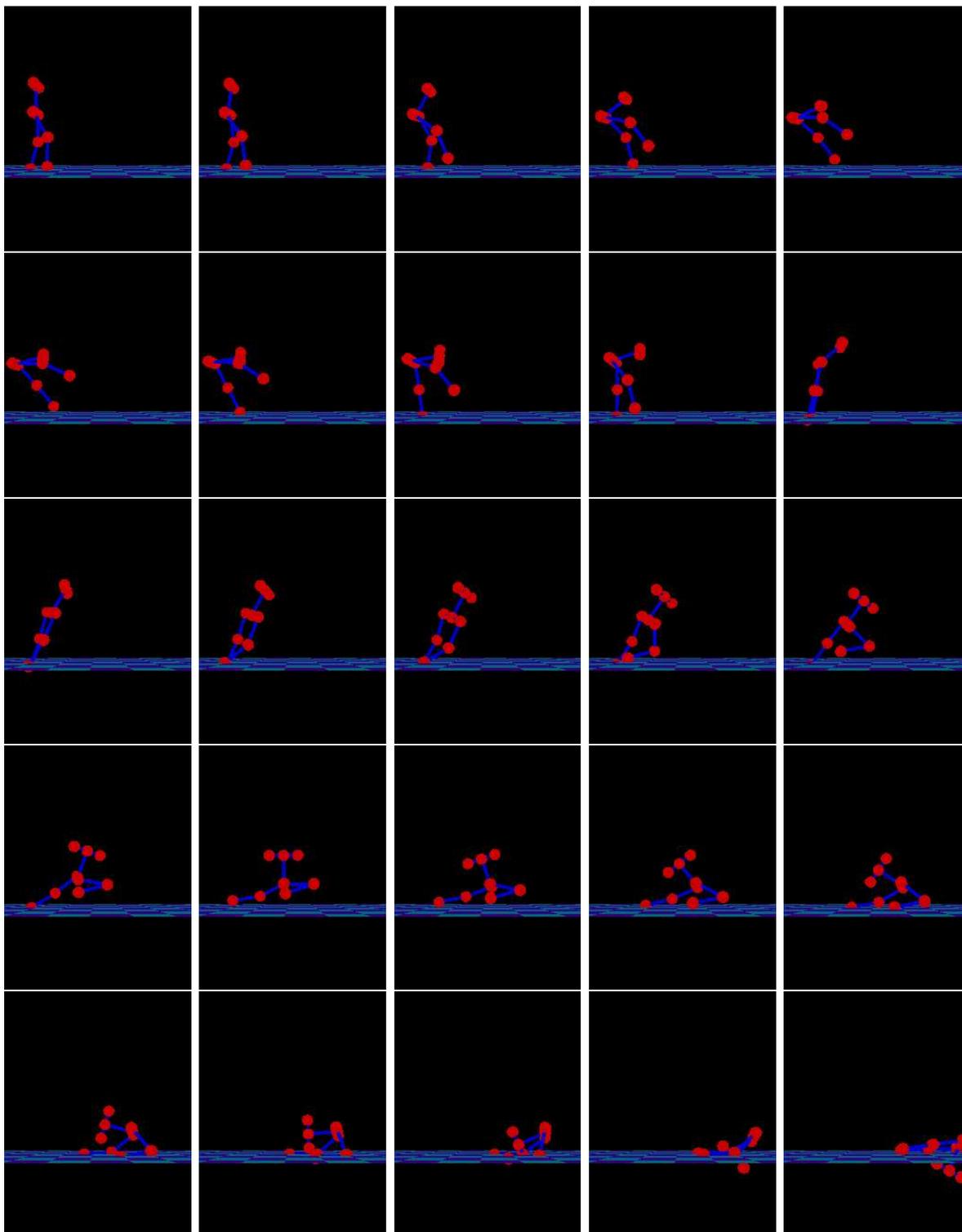


ABBILDUNG VI.7. Masse 4 möglichst weit nach rechts, nun $t_{end} = 0.6s$

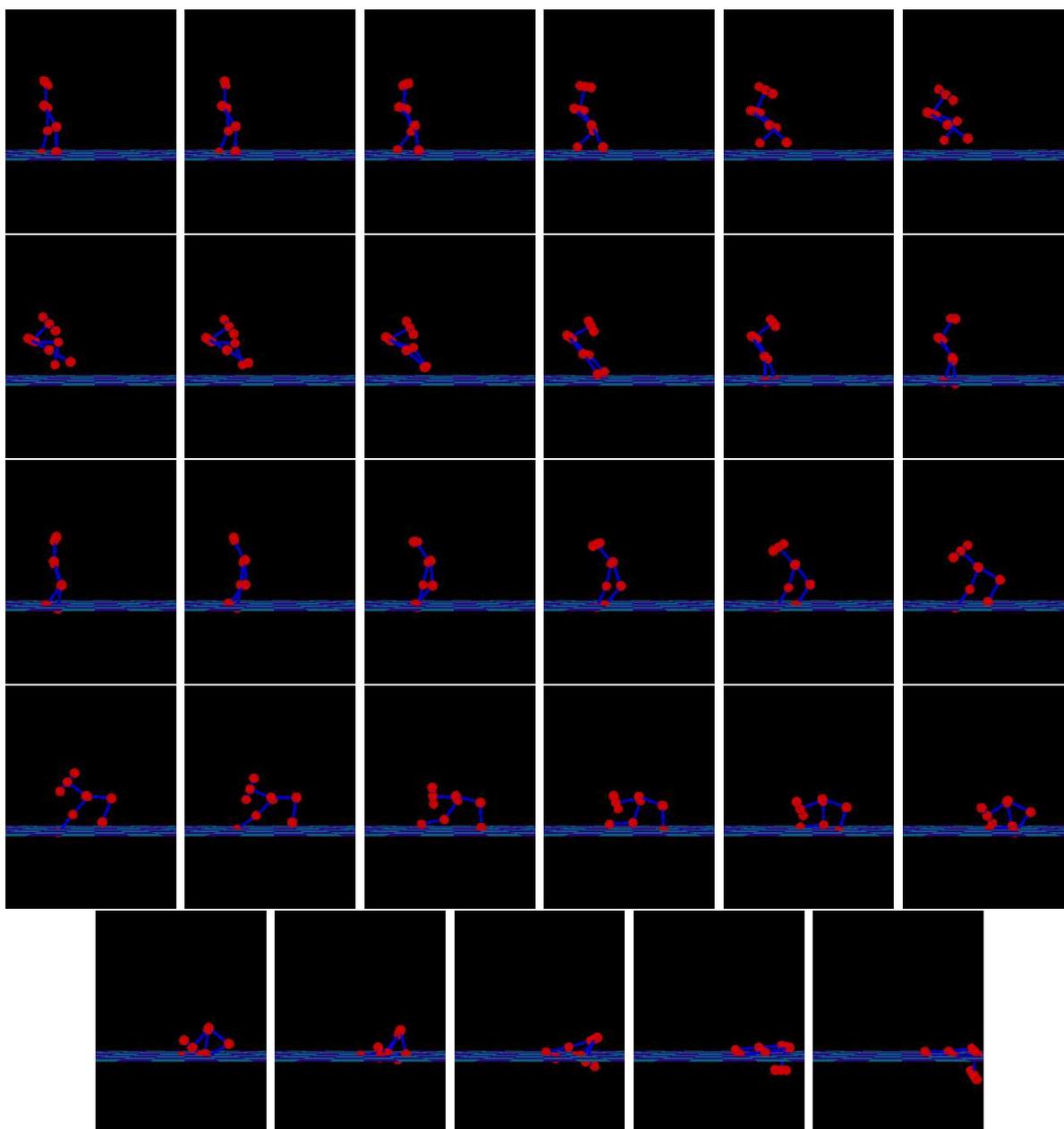


ABBILDUNG VI.8. Masse 4 möglichst weit nach rechts, nun $t_{end} = 0.7s$

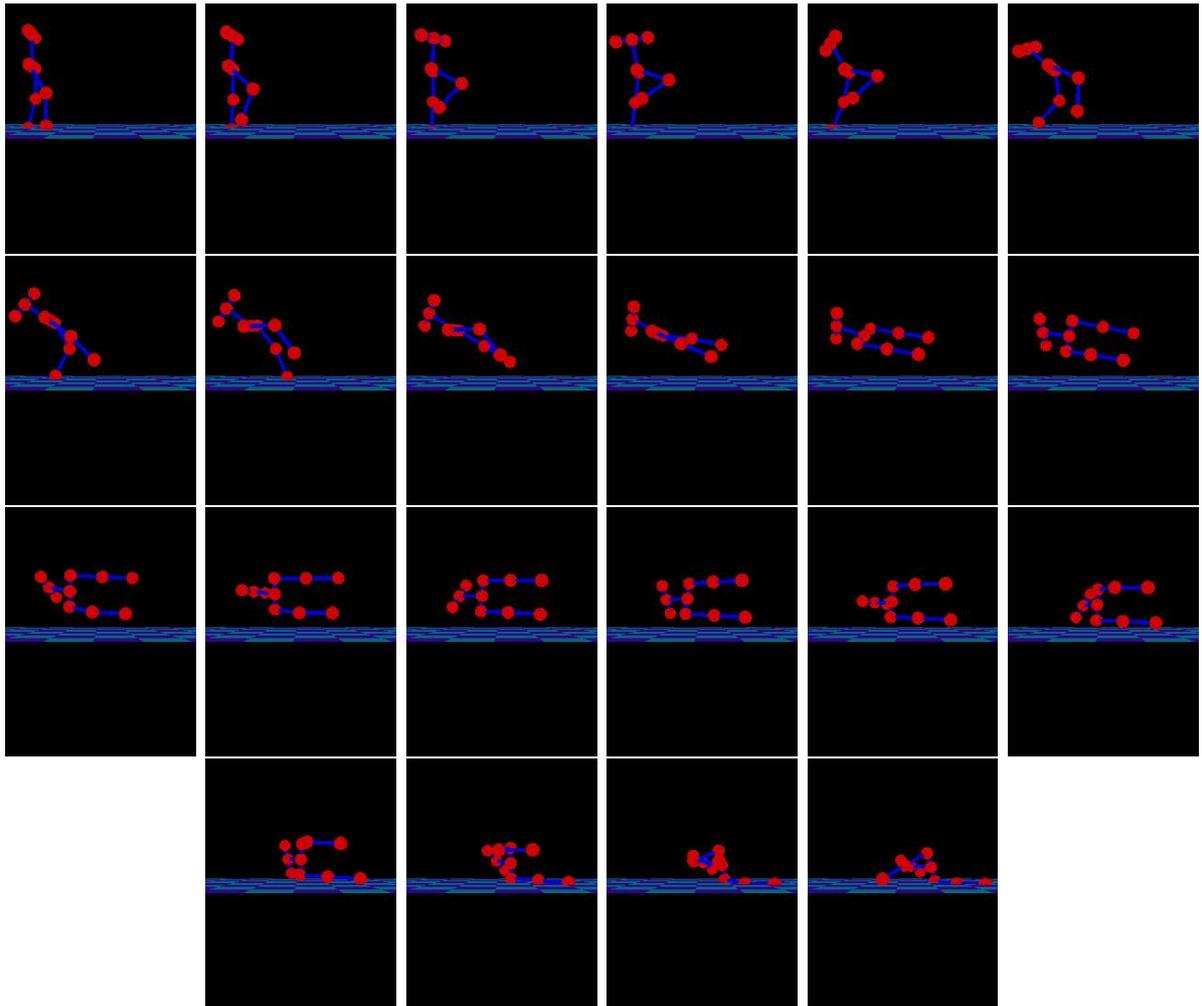


ABBILDUNG VI.9. Schwerpunkt möglichst weit nach rechts

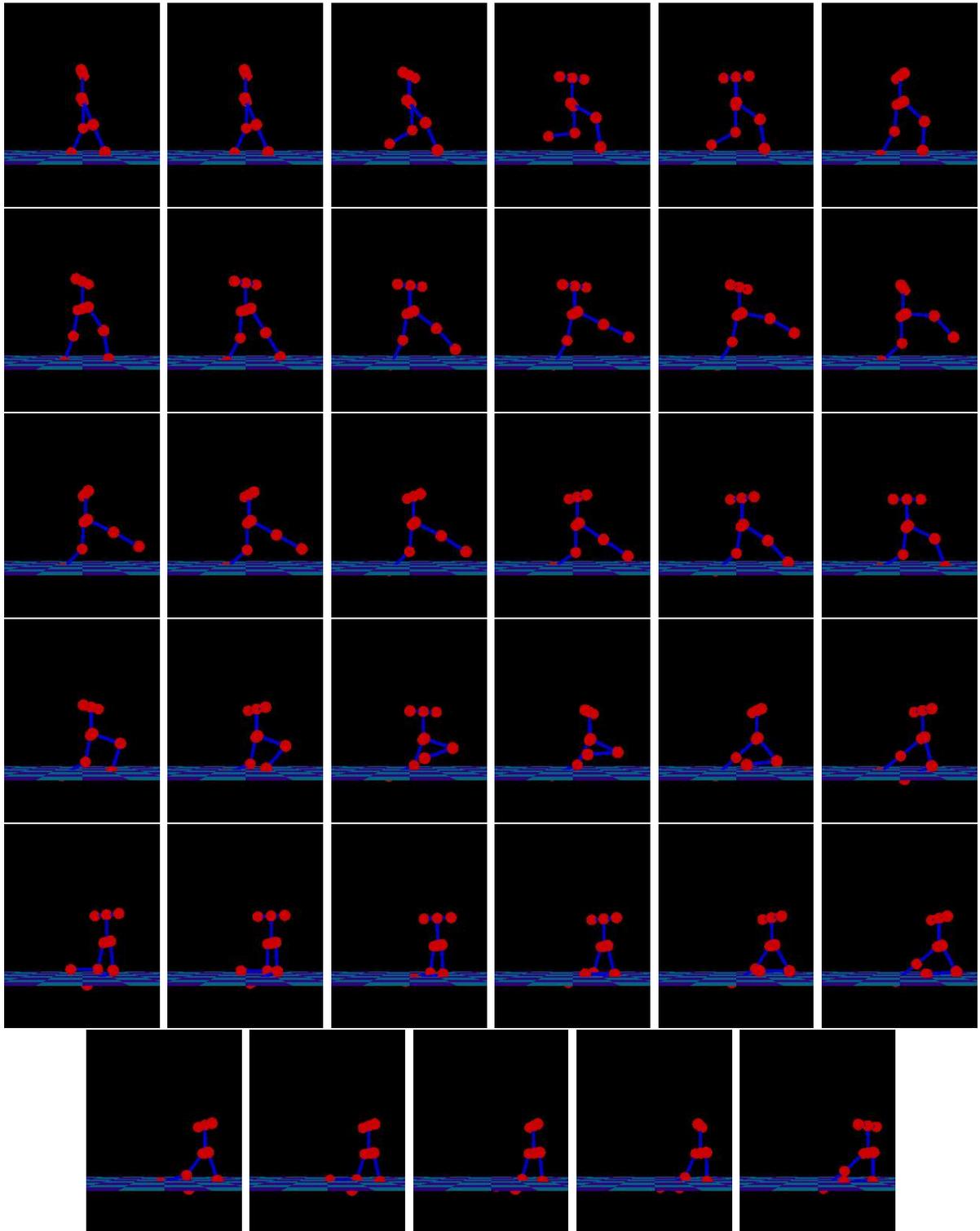


ABBILDUNG VI.10. Erste Schritte

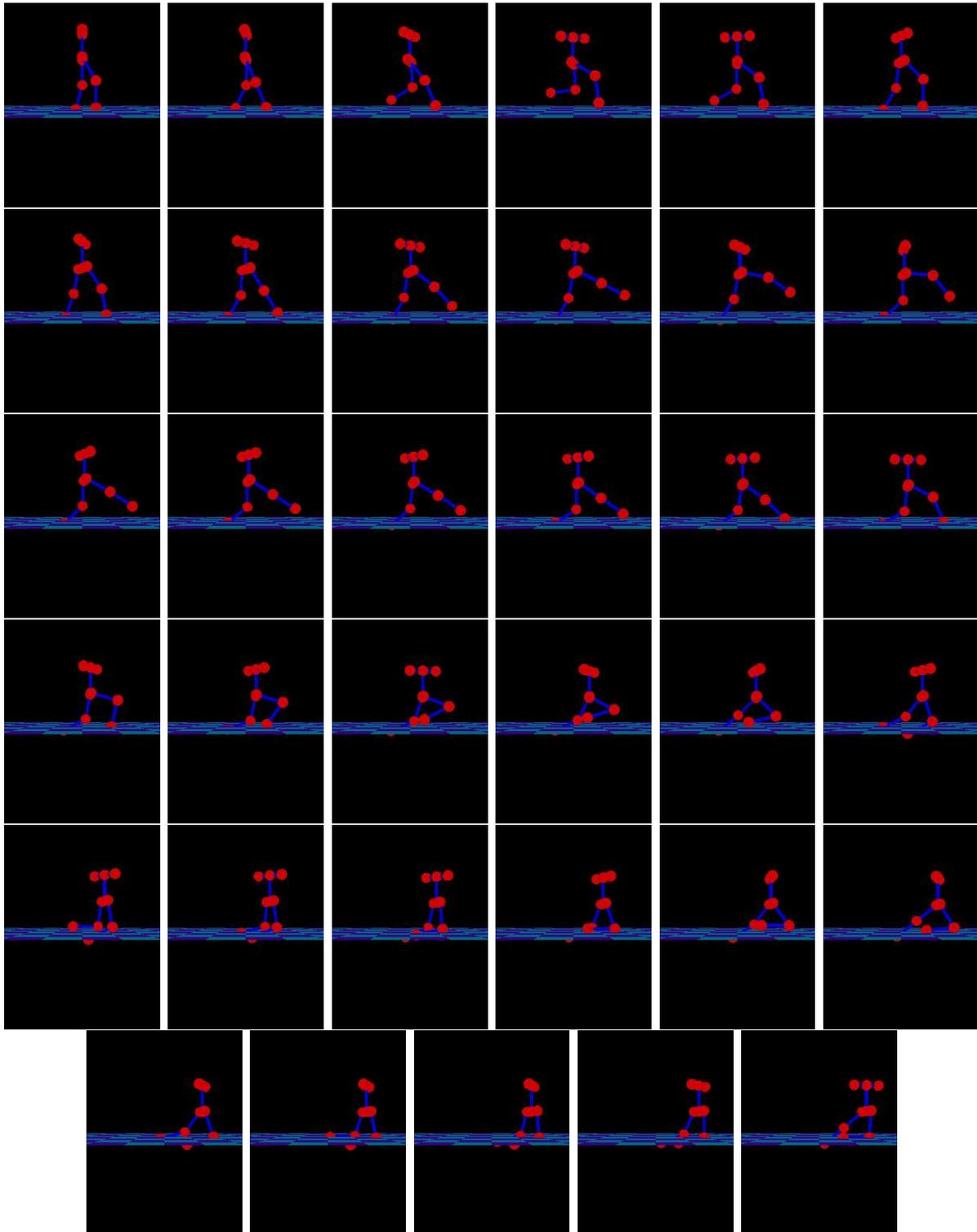


ABBILDUNG VI.11. Tiefergelegter Schwerpunkt

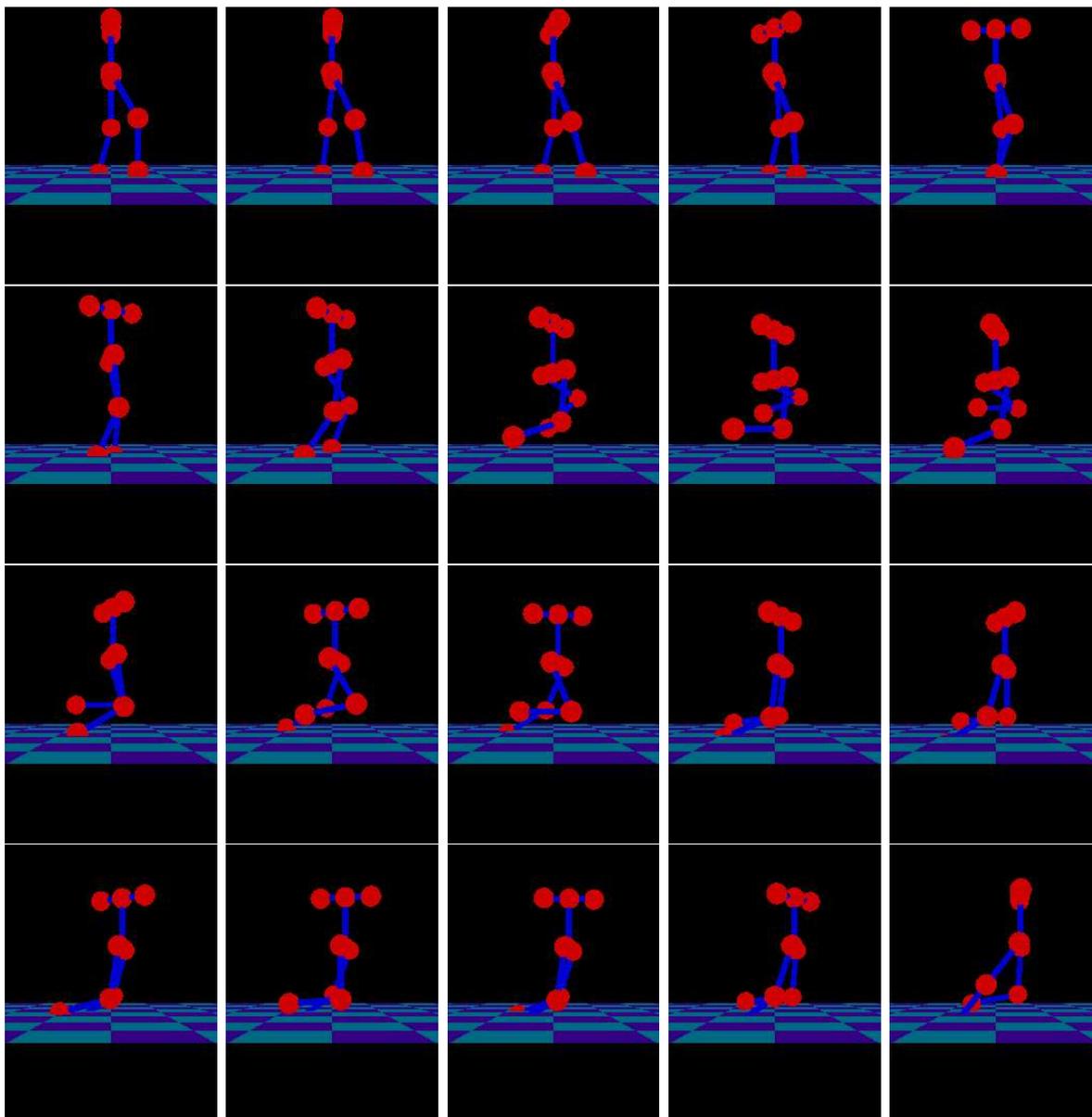


ABBILDUNG VI.12. Strenger tiefergelegter Schwerpunkt

