Simulation interaktiv bewegter Objekte mit Hinderniskontakten

Dissertation

zur Erlangung des Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)
der Technischen Fakultät
der Universität des Saarlandes

von

Matthias Buck

Saarbrücken
1999

Erstgutachter:	Prof. Dr. rer. nat. Dr. h.c.mult. G. Hotz
Zweitgutachter:	Prof. Dr.-Ing. J. Schürmann
Tag des Kolloquiums:	3. Februar 1999

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Daimler-Benz-Forschungsinstitut in Ulm.

Mein ganz besonderer Dank gilt Herrn Prof. Dr. rer. nat. Dr. h.c.mult. G. Hotz für die Übernahme des Hauptgutachtens und seine Bereitschaft, die Arbeit als externe Dissertation an der Universität des Saarlandes zu betreuen und durch zahlreiche fruchtbare Diskussionen und Anregungen zu fördern.

Herzlich bedanken möchte ich mich auch bei Herrn Prof. Dr.-Ing. J. Schürmann für die Übernahme des Zweitgutachtens und für die wertvollen Diskussionen und Hinweise im Verlaufe der Entstehung dieser Arbeit sowie bei deren schriftlicher Ausarbeitung.

Herr Dr. rer. nat. E. Schömer vom Lehrstuhl Prof. Hotz stand im Verlaufe der Arbeit stets als engagierter Diskussionspartner zu Verfügung und trug durch zahlreiche Anregungen und Hinweise zum Gelingen der Arbeit bei. Hierfür bin ich ihm zu großem Dank verpflichtet.

Zahlreichen Kollegen am Daimler-Benz-Forschungsinstitut bin ich dankbar für ihre Kollegialität, viele anregende Diskussionen und konstruktive Kritik während der Entstehung der Arbeit und bei der Durchsicht der Manuskripte. Besonders erwähnen möchte ich hierbei Herrn Dipl.-Inf. Jens Eckstein und Herrn Dipl.-Inf. Jörg Sauer, deren Engagement und Unterstützung mir eine große Hilfe waren.

Für die Anregung zu dieser Arbeit und die stete Ermunterung im Verlaufe ihrer Entstehung möchte ich Herrn Dr.-Ing. F. May herzlich danken. Ihm und Herrn Dr.-Ing. K. Grebner bin ich zudem für die engagierte Förderung meines Vorhabens und die teilweise Entlastung von Routinetätigkeiten dankbar. Ohne diese Unterstützung wäre die Durchführung der vorliegenden Arbeit nicht möglich gewesen.

Herrn Prof. Dr.-Ing. G. Meyer-Brötz verdanke ich den Anstoß, parallel zu meiner Tätigkeit am Daimler-Benz-Forschungsinstitut zu promovieren.

Ein ganz herzlicher Dank gebührt schließlich meiner Frau Sieglinde und meinen Kindern Amelie und Simon für ihre Unterstützung und Geduld während der letzten Jahre, in denen ich viel zu wenig Zeit für sie hatte.

Kurzzusammenfassung

Die interaktive Handhabung starrer Körper in virtuellen Welten erfordert die Simulation eines Objektverhaltens, welches physikalisch plausibel und intuitiv korrekt ist. Andernfalls wäre eine solche Simulation für praktische Anwendungen ungeeignet. Physikalisch plausibles Verhalten impliziert, daß die aus Objektkollisionen resultierenden Bewegungsbeschränkungen im Rahmen der Bewegungssimulation berücksichtigt werden.

Dieser Beitrag stellt ein Verfahren zur Simulation interaktiv bewegter Objekte mit Hinderniskontakten vor, bei welchem alle Kontaktsituationen durch eine minimale Menge automatisch bestimmter Punktkontakte repräsentiert werden. Die daraus abgeleiteten Bewegungsbeschränkungen dienen als Randbedingungen für die Simulation der Objektbewegungen.

Die vorgeschlagene Vorgehensweise erlaubt auch die Nachbildung von Reibungseffekten. Die Approximationseigenschaften dieser Nachbildung werden sowohl durch theoretische als auch durch empirische Untersuchungen diskutiert.

Short Summary

The interactive manipulation of rigid bodies in virtual environments requires the simulation of an object behaviour, which is physically plausible and intuitively correct. Otherwise such a simulation would be useless for practical applications. Physically plausible behaviour implies that motion restrictions due to object collisions are taken into account during the motion simulation.

This contribution presents a method for the motion simulation of interactively controlled objects with obstacle contacts, representing any contact situation by a minimal set of automatically determined point contacts. The motion restrictions derived from this set are the basis for the simulation of object motions.

The proposed method also allows the simulation of friction effects, the approximation properties of which are discussed on the basis of both a theoretical and an empirical analysis.

Inhaltsverzeichnis

1 Einführung

1.1 Direkte Interaktion von Menschen in virtuellen Umgebungen

Durch die rasante Entwicklung immer leistungsfähigerer Computergrafiksysteme rückt die realistische Visualisierung und Echtzeitanimation auch komplexerer computersimulierter Umgebungen in greifbare Nähe. Die Darstellung synthetischer Umgebungen wurde zunächst vorwiegend in Fahr- und Flugsimulatoren sowie zur Architekturvisualisierung eingesetzt. Dabei spielte vorwiegend die Navigation in synthetischen Umgebungen eine Rolle. Mit dem Begriff der 'virtuellen Realität' (VR) verbindet man, zusätzlich zur immersiven Echtzeitvisualisierung, insbesondere auch die Echtzeitsimulation virtueller Umgebungen, und die Möglichkeit, als Anwender direkt in solche Umgebungen einzutauchen, und mit ihnen realistisch zu interagieren. Die neuen Möglichkeiten solcher interaktiver VR-Simulationen wurden zuerst von der Spieleindustrie aufgegriffen. Mittlerweile wird jedoch das Potential interaktiver VR-Techniken auch für industrielle Anwendungen zunehmend erkannt. Das Anwendungsspektrum industrieller VR-Anwendungen mit direkter Interaktion zwischen Anwender und virtuellen Umgebungen reicht von Ergonomieuntersuchungen, Montagesimulationen und Telemanipulation bis hin zu Ausbildungs- und Trainingssystemen.

Bei all diesen VR-Anwendungen steht stets der Mensch im Mittelpunkt. Er bleibt kein unbeteiligter Beobachter der simulierten Abläufe, sondern greift gestaltend in das Geschehen ein. Er ergreift Objekte, benutzt diese als Werkzeuge zur Durchführung von Arbeitsschritten in simulierten Arbeitsumgebungen, er betätigt Schalter und andere Bedienelemente, liest Instrumente ab, und reagiert auf Ereignisse in der virtuellen Welt.

Solche VR-Anwendungen sind jedoch nur dann von praktischer Bedeutung, wenn die virtuellen Umgebungen hinreichend gut mit den realen Vorbildern übereinstimmen. Der Benutzer muß idealerweise mit der virtuellen Welt intuitiv und entsprechend seiner Alltagserfahrung interagieren können. Dies erfordert einerseits, mit simulierten Objekten möglichst genauso umgehen zu können wie dies in der realen Situation der Fall wäre. Umgekehrt müssen sich die Elemente der virtuellen Umgebung intuitiv richtig und physikalisch glaubwürdig verhalten, und entsprechend vorhersagbar auf die Aktionen des Benutzers reagieren.

Die direkte Interaktion von Menschen in virtuellen Umgebungen erfordert daher die Nachbildung der Merkmale realer Umgebungen, die für die jeweilige Anwendung eine wesentliche Rolle spielen. Die klassische Repräsentation virtueller Szenarien beschränkt sich auf die Beschreibung der Objektgeometrien und aller sonstigen zur Visualisierung erforderlichen Merkmale. Physikalische Eigenschaften realer Objekte werden dabei nicht automatisch berücksichtigt, sondern müssen gesondert simuliert werden. Zu den wesentlichen physikalischen Merkmalen gehören grundlegende mechanische Eigenschaften fester Körper, die sich beispielsweise normalerweise nicht gegenseitig durchdringen können, und die der Massenträgheit und der Schwerkraft unterworfen sind. Innerhalb von interaktiven VR-Anwendungen müssen solche Simulationen physikalischer Eigenschaften denselben Echtzeitanforderungen genügen, wie dies für die Visualisierung und die Interaktionskomponenten des VR-Systems gilt.

Die Benutzerschnittstelle zwischen der materiellen, realen Welt des Benutzers und der rein synthetischen, immateriellen virtuellen Welt kann bei interaktiven VR-Anwendungen verschieden ausgeprägt sein. Um intuitive Handhabungssimulationen zu ermöglichen, werden beispielsweise zur Steuerung von Bewegungsabläufen häufig Datenhandschuhe eingesetzt. Diese erlauben zwar dem Anwender die Eingabe von Bewegungsdaten, machen ihm jedoch keine Reaktionskräfte aus der virtuellen Welt erfahrbar. Hierzu wären geeignete haptische Benutzerschnittstellen erforderlich, die derzeit jedoch noch nicht in akzeptabler Form zur Verfügung stehen.

Das Fehlen solcher Kraftrückkopplungsmechanismen beeinträchtigt nicht nur den subjektiven Eindruck des Benutzers von der Realitätstreue virtueller Umgebungen. Ohne die unmittelbare Wahrnehmung von Reaktionskräften ist mit dem Datenhandschuh auch eine korrekte Objekthandhabung kaum möglich, da Bewegungseinschränkungen durch Hinderniskontakte von der kontrollierenden Hand nicht berücksichtigt werden können.

Es ist daher erforderlich, ein intuitiv richtiges, physikalisch glaubwürdiges Objektverhalten durch geeignete Simulationsverfahren innerhalb der virtuellen Umgebung zu realisieren. Eine solche Funktionalität steht bei den derzeit verfügbaren kommerziellen oder universitären VR-Simulationssystemen bisher nicht zur Verfügung. Die Entwicklung dieser Funktionalität ist Gegenstand dieser Arbeit.

1.2 Industrielle Anwendungsfälle

Bei der Entwicklung industrieller Produkte müssen zahlreiche Entwurfsentscheidungen bereits in einer sehr frühen Phase des Produktentwicklungszyklus getroffen werden. Dies betrifft Fragen der Benutzbarkeit und Funktionalität, der mechanischen Konstruktion, der Planung von Produktions- und Montageeinrichtungen, sowie der Wartung und Instandsetzung. Traditionell stützt sich dieser Entscheidungsprozeß auf die Darstellungsmöglichkeiten von Konstruktionswerkzeugen wie CAD-Systemen, auf Funktionsmodelle in verschiedenen Maßstäben, sowie auf Erfahrungswerte aus früheren Entwurfsprozessen. Im Gegensatz hierzu zeigen sich Entscheidungsfehler und hieraus resultierende Probleme leider erst zu sehr viel späteren Zeitpunkten, während der Produktion, bei der Benutzung, oder gar erst bei der Wartung des Produktes. Die dann erforderlichen Korrekturmaßnahmen sind zeitaufwendig und teuer.

Eine frühzeitige Erkennung und Korrektur solcher Fehler bereits während der Konstruktionsphase setzt jedoch voraus, daß die in späteren Phasen erforderlichen Abläufe bis hin zur Handhabung durch den Kunden unter weitgehend realistischen Bedingungen getestet werden können. Die technische Funktion selbst ist dabei weniger problematisch als Fragen der Interaktion zwischen dem Menschen und dem technischen Produkt. Die Einbeziehung des menschlichen Monteurs, Anwenders, oder Wartungsmechanikers ist jedoch mit den traditionellen Entwurfstechniken kaum möglich.

Simulationen in virtuellen Umgebungen bieten neue Möglichkeiten zu Lösung dieser Problematik. Schon in sehr frühen Entwurfsphasen, wenn erste Konstruktionsdaten im Rechner vorliegen, können diese Datensätze im Rahmen virtueller Umgebungen dargestellt und interaktiv getestet werden. Dabei kann sich neben der Geometrie und dem äußeren Erscheinungsbild auch die Funktion des Produktes in der Simulation widerspiegeln, so daß potentielle Probleme bei der Handhabung ebenfalls entdeckt werden können.

Die folgenden Abschnitte stellen einige mögliche industrielle Anwendungsfälle der Interaktion menschlicher Benutzer mit virtuellen Umgebungen dar.

Ergonomieuntersuchungen

Sowohl die Nutzbarkeit eines Produktes, als auch die Kundenzufriedenheit wird wesentlich mitbestimmt durch die Ergonomie eines Produktes, d.h. die Frage ob die Bedienung einfach und intuitiv ist, und ob sie keine unbequemen oder ermüdenden Körperhaltungen erfordert. Solche Fragestellungen lassen sich mit virtuellen Prototypen bereits in sehr frühen Phasen der Produktentwicklung verifizieren, indem die typischen Handgriffe und Bewegungsabläufe bei der Bedienung in unmittelbarer Interaktion mit einer synthetischen Umgebung weitgehend realistisch nachvollzogen werden.

Montagesimulation

Auch bei der Montagesimulation gibt es zunächst einen ergonomischen Aspekt, und die dabei angesprochenen Punkte gelten natürlich auch für Arbeitsvorgänge bei der Montage oder der Wartung. Ein zweiter Aspekt betrifft die generelle Durchführbarkeit einzelner Arbeitsschritte. Oft müssen bei der Montage Teile auf engem Raum montiert werden, und es kann schwierig oder gar unmöglich sein, sie einzubauen, da andere Komponenten den Weg oder die Sicht versperren. Diese Problematik tritt besonders auch bei Reparaturen auf. Wird es erforderlich, nur der Zugänglichkeit wegen andere Komponenten ebenfalls zu demontieren, erhöhen sich Reparaturzeiten und Kosten entsprechend. Der Vorteil, solche Schwierigkeiten durch Simulationen in virtuellen Umgebungen frühzeitig zu erkennen und zu vermeiden, liegt auf der Hand. Ein weiteres Anwendungsfeld stellt die Planung geeigneter Einbaupfade für komplexe Montagesituationen dar, die für die automatisierte Montage mit Handhabungsautomaten erforderlich ist.

Tele-Manipulation

In einer Reihe von Situationen kann ein Bedienablauf nur über ferngesteuerte Roboter durchgeführt werden (z.B. unbemannte Raumstation, Telemedizin, etc.). Beträgt die Signallaufzeit für Steuer- und Videosignale hierbei jedoch mehr als wenige 100ms, ist eine direkte Fernbedienung des Roboters durch den Anwender nicht mehr intuitiv möglich. In solchen Fällen kann der Benutzer in eine virtuelle Kopie des entfernten Arbeitsplatzes versetzt werden, und in dieser ohne störende Verzögerungen arbeiten. Die durchgeführten Aktionen werden dann vom Roboter vor Ort entsprechend verzögert nachvollzogen.

Training

Training unter realen Bedingungen kann sehr gefährlich und sehr teuer sein (man denke an die Ausbildung von Chirurgen oder Piloten). In anderen Beispielen ist die Trainingsumgebung noch nicht real verfügbar, z.B. bei der Ausbildung von Montagepersonal für neuentwickelte Produkte. Training in simulierten Umgebungen hat in solchen Fällen viele Vorteile. Es ist nicht nur weniger kostenintensiv und auch sicherer, sondern oftmals auch effizienter. Kritische Situationen, welche speziell trainiert werden sollen, können kontrolliert eingespielt und dabei individuell auf die Bedürfnisse und Fähigkeiten des Auszubildenden zugeschnitten werden.

1.3 Vorarbeiten und eigene Beiträge

Die Simulation der Kinematik und Dynamik starrer Körper mit Kontaktbedingungen ist ein Themengebiet, welches bereits von zahlreichen Autoren behandelt wurde. Frühe Beiträge lieferten Lötstedt [30] mit grundlegenden Ergebnissen zur Reibungssimulation, Cai und Roth [10], und Hahn [22] mit Ergebnissen zur Dynamiksimulation und zur Modellierung von Kontakten. Eine fundierte Untersuchung der Bewegungssimulation starrer Körper mit Kontakten wurde von Baraff im Rahmen seiner Dissertation [1] vorgelegt, und er trug seitdem weiter wesentlich zur Weiterentwicklung dieses Themengebietes bei ([3, 4]). Cremer, Stewart, und Vaněček [16, 46] verwenden in ihren Simulationssystemen *Newton* und *Isaac* ebenso wie Baraff einen Ansatz mit Kontaktkräften zur Lösung kontinuierlicher Bewegungsgleichungen. Im Gegensatz dazu schlugen Mirtich und Canny [34] ein impulsbasiertes Verfahren vor. Hierbei wird jede Wechselwirkung zwischen sich berührenden Körpern durch eine Vielzahl sogenannter Mikrokollisionen erreicht.

Weitere Formulierungen von kontaktbedingten Bewegungsbeschränkungen finden sich in einer Reihe von Arbeiten [2, 4, 10, 33, 42]. Dabei werden neben Ansätzen mit Kontaktkräften auch das Prinzip des geringsten Widerstandes nach Gauss oder Gradientenabstiegsverfahren in Verbindung mit Potentialfeldern vorgeschlagen.

Baraff stellt in [4] fest, daß unabhängig von diesen Variationen bei allen Ansätzen zur Dynamiksimulation mit Bewegungsbeschränkungen eine grundsätzliche Entscheidung getroffen werden muß: Entweder werden die Beschränkungen dadurch modelliert, daß die verbleibenden Bewegungsmöglichkeiten durch eine reduzierte Anzahl von Koordinaten beschrieben werden. Oder die Einhaltung der Bewegungsbeschränkungen wird durch die Einführung von Zwangskräften sichergestellt. Bei kontaktbedingten Bewegungsbeschränkungen entsprechen diese Zwangskräfte gerade den Kontaktkräften. Baraff diskutiert das Für und Wider dieser beiden Alternativen. Bei der Simulation von Objektbewegungen mit physikalischen Effekten wie Reibung, Schwerkraft und Massenträgheit hat der Ansatz basierend auf Kontaktkräften entscheidende Vorteile. Auch wir werden einen solchen Ansatz wählen.

Zur Lösung der Bewegungsgleichungen mit Kontaktbedingungen, und zur Bestimmung der hierfür erforderlichen Kontaktkräfte werden von den meisten Autoren [1, 4, 18, 44, 45] Randbedingungen für die Kontaktbeschleunigung in Richtung der Berührnormalen aufgestellt. Die resultierenden Ungleichungssysteme nehmen meist die Form eines sogenannten linearen Komplementaritätsproblems (LCP) an [15], zu dessen Lösung vorwiegend der Lemke-Algorithmus [27] herangezogen wird.

Die Formulierung der Bewegungsbeschränkungen auf der Ebene von lokalen Kontaktbeschleunigungen führt auf die gesuchten Zwangskräfte, aus denen sich die globale Beschleunigung des gesamten starren Körpers ergibt. Die resultierende Objektbewegung ergibt sich hieraus durch numerische Integration der Bewegungsgleichungen. In praktischen

Realisierungen wird die Bewegungssimulation in zeitdiskreten Schritten durchgeführt. Daher ergeben sich bei der numerischen Integration der Bewegungsgleichungen kleine Ungenauigkeiten, die sich im Laufe der Simulation akkumulieren, und zu einer zunehmenden Verletzung der Kontaktbedingungen führen können.

Der in dieser Arbeit eingeschlagene Weg vermeidet diese Problematik, indem die Bewegungsbeschränkungen nicht auf der Ebene von Kontaktbeschleunigungen, sondern direkt für die Kontaktabstände formuliert werden. Auch bei der zeitdiskreten Lösung der Bewegungsgleichungen werden durch diesen Ansatz die Kontaktbedingungen stets stabil eingehalten. Im Gegensatz zur klassischen Vorgehensweise erhält man bei diesem Verfahren jedoch kein LCP, sondern ein nichtlineares Komplementaritätsproblem (NCP) zur Bestimmung der Zwangskräfte, welches nicht mehr mit dem Lemke-Algorithmus gelöst werden kann.

In der Literatur zur Bewegungssimulation mit Kontaktbedingungen spielen NCPs keine große Rolle. In Verbindung mit Reibungsproblemen werden NCP-Ansätze zwar erwähnt [45, 39], dann jedoch wieder auf ein LCP zurückgeführt, meist unter Annahme gewisser Einschränkungen.

Die Lösung von Komplementaritätsproblemen wird in einer großen Anzahl von Veröffentlichungen aus dem technischen Bereich, der Mathematik, und nicht zuletzt auch aus den Wirtschaftswissenschaften behandelt [18]. Gemäß den Arbeiten von Mangasarian [31], Fischer [19] sowie Kanzow und Kleinmichel [23, 24, 25, 26] lassen sich LCP's wie NCP's durch geeignete Operatoren auf nichtlineare Gleichungssysteme zurückführen und mit Newton-ähnlichen Verfahren iterativ lösen.

Die Literatur aus dem Bereich Robotik, obwohl thematisch verwandt, ist für unser Problem weniger ergiebig, da dort andere Fragestellungen dominieren. Während wir interaktiv Bewegungen vorgeben, und bei Kollisionen und Kontakten die resultierende Bewegung zu bestimmen suchen, sind in der Robotik Fragen der Bewegungsplanung, der Bewegungssteuerung und der Kollisionsvermeidung von größerer Bedeutung. Kontaktprobleme treten in der Robotik insbesondere bei Greif- und Fügeoperationen auf. Mögliche Konflikte zwischen Bewegungsvorgaben und realen Kontaktrestriktionen werden vorwiegend durch hierfür vorgesehene elastische Elemente in der Greifmechanik ausgeglichen.

Der Frage der Repräsentation von Kontaktsituationen zwischen zwei oder mehreren sich berührenden Körpern wird in der einschlägigen Literatur nur wenig Aufmerksamkeit geschenkt. Die meisten Autoren gehen von einer gegebenen Menge von Punktkontakten aus, und konzentrieren sich auf die Behandlung der resultierenden Kontaktbedingungen und Bewegungsgleichungen. Wie sich die in einer gegebenen Kontaktsituation vorliegenden punktförmigen, linienhaften und flächenhaften Kontaktregionen auf eine adäquate Menge von Punktkontakten abbilden lassen, wird dagegen kaum diskutiert.

Baraff beschreibt in [5] ein Verfahren, nach welchem alle Punktkontakte in die Simulation mit einbezogen werden, welche im Rahmen der Kollisionserkennung registriert werden. Bei der von Baraff beschriebenen Vorgehensweise können zahlreiche redundante Punktkontakte auftreten, welche keinen Beitrag zur Lösung des Problems leisten, sondern nur den numerischen Aufwand erhöhen.

Eine ausführliche Analyse von Kontaktsituationen für physikbasierte Simulationen wird von Bouma und Vaněček [8] dargestellt. Dabei werden die Kontaktregionen zwischen zwei polygonalen Körpern durch eine volle mengentheoretische Schnittbildung der

Objekthüllen ermittelt. Aus den so gewonnenen Kontaktregionen wird anschließend eine endliche Anzahl von Punktkontakten gewonnen. Gemäß [8] beträgt die Komplexität dieser Aufgabe $O(N^2)$, wobei N die Anzahl der Eckpunkte, Kanten und Flächen beider Objekte bezeichne. Der numerische Aufwand hierfür kann bei komplexen Objektgeometrien daher sehr groß werden.

In der vorliegenden Arbeit wird ein Verfahren beschrieben, welches eine solche aufwendige Analyse vermeidet. Es bestimmt automatisch eine kanonische, nichtredundante Menge von Punktkontakten zur Repräsentation der jeweils aktuell vorliegenden Kontaktkonfiguration, und paßt diese Menge dynamisch den wechselnden Kontaktverhältnissen an. Dabei basiert die Auswahl neuer Kontaktpunkte auf Daten, welche im Rahmen der Kollisionserkennung ermittelt werden. Eine geschickte Strategie bei der Auswertung dieser Daten vermeidet dabei das Auftreten redundanter Kontaktpunkte. Hierdurch minimiert sich der numerische Aufwand bei der Aufstellung und Lösung der Bewegungsgleichungen mit Bewegungsbeschränkungen. Zudem wird diese Lösung eindeutig, da keine unterbestimmten Gleichungssysteme auftreten.

Die Thematik der Kollisionserkennung wird in dieser Arbeit nur insoweit gestreift, als sie zur Beschreibung des Aufbaus von Kontaktbedingungen erforderlich ist. Auf Details zur effizienten Detektion von Kollisionen und zur Bestimmung von Kollisionszeitpunkten werden wir daher nicht eingehen. Eine gute Darstellung des Stands der Technik auf diesem Gebiet stammt von Gottschalk, Lin, und Manocha96 [21]. Theoretische Untersuchungen zur numerischen Komplexität der durchzuführenden Kollisionstests finden sich in der Dissertation von Schömer [41]. Die zur Realisierung der Kontaktsimulation verwendeten Kollisionserkennungsmodule werden ausführlich von Eckstein [17] beschrieben.

Zur Beschreibung von Reibung in physikbasierten Simulationen wird gemeinhin das Coulomb'sche Reibungsgesetz zugrundegelegt [1, 34, 44, 45]. Im Falle der Gleitreibung stehen dabei die Beträge der Normalkraftkomponenten und der tangentialen Reibungskraftkomponenten in einem festen Verhältnis zueinander, welches durch den Reibungskoeffizienten bestimmt wird. Im Falle der Haftreibung bestehen für die Komponenten der Kontaktkräfte zusätzliche Freiheitsgrade. Dadurch können bereits in einfachen Mehrkontaktproblemen starrer Körper unterbestimmte Konfigurationen auftreten.

In der vorliegenden Arbeit wird ein neues Verfahren beschrieben, welches Haftreibung nicht wie in der klassischen Vorgehensweise (z.B. [45]) als Attribut der jeweils gegebenen Kontaktpunkte behandelt, sondern durch zusätzliche Kontaktpunkte realisiert. Die Bewegungsbeschränkungen, die sich durch die entsprechenden zusätzlichen Kontaktbedingungen ergeben, unterbinden dabei die tangentiale Gleitbewegung. Diese zusätzlichen Kontaktpunkte werden dynamisch, und nur soweit sie zur Einhaltung der Haftreibungsbedingungen erforderlich sind, automatisch definiert und auch wieder eliminiert. Dadurch werden bei Kontaktsituationen mit Haftreibung unterbestimmte Problemstellungen vermieden.

Die Substitution von Flächenkontakten durch wenige Punktkontakte führt im reibungsfreien Fall zu einem kinematisch korrekten Bewegungsverhalten der simulierten Objekte. Im Falle der Gleitreibung muß jedoch gesondert untersucht werden, inwieweit sich die die Eigenschaften flächiger Gleitreibung durch die Reibung an wenigen Punktkontakten approximieren lassen. Da sich Flächenkontakte durch eine ausreichend große Anzahl von Punktkontakten beliebig approximieren lassen, kann mit entsprechendem numerischem Aufwand auch flächige Reibung approximativ simuliert werden. Eine Bewertung des resultierenden Bewegungsverhaltens bei der Nachbildung von Flächenreibung durch eine geringe Anzahl von Reibungspunkten findet sich in der Literatur jedoch nicht.

Die Literatur aus dem Bereich der physikbasierten Bewegungssimulation starrer Körper beschränkt sich in diesem Zusammenhang auf die Beschreibung von Reibungseffekten an punktförmigen Kontakten. Auf dem Gebiet der Robotik finden sich zahlreiche Untersuchungen zur Mechanik flächiger Gleitreibung. Die Arbeiten von Lynch und Mason [28, 29, 32] behandeln das Bewegungsverhalten realer Objekte, welche von Robotermanipulatoren auf einer ebenen Grundfläche entlanggeschoben werden. Neben theoretischen Untersuchungen werden dort auch experimentelle Anordnungen beschrieben, die der empirischen Bestimmung von Reibungsparametern aus dem gemessenen Bewegungsverhalten der Versuchsobjekte dienen. Eine vergleichende Bewertung des Bewegungsverhaltens von Objekten in Abhängigkeit von der Geometrie der Reibregionen findet jedoch auch dort nicht statt. Da wir uns in der vorliegenden Arbeit mit der Kontaktsimulation unter Verwendung der geringstmöglichen Anzahl von Kontaktpunkten befassen, ist diese Fragestellung von besonderer Relevanz, und wird sowohl theoretisch als auch experimentell untersucht.

Die in dieser Arbeit sowie in [9] beschriebenen Verfahren zur Simulation interaktiv bewegter Objekte mit Hinderniskontakten wurden implementiert und in die von der Daimler-Benz Forschung entwickelte Virtual-Reality-Umgebung *DBView* integriert. Alternativ zu DBView existieren eine Reihe kommerzieller Systeme zur Simulation virtueller Welten.[1] Die meisten dieser Systeme verfügen über Mechanismen zur statischen Kollisionserkennung, und können daher Objektdurchdringungen feststellen. Die erkannten Kollisionen werden dem Benutzer dabei jedoch lediglich durch visuelle beziehungsweise akustische Rückmeldungen angezeigt, oder die betroffenen Objekte verharren in der letzten kollisionsfreien Position. Ein realistisches Objektverhalten durch die Simulation kontaktbedingter Bewegungsbeschränkungen wurde bisher von keinem der verfügbaren Systeme unterstützt. Diese Funktionalität, über die DBView nun verfügt, stellt daher einen signifikanten Fortschritt dar. Sie erweitert die Einsatzmöglichkeiten von VR-Systemen für industrielle Anwendungen wesentlich.

1.4 Überblick über diese Arbeit

Der Gegenstand dieser Arbeit ist die Simulation interaktiv bewegter Körper, deren Bewegungsmöglichkeiten durch Kontakte zwischen diesen Körpern oder durch Kontakte mit feststehenden Hindernissen eingeschränkt sind. Um interaktive Anwendungen zu ermöglichen, soll diese Simulation in Echtzeit erfolgen.

Als Basis für die mathematische Beschreibung der hierfür benötigten Verfahren befaßt sich Kapitel 2 mit den geometrischen und mathematischen Repräsentationen der betrachteten Objekte und der aus diesen Objekten gebildeten Szenen.

Werden in interaktiven Anwendungen vom Benutzer Bewegungsvorgaben spezifiziert, welche wegen kontaktbedingter Bewegungsbeschränkungen von den simulierten Objekten

[1] Beispielhaft seien erwähnt die *Virtuelle Werkstatt*^{TM} von Tecoplan, *OXYGEN*^{TM} von Prosolvia-Clarus, *dVise*^{TM} von Division Inc., *VEGA*^{TM} von Paradigm, *REALAX*^{TM} von Media Systems, und *WorldToolKit*^{TM} von Sense8.

nicht vollzogen werden können, so muß für die betroffenen Objekte eine Ersatzbewegung gefunden werden, welche die Bewegungsbeschränkungen respektiert, die externen Bewegungsvorgaben jedoch so gut wie möglich berücksichtigt. Kapitel 3 befaßt sich zunächst mit verschiedenen Repräsentationsformen von kontaktbedingten Bewegungsbeschränkungen, und beschreibt dann den im weiteren Fortgang verwendeten physikalisch orientierten Ansatz zur Bewegungssimulation. Im Gegensatz zu den meisten in der Literatur beschriebenen Vorgehensweisen schlagen wir eine spezielle Formulierung vor, welche die Kontaktbedingungen nicht als Randbedingung für die Objektbeschleunigung formuliert, sondern als Randbedingung für die Raumlage der beteiligten Objekte, und dadurch die stabile Einhaltung dieser Kontaktbedingungen im Laufe der Simulation sicherstellt. Diese neue Vorgehensweise verlangt allerdings auch besondere Verfahren zur numerischen Lösung der Bewegungsgleichungen, auf die gesondert in Kapitel 6 eingegangen wird.

Die Berücksichtigung von Kontaktbedingungen bei der Bewegungssimulation erfordert natürlich auch eine mathematische Beschreibung der vorkommenden Kontaktsituationen (Kapitel 4). Wir schlagen hierzu ein Verfahren vor, welches alle Kontaktsituationen durch eine geringstmögliche Anzahl von Punktkontakten darstellt. Hierdurch wird nicht nur die Ordnung der resultierenden Gleichungssysteme minimiert. Durch redundanzfreie Punktkontaktkonfigurationen sind die Kontaktkräfte auch eindeutig bestimmbar, was der numerischen Stabilität der Lösungsverfahren zugute kommt.

Ein automatisches Verfahren zur Gewinnung eines solchen minimalen Satzes von Punktkontakten wird in Kapitel 5 beschrieben. Dieses Verfahren hat den Vorteil, daß es ohne eine aufwendige Analyse der jeweiligen geometrischen Kontaktregionen (Flächenkontakte, Linienkontakte und Punktkontakte) zwischen zwei sich berührenden Körpern auskommt. Stattdessen werden die benötigten Kontaktpunkte direkt aus den Ergebnissen der Kollisionserkennung abgeleitet. Desweiteren beschreibt dieses Kapitel detailliert die Änderung von Kontaktsituationen durch die neue Entstehung von Kontakten, durch sich lösende Kontakte, sowie durch direkte Kontaktübergänge.

In Kapitel 6 wird nun das numerische Verfahren zur Lösung der Bewegungsgleichungen mit Kontaktbedingungen vorgestellt. Zunächst wird der Aufbau der Gleichungen noch einmal zusammengefaßt, deren Elemente in den vorangegangenen Kapiteln im Einzelnen beschrieben wurden. Man erhält hierbei ein sogenanntes nichtlineares Komplementaritätsproblem (NCP) zur Bestimmung der an den Punktkontakten wirkenden Kontaktkräfte. Dieses wird mit Hilfe eines speziellen Operators auf ein nichtlineares Gleichungssystem zurückgeführt, welches durch effiziente iterative Verfahren wie den Newton-Algorithmus gelöst werden kann.

Die Ergänzung des bisher beschriebenen Simulationsverfahrens um Reibungseffekte ist der Gegenstand von Kapitel 7. Hierbei wird nicht nur Gleitreibung, sondern auch Haftreibung behandelt. Letztere wird, abweichend von der in der Literatur üblichen Darstellungsweise, durch kinematische Zwangsbedingungen realisiert. Dadurch ist es möglich, das beschriebene Verfahren zur Simulation von kontaktbedingten Bewegungsbeschränkungen auch hierfür einzusetzen. Ein weiterer Schwerpunkt dieses Kapitels ist der Vergleich dieser auf wenigen Punktkontakten basierenden Reibungssimulation mit den realen Verhältnissen flächiger Reibungskontakte.

Die Einbindung der beschriebenen Verfahren zur Bewegungssimulation in ein System zur Realisierung virtueller Welten sowie die Benutzerschnittstelle zur interaktiven Einflußnahme eines Benutzers auf die darin enthaltenen simulierten Objekte diskutieren wir in Kapitel 9.

Schließlich gehen wir in Kapitel 10 auf verschiedene Simulationsbeispiele ein und enden mit einer Zusammenfassung und dem Ausblick auf mögliche Ergänzungen dieser Arbeit.

2 Geometrische und mathematische Repräsentationen

Bevor wir uns mit der Simulation des Bewegungsverhaltens und der kollisionsbedingten Bewegungsbeschränkungen interaktiv gesteuerter Körper befassen, ist es erforderlich, auf die mathematischen und geometrischen Repräsentationen der simulierten Objekte einzugehen.

Was in der Realität als berührbares, ergreifbares und bewegbares Objekt vorliegt, mit dem wir auf intuitive Weise Handhabungsabläufe durchführen, erfordert in computersimulierten Umgebungen entsprechende Repräsentationen und Mechanismen, welche wesentliche Eigenschaften dieser Objekte widerspiegeln, und dem Benutzer ein möglichst ebenso intuitives Interagieren erlauben.

Neben der zunächst abstrakten Repräsentation der simulierten Objekte als solche erfordert eine entsprechende Simulationsumgebung einerseits, daß der Anwender direkt in die synthetische Welt eingreifen kann, und ihm andererseits die Reaktion der simulierten Objekte auf seine Handlungen ohne Verzögerung mitgeteilt wird (Abb. 1).

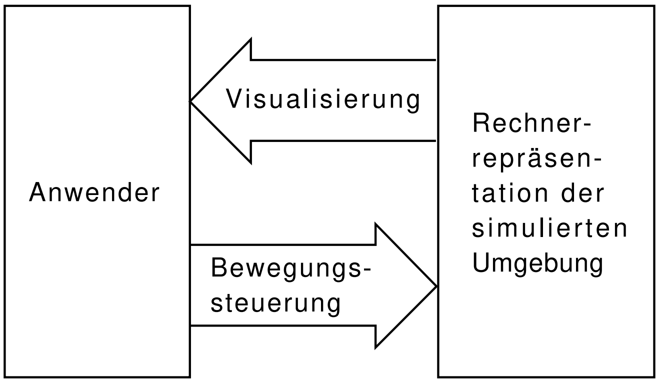


Abbildung 1: Die Interaktion zwischen dem Anwender und der simulierten Umgebung.

Diese Rückmeldung erfolgt vorwiegend durch Echtzeit-Visualisierung der virtuellen Welt. Die hierfür verwendeten Hochleistungsgraphiksysteme bedingen ihrerseits ebenfalls geeignete Datenstrukturen zur Objekt- und Szenenrepräsentation.

2.1 Objektbeschreibungen

Bei der Simulation der Bewegung und des Kontaktverhaltens physikalischer Körper wollen wir uns im Rahmen dieser Arbeit auf starre Objekte beschränken. Auch wenn reale starre Objekte stets ein geringes Maß an Elastizität aufweisen, so sind die auftretenden Verformungen in der Praxis meist so gering, daß sie bei der geometrischen Objektbeschreibung vernachlässigt werden können.

Die Geometrie der betrachteten Körper sei durch die äußere Hülle beschrieben, welche als Polyeder repräsentiert sei. Da sich jede Objektform durch ein Polyeder beliebig genau approximieren läßt, stellt dies in der Praxis keine wesentliche Einschränkung dar. Polyedrische Geometrien haben einerseits den Vorteil, daß sie sich mathematisch einfach beschreiben und handhaben lassen. Andererseits entsprechen sie genau den Oberflächenbeschreibungen, welche von modernen Computergraphik-Systemen effizient visualisiert werden können.

Die Hülle eines Polyeders setzt sich aus polygonalen ebenen konvexen[2] Flächen zusammen, welche an gemeinsamen Kanten und Eckpunkten aneinanderstoßen. Die Objekthüllen lassen sich also durch die geometrischen Grundelemente Fläche, Kante und Ecke darstellen. Diese Elemente seien als abgeschlossene Punktmengen definiert, so daß eine Objektkante der Schnittmenge der beiden angrenzenden polygonalen Flächenelemente entspricht, und ebenso ein Eckpunkt die Schnittmenge aller angrenzenden Flächen darstellt. Diese Eigenschaft wird später für die Behandlung spezieller Kontaktsituationen bedeutsam sein.

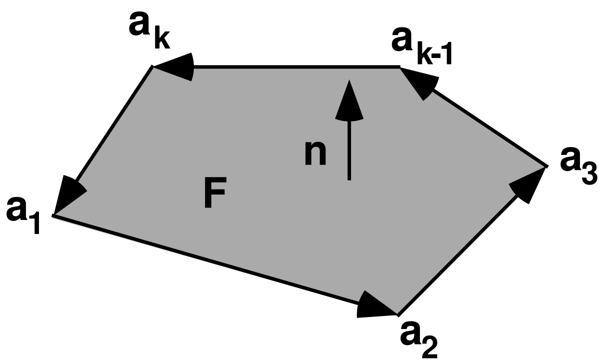


Abbildung 2: Die Beschreibungselemente eines Flächenelementes F.

Jeder Eckpunkt A_j eines polyedrischen Objektes $\mathcal{O}_i$ werde durch seine Koordinaten $\hat{\mathbf{a}}_j \in \mathbb{R}^3$ in einem objektgebundenen Koordinatensystem OKS$_i$ beschrieben. Da sowohl die Kanten als auch die Flächen von $\mathcal{O}_i$ auf diese Eckpunkte zurückgeführt werden, ist auch die Objektgeometrie durch die Eckpunktkoordinaten innerhalb dieses lokalen Koordinatensystems definiert.

Eine Objektkante B_k wird eindeutig durch ihre beiden Endpunkte $\hat{\mathbf{b}}_1$ und $\hat{\mathbf{b}}_2$ beschrieben, sowie durch den resultierenden Richtungsvektor $\hat{\mathbf{u}}_k = \hat{\mathbf{b}}_2 - \hat{\mathbf{b}}_1$.

Ein Flächenelement F der Objekthülle ist gegeben durch die Ebene E: $\hat{\mathbf{n}}^{\mathbf{T}}\mathbf{x} = n_0$ mit dem nach außen gerichteten Normalenvektor $\hat{\mathbf{n}} \in \mathbb{R}^3$, sowie durch die k Eckpunkte $\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \ldots, \hat{\mathbf{a}}_k$, welche von außen betrachtet (d.h. von der Seite der Ebene, in welche die Normale weist) gegen den Uhrzeigersinn angeordnet sind (vgl. Abb. 2). Das Innere des Flächenelementes stellt damit die Menge aller Punkte der Ebene $\mathbf{x} \in$ E dar, für die gilt

$$\forall_i : \quad (\hat{\mathbf{n}} \times (\hat{\mathbf{a}}_{i+1} - \hat{\mathbf{a}}_i))^T (\mathbf{x} - \hat{\mathbf{a}}_i) \geq 0.$$

Durch den Bezug auf ein lokales Koordinatensystem ist diese Geometriebeschreibung unabhängig von der jeweiligen Anordnung des Objektes innerhalb einer Gesamtszene.

Neben der Geometriebeschreibung bedarf es auch der Repräsentation einiger physikalischer Eigenschaften. Zur Aufstellung der Bewegungsgleichungen in Abschnitt 3.4

[2]Nichtkonvexe Hüllflächen lassen sich stets in konvexe Teile zerlegen

benötigen wir insbesondere die Masse m des Objektes, die Lage des Schwerpunktes $\hat{\mathbf{s}}$ in Objektkoordinaten, sowie den Trägheitstensor $\mathbf{I} \in \mathbb{R}^{3\times3}$. Bei topologisch konsistenten Beschreibungen der Objekthülle, die frei von Löchern und Selbstdurchdringungen ist, lassen sich diese Größen unter Verwendung von Annahmen über die Verteilung der Massendichte (z.B. homogene Dichteverteilung) direkt aus den Geometriedaten bestimmen. Ansonsten können auch Werte aus anderen Quellen, z.B. aus den Konstruktionsdaten eines Bauteiles, eingesetzt werden.

Desweiteren sind zur Simulation von Reibungseffekten (siehe Kapitel 7) die Reibungskoeffizienten vorzugeben.

Die zur Visualisierung eines Objektes erforderlichen Merkmale, wie Farben oder Oberflächenbeschaffenheit gehören ebenfalls zur Objektbeschreibung, spielen aber bei der eigentlichen Simulation keine Rolle.

2.2 Szenenbeschreibungen

Zur Beschreibung ganzer Szenen, und insbesondere der Lagebeziehungen verschiedener darin enthaltener Einzelkörper, dient ein gemeinsames Bezugskoordinatensystem, das Weltkoordinatensystem WKS. Die absolute Lage eines Objektes $\mathcal{O}_i$ im Raum wird nun durch eine Transformation T_i^{OW} zwischen OKS_i und WKS beschrieben. Die Koordinaten $\hat{\mathbf{x}} \in \mathbb{R}^3$ eines Objektpunktes im System OKS_i werden damit auf die entsprechenden Koordinaten $\mathbf{x} \in \mathbb{R}^3$ im WKS abgebildet durch die Beziehung

$$\mathbf{x} = T_i^{OW}(\hat{\mathbf{x}}).$$

Die Transformation T_i^{OW} setzt sich aus einem Translationvektor $\mathbf{t} \in \mathbb{R}^3$ und einer Rotationsmatrix $\mathbf{R} \in \mathbb{R}^{3\times3}$ zusammen. Damit erhält man die einfache Transformationsvorschrift

$$\mathbf{x} = \mathbf{R}\hat{\mathbf{x}} + \mathbf{t}. \tag{1}$$

Obwohl Rotationsmatrizen in der Anwendung sehr einfach sind, ist ihre Bestimmung eher unhandlich, da die darin enthaltenen drei freien Parameter (nämlich die drei Rotationswinkel) auf die 9 Matrixelemente verschmiert sind und somit einen hohen Grad an Redundanz enthalten. Zur Spezifikation von Rotationen bietet sich alternativ eine Beschreibung durch die Rotationsachse $\mathbf{r}$ und den Rotationswinkel φ an. Eine besonders geschickte Darstellung erlauben hierbei die Hamiltonschen Quaternionen $\mathfrak{q} = (q_0, \mathbf{q}) \in \mathbb{R}^4$ mit der Zuordnung

$$\mathfrak{q}_{\mathbf{r},\varphi} = \left(\cos\frac{\varphi}{2}, \mathbf{r}\sin\frac{\varphi}{2}\right).$$

Insbesondere entspricht die Verkettung mehrerer Rotationen der Multiplikation der zugehörigen Quaternionen. Details zur Arithmetik mit Quaternionen sind im Anhang A zusammengestellt.

Die zu (1) äquivalente Transformationsvorschrift für T_i^{OW} mit den Komponenten des Quaternions $\mathfrak{q} = (q_0, \mathbf{q})$ lautet

$$\mathbf{x} = \frac{2\mathbf{q}^T\hat{\mathbf{x}}\mathbf{q} + (q_0^2 - \mathbf{q}^2)\hat{\mathbf{x}} + 2q_0(\mathbf{q} \times \hat{\mathbf{x}})}{q_0^2 + \mathbf{q}^2} + \mathbf{t}. \tag{2}$$

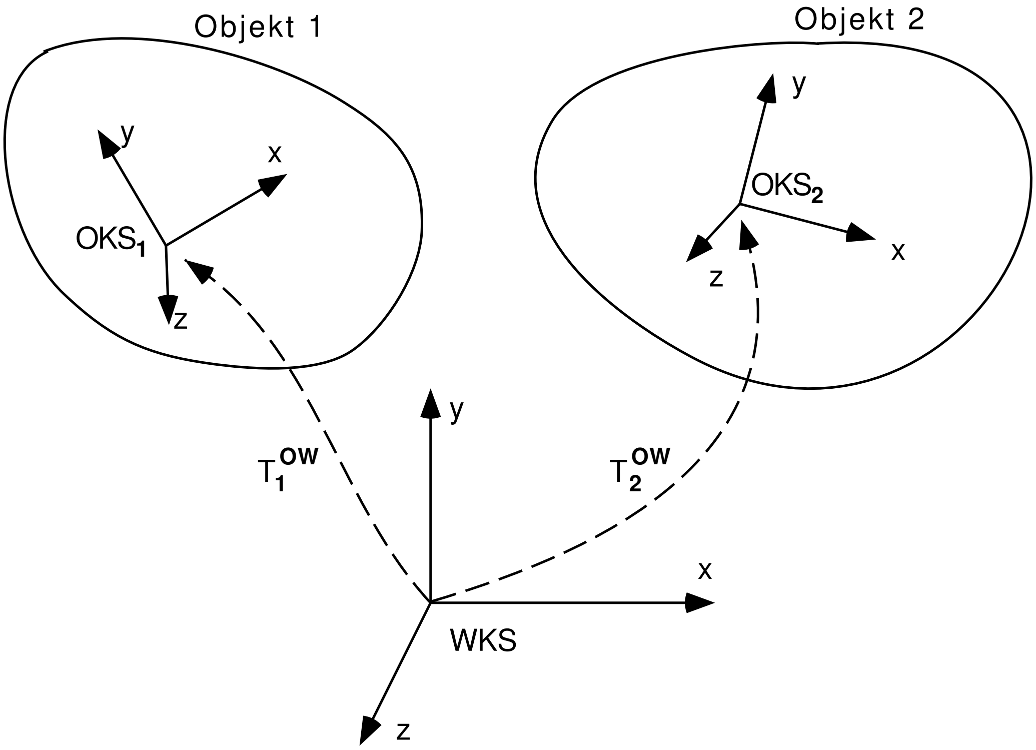


Abbildung 3: Zwei lokale Objektkoordinatensysteme OKS_1 und OKS_2, und ihr Bezug zum globalen Weltkoordinatensystem WKS.

2.3 Verwendete mathematische Darstellungsformen

Schließlich geben wir noch Schriftstilkonventionen sowie die Notation physikalischer Größen an, die wir in dieser Arbeit verwenden.

Schriftstilkonventionen		
t	Kleinbuchstabe	Skalar, Index
$\mathbf{a}$	Kleinbuchstabe fett	Vektor in Weltkoordinaten
$\hat{\mathbf{a}}$	Kleinbuchstabe fett mit Dach	Vektor in Objektkoordinaten
$\mathbf{C}$	Großbuchstabe fett	Matrix
$\boldsymbol{\mathfrak{q}}$	fraktur fett	Quaternion

Hilfsvariable	
$\mathbf{0}$	Nullvektor, Nullmatrix
$\mathbf{E}_n$	$n \times n$ Einheitsmatrix

Notation physikalischer Variablen	
$\mathbf{f}$, $\mathbf{F}$	Kraftvektor
$\mathbf{v}$	Geschwindigkeitsvektor
ω, $\boldsymbol{\omega}$	Winkelgeschwindigkeit
$\mathbf{T}$	Drehmoment
$\mathbf{q}$	verallgemeinerte Koordinaten (vergl. 3.2.1)
$\mathbf{u}$	verallgemeinerter Geschwindigkeitsvektor (vergl. 3.2.2)
m	Masse
$\mathbf{M}$	verallgemeinerte Massenmatrix (vergl. 3.2.2)
$\mathbf{I}$	Trägheitstensor
t	Zeit
$\mathbf{a}$	Beschleunigungsvektor
$\mathbf{p}$	Impulsvektor
μ	Reibungskoeffizient

3 Bewegungssimulation bei interaktiver Handhabung in virtuellen Welten

3.1 Elemente interaktiver Handhabung in virtuellen Welten

Die Simulation virtueller Realitäten (VR) unterscheidet sich von reinen Computergrafikanwendungen vor allem dadurch, daß virtuelle Welten dem Benutzer nicht nur realistisch dargestellt werden, sondern daß der Benutzer selbst aktiv in die virtuelle Welt eingreift und direkt mit ihr interagiert. Direkte Interaktion bedeutet hier, daß es dem Benutzer möglich wird, in der virtuellen Welt enthaltene Objekte mit Hilfe geeigneter Schnittstellen zu ergreifen, sie umherzubewegen, sie als Werkzeuge zu benutzen und damit realistische Handhabungsvorgänge durchzuspielen. Damit dies möglich wird, sind verschiedene Voraussetzungen erforderlich.

Zunächst müssen dem Benutzer Instrumente zur Verbindung seiner physikalischen Realität mit der ja nur als Computerrepräsentation vorliegenden virtuellen Welt zur Verfügung stehen. Hierzu gehören insbesonders Eingabegeräte, mit welchen virtuelle Objekte ausgewählt und die Bewegungen dieser Objekte innerhalb der virtuellen Welt intuitiv gesteuert werden können. Realistische Simulationen erfordern weiterhin, daß virtuelle Objekte gewissen Gesetzmäßigkeiten gehorchen wie ihre realen Vorbilder. Hierzu gehören die grundlegenden physikalischen Eigenschaften massiver Körper, sich im Falle einer Kollision nicht zu durchdringen, sondern stattdessen Ausweichbewegungen zu vollziehen, aneinander entlang zu gleiten, oder die Bewegung zu stoppen.

Für die Simulation interaktiver Handhabung bedeutet dies, daß entweder die vom Benutzer spezifizierte Bewegung eines virtuellen Gegenstandes bei Kontakten innerhalb der virtuellen Welt modifiziert werden muß, oder aber die anderen an dem Kontakt beteiligten Objekte entsprechend ausweichen. Die Bestimmung eines realistischen, den physikalischen Gesetzmäßigkeiten nahekommenden Bewegungsverhaltens in solchen Kontaktsituationen ist der eigentliche Gegenstand dieser Arbeit.

3.2 Repräsentation von Bewegungsbeschränkungen

Die Raumlage eines starren Körpers ist im allgemeinen durch 6 Freiheitsgrade der Bewegung gekennzeichnet, und läßt sich entsprechend durch 6 Parameter bzw. 6 Koordinaten beschreiben. Der zugehörige Parameterraum wird auch als Konfigurationsraum bezeichnet. Bei Vorliegen von Bewegungsbeschränkungen, z.B. aufgrund von Hinderniskontakten, ist die Anzahl der Bewegungsfreiheitsgrade entsprechend reduziert auf $m < 6$. Dies kann nun dadurch repräsentiert werden, daß ein reduzierter Satz von m Koordinaten aufgestellt wird, welcher genau die verbleibenden Bewegungsmöglichkeiten beschreibt. So lassen sich beispielsweise die möglichen Raumlagen, die ein Wagen einer Achterbahn annehmen kann, durch einen einzigen Bahnparameter eindeutig beschreiben, obwohl seine Bahn keineswegs geradlinig verläuft. Eine alternative Darstellung besteht darin, zwar alle 6 Koordinaten beizubehalten, jedoch die zulässigen Wertebereiche, z.B. durch die Einführung von Kontaktbedingungen, entsprechend einzuschränken. Wir wollen diese beiden Ansätze nun vergleichen.