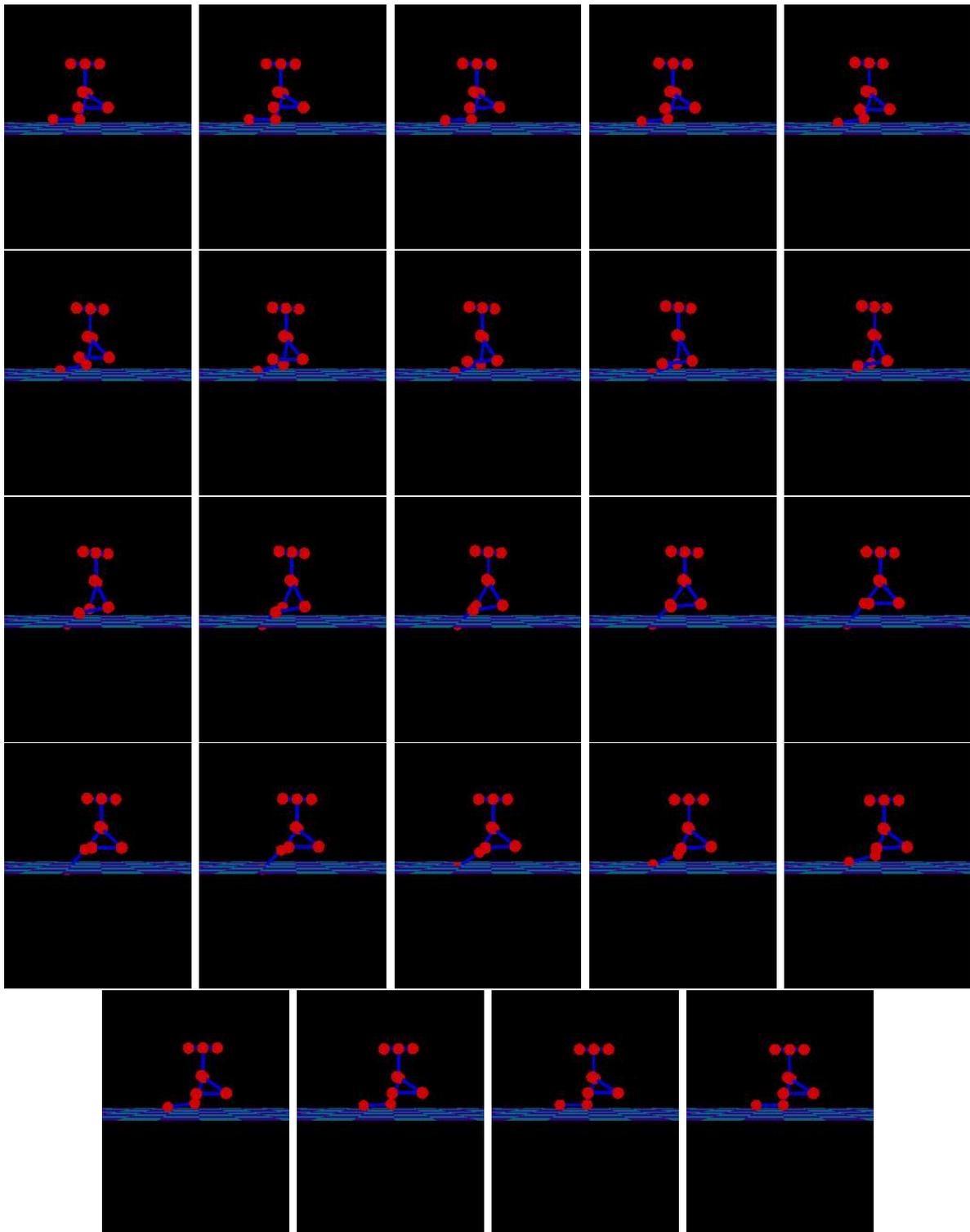


ABBILDUNG VI.13. Periodisch möglichst weit vorwärts kommen, Oberkörper senkrecht

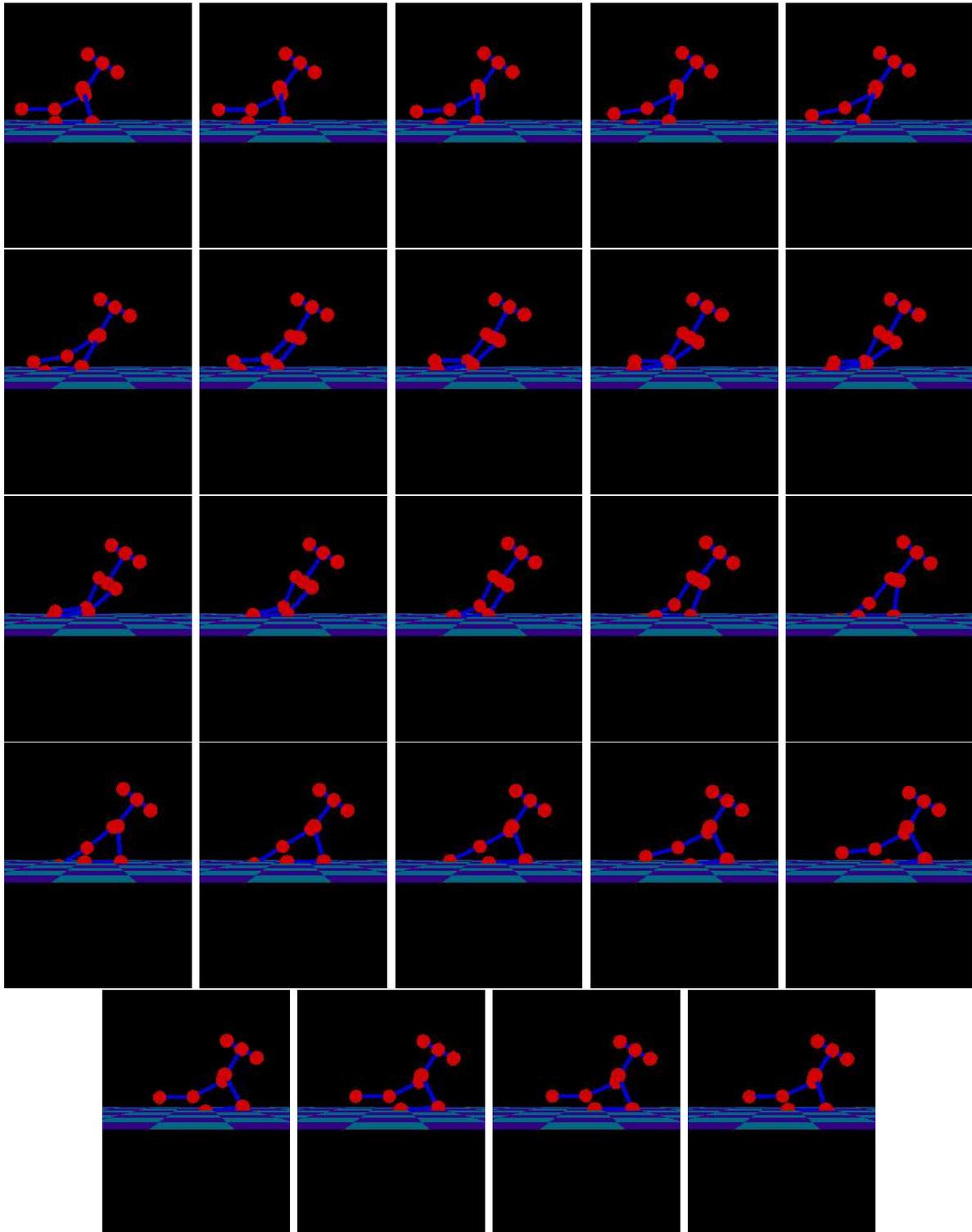


ABBILDUNG VI.14. Periodisch möglichst weit vorwärts kommen

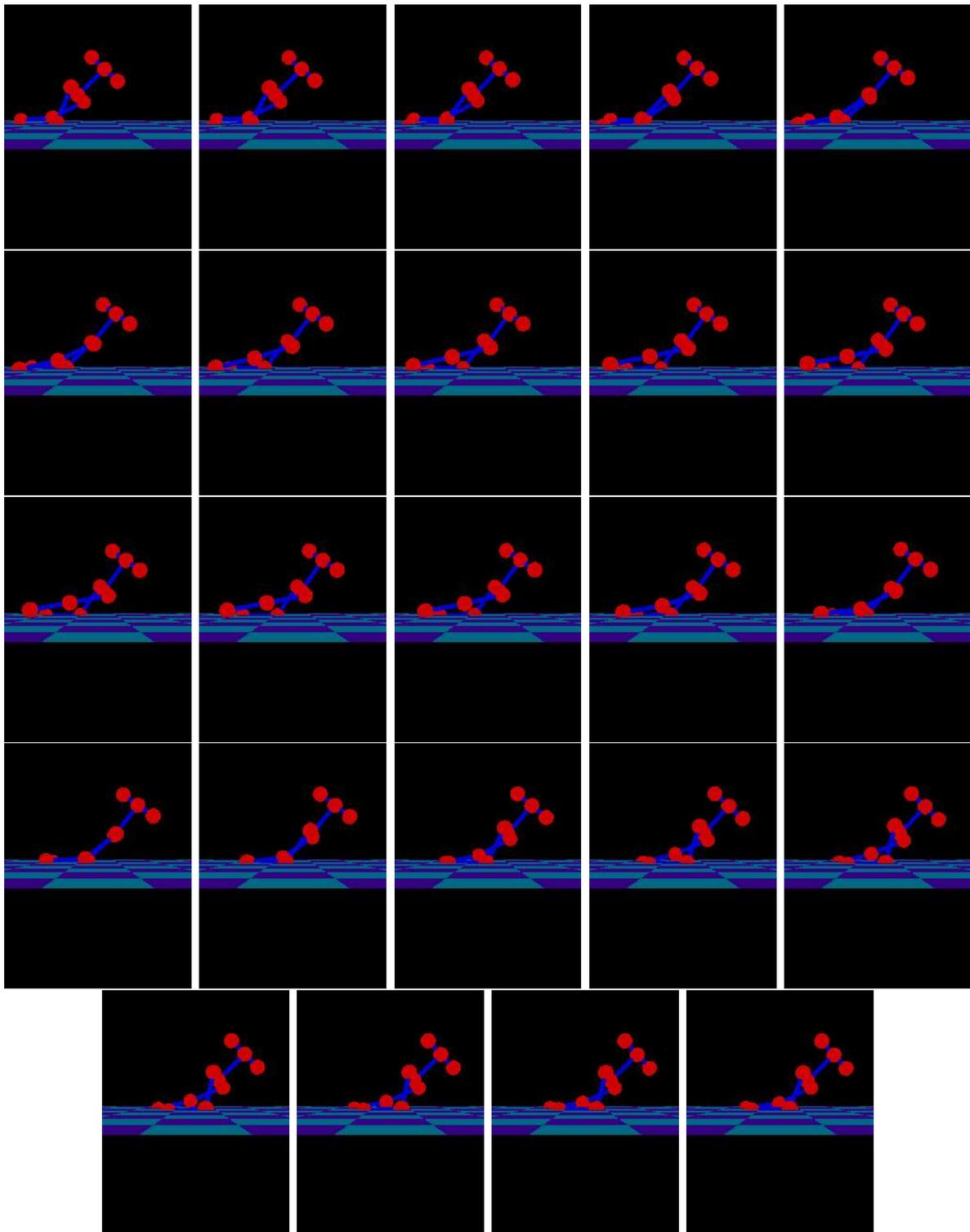


ABBILDUNG VI.15. Periodisch mit möglichst wenig Steuerung

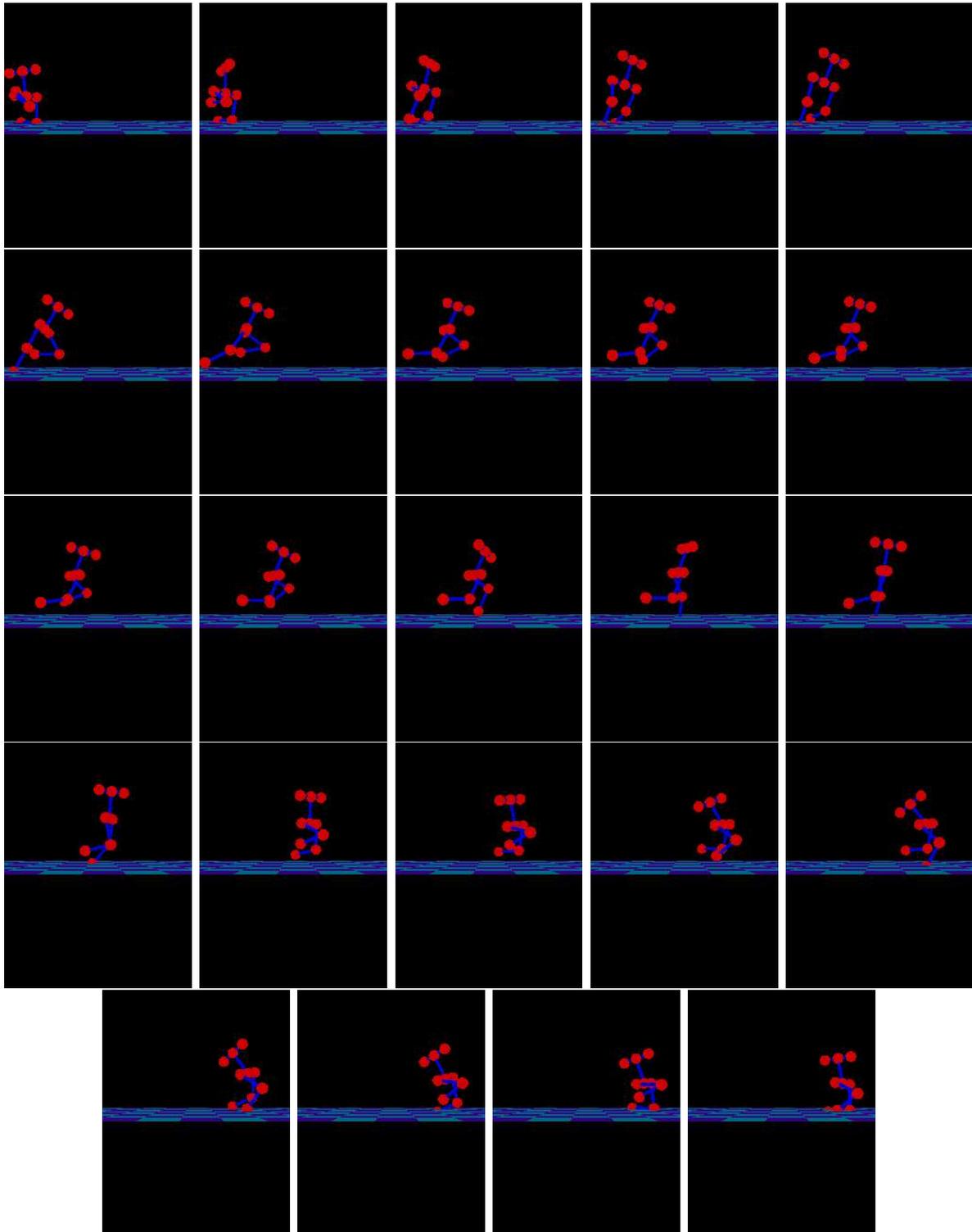


ABBILDUNG VI.16. Periodisch bei freier Endzeit

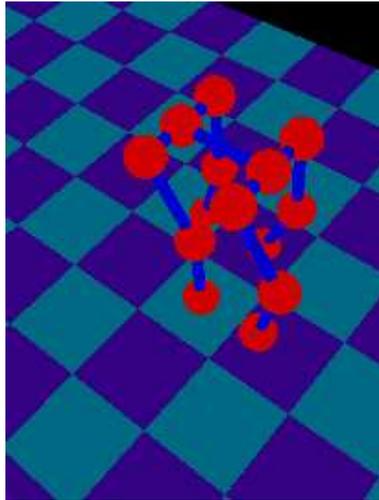


ABBILDUNG VI.17. Das vierbeinige Modell

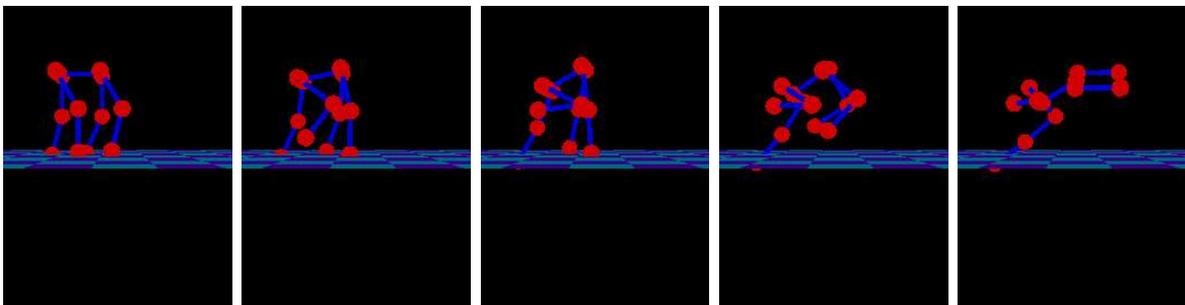


ABBILDUNG VI.18. Masse 4 möglichst weit nach rechts

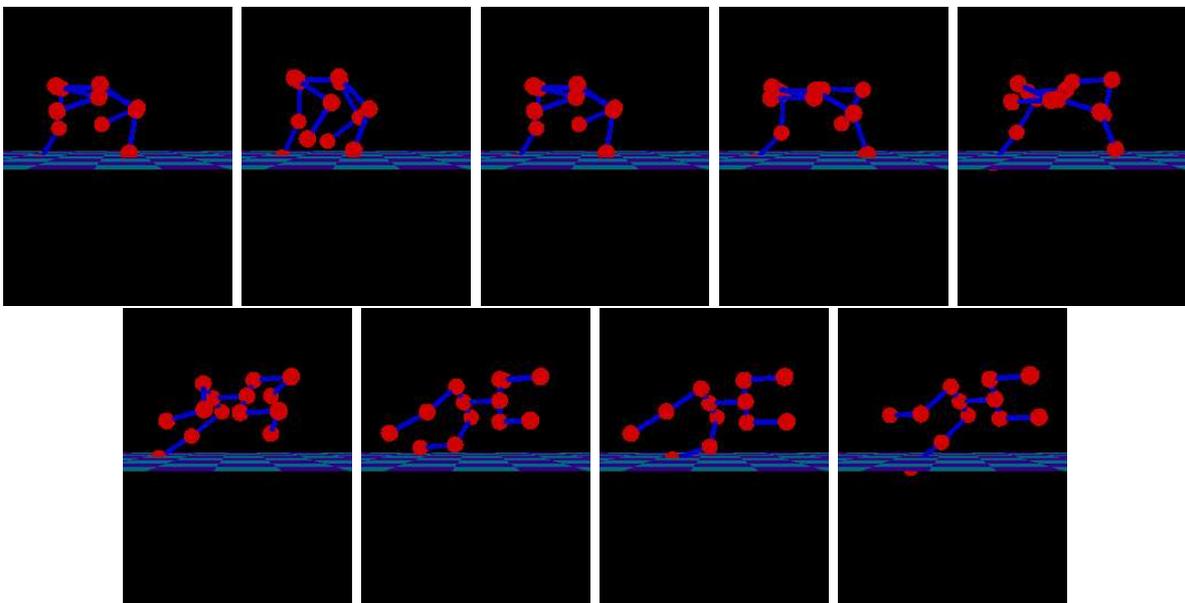


ABBILDUNG VI.19. Masse 4 möglichst weit nach rechts, Rücken waagrecht

KAPITEL VII

Zusammenfassung und Ausblick

§1 Zusammenfassung

Die im vorigen Kapitel vorgestellten optimalen Resultate erwecken in der Tat den Eindruck einer verblüffenden Natürlichkeit der Bewegung. Ausgehend von einer Nullsteuerung ergibt sich ein meist nicht vorhergesehener Bewegungsablauf. Insbesondere ist dabei die Abfolge des Boden-Fuß-Kontakts nicht vorgegeben, sondern resultiert aus der Optimierung. Das ist speziell bei einem Vier- oder Mehrbeiner von Interesse, da hier die Anzahl der