3.2.1 Ansatz mit reduzierten Koordinaten

Die Raumlage eines starren Körpers sei Bewegungsbeschränkungen unterworfen, so daß ihm noch $m < 6$ Bewegungsfreiheitsgrade verbleiben. Dazu sei ein passender Satz von m Bewegungskoordinaten definiert, welcher die Bewegungsmöglichkeiten des Objektes vollständig beschreibt. Diesem Parametersatz entspricht ein m-dimensionaler Unterraum L des allgemeinen, 6-dimensionalen Konfigurationsraumes Ω, d.h. $L \subset \Omega$. Abb. 4 illustriert diesen Zusammenhang aus Darstellungsgründen anhand eines Beispieles niedrigerer Dimension. Hier spannt der Parametersatz $[x, y, z]$ den allgemeinen Konfigurationsraum auf. Der reduzierte Parametersatz $[u, v]$ spannt den Konfigurationsraum eines Objektes mit Bewegungsbeschränkungen auf, d.h. jede zulässige Objektlage läßt sich durch einen Vektor $[u, v]$ eindeutig beschreiben.

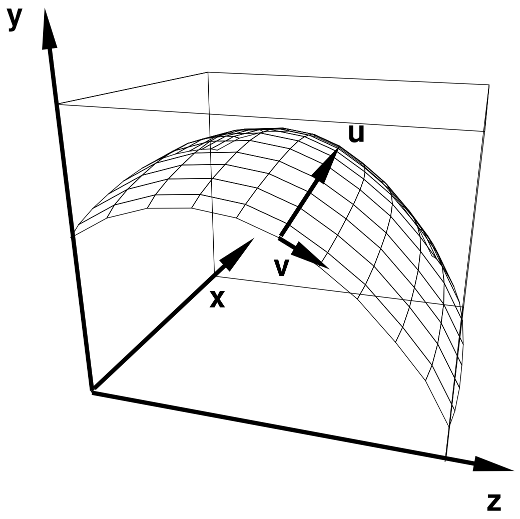


Abbildung 4: Beispiel für reduzierte Koordinaten: Der allgemeine Konfigurationsraum werde durch die Koordinaten $[x, y, z]$ aufgespannt. Ihm entspricht die unbeschränkte Objektbewegung. Kommen Bewegungsbeschränkungen hinzu, wird der Konfigurationsraum auf einen Unterraum geringerer Dimension reduziert. In diesem Beispiel wird der reduzierte Konfigurationsraum durch die Koordinaten $[u, v]$ aufgespannt.

Zur mathematischen Beschreibung des Ansatzes mit reduzierten Koordinaten bezeichne der Vektor $\mathbf{q} \in \mathbb{R}^6$ die Koordinaten im allgemeinen Konfigurationsraum, während $\mathbf{r} \in \mathbb{R}^m, m < 6$ die reduzierten Koordinaten enthalte, welche den Unterraum L parametrisieren.

Wegen $L \subset \Omega$ gibt es eine eindeutige Beziehung zwischen $\mathbf{q}$ und $\mathbf{r}$,

$$\mathbf{q} = \mathbf{F}(\mathbf{r}).$$

Eine zulässige Bewegung innerhalb von L läßt sich in beiden Koordinatensystemen lokal jeweils durch die zeitliche Ableitung darstellen. Die Beziehung zwischen diesen Ableitungen lautet

$$\dot{\mathbf{q}} = \frac{\partial \mathbf{F}(\mathbf{r})}{\partial \mathbf{r}}\,\dot{\mathbf{r}},$$

wobei

$$\mathbf{J} = \frac{\partial \mathbf{F}(\mathbf{r})}{\partial \mathbf{r}} \in \mathbb{R}^{6\times m}$$

die Jakobi-Matrix von $\mathbf{F}(\mathbf{r})$ darstellt.

An dieser Stelle wenden wir uns dem Problem zu, für eine vorgegebene Sollbewegung $\mathbf{u} = \dot{\mathbf{q}}$ eine resultierende Bewegung $\mathbf{u}'$ zu bestimmen, welche den vorliegenden Bewegungsbeschränkungen Rechnung trägt, welche also die Bedingung

$$\mathbf{u}' = \mathbf{J}\dot{\mathbf{r}} \tag{3}$$

erfüllt.

Der Vektor $\dot{\mathbf{r}}$ läßt sich allerdings nicht über $\dot{\mathbf{r}} = \mathbf{J}^{-1}\dot{\mathbf{q}}$ bestimmen, da $\mathbf{J} \in \mathbb{R}^{6\times m}$ nicht quadratisch, und somit auch nicht invertierbar ist. Die physikalische Interpretation dieses Zusammenhangs liegt darin, daß die Lösung dieses Problems ohne weitere Randbedingungen nicht eindeutig ist. Eine Näherungslösung erhält man über die sogenannte Pseudoinverse

$$\mathbf{J}^I = \left(\mathbf{J}^T\mathbf{J}\right)^{-1}\mathbf{J}^T,$$

durch

$$\begin{aligned} \mathbf{u}' &= \mathbf{J}\dot{\mathbf{r}} \\ &= \mathbf{J}\mathbf{J}^I\mathbf{u} \\ &= \mathbf{J}\left(\mathbf{J}^T\mathbf{J}\right)^{-1}\mathbf{J}^T\mathbf{u}. \end{aligned} \tag{4}$$

Diese Lösung minimiert das quadratische Fehlerfunktional

$$\begin{aligned} E &= (\dot{\mathbf{q}} - \mathbf{J}\dot{\mathbf{r}})^T(\dot{\mathbf{q}} - \mathbf{J}\dot{\mathbf{r}}) \\ &= (\mathbf{u} - \mathbf{u}')^T(\mathbf{u} - \mathbf{u}'). \end{aligned} \tag{5}$$

Die resultierende Bewegung $\mathbf{u}'$ ist allerdings nur lokal korrekt, da sie anschaulich einer Projektion der Bewegungsvorgabe $\mathbf{u}$ auf eine Tangentialebene des reduzierten Konfigurationsraumes entspricht [3]. Die Bewegung während eines betrachteten Zeitintervalls $[t, t + \Delta t]$ sollte daher im reduzierten Koordinatensystem integriert werden, um eine Fehlerakkumulation und damit ein Abdriften zu vermeiden. Die Objektlage $\mathbf{q}^{t+\Delta t}$ am Ende des betrachteten Zeitintervalls bestimmt sich dann zu

$$\mathbf{q}^{t+\Delta t} = \mathbf{F}(\mathbf{r}^t + \Delta t\,\mathbf{J}^I\mathbf{u}').$$

[3] Im Beispiel der Abb. 4 würde es sich um eine Tangentialebene an die gekrümmte und durch $[u, v]$ parametrisierte Fläche handeln

Gemessen an der numerischen Komplexität erscheint dieses Verfahren, welches in jedem Simulationsintervall lediglich die Lösung eines linearen Gleichungssystems der Ordnung m erfordert, zunächst attraktiv im Vergleich zu den iterativen Verfahren, die in Abschnitt 6 dargestellt werden.

Tatsächlich versteckt sich der größere Anteil des erforderlichen numerischen Aufwands jedoch in der Aufstellung der Funktion $\mathbf{F}$, die den Zusammenhang zwischen den reduzierten Koordinaten $\mathbf{r}$ und den allgemeinen Koordinaten $\mathbf{q}$ beschreibt. Denn nur in einfachen Situationen läßt sich die einem gegebenen Parametervektor $\mathbf{r}$ entsprechende Objektlage $\mathbf{q} = \mathbf{F}(\mathbf{r})$ geschlossen bestimmen, im allgemeinen Fall ist auch hierzu ein iteratives Näherungsverfahren erforderlich.

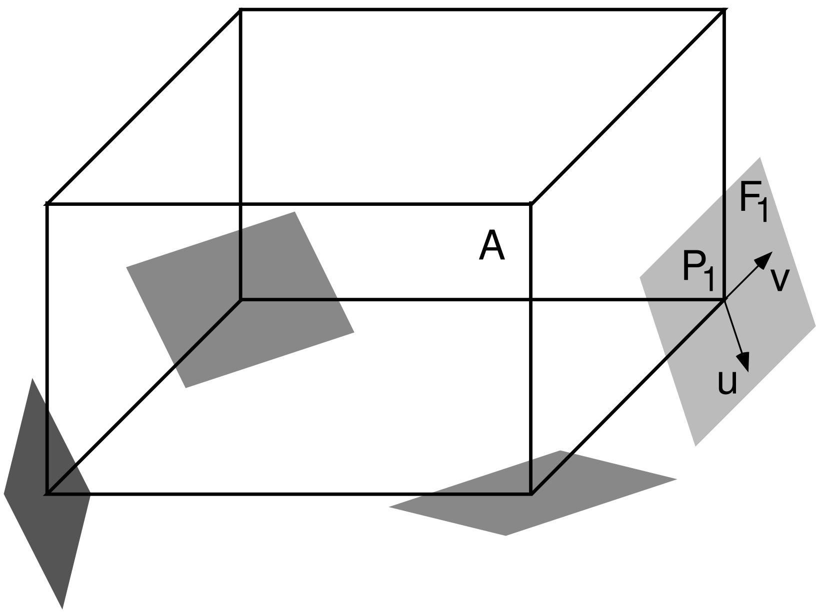


Abbildung 5: Die Lage des Quaders A ist durch vier Punkt-Fläche-Kontakte sowie eine Parametrisierung durch die lokalen Flächenkoordinaten (u, v) eindeutig festgelegt. Diese Lageparameter sind in Weltkoordinaten nicht mehr geschlossen darstellbar.

Ein Beispiel hierfür zeigt Abb. 5. Die Bewegung des Quaders A wird durch vier Hindernisflächen eingeschränkt, welche der Quader mit vier seiner Ecken berührt. Die verbleibenden 2 Bewegungsfreiheitsgrade werden durch die Position der Ecke P_1 innerhalb der Fläche F_1 mit Hilfe der lokalen Koordinaten (u, v) parameterisiert. Für einen gegebenen Parametersatz (u^*, v^*) gibt es bei Einhaltung der vier Kontakte genau eine mögliche Raumlage des Quaders, die sich allerdings nur mit Hilfe numerischer Näherungsverfahren bestimmen läßt.

Ein weiterer Nachteil dieses Verfahrens besteht darin, daß es rein geometrisch definiert ist, und keine physikalischen Größen wie beispielsweise Kräfte oder träge Masse berücksichtigt. Dadurch ist nicht offensichtlich, wie in der Simulation physikalische Effekte, beispielsweise Reibung oder Trägheitseffekte, realisiert werden könnten.

3.2.2 Ansatz mit Kontaktbedingungen und Kontaktkräften

Dieser klassische Ansatz, der in der Literatur auch als Lagrange-Ansatz bezeichnet wird [4], beruht auf der Einführung von Zwangskräften, welche die Berücksichtigung der Bewegungsbeschränkungen sicherstellen. Werden die Bewegungsbeschränkungen durch Hinderniskontakte verursacht, so handelt es sich bei den Zwangskräften um die an den Berührpunkten wirkenden Kontaktkräfte. Die Bewegung eines starren Körpers wird durch die insgesamt auf ihn wirkenden Kräfte bestimmt. Für den reibungsfreien Fall wirken die Kontaktkräfte jeweils in Richtung der Berührnormalen.

Zur mathematischen Beschreibung dieses Ansatzes definieren wir zunächst den verallgemeinerten Geschwindigkeitsvektor $\mathbf{u}$ sowie die Matrix $\mathbf{C}$ der Kontaktgeometrie wie folgt:

$$\mathbf{u} = \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\omega} \end{bmatrix}^T \in \mathbb{R}^6 \tag{6}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \dots & \mathbf{n}_k \\ \mathbf{r}_1 \times \mathbf{n}_1 & \mathbf{r}_2 \times \mathbf{n}_2 & \dots & \mathbf{r}_k \times \mathbf{n}_k \end{bmatrix} \in \mathbb{R}^{6\times k}. \tag{7}$$

Hierbei bezeichnen $\mathbf{u} \in \mathbb{R}^3$ und $\boldsymbol{\omega} \in \mathbb{R}^3$ die translatorische und rotatorische Geschwindigkeit, $\mathbf{n}_i \in \mathbb{R}^3, i = 1 \dots k$ die Normalenvektoren der Berührebenen an den k Kontaktpunkten, und $\mathbf{r}_i \in \mathbb{R}^3, i = 1 \dots k$ die Vektoren vom Massenschwerpunkt des betrachteten Objektes zu den Kontaktpunkten (vgl. Abb. 6).

.

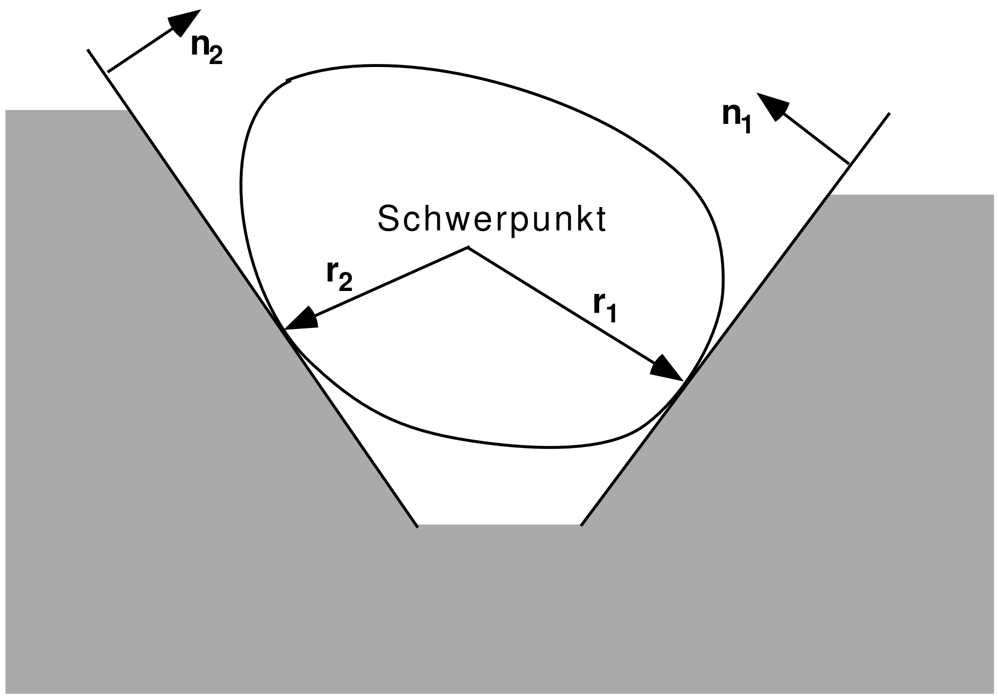


Abbildung 6: Bezeichnungen der Kontaktgeometrie.

Da die Richtungen $\mathbf{n}_i$ der Kontaktkräfte $\mathbf{f}_i = f_i\mathbf{n}_i$ bereits in der Matrix $\mathbf{C}$ repräsentiert sind, können wir die Beträge dieser Kräfte zu einem Vektor $\mathbf{f} = (f_1 \dots f_k)^T \in \mathbb{R}^k$ zusammenfassen.

Schließlich definieren wir die verallgemeinerte Massenmatrix $\mathbf{M}$

$$\mathbf{M} \quad = \quad \begin{bmatrix} m\mathbf{E}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \in \mathbb{R}^{6\times 6},$$

wobei $m \in \mathbb{R}$ die Masse, und $\mathbf{I} \in \mathbb{R}^{3\times 3}$ den Trägheitstensor bezeichnen.

Mit diesen Definitionen läßt sich die Objektbewegung kompakt darstellen als

$$\dot{\mathbf{u}} = \mathbf{M}^{-1}\mathbf{C}\mathbf{f} + \mathbf{a}^e. \tag{8}$$

Dabei seien in $\mathbf{a}^e$ sowohl die Wirkungen von Kreiselkräften als auch von allen sonstigen externen Kräften zusammengefaßt. An den Kontaktpunkten, welche die Bewegungsmöglichkeiten einschränken, darf in Richtung der Berührnormalen keine Bewegung stattfinden. Unter Verwendung der Kontaktgeometrie $\mathbf{C}$ entspricht dies der Bedingung $\mathbf{C}^T\mathbf{u} = \mathbf{0}$. Damit diese Bedingung bei bestehendem Kontakt aufrechterhalten wird, muß an den Kontaktpunkten auch die Beschleunigung in Richtung der Berührnormalen verschwinden, d.h. $\mathbf{C}^T\dot{\mathbf{u}} = \mathbf{0}$. Hieraus folgt nun die Bestimmungsgleichung für die Kontaktkräfte $\mathbf{f}$

$$\mathbf{0} = \mathbf{C}^T\mathbf{M}^{-1}\mathbf{C}\mathbf{f} + \mathbf{C}^T\mathbf{a}^e. \tag{9}$$

Durch Einsetzen in (8) ergibt sich die Objektbeschleunigung, und durch zeitliche Integration schließlich die gesuchte Objektbewegung.

Bei der numerischen Behandlung dieser Vorgehensweise stellt sich die Formulierung $\mathbf{C}^T\dot{\mathbf{u}} = \mathbf{0}$ der Kontaktbedingung jedoch als problematisch heraus. Bei der zeitdiskreten numerischen Integration von (8) akkumulieren sich kleine Fehler, was ohne weitere Maßnahmen dazu führt, daß die Berührbedingungen an den Kontaktpunkten im Laufe der Simulation nicht stabil eingehalten werden.

Um diese Unzulänglichkeit zu vermeiden, wählen wir eine in Abschnitt 3.4 beschriebene Vorgehensweise, welche die exakte Einhaltung der Kontaktbedingungen während der zeitlichen Integration der Bewegungsgleichungen erzwingt.

3.2.3 Kontaktkräfte oder Kontaktimpulse?

Bei der Bewegungssimulation mit Objektberührungen lassen sich zwei Situationen unterscheiden, bei denen jeweils die Durchdringung der beteiligten Objekte verhindert werden muß. Die erste Situation tritt bei einer Kollision genau im Kollisionszeitpunkt auf, wenn die beiden Objekte sich gerade berühren, sich an der Kontaktstelle aber noch aufeinander zubewegen. In diesem Fall kann die Durchdringung nur durch einen Kontaktimpuls $\Delta\mathbf{p} = \Delta t\,\mathbf{f} = m\Delta\mathbf{v}$ verhindert werden, der die erforderliche Geschwindigkeitsänderung bewirkt. Die Zeitspanne $\Delta t > 0$ wird hierbei als sehr kurz angenommen, ohne jedoch ganz zu verschwinden, da eine unendlich große Kontaktkraft $\mathbf{f}$ nicht auftreten kann.

Die zweite Situation betrifft einen bereits andauernden Kontakt, bei der die Relativgeschwindigkeit in Richtung der Kontaktnormalen gleich null ist. In diesem Fall wird kein Impuls ausgetauscht, und es wirkt lediglich die zur Aufrechterhaltung des Kontaktes erforderliche Kontaktkraft.

Der impulsbasierte Ansatz von Mirtich und Canny [35] setzt nun auch in diesem zweiten Fall zur Kontaktsimulation Impulse ein. Anstelle einer kontinuierlich wirkenden Kontaktkraft wirken dann viele kurz hintereinander auftretende Mikrokollisionen mit entsprechenden Impulsen, und die am Kontakt beteiligten Objekte vibrieren unmerklich gegeneinander. Die aufwendige Simulation dieser andauernden mikroskopischen Bewegung auch bei scheinbar ruhenden Gegenständen hat allerdings den Vorteil einer einheitlichen Behandlung von Kollisionen und bereits bestehenden Kontakten.

Es stellt sich nun die Frage, ob eine solche einheitlichen Behandlung auch mit ausschließlicher Verwendung von Kontaktkräften denkbar ist. Bei einer zeitdiskreten Bewegungssimulation ist eine solche Vorgehensweise nicht nur möglich, sondern bietet sich geradezu an. Für die Dauer Δt eines Simulationsschrittes wird dabei jeweils eine konstante Kontaktkraft angesetzt, die im Falle eines andauernden Kontaktes der kontinuierlichen Kontaktkraft entspricht. Die Wirkung dieser Kraft über das Zeitintervall hinweg entspricht aber gleichzeitig auch der Wirkung eines Impulses $\mathbf{p} = \Delta t\,\mathbf{f}$. Setzt man nun die Wirkungsdauer eines bei einer Kollision erforderlichen Kontaktimpulses gleich einem Simulationsschritt, so gehen beide Sichtweisen ineinander über, und eine Unterscheidung ist nicht mehr nötig.

Abb. 7 vergleicht den prinzipiellen zeitlichen Verlauf der Kontaktkraft bei einer Kollision mit anschließend fortdauerndem Kontakt (a) mit dem entsprechenden Verlauf bei einheitlicher Behandlung mit Kontaktimpulsen (b) sowie mit Kontaktkräften in einer zeitdiskreten Bewegungssimulation (c).

3.3 Bewegungsbeschränkungen bei interaktiven Handhabungssimulationen

Die Bewegung eines physikalischen Körpers folgt gemäß den Gesetzen der Mechanik den auf den Körper wirkenden Kräften und Momenten. Wird die Bewegung eines solcher Körpers interaktiv (z.B. durch die Hand eines Benutzers) gesteuert, so tragen hierzu auch die Kräfte und Momente bei, die durch die Handhabung auf den Körper übertragen werden. Umgekehrt wirken Reaktionskräfte vom gegriffenen Objekt auf die Hand zurück, wenn dieses beispielsweise gegen ein Hindernis stößt, und schränken dessen Bewegungsmöglichkeiten ebenfalls ein. Dieser in der physikalischen Welt selbstverständliche Zusammenhang stellt bei der Simulation von Handhabungsabläufen in virtuellen Umgebungen eine Schwierigkeit dar, wenn keine Kraftwechselwirkung zwischen der physikalischen Welt des Benutzers und der virtuellen Welt existiert. Bei der interaktiven Manipulation von Gegenständen in virtuellen Umgebungen soll die Bewegung eines manuell kontrollierten Objektes möglichst genau der Hand des Benutzers folgen. Bestehen in der virtuellen Welt jedoch Bewegungsbeschränkungen, z.B. aufgrund von Hinderniskollisionen, so wirken diese ohne Kraftrückkopplung nicht auf die Hand des Benutzer zurück, und die Bewegungsmöglichkeiten der Hand werden nicht entsprechend eingeschränkt. Folgt das gesteuerte virtuelle Objekt nun weiterhin der Bewegungsvorgabe durch die reale Hand, so werden die innerhalb der virtuellen Welt erforderlichen Bewegungsbeschränkungen bei der Objektbewegung nicht berücksichtigt.

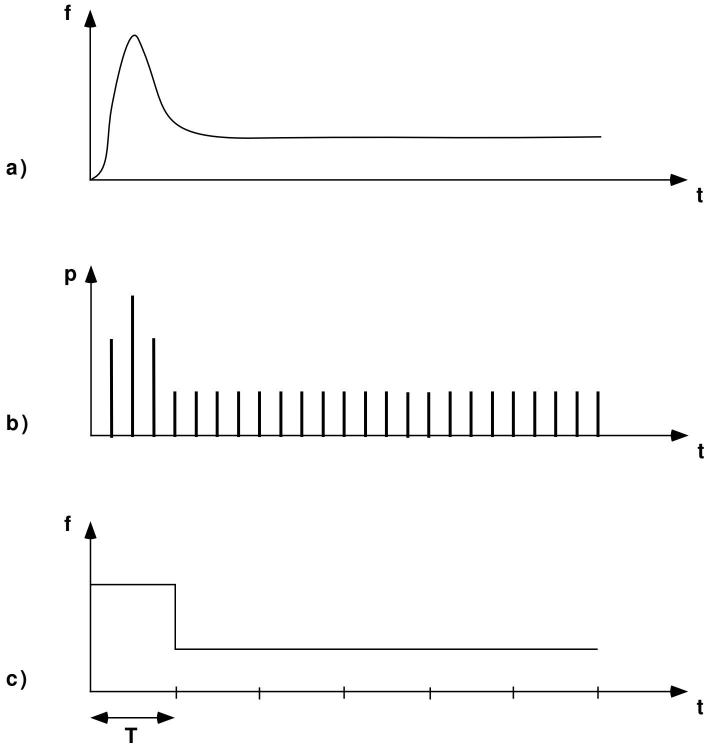


Abbildung 7: Zeitlicher Verlauf der Kontaktkraft bei einer Kollision (a) kontinuierlicher Verlauf, (b) mit Kontaktimpulsen, (c) mit zeitdiskreten Kontaktkräften.

Eine sinnvolle Simulation virtueller Umgebungen ist jedoch nur möglich, wenn das Verhalten der simulierten Objekte weitgehend dem entsprechender physikalischer Objekte entspricht. Dies bedeutet insbesondere, daß sich massive Körper nicht durchdringen dürfen, und kollisionsbedingte Bewegungsbeschränkungen einzuhalten sind.

Die Bewegungssimulation bei interaktiver Handhabung in virtuellen Welten muß daher in der Lage sein, aus der Bewegungsvorgabe durch die reale Hand des Benutzers zusammen mit den Randbedingungen innerhalb der virtuellen Umgebung eine glaubhafte, intuitiv nachvollziehbare, und vom Benutzer aufgrund seiner alltäglichen Erfahrung vorhersagbare Bewegung des gesteuerten Objektes zu bestimmen.

3.3.1 Beschreibung als Minimierungsaufgabe mit Randbedingungen

Gegeben sei ein Körper, welcher an k Kontaktpunkten mit festen Hindernissen in Berührung steht, und für den eine Sollbewegung vorgegeben sei, die sich aufgrund der Kontakte nicht realisieren läßt. Gesucht ist eine Ersatzbewegung, welche die Kontaktbedingungen berücksichtigt. Ein möglicher Ansatz zur Lösung dieser Aufgabe ist die Formulierung einer geeigneten Abstandsfunktion, welche unter gegebenen Randbedingungen minimiert werden soll.

Zur mathematischen Beschreibung dieser Minimierungsaufgabe verwenden wir wieder den in Abschnitt 3.2.2 definierten verallgemeinerten Geschwindigkeitsvektor $\mathbf{u} \in \mathbb{R}^6$ und die Matrix $\mathbf{C} \in \mathbb{R}^{6\times k}$ der Kontaktgeometrie gemäß Gln. (6) und (7).

Weiterhin bezeichne $\mathbf{u}' \in \mathbb{R}^6$ die gesuchte Ersatzbewegung des Objektes und $\mathbf{u}$ die vom Benutzer interaktiv vorgegebene Sollbewegung. Zur Bestimmung von $\mathbf{u}'$ wählen wir ein quadratisches Fehlermaß E der Form

$$E = \frac{1}{2}(\mathbf{u}' - \mathbf{u})^T \mathbf{G}(\mathbf{u}' - \mathbf{u}), \tag{10}$$

welches unter Einhaltung der k Kontaktbedingungen minimiert werden soll. Dabei bezeichnet $\mathbf{G}$ eine Diagonalmatrix, welche durch eine geeignete Gewichtung der Komponenten von $\mathbf{v}$ und $\boldsymbol{\omega}$ eventuelle Skalierungseffekte kompensiert [4].

Die Kontaktbedingungen lassen sich darstellen durch die Beziehung

$$\mathbf{C}^T \mathbf{u}' = \mathbf{0}. \tag{11}$$

Anschaulich bedeutet dies, daß die Normalkomponenten der Kontaktgeschwindigkeiten verschwinden sollen.

Zur Bestimmung einer Geschwindigkeit $\mathbf{u}'$, welche $E(\mathbf{u}')$ minimiert, wählen wir den Ansatz der Lagrange-Multiplikatoren, formulieren die Bedingung

$$\frac{\partial}{\partial \mathbf{u}'}\left(E(\mathbf{u}') - \boldsymbol{\lambda}^T \mathbf{C}^T \mathbf{u}'\right) \quad = \quad \mathbf{0}$$

und erhalten nach einigen Umformungen und unter Verwendung von (11) den Vektor der Lagrange-Multiplikatoren $\boldsymbol{\lambda}$:

$$\begin{aligned} \mathbf{G}(\mathbf{u}' - \mathbf{u}) - \mathbf{C}\boldsymbol{\lambda} &= \mathbf{0} \\ \mathbf{u}' &= \mathbf{u} + \mathbf{G}^{-1}\mathbf{C}\boldsymbol{\lambda} \qquad (12) \\ \mathbf{C}^T\mathbf{u}' = \mathbf{C}^T\mathbf{u} + \mathbf{C}^T\mathbf{G}^{-1}\mathbf{C}\boldsymbol{\lambda} &= \mathbf{0} \\ \boldsymbol{\lambda} &= -\left(\mathbf{C}^T\mathbf{G}^{-1}\mathbf{C}\right)^{-1}\mathbf{C}^T\mathbf{u}. \qquad (13) \end{aligned}$$

Nach Einsetzen von (13) in (12) erhält man schließlich die gesuchte Geschwindigkeit $\mathbf{u}'$:

$$\mathbf{u}' = \mathbf{u} - \mathbf{G}^{-1}\mathbf{C}\left[\mathbf{C}^T\mathbf{G}^{-1}\mathbf{C}\right]^{-1}\mathbf{C}^T\mathbf{u}. \tag{14}$$

3.3.2 Physikalisch orientierter Ansatz

Alternativ zu dieser abstrakten Vorgehensweise läßt sich diesselbe Aufgabe auch in einer eher physikalisch motivierten Weise beschreiben. Dazu führen wir an den Kontaktpunkten Kräfte ein, welche die Kontaktbedingungen aufrechterhalten sollen. Wir beschränken uns im Moment auf den reibungsfreien Fall, so daß die Kontaktkräfte jeweils in Richtung der Berührnormalen $\mathbf{n}_i$ wirken. Die Beträge der Kontaktkräfte seien wieder zu einem Vektor $\mathbf{f} \in \mathbb{R}^k$ zusammengefaßt.

[4] Eine Maßstabsänderung berührt nur die translatorischen Geschwindigkeitskomponenten $\mathbf{v}$, nicht aber die Winkelgeschwindigkeit $\boldsymbol{\omega}$. Um das Fehlermaß unabhängig von einem Skalierungsfaktor s zu machen, muß G proportional zu $\begin{bmatrix} \mathbf{E}_3 & \mathbf{O} \\ \mathbf{O} & s^2\mathbf{E}_3 \end{bmatrix}$ sein.

Um eine kompakte Beschreibung zu erhalten, verwenden wir erneut die Matrix $\mathbf{C}$ der Kontaktgeometrie, sowie die verallgemeinerte Massenmatrix $\mathbf{M} \in \mathbb{R}^{6\times 6}$ gemäß Gl. (8).

Die extern vorgegebene Sollbewegung werde wieder durch den verallgemeinerten Geschwindigkeitsvektor $\mathbf{u}$ bezeichnet und die gesuchte resultierende Bewegung durch $\mathbf{u}'$.

Mit diesen Definitionen erhält man unter Vernachlässigung der Kreiselkräfte die Bewegungsgleichung

$$\mathbf{u}' = \mathbf{u} + \Delta t\, \mathbf{M}^{-1}\mathbf{C}\mathbf{f}. \tag{15}$$

Da die Form von (15) genau derjenigen von (12) enspricht, erhält man analog zu oben das Ergebnis

$$\mathbf{u}' = \mathbf{u} - \mathbf{M}^{-1}\mathbf{C}\left[\mathbf{C}^T\mathbf{M}^{-1}\mathbf{C}\right]^{-1}\mathbf{C}^T\mathbf{u}. \tag{16}$$

Ein Vergleich zwischen (14) und (16) ergibt, daß beide Ergebnisse identisch sind, wenn man die Gewichtungsmatrix $\mathbf{G}$ mit der Massenmatrix $\mathbf{M}$ identifiziert. Der physikalische Ansatz realisiert also genau die geforderte Minimierung des verwendeten Abstandsmasses. Um die Plausibilität dieses Ergebnis zu untermauern, prüfen wir noch, ob die Massenmatrix auch die geforderten Eigenschaften der Gewichtungsmatrix $\mathbf{G}$ besitzt.

Die Komponenten $I_{ij}, i, j = 1\ldots 3$ des Trägheitstensors $\mathbf{I}$ bestimmen sich zu $I_{ij} = \int\limits_V x_i x_j\, \mathrm{d}m$. Bei einer Skalierung der Geometrie mit dem Faktor s sind diese Komponenten alle proportional zu ms^2. Damit hat $\mathbf{M}$ genau die zur Kompensation von Skalierungseffekten erforderliche Struktur.

Ein weiterer Vergleich von (15) und (12) ergibt, daß die Kontaktkräfte $\Delta t\, \mathbf{f}$ gerade den Lagrange-Multiplikatoren $\boldsymbol{\lambda}$ entsprechen. Verfahren, bei welchen Bewegungsbeschränkungen durch Kontaktkräfte realisiert sind, werden daher in der Literatur (z.B.[4]) auch als Lagrange-Ansätze bezeichnet [5].

3.4 Bewegungssimulation mit Stabilisierung der Kontaktbedingungen

Die Fehlerakkumulation bei der in Abschnitt 3.2.2 beschriebenen klassischen Vorgehensweise zur Simulation eingeschränkter Objektbewegungen rührt daher, daß die Randbedingungen für die Objektbeschleunigung formuliert werden, und nicht für die Objektposition. Das nun vorgestellte Verfahren unterscheidet sich hiervon wesentlich dadurch, daß die Randbedingungen der Bewegungsbeschränkungen unmittelbar für die Objektposition formuliert werden, und somit keiner Fehlerakkumulation unterworfen sind.

[5] Baraff und Mattikalli diskutieren in [2] eine verwandte Fragestellung. Sie vergleichen drei Möglichkeiten, die initiale Bewegungsrichtung eines zu Beginn ruhenden Objektes zu bestimmen, welches sich unter dem Einfluß der Schwerkraft entlang gegebener Kontaktflächen zu bewegen beginnt. Die erste Formulierung bestimmt die initiale Beschleunigung aus den Gewichtskräften und den wirkenden Kontaktkräften, die durch ein quadratisches Programm bestimmt werden. Der zweite Ansatz beruht auf Gauss' Prinzip des geringsten Widerstandes. Der dritte Ansatz betrachtet das Schwerefeld als Potentialfunktion und unterstellt, daß Objekte dem steilsten Gradienten dieses Potentials abwärts folgen. Alle drei Vorgehensweisen führen zum selben Resultat.

Um dies zu ermöglichen, zerlegen wir die Ausgangsgleichung (8) zunächst in ihre Komponenten, und erhalten die bekannten Newton-Euler-Gleichungen für die Bewegung eines starren Körpers:

$$\dot{\mathbf{v}} = \frac{1}{m}\sum_{i=1}^{k}\mathbf{f}_i \tag{17}$$

$$\dot{\boldsymbol{\omega}} = \mathbf{I}^{-1}\left(\sum_{i=1}^{k}\mathbf{r}_i \times \mathbf{f}_i - \boldsymbol{\omega}\times\mathbf{I}\boldsymbol{\omega}\right). \tag{18}$$

Dabei bezeichnen $\mathbf{f}_i$ die wirkenden äußeren Kräfte, $\mathbf{r}_i$ die Vektoren vom Massenschwerpunkt des betrachteten Körpers zu den Punkten, an denen die Kräfte ansetzen, m die Masse des Körpers und $\mathbf{I}$ seinen Trägheitstensor. Im folgenden nehmen wir zur Vereinfachung der Darstellung an, daß auf den Körper als äußere Kräfte nur eine Anzahl von k Kontaktkräften wirke. Andere äußere Kräfte, wie z.B. die Schwerkraft, lassen sich problemlos hinzufügen, spielen aber für die weitere Diskussion zur Bestimmung der Kontaktkräfte in diesem Abschnitt keine Rolle.

Die Raumlage des Körpers ergibt sich durch zeitliche Integration der Gleichungen (17) und (18) sowie aus den Beziehungen

$$\dot{\mathbf{c}} = \mathbf{v} \tag{19}$$

$$\dot{\mathfrak{q}} = \frac{1}{2}\boldsymbol{\omega}\mathfrak{q}, \tag{20}$$

wobei der Verschiebungsvektor $\mathbf{c}$ den Ort und das Quaternion $\mathfrak{q}$ die Orientierung in Weltkoordinaten beschreiben. Die Darstellung von Rotationen durch Quaternionen ist im Anhang A genauer beschrieben. Dort findet sich auch eine Herleitung des Quaternionenproduktes $\boldsymbol{\omega}\cdot\mathfrak{q}$ in der Gleichung (20).

Um diese Bewegungsgleichungen numerisch integrieren zu können, diskretisieren wir sie, und erhalten die Differenzengleichungen

$$\mathbf{v}^{t+\Delta t} = \mathbf{v}^t + \Delta t\,\frac{1}{m}\sum_{i=1}^{k} f_i\mathbf{n}_i^t \tag{21}$$

$$\boldsymbol{\omega}^{t+\Delta t} = \boldsymbol{\omega}^t + \Delta t\,(\mathbf{I}^t)^{-1}\left(\sum_{i=1}^{k} f_i\mathbf{r}_i^t\times\mathbf{n}_i^t - \boldsymbol{\omega}^t\times\mathbf{I}^t\boldsymbol{\omega}^t\right). \tag{22}$$

In den Gln. (21) und (22) bezeichnen $\mathbf{n}_i$ die Kraftrichtungen, und f_i die Beträge der Kräfte. Die Integration erfolgt dabei über das Zeitintervall $[t, t+\Delta t]$.

Die Diskretisierung von (23) und (24) führt auf die folgenden Beziehungen für die Objektposition:

$$\mathbf{c}^{t+\Delta t} = \mathbf{c}^t + \Delta t\,\mathbf{v}^{t+\Delta t} \tag{23}$$

$$\mathfrak{q}^{t+\Delta t} = \mathfrak{q}^t + \frac{1}{2}\Delta t\,(0,\,\boldsymbol{\omega}^{t+\Delta t})\cdot\mathfrak{q}^t. \tag{24}$$

Dabei stellt $(0,\,\boldsymbol{\omega}^{t+\Delta t})$ die Verallgemeinerung von $\boldsymbol{\omega}^{t+\Delta t}$ zu einem Quaternion dar. Da jedem Quaternion $\mathfrak{q}$ eine Rotationsmatrix $\mathbf{R}(\mathfrak{q})$ eindeutig zugeordnet ist (vgl. Anhang A), läßt sich Gleichung (24) alternativ auch mit Rotationsmatrizen wie folgt formulieren:

$$\mathbf{R}(\mathfrak{q}^{t+\Delta t}) \quad = \quad \mathbf{R}(1,\, \frac{1}{2}\Delta t\, \boldsymbol{\omega}^{t+\Delta t}) \cdot \mathbf{R}(\mathfrak{q}^{t})$$

oder kurz

$$\mathbf{R}^{t+\Delta t} \quad = \quad \mathbf{R}(1,\, \frac{1}{2}\Delta t\, \boldsymbol{\omega}^{t+\Delta t}) \cdot \mathbf{R}^{t}. \tag{25}$$

Man beachte, daß bei der Diskretisierung für (21) und (22) Vorwärtsdifferentiation, dagegen für (23) und (24) bzw. (25) Rückwärtsdifferentiation gewählt wurde. Dadurch ist es möglich, (21) in (23) und (22) in (25) direkt einzusetzen. Insgesamt erhält man damit ein Gleichungssystem für die Position $\mathbf{c}^{t+\Delta t}$ und die Orientierung $\mathbf{R}^{t+\Delta t}$ des Körpers am Ende des Zeitintervalles Δt als Funktion der unbekannten Kontaktkräfte $\mathbf{f} = (f_1 \ldots f_k)^T$.

Die Kontaktbedingungen lassen sich damit, wie gefordert, für die Raumlage des Körpers formulieren, anstatt für dessen Geschwindigkeit oder Beschleunigung. Die Kontaktbedingungen $g_i, i = 1 \ldots k$ beziehen sich dabei direkt auf die Kontaktabstände w_i, wie sie in Kapitel 4.5 dargestellt werden, und nehmen die allgemeine Form

$$w_i = g_i(\mathbf{R}^{t+\Delta t},\, \mathbf{c}^{t+\Delta t}) = 0, \qquad i = 1 \ldots k \tag{26}$$

an. Setzt man die Bewegungsgleichungen in diese Kontaktbedingungen ein, und faßt diese zu einem Vektor $\mathbf{g}$ zusammen, so erhält man schließlich ein nichtlineares Gleichungssystem

$$\mathbf{g}(\mathbf{f}) = \mathbf{0}, \qquad \mathbf{g} : \mathbb{R}^k \to \mathbb{R}^k \tag{27}$$

zur Bestimmung der Kontaktkräfte $\mathbf{f}$.

In den folgenden Kapiteln 4 und 5 wird diese Beziehung erweitert werden auf ein Gleichungssystem der Form

$$\mathbf{g}(\mathbf{f}) \geq \mathbf{0}, \qquad \mathbf{f} \geq \mathbf{0}, \qquad \mathbf{g}^T\mathbf{f} = 0.$$

In Kapitel 6 wird auf die Lösung dieses Systems detailliert eingegangen. Durch die dargestellte spezielle Formulierung wird dabei erreicht, daß die Einhaltung der Kontaktbedingungen im Laufe der Simulation stets gewährleistet ist.

3.5 Erweiterung auf mehrere bewegliche Objekte

Die vorangegangenen Darstellungen beschränkten sich auf den Fall eines einzelnen beweglichen Objektes, welches mit feststehenden Hindernissen in Kontakt steht und dadurch in seiner Bewegungfreiheit eingeschränkt ist. Wir betrachten nun die Erweiterung dieser Situation, indem wir mehrere bewegliche Objekte in Betracht ziehen, welche sich insbesondere auch gegenseitig berühren können. Dieser Fall ist für viele praktische Anwendungen von großer Bedeutung. Man denke beispielsweise an die Simulation eines interaktiv geführten Werkzeuges, welches auf bewegliche Werkstücke einwirkt (z.B. Schraubenzieher und Schraube).

Solange die beweglichen Objekte sich nicht gegenseitig berühren, sondern jeweils nur feststehende Hindernisse, läßt sich die eingeschränkte Bewegung jedes Objektes unabhängig betrachten. Es genügt in diesem Fall, die oben beschriebene Vorgehensweise auf jedes der Objekte anzuwenden. Der numerische Aufwand ist dabei direkt proportional zur Anzahl der beweglichen Objekte.

Wenn sich bewegliche Objekte auch gegenseitig berühren, schränken sie ihre Bewegungsmöglichkeiten wechselseitig ein, und können daher nicht mehr unabhängig betrachtet werden. Wir bezeichnen eine Menge von beweglichen Objekten, welche entweder direkt oder über andere bewegliche Objekte miteinander in Kontakt stehen, als Kontaktcluster (vgl. Abb. 8).

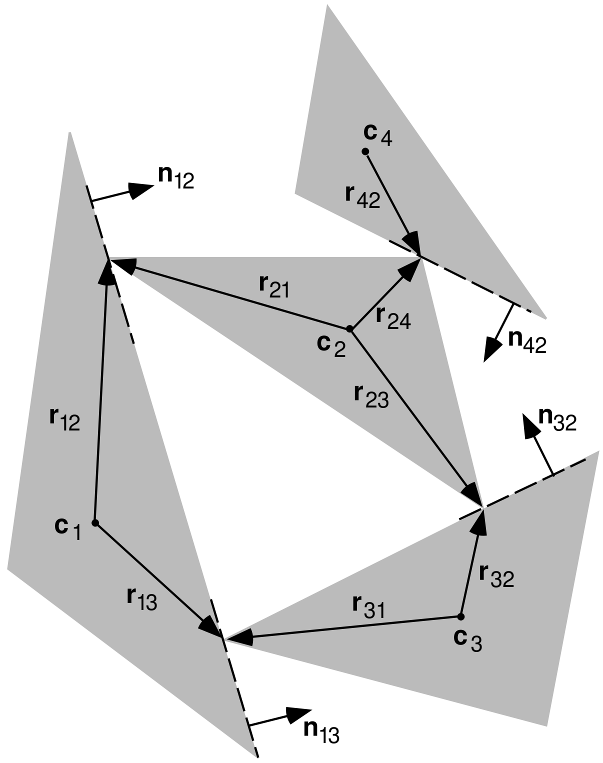


Abbildung 8: Darstellung der zur Beschreibung eines Kontaktclusters verwendeten Größen (siehe Text).

Zur Beschreibung der Simulation eines Kontaktclusters aus n Objekten mit insgesamt k Kontakten verallgemeinern wir einige der eingeführten Größen wie folgt.

Die Matrix der Kontaktgeometrien $\mathbf{C} \in \mathbb{R}^{6n\times k}$ hat nun die Form

$$\mathbf{C}^T = \begin{array}{c} \begin{array}{ccccccc} & i & & & j & & \\ & \downarrow & & & \downarrow & & \end{array} \\ \left[\begin{array}{ccccccc} 0\ldots0 & \mathbf{n}_{ij}^T & (\mathbf{r}_{ij}\times\mathbf{n}_{ij})^T & 0\ldots0 & \mathbf{n}_{ji}^T & (\mathbf{r}_{ji}\times\mathbf{n}_{ji})^T & 0\ldots0 \end{array} \right] \end{array} .$$

Dabei bezeichnet $\mathbf{r}_{ij}$ den Vektor vom Schwerpunkt $\mathbf{c}_i$ des Körpers i zum Kontaktpunkt zwischen den Körpern i und j und $\mathbf{n}_{ij} = -\mathbf{n}_{ji}$ steht für die Normale der Kontaktebene zwischen diesen beiden Körpern. Jeder Spaltenvektor von $\mathbf{C}$ hat im Falle eines Kontaktes zwischen den Körpern i und j je einen Eintrag der Form $[\mathbf{n}^T, (\mathbf{r}\times\mathbf{n})^T]^T$ mit Indices (ij) und (ji), oder aber im Falle eines Kontaktes zwischen Körper i und einem feststehenden Hindernis nur einen solchen Eintrag mit Index i.

Weiterhin sei $\mathbf{u} \in \mathbb{R}^{6n}$ der Vektor der verallgemeinerten Geschwindigkeiten

$$\mathbf{u}^T = [\mathbf{v}_1, \boldsymbol{\omega}_1, \ldots, \mathbf{v}_n, \boldsymbol{\omega}_n]^T$$

und $\mathbf{M} \in \mathbb{R}^{6n\times 6n}$ die verallgemeinerte Massenmatrix

$$\mathbf{M} = \begin{bmatrix} m_1\mathbf{E} & & & & 0 \\ & \mathbf{I}_1 & & & \\ & & \ddots & & \\ & & & m_n\mathbf{E} & \\ 0 & & & & \mathbf{I}_n \end{bmatrix}.$$

Mit $\mathbf{f} \in \mathbb{R}^k$ sei der Vektor der Kontaktkräfte bezeichnet und mit $\mathbf{a}^e$ der Vektor der externen Beschleunigungen,

$$(\mathbf{a}^e)^T = [\mathbf{g}, -\mathbf{I}_1^{-1}\boldsymbol{\omega}_1 \times \mathbf{I}_1\boldsymbol{\omega}_1, \ldots, \mathbf{g}, -\mathbf{I}_n^{-1}\boldsymbol{\omega}_n \times \mathbf{I}_n\boldsymbol{\omega}_n]^T.$$

Mit Hilfe dieser Bezeichnungen können wir die Bewegungsgleichung (8) für das Mehrkörperproblem unverändert formulieren:

$$\dot{\mathbf{u}} = \mathbf{M}^{-1}\mathbf{C}\mathbf{f} + \mathbf{a}^e.$$

Die Diskretisierung des Gleichungssystems und die Aufstellung des nichtlinearen Gleichungssystems zur Bestimmung der Kontaktkräfte geschieht ganz analog zu der für ein einzelnes bewegtes Objekt beschriebenen Vorgehensweise. Man erhält auch hier ein System

$$\mathbf{g}(\mathbf{f}) = \mathbf{0}, \mathbf{g} : \mathbb{R}^k \to \mathbb{R}^k,$$

welches iterativ zu lösen ist.

Identifikation von Kontaktclustern

Die Körper, welche zu einem gemeinsamen Kontaktcluster gehören, lassen sich effizient mit einem Verfahren zur Bestimmung von Zusammenhangsgraphen bestimmen ([13]). Alle in Frage kommenden Objekte stellen dabei die Knoten des Graphen dar, und alle Kontaktpunkte entsprechen den Kanten des Graphen. Gesucht sind nun diejenigen Gruppen von Knoten, welche direkt oder indirekt über Kanten miteinander in Verbindung stehen.

Wir verwenden hierzu eine depth-first Strategie (vgl. untenstehenden Pseudocode). Jeder Knoten durchläuft dabei die Stadien *unbekannt*, *bekannt* und *fertig*. Zu Beginn sind alle Knoten unbekannt. Ausgehend von einem Startknoten K_i werden alle mit diesem über eine Kante verbundenen adjazenten Knoten $K_j \in Adj[K_i]$ untersucht. Wird dabei ein noch unbekannter Knoten entdeckt, so wird dieser in den Zustand *bekannt* versetzt, und die depth-first-Suche wird rekursiv bei K_j fortgesetzt. Hat ein Knoten keine adjazenten unbekannten Knoten mehr, so geht er in den Status *fertig* über, und die Suche kehrt zur nächsthöheren Rekursionsebene zurück. Während des Suchprozesses werden alle neu entdeckten Cluster-Elemente in eine Cluster-Liste aufgenommen. Ein Kontaktcluster ist vollständig identifiziert, wenn der Startknoten K_i in den Zustand *fertig* übergeht.

Anschliessend wird der gesamte Graph erneut nach unbekannten Knoten durchsucht. Wird ein solcher entdeckt, dient er als Startknoten zur Identifikation eines weiteren Kontaktclusters. Die Analse wird solange fortgesetzt, bis sich alle Knoten des Graphen im Zustand *fertig* befinden.

```
clusterAnalyse {
    for each K_i {
        status[K_i] = unbekannt
    }
    for each K_i {
        if (status[K_i] == unbekannt) {
            initClusterList()
            depthFirstSearch (K_i)
            processClusterList()
        }
    }
}
```

```
depthFirstSearch (K_i) {
    addToClusterList(K_i)
    status[K_i] = bekannt
    for each K_j ∈ Adj[K_i] {
        if (status[K_j] == unbekannt) {
            depthFirstSearch (K_j)
        }
    }
    status[K_i] = fertig
}
```

4 Beschreibung von Kontakten zwischen Objekten

4.1 Allgemeines

Aufgrund der polyhedralen Gestalt der betrachteten Körper lassen sich alle Kontaktsituationen zwischen zwei solchen Körpern auf die 9 Kontaktarten zurückführen, welche sich durch die Kombination der 3 Geometrieelemente Ecke, Kante und Fläche ergeben. Diese Kombinationen lassen sich, wie in Tabelle 1 dargestellt, durch die Dimension der zugehörigen Kontaktregionen klassifizieren.

Dimension der Kontaktregionen			
	Fläche	Kante	Ecke
Fläche	2	1	0
Kante	1	0	(0)
Ecke	0	(0)	(0)

Tabelle 1: Klassifikation der Kontaktpaare nach der Dimension der zugehörigen Kontaktregionen. Die mit Klammern markierten degenerierten Kontaktpaare werden im Text näher erläutert.

Neben der Dimension der Kontaktregion läßt sich jedem Kontaktpaar auch eine Kontaktnormale $\mathbf{n}$ sowie die zugehörige Berührebene zuordnen, welche die betrachteten Objekte im Bereich der Kontaktregion separiert. Wenn an einem Kontaktpaar eine Fläche beteiligt ist, fällt die Kontaktnormale mit der Flächennormale zusammen (Abb. 9a). Im Falle eines Kante-Kante-Kontaktes (Abb. 9b) bestimmt sich die Kontaktnormale aus den beiden Kantenrichtungen $\mathbf{u}_1$, $\mathbf{u}_2$ durch $\mathbf{n} = \mathbf{u}_1 \times \mathbf{u}_2$. In den verbleibenden Fällen, welche in Tabelle 1 durch Klammern markiert sind, ist zunächst keine eindeutige Kontaktnormale angebbar.

Um eine einheitliche Vorgehensweise zur Berücksichtigung von Kontakten im Rahmen der Bewegungssimulation zu ermöglichen, zeigen wir in den folgenden Abschnitten zunächst, daß sich alle in Tabelle 1 aufgeführten Kontakttypen durch die Punktförmigen Kontaktarten Ecke-Fläche bzw. Fläche-Ecke und Kante-Kante substituieren lassen.

4.2 Substitution von Linien- und Flächenkontakten durch Punktkontakte

Eine linienhafte Kontaktregion entsteht bei der Berührung einer Kante K_1 eines Objektes O_1 mit einer konvexen Fläche F_2 eines weiteren Objektes O_2 (Abb.10a) als Schnittmenge $S = K_1 \cap F_2$. Seien $\mathbf{a}_1$ und $\mathbf{a}_2$ zwei verschiedene Punkte in K_1, und die Fläche F_2 liege in der Ebene $E : \mathbf{n}_E^T\mathbf{x} = d_E$. Damit sind die Aussagen

$$\mathbf{n}_E^T\mathbf{a}_1 - d_E = 0 \wedge \mathbf{n}_E^T\mathbf{a}_2 - d_E = 0$$

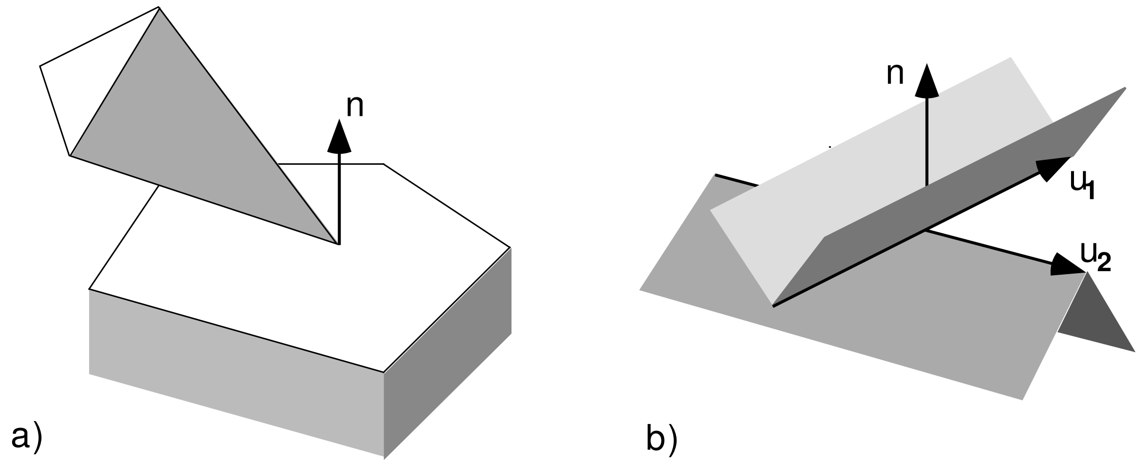


Abbildung 9: Bei einem Punkt-Fläche-Kontakt ist die Berührnormale durch die Normale der Kontaktebene definiert (a). Bei einem Kante-Kante-Kontakt steht die Berührnormale $\mathbf{n}$ orthogonal zu den beiden Kantenrichtungen $\mathbf{u}_1$ und $\mathbf{u}_2$ (b).

und

$$\mathbf{n}_E^T(\mathbf{a}_1 + \lambda(\mathbf{a}_2 - \mathbf{a}_1)) - d_E = 0 \quad \forall \lambda \in [0,1]$$

äquivalent. Hieraus folgt, daß S genau dann eine linienhafte Kontaktregion darstellt, wenn es zwei verschiedene Punkte $\mathbf{a}_1$, $\mathbf{a}_2 \in K_1$ gibt, für die gilt $\mathbf{a}_1, \mathbf{a}_2 \in F_2$. Ein Linienkontakt kann somit durch zwei Punktkontakte substituiert werden, welche innerhalb der Kontaktregion liegen, ohne daß sich an der Kinematik der beteiligten Polyeder etwas ändert (Abb. 10b).

Eine analoge Überlegung läßt sich für einen flächenhaften Kontakt (Abb. 10c) anstellen, und man erhält das Ergebnis, daß ein solcher Kontakt durch drei Punktkontakte ersetzbar ist, welche innerhalb der flächigen Kontaktregion liegen (Abb. 10d).

Dieses Ergebnis ist sowohl intuitiv nachvollziehbar, als auch physikalisch nicht unvernünftig, wenn man berücksichtigt, daß bei realen Objekten vollkommen ebene Flächen und vollkommen gerade Kanten nicht zu erwarten sind. Diese idealisierten Strukturen sind das Ergebnis der Formapproximation durch Polyeder, durch welche mikroskopische Unregelmäßigkeiten der äußeren Gestalt unterdrückt werden. Solche Rauhigkeiten führen dazu, daß auch bei realen starren Objekten keine wirklichen Linienkontakte oder Flächenkontakte auftreten, sondern stattdessen eine beschränkte Anzahl von Punktkontakten (vgl. Abb. 11).

Wir wenden uns nun der Frage zu, welche Kontaktpunkte zur Repräsentation der betrachteten Kontaktregionen herangezogen werden sollen. Streng mathematisch spielt es keine Rolle, welche drei verschiedenen Punkte einer Kontaktfläche bzw. welche zwei verschiedenen Punkte eines linienhaften Kontaktsegments ausgewählt werden. Liegen die Kontaktpunkte jedoch sehr nahe beieinander, so führen kleine numerische Ungenauigkeiten bei der Bestimmung des Kontaktabstandes zu einem relativ großen Winkelfehler zwischen den sich berührenden Körpern (vgl. Abb. 23). In diesem Sinne ist beispielsweise die Konfiguration in Abb. 12b derjenigen in Abb. 12a vorzuziehen.

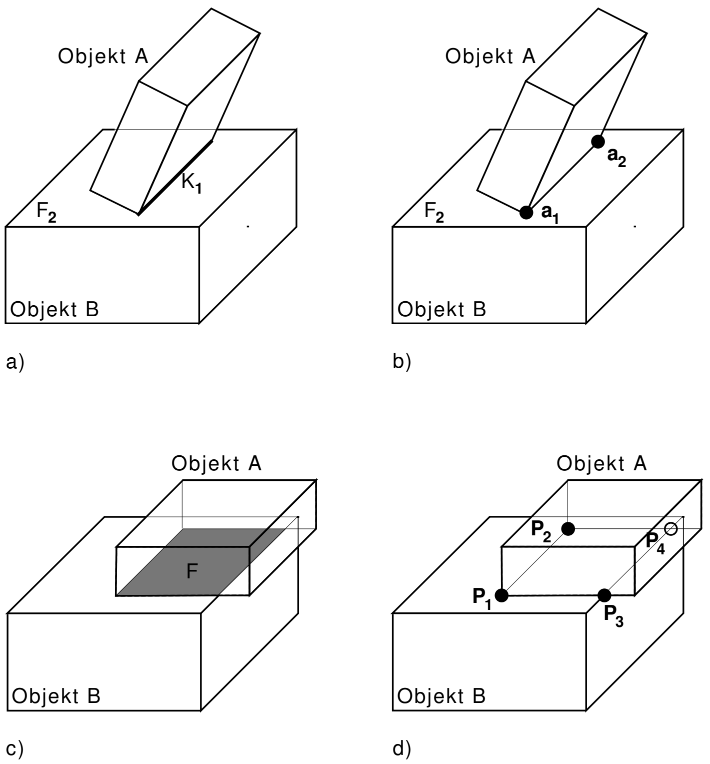


Abbildung 10: Ein Ecke-Fläche-Kontakt des Liniensegmentes K_1 (a) ist äquivalent zu zwei Punkt-Fläche-Kontakten bei $\mathbf{a}_1$ und $\mathbf{a}_2$ (b). Analog kann ein Fläche-Fläche-Kontakt F (c) durch drei Punkt-Fläche-Kontakte ersetzt werden (d).

Abbildung 11: Bei realen Objekten existieren aufgrund der Feinstruktur der Materialien keine perfekt ebenen Flächen und keine perfekt geraden Kanten. Unter der Annahme starrer Objekte ist die Realisierung von Flächenkontakten und Linienkontakten durch jeweils wenige Punktkontakte auch physikalisch plausibel.

Bei einem Linienkontakt bietet es sich daher unmittelbar an, die beiden Randpunkte der Kontaktregion als Kontaktpunkte zu wählen. In diesem Zusammenhang betrachten wir verschiedene Konfigurationen von Linienkontakten näher. In Abb. 13a sind die drei möglichen Fälle dargestellt, welche bei einem Kante-Fläche-Kontakt auftreten können. Hierbei sind die Kontaktsegmente jeweils fett markiert. Da die Endpunkte einer Objektkante gleichzeitig auch Eckpunkte des Polyeders sind, ist offensichtlich, daß die Rand-

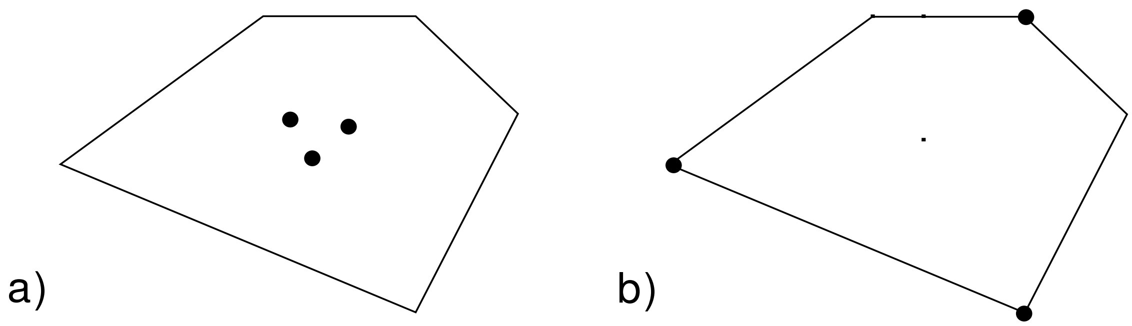


Abbildung 12: Ein Fläche-Fläche-Kontakt läßt sich durch drei nicht kollineare Punkt-Fläche-Kontakte nachbilden. Liegen diese sehr nahe beieinander (a), so können numerische Ungenauigkeiten die Stabilität der Kontaktsimulation beeinträchtigen. Besser wäre beispielsweise die Anordnung (b).

punkte der Kontaktsegmente stets mit Ecke-Fläche-Kontakten (a,b,c) oder Kante-Kante-Kontakten (d,e,f) zusammenfallen.

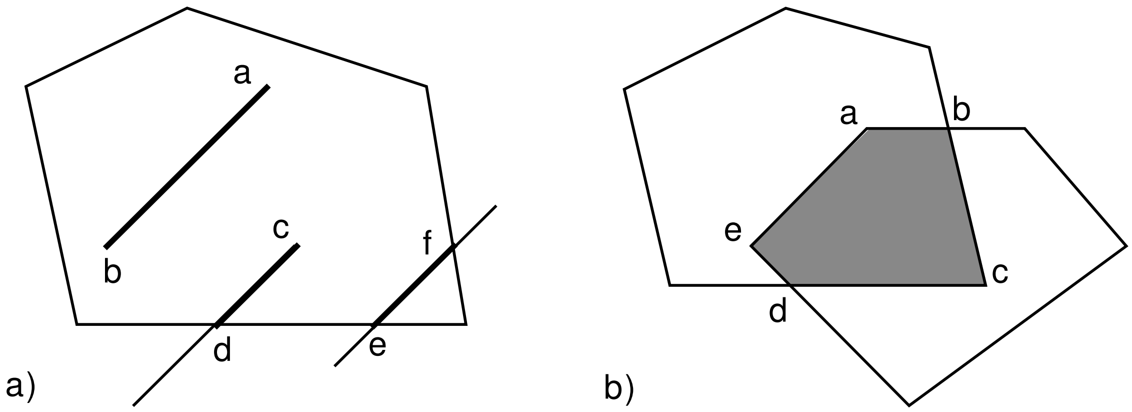


Abbildung 13: Links: Realisierung eines Kante-Fläche-Kontaktes durch Punkt-Kontakte (bei a,b,c als Punkt-Fläche-Kontakt, bei d,e,f als Kante-Kante-Kontakt). Die Kontaktregionen sind fett markiert. Rechts: Realisierung eines Fläche-Fläche-Kontaktes durch Punkt-Kontakte (bei a,c,e als Punkt-Fläche-Kontakt, bei b,d als Kante-Kante-Kontakt). Die Kontaktregion ist schattiert dargestellt.

Im Falle eines Flächenkontaktes gelten ähnliche Überlegungen. Zunächst läßt sich leicht zeigen, daß die Schnittmenge zweier konvexer Polygone ebenfalls ein konvexes Polygon darstellt. Dies gilt somit auch für die Kontaktregion eines Flächenkontaktes. Wie sich aus Abb. 13b ersehen läßt, fallen die Eckpunkte der schattiert dargestellten polygonalen Kontaktregion wiederum entweder mit Ecke-Fläche-Kontakten (a,c,e) oder Kante-Kante-Kontakten (b,d) zusammen.

Handelt es sich bei der Kontaktfläche um ein Dreieck, so stellen dessen Eckpunkte diejenigen Punkte mit dem größten gegenseitigen Abstand dar und sind somit optimal als Punktkontakte zur Substitution des Flächenkontaktes geeignet. Wird die Kontaktfläche von mehr als drei Kanten begrenzt, so sind die Eckpunkte dieses Polygons zumindest gute Kandidaten zur Auswahl der drei benötigten Punktkontakte.

Zusammenfassend läßt sich feststellen, daß sich sowohl Linienkontakte als auch Flächenkontakte durch Punktkontakte ersetzen lassen. Dabei sind die Eckpunkte der jeweiligen Kontaktregion als Kontaktpunkte besonders gut geeignet. Diese Punkte fallen jeweils genau mit den an der Kontaktregion beteiligten Ecke-Fläche-Kontakten und Kante-Kante-Kontakten zusammen.

In diesem Zusammenhang ist entscheidend, daß die verfügbaren Verfahren zur Feststellung von Kollisionen zwischen polyhedralen Körpern gerade auf die Erkennung von Ecke-Fläche-Kontakten und Kante-Kante-Kontakten zugeschnitten sind, und somit genau die von uns benötigten Kontaktelemente bereitstellen. Hierauf wird bei der Diskussion der Entstehung neuer Punktkontakte in Kapitel 5.2 noch näher eingegangen.

4.3 Substitution von degenerierten Kontakten durch Punktkontakte

Als degeneriert bezeichnen wir im folgenden die Kontakttypen Ecke-Ecke und Ecke-Kante. Diese Fälle, welche in Tabelle 1 durch Werte in Klammern markiert sind, unterscheiden sich von den restlichen (regulären) Kontakttypen durch folgende Merkmale:

- Für sie ist keine eindeutige Kontaktebene bzw. Kontaktnormale definiert
- Eine Relativbewegung am Kontaktpunkt kann entweder gar nicht (bei Ecke-Ecke) oder nur in einer einzigen Richtung (bei Ecke-Kante) erfolgen, ohne daß der Kontakt sich löst oder in einen anderen Kontakttyp (z.B. Ecke-Fläche oder Kante-Kante) übergeht. Bei den regulären Kontaktenarten stehen für die tangentiale Kontaktgeschwindigkeit beide Dimensionen der Kontaktebene zur Verfügung.

Insbesondere der erste Punkt ist für die Definition von Kontaktkräften, denen auch eine Richtung zugeordnet werden muß , problematisch. Eine von Bouma und Vaněček [7] vorgeschlagene Lösung geht davon aus, daß die degenerierten Kontaktarten aus regulären Kontaktarten heraus entstehen und übernehmen im Sinne zeitlicher Kohärenz die Kontaktnormale dieses zeitlich vorausgehenden Kontaktes auch für den degenerierten Kontakt. Dieser Vorschlag geht in die richtige Richtung. Es erscheint aber konsequenter, die Situationen Ecke-Ecke und Ecke-Kante gar nicht als eigenständige Kontaktarten zu repräsentieren, sondern grundsätzlich durch reguläre Punktkontakte zu ersetzen.

Dies läßt sich aufgrund der gewählten Topologie auch formal gut begründen. Es wurde bereits in Abschnitt 2.1 dargestellt, daß die Flächen und Kanten der betrachteten polyhedralen Objekthüllen als abgeschlossene Punktmengen definiert sind. Damit kann ein Ecke-Ecke-Kontakt gleichermaßen als eine der Kontaktarten Ecke-Fläche oder Kante-Kante interpretiert werden. Analog kann ein Ecke-Kante-Kontakt als Ecke-Fläche-Kontakt oder Kante-Kante-Kontakt aufgefaßt werden (siehe Abb. 14,15).

Welche der genannten Alternativen jeweils zum Zuge kommt, ergibt sich automatisch aus der jeweiligen Vorgeschichte des Kontaktes, und damit letztlich aus den näheren

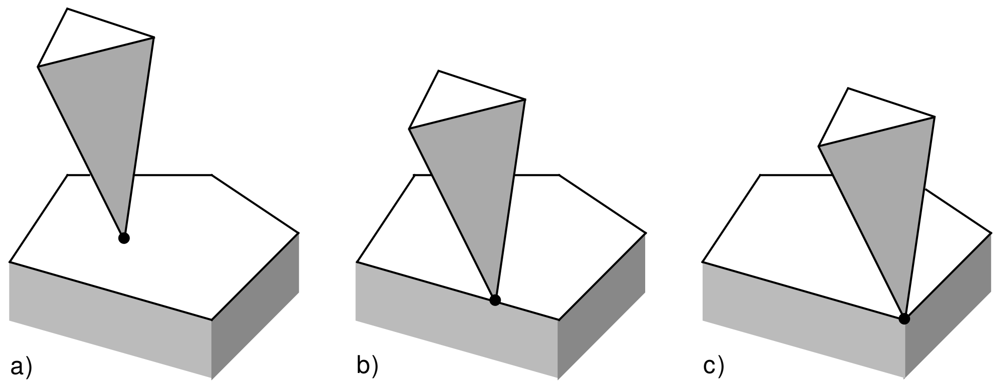


Abbildung 14: Ein Ecke-Fläche-Kontakt (a) und seine degenerierten Fälle Ecke-Kante (b) und Ecke-Ecke (c).

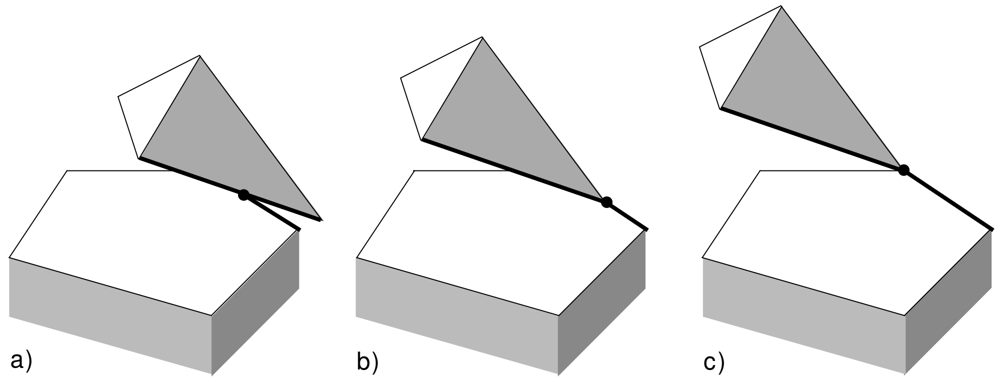


Abbildung 15: Ein Kante-Kante-Kontakt (a) und seine degenerierten Fälle Ecke-Kante (b) und Ecke-Ecke (c).

Umständen bei der Entstehung des Kontaktes, z.B. aufgrund einer Kollision.[6]

In Abb. 14 wandert ein Ecke-Fläche-Kontakt (14a) zu einer Kante (14b) oder Ecke (14c) der Kontaktfläche. In Abb. 15 wandert ein Kante-Kante-Kontakt (15a) zunächst zu einem Endpunkt einer der Kanten (15b), und schließlich auch noch zu einem Endpunkt der zweiten beteiligten Kante (15c). Die beiden Ecke-Kante-Konfigurationen (14b,15b) und die beiden Ecke-Ecke-Konfigurationen (14c,15c) sind aufgrund ihrer Vorgeschichte jeweils eindeutig als Ecke-Fläche-Kontakt beziehungsweise als Kante-Kante-Kontakt repräsentiert.

Die zur Repräsentation degenerierter Kontaktarten durch reguläre Punktkontakte benötigte Anzahl von Kontaktpunkten hängt davon ab, ob an der Kontaktsituation konvexe und/oder konkave Ecken und Kanten beteiligt sind. Stellvertretend sei dies am Beispiel des Kontaktes zwischen der Ecke E eines Körpers und der Kante K eines weiteren Körpers diskutiert. Ist die durch die Flächen F_1 und F_2 gebildete Kante K konvex, so wird die Situation durch einen einzigen regulären Punktkontakt beschrieben, wie in Abb. 16a dargestellt. Die Ecke E kann nicht mit F_2 kollidieren, solange der Kontakt mit F_1 bestehen bleibt.

[6]Bei Kollisionen werden, wie in Abschnitt 5.2 beschrieben, genau die in Betracht kommenden Punktkontaktarten Ecke-Fläche und Kante-Kante registriert.

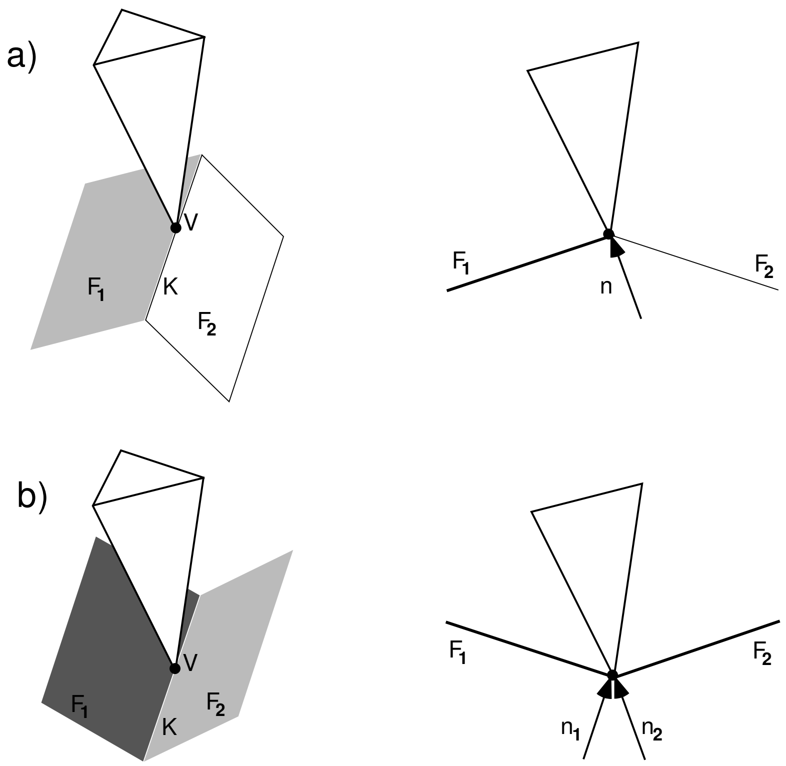


Abbildung 16: Ein Ecke-Kante-Kontakt wird im konvexen Fall (a) durch einen Punkt-Fläche-Kontakt realisiert, im konkaven Fall (b) durch zwei Punkt-Fläche-Kontakte.

Handelt es sich jedoch um eine konkave Kante (vgl. Abb. 16b), so kann K durchaus mit F_2 kollidieren, solange der Kontakt mit F_1 bestehen bleibt. Die Situation Ecke-Kante wird in diesem Fall also durch zwei reguläre Punktkontakte beschrieben. In der Situation konvexe Ecke - konkave Ecke sind entsprechend sogar drei Punktkontakte erforderlich.

Die Behandlung aller anderen Varianten degenerierter Kontaktsituationen geschieht vollkommen analog. In Abb. 17 sind alle Fälle zusammengestellt:

- Konvexe Ecke - konvexe Ecke (Abb. 17a)
- Konvexe Ecke - konkave Ecke (Abb. 17b)
- Konvexe Ecke - konvexe Kante (Abb. 17c)
- Konvexe Ecke - konkave Kante (Abb. 17d)
- Konvexe Kante - konvexkonkave Ecke (Abb. 17e)
- Konvexkonkave Ecke - konvexkonkave Ecke (Abb. 17f)

Tabelle 2 gibt einen Überblick über die Zuordnung zwischen degenerierten Kontaktsituationen und den dazu äquivalenten Konfigurationen regulärer Punktkontakte.

Es sollte betont werden, daß in all diesen Fällen keinerlei spezielle Mechanismen der Sonderbehandlung benötigt werden.

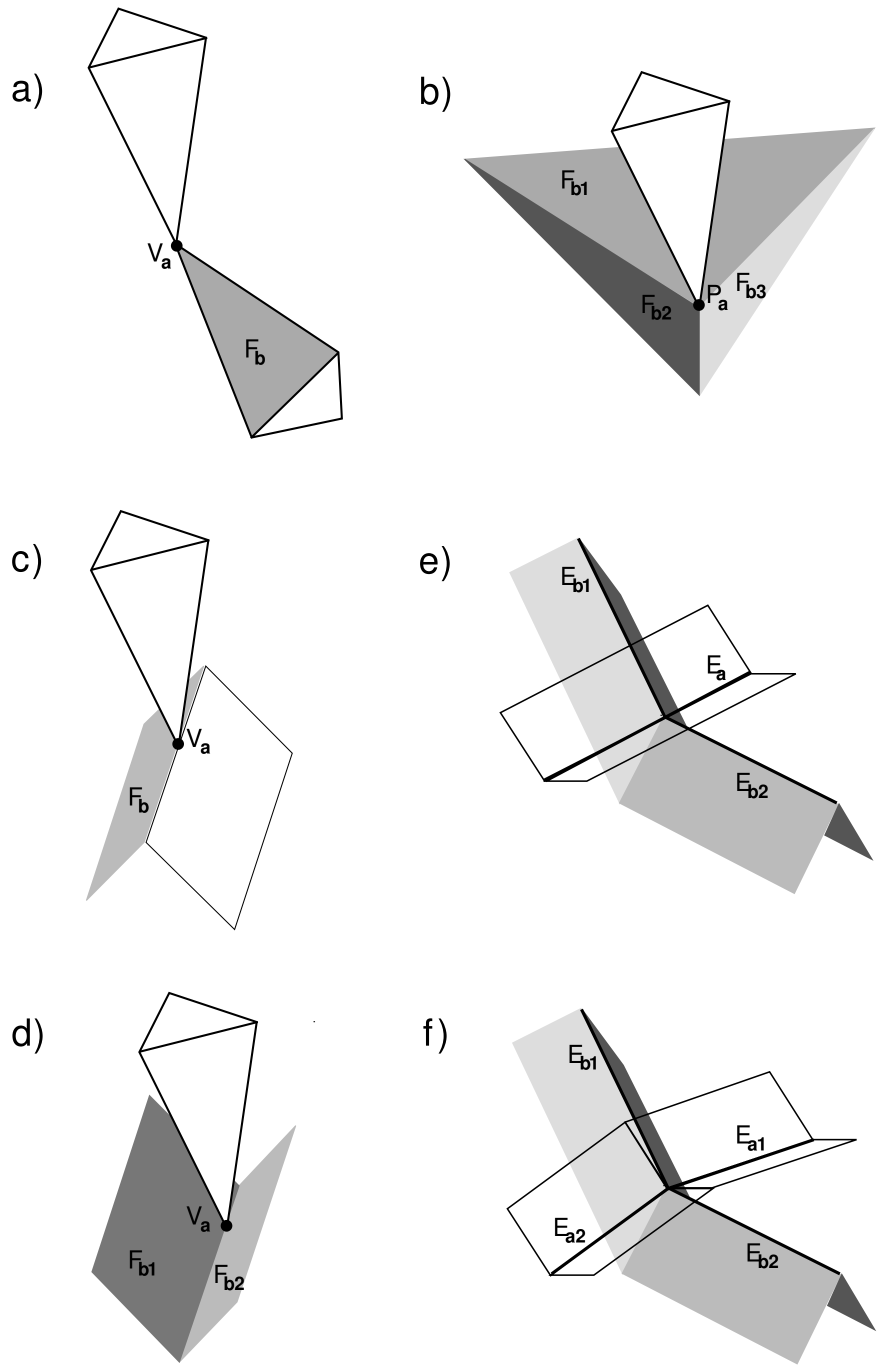


Abbildung 17: Degenerierte Kontaktfälle. Ein Ecke-Ecke-Kontakt kann ersetzt werden durch einen (a) oder drei (b) Ecke-Fläche-Kontakte, oder durch vier Kante-Kante-Kontakte (f). Ein Ecke-Kante-Kontakt kann ersetzt werden durch einen (c) oder zwei (d) Ecke-Fläche-Kontakte, oder durch zwei Kante-Kante-Kontakte (e). Die an den Kontakten beteiligten Ecken sind durch Punkte markiert, die beteiligten Kanten sind fett dargestellt, und die beteiligten Flächen schattiert.

Kontaktkonfiguration	Abb.	Ersatz durch äquivalente Punktkontakte
Ecke – Fläche	14a	EF
Kante – Kante	15a	KK
Fläche – Kante	10a,b	2 Punktkontakte (EF oder KK)
Fläche – Fläche	10c,d	3 Punktkontakte (EF oder KK)
konvexe Ecke – konvexe Ecke	17a	EF, KK
konvexe Ecke – konkave Ecke	17b	EF + EF + EF
konvexe Ecke – konvexe Kante	17c	EF, KK
konvexe Ecke – konkave Kante	17d	EF + EF
Sattelpunkt – konvexe Kante	17e	KK + KK
Sattelpunkt – Sattelpunkt	17f	KK + KK + KK

Tabelle 2: Repräsentation der verschiedenen Kontaktkonfigurationen durch Punktkontakte. Die Abkürzungen in der rechten Spalte bedeuten: EF = Ecke – Fläche, KK = Kante – Kante.

4.4 Mimimale Punktkontaktkonfigurationen

Wir betrachten zunächst den Zusammenhang zwischen gewissen Kontaktsituationen und dem damit verbundenen Verlust an Bewegungsfreiheitsgraden der beteiligten Objekte.

Ein unbeschränkt bewegliches starres Objekt hat bekanntlich 6 Freiheitsgrade. Besteht ein Punktkontakt mit einem feststehenden Hindernis, so reduziert sich diese Zahl auf 5. Bei einem Linienkontakt (oder äquivalent 2 Punktkontakten) verbleiben noch 4 Freiheitsgrade, und bei einem Flächenkontakt (oder äquivalent 3 nicht kollinearen Punktkontakten) verbleiben 3 Freiheitsgrade. Offensichtlich reduziert in diesen Beispielen jeder nicht-redundante Punktkontakt die Anzahl der verbleibenden Bewegungsfreiheitsgrade um 1. Als redundante Punktkontakte bezeichnen wir dabei solche, welche hinzugefügt oder entfernt werden können, ohne die Bewegungsmöglichkeiten der an dem Kontakt beteiligten Objekte zu verändern.

Dieser Zusammenhang zwischen der Anzahl nicht-redundanter Punktkontakte und der Anzahl der durch diese Kontakte entfernter Bewegungsfreiheitsgrade gilt allgemein. Er spiegelt sich auch in der mathematischen Behandlung der Kontaktsimulation wie folgt wider.

Die allgemeine Bewegung eines starren Körpers wird durch 6 unabhängige Bewegungsparameter, die zu einem Vektor $\mathbf{x} = (x_1, \ldots, x_6)^T$ zusammengefaßt seien, beschrieben. Jedem (nicht-redundanten) Punktkontakt des Körpers entspricht eine skalare Bedingungsgleichung $g(\mathbf{x}) = 0$ für diese Parameter, welche die Anzahl der unabhängigen Bewegungsparameter jeweils um 1 reduziert. Redundanten Punktkontakten entsprechen dabei linear abhängige Bedingungsgleichungen.

Bestehen für einen Körper 6 linear unabhängige Bedingungen, dann liegen alle Bewegungsparameter fest, d.h. die Raumlage des Objektes ist vollkommen bestimmt.

Aus diesen Zusammenhängen kann man entnehmen, daß es zur Beschreibung der Kinematik eines starren Körpers mit kontaktbedingten Bewegungsbeschränkungen ausreicht eine minimale, nicht-redundante Konfiguration von Punktkontakten zu formulieren. Über einen solchen minimalen (kanonischen) Satz hinausgehende Punktkontakte leisten keinen Beitrag zur Kontakt- und Bewegungssimulation. Sie blähen lediglich die zu behandelnden Gleichungssysteme auf, was einerseits zu Laufzeiterhöhungen führt, und andererseits deren eindeutige Lösbarkeit beeinträchtigen kann. Wir streben daher an, redundante Kontaktkonfigurationen zu vermeiden. Wie dies im Rahmen der Entstehung neuer Kontakte gewährleistet werden kann, wird in Abschnitt 5.5 diskutiert.

Wie bereits erwähnt, lassen sich linienhafte oder flächige Kontakte jeweils durch mehrere Punktkontakte substituieren. Im Sinne einer minimalen Punktkontaktkonfiguration ist es jedoch nicht ausreichend, jeden Linienkontakt durch 2 Punktkontakte und jeden Flächenkontakt durch 3 Kontaktpunkte zu ersetzen. Dies sei an folgendem Beispiel erläutert (Abb. 18).

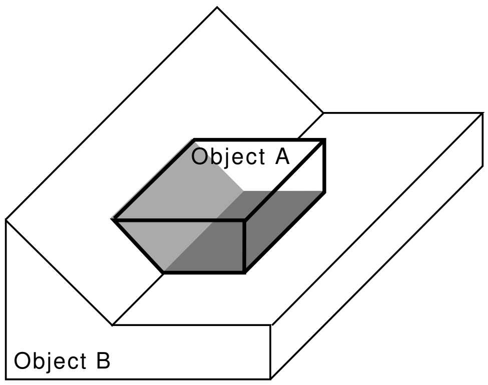


Abbildung 18: Der Keil A hat zwei Flächenkontakte mit Objekt B. Jeder der Flächenkontakte kann bis zu 3 Freiheitsgrade entfernen. Beide zusammen entfernen in diesem Fall jedoch nur 5 Freiheitsgrade.

Die beiden Objekte A und B berühren sich an zwei Flächenkontakten. Jeder dieser Flächenkontakte würde, isoliert betrachtet, dem System der beiden Körper drei Bewegungsfreiheitsgrade entziehen. Gemeinsam reduzieren beide Flächenkontakte die Anzahl der Freiheitsgrade jedoch nur um 5, denn es verbleibt ja noch eine eindimensionale Bewegungsmöglichkeit entlang der gemeinsamen Kante. Entsprechend sind auch nur 5 Punktkontakte erforderlich, um diese Konfiguration zu repräsentieren.

Das in Abschnitt 5.5 beschriebene Verfahren bestimmt automatisch jeweils nur genau die zur Erzwingung einer Bewegungsbeschränkung minimal erforderliche Anzahl von Punktkontakten.

4.5 Formulierung der Kontaktbedingungen für Punktkontakte

Gegeben seien zwei polyedrische Körper, die sich an einem oder an mehreren Kontaktpunkten berühren. Die Raumlage jedes der Körper wird durch die jeweilige Transformation zwischen seinem lokalen Objektkoordinatensystem und dem beiden Objekten gemeinsamen Weltkoordinatensystem bestimmt. Diese Transformationen definieren somit auch die Raumlage der an den Punktkontakten unmittelbar beteiligten Flächen, Kanten und Eckpunkten.

Die Berührbedingungen an jedem Kontaktpunkt lassen sich entweder als Schnittbedingung eines Punktes mit einer Ebene (Ecke-Fläche-Kontakt) oder aber als Schnittbedingung zweier Geraden (Kante-Kante-Kontakt) in Weltkoordinaten formulieren. Bei gegebenen Objektgeometrien der sich berührenden starren Körper lassen sich somit aus diesen Berührbedingungen jeweils Randbedingungen für die Raumlagen dieser Körper ableiten. Dies wollen wir nun für die beiden genannten elementaren Punktkontaktarten durchführen.

4.5.1 Kontakt zwischen Ecke und Fläche

Ein Ecke-Fläche-Kontakt werde gebildet durch den Eckpunkt $\hat{\mathbf{a}}$ eines Objektes $\mathcal{O}_1$ und eine Fläche des Objektes $\mathcal{O}_2$, welche in der Ebene $\hat{\mathbf{n}}^T\mathbf{x} = \hat{n}_0$ liegt. Die in Objektkoordinaten gegebenen Größen $\hat{\mathbf{a}}, \hat{\mathbf{n}}, \hat{n}_0$ werden in das Weltkoordinatensystem transformiert durch die Beziehungen

$$\mathbf{a} = \mathbf{R}_1\hat{\mathbf{a}} + \mathbf{c}_1, \qquad \mathbf{n} = \mathbf{R}_2\hat{\mathbf{n}}, \qquad n_0 = \hat{n}_0 + \mathbf{c}_2^T\mathbf{n},$$

wobei $\mathbf{R}_i \in \mathbb{R}^{3\times3}$ und $\mathbf{c}_i \in \mathbb{R}^3$ die jeweiligen Rotationsmatritzen und Verschiebungsvektoren der beiden Transformationen bezeichnen.

Die Berührnormale eines Ecke-Fläche-Kontaktes wird durch die Flächennormale $\mathbf{n}$ bestimmt, und der Kontaktabstand w ergibt sich zu

$$w = \hat{\mathbf{n}}^T\mathbf{R}_2^T(\mathbf{R}_1\hat{\mathbf{a}} + \mathbf{c}_1 - \mathbf{c}_2) - \hat{n}_0. \tag{28}$$

Die Kontaktbedingung lautet damit

$$w = 0. \tag{29}$$

4.5.2 Kontakt zwischen zwei Kanten

Eine zu Objekt $\mathcal{O}_1$ gehörende Kante mit den Endpunkten $\hat{\mathbf{a}}_1$ und $\hat{\mathbf{b}}_1$ stehe in Kontakt mit einer zu Objekt $\mathcal{O}_2$ gehörenden Kante mit den Endpunkten $\hat{\mathbf{a}}_2$ und $\hat{\mathbf{b}}_2$. Diese in Objektkoordinaten gegebenen Endpunkte werden in das Weltkoordinatensystem transformiert durch die Beziehungen

$$\mathbf{a}_i = \mathbf{R}_i\hat{\mathbf{a}}_i + \mathbf{c}_i, \qquad \mathbf{b}_i = \mathbf{R}_i\hat{\mathbf{b}}_i + \mathbf{c}_i, \tag{30}$$

wobei die jeweiligen Objekttransformationen wiederum durch $\mathbf{R}_i, \mathbf{c}_i$ beschrieben seien.

Die Berührnormale $\mathbf{n}$ eines Kante-Kante-Kontaktes wird durch die beiden Kantenorientierungen bestimmt:

$$\mathbf{n} = \mathbf{R}_1(\hat{\mathbf{b}}_1 - \hat{\mathbf{a}}_1) \times \mathbf{R}_2(\hat{\mathbf{b}}_2 - \hat{\mathbf{a}}_2). \tag{31}$$

Denkt man sich die beiden Kanten eingebettet in zwei parallele, zu $\mathbf{n}$ orthogonale Ebenen, so entspricht der Abstand dieser Ebenen gerade dem Kontaktabstand. Dieser läßt sich bestimmen durch

$$w \;=\; \frac{\mathbf{n}^T(\mathbf{a}_1 - \mathbf{a}_2)}{|\mathbf{n}|}. \tag{32}$$

Aus der Berührbedingung $w = 0$ folgt aus Gl.(32) zusammen mit (30) und (31) die entsprechende Randbedingung für die Objekttransformationen.

Es soll noch bemerkt werden, daß diese Bestimmung des Kontaktabstandes bezüglich des Vorzeichens von w nicht eindeutig ist, da die Orientierung der Berührnormalen gemäß Gl.(31) ebenfalls nicht eindeutig ist. In Verbindung mit der Bedingung $w = 0$ ist diese Mehrdeutigkeit zunächst bedeutungslos. Im nächsten Abschnitt wird die Auflösung dieser Mehrdeutigkeit beschrieben, da dort das Vorzeichen des Kontaktabstandes eine wesentliche Rolle spielen wird.

4.5.3 Verallgemeinerte Formulierung

Der Vollständigkeit halber geben wir noch die Funktion des Kontaktabstandes für allgemeine Kontaktgeometrien an. Dazu betrachten wir zwei Objekte $\mathcal{O}_1$ und $\mathcal{O}_2$, welche sich in den Punkten

$$\mathbf{a}_1 = \mathbf{R}_1\hat{\mathbf{a}}_1 + \mathbf{c}_1, \qquad \mathbf{a}_2 = \mathbf{R}_2\hat{\mathbf{a}}_2 + \mathbf{c}_2$$

berühren. Die Berührnormale sei $\mathbf{n}$, $|\mathbf{n}| = 1$. Damit bestimmt sich der Kontaktabstand zu

$$w = \mathbf{n}^T(\mathbf{R}_1\hat{\mathbf{a}}_1 + \mathbf{c}_1 - \mathbf{R}_2\hat{\mathbf{a}}_2 - \mathbf{c}_2), \tag{33}$$

wobei zu beachten ist, daß die Berührnormale $\mathbf{n}$ im Allgemeinen eine Funktion der Objektbewegungen ist. Dies läßt sich anhand der Herleitung der Beziehung (28) aus (33) darstellen. Die am Punkt-Fläche-Kontakt beteiligte Fläche gehöre dabei zu Objekt $\mathcal{O}_2$. In diesem Fall ist die Berührnormale gegeben durch $\mathbf{n} = \mathbf{R}_2\hat{\mathbf{n}}$, und somit $\mathbf{n}^T = \hat{\mathbf{n}}^T\mathbf{R}_2^T$. Damit erhalten wir

$$\begin{aligned} w &= \hat{\mathbf{n}}^T\mathbf{R}_2^T(\mathbf{R}_1\hat{\mathbf{a}}_1 + \mathbf{c}_1 - \mathbf{R}_2\hat{\mathbf{a}}_2 - \mathbf{c}_2) \\ &= \hat{\mathbf{n}}^T\mathbf{R}_2^T(\mathbf{R}_1\hat{\mathbf{a}}_1 + \mathbf{c}_1 - \mathbf{c}_2) - \hat{\mathbf{n}}^T\mathbf{R}_2^T\mathbf{R}_2\hat{\mathbf{a}}_2 \\ &= \hat{\mathbf{n}}^T\mathbf{R}_2^T(\mathbf{R}_1\hat{\mathbf{a}}_1 + \mathbf{c}_1 - \mathbf{c}_2) - \hat{\mathbf{n}}^T\hat{\mathbf{a}}_2 \\ &= \hat{\mathbf{n}}^T\mathbf{R}_2^T(\mathbf{R}_1\hat{\mathbf{a}}_1 + \mathbf{c}_1 - \mathbf{c}_2) - \hat{n}_0. \end{aligned}$$

Der letzte Schritt ergibt sich daraus, daß der Punkt $\hat{\mathbf{a}}_2$ in der Kontaktfläche liegt, und somit die Ebenengleichung $\hat{\mathbf{n}}^T\mathbf{x} = \hat{n}_0$ erfüllen muß.

4.6 Unilaterale Kontaktbedingungen

Die Einschränkung der Bewegungsmöglichkeiten eines starren Körpers durch Hinderniskontakte beruht physikalisch auf den an der Kontaktregion wirkenden Kontaktkräften. Abgesehen von Adhäsionseffekten können diese Kontaktkräfte nur abstoßend wirken, niemals anziehend. Entsprechend müssen die Kontaktbedingungen, die wir eben als skalare Bedingungsgleichungen für die Objekttransformationen hergeleitet haben, eigentlich als Ungleichungen formuliert werden.

Zur Bestimmung dieser Ungleichungen ist es zunächst erforderlich, eine Konvention bezüglich des Vorzeichens des Abstandes w zwischen zwei Objekten zu treffen. Wir legen daher fest, daß ein negativer Kontaktabstand die Durchdringung zweier Körper kennzeichne. Für unilaterale Kontaktbedingungen muß daher gelten

$$w \geq 0.$$

Die im letzten Abschnitt dargestellten Kontaktbedingungen legen den Kontaktabstand zunächst nur betragsmäßig fest. Für unilaterale Kontaktbedingungen muß der Kontaktabstand jedoch auch vorzeichenrichtig bestimmt werden. Hierzu stellen wir nachfolgend zwei Möglichkeiten dar.

Herleitung der Normalenrichtung durch Auswertung topologischer Information

Liegt für die betrachteten Körper eine topologisch konsistente Geometriebeschreibung vor, dann bilden die zu einem Körper gehörigen Flächenelemente eine geschlossene Hülle, die das Objektinnere eindeutig vom Objektäußeren unterscheidet. In diesem Fall kann somit auch jedem Flächenelement eine eindeutig orientierte Normale zugeordnet werden, die üblicherweise zum Objektäußeren hin gerichtet ist. Da die Objekthülle in sich geschlossen ist, kann zudem jeder Objektkante ein Paar angrenzender Flächenelemente zugeordnet werden. Die Auswertung der Flächennormalen erlaubt nun die eindeutige Bestimmung der Berührnormalen an den auftretenden Kontaktpunkten, und damit die vorzeichenrichtige Bestimmung des Kontaktabstandes.

Da im Falle eines Ecke-Fläche-Kontaktes die Berührnormale eindeutig durch die Orientierung der zugehörigen Flächennormale festgelegt ist, kann Gleichung (29) mit obiger Vorzeichenkonvention unmittelbar als Ungleichung umformuliert werden:

$$\hat{\mathbf{n}}^T \mathbf{R}_2^T (\mathbf{R}_1 \hat{\mathbf{a}} + \mathbf{c}_1 - \mathbf{c}_2) - \hat{n}_0 \quad \geq \quad 0. \tag{34}$$

Im Falle des Kante-Kante-Kontaktes liegen die Dinge komplizierter, da zwei Kanten keine Normalenorientierung unmittelbar zugeordnet ist. Die Bestimmungsgleichung (31) der Kontaktnormale legt die Richtung der Normale nur bis auf das Vorzeichen fest. Zur Herleitung der unilateralen Kontaktbedingung zwischen zwei Kanten muß daher weitere Information über die Kontaktgeometrie einfließen.

Seien e_1 und e_2 zwei Kanten mit den Endpunkten $\mathbf{a}_1$, $\mathbf{b}_1$ und $\mathbf{a}_2$, $\mathbf{b}_2$.

Zur Bestimmung des gerichteten Abstandes zwischen e_1 und e_2 sollen nun neben diesen Kanten selbst auch die jeweils angrenzenden Flächensegmente f_i und f_i' (vgl. Abb. 19) berücksichtigt werden. Hierzu führen wir die Normalenvektoren $\mathbf{n}_{f_i}$ und $\mathbf{n}'_{f_i}$ dieser Flächen ein, sowie die Vektoren

$$\begin{aligned} \mathbf{n}_{e_i} &= \mathbf{n}_{f_i} \times (\mathbf{b}_i - \mathbf{a}_i) \\ \mathbf{n}'_{e_i} &= (\mathbf{b}_i - \mathbf{a}_i) \times \mathbf{n}'_{f_i}. \end{aligned}$$

Ein Kontakt zwischen e_1 und e_2 kann nur bestehen, wenn beide Kanten konvex sind, d.h. wenn

$$\det[\mathbf{b}_i - \mathbf{a}_i, \mathbf{n}_{f_i}, \mathbf{n}'_{f_i}] > 0.$$

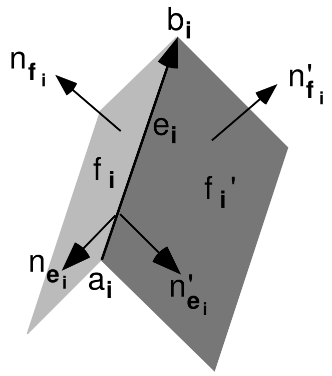


Abbildung 19: Notation für eine Kante e_i und die angrenzenden Flächen f_i, f_i'.

Die Berührebene H separiert die Keile, welche zu den Kanten e_1 und e_2 gehören (vgl. Abb. 20). Die Normale $\mathbf{n}$ der Ebene H bestimmt sich zu

$$\mathbf{n} = (\mathbf{b}_1 - \mathbf{a}_1) \times (\mathbf{b}_2 - \mathbf{a}_2).$$

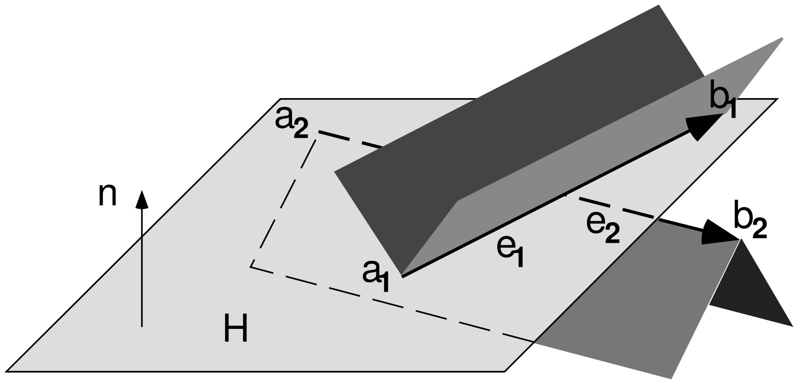


Abbildung 20: Die Berührebene H separiert die Keile, welche zu den Kanten e_1 und e_2 gehören.

Um entscheiden zu können, auf welcher Seite von H die Flächensegmente f_i und f_i' liegen, definieren wir die Orientierungsparameter s_i und s_i',

$$s_i = \mathbf{n}^T \mathbf{n}_{e_i} \qquad s_i' = \mathbf{n}^T \mathbf{n}_{e_i}'.$$

die positiv werden, wenn die Normale $\mathbf{n}$ in Richtung des jeweiligen Flächensegments f_i und f_i' zeigt.

Mit diesen Größen kann nun der gerichtete Abstand w zwischen den Kanten e_1 und e_2 bestimmt werden durch

$$w = \operatorname{sign}(s_1 + s_1') \frac{\mathbf{n}^T(\mathbf{a}_1 - \mathbf{a}_2)}{|\mathbf{n}|},$$

die zugehörige unilaterale Kontaktbedingung lautet entsprechend

$$\operatorname{sign}(s_1 + s_1') \frac{\mathbf{n}^T(\mathbf{a}_1 - \mathbf{a}_2)}{|\mathbf{n}|} \quad \geq \quad 0.$$

Herleitung der Normalenrichtung durch Auswertung der Kollisionsdaten

In praktischen Anwendungen stehen häufig keine topologisch konsistenten Geometriebeschreibungen mit geschlossenen Objekthüllen zur Verfügung. Zwischen benachbarten Flächenelementen treten oft spaltförmige Löcher in der Objekthülle auf, so daß die Hülle topologisch nicht geschlossen ist. Dünne Bleche werden oft nur einwandig modelliert, so daß eine Unterscheidung zwischen dem Objektinneren und dem Objektäußeren nicht möglich ist. Um auch in diesen Fällen zu vorzeichenrichtigen Bestimmungen des Kontaktabstandes zu kommen, kann also auf die Flächennormalen nicht zurückgegriffen werden. Die Auswertung der Daten, welche bei der Kollisionserkennung anfallen, erlaubt jedoch ebenfalls die eindeutige Orientierung der Berührnormalen.

Eine erkannte Kollision geht stets mit einem Vorzeichenwechsel des Kontaktabstandes einher. Da davon ausgegangen werden kann, daß die an der Kollision beteiligten Objekte vor der Kollision durchdringungsfrei waren, muß der vorzeichenrichtige Kontaktabstand zu Beginn des Zeitintervalles, in dem die Kollision erkannt wurde, positiv sein. Ergibt die Auswertung der Beziehungen (28) beziehungsweise (32) zu diesem Zeitpunkt einen negativen Kontaktabstand, so muß die zugrundeliegende Berührnormale gespiegelt werden. Auf diese Weise erhält man auch bei topologisch inkonsistenten Geometriebeschreibungen vorzeichenrichtige unilaterale Kontaktbedingungen.