möglichen Phasen deutlich größer ist als bei einem Zweibeiner und die Vorgabe der Phasen eine starke Beschränkung möglicher Bewegungen darstellen würde.

Die ermittelten Bewegungen sind zudem von hoher Durchschnittsgeschwindigkeit¹, in der Größenordnung von 10 m/s. Die Bewegungen sind also sehr „dynamisch“ und gleichen mehr einem Rennen als einem Laufen. Sehr bemerkenswert ist die Ausnutzung aller anatomischen Bewegungsfreiheiten. Dies führt meist zu einem grotesk wirkenden und unerwarteten Bewegungsablauf, zeigt aber zugleich die Stärke der Optimierung, Verbesserungen jenseits des Erwarteten zu liefern.

Das Zielfunktional der Optimierung kann dabei recht frei gewählt werden, so z. B. als zurückgelegter Weg eines Körpers, zurückgelegter Weg des Schwerpunkts oder aufintegrierte quadrierte Kontrollen. Wahlweise können Randbedingungen wie Periodizität hinzugenommen werden.

Die Modellformulierung kann komfortabel als DAE inklusive algebraischer Nebenbedingungen formuliert werden. Die Gleichungen werden vergleichbar einem Baukasten zusammengestellt und bei Bedarf erweitert. Der numerische Aufwand ist dabei nicht größer als bei einer Formulierung in Minimalkoordinaten, i. allg. sogar geringer.