5 Änderungen von Kontaktsituationen

Jede Kontaktsituation zwischen zwei polygonalen Körpern kann durch eine endliche Menge $\mathcal{K}$ von Punktkontakten beschrieben werden, wie in Kapitel 4 dargestellt wurde. Ändert sich durch Objektbewegungen die Kontaktsituation, so muß auch der zu ihrer Beschreibung erforderliche Satz von Punktkontakten entsprechend angepaßt werden. Solche Änderungen der Kontaktsituation erfolgen stets dann, wenn

- aufgrund einer Kollision ein neuer Kontaktpunkt entsteht.
- ein Kontaktpunkt sich löst. Da die betrachteten Objektgeometrien als Polyeder angenommen sind, gibt es hierbei zwei verschiedene, unterschiedlich zu handhabende Fälle, die Ablösung in Richtung der Berührnormalen und die tangentiale Ablösung parallel zur Berührebene.
- ein Kontaktübergang an Flächengrenzen oder Kantengrenzen erfolgt. Hierbei wird ein bestehender Punktkontakt durch einen neuen Punktkontakt ersetzt.

Diese Ereignisse sind beispielhaft in Abb. 21 schematisch dargestellt. In Abb. 21b und c entstehen die Kontakte k_1 und k_2 durch Kollisionen zwischen den Objekten A und B. In Abb. 21d erfolgt der Kontaktübergang von k_1 in k_3, in Abb. 21f ein weiterer Kontaktübergang von k_2 in k_4. Der Kontakt k_3 löst sich in Richtung der Berührnormale (vgl. Abb. 21g), wogegen sich der Kontakt k_4 parallel zur Berührebene löst (vgl. Abb. 21h).

Die in Abschnitt 4.4 aufgestellte Forderung nach minimalen Punktkontaktkonfigurationen muß natürlich insbesonders bei Änderungen der Kontaktsituation berücksichtigt werden. Bei allen Änderungen der Punktkontaktmenge $\mathcal{K}$ müssen redundante Kontaktpunkte vermieden werden. In Abschnitt 5.1 wird eine Vorgehensweise erläutert, wie dies sichergestellt werden kann.

Die weiteren Abschnitte dieses Kapitels diskutieren die Auffindung neuer Kontakte, die Ablösung von Kontakten, sowie Kontaktübergänge im Detail.

5.1 Bestimmung einer nicht-redundanten Menge von Punktkontakten

Bouma und Vaněček geben in [8] an, daß zur Analyse der Kontaktsituation zwischen zwei polygonalen Körpern eine volle mengentheoretische Schnittbildung zur Bestimmung der Kontaktregionen erforderlich sei. Aus der Analyse dieser Kontaktregionen lasse sich anschließend die zu ihrer Repräsentation ausreichende endliche Menge von Punktkontakten gewinnen. Gemäß [8] beträgt die Komplexität dieser Aufgabe $O(N^2)$, wobei N die Anzahl der Eckpunkte, Kanten und Flächen beider Objekte bezeichne. Der numerische Aufwand hierfür kann bei komplexen Objektgeometrien daher sehr groß werden.

Diese Betrachtung geht von einer statischen Sichtweise aus, und berücksichtigt nicht die Entstehungsgeschichte der jeweiligen Kontaktsituation. Zieht man diese mit in Betracht, so wird eine sehr viel effizientere Vorgehensweise möglich. Die Grundidee hierbei ist die folgende.

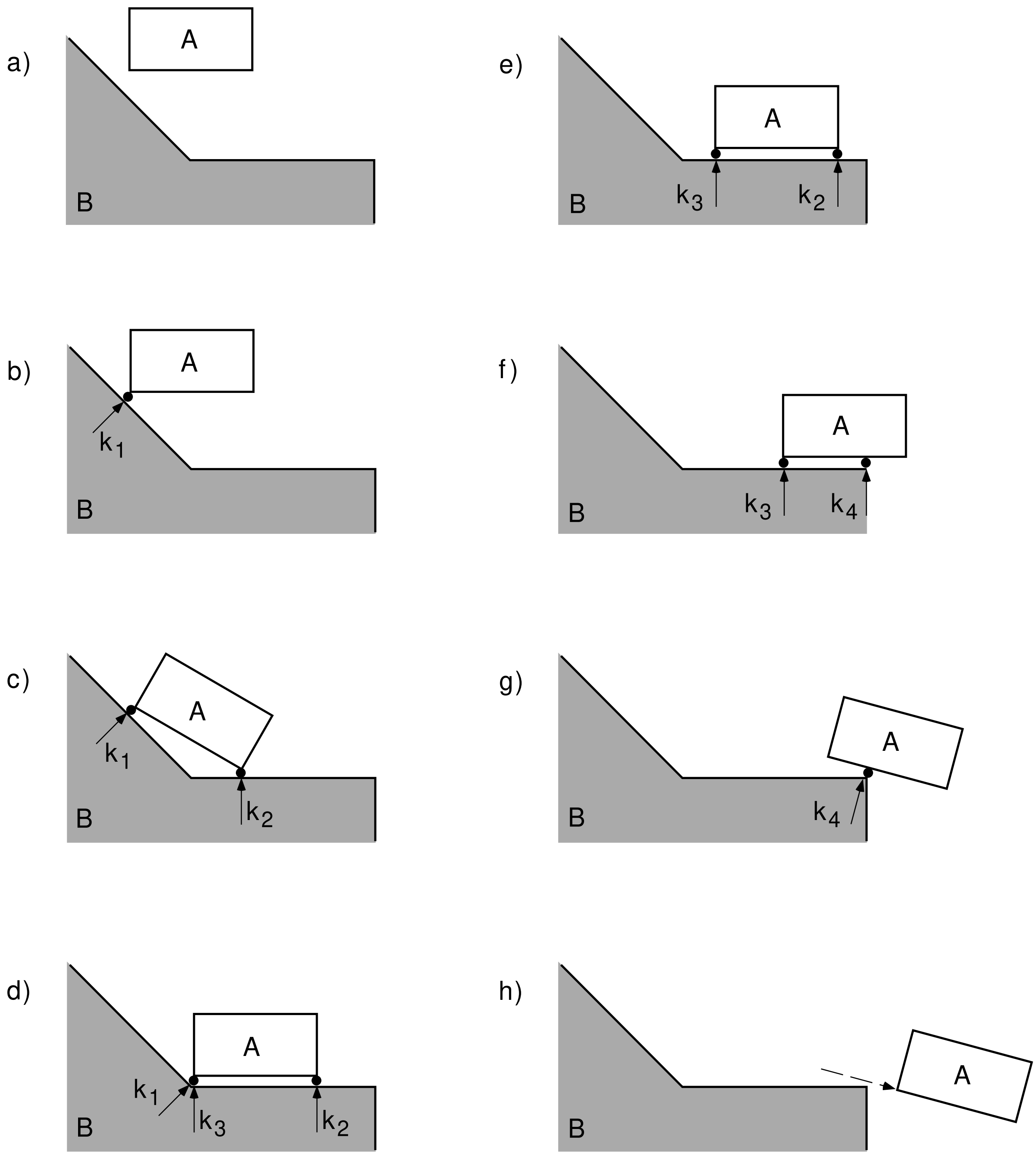


Abbildung 21: Verschiedene Phasen einer sich mehrfach verändernden Kontaktsituation zwischen den Objekten A und B. Die jeweils bestehenden Kontakte sind durch schwarze Punkte, die Kontaktkräfte sind durch Pfeile markiert.

Die früheste Berührung zwischen zwei kollidierenden polygonalen Körpern findet, von Sonderfällen abgesehen, an einem einzelnen Berührpunkt statt (der Fall, daß die früheste Berührung an mehreren Punkten exakt gleichzeitig geschieht, wird in Abschnitt 5.2.3 betrachtet). Um die Durchdringung zu verhindern, muß an diesem Kontaktpunkt eine ausreichend große Kontaktkraft angesetzt werden. Wir betrachten nun das Auftreten einer solchen Kollision selbst als Nachweis dafür, daß der aus ihr resultierende Kontaktpunkt notwendig, und somit nicht redundant ist. Umgekehrt formuliert wäre ein redundanter neuer Punktkontakt dadurch gekennzeichnet, daß er keinen zusätzlichen Einfluß auf die Bewegungsmöglichkeiten der betrachteten Objekte hat, da die erforderlichen Bewegungsbeschränkungen bereits von anderen Kontaktpunkten bewirkt werden. Aufgrund eben dieser Bewegungsbeschränkungen könnte an dem redundanten Kontaktpunkt zwar der Abstand zwischen den betrachteten Objekten gleich Null sein, es könnte jedoch keine Kollision mit Objektdurchdringung stattfinden.

Dieser Mechanismus funktioniert allerdings nur dann, wenn nicht mehrere neue Kontaktpunkte gleichzeitig hinzugenommen werden. Andernfalls wäre nur einer dieser neuen Kontaktpunkte sicher nicht-redundant, für alle übrigen Punkte könnte Redundanz nicht ausgeschlossen werden. Von mehreren exakt gleichzeitig auftretenden Berührpunkten kann jedoch schon aufgrund geringer numerischer Ungenauigkeiten bei der Bestimmung der Kollisionszeitpunkte kaum ausgegangen werden. Daher liegt es nahe, in diesem Falle einen der anscheinend gleichzeitig auftretenden Berührpunkte als neuen Kontaktpunkt auszuwählen. Dabei kann einer der Punkte entweder zufällig ausgewählt werden, oder gemäß einem geeignet erscheinenden Auswahlkriterium. In der vorliegenden Implementierung wird in solchen Fällen der als erstes registrierte Punkt ausgewählt. Waren weitere dieser Punkte ebenfalls nicht-redundant, so werden sie bei darauffolgenden Kollisionen in Erscheinung treten.

Dieser Zusammenhang soll anhand eines konkreten Beispiels veranschaulicht werden. Dazu betrachten wir einen Tisch mit vier exakt gleich langen Beinen, welcher bereits mit zwei Beinen Kontakt zur ebenen Grundfläche hat und nun mit den restlichen Beinen aufsetzt (Abb. 22a). Da im Rahmen der Kollisionserkennung nur höchstens ein neuer Punktkontakt registriert wird, steht der Tisch anschließend auf dreien der vier Beine, während das vierte Bein mit Abstand 0, aber ohne registrierten Kontakt über der Grundfläche 'schwebt' (Abb. 22b). Solange diese drei Beine ihren Kontakt beibehalten, kann der Tisch nicht kippen, und das vierte Bein kann nicht um das für eine Kollision erforderliche Maß in die Grundfläche eintauchen.

Ausgehend von einer redundanzfreien (eventuell auch leeren) Punktkontaktmenge können aufgrund dieser Überlegungen nur nicht-redundante Punktkontakte hinzukommen, so daß die Menge insgesamt redundanzfrei bleibt. Lösen sich einer oder mehrere der Punktkontakte wieder, so wird die reduzierte Punktkontaktmenge selbstverständlich ebenfalls redundanzfrei sein.

5.1.1 Schutzabstände zur Vermeidung redundanter Punktkontakte

Die Lösung der in Kapitel 3.4 aufgestellten Gleichungssysteme zur Bewegungssimulation sorgt für die Einhaltung der vorgegebenen Kontaktbedingungen und verhindert Objektdurchdringungen an den Kontaktpunkten. Die hierbei eingesetzten iterativen Verfahren liefern jedoch keine mathematisch exakten Lösungen, sondern nur sehr gute Näherungen mit einem kleinen Restfehler. Anschaulich bedeutet dies, daß der Kontaktabstand an jedem der Kontaktpunkte nicht stets exakt gleich Null ist, sondern im Laufe der Simulation geringfügig um Null schwankt. Daher würden auch an bereits etablierten Kontakten unter

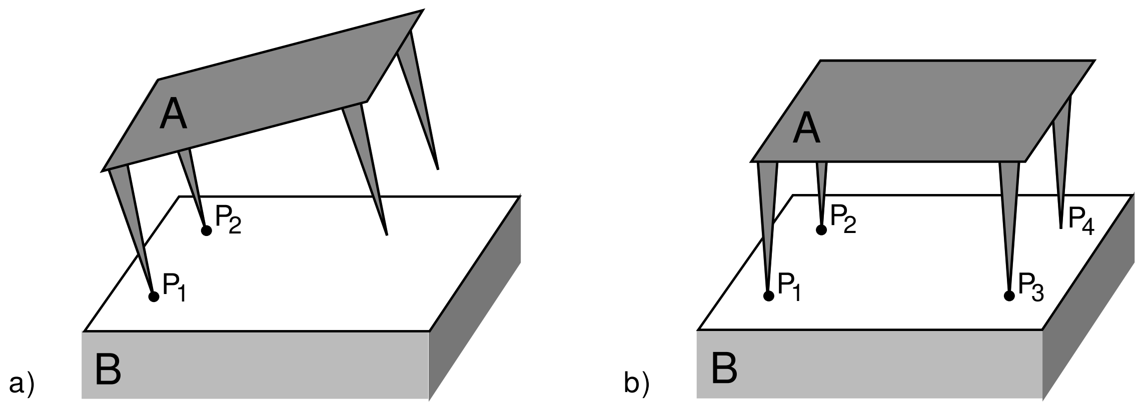


Abbildung 22: (a) Der Tisch A berührt die Unterlage B in den Kontaktpunkten bei P_1 und P_2. (b) Die Kollisionen bei P_3 und P_4 erfolgen zwar theoretisch zeitgleich, es wird jedoch zunächst nur ein neuer Kontaktpunkt bei P_3 zugelassen. Solange die drei Kontaktpunkte bei P_1, P_2 und P_3 bestehen, kann das vierte Tischbein nicht in die Unterlage eindringen.

Umständen häufig neue Kollisionen registriert. Ein Beispiel hierfür zeigt Abb. 23a. Ein Flächenkontakt zwischen den beiden Objekten A und B werde durch drei Punktkontakte an den zu Objekt A gehörenden Eckpunkten P_1, P_2 und P_3 repräsentiert. In einer gemeinsamen Ebene mit diesen Punkten liegen weitere Eckpunke der zu A gehörenden Fläche F_A. All diese Eckpunkte kollidieren aufgrund der erwähnten numerischen Ungenauigkeiten regelmäßig mit der zu B gehörenden Fläche F_B, obwohl dies gemäß den obigen Überlegungen zur Vermeidung redundanter Kontaktpunkte nicht der Fall sein dürfte (Abb. 23b).

Diese Problematik wird vermieden, wenn die Kontaktbedingungen an den einzelnen Kontaktpunkten so modifiziert werden, daß ein geringer Sicherheitsabstand ε_s an den Kontaktstellen eingehalten wird. Zwei eigentlich in Kontakt stehende Objekte berühren sich dann nicht mehr tatsächlich, sondern halten einen geringen Abstand zueinander ein (Abb. 23c). Die hierfür erforderlichen Kontaktbedingungen lauten $w = \varepsilon_s$ anstelle $w = 0$, wobei sich der Kontaktabstand w gemäß den Gleichungen (28) beziehungsweise (32) bestimmt.

Der Betrag von ε_s sollte um mindestens eine Größenordnung größer sein als die numerische Ungenauigkeit bei der Einhaltung des Kontaktabstandes, um redundante Kollisionen auch an abseits liegenden Objektpunkten zu vermeiden (vgl. P_4 in Abb. 23c).

5.2 Neue Kontakte aufgrund von Kollisionen

Nach jedem Simulationsschritt muß geprüft werden, ob zwischen zwei oder mehreren Objekten in der Szene Kollisionen aufgetreten sind, d.h. ob die Objekthüllen der betrachteten Objekte beginnen, sich zu durchdringen. In diesem Fall darf die Bewegung zunächst nur bis zu jenem Zeitpunkt der ursprünglich berechneten Bahn folgen, an dem die Durchdringung beginnen würde. Ab diesem Zeitpunkt muß der neue Kontakt und die hieraus resultierende zusätzliche Bewegungsbeschränkung für die weitere Bewegungssimulation berücksichtigt werden.

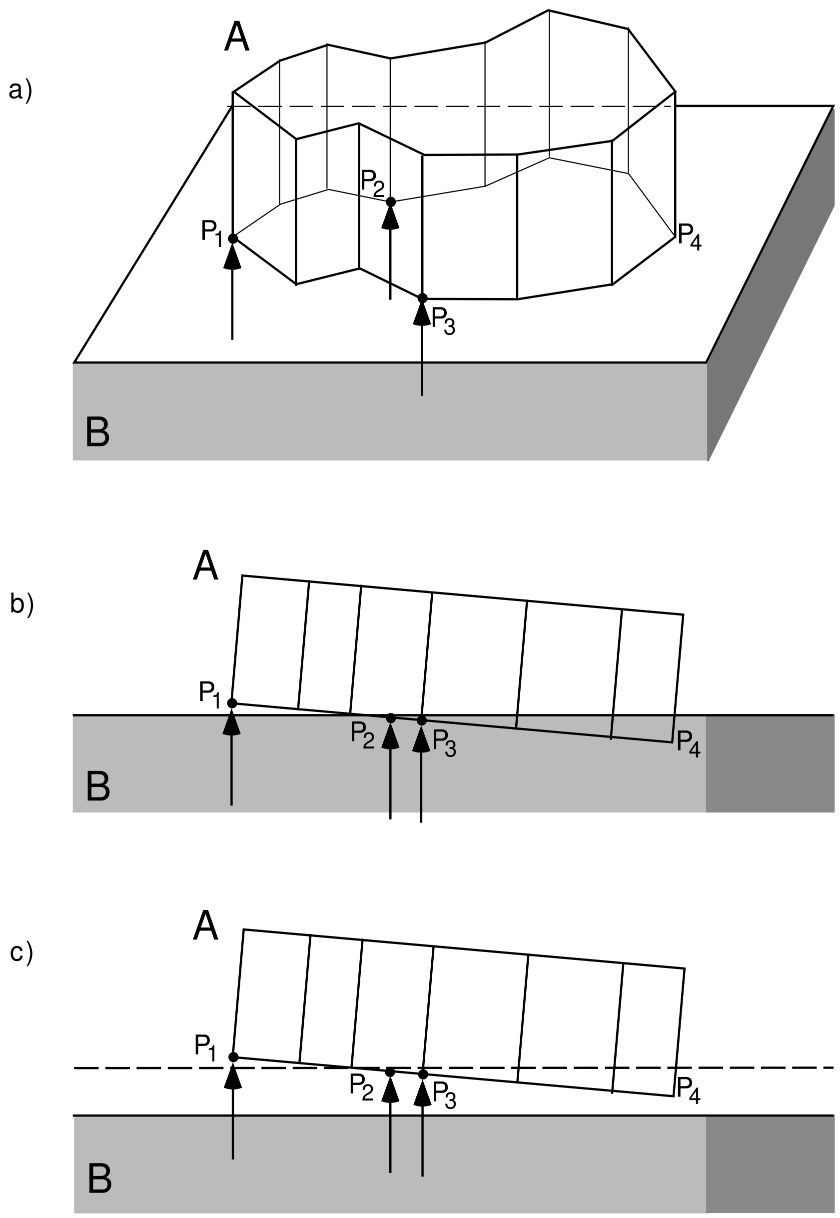


Abbildung 23: Schutzabstand zur Vermeidung redundanter Punktkontakte. (a) Der Flächenkontakt zwischen Prisma A und der Unterlage B wird durch die drei Punktkontakte P_1, P_2 und P_3 repräsentiert. (b) Durch numerische Ungenauigkeiten bei der Bestimmung des Kontaktabstandes kommt es ohne einen Schutzabstand zu Kollisionen an redundanten Kontaktpunkten. (c) Durch den Schutzabstand werden redundante Kontaktpunkte vermieden.

Als Kollisionen bezeichnen wir dabei nicht nur den Fall der ersten Berührung zwischen zwei Körpern. Auch wenn zwei Objekte, die sich bereits in Kontakt befinden, an einer weiteren Stelle aufeinanderprallen (wie beispielsweise k_2 in Abb. 21c), soll dies als Kollision behandelt werden.

Im Einzelnen sind nach dem Auftreten einer Kollision folgende Schritte notwendig:

1. Bestimmung des frühesten Kollisionszeitpunktes t_c. Die simulierte Bewegung der an der Kollision direkt oder indirekt beteiligten Objekte ist nur bis zu diesem Zeitpunkt t_c gültig. Ab diesem Zeitpunkt muß der neu entstandene Kontakt mit berücksichtigt werden.

2. Bestimmung der Raumlage zum Zeitpunkt t_c aller beteiligten Objekte, insbesondere aber diejenige der beiden an der Kollision unmittelbar beteiligten Objekte.

3. Analyse der Merkmale der neuen Kontaktsituation und Hinzunahme der neuen Kontaktbedingung in die Menge $\mathcal{K}$.

Bevor wir auf diese Punkte näher eingehen, zunächst einige Bemerkungen zur Kollisionserkennung selbst. Unter Kollisionserkennung wird meist die sogenannte statische Kollisionserkennung verstanden, welche lediglich prüft, ob die Schnittmenge zweier Objekte leer ist. Hierzu wird untersucht, ob sich die Hüllen der betrachteten Objekte in einer gegebenen Lage schneiden. Die demgegenüber erweiterte Fragestellung der dynamischen Kollisionserkennung lautet, in welcher Objektlage und mit welchen Teilen der Objektoberfläche der früheste Kontakt zwischen den kollidierenden Körpern auftrat. Da im Falle polyedrischer Körper die Oberflächengeometrien durch Flächen, Kanten und Ecken repräsentiert werden, sind Intersektionstests zwischen diesen Elementen erforderlich. Dabei ist es ausreichend, die Kontaktfälle Ecke-Fläche bzw. Fläche-Ecke und Kante-Kante zu berücksichtigen (vgl. [41]). Im Falle einer Kollision ist die neu entstandene Kontaktregion damit stets punktförmig, was für die weitere Kontaktsimulation sehr vorteilhaft ist. Die Problematik der Intersektionstests liegt weniger in ihrer Durchführung an sich, als in der zu bewältigenden numerischen Komplexität, die näherungsweise quadratisch mit der Komplexität der Einzelgeometrien (gemessen in der Anzahl N von Flächen, Kanten, und Ecken) wächst[7]. Um auch bei großen N noch effiziente Verfahren zu ermöglichen, werden häufig schnelle Vortests mit einfachen Hüllkörpern (beispielsweise Quadern oder Kugeln) durchgeführt.

Obgleich eine effiziente Kollisionserkennung für die Kontaktsimulation eine wichtige Voraussetzung darstellt, wollen wir Verfahren und Algorithmen hierfür im folgenden nicht weiter vertiefen, sondern ein ausreichend effizientes Verfahren zur Kollisionserkennung als verfügbar annehmen. Allerdings ist es notwendig, einige hierbei auftretende Probleme sowie gewisse Anforderungen an ein Kollisionserkennungsmodul zu erörtern, welche für die Behandlung des Themas Kontaktsimulation relevant sind.

5.2.1 Bemerkungen zur Bestimmung des Kollisionszeitpunktes

Wir betrachten eine Kollision zwischen zwei bewegten Polyedern in einem Zeitintervall $[t_1, t_2]$. Der Kollisionszeitpunkt $t_c \in [t_1, t_2]$ entspricht dem Zeitpunkt der frühesten Berührung zwischen einem Ecke-Fläche-Paar oder einem Kante-Kante-Paar der beiden betrachteten Körper. Zur Bestimmung von t_c ist es erforderlich, alle Ecke-Fläche-Kollisionen und alle Kante-Kante-Kollisionen innerhalb des betrachteten Zeitraumes zu bestimmen, und daraus die früheste dieser Kollisionen zu ermitteln.

[7] In Schömer [41] wird gezeigt, daß für die Fälle rein translatorischer Bewegung sowie rein rotatorischer Bewegung Verfahren zur Kollisionserkennung mit der Komplexität $N^{8/5}$ existieren.

Für jedes Ecke-Fläche-Paar und jedes Kante-Kante-Paar kann der zeitliche Verlauf des Abstandes in Form einer vorzeichenbehafteten Abstandsfunktion $d(t)$ beschrieben werden. Die Nullstellen dieser Abstandsfunktionen entsprechen dann den jeweiligen Kollisionszeitpunkten.

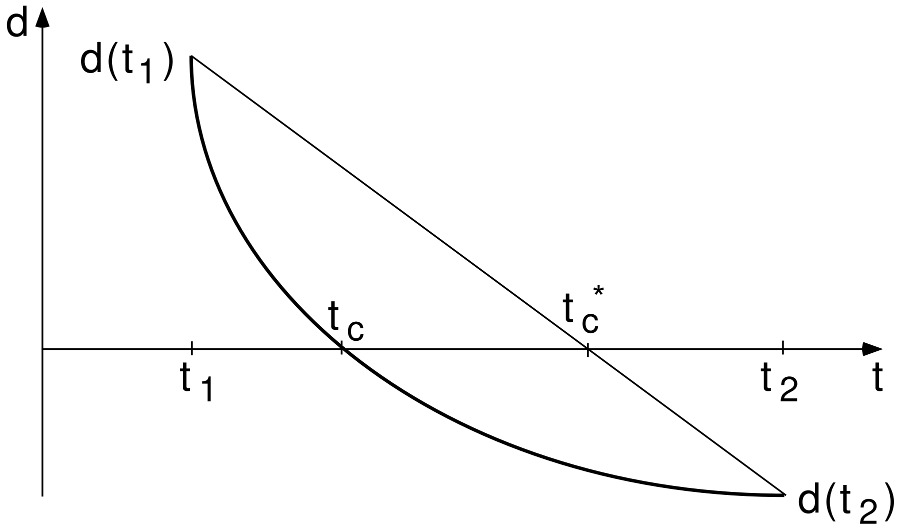


Abbildung 24: Wird der Kollisionszeitpunkt mit Hilfe einer linearen Approximation an den tatsächlichen Verlauf der Abstandsfunktion $d(t)$ ermittelt, so kann der so bestimmte Zeitpunkt t_c^* vom korrekten Wert t_c abweichen.

Der genaue Verlauf der Abstandsfunktion $d(t)$ ist dem Kollisionserkennungsmodul jedoch nicht unmittelbar bekannt. Im Rahmen der Bewegungssimulation werden die Objektlagen zu diskreten Zeitpunkten explizit bestimmt und dem Kollisionserkennungsmodul mitgeteilt. Die Abstandsfunktion $d(t)$ kann daher zunächst ebenfalls nur zu diesen Zeitpunkten ausgewertet werden. Damit kann vom Kollisionserkennungsmodul zwar ein Vorzeichenwechsel von $d(t)$ innerhalb des Zeitintervalles $[t_1, t_2]$ festgestellt werden, nicht jedoch der exakte Zeitpunkt des Nulldurchganges, der dem Kollisionszeitpunkt entspricht. Zur Behandlung dieser Problematik stellt das verwendete Kollisionserkennungsmodul verschiedene Möglichkeiten zur Interpolation beziehungsweise Approximation des Kollisionszeitpunktes zur Verfügung.

Die einfachste Vorgehensweise besteht darin, den unbekannten Verlauf von $d(t)$ linear zwischen den bekannten Stützpunkten $d(t_1)$ und $d(t_2)$ zu interpolieren. Damit erhält man eine gute Approximation an den tatsächlichen Kollisionszeitpunkt, wenn der Verlauf von $d(t)$ näherungsweise eine lineare Funktion darstellt. Dies ist jedoch nur dann der Fall, wenn die relative Rotationsbewegung der betrachteten Körper im jeweiligen Zeitintervall hinreichend klein ist.

Im Allgemeinen weicht die lineare Interpolation jedoch vom tatsächlichen Verlauf von $d(t)$ ab, und erlaubt somit nur eine approximative Bestimmung des Kollisionszeitpunktes t_c. Dieser Fall ist in Abb. 24 dargestellt. Der mit der linearen Interpolation bestimmte Zeitpunkt t_c^* weicht in diesem Beispiel vom tatsächlichen Wert t_c ab. Eine solche ungenaue Bestimmung des Kollisionszeitpunktes kann dazu führen, daß im Rahmen der

Kontaktsimulation Objektdurchdringungen auftreten. Dies kann insbesondere dann geschehen, wenn fälschlicherweise ein zu großer Wert für t_c bestimmt wird, so daß sich die kollidierenden Polyeder zu diesem Zeitpunkt bereits durchdringen (vgl. Abb. 26b).

Eine bessere Approximation des Kollisionszeitpunktes wird möglich, wenn von der Bewegungssimulation zusätzliche Daten über den Verlauf der Objektbewegungen zur Verfügung gestellt werden können, so daß sich zusätzliche Werte von $d(t)$ bestimmen lassen. Das verwendete Kollisionserkennungsmodul bietet hierfür mehrere Möglichkeiten:

- Quadratische Interpolation unter Verwendung einer weiteren Stützstelle zwischen t_1 und t_2.
- Äquidistante Abtastung unter Verwendung einer vorgebbaren Anzahl weiterer Stützstellen zwischen t_1 und t_2.
- Intervallschachtelung mit Regula Falsi (vgl. Abb. 25).

Die beiden letztgenannten Verfahren erlauben eine beliebig genaue Approximation des Kollisionszeitpunktes. Da die Intervallschachtelung mit Regula Falsi jedoch in aller Regel die Auswertung einer sehr viel kleineren Anzahl von Stützstellen erfordert als die äquidistante Abtastung, bis ein vorgegebener Approximationsfehler unterschritten wird, ist diesem Verfahren der Vorzug zu geben.

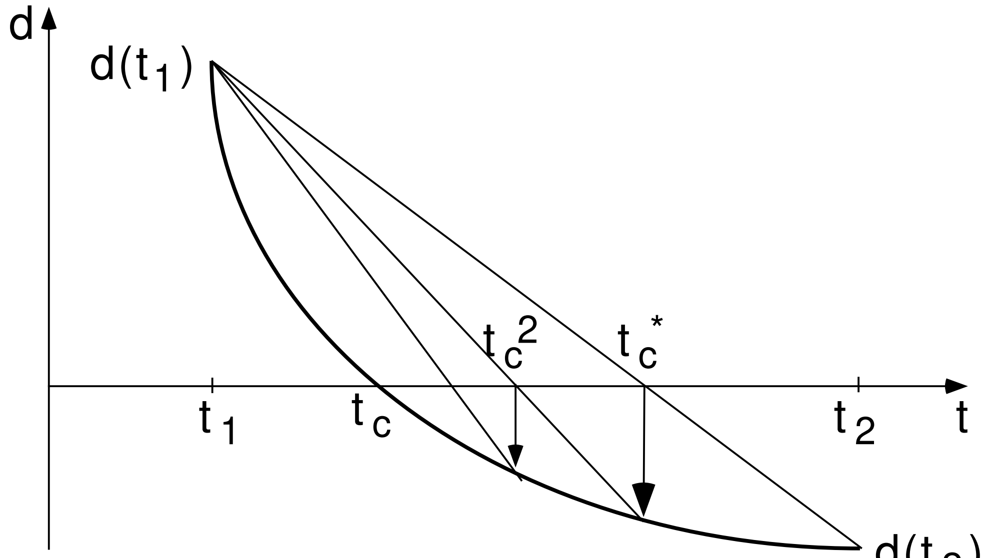


Abbildung 25: Iterative Approximation des Kollisionszeitpunktes t_c mit der Regula falsi. Die gekrümmte Linie entspricht dem tatsächlichen Verlauf der Abstandsfunktion $d(t)$ im Intervall $[t_1, t_2]$.

Auch bei diesen verbesserten Approximationen bleibt der ermittelte Kollisionszeitpunkt jedoch mit einem Restfehler behaftet, was zu Objektdurchdringungen führen könnte. Dieser Problematik kann durch die bereits beschriebene Einführung eines Schutzabstandes wie folgt begegnet werden. Zur Erkennung einer Kollision wird, wie bereits beschrieben, der Vorzeichenwechsel der Abstandsfunktion $d(t)$ herangezogen. Als Kollisionszeitpunkt t_s wird nun allerdings nicht der Zeitpunkt t_c dieses Nulldurchganges verwendet, sondern ein geringfügig früherer Zeitpunkt, $t_s = t_c - t_\varepsilon$. Dieser wird so bestimmt, daß die beteiligten Geometrien noch einen geringen Abstand, eben den erwähnten Schutzabstand, voneinander haben. Bleibt der Restfehler bei der Bestimmung des Kollisionszeitpunktes kleiner als t_ε, so kann es zu keiner Objektdurchdringung kommen, sondern lediglich zu einer Unterschreitung des Sicherheitsabstandes zwischen den Objekten.

Da die Bestimmung von t_ε nicht mit hoher Genauigkeit erfolgen muß , genügt es hierfür eine lineare Interpolation zwischen t_1 und t_c heranzuziehen. Es muß allerdings sichergestellt werden, daß $t_s \geq t_1$ bleibt. Somit erhält man folgende Beziehungen zur Bestimmung von t_s:

$$\begin{aligned} t_\varepsilon &= \frac{(t_c - t_1)\varepsilon_s}{|d(t_c) - d(t_1)|}, \\ t_s &= \max(t_1, t_c - t_\varepsilon). \end{aligned}$$

5.2.2 Die sichere Erkennung aller neuen Kontakte

Im Falle einer Kollision zwischen zwei Körpern ist es zur Vermeidung von Durchdringungen erforderlich, den frühesten Kontaktpunkt sicher zu erkennen. Im vorigen Abschnitt haben wir als eine wesentliche Voraussetzung hierfür die möglichst genaue Bestimmung der Kollisionszeitpunkte bei den einzelnen Basiskollisionstests (Ecke-Fläche-Tests und Kante-Kante-Tests) diskutiert. Dabei diente die Detektion eines Vorzeichenwechsels der jeweiligen Abstandsfunktion $d(t)$ zwischen t_1 und t_2 als Kriterium zur Erkennung einer Kollision, d.h. $d(t_1)d(t_2) < 0$. Dieses Kriterium ist allerdings nicht hinreichend zur sicheren Erkennung aller Kollisionen, da $d(t)$ zwei Nulldurchgänge besitzen könnte. In diesen Fällen ist $d(t_1)d(t_2) > 0$. Anschaulich bedeutet dies, daß zwei Körper zu Beginn und am Ende eines betrachteten Zeitintervalls durchdringungsfrei sind, sich jedoch dazwischen kurzzeitig durchdringen (vgl. Abb. 26c).

Eine Möglichkeit zur Erkennung solcher Kollisionen besteht darin, die Abstandsfunktion an weiteren Stellen im Inneren des Zeitintervalles auszuwerten. Das zur Verfügung stehende Kollisionserkennungsmodul bietet hierfür den Modus der äquidistanten Abtastung der Abstandsfunktion an. Durch eine ausreichend dichte Abtastung kann somit auch in diesem Fall eine Kollision sicher erkannt werden. Dies ist allerdings mit einem entsprechenden eventuell unnötig hohen numerischen Aufwand verbunden.

Um diesen Aufwand so gering wie möglich zu halten, benötigen wir ein Entscheidungskriterium, welches uns anzeigt, ob die Abstandsfunktion zwischen zwei Stellen gleichen Vorzeichens überhaupt eine Nullstelle haben kann. Ist dies der Fall, so teilen wir das Intervall in zwei Hälften. Hat die Abstandsfunktion in der Mitte des Intervalls dasselbe Vorzeichen wie an den Intervallgrenzen, so setzen wir das Verfahren rekursiv für beide Intervallhälften fort. Das Verfahren bricht ab, wenn für alle Teilintervalle aufgrund des noch zu besprechenden Entscheidungskriteriums ein Vorzeichenwechsel ausgeschlossen werden kann, oder aber wenn ein solcher Vorzeichenwechsel erkannt wird.

Zur Erläuterung des Entscheidungskriteriums betrachten wir Abb. 27. Das Maß h bezeichnet die maximale Abweichung zwischen dem Verlauf der tatsächlichen Abstands-

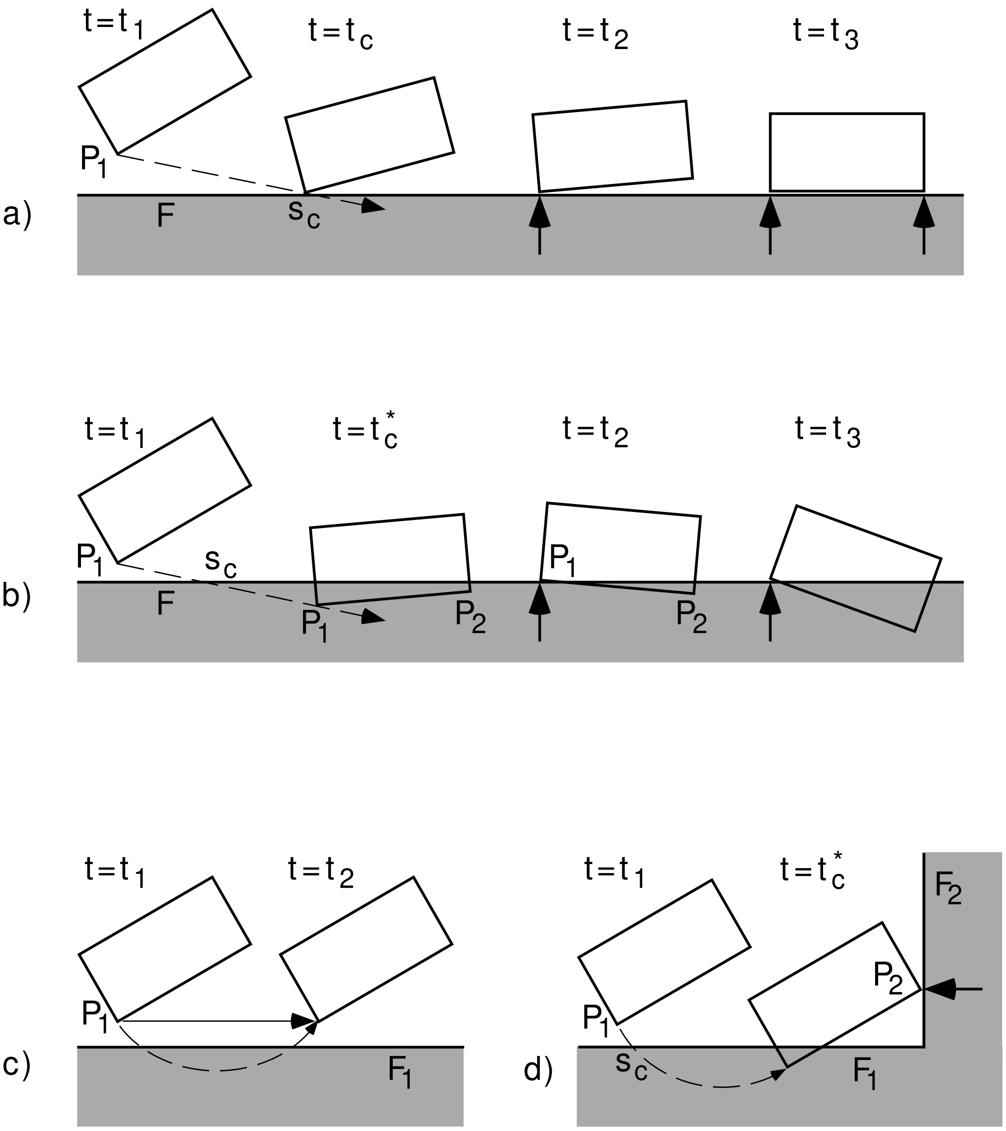


Abbildung 26: (a) Bei korrekter Bestimmung des Kollisionszeitpunktes t_c werden Objektdurchdringungen sicher vermieden. Die hierzu eingeführten Punktkontakte sind durch Pfeile markiert. (b) Wird der Kollisionszeitpunkt zu spät bestimmt, so können Durchdringungen auftreten, wie hier bei P_2 dargestellt. (c) Liegen bei $t = t_1$ und bei $t = t_2$ keine Durchdringungen vor, so könnte doch dazwischen eine Kollision stattfinden (gestrichelter Pfeil), die jedoch bei linearer Approximation der Trajektorie (durchgezogener Pfeil) unbemerkt bleibt. (d) Die Situation in (c) kann zu unerwünschten Durchdringungen führen, wenn während des Intervalles an anderer Stelle eine Kollision stattfindet.

funktion $d(t)$ und der linearen Interpolation $\bar{d}(t)$. Gegeben sei nun eine obere Schranke $\hat{h} \geq h$ dieser maximalen Abweichung. Damit kann eine Nullstelle von $d(t)$ im Intervall $[t_1, t_2]$ nur auftreten, wenn entweder $d(t_1) \leq \hat{h}$ oder $d(t_2) \leq \hat{h}$. Da wir den genauen Verlauf der Abstandsfunktion $d(t)$ nicht kennen, können wir den Wert von h nicht unmittelbar bestimmen, und müssen daher eine Abschätzung $\hat{h}$ heranziehen.

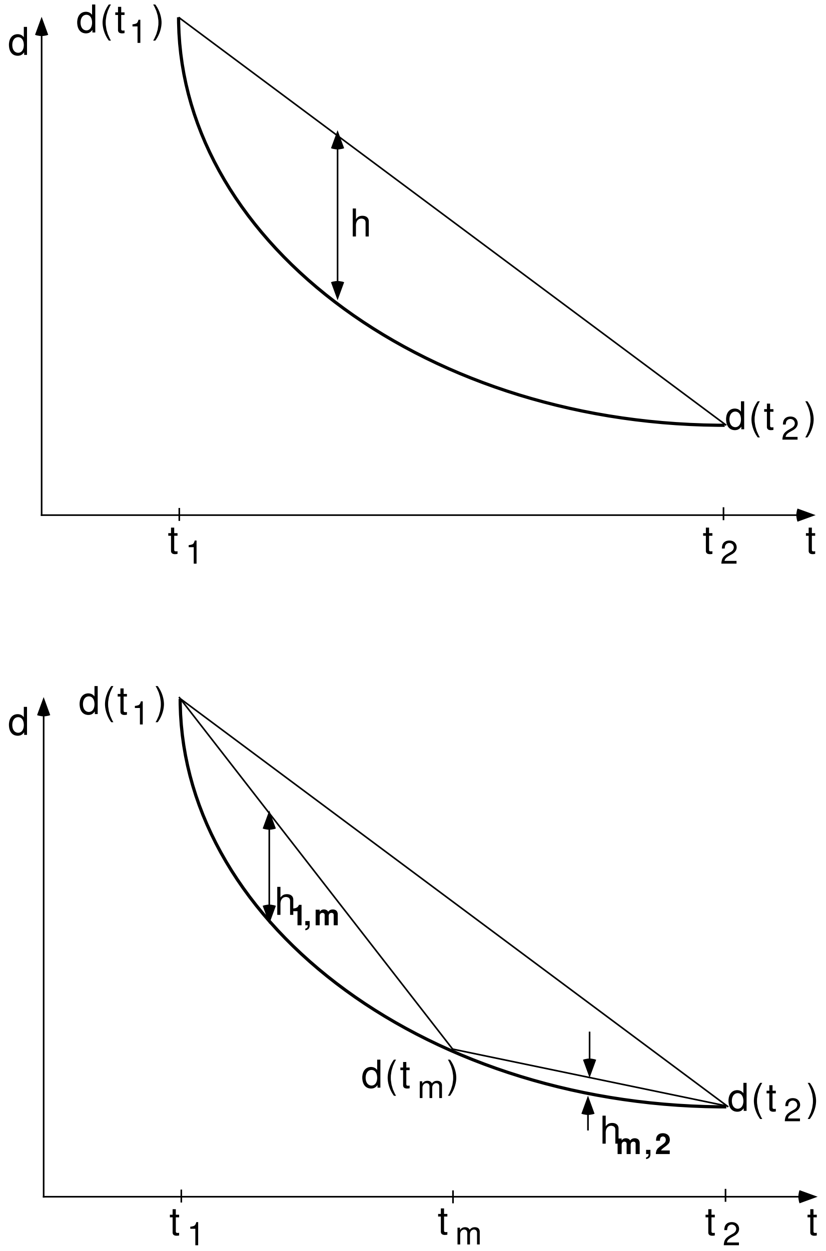


Abbildung 27: Die maximale Abweichung h zwischen der Abstandsfunktion $d(t)$ und ihrer linearen Approximation dient als Kriterium für eine mögliche Kollision im Intervall $[t_1, t_2]$, wenn $d(t_1)$ und $d(t_2)$ vorzeichengleich sind. Wegen $h > d(t_2)$ erfolgt in diesem Beispiel eine Intervallhalbierung. In den resultierenden Teilintervallen sind die entsprechenden Abweichungen klein genug ($h_{1,m} < d(t_m)$, $h_{m,2} < d(t_2)$), so daß eine Kollision ausgeschlossen werden kann.

Abschätzung der maximalen Abweichung h

Zur Abschätzung der maximalen Abweichung h bestimmen wir zunächst die lineare Approximation $\bar{d}(t)$ der tatsächlichen Abstandsfunktion $d(t)$ bezogen auf die Endpunkte des Zeitintervalls $[t_1, t_2]$:

$$\bar{d}(t) = d(t_1) + \frac{t - t_1}{t_2 - t_1} \left(d(t_2) - d(t_1)\right). \tag{35}$$

Gemäß den Bewegungsgleichungen (21) ... (25) läßt sich die Objektposition zu jedem Zeitpunkt $t \in [t_1, t_2]$ eindeutig bestimmen, und hieraus in Verbindung mit der Kontaktgeometrie (vgl. Gleichungen (28), (32), (33)) ein Kontaktabstand $d(t)$ berechnen. Wir nehmen hierzu vereinfachend an, daß die Kontaktkräfte für die Dauer des betrachteten kurzen Zeitintervalls als näherungsweise konstant angenommen werden können.

Zur Bestimmung von h betrachten wir zunächst die mögliche Abweichung $\Delta d(t) = d(t) - \bar{d}(t)$ am Fall eines Punkt-Fläche-Kontaktes.

Zur Auswertung der Kontaktabstandsfunktion (vgl. Gleichung (28))

$$d(t) = \hat{\mathbf{n}}^T \mathbf{R}_2^T(t) \left(\mathbf{R}_1(t)\hat{\mathbf{a}} + \mathbf{c}_1(t) - \mathbf{c}_2(t)\right) - \hat{n}_0$$

leiten wir aus (25) und (80) sowie mit $\Delta t = t - t_1$ die Beziehung

$$\begin{aligned} \mathbf{R}(t) &= \mathbf{R}(1, \frac{1}{2}\Delta t\, \boldsymbol{\omega}) \cdot \mathbf{R}(t_1) \\ &= \frac{\left(1 - \frac{1}{4}\Delta t^2 \boldsymbol{\omega}^2\right) \mathbf{E} + \frac{1}{2}\Delta t^2 \boldsymbol{\omega}\boldsymbol{\omega}^T + \Delta t \boldsymbol{\omega}^\times}{1 + \frac{1}{4}\Delta t^2 \boldsymbol{\omega}^2} \cdot \mathbf{R}(t_1) \end{aligned} \tag{36}$$

her. Dabei entspricht $\boldsymbol{\omega}^\times$ der Matrix

$$\boldsymbol{\omega}^\times = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

und es gilt allgemein $[\boldsymbol{\omega}^\times]\mathbf{x} = \boldsymbol{\omega} \times \mathbf{x}$. Wir nehmen die Länge des Zeitintervalls Δt als hinreichend kurz an, so daß wir Gl.(36) approximieren können durch

$$\mathbf{R}(t_1 + \Delta t) \approx \left[\mathbf{E} + \Delta t \boldsymbol{\omega}^\times + \frac{1}{2}\Delta t^2 \left(\boldsymbol{\omega}\boldsymbol{\omega}^T - \boldsymbol{\omega}^2 \mathbf{E}\right)\right] \cdot \mathbf{R}(t_1). \tag{37}$$

Mit

$$\begin{aligned} \hat{\mathbf{n}}^T \mathbf{R}_2^T(t) &= \hat{\mathbf{n}}^T \mathbf{R}_2^T(t_1) \cdot \mathbf{R}(1, -\frac{1}{2}\Delta t\, \boldsymbol{\omega}_2) \\ &= \mathbf{n}^T \cdot \mathbf{R}(1, -\frac{1}{2}\Delta t\, \boldsymbol{\omega}_2) \end{aligned}$$

und der Abkürzung $\mathbf{r}_1 = \mathbf{R}_1(t_1)\hat{\mathbf{a}}$ ergibt sich damit der approximative Kontaktabstand

$$\begin{aligned} d(t) &= \mathbf{n}^T \left(\mathbf{E} - (t-t_1)\boldsymbol{\omega}_2^\times + \frac{1}{2}(t-t_1)^2\left(\boldsymbol{\omega}_2\boldsymbol{\omega}_2^T - \boldsymbol{\omega}_2^2\mathbf{E}\right)\right) \cdot \\ &\quad \left[\left(\mathbf{E} + (t-t_1)\boldsymbol{\omega}_1^\times + \frac{1}{2}(t-t_1)^2\left(\boldsymbol{\omega}_1\boldsymbol{\omega}_1^T - \boldsymbol{\omega}_1^2\mathbf{E}\right)\right) \cdot \mathbf{r}_1 \right. \\ &\quad \left. + \mathbf{c}_1(t_1) - \mathbf{c}_2(t_1) + (t-t_1)(\mathbf{v}_1 - \mathbf{v}_2)\right] - \hat{n}_0 \\ d(t_1) &= \mathbf{n}^T\left[\mathbf{r}_1 + \mathbf{c}_1(t_1) - \mathbf{c}_2(t_1)\right] - \hat{n}_0. \end{aligned}$$

Bei der Auswertung dieses Ausdrucks vernachlässigen wir alle Terme mit $(t-t_1)^3$ oder höherer Ordnung.

Die uns interessierende Größe $\Delta d(t) = d(t) - \bar{d}(t)$ erhalten wir hieraus durch Einsetzen und wenige Umformungen zu

$$\begin{aligned} \Delta d(t) &= d(t) - d(t_1) - \frac{t-t_1}{t_2-t_1}\left[d(t_2) - d(t_1)\right] \\ &\approx \frac{1}{2}(t-t_1)(t_2-t)\mathbf{n}^T\left[\left(\boldsymbol{\omega}_1^2\mathbf{E} - \boldsymbol{\omega}_1\boldsymbol{\omega}_1^T\right)\cdot\mathbf{r}_1 \right. \\ &\quad \left. + \left(\boldsymbol{\omega}_2^2\mathbf{E} - \boldsymbol{\omega}_2\boldsymbol{\omega}_2^T\right)\cdot(\mathbf{r}_1 + \mathbf{c}_1(t_1) - \mathbf{c}_2(t_1)) + 2\boldsymbol{\omega}_2 \times (\boldsymbol{\omega}_1 \times \mathbf{r}_1 + \mathbf{v}_1 - \mathbf{v}_2)\right]. \end{aligned}$$

Der Betrag dieses Terms läßt sich approximativ nach oben abschätzen mit

$$\begin{aligned} |\Delta d(t)| &< \frac{1}{2}(t-t_1)(t_2-t)\Big[\boldsymbol{\omega}_1^2\cdot|\mathbf{r}_1| + \boldsymbol{\omega}_2^2\cdot|\mathbf{r}_1 + \mathbf{c}_1(t_1) - \mathbf{c}_2(t_1)| \\ &\quad + \ 2|\boldsymbol{\omega}_1|\cdot|\boldsymbol{\omega}_2|\cdot|\mathbf{r}_1| + 2|\boldsymbol{\omega}_2|\cdot(|\mathbf{v}_1| + |\mathbf{v}_2|)\Big], \end{aligned}$$

wobei $\mathbf{r}_1$ und $\mathbf{r}_2 = \mathbf{r}_1 + \mathbf{c}_1(t_1) - \mathbf{c}_2(t_1)$ die Vektoren vom jeweiligen Rotationszentrum der beiden Objekte zum Eckpunkt A bezeichnen (vgl. Abb. 28).

Wie zu erwarten, sind sowohl $\Delta d(t_1) = 0$ als auch $\Delta d(t_2) = 0$. Offensichtlich erhält man die größten Beträge von $\Delta d(t)$ bei $t \approx \frac{1}{2}(t_1 + t_2)$. Die maximale Differenz zwischen der Abstandsfunktion $d(t)$ und ihrer linearen Approximation $\bar{d}(t)$ läßt sich für den Punkt-Fläche-Kontakt somit approximieren durch

$$|\Delta d(t)| \ < \ \frac{1}{8}\Delta t^2\left[\boldsymbol{\omega}_1^2 r_1^{max} + \boldsymbol{\omega}_2^2 r_2^{max} + 2|\boldsymbol{\omega}_1|\cdot|\boldsymbol{\omega}_2|r_1^{max} + 2|\boldsymbol{\omega}_2|\cdot(|\mathbf{v}_1| + |\mathbf{v}_2|)\right], \quad (38)$$

wenn $r_1^{max} = \max_i |\hat{\mathbf{a}}_i|$ den maximalen Abstand aller Eckpunkte vom Zentrum des lokalen Koordinatensystems von Objekt 1 bezeichnet, und r_2^{max} den entsprechenden maximalen Eckpunktabstand bei Objekt 2. Die Unsymmetrie dieser Formel ist auf die ebenfalls unsymmetrische Beziehung (28) zurückzuführen.