Das mögliche Auftreten von Singularitäten kann durch Verwendung geeigneter Relaxierungen verhindert werden. Als weiteres Produkt ist für diese Arbeit ein Algorithmus zur Integration durch Singularitäten hindurch entwickelt worden. Die zugrundeliegende Methode, die Aufblasung, ist dabei in der algebraischen Geometrie schon lange bekannt. Doch konnte für den interessierenden Spezialfall, die Singularität auf einer Kurve zu durchqueren, ein sehr einfacher Beweis der Endlichkeit des Aufblasungs-Algorithmus gefunden werden. Dem Autor sind bislang nur solche Verfahren zur Lösung des Problems, durch Singularitäten integrieren zu müssen, bekannt, bei denen die Singularität als Störung betrachtet wird. Durch entsprechend gewählte zusätzliche Störungen wird dabei versucht, die Singularität wieder zu beseitigen. Bei diesem Vorgehen werden sowohl die Natur der Singularität ignoriert, denn die Singularität von $xy = 0$ kann nicht einfach als Entartung von $xy = \varepsilon$ betrachtet werden², als auch die Erkenntnisse des Zweigs der Mathematik, der Singularitäten auf Varietäten als zentralen Gegenstand erforscht. Diese spezielle Verbindung von Numerik und algebraischer Geometrie ist nach Wissensstand des Autors neu.

¹Dabei bezieht sich diese Durchschnittsgeschwindigkeit auf die im Zielfunktional vorkommenden Größen, d. h. es kann sich in einem Fall um die Durchschnittsgeschwindigkeit des Schwerpunkts handeln, im anderen dagegen um die eines einzelnen Körpers.

²siehe auch das Beispiel der liegenden Acht in §3 des Kapitel IV

Das hauptsächlich behandelte Modell verfügt über 22 Freiheitsgrade und fünf Kontrollen. Ausgehend von der Nullkontrolle als Starttrajektorie konnten nichtperiodische und periodische Optima bestimmt werden. Die Rechenzeiten dafür waren auf einem für heutige Maßstäbe schon schwachen PC (PentiumII, 266 MHz) in der Größenordnung von Tagen und damit durchaus akzeptabel.

§2 Vergleich

Im Bereich des digitalen Films werden natürliche Bewegungen ausschließlich von Künstlern produziert. Moderne Animationstools wie `SOFTIMAGE3D` oder `MAYA` verfügen zwar über Methoden, um den Animateur zu unterstützen, diese Methoden beschränken sich aber auf Inverse Kinematik oder Inverse Dynamik³. Dabei wird in einem hierarchisch angeordneten Mehrkörpersystem die Bewegung eines bestimmten Körpers in seiner Bahn vorgegeben, die Programme errechnen dann daraus kinematisch bzw. dynamisch zulässige Bewegungen der anderen Körper. Bei der ersten Methode ergeben sich meist physikalisch unzulässige Bewegungen, bei der zweiten (schwierigeren) Methode der Inversen Dynamik sind die errechneten Kraft- und Momentverläufe häufig unrealistisch. Außerdem kann i. allg. nicht die Existenz einer Lösung des so gestellten inversen Problems gesichert werden.

Der überwältigende Anteil von Untersuchungen, die sich mit Laufbewegungen von Mensch, Tier oder insbesondere Roboter beschäftigen, befaßt sich mit regelungstechnischen Fragen. Es wird nach geeigneten Sensoren gesucht sowie nach brauchbaren Methoden, aus den Messungen dieser Sensoren den Roboter zu steuern. Bevorzugtes Ziel der Steuerung ist dabei Stabilität in einem allgemeinen Sinn. So soll der Roboter nicht umfallen, aber zugleich eine gewisse Wegstrecke zurücklegen. Seine Bewegung soll zudem noch „fortsetzbar“ sein, d. h. er soll nicht in ein Endstadium geraten, aus dem ein Weiterlaufen nicht mehr möglich ist. Wesentliche Kriterien sind hier Echtzeitfähigkeit, Einfachheit und Unversehrtheit.⁴ Diese regelungstechnischen Themen führen zu anderen als den hier behandelten Fragestellungen, da die vorliegende Arbeit nicht-rückgekoppelte Systeme betrachtet. Aber es gibt einen gewissen Berührungspunkt, denn sind die optimierten Laufbewegungen eines Mechanismus bekannt⁵, so könnte man versuchen diese zu benutzen, um daraus Steuerungsregeln abzuleiten. Die Literatur bezüglich dieser rückgekoppelten Systeme ist sehr umfangreich. Es sei daher lediglich auf die Netzadresse des DFG-SCHWERPUNKTS „AUTONOMES LAUFEN“ [17] verwiesen, um dort umfangreiche Literaturverweise zu finden. In den Arbeiten von CHEVALLEREAU UND Aoustin [14] sowie von HARDT ET AL. [31] wird die Phasen-Sequenz der Fuß-Boden-Kontakte fest vorgegeben. Bei Zweibeinern ist die Anzahl der Möglichkeiten nicht allzu hoch, bei Vierbeinern jedoch bereits beträchtlich. Das Problem, das Optimum über alle möglichen Phasen-Sequenzen zu finden, konnte bisher nicht gelöst werden⁶. Hier behilft man sich damit, daß man (im Falle eines Vierbeiners) die von der Natur gefundenen Abläufe bei Tieren als Ausgangspunkt nimmt. Eine zum vorliegenden Thema verwandte Untersuchung findet sich bei HARDT [30]. Hier

³siehe z. B. IMMLER [36]

⁴Diese Unversehrtheit hat höchste Priorität in einem der Projekte des DFG-SCHWERPUNKTS „AUTONOMES LAUFEN“ [17], der Entwicklung einer Stand-Gang-Prothese. Diese soll einem querschnittsgelähmten Menschen z. B. das Aufstehen aus einem Stuhl ermöglichen. Hierbei muß jedwedem Umfallen verhindert werden.

⁵Insbesondere die in Hinsicht auf Stabilität optimierten Bewegungen, zu deren Ermittlung man die von Mombauro [48] erarbeiteten Methoden verwenden kann.

⁶siehe HARDT UND VON STRYK [32]

wird die optimale Bewegung eines Zweibeiners untersucht. Dabei wird die Mehrzielmethode und ein SQP-Verfahren angewandt. Allerdings ist das Modell nur zweidimensional, die Phasen des Fuß-Boden-Kontakts sind fest vorgegeben (nicht nur in ihrem Ablauf, sondern auch in ihrer relativen Länge) und mit wechselnden algebraischen Nebenbedingungen formuliert. Für den Optimierer werden die Bewegungsgleichungen durch eine rekursive Formulierung in ein gewöhnliches Differentialgleichungssystem verwandelt. Das Zielfunktional besteht aus den aufintegrierten, quadrierten Kontrollen und Drehmomenten und ist sehr konvex. Als Ergebnis resultiert eine Laufbewegung, deren Durchschnittsgeschwindigkeit auf 0.84 m/s festgelegt ist. Die Bewegungsgeschwindigkeit liegt damit um eine Größenordnung niedriger als die Ergebnisse in Kapitel VI.

Bei HARDT UND VON STRYK[33] werden periodische Bewegungen eines Vierbeiners in Hinsicht auf Stabilität optimiert. Die Bewegungsgleichungen sind dort wie hier DAEs, für die Optimierung allerdings werden die Zustandsvariablen mit Hilfe einer festen Transformation auf unanhängige Variablen reduziert („reduced dynamics“, „coordinate partitioning“, siehe dazu auch WEHAGE[67]). Dies hat den Zweck, für die Optimierung die DAE in eine ODE umzuwandeln. Dieses Auffinden von unabhängigen Koordinaten kann jedoch nicht algorithmisch geschehen und ist zudem noch mit numerischen Schwierigkeiten (Fehlerfortpflanzung, Umpartitionierung) verbunden. Die initiale Projektion erfüllt daher diesen Zweck besser. Als Ergebnis der Optimierung bezüglich eines komplizierten Stabilitätskriteriums von KOO UND YOON[37] resultiert eine Durchschnittsgeschwindigkeit von 0.67 m/s. Die im Kapitel VI gefundenen Durchschnittsgeschwindigkeiten bei Zwei- oder Vierbeiner (aber bezüglich anderer Zielfunktionale) liegen dagegen bei etwa 6-15 m/s und weisen damit eine andere Größenordnung auf.

In der Arbeit von GERDTS[26] werden Optimierungsprobleme mit darunterliegender DAE ebenfalls mit Hilfe von Mehrzielmethode und SQP-Verfahren behandelt. Es wird eine Projektionsmethode angewandt, um konsistente Initialwerte für die Integration der DAE zu erhalten. Allerdings wird diese Methode nicht dazu benutzt, die Anzahl der durchzuführenden Integrationen für die Berechnung der Sensitivitätsmatrix zu reduzieren. Die Anwendungen beziehen sich dann aber auf ein ganz anderes Gebiet, nämlich Fahrzeugsimulationen.