Die im Ausdruck (38) enthaltenen Terme der Form $\frac{1}{8}\Delta t^2\omega^2 r$ und $\frac{1}{8}\Delta t^2\omega v$ lassen sich auch anschaulich gut nachvollziehen. Der maximale Abstand h zwischen Sehne und Bogen eines Kreissegmentes mit dem Winkel $\Delta t\omega$ und dem Radius r bestimmt sich zu $h =$

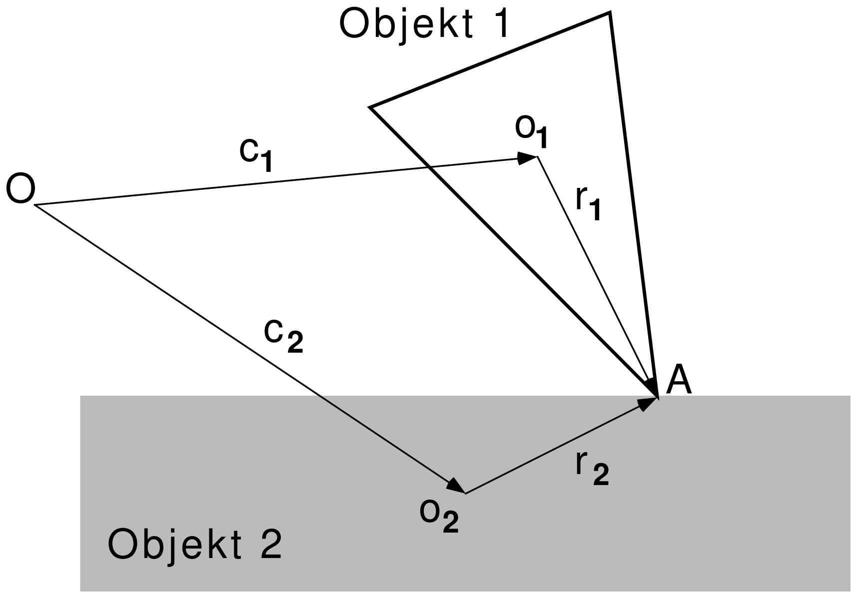


Abbildung 28: Erläuterung der in Gleichung (38) verwendeten Größen r_1 und r_2.

$r - r\cos(\frac{1}{2}\Delta t\omega)$. Ersetzt man für kleine Argumentwerte die Winkelfunktion durch ihre Reihenentwicklung bis zum quadratischen Term, $\cos x \approx 1 - \frac{1}{2}x^2$, so ergibt sich mit

$$\begin{aligned} h &= r - r\cos(\frac{1}{2}\omega\Delta t) \\ &\approx r - r\left(1 - \frac{1}{2}\left(\frac{1}{2}\omega\Delta t\right)^2\right) \\ &\approx \frac{1}{8}\Delta t^2\omega^2 r \end{aligned}$$

gerade ein Ausdruck der fraglichen Form. Berücksichtigt man weiter den Zusammenhang $v = \omega r$ für die Bahngeschwindigkeit entlang des Bogens, so erhält man entsprechend $h = \frac{1}{8}\Delta t^2\omega v$.

Mit Gl.(38) haben wir für den Fall des Ecke-Fläche-Kontaktes eine Abschätzung der im Zeitpunkt t_c^* zu erwartenden Abweichung zwischen der Abstandsfunktion $d(t)$ und ihrer linearen Approximation $\bar{d}(t)$ gewonnen, welche nur von den Objektdimensionen und den Beträgen der Objektgeschwindigkeiten abhängt.

Eine entsprechende Beziehung soll nun auch für den Fall des Kante-Kante-Kontaktes hergeleitet werden. Hiezu gehen wir von Gl.(32) aus. Die Orientierung der Berührnormale $\mathbf{n}$ hängt sowohl von der Rotation $\mathbf{R}_1$ als auch von der Rotation $\mathbf{R}_2$ ab. Wir approximieren den zeitlichen Verlauf $\mathbf{n}(t)$ durch eine gleichmäßige Rotationsbewegung wie folgt:

$$\begin{aligned} \mathbf{n}(t) &= \mathbf{R}_n(t)\hat{\mathbf{n}} \\ &= \mathbf{R}(1, \frac{1}{2}\Delta t\,\boldsymbol{\omega}_n)\,\mathbf{R}_n(t_1)\,\hat{\mathbf{n}} \end{aligned}$$

$$= \mathbf{R}(1, \frac{1}{2}\Delta t\, \boldsymbol{\omega}_n)\, \mathbf{n}(t_1).$$

Mit (37) und den Abkürzungen $\mathbf{r}_1 = \mathbf{R}_1(t_1)\hat{\mathbf{a}}_1$ sowie $\mathbf{r}_2 = \mathbf{R}_2(t_1)\hat{\mathbf{a}}_2$ erhält man somit den approximativen Kontaktabstand

$$\begin{aligned} d(t) &= \mathbf{n}^T(t_1)\left(\mathbf{E} - (t-t_1)\boldsymbol{\omega}_n^\times + \frac{1}{2}(t-t_1)^2\left(\boldsymbol{\omega_n}\boldsymbol{\omega}_n^T - \boldsymbol{\omega}_n^2\mathbf{E}\right)\right) \cdot \\ &\quad \left[\left(\mathbf{E} + (t-t_1)\boldsymbol{\omega}_1^\times + \frac{1}{2}(t-t_1)^2\left(\boldsymbol{\omega_1}\boldsymbol{\omega}_1^T - \boldsymbol{\omega}_1^2\mathbf{E}\right)\right) \cdot \mathbf{r}_1 \right. \\ &\quad -\left(\mathbf{E} + (t-t_1)\boldsymbol{\omega}_2^\times + \frac{1}{2}(t-t_1)^2\left(\boldsymbol{\omega_2}\boldsymbol{\omega}_2^T - \boldsymbol{\omega}_2^2\mathbf{E}\right)\right) \cdot \mathbf{r}_2 \\ &\quad \left. +\mathbf{c}_1(t_1) - \mathbf{c}_2(t_1) + (t-t_1)(\mathbf{v}_1 - \mathbf{v}_2)\right] \\ d(t_1) &= \mathbf{n}^T(t_1)\left[\mathbf{r}_1 + \mathbf{c}_1(t_1) - \mathbf{r}_2 - \mathbf{c}_2(t_1)\right]. \end{aligned}$$

Erneut vernachlässigen wir bei der Auswertung dieses Ausdrucks alle Terme mit $(t - t_1)^3$ oder höherer Ordnung und erhalten wir hieraus die uns interessierende Größe $\Delta d(t) = d(t) - \bar{d}(t)$ durch Einsetzen und wenige Umformungen zu

$$\begin{aligned} \Delta d(t) &= d(t) - d(t_1) - \frac{t-t_1}{t_2-t_1}\left[d(t_2) - d(t_1)\right] \\ &\approx \frac{1}{2}(t-t_1)(t_2-t)\mathbf{n}^T(t_1)\Big[\left(\boldsymbol{\omega}_1^2\mathbf{E} - \boldsymbol{\omega_1}\boldsymbol{\omega}_1^T\right) \cdot \mathbf{r}_1 - \left(\boldsymbol{\omega}_2^2\mathbf{E} - \boldsymbol{\omega_2}\boldsymbol{\omega}_2^T\right) \cdot \mathbf{r}_2 \\ &\quad + \left(\boldsymbol{\omega}_n^2\mathbf{E} - \boldsymbol{\omega_n}\boldsymbol{\omega}_n^T\right) \cdot (\mathbf{r}_1 + \mathbf{c}_1(t_1) - \mathbf{r}_2 - \mathbf{c}_2(t_1)) \\ &\quad + 2\boldsymbol{\omega}_n \times \boldsymbol{\omega}_1 \times \mathbf{r}_1 - 2\boldsymbol{\omega}_n \times \boldsymbol{\omega}_2 \times \mathbf{r}_2 + 2\boldsymbol{\omega}_n \times (\mathbf{v}_1 - \mathbf{v}_2)\Big]. \end{aligned}$$

Der Betrag dieses Terms läßt sich approximativ nach oben abschätzen mit

$$\begin{aligned} |\Delta d(t)| &< \frac{1}{2}(t-t_1)(t_2-t)\Big[\boldsymbol{\omega}_1^2 \cdot |\mathbf{r}_1| + \boldsymbol{\omega}_2^2 \cdot |\mathbf{r}_2| + \boldsymbol{\omega}_n^2 \cdot |\mathbf{r}_1 + \mathbf{c}_1(t_1) - \mathbf{r}_2 - \mathbf{c}_2(t_1)| \\ &\quad + 2|\boldsymbol{\omega}_n| \cdot (|\boldsymbol{\omega}_1|\,|\mathbf{r}_1| + |\boldsymbol{\omega}_2||\mathbf{r}_2|) + 2|\boldsymbol{\omega}_n| \cdot (|\mathbf{v}_1| + |\mathbf{v}_2|)\Big]. \end{aligned}$$

An dieser Stelle schätzen wir nun die Rotationsgeschwindigkeit $|\boldsymbol{\omega}_n|$ der Kontaktnormale nach oben ab durch

$$|\boldsymbol{\omega}_n| \leq |\boldsymbol{\omega}_1| + |\boldsymbol{\omega}_2| \tag{39}$$

und erhalten den Gl.(38) entsprechenden Ausdruck für den Kante-Kante-Kontakt zu

$$\begin{aligned} |\Delta d(t)| &< \frac{1}{8}\Delta t^2\Big[\boldsymbol{\omega}_1^2 r_1^{max} + \boldsymbol{\omega}_2^2 r_2^{max} + (|\boldsymbol{\omega}_1| + |\boldsymbol{\omega}_2|)^2\,|d(t_1)| \\ &\quad + 2\,(|\boldsymbol{\omega}_1| + |\boldsymbol{\omega}_2|) \cdot (|\boldsymbol{\omega}_1|\, r_1^{max} + |\boldsymbol{\omega}_2| r_2^{max} + |\mathbf{v}_1| + |\mathbf{v}_2|)\Big]. \end{aligned} \tag{40}$$

Vergleicht man den Aufbau der Beziehungen (32) und (33), so erkennt man, daß wir mit (40) zugleich den allgemeinen Fall gelöst haben, wenn die Abschätzung (39) gilt. Ein Vergleich mit (38) bestätigt, daß der Fall des Ecke-Fläche-Kontaktes mit (40) ebenfalls abgedeckt ist.

Diese Beziehung verifizieren wir nun empirisch, indem wir eine Stichprobe stochastisch erzeugter Kollisionen betrachten, und die im jeweiligen Kollisionszeitpunkt t_c^* gemessenen Abweichungen $\Delta d(t)$ mit dem gemäß Gl. (40) berechneten Wert in Beziehung setzen. In Abb. 29 sind die resultierenden Quotienten $\frac{|\Delta d(t)|}{h(t)}$ einer solchen Stichprobe dargestellt. Erwartungsgemäß sind alle in dieser Stichprobe enthaltenen Werte betragsmäßig kleiner als 1.

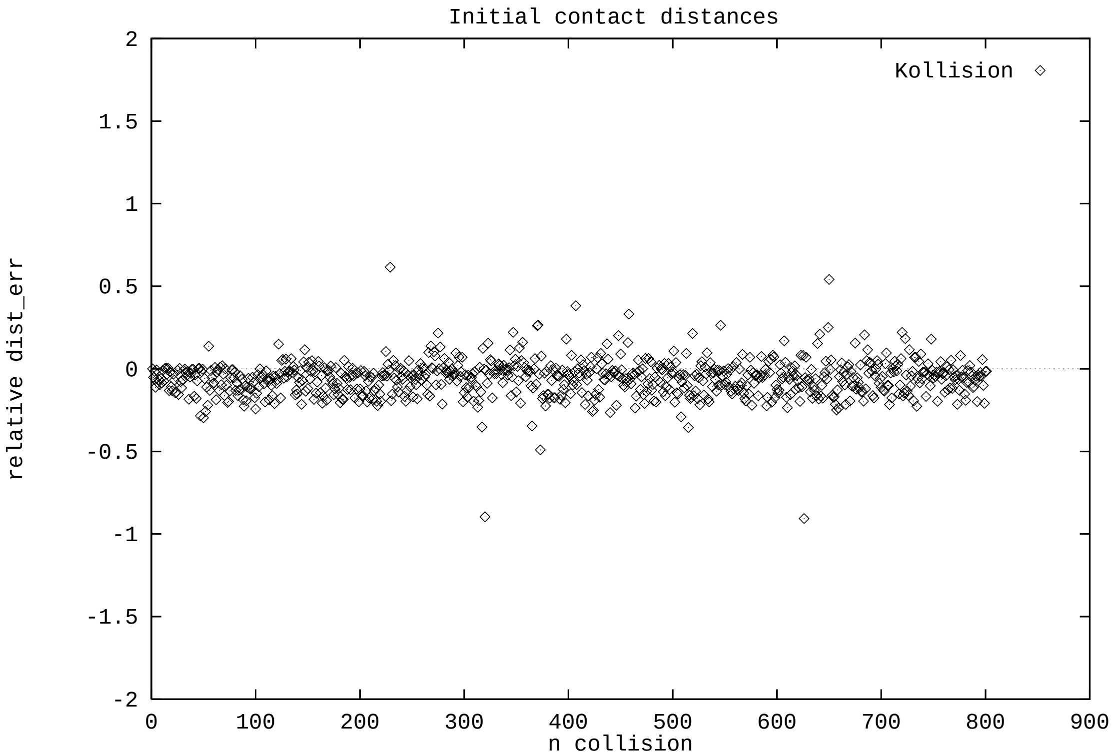


Abbildung 29: Darstellung des Quotienten $\frac{|\Delta d(t)|}{h(t)}$ für eine Stichprobe stochastisch erzeugter Kollisionen.

5.2.3 Zur Eindeutigkeit der Kontaktpunktbestimmung

Wie bereits dargestellt, werden zur Erkennung von Kollisionen viele Intersektionstests zwischen einzelnen geometrischen Elementen (Flächen, Kanten oder Ecken) durchgeführt. Entscheidend für das Resultat ist die früheste der dabei erkannten Kollisionen. Wie ist nun der Fall zu behandeln, daß der früheste Kontakt an mehreren Stellen exakt gleichzeitig eintritt? Zunächst ist festzustellen, daß aufgrund der bereits erwähnten numerischen Ungenauigkeiten (vgl. Abschnitt 5.2.1) von einer exakten Gleichheit mehrerer Kollisionszeitpunkte nicht zuverlässig gesprochen werden kann. Wenn aber nicht entschieden werden kann, ob mehrere Kollisionen exakt gleichzeitig stattfinden, so ist es auch sinnlos, mehr als einen Kontaktpunkt als Ort der frühesten Kollision zu registrieren. Sollten tatsächlich für mehrere Kontaktpunkte identische Kollisionszeitpunkte bestimmt werden, so ist es ausreichend, einen davon als den frühesten Kontaktpunkt auszuwählen. Dies kann entweder zufällig geschehen (etwa abhängig von der Reihenfolge der verschiedenen Intersektionstests), oder nach einem sonstigen Kriterium (ein solches ist jedoch nicht offensichtlich).

Es soll noch daran erinnert werden, daß es auch aus der Perspektive der Kontaktsimulation sehr vorteilhaft sein kann, im Rahmen der Kollisionserkennung jeweils nur höchstens einen neuen Punktkontakt während eines Simulationsschrittes zu registrieren. Dies wurde in Abschnitt 5.1 erläutert.

5.3 Kontaktablösung

Die mathematischen Formulierungen der beiden elementaren Kontaktarten Ecke-Fläche und Kante-Kante beinhalten jeweils zwei Kontaktbedingungen, die gleichzeitig erfüllt sein müssen:

1. **Berührbedingung**
   Der Kontakt muß die Berührbedingung für die übergeordneten Punktmengen (d.h. die Kontaktbedingung Punkt-Ebene bzw. Gerade-Gerade, vgl. Gln. (28, 32) erfüllen.

2. **Segmentbedingung**
   Der Kontaktpunkt muß innerhalb der am Kontakt beteiligten Flächensegmente bzw. Kantensegmente liegen (Gln. 44, 45).

Diesen kombinierten Bedingungen entsprechen jeweils zwei unterschiedlich zu handhabende Formen der Kontaktablösung, nämlich

1. die Ablösung in Richtung der Berührnormale.

2. die Ablösung durch Verlassen des berührten Kantensegmentes oder Flächensegmentes parallel zur Berührebene (Abb. 31).

Während eines Simulationsintervalls müssen beide Kontaktbedingungen eines jeden Kontaktes geprüft werden. Ist für einen Punktkontakt mindestens eine der Bedingungen nicht mehr erfüllt, wird dieser Kontakt eliminiert, und die zugehörige Randbedingung aus dem System der Bewegungsgleichungen entfernt.

5.3.1 Kontaktablösung in Normalenrichtung

Anhand welcher Kriterien läßt sich nun entscheiden, ob ein Kontakt sich in Richtung der Berührnormalen ablöst? Zunächst gilt, daß die resultierende Bewegung eines Körpers sich aus allen auf ihn wirkenden Kräften ergibt. Hierzu zählen bei Vorliegen kontaktbedingter Bewegungsbeschränkungen auch die an den Kontaktstellen wirkenden Kräfte, zusätzlich zu allen sonstigen äußeren Kräften, wie z.B. der Schwerkraft.

Ob ein Kontakt sich löst, ergibt sich aus der Gesamtbewegung der beiden beteiligten Objekte, und hängt somit auch von allen auf diese Objekte wirkenden Kräften ab.

Diese Zusammenhänge werden im Folgenden für zwei verschiedene Formulierungen der Kontaktbedingungen dargestellt.

Behandlung mit bilateralen Kontaktbedingungen

Eine einfache Möglichkeit zur Simulation von Kontakten stellt die Aufstellung bilateraler Bewegungseinschränkungen dar, obwohl die Kontaktbedingungen ihrer Natur nach eigentlich unilaterale Beschränkungen darstellen. Der große Vorteil bilateraler Kontaktbedingungen liegt darin, daß sie durch Gleichungssysteme darstellbar sind, während unilaterale Kontaktbedingungen durch Ungleichungssysteme beschrieben werden, deren Lösung wesentlich aufwendiger ist. Diesen Vorteil erkauft man sich jedoch mit dem Nachteil, daß an den Kontakten negative (d.h. anziehende) Kontaktkräfte auftreten können, was physikalisch bei Vernachlässigung von Adhäsionseffekten nicht der Fall ist. Ohne weitere Maßnahmen würden solche bilateralen Kontakte sich nicht mehr lösen, was absolut unrealistisch wäre. Allerdings läßt sich gerade das Auftreten negativer Kontaktkräfte als Indiz dafür ansehen, daß sich ein Kontakt lösen muß. Die Nachbildung unilateraler Kontakte durch bilaterale Kontaktbedingungen besteht also darin, Kontakte stets dann zu lösen, wenn negative Kontaktkräfte auftreten. Da die Kontaktkräfte beim Lösen des Systems der bedingten Bewegungsgleichungen als Zwischenergebnis anfallen, ist der zusätzliche Aufwand zur Entscheidung über sich lösende Kontakte minimal. Aufgrund der kurzzeitig (maximal ein Simulationsintervall) wirkenden negativen Kontaktkräfte ergeben sich jedoch auch gewisse störende Effekte.

Dies soll das folgende Beispiel veranschaulichen. Bei der in Abb. 30 dargestellten Situation preßt zum Zeitpunkt t_1 das interaktiv kontrollierte Objekt A ein zweites bewegliches Objekt B gegen das feststehende Hindernis C. Es besteht je ein Kontaktpunkt zwischen A und B sowie zwischen B und C. Zum Zeitpunkt t_2 kehrt sich die Bewegungstendenz von Objekt A um. Da die bilateralen Kontaktbedingungen zunächst jedoch noch gültig sind, wird der Kontakt an beiden Berührpunkten aufrechterhalten, allerdings jeweils mit negativen Kontaktkräften. Aufgrund dieser Tatsache werden nun beide Kontaktbedingungen eliminiert. Zum Zeitpunkt t_3 kann sich Objekt A nun frei von Objekt B lösen. Objekt B selbst verbleibt nach wie vor in Ruhe, da keine Kräfte darauf wirken.

In diesem Fall ist das resultierende Bewegungsverhalten aller beteiligten Objekte korrekt. Die kurzzeitig wirkenden negativen Kontaktkräfte führen zu einer Verzögerung der Ablösebewegung um ein Simulationsintervall (in der Größenordnung von 10ms), was äußerlich kaum bemerkbar ist. Allerdings kann es in vergleichbaren Situationen geschehen, daß aufgrund negativer Kontaktkräfte Kontakte gelöst werden, welche eigentlich bestehen bleiben sollten. Man denke an einen Stapel aufeinander geschichteter Quader, welche durch die Schwerkraft aufeinandergedrückt werden ($t = t_1$). Nun wird der oberste, interaktiv kontrollierte Quader ruckartig nach oben gezogen. Wie bei einer gespannten Kette nehmen nun alle Kontaktkräfte zwischen den Quadern negative Werte an, und die Kontakte lösen sich ($t = t_2$). Tatsächlich bewegt sich anschließend nur der oberste Quader nach oben weg. Die Kontakte zwischen allen anderen Quadern müssen sich, wegen der wirkenden Schwerkraft, wieder neu aufbauen. Dieser Ablauf wirkt zwar äußerlich ebenfalls korrekt. Da in jedem Simulationsintervall jedoch nur ein neuer Kontaktpunkt hinzukommen kann (vgl. Abschnitt 5.1), vergehen unter Umständen zahlreiche Intervalle, bis alle Kontaktpunkte wieder hergestellt sind.

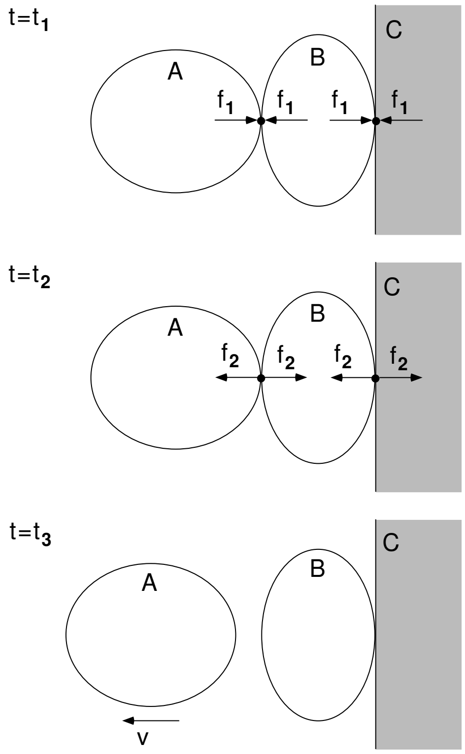


Abbildung 30: Beispiel einer Kontaktablösung mit bilateralen Kontaktbedingungen.

Behandlung mit unilateralen Kontaktbedingungen

Die physikalisch korrekte Formulierung von unilateralen Kontaktbedingungen läßt sich als sogenanntes lineares Komplementaritätsproblem darstellen. Die in der Literatur (z.B. [3, 44]) vorherrschende Beschreibung beruht auf folgenden Überlegungen.

Analog zu Gleichung (8) beschreibe $\mathbf{a} = \mathbf{M}^{-1}\mathbf{C}\mathbf{f} + \mathbf{a}_{ext}$ den Einfluß der Kontaktkräfte $\mathbf{f}$ auf die Objektbeschleunigung. Die Normalkomponenten der Kontaktbeschleunigung $\mathbf{a}_\perp = \mathbf{C}^T\mathbf{a}$ sind positiv, wenn sich Kontakte zu lösen beginnen, und negativ, wenn sich die einander berührenden Objekte zu durchdringen beginnen. Daher lautet die zugehörige Kontaktbedingung $\mathbf{a}_\perp \geq \mathbf{0}$. Für die Kontaktkräfte gilt gleichermaßen $\mathbf{f} \geq \mathbf{0}$. Da an einem sich lösenden Kontakt die Kontaktkraft verschwindet, und bei positiver Kontaktkraft an einem bestehenden Kontakt die Normalkomponente der Kontaktbeschleunigung verschwindet, gilt für f und $\mathbf{a}_\perp$ zusätzlich die Komplementaritätsbedingung $\mathbf{a}_\perp^T\mathbf{f} = 0$. Insgesamt erhält man damit die Bedingungen

$$\mathbf{a}_\perp \geq \mathbf{0}, \qquad \mathbf{f} \geq \mathbf{0}, \qquad \mathbf{a}_\perp^T\mathbf{f} = 0. \tag{41}$$

Zusammen mit der linearen Beziehung zwischen f und $\mathbf{a}_\perp$,

$$\mathbf{a}_\perp = \mathbf{C}^T\mathbf{M}^{-1}\mathbf{C}\mathbf{f} + \mathbf{C}^T\mathbf{a}_{ext}, \tag{42}$$

stellt die Bestimmung von f und $\mathbf{a}_\perp$ ein lineares Komplementaritätsproblem (LCP, siehe auch [14, 15]) dar. Zur Lösung dieses LCP wird häufig der dem Simplex-Algorithmus verwandte Lemke-Algorithmus [27] herangezogen, dessen Konvergenz unter den gegebenen Bedingungen ($\mathbf{A} = \mathbf{C}^T\mathbf{M}^{-1}\mathbf{C}$ ist zumindest positiv semi-definit) gesichert ist.

Dieses Verfahren hat allerdings den Nachteil, daß es nicht auf einer unmittelbaren Randbedingung für die Objektlage basiert, sondern auf einer Randbedingung für die Objektbeschleunigung. Im Laufe der Simulation ist daher mit der Akkumulation von Integrationsfehlern zu rechnen, was zu einem zunehmenden Abdriften von der korrekten Berührsituation führen kann. In Abschnitt 3.4 wurde eine Vorgehensweise zur Stabilisierung der Kontaktbedingungen dargestellt, die solche Driftprobleme vermeidet, indem sie die Kontaktbedingungen direkt auf die Kontaktabstände $\mathbf{w} = \mathbf{g}(\mathbf{f}), \mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ bezieht. Die Funktion $\mathbf{g}(\mathbf{f})$ ist jedoch nichtlinear in $\mathbf{f}$, so daß sich die zugehörigen Komplementaritätsbedingungen (vgl. Gl. (41))

$$\mathbf{w} \geq \mathbf{0}, \qquad \mathbf{f} \geq \mathbf{0}, \qquad \mathbf{w}^T\mathbf{f} = 0 \tag{43}$$

nicht mehr mit dem Lemke-Algorithmus oder anderen Verfahren lösen lassen, welche für den linearen Fall konzipiert sind. Ein Verfahren zur numerische Lösung des nichtlinearen Komplementaritätsproblems (43) wird in Kapitel 6 vorgestellt.

5.3.2 Tangentiale Kontaktablösung

Die mathematische Formulierung der Kontaktbedingungen in Abschnitt 4.5 entspricht eigentlich den Bedingungen für einen Punkt-Ebene-Schnitt beziehungsweise für den Schnitt zweier Geraden. Die an einem Kontakt beteiligten Flächen und Kanten sind jedoch nicht unendlich ausgedehnt, sondern stellen endliche Ebenensegmente beziehungsweise Geradensegmente dar. Verläßt ein Kontaktpunkt nun das jeweilige Ebenensegment oder Geradensegment, so löst sich dieser Kontakt entweder, oder er geht unmittelbar in einen neuen Kontakt mit einem angrenzenden Kontaktelement über (siehe Abschnitt 5.4).

Wir betrachten nun die Segmentbedingungen für eine solche tangentiale Kontaktablösung.

Segmentbedingungen eines Punkt-Fläche-Kontaktes

Ein Punkt-Fläche-Kontakt werde durch den Punkt P und das in der Ebene S liegende Flächensegment F gebildet. Gemäß den Bezeichnungen in Abb. 2 entspricht das Innere der Fläche F der Menge aller Punkte $\mathbf{x} \in S$, für die gilt

$$\forall_i : \quad (\mathbf{n} \times (\mathbf{a}_{i+1} - \mathbf{a}_i))^T(\mathbf{x} - \mathbf{a}_i) \geq 0. \tag{44}$$

Verletzt der Kontaktpunkt P diese Bedingung, so liegt er außerhalb des Flächensegmentes, und der Kontakt zwischen P und F löst sich (vgl. Abb. 31).

Segmentbedingungen eines Kante-Kante-Kontaktes

Ein Kante-Kante-Kontakt werde gebildet durch zwei Kanten mit den Endpunkten $\mathbf{a}_1$, $\mathbf{b}_1$ beziehungsweise $\mathbf{a}_2$, $\mathbf{b}_2$. Der Schnittpunkt der beiden Geraden, in welche diese Kanten eingebettet sind, ergibt sich aus der Schnittbedingung

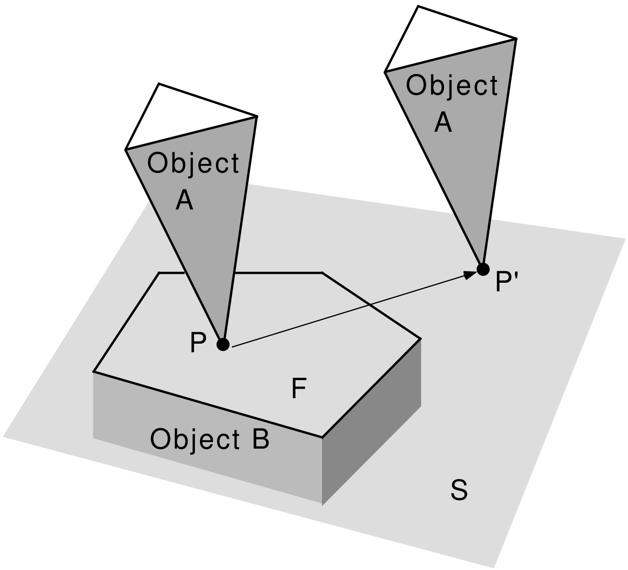


Abbildung 31: Der Punkt-Fläche-Kontakt zwischen der Ecke P und dem Flächensegment $F \subset S$ löst sich, wenn P entlang der Ebene S gleitet, und dabei das Flächensegment verläßt.

$$\mathbf{a}_1 + \lambda_1(\mathbf{b}_1 - \mathbf{a}_1) = \mathbf{a}_2 + \lambda_2(\mathbf{b}_2 - \mathbf{a}_2).$$

Wenn sich die Geraden schneiden, so hat dieses Gleichungssystem die Lösungen

$$\begin{aligned} \lambda_1 &= \frac{(\mathbf{b}_2 - \mathbf{a}_2)^T(\mathbf{n} \times (\mathbf{a}_2 - \mathbf{a}_1))}{\mathbf{n}^2}, \\ \lambda_2 &= \frac{(\mathbf{b}_1 - \mathbf{a}_1)^T(\mathbf{n} \times (\mathbf{a}_2 - \mathbf{a}_1))}{\mathbf{n}^2}. \end{aligned}$$

Der so berechnete Schnittpunkt liegt aber nur dann im Inneren beider Kantensegmente, wenn die folgenden Segmentbedingungen erfüllt sind:

$$0 \leq \lambda_1 \leq 1 \quad \wedge \quad 0 \leq \lambda_2 \leq 1. \tag{45}$$

Andernfalls schneiden sich die beiden Geradensegmente nicht, und der Kante-Kante-Kontakt löst sich.

5.4 Kontaktübergänge

Wir betrachten zwei sich berührende Objekte A und B, welche aneinander entlang gleiten (Abb. 32). Der zu Objekt A gehörende Punkt P bildet zunächst einen Ecke-Fläche-Kontakt mit der zu Objekt B gehörenden Fläche F_1. Im Laufe der Gleitbewegung wechselt der Kontaktpunkt auf die zu F_1 benachbarte Objektfläche F_2 über. Die Gestaltung dieses

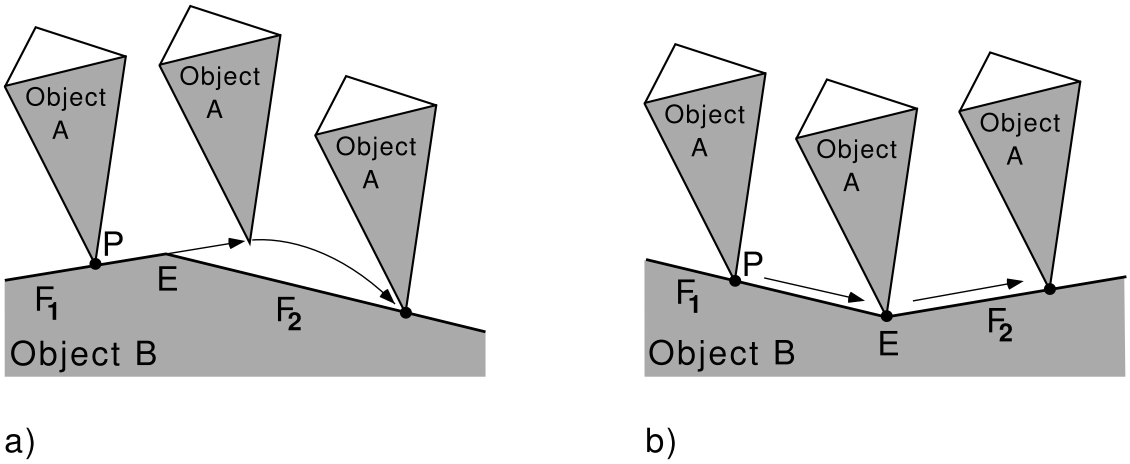


Abbildung 32: Übergang des Kontaktes zwischen Eckpunkt P und Fläche F_1 auf den Kontakt zwischen P und F_2. Im konvexen Fall (a) löst sich der Kontakt während des Überganges vollkommen. Im konkaven Fall (b) berührt P während des Überganges zeitweilig beide Flächen.

Kontaktübergangs hängt nun davon ab, ob die den Flächen F_1 und F_2 gemeinsame Kante E konvex oder konkav ist.

Ist E konvex (Abb. 32a), löst sich P zuerst von F_1 aufgrund einer Verletzung der Segmentbedingung, bevor nach einer Kollision mit F_2 ein neuer Kontakt entstehen kann. Kurzzeitig besteht also keine Verbindung zwischen dem Punkt P und Objekt B.

Ist E dagegen konkav (Abb. 32b), so kollidiert P mit der Fläche F_2, während der Kontakt mit F_1 noch besteht. Zu diesem Zeitpunkt bestehen damit zwei Ecke-Fläche-Kontakte zwischen P und Objekt B. Wenn sich die Gleitbewegung zwischen den beiden Körpern in der ursprünglichen Richtung fortsetzt, so löst sich P von der Fläche F_1 und gleitet entlang F_2 weiter.

Diese Abläufe wurden hier am Beispiel eines Ecke-Fläche-Kontaktes dargestellt, der in einen neuen Ecke-Fläche-Kontakt übergeht. Ganz analoge Abläufe treten beim Kontaktübergang an Kante-Kante-Kontakten auf. Zudem kann bei einem Kontaktübergang ein Ecke-Fläche-Kontakt in einen Kante-Kante-Kontakt übergehen und umgekehrt (Abb. 33).

Bei allen diesen Abläufen entsteht der neue Kontaktpunkt jeweils aufgrund einer Kollision. Es sind jedoch auch fließende Kontaktübergänge möglich, bei denen keine Kollisionen auftreten.

Fließende Kontaktübergänge

Neue Kontakte entstehen typischerweise aufgrund von lokalen Kollisionen. Dabei wird die Objektbewegung im Kollisionszeitpunkt unterbrochen, und eine neue Kontaktbedingung in die Liste der Bewegungsbeschränkungen aufgenommen. Anschließend wird die Bewegungssimulation mit diesen erweiterten Randbedingungen fortgesetzt. Der in Abschnitt 5.1 beschriebene Mechanismus zur Erzwingung minimaler Punktkontaktkonfigu-

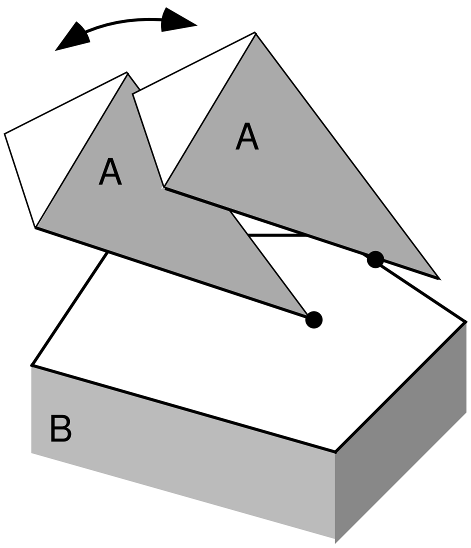


Abbildung 33: Kontaktübergang zwischen einem Punkt-Fläche-Kontakt und einem Kante-Kante-Kontakt.

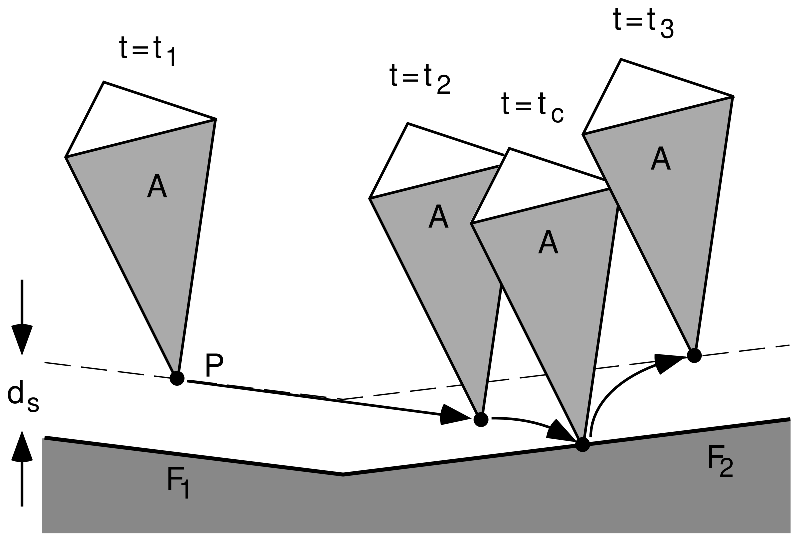
rationen erfordert jedoch, daß in jedem Simulationsintervall nur höchstens eine neue Kollision berücksichtigt werden kann. Zwischen zwei Kollisionsvorgängen muß daher jeweils mindestens ein Simulationsschritt durchgeführt werden, welcher die iterative Lösung eines Gleichungssystems und somit einen gewissen Zeitaufwand erfordert. Treten nun viele Kollisionen in rascher Folge auf, so steigt der Simulationsaufwand, und die Simulation wird insgesamt verlangsamt.

Bei Kontaktübergängen kann jedoch in bestimmten Fällen der neue Kontakt auch ohne eine Kollision realisiert werden. Dies sei anhand von Abb. 34 erläutert. Ein Eckpunkt P von Objekt A gleite entlang einer Fläche F_1 von Objekt B, unter Einhaltung des in Abschnitt 5.1.1 beschriebenen Schlupfabstandes ε_s. Am Ende des Simulationsintervalles $[t_1, t_2]$ liegt P außerhalb des Flächensegmentes F_1, und dieser Kontakt löst sich. Gleichzeitig liege P innerhalb des Flächensegmentes F_2. Ein fließender Kontaktübergang wird nun dadurch realisiert, daß der alte Kontakt mit F_1 unmittelbar durch den neuen Kontakt zwischen P und F_2 ersetzt wird, ohne eine Kollision zwischen diesen Elementen abzuwarten.

Dieser Mechanismus greift jedoch nur dann, wenn P zum Zeitpunkt t_2 einen Abstand $0 \leq d(t_2) \leq 2\varepsilon_s$ von der Fläche F_2 hat (vgl. Abb. 34). Bei negativen Werten von $d(t_2)$ tritt bereits bei $t_c < t_2$ eine Kollision auf. Bei Werten von $d(t_2) > 2\varepsilon_s$ wird kein fließender Kontaktübergang durchgeführt, da bei großen Werten von $d(t_2)$ ein späterer Kontakt zwischen P und F_2 zunehmend unwahrscheinlich wird. Insgesamt beschränken sich fließende Kontaktübergänge im wesentlichen auf solche Situationen, bei denen die betrachteten benachbarten Flächen F_1 und F_2 nur geringfügig verschiedene Orientierungen besitzen.

5.5 Entstehung von Gleichgewichtslagen

Gegeben sei die Anordnung in Abb. 35a, ein Prisma A, dessen Flächenkontakt mit der Unterlage B durch drei Punktkontakte an den Ecken P_1, P_2 und P_3 realisiert ist. Außer



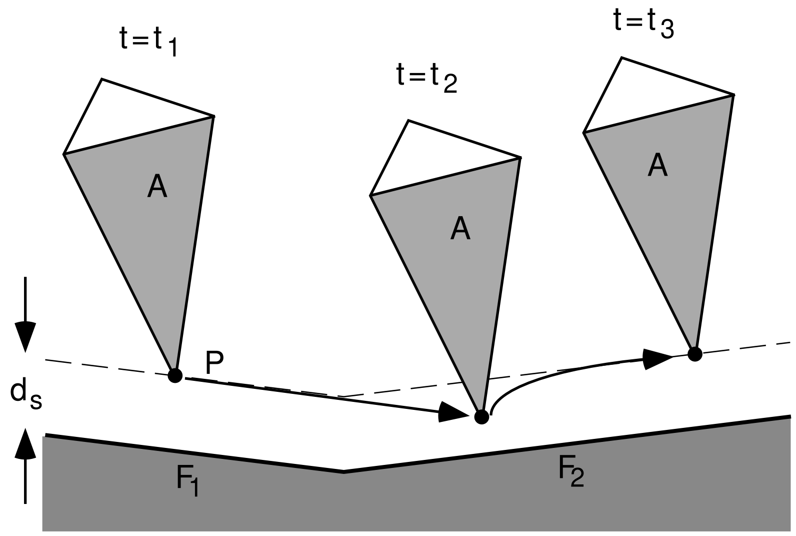


Abbildung 34: Fließender Kontaktübergang zwischen zwei Punkt-Fläche-Kontakten.

den Kontaktkräften wirke auf A nur die Gewichtskraft, die im Schwerpunkt S ansetzt. Zwar definieren die drei Punktkontakte eindeutig die parallele Lage der sich berührenden Flächensegmente. Da an ihnen keine negativen (anziehenden) Kontaktkräfte wirken können, der Lotfußpunkt des Schwerpunkts S jedoch außerhalb des Dreiecks (P_1 P_2 P_3) liegt, ist die Anordnung statisch nicht stabil.

Dies hat zur Folge, daß sich der Kontakt bei P_1 löst, und Objekt A um die Achse $\overline{P_2 P_3}$ kippt, bis es bei dem Eckpunkt P_4 mit Objekt B kollidiert. Diese Situation ist in Abb. 35b dargestellt. Durch den Sicherheitsabstand ε_s, der an den Kontaktpunkten eingehalten wird, sind die Objekte in dieser Lage leicht gegeneinander geneigt (der Neigungswinkel ist in der Abbildung stark übertrieben dargestellt). Für den neuen Kontaktpunkt bei P_4 gilt ebenfalls der Kontaktabstand ε_s, so daß die Objekte anschließend wieder parallel ausgerichtet sind (Abb. 35c).

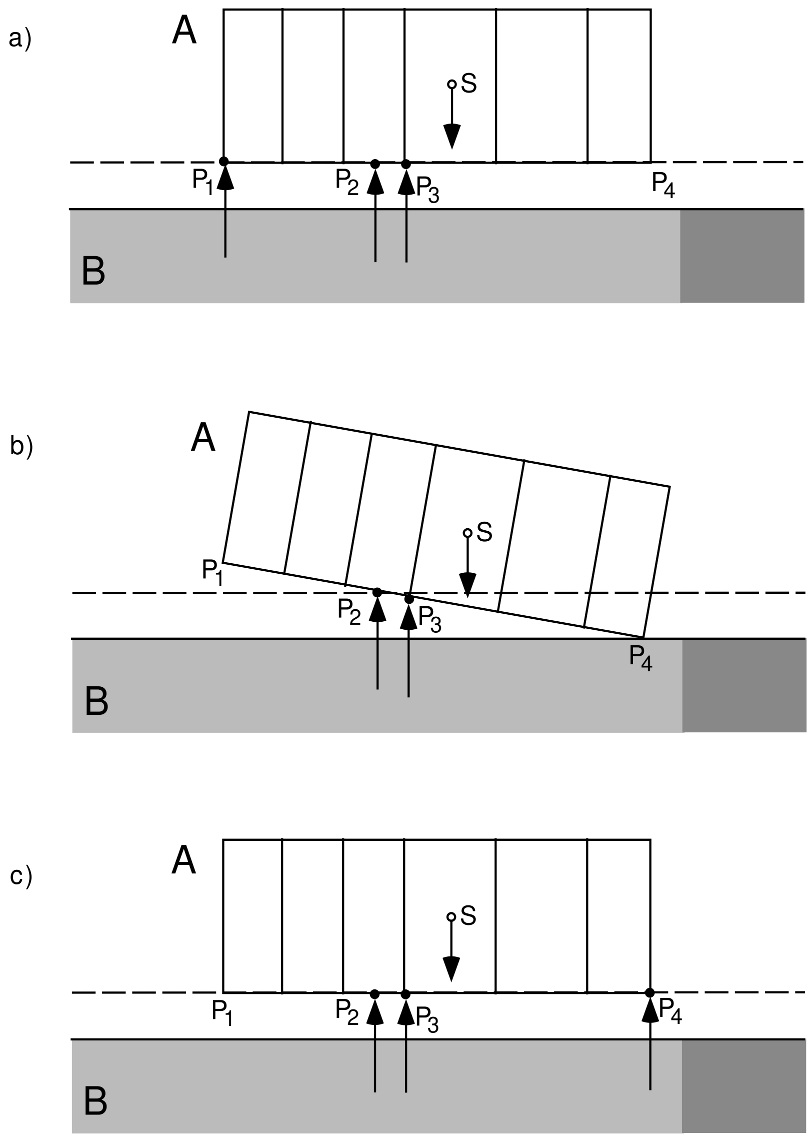


Abbildung 35: Kontaktwechsel zur Erzielung von Gleichgewichtslagen. (a) Das in instabiler Lage befindliche Objekt A löst sich am Kontaktpunkt P_1 von der Unterlage B. (b) Die Objekte A und B kollidieren bei P_4 miteinander. (c) Der neue Punktkontakt bei P_4 stellt die stabile Gleichgewichtslage her.

Geht nun die Schwerpunktsachse durch das Dreieck der Kontaktpunkte, so ist eine Gleichgewichtslage erreicht. Andernfalls wechselt die Kontaktpunktkonfiguration erneut, bis sich das Gleichgewicht einstellt (vgl. Abb. 36). Dieser Mechanismus, der hier am Beispiel der wirkenden Schwerkraft dargestellt wurde, funktioniert selbstverständlich ebenso, wenn andere externe Kräfte auf die sich berührenden Objekte wirken. Generell entwickelt sich die Anordnung der Kontaktpunkte so, daß ein statisches beziehungsweise dynamisches Gleichgewicht aufgebaut wird.

Abschließend soll noch auf ein interessantes Merkmal dieses Mechanismus eingegan-

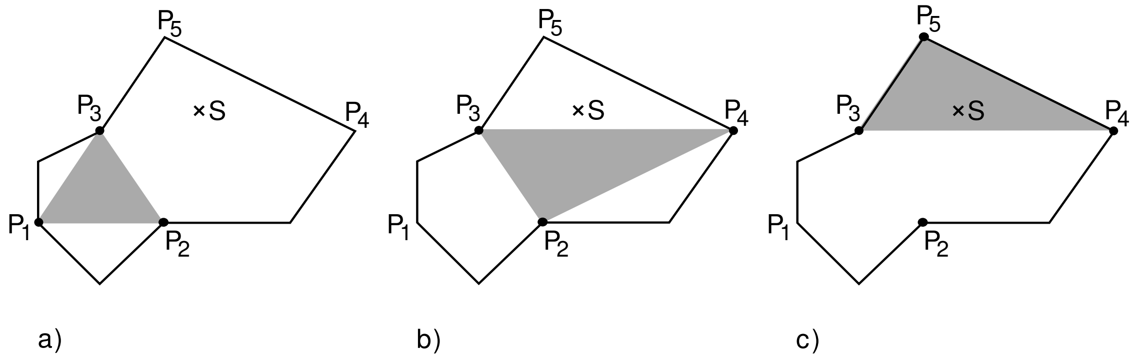


Abbildung 36: (a) Die polygonale Kontaktregion wird zunächst durch die drei Kontaktpunkte P_1, P_2 und P_3 repräsentiert. Nun greift eine externe Kraft an, deren Kraftlinie durch den Punkt S geht. (b) Durch eine geringfügige Kippbewegung um die Achse $\overline{P_2P_3}$ löst sich der Kontaktpunkt P_1, und ein neuer Kontaktpunkt entsteht bei P_4. Da S nicht im Dreieck (P_2, P_3, P_4) liegt, ist diese Situation ebenfalls instabil. (c) Durch eine erneute Kippbewegung löst sich der Kontaktpunkt P_2, und ein neuer Kontaktpunkt entsteht bei P_5. Da S im Dreieck (P_3, P_4, P_5) liegt ist diese Situation stabil.

gen werden. Wie aus Abb. 35b ersichtlich ist, entsteht der neue Kontaktpunkt genau an jenem Eckpunkt der Grundfläche des Prismas, welcher den größten senkrechten Abstand von der Kippachse $\overline{P_2P_3}$ hat. Dadurch besitzt das neue Kontaktpunktdreieck, bezogen auf die Basis $\overline{P_2P_3}$, die maximal mögliche Fläche. Dies kommt der numerischen Stabilität der Kontaktsimulation zugute (vgl. hierzu Abb. 12). Die Fläche des von den Kontaktpunkten gebildeten Dreiecks wird durch diesen Mechanismus jedoch nicht unbedingt global maximiert. Das flächengrößte Kontaktpunktdreieck in Abb. 36c wäre (P_1, P_4, P_5).

5.6 Zusammenfassung

Die Anpassung der Kontaktbedingungen an Veränderungen der Kontaktsituation fassen wir noch einmal anhand des Struktogrammes in Abb. 37 zusammen.

Wir gehen aus von einer simulierten Objektbewegung, welche sich durch die iterative Lösung der Bewegungsgleichungen unter Berücksichtigung der aktuell bestehenden Bewegungsrestriktionen ergibt. Nachdem für alle beweglichen Objekte auf diese Weise eine neue Lage bestimmt wurde, prüft ein externes Modul zur Kollisionserkennung, ob im letzten Bewegungsschritt Kollisionen aufgetreten sind. Werden Kollisionen registriert, so liefert das Kollisionserkennungsmodul auch approximative Kollisionszeitpunkte t_c^* hierfür. Die Kollisionszeitpunkte werden dabei durch lineare Interpolation des Kontaktabstandes zwischen Beginn und Ende des Simulationsintervalles approximiert. Unter Verwendung der Bewegungsgleichungen (21,22) kann t_c^* mit Hilfe der Regula Falsi iterativ verbessert werden. Die simulierte Objektbewegung wird im Kollisionsfall nur bis zum derart bestimmten Kollisionszeitpunkt durchgeführt, und der neue Kontaktpunkt in die Liste der Kontaktbedingungen aufgenommen.

Die resultierende neue Objektlage wird schließlich daraufhin geprüft, ob sich beste-

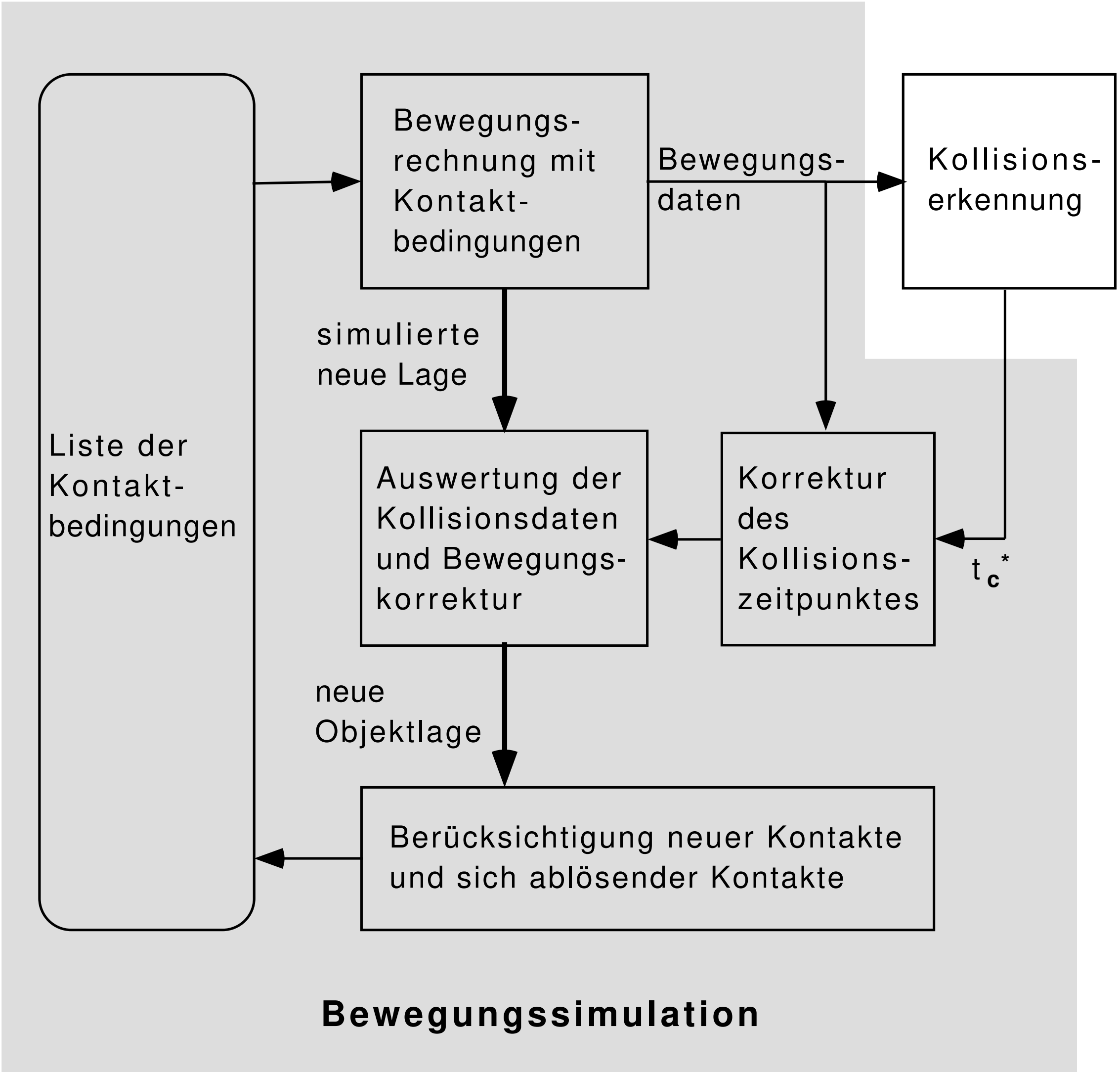


Abbildung 37: Struktogramm zur Anpassung der Kontaktbedingungen an Veränderungen der Kontaktsituation.

hende Kontakte gelöst haben, oder ob Kontaktübergänge stattfanden. Auch solche Änderungen der Kontaktsituation werden in der Liste der Kontaktbedingungen berücksichtigt.

6 Numerische Lösung des Gleichungssystems

Bevor wir auf die Lösung des Gleichungssystems zur Simulation kontaktbedingter Bewegungseinschränkungen genauer eingehen, fassen wir zuächst noch einmal zusammen, aus welchen in früheren Kapiteln beschriebenen Bestandteilen sich dieses Gleichungssystem zusammensetzt.

In Abschnitt 3.4 haben wir die allgemeinen Bewegungsgleichungen aufgestellt, gemäß denen sich ein simuliertes Objekt unter dem Einfluß externer Kräfte während eines Zeitintervalls $[t, t+\Delta t]$ bewegt. Man erhält dabei ein Gleichungssystem für die Position $\mathbf{c}^{t+\Delta t}(\mathbf{f})$ und die Orientierung $\mathbf{R}^{t+\Delta t}(\mathbf{f})$ des Körpers am Ende des Zeitintervalles Δt als Funktionen der unbekannten Kontaktkräfte $\mathbf{f} = (f_1 \dots f_n)^T$. Die Komponenten von $\mathbf{f}$ stellen dabei die Beträge der n Kontaktkräfte dar, während sich die Kraftrichtungen aus der zum Zeitpunkt t bekannten Kontaktgeometrie ergeben.

In Abschnitt 4.5 wurde der Kontaktabstand w an einem Punktkontakt als Funktion der Raumlagen der beteiligten Körper hergeleitet (Gln. (28),(32)), der sich allgemein darstellen läßt als $w^{t+\Delta t} = g(\mathbf{R}_1(\mathbf{f}), \mathbf{c}_1(\mathbf{f}), \mathbf{R}_2(\mathbf{f}), \mathbf{c}_2(\mathbf{f}))$, oder kurz $w^{t+\Delta t} = g(\mathbf{f})$. Aufgrund des Rotationsanteils der Bewegungen ist die Funktion g nichtlinear in den Kraftkomponenten $\mathbf{f}$.

Für ein System von Körpern, welche sich an insgesamt n Kontaktpunkten berühren, erhält man somit n Kontaktbedingungen der Form

$$w_i^{t+\Delta t} = g_i(\mathbf{f}) = 0, \qquad i = 1 \dots n,$$

oder, kompakt zusammengefaßt, $\mathbf{w}^{t+\Delta t} = \mathbf{g}(\mathbf{f}) = \mathbf{0}, \mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$.

Da wir die Kontakte als unilaterale Bewegungsbeschränkungen betrachten, gelten an den n Kontaktpunkten nun die in Abschnitt 5.3.1 beschriebenen Komplementaritätsbedingungen

$$w_i \geq 0, \qquad f_i \geq 0, \qquad w_i f_i = 0, \qquad i = 1 \dots n,$$

beziehungsweise in der kompakten Vektorschreibweise

$$\mathbf{w} \geq \mathbf{0}, \qquad \mathbf{f} \geq \mathbf{0}, \qquad \mathbf{w}^T\mathbf{f} = 0. \tag{46}$$

Die Ungleichungen sind dabei komponentenweise zu verstehen.

Wegen der nichtlinearen Beziehung $\mathbf{w}^{t+\Delta t} = \mathbf{g}(\mathbf{f})$ stellt Gleichung (46) ein sogenanntes nichtlineares Komplementaritätsproblem (nonlinear complementarity problem, NCP) zur Bestimmung der Größen $\mathbf{w}^{t+\Delta t}$ und $\mathbf{f}$ dar. Die aus der Lösung dieses NCP resultierenden Kontaktkräfte werden schließlich wieder in die Bewegungsgleichungen eingesetzt, woraus sich die Bewegungsgrößen und Objektlagen der beteiligten Objekte am Ende des betrachteten Zeitintervalls $t + \Delta t$ ergeben.

6.1 Die Lösung des nichtlinearen Komplementaritätsproblems

Ein nichtlineares Komplementaritätsproblem (NCP) der Form

$$\mathbf{x} \geq \mathbf{0}, \quad \mathbf{F}(\mathbf{x}) \geq \mathbf{0}, \quad \mathbf{x}^T\mathbf{F}(\mathbf{x}) = 0$$

kann mit Hilfe eines Operators $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ auf ein nichtlineares Gleichungssystem zurückgeführt werden, wenn gilt

$$\mathbf{\Psi}(\mathbf{x}^*) = \mathbf{0} \qquad \Longleftrightarrow \qquad \mathbf{x}^* \geq \mathbf{0}, \quad \mathbf{F}(\mathbf{x}^*) \geq \mathbf{0}, \quad \mathbf{x}^{*T}\mathbf{F}(\mathbf{x}^*) = 0.$$

Da $\mathbf{\Psi}(\mathbf{x}^*) = \mathbf{0}$ ein gewöhnliches nichtlineares Gleichungssystem darstellt, kann es mit dem Newton-Verfahren oder verwandten Ansätzen gelöst werden. Diese Vorgehensweise zur Behandlung eines NCP geht auf Mangasarian [31] zurück. Seitdem wurden von verschiedenen Autoren Varianten zur Wahl von $\mathbf{\Psi}$ vorgeschlagen. Eine gute Übersichtsdarstellung findet sich in [25, 26].

NCP-Funktionen

Nach Kanzow und Kleinmichel [23] ist die Klasse der NCP-Funktionen $\varphi(a,b) : \mathbb{R}^2 \to \mathbb{R}$ definiert durch die Eigenschaft

$$\varphi(a,b) = 0 \qquad \Longleftrightarrow \qquad a \geq 0, \quad b \geq 0, \quad ab = 0. \tag{47}$$

Diese Funktionen transformieren ein 1-dimensionales Komplementaritätsproblem in eine äquivalente nichtlineare skalare Gleichung, welche beispielsweise mit dem Newton-Verfahren gelöst werden kann.

Besonders günstige Eigenschaften bei der Lösung dieser nichtlinearen Gleichung ergeben sich bei Verwendung der folgenden NCP-Funktion, die auch als Fischer-Funktion [19] bezeichnet wird:

$$\varphi(a,b) = \sqrt{a^2 + b^2} - a - b. \tag{48}$$

Die Eigenschaft (47) dieser Funktion läßt sich leicht verifizieren anhand der 9 Fälle, die sich durch die Kombination von $a > 0, \quad a = 0, \quad a < 0$ mit $b > 0, \quad b = 0, \quad b < 0$ ergeben.

Lösung des Gleichungssystems mit dem Newton-Verfahren

Mit dieser Hilfsfunktion definiert man nun den Operator $\mathbf{F}_\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ wie folgt:

$$\mathbf{F}_\varphi(\mathbf{f},\mathbf{w}) = \begin{pmatrix} \mathbf{g}(\mathbf{f}) - \mathbf{w} \\ \mathbf{\Phi}(\mathbf{f},\mathbf{w}) \end{pmatrix}, \tag{49}$$

wobei

$$\mathbf{\Phi}(\mathbf{f},\mathbf{w}) = \left(\varphi(f_1,w_1),\ldots,\varphi(f_n,w_n)\right)^T.$$

Die Werte $(\mathbf{f}^*,\mathbf{w}^*)$ lösen also genau dann das obige NCP, wenn gilt $\mathbf{F}_\varphi(\mathbf{f}^*,\mathbf{w}^*) = \mathbf{0}$. Das Gleichungssystem $\mathbf{F}_\varphi(\mathbf{f},\mathbf{w}) = \mathbf{0}$ ist vom Grad $2n$. Wie in [25] dargestellt läßt sich die Lösung dieses Systems jedoch auf die Lösung eines Gleichungssystems vom Grade n reduzieren, da die n Komponenten von $\mathbf{w}$ letztlich keine unabhängigen Variablen darstellen.

Mit der Jakobi-Matrix

$$\frac{\partial \mathbf{F}_\varphi}{\partial(\mathbf{f},\mathbf{w})}(\mathbf{f}^k,\mathbf{w}^k) = \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \mathbf{f}}(\mathbf{f}^k) & -\mathbf{E}_k \\ \frac{\partial \mathbf{\Phi}}{\partial \mathbf{f}}(\mathbf{f}^k,\mathbf{w}^k) & \frac{\partial \mathbf{\Phi}}{\partial \mathbf{w}}(\mathbf{f}^k,\mathbf{w}^k) \end{bmatrix}$$

läßt sich der Newton-Ansatz

$$\frac{\partial \mathbf{F}_\varphi}{\partial(\mathbf{f},\mathbf{w})}(\mathbf{f}^k,\mathbf{w}^k) \begin{pmatrix} \Delta\mathbf{f}^k \\ \Delta\mathbf{w}^k \end{pmatrix} = -\mathbf{F}_\varphi(\mathbf{f}^k,\mathbf{w}^k)$$

umformen in das System

$$\left(\frac{\partial \mathbf{\Phi}}{\partial \mathbf{f}}(\mathbf{f}^k,\mathbf{w}^k) + \frac{\partial \mathbf{\Phi}}{\partial \mathbf{w}}(\mathbf{f}^k,\mathbf{w}^k)\frac{\partial \mathbf{g}}{\partial \mathbf{f}}(\mathbf{f}^k)\right)\Delta\mathbf{f}^k = -\mathbf{\Phi}(\mathbf{f}^k,\mathbf{w}^k) - \frac{\partial \mathbf{\Phi}}{\partial \mathbf{w}}(\mathbf{f}^k,\mathbf{w}^k)\left(\mathbf{g}(\mathbf{f}^k) - \mathbf{w}^k\right) \tag{50}$$

$$\Delta\mathbf{w}^k = \frac{\partial \mathbf{g}}{\partial \mathbf{f}}(\mathbf{f}^k)\Delta\mathbf{f}^k + \mathbf{g}(\mathbf{f}^k) - \mathbf{w}^k. \tag{51}$$

Zur Bestimmung von $\Delta\mathbf{f}^k$ aus (50) ist also ein lineares Gleichungssystem vom Grade n zu lösen, während sich $\Delta\mathbf{w}^k$ im Anschluß daran durch direktes Einsetzen von $\Delta\mathbf{f}^k$ in Gleichung (51) ergibt.

Die Iteration

$$\begin{aligned} \mathbf{f}^{k+1} &= \mathbf{f}^k + \Delta\mathbf{f}^k \\ \mathbf{w}^{k+1} &= \mathbf{w}^k + \Delta\mathbf{w}^k \end{aligned}$$

wird schließlich solange fortgesetzt, bis $\|\mathbf{\Phi}(\mathbf{f}^k,\mathbf{w}^k)\|$ einen vorgebbaren Schwellwert unterschreitet.

Bemerkungen zur Implementierung

Die Jacobi-Matrix $\frac{\partial \mathbf{F}_\varphi}{\partial(\mathbf{f},\mathbf{w})}$ wird in der vorliegenden Implementierung nicht analytisch bestimmt, sondern mit finiten Differenzen numerisch approximiert (siehe [38]). Der Differentialquotient $\frac{\partial F_i(x_j)}{\partial x_j}$ an der Stelle $x_j = \xi$ wird dabei durch den Differenzenquotienten $\frac{F_i(\xi+h)-F_i(\xi)}{h}$ angenähert, wobei h einen sehr kleinen Wert annimmt.

Die zur Durchführung eines Iterationsschrittes erforderliche Lösung eines linearen Gleichungssystems wurde effizient mit Hilfe von Standardbibliotheken (siehe [38]) realisiert.

Zur Gewinnung des Abbruchkriteriums wurde anstelle der Norm

$$\|\mathbf{\Phi}(\mathbf{f}^k,\mathbf{w}^k)\| = \sqrt{\sum_{i=1}^{n} \varphi^2(f_i^k, w_i^k)}$$

die Summe der Beträge der Komponenten von $\mathbf{\Phi}(\mathbf{f}^k,\mathbf{w}^k)$ bewertet. Die Iteration wird abgebrochen, wenn die Bedingung

$$\sum_{i=1}^{n} |\varphi(f_i^k, w_i^k)| < \varepsilon$$

erfüllt ist. Als Wert für ε wurde 10^{-6} gewählt.

6.2 Laufzeitverhalten

Der numerische Aufwand zur Durchführung eines Iterationsschritts hängt im wesentlichen vom Grad n des zu lösenden linearen Gleichungssystem ab, und somit von der Anzahl n der Kontaktpunkte in einem Kontaktcluster. Die zur Lösung des gesamten NCP erforderliche Laufzeit hängt außer von n auch ganz wesentlich ab von der Wahl der Startwerte für die Iteration, sowie von der konkreten Konfiguration der Kontaktpunkte. Letztere entscheidet über die tatsächliche Konvergenzgeschwindigkeit, und somit auch über die erforderliche Anzahl von Iterationsschritten. Dieser Einfluß der jeweiligen Kontaktsituation läßt sich nur schwer analytisch erfassen, und soll daher empirisch untersucht werden.

Dazu wurde eine spezielle Versuchsanordnung gewählt, bei der die Bewegung mehrerer sich berührender Objekte stochastisch gesteuert wurde. Konkret wurden zwei polygonale Körper mit zufällig bestimmten externen Kräften und Drehmomenten beaufschlagt. Der Erwartungswert der Kräfte war dabei jedoch ungleich Null, so daß dadurch eine bevorzugte Bewegungsrichtung erzeugt wurde. Die Bewegungstendenz wurde so gewählt, daß sich alle Objekte in Richtung einer konkaven Hindernisecke bewegten und dort gegeneinanderstießen (siehe Abb. 38). Durch die zufällige Variation der Kräfte und Drehmomente stellte sich kein stationärer Zustand ein. Stattdessen ergab sich der für unsere Fragestellung angestrebte ständige Wechsel der Kontaktsituationen.

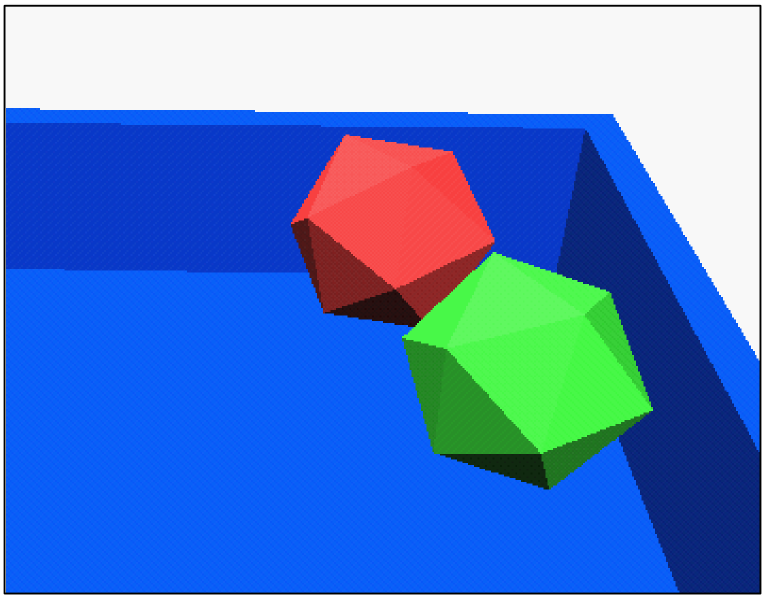


Abbildung 38: Versuchsanordnung zur Bestimmung des Laufzeitverhaltens.

Die zur Aufstellung und Lösung des NCP während einer Meßreihe benötigten Rechenzeiten [8] wurden zusammen mit der jeweiligen Anzahl von Kontaktpunkten registriert, und sind in Abb. 39 dargestellt.

Man erkennt, daß die Laufzeiten für eine gegebene Anzahl von Kontaktpunkten stark schwanken, dies spiegelt die unterschiedliche Konvergenzrate bei verschiedenen Kontakt-

[8] auf einer SGI O^2 mit Prozessor R10000 bei 150 MHz

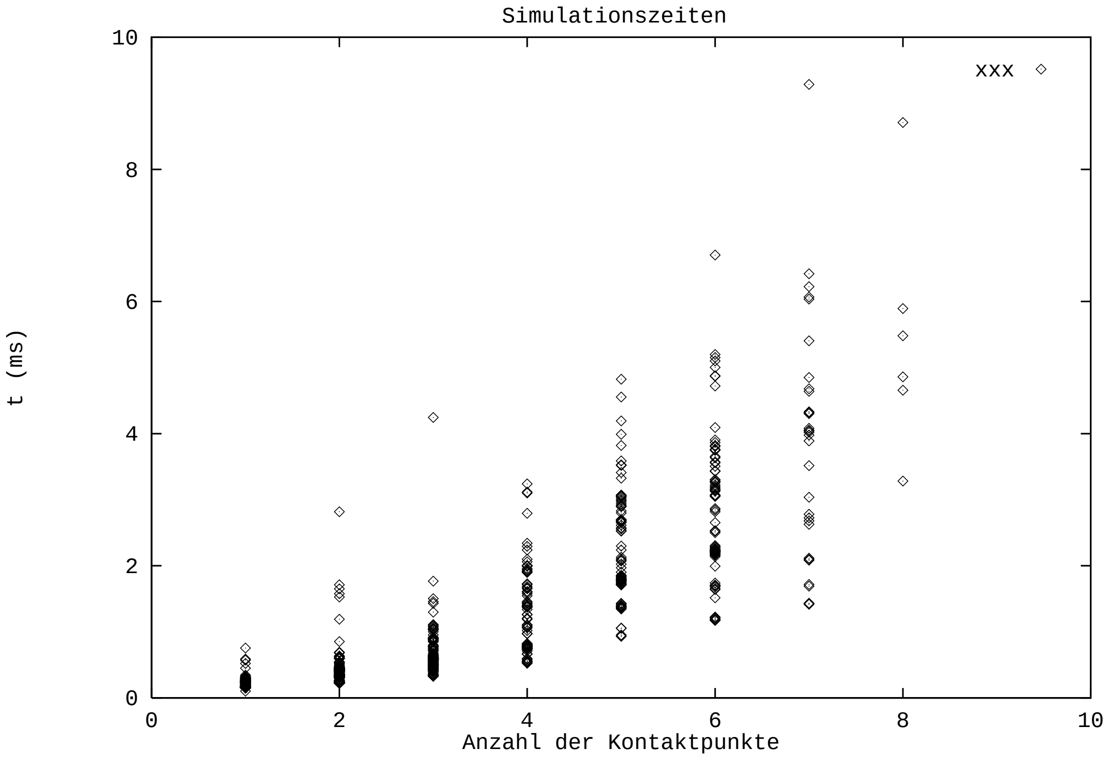


Abbildung 39: Darstellung der zur Lösung des NCP bestimmten Laufzeiten, in Abhängigkeit von der Ordnung des Gleichungssystems, die mit der Anzahl der Kontaktpunkte übereinstimmt.

situationen wider. Die mittleren Laufzeiten liegen bis $n = 8$ unter 6ms, die Maximalwerte unter 10ms. Diese Werte liegen somit in einem Bereich, der Echtzeitanwendungen zuläßt.

6.3 Gewinnung von Startwerten für die Newton-Iteration

Das Newton-Verfahren weist in der Nähe einer Lösung quadratische Konvergenz auf. Die Anzahl der erforderlichen Iterationen ist daher um so kleiner, je näher der Startwert an einer Lösung des Gleichungssystems liegt. Für die eben beschriebene Newton-Iteration zur Bestimmung der Kontaktkräfte sollen nun verschiedene Möglichkeiten zur Gewinnung passender Startwerte diskutiert werden.

1. **Initialisierung mit Null**
   Wenn keine qualifizierte Prädiktion der zu erwartenden Kontaktkräfte vorliegt, bietet es sich an, die Newton-Iteration mit Null zu initialisieren.

2. **Initialisierung mit den zuletzt bestimmten Kontaktkräften**
   Dieser Vorgehensweise liegt die Annahme zugrunde, daß sich die Kräfteverhältnisse zwischen aufeinanderfolgenden Simulationsintervallen nur unwesentlich ändern. Diese Annahme ist plausibel, solange sich die Konfiguration der Kontaktpunkte nicht wesentlich (beispielsweise durch eine Kollision) ändert.

3. **Initialisierung mit linearer Approximation**
   Unter der idealisierenden Annahme, daß alle bestehenden Kontaktpunkte während des aktuellen Simulationsintervalls erhalten bleiben, kann Gleichung (42) zur Gewinnung einer Schätzung der Kontaktkräfte verwendet werden. Wegen der Kontaktbedingung $\mathbf{a}_\perp = \mathbf{0}$ führt dies auf ein lineares Gleichungssystem der Form

   $$\mathbf{C}^T\mathbf{M}^{-1}\mathbf{C}\mathbf{f} + \mathbf{C}^T\mathbf{a}_{ext} = \mathbf{0}$$

   zur näherungsweisen Bestimmung von $\mathbf{f}$.

4. **Initialisierung mit LCP**
   Gemäß Gleichung (41) kann die Bestimmung der Kontaktkräfte auch als lineares Komplementaritätsproblem (LCP) formuliert werden, wenn dies auch mit gewissen Nachteilen verbunden ist (vgl. Abschnitt 5.3.1). Die Lösung dieses LCP liefert ebenfalls Kontaktkräfte, die zur Initialisierung der Newton-Iteration herangezogen werden können.

Zur Bewertung und zum Vergleich dieser Varianten ziehen wir die Auswertung einer großen Anzahl von Simulationsschritten heran, die unter zufällig gewählten Randbedingungen mit verschiedenen Kontaktkonfigurationen und externen Kräften durchgeführt wurden. Dabei wurden sowohl die jeweils erforderliche Anzahl von Iterationsschritten als auch die zur Lösung des NCP erforderliche Rechenzeit festgehalten. Die Ergebnisse sind in den Diagrammen Abb. 40 und 41 jeweils in Abhängigkeit von der Ordnung des Gleichungssystems dargestellt.

Der Abb. 40 entnehmen wir wie erwartet, daß die benötigte Anzahl der Iterationen um so kleiner wird, je näher die initialen Kraftwerte an der Lösung des Gleichungssystems liegen. Die Vermutung, daß die Kraftwerte des jeweils vorherigen Simulationsintervalls eine bessere Prädiktion darstellen als die Initialisierung mit Null, bestätigt sich. Allerdings wird die Prädiktion noch besser, wenn die Verhältnisse des jeweils aktuellen Simulationsintervalls mitberücksichtigt werden, und in Form einer linearen Approximation oder durch die Lösung eines LCP zur Bestimmung der Startwerte herangezogen werden.

Bei einem Vergleich der betrachteten Verfahren zur Startwertgewinnung ist allerdings nicht nur der mögliche Zeitgewinn bei der Newton-Iteration selbst von Bedeutung, sondern auch der zur Gewinnung dieser Startwerte zusätzlich erforderliche Zeitaufwand. Insbesondere bei den beiden letzten Verfahren kann dieser Zeitaufwand nicht mehr vernachlässigt werden. In Abb. 41 sind daher die im Rahmen der Stichprobe im Mittel benötigten Rechenzeiten zur iterativen Lösung des NCP zuzüglich des Zeitaufwandes zur jeweiligen Gewinnung der Startwerte aufgetragen. Dabei fällt auf, daß die mittleren Rechenzeiten sich bei den betrachteten Prädiktionsverfahren nur wenig unterscheiden. Der größere Rechenzeitaufwand der Prädiktion durch lineare Approximation oder durch Lösung eines LCP bringt insgesamt betrachtet keinen signifikanten Vorteil gegenüber der Verwendung der Kräfte des vergangenen Simulationsintervalls.

6.4 Bemerkungen zur Konvergenz des NCP

Damit das Newton-Verfahren bei der Lösung des nichtlinearen Gleichungssystems $\mathbf{F}_\varphi(\mathbf{f}) = \mathbf{0}$ konvergiert, darf die Jacobi-Matrix von $\mathbf{F}_\varphi(\mathbf{f})$ an einer Lösungsstelle $\mathbf{f} = \mathbf{f}^*$ des NCP

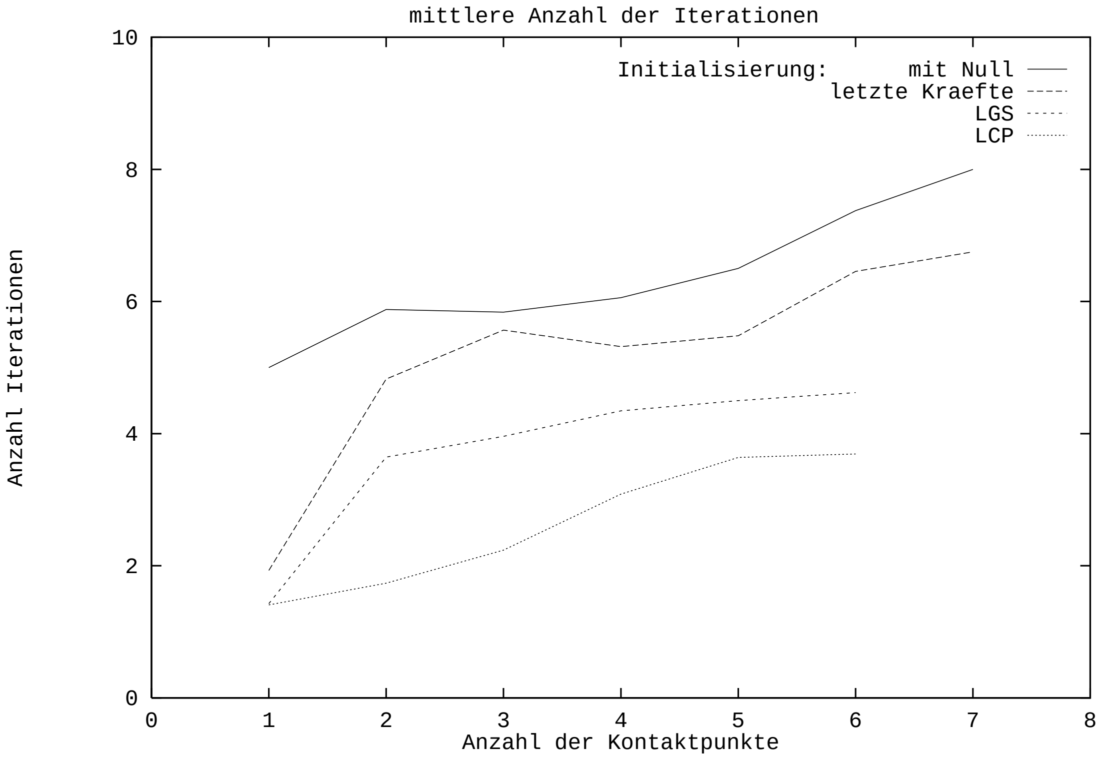


Abbildung 40: Darstellung der zur Lösung des NCP erforderlichen Anzahl von Iterationen, welche bei einer Simulationsstichprobe ermittelt wurden. Die Werte sind in Abhängigkeit von der Ordnung des Gleichungssystems dargestellt.

nicht singulär sein. Kanzow [24] zeigt, daß diese Bedingung unter folgenden Voraussetzung erfüllt ist:

1. $\mathbf{g}(\mathbf{f})$ ist in einer Umgebung von $\mathbf{f}^*$ differenzierbar, und die Ableitung ist dort Lipschitz-stetig.

2. Die Jacobi-Matrix von $\mathbf{g}(\mathbf{f})$ ist an der Stelle $\mathbf{f}^*$ nicht singulär.

3. Die strikte Komplementaritätsbedingung $\mathbf{f} + \mathbf{w} > \mathbf{0}$ ist erfüllt, d.h. an einem bestehenden Kontaktpunkt verschwindet die Kontaktkraft nicht.

Voraussetzung 1 ist erfüllt, wie sich anhand der Herleitung von $\mathbf{g}(\mathbf{f})$ in Abschnitt 3.4 leicht nachprüfen läßt. Dies ist auch physikalisch plausibel, da bei massebehafteten Körpern eine endliche Änderung des Kontaktabstandes nur durch eine nicht-verschwindende Änderung der Kontaktkräfte erzielt werden kann. Die Voraussetzungen 2 und 3 sind erfüllt, wenn die Kontaktbedingungen nicht redundant sind. Die in Abschnitt 5 dargestellten Mechanismen zur Verwaltung der Kontaktpunkte vermeiden redundante Kontaktpunktkonfigurationen. Daher sollten die Kontaktkräfte in der Regel eindeutig bestimmbar sein. Die vorgestellten empirischen Ergebnisse zeigen die rasche Konvergenz des Verfahrens und bestätigen dies ebenfalls.

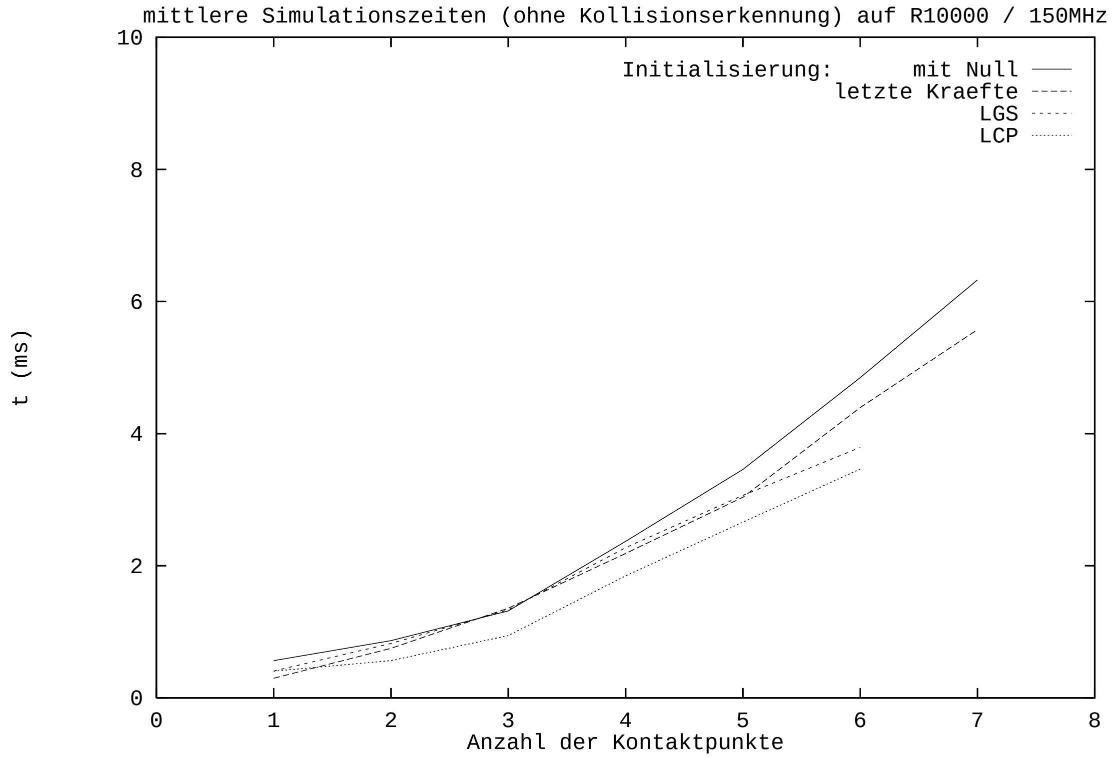


Abbildung 41: Darstellung der zur Lösung des NCP benötigten Laufzeit, welche bei einer Simulationsstichprobe ermittelt wurde. Die Werte sind in Abhängigkeit von der Ordnung des Gleichungssystems dargestellt.

Dennoch sind Situationen möglich, in denen das Verfahren nicht oder nur sehr langsam konvergiert. Solche Fälle können auftreten, wenn der Startwert der Iteration zu weit von der Lösung entfernt ist.

Ein Beispiel hierfür ist in Abb. 42 dargestellt, bei dem sich Objekt A in eine keilförmige Öffnung von Objekt B hineinbewegt. Dabei gleite Objekt A entlang der Fläche F_2 von Objekt B, unter Einhaltung des Sicherheitsabstandes ε_s. Aufgrund des kleinen Winkels zwischen den Flächen F_1 und F_2 dringt Objekt A bereits sehr viel früher in das Sicherheitsintervall der Fläche F_1 ein, bevor eine Kollision an dieser Stelle auftreten kann. Nach einer solchen Kollision zwischen Eckpunkt P und Fläche F_1 (Abb. 42b) fordert die neu entstandene Kontaktbedingung die Einhaltung des Sicherheitsabstandes ε_s auch an diesem Kontakt. Dies kann nur dadurch erreicht werden, daß Objekt A um die Strecke h nach links bewegt. Bei einem sehr kleinen Winkel zwischen F_1 und F_2 kann h sehr groß werden, und Objekt A müßte sichtbar nach links springen, was allerdings keinem physikalisch plausiblen Objektverhalten entspräche.

Tatsächlich konvergiert das iterative Lösungsverfahren zur Bestimmung der Kontaktkräfte in solchen Fällen nur sehr langsam oder gar nicht. Nach einer vorgegebenen Maximalzahl von Iterationen wird das Verfahren daher abgebrochen, und die beteiligten Objekte werden in die vorhergehende Situation zurückversetzt (Abb. 42a). Diese Vorgehensweise führt zu einem physikalisch plausiblen Objektverhalten, welches sich im dargestellten Beispiel als Verklemmung umschreiben läßt.

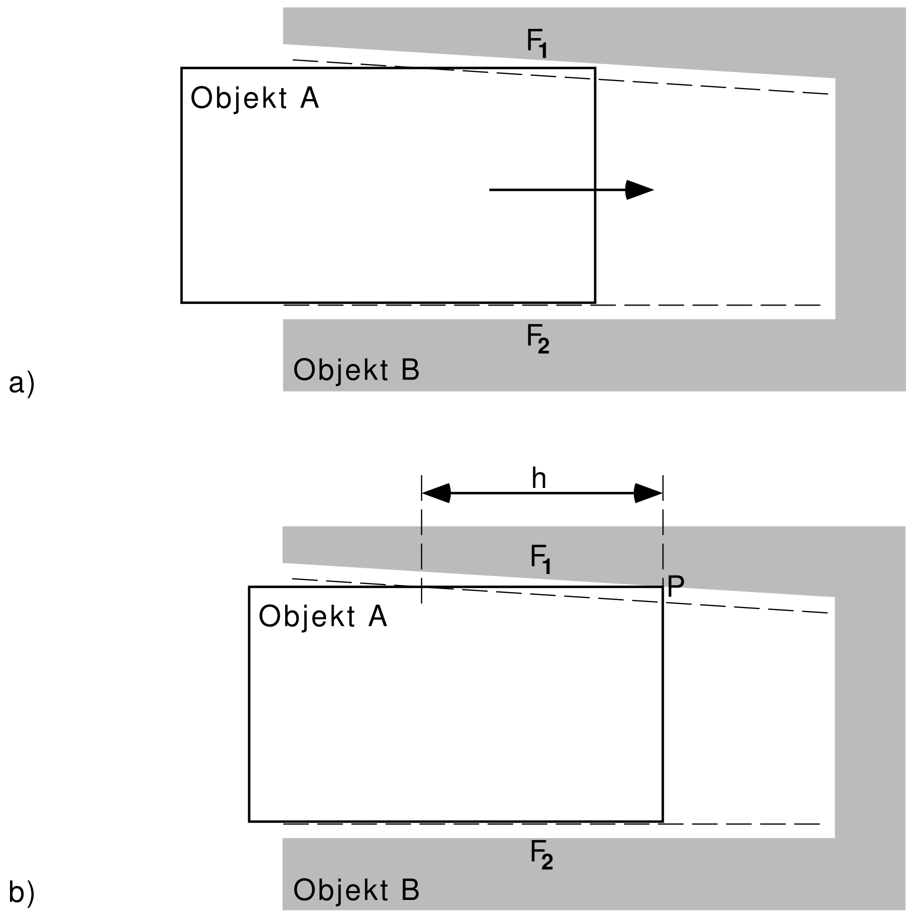


Abbildung 42: Das Objekt A dringt in eine keilförmige Öffnung von Objekt B ein (a), und kollidiert schließlich mit diesem (b). Wenn das Verfahren zur Bestimmung der Kontaktkräfte in dieser Situation nicht konvergiert, so kehrt die Simulation zurück zum letzten konsistenten Zustand (a).

7 Reibungseffekte

Die bei der Simulation von Kontakten zwischen virtuellen Objekten auftretenden Kontaktkräfte beschränkten sich in der bisherigen Darstellung ausschließlich auf die in Richtung der jeweiligen Kontaktnormalen wirkenden Kräfte. Übersetzt in physikalische Begriffe wurde damit lediglich der reibungsfreie Fall behandelt. In praktischen Anwendungsszenarien sind jedoch reibungsbehaftete Kontakte die Regel und müssen daher im Rahmen einer realistischen Kontaktsimulation berücksichtigt werden. Hierzu müssen die an den Kontaktstellen wirkenden Kräfte um tangentiale Komponenten ergänzt werden.

In den folgenden Abschnitten wird die Erweiterung der Kontaktsimulation um Reibungseffekte beschrieben. Dabei werden die beiden Reibungsarten der Gleitreibung und der Haftreibung ihrer unterschiedlichen Eigenschaften wegen getrennt behandelt.

7.1 Simulation von Gleitreibung

Die physikalische Beschreibung der Gleitreibung zwischen festen Körpern wird häufig in drei zentrale empirische Gesetzmäßigkeiten gefaßt, die auch als Coulomb'sches Reibungsgesetz bezeichnet werden:

- Die Gleitreibungskraft $\mathbf{f}_R$ ist der tangentialen Relativbewegung $\mathbf{v}_T$ der beteiligten Körper stets entgegengerichtet. Ihr Betrag ist proportional zum Betrag f_N der zugehörigen Normalkraft.
- Die Gleitreibungskraft ist unabhängig von der Fläche der Kontaktregion.
- Die Gleitreibungskraft ist unabhängig vom Betrag der Gleitgeschwindigkeit, solange die Bewegung nicht zur Ruhe kommt.

Die entsprechende mathematische Formulierung dieses Reibungsgesetzes lautet also

$$\mathbf{f}_R = -\mu f_N \frac{\mathbf{v}_T}{|\mathbf{v}_T|}, \qquad (52)$$

wobei der Reibungskoeffizient μ die Proportionalitätskonstante zwischen Reibungskraft und Normalkraft bezeichnet.

Diese Darstellung ist zunächst nur für punktförmige Kontaktregionen vollständig. Flächenhafte Kontakte lassen sich jedoch als Menge vieler Punktkontakte auffassen, für die jeweils lokal das obige Reibungsgesetz gilt.

Da bei der Simulation von Kontaktsituationen auch rotatorische Gleitbewegungen auftreten können, ist die lokale Tangentialgeschwindigkeit an den verschiedenen Reibungspunkten als ortsabhängig verschieden anzunehmen. Zudem muß berücksichtigt werden, daß zwischen den beteiligten Körpern nicht nur Kontaktkräfte, sondern auch Momente ausgetauscht werden können, so daß die lokalen Normalkräfte ebenfalls ortsabhängig sind.

Die explizite Berücksichtigung einer großen Anzahl ortsabhängig verschiedener lokaler Kontaktkräfte innerhalb einer Kontaktfläche im Rahmen der bisher beschriebenen Vorgehensweise zur Kontaktsimulation erscheint aus verschiedenen Gründen fragwürdig:

- Das verwendete Kontaktmodell beruht auf einer minimalen Anzahl von Punktkontakten. Nur für diese Kontaktpunkte werden Kontaktkräfte bestimmt.
- Selbst wenn eine große Anzahl von Kontaktkräften innerhalb einer Kontaktfläche realisiert werden könnte, wäre das resultierende Kräftesystem unter Voraussetzung starrer Körper statisch und dynamisch unterbestimmt und somit redundant.
- Würden genügend viele Randbedingungen zur Verfügung stehen, wäre das resultierende Gleichungssystem sehr umfangreich und entsprechend aufwendig zu lösen.

Wir behalten daher die Strategie bei, redundante Kontaktpunkte zu vermeiden, und beschränken uns bei der Modellierung von Reibungseffekten auf diesselben wenigen Punktkontakte, welche zur Befriedigung der kinematischen Kontaktbedingungen erforderlich sind.

In welchem Maße durch diese Vorgehensweise das Reibungsverhalten der Kontaktsimulation von den realen Verhältnissen abweicht, wird gesondert in Abschnitt 8.3 untersucht.

Im folgenden beschreiben wir, wie das Coulomb'sche Reibungsgesetz in die bisher beschriebene Vorgehensweise zur Kontaktsimulation eingearbeitet werden kann. Im reibungsfreien Fall wurde die Richtung jeder Kontaktkraft $\mathbf{f}_i$ durch die Kontaktnormale $\mathbf{n}_i$ vorgegeben. Mit Hilfe der kinematischen Kontaktbedingungen wurden jeweils die Beträge f_i der Kontaktkräfte bestimmt. Nun kommen zu den Normalkräften $\mathbf{f}_{N_i}$ additiv die Reibungskräfte $\mathbf{f}_{R_i}$ hinzu (vgl. Abb.43):

$$\begin{aligned}\mathbf{f}_i &= \mathbf{f}_{N_i} + \mathbf{f}_{R_i} \\ &= f_i \mathbf{n}_i - \mu f_i \frac{\mathbf{v}_i}{|\mathbf{v}_i|} \\ &= f_i \left(\mathbf{n}_i - \mu \frac{\mathbf{v}_i}{|\mathbf{v}_i|} \right).\end{aligned}$$

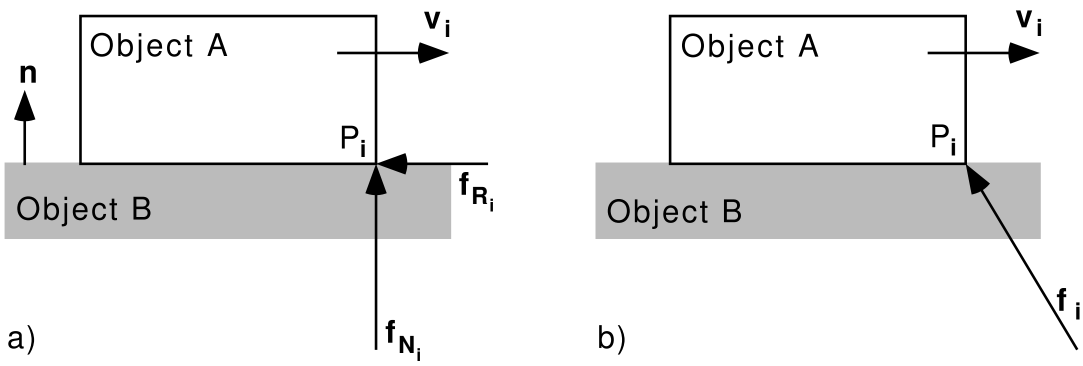


Abbildung 43: Die Normalkraft $\mathbf{f}_{N_i}$ und die tangentiale Reibungskraft $\mathbf{f}_{R_i}$ können durch eine geneigte resultierende Reibungskraft $\mathbf{f}_i$ ersetzt werden.

Bei der Bestimmung der Beträge der Kontaktkräfte kann nun in den Gleichungen (21) und (22) jeweils die resultierende Kraftrichtung $\mathbf{n}_i - \mu \frac{\mathbf{v}_i}{|\mathbf{v}_i|}$ anstelle der Kontaktnormale $\mathbf{n}_i$ eingesetzt werden.

Bestimmung der Richtung der Gleitreibungskraft

Gemäß dem Coulomb'schen Reibungsgesetz ist die Gleitreibungskraft stets der relativen Kontaktgeschwindigkeit entgegengerichtet. Wie bereits in Abschnitt 3.4 soll die Kontaktgeometrie zu Beginn des betrachteten Simulationsintervalls zugrunde gelegt werden. Diese Vorgehensweise entspricht der in den Gleichungen (21) und (22) verwendeten Vorwärtsdifferentiation. Daher bestimmen wir auch nun die relative Kontaktgeschwindigkeit $\mathbf{v}_i^t$ des i-ten Kontaktes am Anfang des Zeitintervalles zu

$$\mathbf{v}_i^t \quad = \quad \mathbf{v}_1^t + \boldsymbol{\omega}_1^t \times \mathbf{r}_{i1}^t - \mathbf{v}_2^t - \boldsymbol{\omega}_2^t \times \mathbf{r}_{i2}^t.$$

7.2 Simulation von Haftreibung

Tritt an einer Kontaktstelle zwischen zwei starren Körpern keine tangentiale Gleitbewegung auf, haften sie also aneinander, so muß eine orthogonal zur Kontaktnormale wirkende Widerstandskraft überwunden werden, bevor eine Gleitbewegung einsetzen kann. Diese Widerstandskraft wird als Haftreibungskraft bezeichnet. Im Gegensatz zur Gleitreibungskraft ist sie nicht proportional zur Normalkraft an der Berührstelle, sondern gerade so groß, daß eine Gleitbewegung verhindert wird, bis ein Grenzwert, die maximale Haftreibungskraft, überschritten wird. Dieser Grenzwert allerdings ist proportional zur Normalkraft, so daß insgesamt gilt $f_{haft} \leq \mu_{haft} f_N$, wobei typischerweise $\mu_{gleit} < \mu_{haft}$.

Da die Haftreibungskraft in keinem festen Verhältnis zur Normalkraft steht, sind die beiden tangentialen Komponenten dieser Kraft als zusätzliche variable Kräfte zu behandeln. Zur vollständigen Bestimmung der Kontaktkräfte bei Haftreibung sind also an jedem Kontaktpunkt neben dem Betrag der Normalkraft zusätzlich die zwei Komponenten der Reibungskraft zu bestimmen. Es ist leicht einzusehen, daß die Haftreibungskräfte bei mehreren Kontaktpunkten zwischen zwei als starr angenommenen Körpern nicht eindeutig bestimmbar sind, selbst wenn die Normalkräfte eindeutig bestimmt sind.

Bestehen zwischen zwei Körpern beispielsweise drei nicht-redundante Kontaktpunkte, so verbleiben für eine Gleitbewegung noch $6 - 3 = 3$ Freiheitsgrade. Sollen diese durch Haftreibungskräfte blockiert werden, so sind genau drei nicht-redundante zusätzliche Kräfte eindeutig bestimmbar. An jedem der drei Kontaktpunkte wirken jedoch zwei tangentiale Kraftkomponenten, insgesamt also 6 Stück.

Zur Lösung solch unterbestimmter Problemstellungen stehen verschiedene mathematische Methoden bereit, wie zum Beispiel Singular Value Decomposition (SVD, [38]). Um den zusätzlichen numerischen Aufwand solcher Verfahren zu vermeiden, ziehen wir jedoch eine Vorgehensweise vor, welche das Auftreten redundanter Kontaktkräfte verhindert und eine eindeutige Lösung ermöglicht. Zudem treffen die in der Literatur ([1, 35, 45] u.a.) vorgestellten Verfahren meist die vereinfachende und nicht zutreffende Annahme $\mu_{gleit} = \mu_{haft}$. Der im folgenden dargestellte Ansatz zur Realisierung von Haftreibung bedarf dieser Vereinfachung nicht.

7.2.1 Haftreibung als kinematische Zwangsbedingung

Im Gegensatz zur Gleitreibung, die einer Tangentialbewegung zwar entgegenwirkt, sie aber nicht verhindert, blockiert die Haftreibung jede tangentiale Relativbewegung, solange die maximale Haftreibungskraft nicht überschritten wird. Die Haftreibung erzwingt also die Einhaltung einer kinematischen Zwangsbedingung, und ähnelt darin der in Normalenrichtung wirkenden Kontaktkraft, welche die Durchdringung der sich berührenden Körper an der Kontaktstelle verhindert.

Dieser Vergleich ist durchaus auch physikalisch motiviert: man kann sich die Haftreibung als durch mikroskopisch kleine Rauhigkeiten an den Berührflächen verursacht vorstellen, an denen die beiden Körper sich verhaken. Die Haftreibungskraft stellt sich damit als Kontaktkraft zur Befriedigung von zusätzlichen Kontaktbedingungen dar.

Gemäß dieser Sichtweise fügt sich die Realisierung von Haftreibung als kinematische Zwangsbedingung nahtlos in das bereits beschriebene Konzept zur Kontaktsimulation ein. Erforderlich ist hierzu lediglich die Einführung von geeigneten Gleithindernissen an den Kontaktpunkten, die für die zusätzlichen tangentialen Kontaktkräfte sorgen.

7.2.2 Realisierung der Haftreibung an Punkt-Fläche-Kontakten

Ein Punkt-Fläche-Kontakt sei gegeben durch den zu Objekt 1 gehörenden Punkt P, der die zu Objekt 2 gehörende Fläche F im Punkt B berührt. Geht dieser Kontakt in Haftreibung über (vgl. Abschnitt 7.3), so werden an Objekt 2 zusätzliche Hilfsflächen definiert, welche jeweils orthogonal zu F stehen und in sehr geringem Abstand rings um B angeordnet sind (Abb. 44). Diese Hilfsflächen sind nur für den Punkt P kollisionssensitiv.

Bewegen sich die beteiligten Objekte tangential zueinander, so kollidiert P sofort mit einer dieser Hilfsflächen, und es entsteht eine neue Kontaktbedingung, welche eine weitere Gleitbewegung von P relativ zu F unterbindet.

Ein solcher Hilfskontakt wird in der Simulation als unilateraler, reibungsfreier Punktkontakt realisiert und wird entsprechend wieder gelöst, wenn die Kontaktkraft verschwindet.

Der Raum zwischen den vier Hilfsflächen erlaubt dem Kontaktpunkt P ein geringes Spiel, welches aber wegen der geringen Distanzen zwischen den Hilfsflächen nicht sichtbar wird. Dieser Spielraum ist jedoch erforderlich, damit sich solche Hilfskontakte bei Bedarf auch wieder lösen und gegebenenfalls an einer der anderen Hilfsflächen neu entstehen können.