Methodisch am nächsten verwandt sind die Arbeiten von MOMBAUR[48]-[52]. Hierbei werden Laufbewegungen von 2D- und 3D-Modellen in Hinsicht auf Stabilität untersucht, so ein 2D-Einbein (aktiv gesteuert), ein 2D-Zweibein (aktiv gesteuert) und ein 3D-Zweibein (passiv, d. h. es gibt keine Kontrollen). Diese Ein- und Zweibeiner sind nicht-rückgekoppelt. Im Unterschied zur vorliegenden Arbeit werden die Modelle allerdings in Minimalkoordinaten formuliert. Betrachtet werden ausgehend von einer periodischen Bewegung $x(t)$ mit Periode T die Eigenwerte der Monodromiematrix

$$C = \frac{\partial x(t+T)}{\partial x(t)}.$$

Ist der größte Eigenwert dieser Matrix C oder einer Potenz C^n vom Betrag her kleiner als 1, so liegt eine stabile Bewegung vor, d. h. eine Bewegung, bei der Störungen der Anfangswerte abklingen. Das Optimierungsproblem besteht nun darin, Kontrollen bzw. Designparameter zu finden, so daß dieser größte Eigenwert vom Betrag her kleiner 1 ist. Der Fuß-Boden-Kontakt ist nicht als Feder-Dämpfer-Kombination modelliert, sondern als Phase mit geringerer Anzahl von Freiheitsgraden. Da es dadurch zu Undifferenzierbarkeiten in den Zustandsvariablen kommt, ist es notwendig, den Phasenablauf des Bodenkontakts bereits fest vorzugeben (wenn auch nicht in seiner zeitlichen Dauer).

Das nicht-differenzierbare Zielfunktional führt zu anderen Fragestellungen und erfordert

andere Optimierungsmethoden mit einer äußeren und einer inneren Optimierungsschleife und die Verwendung des Nelder-Mead-Algorithmus⁷.

§3 Ausblick

Die vorliegende Untersuchung kann in mehrere Richtungen vertieft werden.

So besteht die Möglichkeit die numerische Integration zu parallelisieren. Das simulierte Zeitintervall ist durch den Mehrzielansatz in mehrere Knoten unterteilt, somit kann die Integration innerhalb eines solchen Knotens von jeweils einem Prozessor übernommen werden. Da die Integration den größten Teil der Rechenzeit ausmacht, ist der zeitliche Gewinn enorm, der Beschleunigungsfaktor entspricht etwa der Anzahl der eingesetzten Prozessoren.

In der vorliegenden Formulierung hat der Boden einen Teil der Funktion des Fußes übernommen, nämlich die elastische Deformierung. Die Parameter des Bodens als Feder-Dämpfer-Element könnten variiert werden, um die Auswirkungen von Weichheit oder Härte auf die optimale Bewegung zu untersuchen. Hier kann man auch versuchen Parameter zu finden, die die Elastizität von realen Füßen besser wiedergeben.

Die vielfältigsten Möglichkeiten bieten natürlich Veränderungen am Modell und Zielfunktional. Mechanismen mit mehr Beinen, mit oder ohne Knie, mit verschiedenen orientierten Bewegungsbeschränkungen z. B. des Knies, komplexere Aktuatoren mit realistischeren Kräfteverläufen verknüpft mit einem reichhaltigen Spektrum an Zielfunktionalen versprechen weitere ungewöhnliche und unerwartete Bewegungsabläufe, die das Wissen um die Natur des Laufens bereichern könnten.

⁷siehe auch NELDER UND MEAD[53]

Literaturverzeichnis

- [1] Bauer, I., *Numerische Behandlung Differentiell-Algebraischer Gleichungen mit Anwendungen in der Chemie*, Diplomarbeit, Universität Augsburg, 1994
- [2] Bauer, I., *Numerische Verfahren zur Lösung von Anfangswertaufgaben und zur Generierung von ersten und zweiten Ableitungen mit Anwendungen bei Optimierungsaufgaben in Chemie und Verfahrenstechnik*, IWR Preprint 2001-28, Universität Heidelberg, 2001
- [3] Baumgarte, J., *Stabilization of Constraints and Integrals of Motion in Dynamical Systems*, in Computer Methods in Applied Mechanics and Engineering, Ausgabe 1, 1972
- [4] Becker, T. und Weispfennig, V., *Gröbner Bases*, Springer Verlag, 1993
- [5] Berns, K., *Walking Machines Catalogue*, Internetquelle <http://www.fzi.de/ipt/WMC/walking-machines-katalog/walking-machines-katalog.html>, 2000
- [6] Bock, H.-G., *Zur numerischen Behandlung zustandsbeschränkter Steuerungsprobleme mit Mehrzielmethode und Homotopieverfahren*, Zeitschrift für angewandte Mathematik und Mechanik 57, 1977
- [7] Bock, H.-G., *Numerical Solution of Nonlinear Multipoint Boundary Value Problems with Applications to Optimal Control*, Z. Angew. Math. Mech. 58, 1978
- [8] Bock, H.-G., *Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen*, Dissertation, Universität Bonn, 1985
- [9] Bock, H.-G. und Plitt, K.-J., *A Multiple Shooting Algorithm for Direct Solution of Optimal Control Problems*, International Federation of Automatic Control, 9th World Congress, Budapest, 1984
- [10] Bock, H.-G., Schlöder, J., von Schwerin, R., Winckler, M., Lindner, J., *Entwicklung mathematischer Software für die Gesamtfahrzeugsimulation*, in HOFFMANN ET AL.[35]
- [11] Bolotnik, N.N., Kostin, G.V., Chernousko, F.L., *Modeling and Optimization of Motion of a Tube-crawling Robot*, Mechanics of Solids, Vol. 31, No. 3, 1996
- [12] Buchberger, B., *Groebner Bases: an Algorithmic Method in Polynomial Ideal Theory*, in *Multidimensional Systems Theory*, N. K. Bose (ed.), D. Reidel Publishing Company, Dordrecht, 1985
- [13] Bulirsch, R., *Die Mehrzielmethode zur numerischen Lösung von nichtlinearen Randwertproblemen bei Aufgaben der optimalen Steuerung*, Manuskript, Carl-Cranz-Gesellschaft, 1971
- [14] Chevallererua, C., Aoustin, Y., *Optimal reference trajectories for walking and running of a biped robot*, Robotica 19, 2001
- [15] Cox, D., Little, J. und O'Shea, D., *Ideals, Varieties and Algorithms*, Second Edition, Springer, 1996
- [16] Denoth, J., *Load on the Locomotor System and Modelling*, in: Nigg, B.M.(Hrsg.), *Biomechanics of Running Shoes*, Champaign 1986
- [17] DFG-Schwerpunkt „Autonomes Laufen“, http://www.fzi.de/ipt/dfg_schwerpunkt_laufen/start_page.html, 1998
- [18] Encinas, S., *A New Theorem of Desingularization over Fields of Characteristic Zero*, Universidad de Valladolid, 2001
- [19] Featherstone, R., *Robot Dynamics Algorithms*, Kluwer Academic, Publ., Boston, 1987
- [20] Fletcher, R., *Practical Methods of Optimization*, Wiley, 1987
- [21] Freitag, E., Busam, R., *Funktionentheorie*, Springer, 1993
- [22] Gallitzendörfer, J., *Parallele Algorithmen für Optimierungsrandwertprobleme*, Dissertation, Universität Heidelberg, 1996
- [23] Gao, J., *Ein Beitrag zur stoßfreien Gehbewegung*, Fortschritts-Berichte VDI Reihe 18, Nr. 114, VDI-Verlag 1992
- [24] Garcia de Jalón, J., Bayo, E., *Kinematic and Dynamic Simulation of Multibody Systems*, Springer, 1994
- [25] Garcia de Jalón, J., Unda, J., Avello, A., *Natural Coordinates for the Computer Analysis of Multibody Systems*, Computer Methods in Applied Mechanics and Engineering, Ausgabe 56, 1986