Ein solcher Hilfskontakt entsteht jedoch nur aufgrund einer Kollision des Punktes P mit einer der Hilfsflächen. Eine solche Kollision wiederum kann nur eintreten, wenn die entsprechende Bewegung nicht bereits durch andere Zwangsbedingungen blockiert war. Dadurch wird sichergestellt, daß keine redundanten Hilfskontakte entstehen, ganz entsprechend wie dies auch für reguläre Kontaktbedingungen gilt. Dieser Mechanismus zur Vermeidung redundanter Kontaktkonfigurationen wurde bereits in Abschnitt 5.1 beschrieben.

Diese Vorgehensweise sei anhand eines einfachen zweidimensionalen Beispieles illustriert (vgl. Abb 45). Für einen schräg an einer Wand lehnenden Stab lassen sich drei kinematische Gleichgewichtsbedingungen aufstellen, und somit drei nicht-redundante Kraftkomponenten bestimmen. An den beiden Kontaktstellen gibt es jeweils eine Normalkraft und eine Haftreibungskraft (vgl. Abb. 45a). Diese 4 Kräfte lassen sich somit ohne weitere Randbedingungen nicht eindeutig bestimmen. Mit der Einführung von Hilfskontakten zur Realisierung der Haftreibung stellt sich diesselbe Anordnung beispielsweise wie in Abb.

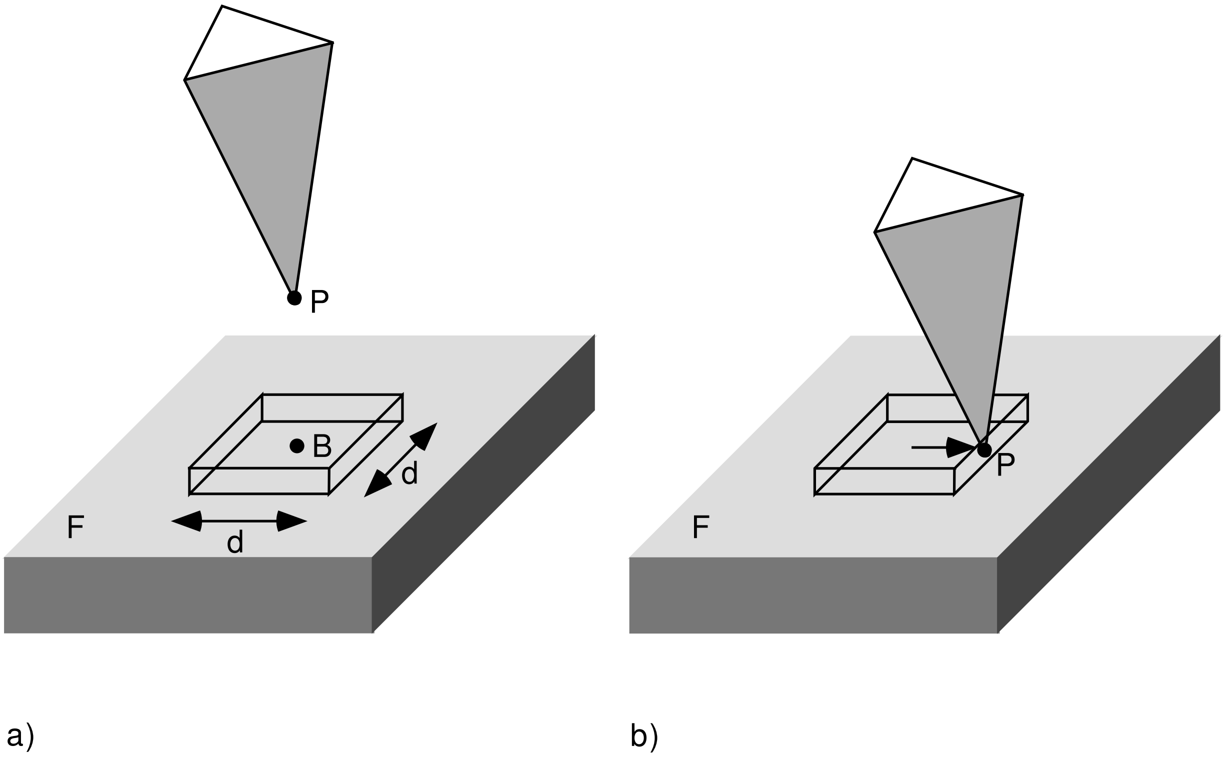


Abbildung 44: Realisierung der Haftreibung an einem Punkt-Fläche-Kontakt: Die Gleitbewegung des Punktes P entlang der Kontaktfläche F wird durch vier Hilfsflächen verhindert, die im Abstand d um den Kontaktpunkt B herum angeordnet sind. Links sind die beiden am Kontakt beteiligten Elemente separat dargestellt, rechts nachdem P mit einer der Hilfsflächen kollidiert ist.

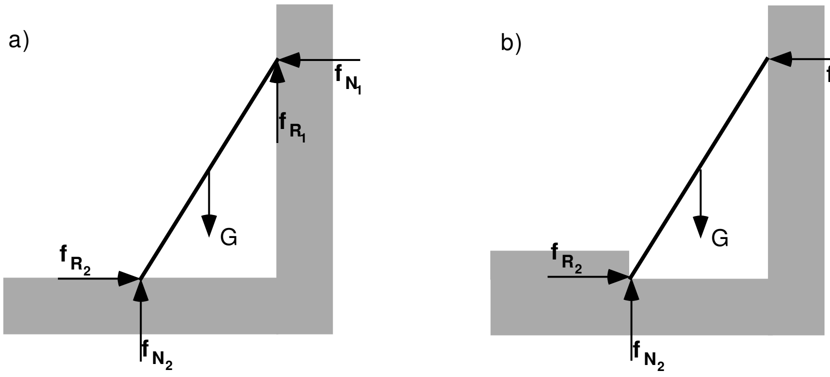


Abbildung 45: Die vier auf den schräg an einer Wand lehnenden Stab wirkenden Kraftkomponenten ($f_{N_1}, f_{R_1}, f_{N_2}, f_{R_2}$) lassen sich bei Haftreibung ohne weitere Annahmen nicht eindeutig bestimmen (links). Durch die Wirkung eines der im Text beschriebenen Hilfskontakte ergibt sich eine kinematisch eindeutig bestimmte Situation (rechts).

45b dar. Die Redundanz wird vermieden, da bereits eine der beiden Tangentialkräfte ausreicht, um die Gleitbewegung zu blockieren.

Es soll noch bemerkt werden, daß durch diese Realisierung der Haftreibung den am Kontakt beteiligten Körpern nur genau die erforderlichen Bewegungsfreiheitsgrade genommen werden. Dies zeigt sich am Beispiel einer Bleistiftspitze, an deren Kontakt zur Unterlage Haftreibung herrsche. Dadurch bleiben die Rotationsmöglichkeiten des Bleistiftes um seine Spitze, in der Simulation wie in der Realität, vollkommen erhalten.

7.2.3 Realisierung der Haftreibung an Kante-Kante-Kontakten

Die Realisierung der Haftreibung an Kante-Kante-Kontakten folgt demselben Prinzip wie eben für Punkt-Fläche-Kontakte beschrieben. Der Unterschied besteht lediglich in der besonderen Anordnung der zur Nachbildung der Haftreibung erforderlichen Hilfsflächen.

Da diese Hilfsflächen eine tangentiale Gleitbewegung blockieren sollen, müssen sie jeweils senkrecht zur Berührebene orientiert werden. Im Falle eines Punkt-Fläche-Kontaktes werden sie daher senkrecht zu der Fläche F angeordnet und sind fest mit ihr verbunden solange die Haftreibung an diesem Kontaktpunkt fortbesteht. Im Falle eines Kante-Kante-Kontaktes ergibt sich die Berührebene aus beiden beteiligten Kantenrichtungen, entsprechend müssen je zwei der Hilfsflächen so mit jeder der beiden Kanten verbunden werden, daß die Hilfsflächen orthogonal zur jeweiligen Kantenrichtung stehen.

Der Kante-Kante-Kontakt sei gegeben durch die zu Objekt 1 gehörende Kante K_1 mit dem Punkt P_1, sowie durch die zu Objekt 2 gehörende Kante K_2 mit dem Punkt P_2. P_1 und P_2 fallen im Kontaktpunkt zusammen. Geht dieser Kontakt in Haftreibung über, so werden gemäß Abb. 46 die zu K_1 orthogonalen Hilfsflächen F_{11} und F_{12} beiderseits von P_1 angeordnet, sowie die zu K_2 orthogonalen Hilfsflächen F_{21} und F_{22} beiderseits von P_2.

Gleitbewegungen zwischen den beteiligten Kanten werden nun blockiert durch Kollisionen von P_1 mit einer der Flächen F_{21} oder F_{22}, beziehungsweise durch Kollisionen von P_2 mit F_{11} oder F_{12}. Auch in diesem Fall bleibt durch den geringen Abstand zwischen jedem Flächenpaar ein unmerklicher Spielraum bestehen, so daß auch hier der Mechanismus zur Vermeidung redundanter Kontaktsituationen greifen kann.

7.3 Realisierung des Übergangs zwischen Gleiten und Haften

Kommt die Gleitbewegung zwischen zwei sich berührenden Körpern zum Stillstand, so geht die Gleitreibung in Haftreibung über. In einem zeitdiskreten System zur Kontaktsimulation sind die Kontaktgeschwindigkeiten jedoch nur zu diskreten Zeitpunkten explizit bekannt, und in der Regel erfolgt der Übergang von Gleitreibung in Haftreibung zwischen zwei solchen Zeitpunkten. Dies hat zur Konsequenz, daß die Gleitreibungskraft, deren Richtung sich aus der Gleitgeschwindigkeit $\mathbf{v}_G(t_1)$ zu Beginn des betreffenden Zeitintervalls ergibt, über das ganze Zeitintervall hinweg wirkt, und somit auch über den Zeitpunkt t^* hinaus, an dem die Gleitbewegung eigentlich zum Stillstand kommen müßte (vgl. Abb. 47). Die Gleitreibungskraft, die eigentlich nur bremsend wirken sollte, beschleunigt nun die Bewegung entgegen der ursprünglichen Richtung bis zum Ende des Simulationsintervalls t_2 auf die Geschwindigkeit $\mathbf{v}_G(t_2)$, die Bewegung schwingt gewissermaßen über. Entsprechend dem Wechsel der Bewegungsrichtung ändert sich damit auch die Richtung der Gleitreibungskraft für das folgende Zeitintervall $[t_2, t_3]$. Es besteht nun die Möglichkeit, daß sich derselbe Vorgang mit alternierenden Bewegungsrichtungen wiederholt, und das System ins Schwingen kommt. Da dabei zu keinem der diskreten Zeitpunkte die Gleitgeschwindigkeit gleich null wird, würde der Übergang zur Haftreibung auf diese Weise nicht erfolgen.

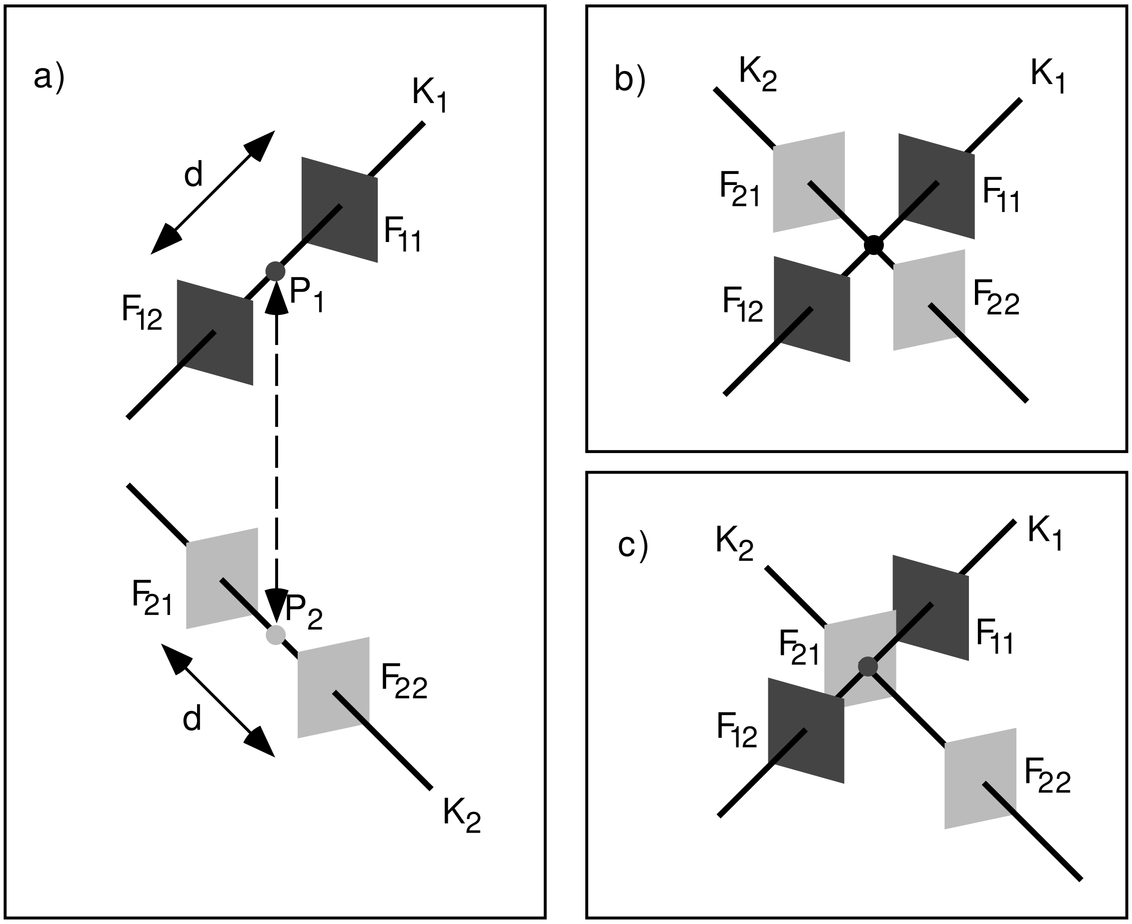


Abbildung 46: Realisierung der Haftreibung an einem Kante-Kante-Kontakt: Jeder der beiden Kanten K_1 und K_2 sind je zwei Hilfsflächen zugeordnet (links separat dargestellt, rechts oben in der Kontaktsituation). Eine Gleitbewegung wird dadurch unterbunden, daß P_1 mit einer der Flächen F_{21} oder F_{22} kollidiert, bzw. P_2 mit F_{11} oder F_{12} (rechts unten).

Um dies zu vermeiden, lassen wir den Übergang zur Haftreibung stets dann erfolgen, wenn sich die Richtung der Gleitbewegung umkehrt, d.h. wenn $\mathbf{v}_G(t)^T\mathbf{v}_G(t+\Delta t) < 0$. Diese Vorgehensweise ist in Abb. 48 dargestellt. Um auch den Fall der schleichenden Annäherung der Gleitgeschwindigkeit an Null abzudecken, erfolgt der Übergang zur Haftreibung weiterhin ebenfalls dann, wenn eine Mindestgeschwindigkeit v_{min} unterschritten wird, d.h. wenn $|\mathbf{v}_G(t+\Delta t)| < v_{min}$.

Besteht zwischen zwei Körpern eine linienhafte oder flächenhafte Kontaktregion, so

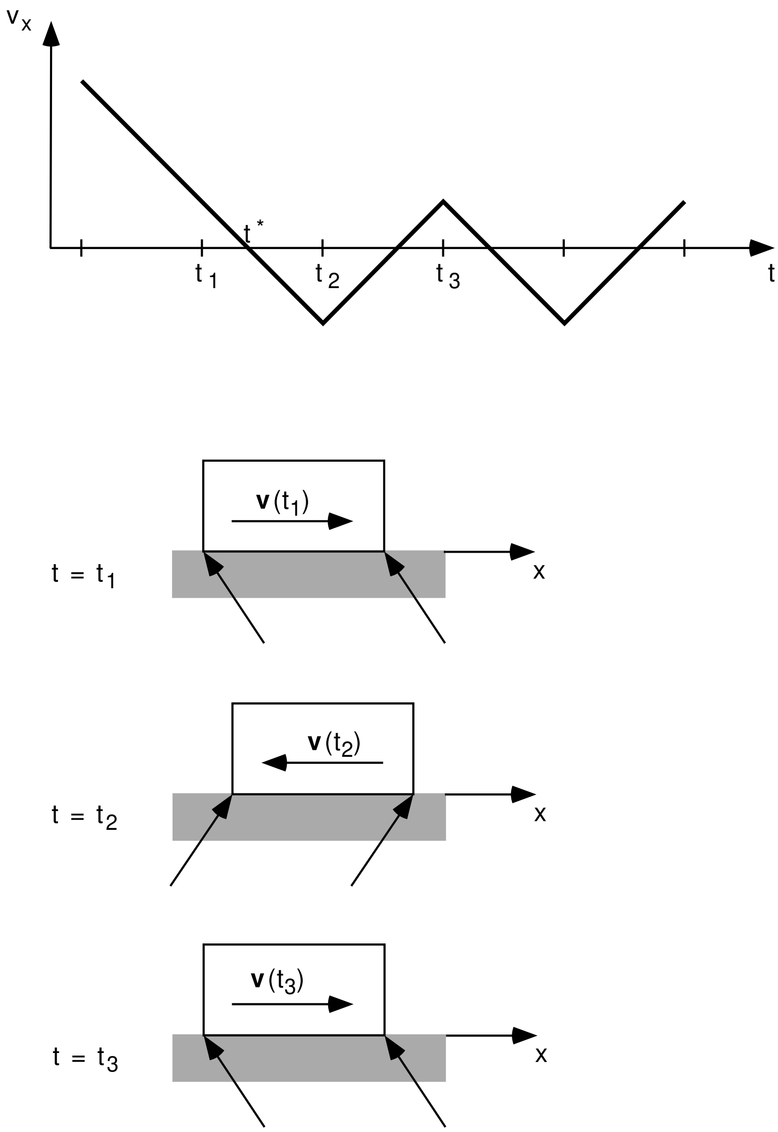


Abbildung 47: In zeitdiskreten Simulationen kann die Reibungskraft über das Ende der Gleitbewegung bei $t = t^*$ hinaus wirken, und beschleunigt dann das Objekt in Gegenrichtung. Dadurch kann das System ins Schwingen kommen.

wechselt stets die gesamte Kontaktregion gemeinsam zwischen Gleiten und Haften [9]. Eine solche Kontaktregion wird gemäß unserem Kontaktmodell durch zwei beziehungsweise drei Punktkontakte repräsentiert. Diese Kontakte müssen entsprechend auch gemeinsam zwischen Gleiten und Haften wechseln.

Aus diesem Grunde wechseln in der realisierten Implementierung stets alle Punktkontakte zwischen zwei Körpern gemeinsam zwischen Gleiten und Haften [10].

[9] Bei einer Gleitbewegung mit Rotationskomponente kann an einem einzelnen Punkt der Kontaktregion die Gleitgeschwindigkeit gleich null sein. Da der Beitrag eines solchen Punktes verglichen mit der gesamten Kontaktregion infinitesimal klein ist, muß dieser Fall nicht gesondert berücksichtigt werden.

[10] Diese Vorgehensweise ist auch dann gerechtfertigt, wenn mehrere Punktkontakte zwischen zwei

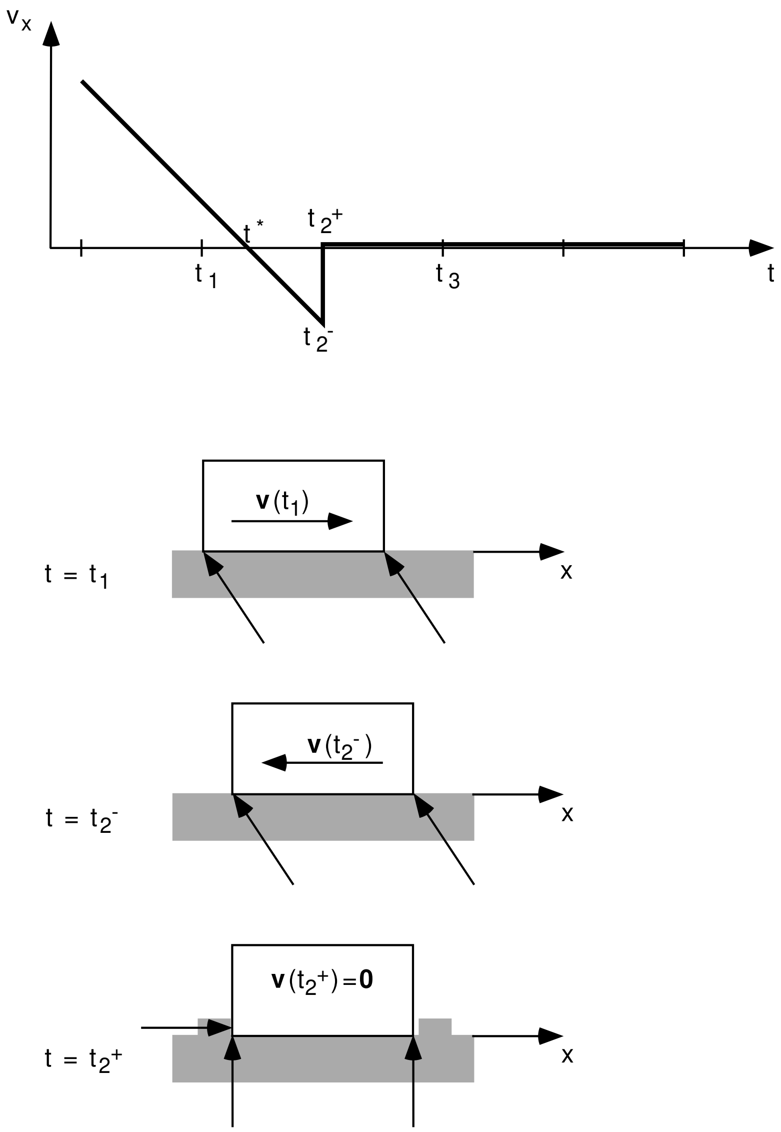


Abbildung 48: Kehrt sich die lokale Gleitgeschwindigkeitsrichtung durch die über $t = t^*$ hinaus wirkende Gleitreibungskraft um, so erfolgt der Übergang in Haftreibung. Durch die Einführung von zusätzlichen kinematischen Zwangsbedingungen wird eine weitere Gleitbewegung unterbunden, solange Haftreibung herrscht.

Umgekehrt geht die Haftreibung wieder in Gleitreibung über, wenn der Betrag der Haftreibungskraft den Maximalwert erreicht, d.h. wenn $f_{haft} \geq \mu_{haft} f_N$. Bestehen mehrere Kontaktpunkte zwischen zwei Körpern, so stellt sich die Frage, wie diese Bedingung nun angewandt werden muß.

Körpern keine linienhafte oder flächenhafte Kontaktregion repräsentieren, wie zum Beispiel beim Kontakt zwischen den Beinen eines Tisches mit dem Fußboden. Das resultierende Objektverhalten bleibt auch in solchen Fällen realistisch, solange nicht zwingend erwartet wird, daß ein Teil der Punktkontakte haftet und andere gleiten. Dies ist nur in Ausnahmesituationen der Fall, beispielsweise bei einem Zirkel, dessen Nadel haften muß während die Graphitspitze gleitet.

Da redundante Haftreibungskräfte vermieden werden, sind nicht alle Richtungskomponenten der Haftreibungskraft an jedem Kontaktpunkt vertreten, an einem der Kontaktpunkte eventuell sogar gar keine (vgl. Abb. 49). Der Übergang zwischen Haften und Gleiten kann daher für diese Kontaktpunkte nicht isoliert entschieden werden. Da zudem alle Punktkontakte zwischen zwei Körpern ihren Zustand gemeinsam wechseln, bietet es sich an, die obige Bedingung für einen Kontaktpunkt auf N Kontaktpunkte wie folgt zu erweitern:

$$\sum_{i=1}^{N} f_{haft_i} \quad \geq \quad \mu_{haft} \sum_{i=1}^{N} f_{N_i}.$$

Dabei entspricht f_{haft_i} der resultierenden Haftreibungskraft am i-ten Kontaktpunkt.

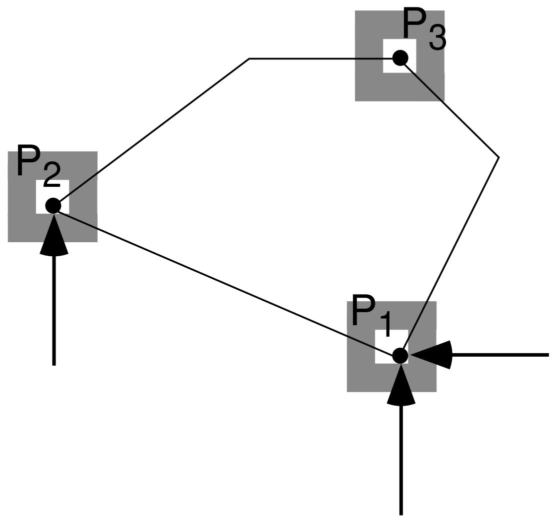


Abbildung 49: Eine ebene Gleitbewegung hat 3 Bewegungsfreiheitsgrade. Die Realisierung der Haftreibung durch kinematische Zwangsbedingungen erfordert also genau drei Zwangskräfte, die zusätzlich an den drei Kontaktpunkten ansetzen. An jedem Kontaktpunkt können bis zu zwei solche Kräfte ansetzen (P_1), eventuell aber auch keine (P_3).

8 Die Eigenschaften der Approximation von ebenen Flächenkontakten durch Punktkontakte

8.1 Physikalische Merkmale realer Flächenkontakte

Wir wollen uns bei den folgenden Betrachtungen auf die Merkmale von ebenen Flächenkontakten zwischen starren Körpern beschränken, da wir uns im Rahmen der Kontaktsimulation ebenfalls nur mit starren Körpern befassen. Die Beschränkung auf ebene Kontaktflächen ist durch die von uns verwendete polyhedrale Geometriebeschreibung motiviert.

Bei der Betrachtung realer Flächenkontakte zwischen starren Körpern auf der mikroskopischen Ebene stellt man zum einen fest, daß feste Körper, welche makroskopisch als starr erscheinen (d.h. nicht offensichtlich plastisch oder elastisch sind), auf der mikroskopischen Ebene stets eine gewisse Elastizität aufweisen. Desweiteren besitzt jede Oberfläche eines festen Körpers, auch wenn sie makroskopisch als glatt erscheint, auf mikroskopischer Ebene eine Oberflächenrauhigkeit, die sich alleine schon durch die atomare Struktur des jeweiligen Materials ergibt.

Dadurch folgt, daß flächenhafte Kontakte im mathematisch strengen Sinne gar nicht existieren, sondern stets durch eine mehr oder minder große Anzahl von Punktkontakten gebildet werden. Die Anzahl und die geometrische Anordnung dieser Punktkontakte hängen stark ab von der Mikrostruktur der beteiligten Oberflächen, von der Elastizität der jeweiligen Materialien, sowie von dem an der Kontaktfläche herrschenden Flächendruck. Insbesondere der letzte Faktor hat sich in jüngeren Ergebnissen der Nanotribologie[11] als interessant herausgestellt. Der näherungsweise lineare Zusammenhang zwischen Gleitreibungskraft und Normalkraft ist bereits lange bekannt, und wird allgemein als Coulomb'sches Reibungsgesetz bezeichnet[12]. Erst in jüngerer Zeit wurde der physikalische Hintergrund dieses Gesetzes näher beleuchtet [43]. Dabei wurde festgestellt, daß die Gleitreibungskraft eines Punktkontaktes unabhängig von der wirkenden Normalkraft ist. Mit zunehmender Normalkraft nimmt jedoch aufgrund der Mikroelastizität die Anzahl der Kontaktpunkte zu, und damit auch die insgesamt wirkende Gleitreibungskraft. Die Anzahl und die geometrische Verteilung der im Verlaufe einer Gleitbewegung jeweils bestehenden Punktkontakte kann als stochastischer Prozeß aufgefaßt werden. Das makroskopische Objektverhalten ergibt sich dann durch die Integration der Beiträge der zeitlich variierenden mikroskopischen punktförmigen Reibungskontakte.

8.1.1 Modellierung realer Flächenkontakte

Als Basis zur Bewertung unseres Ansatzes der Beschreibung von Flächenkontakten durch wenige Punktkontakte im Vergleich zu realen Flächenkontakten benötigen wir eine Modellierung der realen Verhältnisse.

[11] Die Nanotribologie befaßt sich mit den Ursachen der Reibung und des Gleitwiderstandes auf atomarer Ebene.

[12] Dieses Reibungsgesetz wurde bereits von Leonardo da Vinci (1452-1519) formuliert, und unabhängig davon im 17. Jahrhundert von Guillaume Amontons (1663-1806).

Wir unterstellen, die Anzahl der Punktkontakte im realen Fall sei ausreichend groß und gleichmäßig verteilt, so daß der makroskopische Effekt aufgrund der sich stochastisch ändernden räumlichen Verteilung vernachlässigt werden kann. Dann ist es zulässig, reale Flächenkontakte durch eine gleichmäßig verteilte, gleichbleibende Anordnung einer festen, ausreichend großen Anzahl von Punktkontakten zu modellieren. Die Punktkontaktdichte wird so gewählt, daß eine weitere Erhöhung das resultierende makroskopische Verhalten nicht mehr nennenswert beeinflußt.

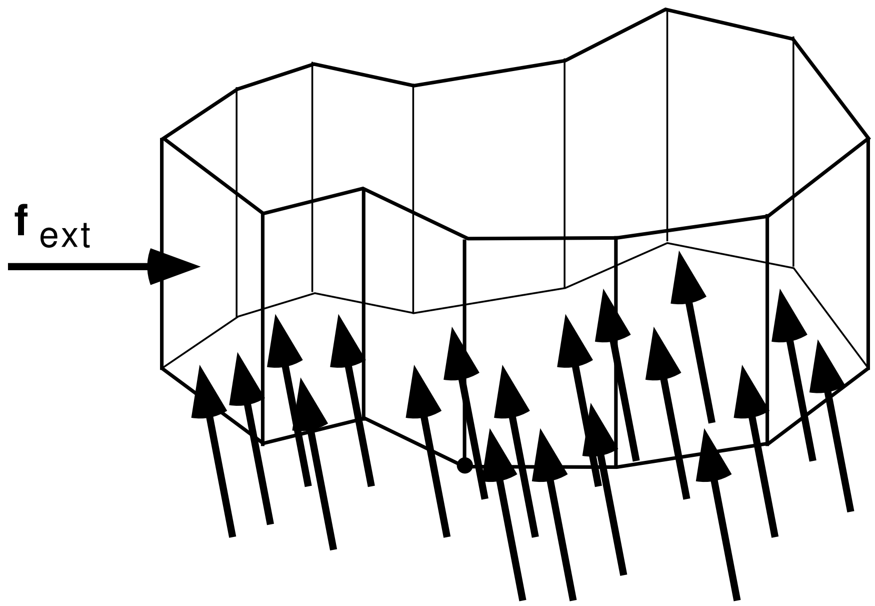


Abbildung 50: Versuchsanordnung zur Bewertung verschiedener Anzahlen von zufällig angeordneten Punktkontakten als Approximation eines flächigen Reibungskontaktes. Der prismenförmige Körper wird von einer externen Kraft konstanter Richtung angeschoben, die an einem festen Punkt des Körpers ansetzt.

Zur Klärung dieser Fragestellung betrachten wir prismatische Körper mit ebener Grundfläche, die auf einer ebenen Unterlage ruhen, und durch eine horizontal wirkende externe konstante Kraft in Bewegung gesetzt werden. Diese Situation ist in Abb. 50 dargestellt. Innerhalb der polygonalen Auflagefläche des Prismas ordnen wir eine zufällig verteilte Menge von Punktkontakten an. In einer Serie von Experimenten simulieren wir nun das Bewegungsverhalten der Prismen für zunehmend dichtere Punktkontaktverteilungen. Als Anzahl der Kontaktpunkte wählen wir dabei 10, 30, 100, 300, 1000 und 3000. Als Vergleichskriterium für das Bewegungsverhalten verwenden wir den zeitlichen Verlauf des Rotationswinkels und den zeitlichen Verlauf der Bewegung des Schwerpunktes der Prismen bei nichtzentralem Anschieben, wie es schematisch in Abb. 66 dargestellt ist. Der zeitliche Verlauf des Rotationswinkels und die Bahn des Schwerpunktes sind in den Diagrammen 52 und 53 dargestellt. Es läßt sich feststellen, daß das Bewegungsverhalten bei zunehmender Dichte der Kontaktpunkte rasch konvergiert, und sich bereits ab 100 Punktkontakten keine erkennbaren Veränderungen mehr zeigen. Entsprechende Untersuchungen mit verschiedenen Geometrien der Auflagefläche liefern vergleichbare Ergebnisse. Hierzu werden als Testkörper Prismen gewählt, deren polygonale Grundflächen durch einen Zufallsprozess erzeugt wurden. Dazu wurde zunächst die Anzahl n_K der Kanten der jeweiligen Grundfläche festgelegt. Dann wurden die Koordinaten des i-ten Eckpunktes wie folgt bestimmt:

$$
\begin{aligned}
r_i &= random(0,1) \\
x_i &= r_i \cos \frac{2\pi i}{n_K} \\
y_i &= r_i \sin \frac{2\pi i}{n_K}.
\end{aligned}
$$

Ein Beispiel für die derart erzeugten Geometrien zeigt Abb. 51.

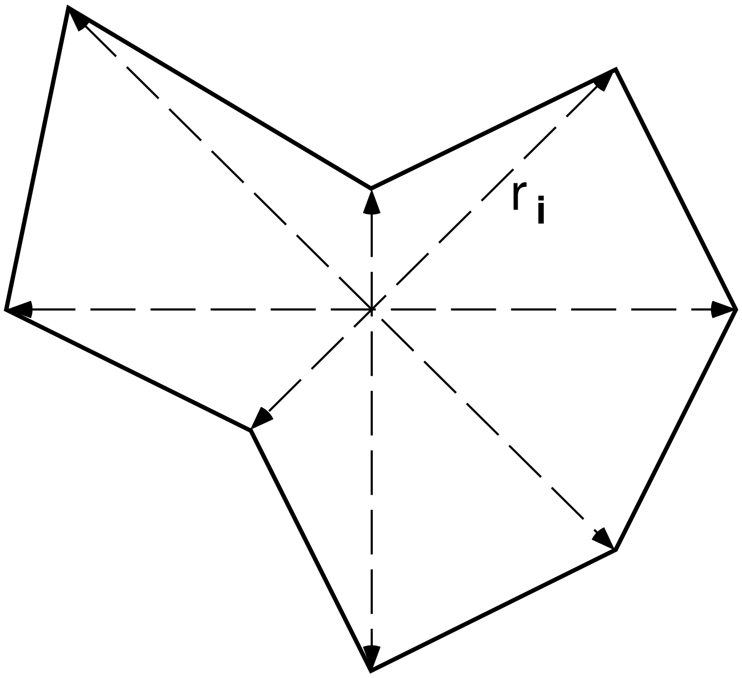


Abbildung 51: Beispiel für die Geometrie der Grundfläche eines Prismas, welches mit dem im Text beschriebenen Verfahren zufällig erzeugt wurde.

Neben der geometrischen Verteilung der Kontaktpunkte spielt auch die Verteilung der Normalkräfte auf die einzelnen Punktkontakte eine Rolle für das makroskopische Reibungsverhalten. Insbesondere kann nicht angenommen werden, diese Verteilung (und damit der Flächendruck) sei in der gesamten Kontaktfläche konstant.

Formal läßt sich dies damit begründen, daß ein Flächenkontakt einer dreiwertigen kinematischen Lagebedingung entspricht, welche durch die wirkenden Kontaktkräfte erzwungen wird. Daher muß die Verteilung der Normalkräfte mindestens drei unabhängige Variablen besitzen. Im Falle der minimalen Anzahl von genau drei Kontaktpunkten entsprechen diese Variablen gerade den drei Kontaktkräften selbst. Anschaulich ist dies ebenfalls einsichtig, da die Kontaktkräfte neben der pauschalen Normalkraft auch Momente zwischen den sich berührenden Körpern realisieren müssen, um den Kontakt ohne Durchdringungen aufrecht zu erhalten.

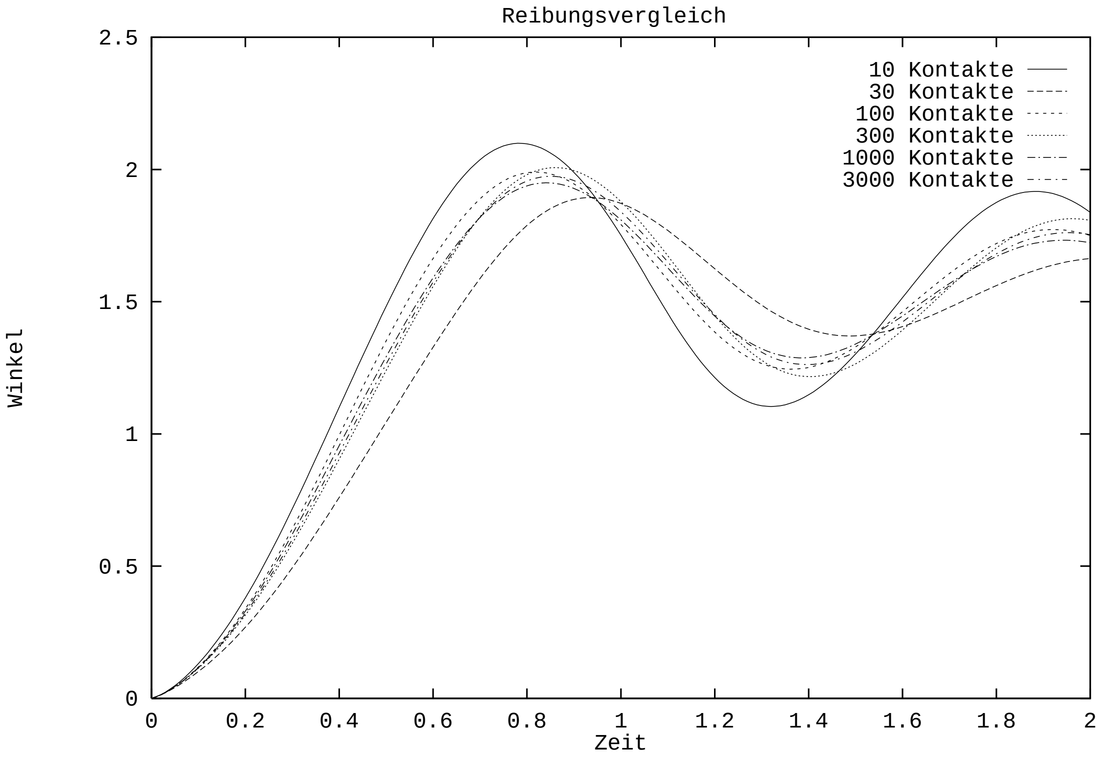


Abbildung 52: Vergleich der Rotationsbewegungen eines exzentrisch angeschobenen Körpers bei Variation der Punktkontaktdichte. Ab 100 Kontaktpunkten ergeben sich nur noch geringe Variationen der resultierenden Bewegung.

Für eine exakte Bestimmung der Verteilung der Normalkräfte müßte das Elastizitätsverhalten der beteiligten Materialien genau berücksichtigt werden, was jedoch den Rahmen unserer Untersuchung sprengen würde. Daher beschränken wir uns darauf, für die Verteilung der Normalkräfte innerhalb der Kontaktregion eine systematische Gesetzmäßigkeit anzunehmen. Beschreibt man die Lage eines Kontaktpunktes in einem ebenen orthogonalen Koordinatensystem durch die Koordinaten (u, v), so stellen wir die auf diesen Kontaktpunkt entfallende Normalkraft $f_N(u, v)$ dar als

$$f_N(u,v) \quad = \quad f_0 + f_{\lambda_u}(u) + f_{\lambda_v}(v). \tag{53}$$

Die zugrundegelegten Funktionen f_{λ_u} und f_{λ_v} wählen wir dabei wie folgt:

$$\begin{aligned} f_{\lambda_u}(u) &= \lambda_u \cdot \hat{u} \cdot |\hat{u}|^s, \\ f_{\lambda_v}(v) &= \lambda_v \cdot \hat{v} \cdot |\hat{v}|^s, \end{aligned}$$

wobei die Koordinaten u und v normiert werden auf die Dimensionen u_{max}, v_{max} des umschreibenden Rechteckes der polygonalen Kontaktregion, d.h.

$$\hat{u} \quad = \quad \frac{u}{u_{max}},$$

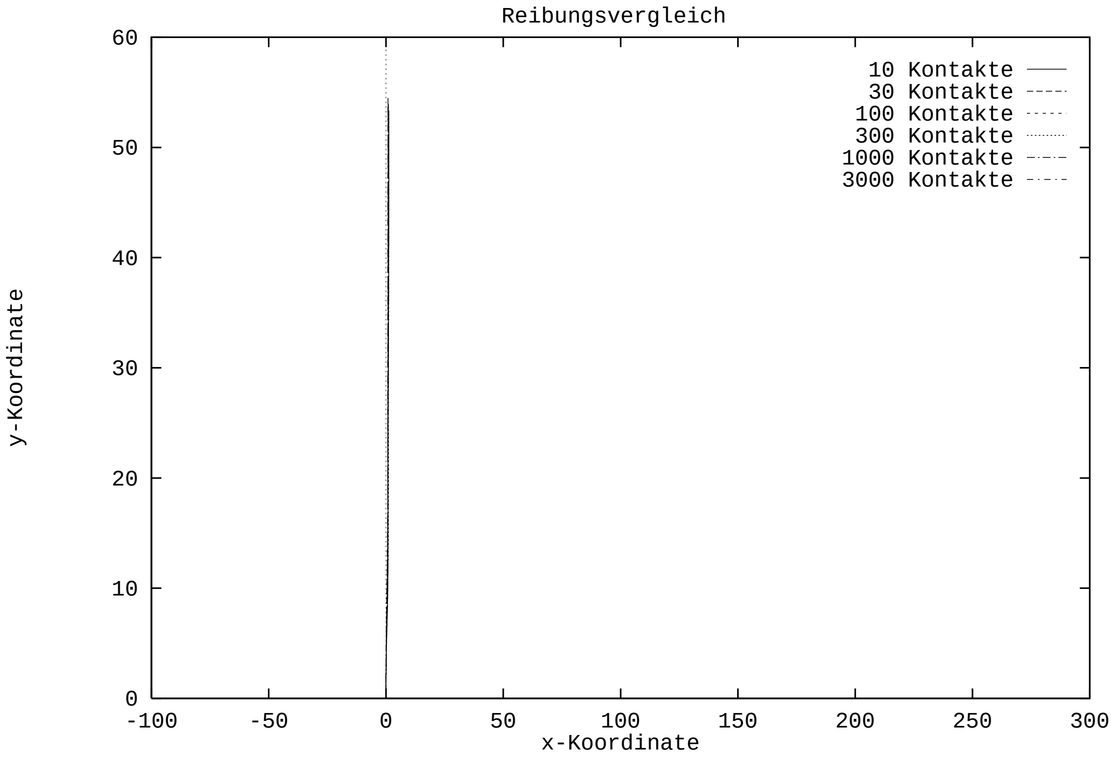


Abbildung 53: Vergleich der Bewegungsbahnen des Schwerpunktes eines exzentrisch angeschobenen Körpers bei Variation der Punktkontaktdichte. Es sind nur sehr geringe Abweichungen von der geradlinigen Bewegung zu beobachten.

$$\hat{v} = \frac{v}{v_{max}}.$$

Der Parameter $s \in (-1, \infty)$ legt die Nichtlinearität des jeweiligen Funktionsverlaufes fest. Dieser Verlauf ist in Abb. 54 für einige Werte von s aufgetragen.

Die 3 Parameter $f_0, \lambda_u, \lambda_v$ in Gleichung (53) müssen sich aus den kinematischen und dynamischen Kontaktbedingungen des Flächenkontaktes ergeben. Hierzu werden an drei der Kontaktpunkte die üblichen Kontaktbedingungen formuliert. Zur Einhaltung dieser Kontaktbedingungen werden nun aber nicht speziell die an diesen Punkten wirkenden Kontaktkräfte angesetzt, sondern die gesamte durch $f_0, \lambda_u, \lambda_v$ parametrisierte Menge von Kontaktkräften (vgl. Abb. 56). Bei der Lösung des Gleichungssystems zur Bestimmung dieser Kontaktkräfte dienen die drei Parameter $f_0, \lambda_u, \lambda_v$ als Unbekannte. Die Kontaktkräfte selbst folgen daraus gemäß Gleichung (53). Dabei ergeben sich vergleichbare Ergebnisse weitgehend unabhängig von Variationen des Parameters s (siehe Abb. 55).

Die kinematischen Eigenschaften und die Reibungseigenschaften realer Flächenkontakte, die wie eben dargestellt modelliert werden, sollen nun in den folgenden beiden Abschnitten mit den Eigenschaften unseres Modelles mit weniger (im konkreten Fall dreier) Punktkontakte verglichen werden.

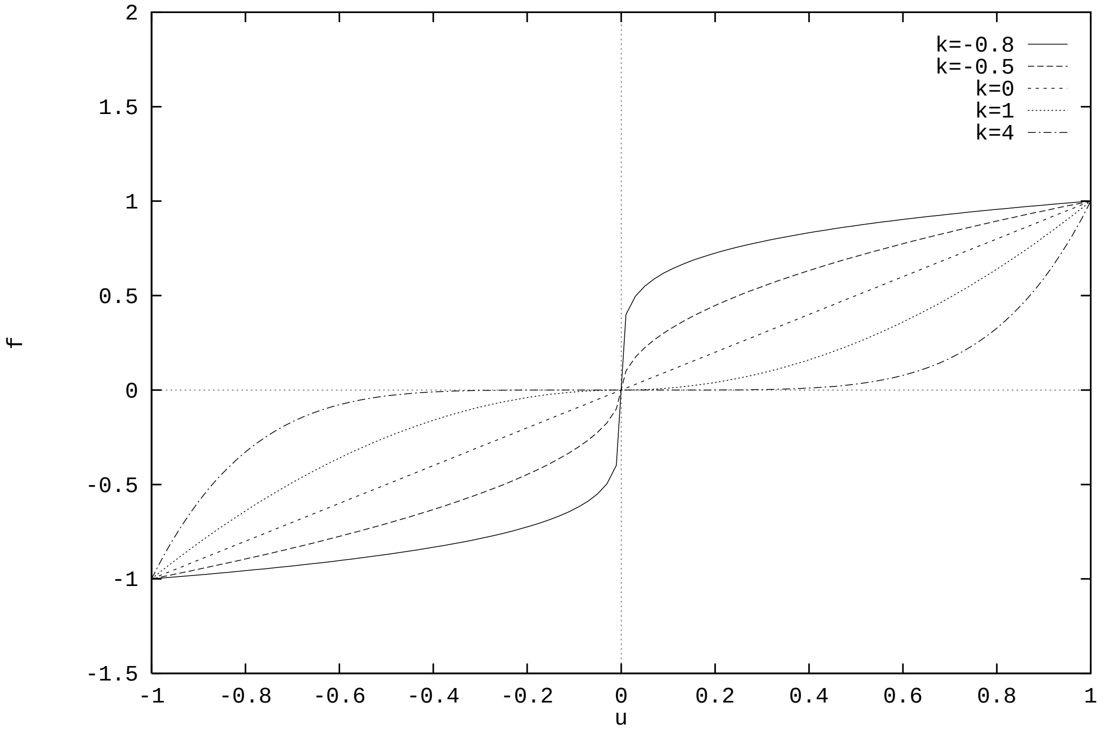


Abbildung 54: Der Verlauf der Funktionen f_λ für einige Werte des Parameters s, welcher die Nichtlinearität der Funktion bestimmt. Der Parameter u ist normiert auf die Abmessungen des umschreibenden Rechtecks der polygonalen Kontaktregion.

8.2 Vergleich der kinematischen Eigenschaften

Die kinematischen Kontaktbedingungen eines Flächenkontaktes werden bereits durch drei geeignet angeordnete, insbesondere nicht kollineare, Punktkontakte exakt repräsentiert. Um dies zu zeigen, formulieren wir zunächst folgenden Hilfssatz:

Gegebenen sei eine von einem beliebigen Polygon berandete ebene Kontaktfläche, innerhalb deren n Punktkontaktkräfte mit nichtnegativer Normalkomponente $f_i, i = 1, ...n$ wirken. Diese Normalkräfte lassen sich durch eine einzige Normalkraft ersetzen, die an geeigneter Stelle innerhalb der ebenen konvexen Hülle der Kontaktfläche angreift.

Beweis:

Sollte die Kontaktfläche nichtkonvex sein, so ersetzen wir sie durch ihre konvexe Hülle. In den dadurch hinzugenommenen Flächenanteilen wirken natürlich keine weiteren Kontaktkräfte. In der Kontaktebene wählen wir ein orthogonales Koordinatensystem (u, v). Die Behauptung lautet damit, es existiere ein Koordinatenpaar (u^*, v^*) im Inneren der Fläche und eine Normalkraft g, so daß gilt:

$$g = \sum_{i=1}^{n} f_i \tag{54}$$

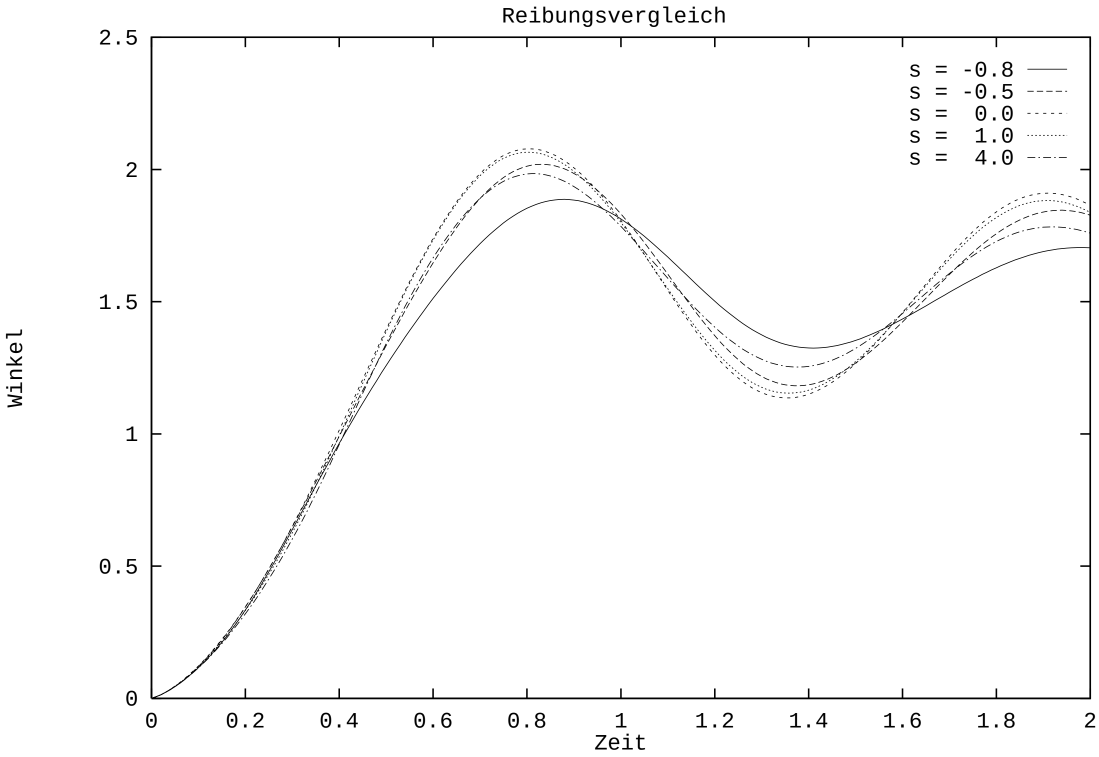


Abbildung 55: Vergleich der Rotationsbewegungen eines exzentrisch angeschobenen Körpers bei Variation des Parameters s. Die Simulationen wurden mit jeweils 100 stochastisch gewählten Kontaktpunkten und einer relativen Exzentrizität der anschiebenden Kraft von 0.5 durchgeführt.