- [26] Gerdts, M., *Numerische Methoden optimaler Steuerprozesse mit differential-algebraischen Gleichungssystemen höheren Indexes und ihre Anwendungen in der Kraftfahrzeugsimulation und Mechanik*, Bayreuther Mathematische Schriften 61, 2001
- [27] Gill, P.E., Murray, W., Wright, M.H., *Practical Optimization*, Academic Press, 1981
- [28] Großmann, Ch., Terno, J., *Numerik der Optimierung*, B. G. Teubner, Stuttgart, 1997
- [29] Hairer, E., Wanner, G., *Solving Ordinary Differential Equations II*, Band 14 Springer Series in Computational Mathematics, Springer, Heidelberg, 1991
- [30] Hardt, M., *Multibody Dynamical Algorithms, Numerical Optimal Control, with Detailed Studies in the Control of Jet Engine Compressors and Biped Walking*, Dissertation, University of California, San Diego, 1999
- [31] Hardt, M., Helton, J.W., Kreutz-Delgado, K., *Optimal Biped Walking with a Complete Dynamical Model*, IEEE Conference on Decision and Control, 1999
- [32] Hardt, M., von Stryk, O., *Towards Optimal Hybrid Control Solutions for Gait Patterns of a Quadruped*, CLWAR: International Conference on Climbing and Walking Robots, 2000
- [33] Hardt, M., von Stryk, O., *Increasing Stability in Dynamic Gaits Using Numerical Optimization*, IFAC World Congress on Automatic Control, Barcelona, Spain, 2002, <http://www-m2.mathematik.tu-muenchen.de/hardt/publications.html>
- [34] Hartshorne, R., *Algebraic Geometry*, Springer, Heidelberg, 1977
- [35] Hoffmann, K.-H., Jäger, W., Lohman, T., Schunck, H. (Hrsg.), *Mathematik - Schlüsseltechnologie für die Zukunft*, Springer, Heidelberg, 1996
- [36] Immler, C., *3D StudioMax 3*, Data Becker, Düsseldorf 1999
- [37] Koo, T.W., Yoon, Y.S., *Dynamic Instant Gait Stability Measure for Quadruped Walking robot*, Robotica 17, 1999
- [38] Krabbe, B., *Zur Belastung des Bewegungsapparates beim Laufen*, Shaker, Aachen 1994
- [39] Kraus, C., *Modellierung und Rekursive Algorithmen für Mehrkörpersysteme in Natürlichen Koordinaten*, Diplomarbeit, IWR Universität Heidelberg, 1997
- [40] Kraus, C., Winckler, M. und Bock, H.-G., *Modeling Mechanical DAE Using Natural Coordinates*, IWR Preprint, Universität Heidelberg, 2000
- [41] Kraus, C., Winckler, M., *Natural Coordinates and Mechanical DAE*, Proceedings of 3rd MATHMOD VIENNA, 2000
- [42] Leineweber, D.B., *Analyse und Restrukturierung eines Verfahrens zur direkten Lösung von Optimal-Steuerungsproblemen*, Diplomarbeit, Universität Heidelberg, 1995
- [43] Leineweber, D.B., *Efficient Reduced SQP Methods for the Optimization of Chemical Processes Described by Large Sparse DAE Models*, volume 3, Fortschrittsberichte, VDI, 1999
- [44] Lucas, E., *Huitieme recreation - La machine à marcher*, Recreat. Math. 4, 1894
- [45] Magnenat-Thalmann, N., *An Object-oriented Methodology Using Dynamic Variables for Animation and Scientific Visualization*, in: *CG International '90*
- [46] Metzger, K.D., *Einsatz eines dynamischen Optimierungsverfahrens zur Muskelkraftberechnung menschlicher Bewegungen*, Diplomarbeit, Institut A für Mechanik, Universität Stuttgart, 1995
- [47] MIT Leg Lab, *Leg Lab Robots*, Internetadresse: <http://www.ai.mit.edu/projects/leglab/robots/robotsmain.html>, 2001
- [48] Mombaur, K.D., *Stability Optimization of Open-Loop Controlled Walking Robots*, Dissertation, Fortschritts-Berichte VDI Reihe 8 Nr. 922, Düsseldorf, VDI Verlag, 2002
- [49] Mombaur, K.D., Bock, H.-G., Longman, R.W., *Stable, Unstable and Chaotic Motions of Bipedal Walking Robots without Feedback*, In: Chernousko, F.L., Fradkov, A.L., editors, *Proceedings of 2nd International Conference on Control of Oscillations and Chaos, St. Petersburg*, Volume 2, pages 282-285, 2000
- [50] Mombaur, K.D., Bock, H.-G., Schlöder, J.P., *Numerical generation and stabilization of periodic gaits*, Technical report 2000-39, IWR Preprint, Universität Heidelberg, 2000
- [51] Mombaur, K.D., Bock, H.-G., Schlöder, J.P., Longman, R.W., *Human-like actuated walking that is asymptotically stable without feedback*, Proceedings of IEEE International Conference on Robotics and Automation, 2001
- [52] Mombaur, K.D., Bock, H.-G., Schlöder, J.P., Winckler, M.J., Longman, R.W., *Open-loop Stable Control of Running Robots - a Numerical Method for Studying Stability in the Context of Optimal Control Problems*, Proceedings of Clawar '98 Brussels, 1998
- [53] Nelder, J. A., Mead, R., *A simplex method for function minimization*, Computer Journal 7, 1965

- [54] Pai, Dinesh K., Barman, Roderick A., Ralph, Scott K., *Platonic Beasts: a New Family of Multilimbed Robots*, Department of Computer Science, University of British Columbia, 1995
- [55] Rabier, P.J., Rheinboldt, W.C., *On Impasse Points of Quasilinear Differential-Algebraic Equations*, J. Math. Anal. Appl. 181(2):429-454, 1994
- [56] Rabier, P.J., Rheinboldt, W.C., *On the Computation of Impasse Points of Quasilinear Differential-Algebraic Equations*, Math. Comp., 62(101):133-154, 1994
- [57] Raibert, M.H., *Legged Robots*, Communications of the ACM 29
- [58] Raibert, M.H., Sutherland, I.E., *Machines That Walk*, Scientific American, 248(1):32-41, Jan. 1983
- [59] Ruder, H., Ertl, T., Gruber, K., Günther, M., Hospach, F., Ruder, M., Subke, J., Widmayer, K., *Kinematics and Dynamics for Computer Animation*, Institut für theoretische Astrophysik, Universität Tübingen, Eurographics Technical Report Series, 1991
- [60] Schlöder, J.P., *Numerische Methoden zur Behandlung hochdimensionaler Aufgaben der Parameteridentifizierung*, Bonner Mathematische Schriften 187, 1988
- [61] Schulz, V. H., Bock, H.-G. und Steinbach, M. C., *Exploiting Invariants in the Numerical Solution of Multipoint Boundary Value Problems for DAE*, IWR-Preprint Universität Heidelberg, 1993
- [62] Science Magazine, <http://www.sciencemag.org/cgi/content/full/278/5344/1743>, Vol. 278, Nr. 5344, Dez. 1997, S. 1743-1748
- [63] Simeon, B., *Numerische Integration mechanischer Mehrkörpersysteme: Projizierende Deskriptorform, Algorithmen und Rechenprogramme*, Bd. 130 der Fortschrittsberichte, Reihe 20, VDI, 1994
- [64] Simon, J., *Modellierung von Kontakt ereignissen bei der Simulation von Laufvorgängen*, Diplomarbeit, Universität Heidelberg, 1998
- [65] Villamayor, O., *Constructiveness of Hironaka's Resolution*, Ann. Scient. Ec. Norm. Sup., 4^e serie, 22, 1989
- [66] von Schwerin, R., *Numerical Methods, Algorithms and Software for Higher Index Nonlinear Differential-Algebraic Equations in Multibody System Simulation*, Dissertation, Universität Heidelberg, 1997
- [67] Wehage, R.A., Haug, E.J., *Generalized Coordinate Partitioning for Dimension Reduction in Analysis of Constrained Dynamic Systems*, Journal of Mechanical Design, Ausgabe 134, 1982
- [68] Winckler, M., *Numerische Werkzeuge zur Simulation, Visualisierung und Optimierung un stetiger dynamischer Prozesse*, Dissertation, Universität Heidelberg, 2000
- [69] Winckler, M., Huber, J., *JAFV - an OpenGL Visualization Testbed for Dynamic Models*, Technical report, IWR Preprint, 1999
- [70] Winckler, M., von Schwerin, R., *Simulating Vehicle Systems with Discontinuous Effects using MBS-SIM*, in ZAMM96, 1996

Anhang

In diesem Anhang sind einige der Optimierungsergebnisse wiedergegeben. Die Werte sind dabei auf die vierte Stelle gerundet worden.

Die Auflistung der Zustände und Kontrollen zu den Zeitpunkten der Mehrzielknoten kann unter <http://www.iwr.uni-heidelberg.de/~thorsten.stossmeister> eingesehen werden. Dort befindet sich auch jeweils eine Visualisierung der Bewegung.

Optimierungsergebnis „Strenger tiefergelegter Schwerpunkt“

Abschnitt	§6.3
Abbildung	VI.12
Ergebnis Zielfunktional:	5.344
Zeitliche Länge des Intervalls:	$4.75 \cdot 10^{-1}$ (Zeitdauer fixiert auf diesen Wert)
Karush-Kuhn-Tucker-Bedingung:	$3.792 \cdot 10^{-6}$

Optimierungsergebnis „Periodisch vorwärts kommen, aufrecht bleiben“

Abschnitt	§7.1
Abbildung	VI.13
Ergebnis Zielfunktional:	-1.144
Zeitliche Länge des Intervalls:	$1.5 \cdot 10^{-1}$ (Zeitschranke $0.1 \leq T \leq 0.15$)
Karush-Kuhn-Tucker-Bedingung:	$8.887 \cdot 10^{-6}$

Optimierungsergebnis „Periodisch vorwärts kommen“

Abschnitt	§7.2
Abbildung	VI.14
Ergebnis Zielfunktional:	-2.305
Zeitliche Länge des Intervalls:	$1.5 \cdot 10^{-1}$ (Zeitschranke $0.1 \leq T \leq 0.15$)
Karush-Kuhn-Tucker-Bedingung:	$7.005 \cdot 10^{-7}$

Optimierungsergebnis „Periodisch mit minimaler Kontrolle“

Abschnitt	§7.3
Abbildung	VI.15
Ergebnis Zielfunktional:	$1.337 \cdot 10^1$
Zeitliche Länge des Intervalls:	$1.5 \cdot 10^{-1}$ (Zeitschranke $0.1 \leq T \leq 0.15$)
Karush-Kuhn-Tucker-Bedingung:	$8.893 \cdot 10^{-6}$

Optimierungsergebnis „Periodisch bei freier Endzeit“

Abschnitt	§7.4
Abbildung	VI.16
Ergebnis Zielfunktional:	-4.617
Zeitliche Länge des Intervalls:	$6.407 \cdot 10^{-1}$ (Zeitschranke $0.1 \leq T \leq 0.15$)
Karush-Kuhn-Tucker-Bedingung:	$6.011 \cdot 10^{-7}$