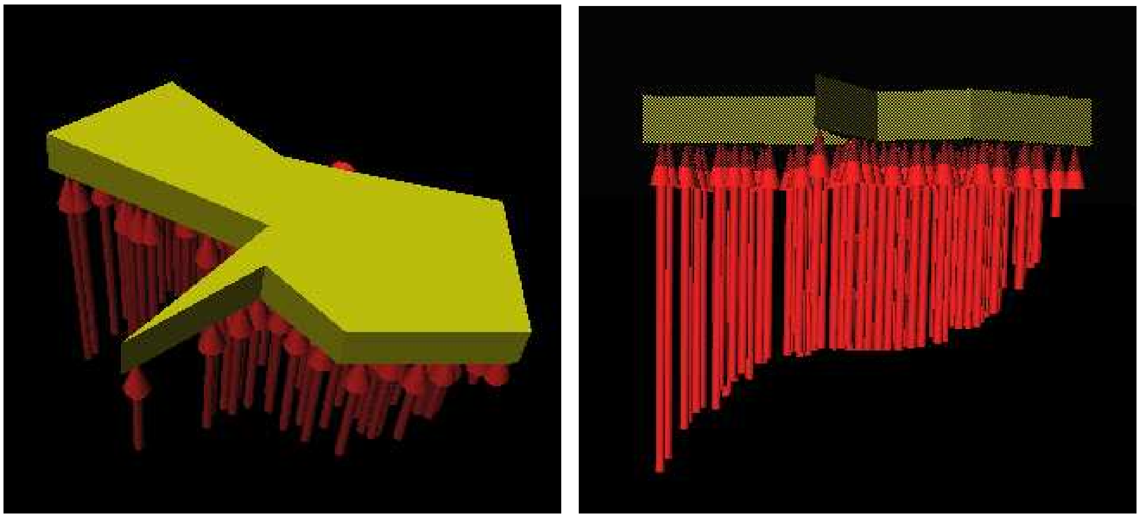


Abbildung 56: Links: Beispiel eines Prismas mit zufällig erzeugter Grundflächengeometrie. Rechts: Beispiel einer nichtlinearen Verteilung der Beträge der Kontaktkräfte gemäß Gl. (53). Der Wert des Verteilungsparameters beträgt in diesem Beispiel $s = 2.0$.

$$u^* g = \sum_{i=1}^{n} u_i f_i \tag{55}$$

$$v^* g = \sum_{i=1}^{n} v_i f_i, \tag{56}$$

wobei Gl.(54) die Kraftbedingung und Gln.(55,56) die Momentenbedingungen darstellen.

Durch Auflösen nach u^* und v^* folgt unmittelbar

$$u^* = \frac{\sum_{i=1}^{n} u_i f_i}{\sum_{i=1}^{n} f_i} \tag{57}$$

$$v^* = \frac{\sum_{i=1}^{n} v_i f_i}{\sum_{i=1}^{n} f_i}. \tag{58}$$

Die Koordinaten (u^*, v^*) stellen also das gewichtete Mittel der jeweiligen Kontaktpunkte (u_i, v_i) dar, und liegen somit im Innern der konvexen Hülle der Kontaktfläche. Es ist nun leicht zu zeigen, daß es drei Normalkräfte g_j an den Kontaktpunkten $(u_j, v_j), j = 1, ...3$ im Innern der konvexen Hülle der Kontaktfläche gibt, für die gilt

$$g = \sum_{i=1}^{3} g_j$$

$$u^* g = \sum_{i=1}^{3} u_j g_j$$

$$v^* g = \sum_{i=1}^{3} v_j g_j,$$

sofern das durch die Punkte (u_j, v_j) bestimmte Dreieck den Punkt (u^*, v^*) einschließt. Diese drei Punktkontaktkräfte sind somit ebenfalls äquivalent zu den Punktkontaktkräften f_i.

Es bleibt noch zu zeigen, daß die drei Kontaktpunkte so gewählt werden können, daß sie nicht nur im Inneren der konvexen Hülle der Kontaktfläche liegen, sondern in Innern der Kontaktfläche selbst.

Um dies zu zeigen stellen wir fest, daß alle Eckpunkte der konvexen Hülle auch Eckpunkte der Kontaktfläche darstellen. Zerlegt man das Polygon der konvexen Hülle in Dreiecke, so liegt der Punkt (u^*, v^*) in genau einem dieser Dreiecke (siehe Abb. 57). Die Eckpunkte dieses Dreiecks erfüllen alle Forderungen an die gesuchten Punkte $(u_j, v_j), j = 1, ...3$.

Wir haben also gezeigt, daß die kinematischen Randbedingungen eines ebenen Flächenkontaktes durch drei geeignet gewählte Punktkontaktkräfte ebenso gewährleistet werden können wie durch die vielen Kontaktkräfte des realen Flächenkontaktes.

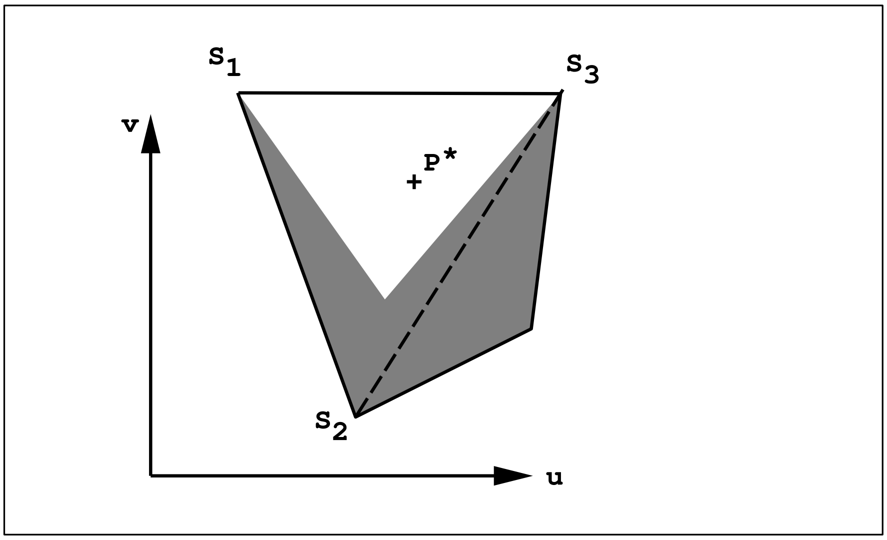


Abbildung 57: Eine konkave polygonale Kontaktregion (schattiert) und ihre konvexe Hülle, P^* bezeichnet den stets im Inneren der konvexen Hülle befindlichen Ansatzpunkt der resultierenden Kontaktkraft. Aus den Eckpunkten der polygonalen Kontaktregion lassen sich stets drei Eckpunkte $S_j, j = 1, ...3$ so auswählen, daß P^* im Inneren des dadurch definierten Dreicks liegt.

8.3 Approximation der Reibungseigenschaften

8.3.1 Theoretischer Vergleich der Reibung bei flächiger und punktueller Auflage

Allgemeines

Wir betrachten im folgenden einen Körper, der mit einer ebenen Fläche auf einer horizontalen ebenen Unterlage aufliegt und sich entlang dieser bewegt. Zur Beschreibung dieser Bewegung und der auftretenden Kräfte und Momente wählen wir das Koordinatensystem so, daß die xz-Ebene mit der Berührebene zusammenfällt, und die y-Achse zum betrachteten Zeitpunkt durch den Schwerpunkt des bewegten Körpers geht (Abb. 58). Auf den Körper wirke die Gewichtskraft $\mathbf{G} = [0, -f_g, 0]^T$. Die Bewegung des Körpers zum betrachteten Zeitpunkt wird beschrieben durch die Translation $\mathbf{v}$ und die Rotation $\boldsymbol{\omega}$ relativ zur Unterlage:

$$\begin{aligned} \mathbf{v} &= [v_x, 0, v_z]^T, \\ \boldsymbol{\omega} &= [0, \omega_y, 0]^T. \end{aligned}$$

An den Berührstellen zwischen dem bewegten Körper und der Unterlage treten Reibungskräfte auf, die sich zu einer resultierenden Reibungskraft und einem resultierenden Reibmoment summieren, und wesentlich auf das Bewegungsverhalten des Körpers Einfluß nehmen. Wir wollen nun untersuchen, inwieweit dieser Einfluß von der Geometrie der Berührstellen abhängt, insbesondere davon, ob eine flächenhaft verteilte Reibregion vorliegt (Abb. 58) oder diskrete Auflagepunkte (Abb. 59).

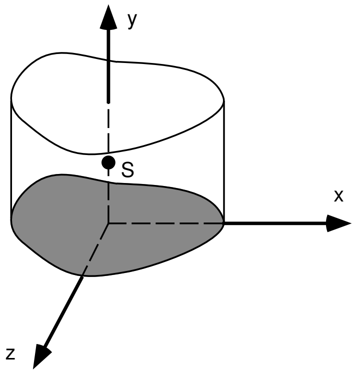


Abbildung 58: Koordinatensystem. Die ebene Berührfläche (grau schattiert) liegt in der xz-Ebene, die y-Achse geht durch den Schwerpunkt S des bewegten Körpers.

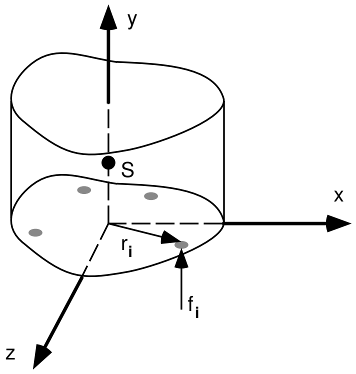


Abbildung 59: Diskrete Punktkontakte mit Ortsvektor $\mathbf{r}_i$ und Normalkraft $\mathbf{f}_i$.

Den Beitrag $d\mathbf{f}_R$ zur Reibungskraft eines Flächenelements dA an der Stelle $\mathbf{x} = [x, y, z]^T$ bei Gleitreibung werde gemäß dem Reibungsmodell nach Coulomb beschrieben als

$$\mathrm{d}\mathbf{f}_R(\mathbf{x}) = -\mu \frac{\mathbf{v}(\mathbf{x})}{|\mathbf{v}(\mathbf{x})|} p_A(\mathbf{x}) \mathrm{d}A,$$

wobei $\mathbf{v}(\mathbf{x})$ die lokale Geschwindigkeit des bewegten Körpers relativ zur Unterlage bezeichnet, und der Gleitreibungskoeffizient μ in der gesamten Berührfläche gleich sei.

Daraus resultieren die Kraft $\mathbf{F}_R$ und das Moment $\mathbf{M}_R$ durch Reibung:

$$\mathbf{F}_R = -\mu \int_A \frac{\mathbf{v}(\mathbf{x})}{|\mathbf{v}(\mathbf{x})|} p_A(\mathbf{x})\, \mathrm{d}A, \tag{59}$$

$$\mathbf{M}_R = -\mu \int_A \mathbf{r} \times \frac{\mathbf{v}(\mathbf{x})}{|\mathbf{v}(\mathbf{x})|} p_A(\mathbf{x})\, \mathrm{d}A. \tag{60}$$

Die x- und z-Komponenten von $\mathbf{M}_R$ sind stets Null.

Die Normalkräfte der Auflagefläche bewirken ebenfalls einen Beitrag zu der Kraft und dem Moment welche insgesamt auf den Körper wirken:

$$\begin{aligned}\mathbf{F}_N &= \int\limits_A p_A(\mathbf{x})\,\mathrm{d}A,\\ \mathbf{M}_N &= \int\limits_A \mathbf{x}\times\hat{\mathbf{e}}_y p_A(\mathbf{x})\,\mathrm{d}A,\end{aligned}$$

wobei $\hat{\mathbf{e}}_y = [0,1,0]^T$ den Einheitsvektor in Richtung der y-Achse bezeichnet. Die y-Komponente von $\mathbf{M}_N$ ist stets Null.

Zusätzlich lassen wir noch eine externe Kraft $\mathbf{F}_{ext} = [f_x, 0, f_z]^T$ zu, die horizontal am Punkt $\mathbf{r} = [r_x, r_y, r_z]^T$ ansetzt, und ein zusätzliches Moment $\mathbf{r}\times\mathbf{F}_{ext}$ hervorruft. Damit der Körper die Unterlage nicht verläßt muß für die resultierende Kraft $\mathbf{F} = [F_x, F_y, F_z]^T = \mathbf{F}_R + \mathbf{F}_N + \mathbf{F}_{ext}$ und das resultierende Moment $\mathbf{M} = [M_x, M_y, M_z]^T = \mathbf{M}_R + \mathbf{M}_N + \mathbf{r}\times\mathbf{F}_{ext}$ gelten:

$$F_y = 0 \qquad M_x = 0 \qquad M_z = 0. \tag{61}$$

Berühren sich der bewegte Körper und die ebene Auflagefläche nur in einer Anzahl N diskreter Punkte, so gehen die Intergrale (59) und (60) in diskrete Summen über:

$$\begin{aligned}\mathbf{F}_R &= -\mu\sum_{i=1}^{N}\frac{\mathbf{v}_i}{|\mathbf{v}_i|}f_i, && (62)\\ \mathbf{M}_R &= -\mu\sum_{i=1}^{N}\mathbf{r}_i\times\frac{\mathbf{v}_i}{|\mathbf{v}_i|}f_i. && (63)\\ & && (64)\end{aligned}$$

Rein translatorische Bewegung

Nun betrachten wir den Fall rein translatorischer Bewegung, bei der die lokale Geschwindigkeit $\mathbf{v} = [v_x, 0, v_z]^T$ überall gleich ist. Der Körper werde durch $\mathbf{F}_{ext}$ zentral und parallel zur Bewegungsrichtung angeschoben, d.h. es gelte $f_x r_z = f_z r_x$ und $f_x v_z = f_z v_x$. Anschaulich bedeutet dies, daß die Kraftlinie von $\mathbf{F}_{ext}$ die y-Achse schneidet.

Die Komponenten des resultierenden Momentes $\mathbf{M}$ ergeben sich damit zu

$$\begin{aligned}M_x &= -\int\limits_A z p_A(\mathbf{x})\,\mathrm{d}A + r_y f_z && (65)\\ M_y &= \frac{\mu}{|\mathbf{v}|}\int\limits_A (x v_z - z v_x) p_A(\mathbf{x})\,\mathrm{d}A && (66)\\ M_z &= \int\limits_A x p_A(\mathbf{x})\,\mathrm{d}A - r_y f_x. && (67)\end{aligned}$$

Setzt man (61), (65) und (67) in (66) ein, so ergibt sich für M_y

$$M_y = \frac{\mu r_y}{|\mathbf{v}|}(M_x\, v_x + M_z\, v_z) = 0.$$

Aus den Bedingungen $M_x = M_z = 0$ folgt damit unmittelbar, daß auch das Moment um die y-Achse M_y verschwindet. Solange kein zusätzliches äußeres Moment angreift bleibt damit die Bewegung rein translatorisch. Dies gilt unabhängig von der Geometrie der Auflagefläche, und somit auch für diskrete Stützpunkte, unabhängig von deren geometrischer Anordnung.

Mason [32] zeigt, daß bei rein translatorischer Bewegung sich die über die Kontaktfläche verteilten Reibungskräfte durch eine einzige resultierende Punktkraft ersetzen lassen, welche in der Projektion des Schwerpunktes auf die Berührebene angreift. Dies bestätigt unser Ergebnis, da diese resultierende Reibungskraft somit kein Moment um die y-Achse erzeugt.

Bewegung mit rotatorischen Komponenten

Jede nicht rein translatorische Objektbewegung in der xz-Ebene läßt sich zu einem gegebenen Zeitpunkt als Rotation um eine zur y-Achse parallele Achse darstellen, welche durch ein Rotationszentrum $\mathbf{p}_r = [p_{rx}, 0, p_{rz}]^T$ geht. Dieses ergibt sich aus den Bedingungen $\mathbf{v} + \boldsymbol{\omega} \times \mathbf{p}_r = \mathbf{0}$ und $\boldsymbol{\omega}\mathbf{p}_r = 0$ zu

$$\mathbf{p}_r = \frac{\boldsymbol{\omega} \times \mathbf{v}}{|\boldsymbol{\omega}|^2}. \tag{68}$$

Der Fall reiner Translation, d.h. $\boldsymbol{\omega} = \mathbf{0}$, kann hierbei als Grenzfall einer Rotation mit unendlichem Rotationsradius betrachtet werden.

Wir beschränken uns im folgenden auf das ebene Problem, so daß die Momente M_x und M_z nicht weiter betrachtet werden. Desweiteren sollen die Bewegungen quasi-statisch verlaufen, so daß Trägheitskräfte vernachlässigbar sind. Der Flächendruck $p_A(\mathbf{x})$ muß nicht überall in der Berührfläche gleich sein, die Resultierende der Normalkräfte muß jedoch aus statischen Gründen auf den Schwerpunkt zeigen, d.h. es gilt

$$\int_A \mathbf{x}\, p_A(\mathbf{x})\, \mathrm{d}A \quad = \quad [s_x, 0, s_y] = \mathbf{0}. \tag{69}$$

Unter diesen Voraussetzungen ergeben sich gemäß [28] die resultierende Reibungskraft $\mathbf{F}_R$ und das resultierende Reibungsmoment $\mathbf{M}_r$ bezüglich des Rotationszentrums $\mathbf{p}_r$ zu

$$\mathbf{F}_R \quad = \quad -\mu \operatorname{sgn}(\omega_y)\hat{\mathbf{e}}_y \times \int_A \frac{\mathbf{x} - \mathbf{p}_r}{|\mathbf{x} - \mathbf{p}_r|} p_A(\mathbf{x})\, \mathrm{d}A, \tag{70}$$

$$\mathbf{M}_r \quad = \quad -\mu \operatorname{sgn}(\omega_y)\hat{\mathbf{e}}_y \int_A |\mathbf{x} - \mathbf{p}_r| p_A(\mathbf{x})\, \mathrm{d}A. \tag{71}$$

Ähnlich wie die Wirkung der Gravitation auf einen räumlich ausgedehnten Körper durch eine äquivalente Gewichtskraft beschrieben werden kann, die punktuell im Schwerpunkt angreift, so können auch die auf die Berührfläche verteilten Reibungskräfte durch eine äquivalente Ersatzkraft $\mathbf{F}_R$ dargestellt werden, welche entlang einer Wirkungslinie L wirkt (siehe Abb. 60). Entsprechend muß eine externe Kraft $\mathbf{F}_{ext} = -\mathbf{F}_R$ entlang dieser Kraftlinie angreifen, um genau diese Rotationsbewegung auszulösen.

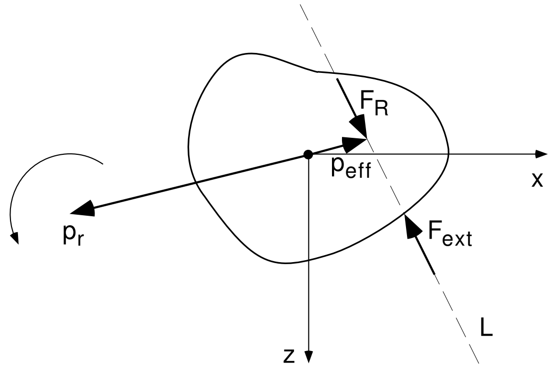


Abbildung 60: Das effektive Reibzentrum $\mathbf{p}_{eff}$. Die resultierende Reibungskraft F_R entlang der Kraftlinie L durch $\mathbf{p}_{eff}$ erzeugt dasselbe Reibmoment bezüglich des Rotationszentrums $\mathbf{p}_r$ wie die Reibungskräfte in der Berührfläche.

An welchem Punkt $\mathbf{p}_{eff} \in L$ die Kraft $\mathbf{F}_R$ ansetzt, spielt dabei keine Rolle. Wir wählen daher einen für die weitere Diskussion geschickten Punkt so, daß $\mathbf{p}_{eff}$ und $\mathbf{p}_r$ kollinear sind. Da die Kraft $\mathbf{F}_R$ bezüglich des Rotationszentrums $\mathbf{p}_r$ das Reibmoment $\mathbf{M}_R$ hervorrufen soll, gelten also die Beziehungen

$$\begin{aligned} (\mathbf{p}_{eff} - \mathbf{p}_r) \times \mathbf{F}_R &= \mathbf{M}_r, & (72)\\ \lambda \mathbf{p}_r &= \mathbf{p}_{eff}, & (73) \end{aligned}$$

wobei offensichtlich $\lambda < 1$ gelten muß.

Setzt man (73) in (72) ein, so folgen daraus die Beziehungen

$$\begin{aligned} (\lambda - 1)\mathbf{p}_r \times \mathbf{F}_R &= \mathbf{M}_r \\ |1-\lambda| \cdot |\mathbf{p}_r \times \mathbf{F}_R| &= |\mathbf{M}_r| \\ |1-\lambda| &= \frac{|\mathbf{M}_r|}{|\mathbf{p}_r \times \mathbf{F}_R|} \\ &\geq \frac{|\mathbf{M}_r|}{|\mathbf{p}_r| \cdot |\mathbf{F}_R|}. \end{aligned}$$

Mit der Dreiecksungleichung

$$\left| \int_A \frac{\mathbf{x} - \mathbf{p}_r}{|\mathbf{x} - \mathbf{p}_r|} p_A(\mathbf{x}) \, \mathrm{d}A \right| \leq \int_A \left| \frac{\mathbf{x} - \mathbf{p}_r}{|\mathbf{x} - \mathbf{p}_r|} p_A(\mathbf{x}) \right| \mathrm{d}A = \int_A p_A(\mathbf{x}) \, \mathrm{d}A = f_g$$

folgt aus (70) die Ungleichung

$$|\mathbf{F}_R| \leq \mu f_g. \tag{74}$$

Desweiteren folgt mit den Dreiecksungleichungen

$$\int_A |\mathbf{x} - \mathbf{p}_r| \, p_A(\mathbf{x}) \, \mathrm{d}A \quad \geq \quad \left| \int_A (\mathbf{x} - \mathbf{p}_r) p_A(\mathbf{x}) \, \mathrm{d}A \right|$$

$$\begin{aligned}
&\geq \left| \int\limits_A \mathbf{x} p_A(\mathbf{x})\, dA - \int\limits_A \mathbf{p}_r p_A(\mathbf{x})\, \mathrm{d}A \right| \\
&\geq |\mathbf{0} - \mathbf{p}_r f_g| \\
&\geq |\mathbf{p}_r|\; f_g
\end{aligned}$$

aus Gleichung (71) die Ungleichung

$$|\mathbf{M}_R| \geq \mu |\mathbf{p}_r| f_g. \tag{75}$$

Damit ergibt sich schließlich die Beziehung

$$|1-\lambda| \;\geq\; \frac{\mu|\mathbf{p}_r|f_g}{|\mathbf{p}_r|\mu f_g} = 1 \tag{76}$$

$$\lambda \;\leq\; 0, \tag{77}$$

welche aussagt, daß das Rotationszentrum $\mathbf{p}_r$ und der Schwerpunkt S in derselben Halbebene bezogen auf die Kraftlinie L liegen (vgl. Abb. 60). Eine externe Kraft $\mathbf{F}_{ext} = -\mathbf{F}_R$ erzeugt somit bezüglich des Schwerpunktes S ein Drehmoment, dessen Orientierung der Rotation mit $\boldsymbol{\omega}$ entspricht. Während bei zentralem Anschieben also eine Bewegung ohne Rotationskomponente erfolgt, ergibt sich bei nichtzentralem Anschieben eine Drehbewegung der zu erwartenden Orientierung. Das Entscheidende an diesem Ergebnis für unsere Fragestellung ist, daß diese Aussage unabhängig von der Geometrie der Kontaktregion gilt, und somit auch für den Fall einzelner Punktkontakte.

Rotation um den Schwerpunkt

Schließlich wollen wir die auftretenden Reibungskräfte und Reibungsmomente bei der ebenen Rotation um den Schwerpunkt betrachten. Analog zu (70, 71) erhält man mit $\mathbf{p}_r = \mathbf{0}$:

$$\mathbf{F}_R \;=\; -\mu\,\mathrm{sgn}(\omega_y)\hat{\mathbf{e}}_y \times \int\limits_A \frac{\mathbf{x}}{|\mathbf{x}|} p_A(\mathbf{x})\,\mathrm{d}A, \tag{78}$$

$$M_y \;=\; -\mu\,\mathrm{sgn}(\omega_y) \int\limits_A |\mathbf{x}| p_A(\mathbf{x})\,\mathrm{d}A. \tag{79}$$

Im flächigen Fall können wir konstanten Flächendruck $p_A = \frac{f_g}{A}$ annehmen, und erhalten

$$M_y \;=\; -\mu\,\frac{f_g}{A} \int\limits_A |\mathbf{x}|\,\mathrm{d}A,$$

wobei A den Flächeninhalt der Auflagefläche bezeichnet. Wir bestimmen dieses Integral nun allgemein für ein Dreieck gemäß Abb. 61. Beliebig geformte Auflageflächen können durch solche Dreiecke zusammengesetzt bzw. approximiert werden:

$$
\begin{aligned}
M_y &= -\mu \frac{2 f_g}{m L^2} \int_{z=0}^{L} \int_{x=0}^{mz} \sqrt{x^2+z^2}\, \mathrm{d}x\, \mathrm{d}z \\
&= -\mu \frac{2 f_g}{m L^2} \int_{z=0}^{L} \frac{1}{2} z^2 \left[m\sqrt{m^2+1} + \ln\left(m+\sqrt{m^2+1}\right)\right] \mathrm{d}z \\
&= -\mu\, f_g\, \frac{L}{3} \left[\sqrt{m^2+1} + \frac{\ln\left(m+\sqrt{m^2+1}\right)}{m} \right].
\end{aligned}
$$

Bei diskreten Auflagepunkten anstelle flächiger Berührung bestimmt sich M_y zu

$$
M_y = -\mu \sum_{i=1}^{N} f_i\, |\mathbf{x}_i|.
$$

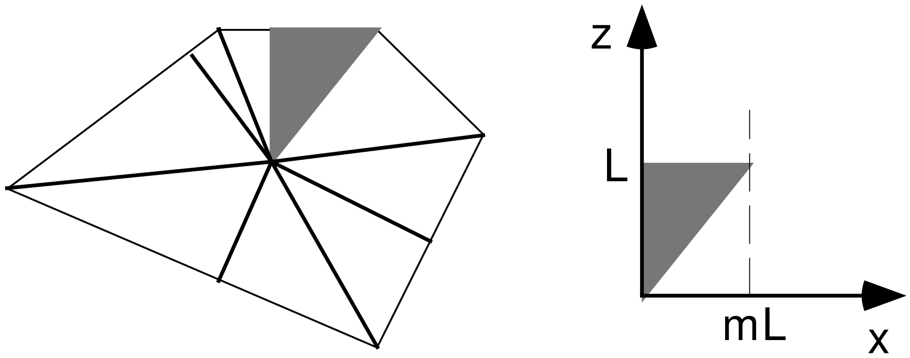


Abbildung 61: Das Integral über beliebige Polygone läßt sich durch Summation der (positiven oder negativen) Beiträge mehrerer rechtwinkliger Dreiecke wie dem schattiert markierten zusammensetzen.

Wir vergleichen nun die Reibmomente M_y bei flächiger Berührung mit denjenigen bei diskreten Auflagepunkten für verschiedene Geometrien. Als Beispiele wählen wir gleichseitige N-Ecke mit Umkreisradius R als Berührflächen. Die diskreten Auflagepunkte legen wir jeweils auf Eckpunkte der Grundfläche, und erhalten das Reibmoment M_y zu

$$
\begin{aligned}
M_y &= -\mu \sum_{i=1}^{N} f_i\, R \\
&= -\mu\, R \sum_{i=1}^{N} f_i \\
&= -\mu\, f_g\, R.
\end{aligned}
$$

Das Reibmoment M_y bei diskreten Auflagepunkten mit gleichen Schwerpunktabständen ist also unabhängig sowohl von der Anzahl N der Auflagepunkte als auch von deren genauer Lage!

Im Fall flächiger Berührung sind einige Werte in der folgenden Tabelle dargestellt:

Form der Grundfläche	M_y
gleichseitiges Dreieck	$-0.460\,\mu f_g R$
Quadrat	$-0.541\,\mu f_g R$
gleichseitiges Sechseck	$-0.608\,\mu f_g R$
Kreis	$-0.667\,\mu f_g R$

Man erkennt, daß das Reibmoment in diesen Beispielen mit peripheren Punktkontakten im radialen Abstand R etwa um den Faktor 2 zu groß ausfällt, verglichen mit der physikalisch korrekteren Behandlung in Form von Flächenkontakten.

Für die genannten Beispiele verschwindet bei konstantem Flächendruck die resultierende Reibungskraft $\mathbf{F}_R$ aus Symmetriegründen. Im Falle diskreter Auflagepunkte erhalten wir aus Gleichung (78) unter Berücksichtigung von (69) dasselbe Ergebnis, wenn alle Stützpunkte auf einem Kreis um den Ursprung liegen. Diese Voraussetzung ist gegeben, wenn die Stützpunkte auf beliebigen Eckpunkten eines gleichseitigen N-Ecks liegen.

Zusammenfassung

Insgesamt zeigen die obigen Ergebnisse, daß sich durch die Annahme peripherer punktförmiger Berührung mit Reibung in der Ebene im wesentlichen ein intuitiv richtiges Bewegungsverhalten ergeben sollte. Durch die periphere Lage der Berührpunkte ergeben sich im Vergleich mit idealer homogener Flächenberührung zu große Reibmomente, was z.B. beim nichtzentralen Anschieben in einer zu geringen Rotationsbewegung resultiert. Die Rotationsrichtung bleibt jedoch stets richtig.

8.3.2 Experimenteller Vergleich

Wie sich aus den theoretischen Untersuchungen ergibt, weichen die Reibungseigenschaften einer Kontaktregion, die durch wenige Punktkontakte modelliert ist, von denen einer flächig modellierten Kontaktregion ab. Sollen Flächenkontakte durch solche Punktkontakte approximiert werden, so ist eine Bewertung dieser Abweichung erforderlich. Die theoretischen Ergebnisse aus Abschnitt 8.3.1 lassen vermuten, daß das globale Bewegungsverhalten durch eine solche Approximation lediglich quantitativ abweicht, jedoch qualitativ unverändert, und somit intuitiv glaubwürdig bleibt. Da die dort gewonnenen Ergebnisse nur punktuell spezifische Bewegungssituationen vergleichen, kann daraus noch keine fundierte Aussage über das Bewegungsverhalten entlang ausgedehnter Bewegungsbahnen abgeleitet werden. In diesem Abschnitt soll daher das resultierende Bewegungsverhalten verschiedener Approximationen flächiger Kontaktregionen durch Punktkontakte anhand von Bewegungssimulationen bewertet werden. Dabei wird insbesondere der Einfluß von Anzahl und Anordnung der Punktkontakte auf das Bewegungsverhalten untersucht.

Als Testumgebung zur Untersuchung dieser Fragen verwenden wir Variationen der bereits in Abschnitt 8.1.1 beschriebenen Anordnung. Das Prisma wird von einer konstanten Kraft beschleunigt, die jeweils an einer festen Stelle des Prismas ansetzt, und nur geringfügig stärker ist als die gesamte Gleitreibungskraft bei rein translatorischer Bewegung. Dadurch wirken sich Trägheitskräfte nur sehr wenig aus, und das Bewegungsverhalten wird im wesentlichen durch die Reibungsverhältnisse bestimmt. Als Referenzmodell zur Nachbildung des idealisierten flächenhaften Kontaktes wählen wir, gemäß den Ergebnissen von Abschnitt 8.1.1, eine zufällig über die Auflagefläche verteilte Anordnung von 100 Kontaktpunkten.

Mit ganz ähnlichen Versuchsanordnungen wurde von Lynch [29] das Reibungsverhalten von Objekten untersucht, die durch eine punktförmige externe Kraft in eine ebene Gleitbewegung versetzt wurden. Ziel der dort beschriebenen Experimente war es, aus dem resultierenden Bewegungsverhalten verschiedene Reibungsparameter (Reibungskoeffizient, geometrische Anordnung von Auflagepunkten) zu rekonstuieren. Die uns interessierende Frage, inwieweit das Bewegungsverhalten davon abhängt, ob die Reibregionen flächig oder durch wenige Punktkontakte modelliert werden, wurde von Lynch jedoch nicht betrachtet.

Zentrales Anschieben mit stochastisch verteilten Kontaktpunkten

Die erste Serie von Experimenten untersucht die Frage, ob sich ein zentral angeschobener Körper bei verschiedenen Kontaktkonfigurationen geradlinig und rein translatorisch fortbewegt, oder ob er seitlich abdriftet. Die Versuchsanordnung und das theoretisch erwartete Verhalten sind in Abb. 62 schematisch dargestellt.

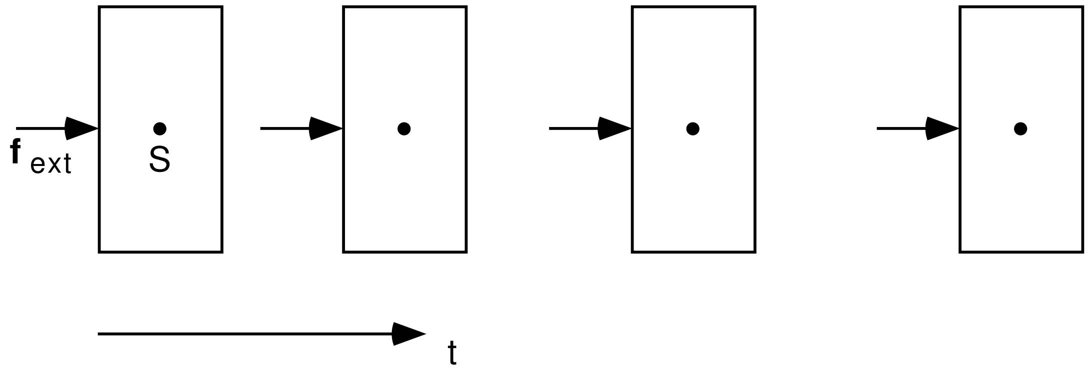


Abbildung 62: Prinzipielle Darstellung verschiedener Phasen der Bewegung eines Objektes bei zentralem Anschieben.

Es sollte noch bemerkt werden, daß der gewählte Ansatzpunkt der externen Kraft hinter dem Schwerpunkt liegt, so daß bereits die geringste seitliche Abweichung des Prismas von der geradlinigen Bahn infolge einer Rotation zu einem zusätzlichen Drehmoment führt, welches diese Rotation zunehmend verstärkt. Diese Anordnung reagiert somit sehr empfindlich auch auf kleinste Abweichungen vom idealen Bewegungsverlauf.

Wir interessieren uns besonders für das Bewegungsverhalten bei verschiedenen unsymmetrischen Anordnungen von jeweils drei Kontaktpunkten. Dazu wählen wir die drei Kontaktpunkte jeweils zufällig innerhalb der rechteckigen Auflagefläche des Prismas. Als einzige Randbedingung fordern wir, daß der Lotfußpunkt S des Körperschwerpunktes auf die Berührebene im Inneren des durch die drei Kontaktpunkte gebildeten Dreieckes liegt (vgl. Abb. 63).

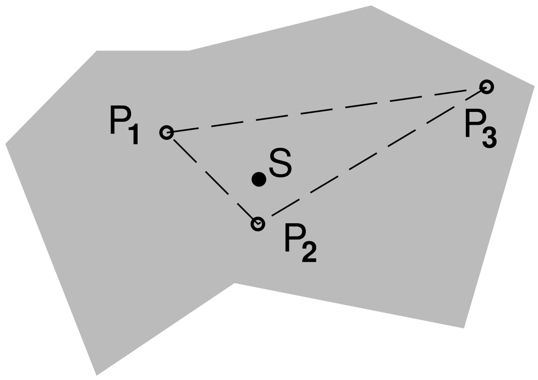


Abbildung 63: Approximation einer Kontaktfläche (schattiert dargestellt) durch drei Punktkontakte (P_1, P_2, P_3). Der Einfluß dieser Approximation auf das Reibungsverhalten kann anhand von zufällig innerhalb der Kontaktregion gewählten Kontaktpunkttripeln untersucht werden. Zur Gewährleistung des statischen Gleichgewichts müssen die Kontaktpunkte allerdings so gewählt werden, daß der Lotfußpunkt S des Schwerpunktes im Inneren des Dreieckes (P_1, P_2, P_3) liegt.

Die experimentellen Ergebnisse bestätigen die theoretischen Resultate vollkommen. In allen Fällen bleibt die Bewegung des Objektes zunächst vollkommen geradlinig, auch wenn die Anordnung der Punktkontakte stark unsymmetrisch ist. Erst nach mehreren hundert Simulationsintervallen (je 10ms) haben sich kleinste numerische Fehler soweit akkumuliert, daß eine sichtbare Abweichung von der Ideallinie erkennbar wird (vgl. Abb. 64). Das Referenzmodell mit 100 Kontaktpunkten zeigt selbstverständlich ebenfalls dieses Verhalten. Die Bahn des Schwerpunktes weicht in beiden Fällen nur minimal voneinander ab (vgl. Abb. 65).

Aus dem bei zentralem Anschieben rein translatorischen Bewegungsverhalten verschiedener Prismen mit allen zufällig gewählten Kontaktpunkttripeln läßt sich schließen, daß dieses Verhalten allgemein gilt, und insbesondere auch für alle im Rahmen der Kontaktsimulation auftretenden minimalen Punktkontaktkonfigurationen zu erwarten ist.

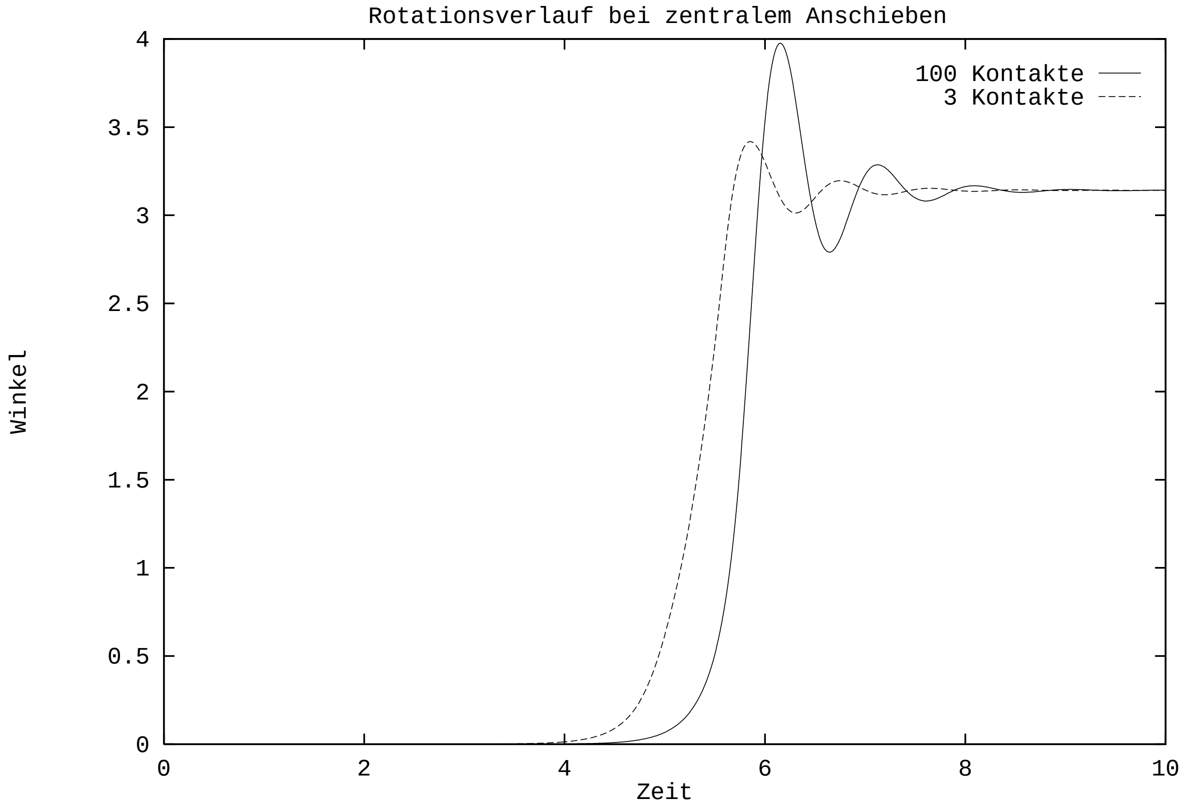


Abbildung 64: Beispielhafter Verlauf des Rotationswinkels der Bewegung eines zentral angeschobenen Körpers bei ebener Gleitbewegung. Verglichen wird der Verlauf bei drei Punktkontakten mit dem Verlauf bei vielen Kontakten, welche den Fall flächiger Reibung approximieren.

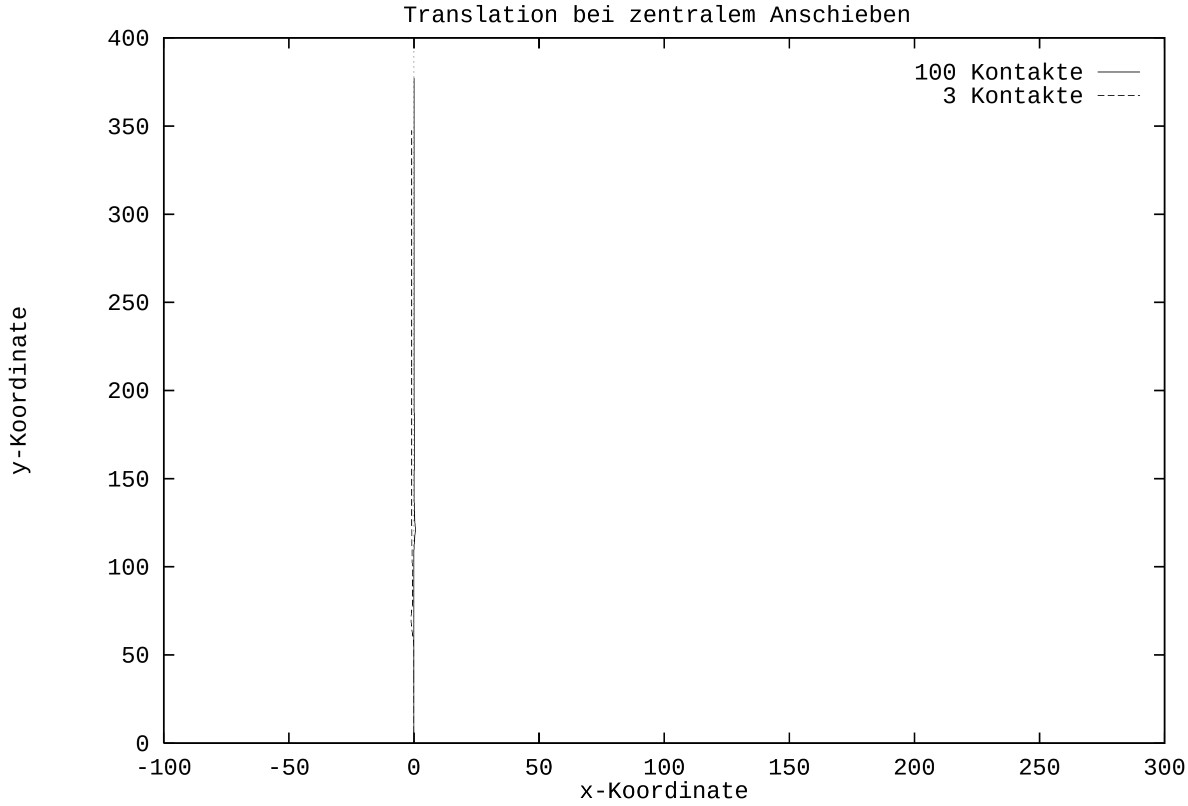


Abbildung 65: Die translatorische Bewegung des Körperschwerpunktes der in Abb. 64 beschriebenen Bewegung.

Nichtzentrales Anschieben

Tritt an einem Flächenkontakt eine Gleitbewegung mit rotatorischer Bewegungskomponente auf, so zeigen die theoretischen Ergebnisse, daß die an der Kontaktfläche wirkenden Reibmomente sich bei verschiedenen Anordnungen der Kontaktpunkte unterscheiden. Liegen die wirksamen Kontaktpunkte stets am Rande der Kontaktregion, so sind größere Reibmomente zu erwarten als bei gleichmäßiger Verteilung innerhalb dieser Region. Umgekehrt erwarten wir, daß bei vorgegebenen externen Kräften und Momenten die rotatorische Bewegungskomponente mit randständigen Punktkontaktkonfigurationen stärker gedämpft wird als bei gleichmäßiger Verteilung in der Fläche.

Um dies zu prüfen wählen wir eine ähnliche Versuchsanordnung wie im letzten Abschnitt beschrieben und setzen die externe Kraft jedoch seitlich versetzt an, so daß eine Gleitbewegung mit translatorischen und rotatorischen Bewegungskomponenten einsetzt (siehe Abb. 66).

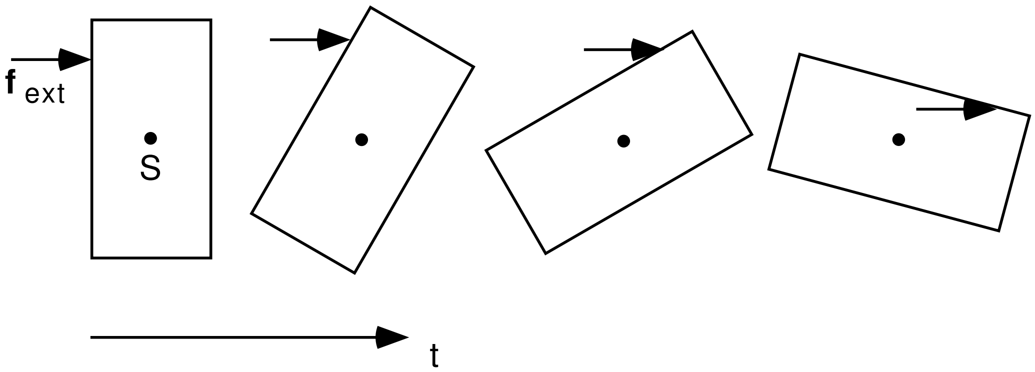


Abbildung 66: Prinzipielle Darstellung verschiedener Phasen der Bewegung eines Körpers bei nichtzentralem Anschieben.

Mit dieser Anordnung vergleichen wir nun das resultierende Bewegungsverhalten bei verschiedenen Punktkontaktkonfigurationen. Wir sind uns dabei bewußt, daß mit solchen Simulationen die Vielfalt möglicher Bewegungsabläufe keinesfalls erschöpfend untersucht und verglichen werden kann. Die gewählte Anordnung erscheint jedoch repräsentativ zur Bewertung des Ausmaßes der auftretenden Abweichungen zwischen realen Flächenkontakten und deren Approximation durch minimale Punktkontaktkonfigurationen.

In Abb. 67 ist der zeitliche Verlauf des Rotationswinkels relativ zur Ausgangslage für einen typischen Bewegungsablauf gemäß Abb. 66 dargestellt. Dabei zeigt die durchgezogene Linie den Verlauf, wenn die flächige Reibung durch viele (100) Punktkontakte approximiert wird. Die durchbrochene Linie zeigt den entsprechenden Verlauf bei Verwendung von nur drei Punktkontakten, die an Eckpunkten des Prismas liegen.

Beide Kurven zeigen prinzipiell denselben Ablauf. Durch die exzentrisch angreifende externe Kraft wird das Objekt translatorisch und rotatorisch beschleunigt. Durch die Massenträgheit dreht sich der Körper bis zu einem maximalen Winkel, und pendelt sich dann auf den Drehwinkel ein, bei dem die externe Kraft kein Drehmoment bezüglich des Schwerpunktes erzeugt. Der genaue Verlauf hängt bei einer gegebenen Punktkontaktkonfiguration vorwiegend vom Betrag der externen Kraft, von der Exzentrität des Kraftansatz-

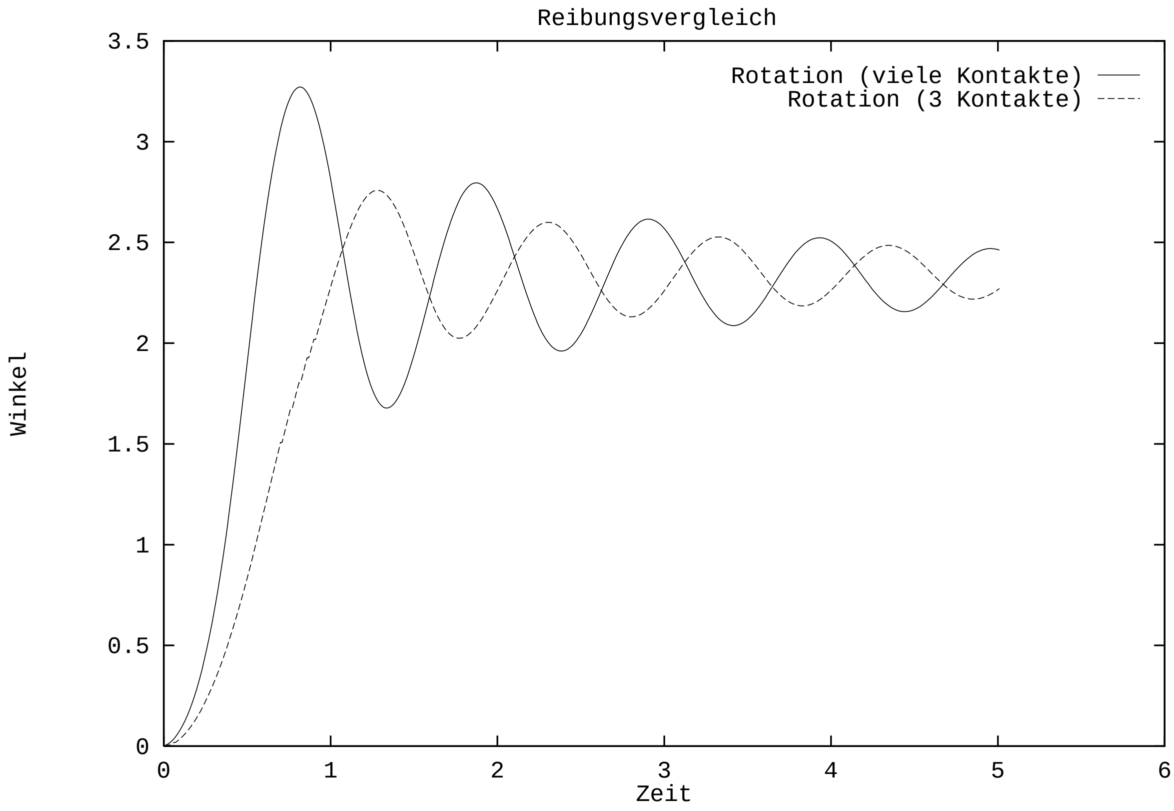


Abbildung 67: Beispielhafter Verlauf des Rotationswinkels der Bewegung eines exzentrisch angeschobenen Körpers bei ebener Gleitbewegung. Verglichen wird der Verlauf bei drei Punktkontakten mit dem Verlauf bei vielen Kontakten, welche den Fall flächiger Reibung approximieren.

punktes, sowie vom Gleitreibungskoeffizienten ab. Der qualitative Verlauf bleibt jedoch stets gleich.

Die Unterschiede zwischen den beiden dargestellten Kurven liegen vorwiegend in der Amplitude der Schwingbewegung. Bei wenigen Punktkontakten, die alle am Rand der Auflagefläche angeordnet sind, verläuft die Rotationsbewegung stärker gedämpft, da die Reibungskräfte ein im Mittel größeres bremsendes Drehmoment erzeugen als bei flächiger Reibung. Dies entspricht genau den theoretischen Ergebnissen aus Abschnitt 8.3.1.

Neben dem Verlauf der Rotation soll auch die translatorische Bewegung des Körpers bei beiden Kontaktmodellen verglichen werden. In Abb. 68 sind jeweils die Trajektorien des Körperschwerpunktes für die selben Bedingungen wie in Abb. 67 aufgetragen. Es ist zu erkennen, daß abgesehen von der ersten Beschleunigungsphase die translatorische Bewegung in beiden Fällen genau in Richtung der externen Antriebskraft erfolgt. Auch dies entspricht dem erwarteten Ergebnis.

Das unterschiedliche Bewegungsverhalten soll nun durch Variation der Anordnung der Punktkontakte sowie der Exzentrizität des Kraftansatzpunktes weiter untersucht werden. Wie im letzten Abschnitt bestimmen wir auch hier für jeden Versuchslauf drei zufällig angeordnete Punktkontakte, und simulieren die Bewegung bei exzentrischem Anschie-

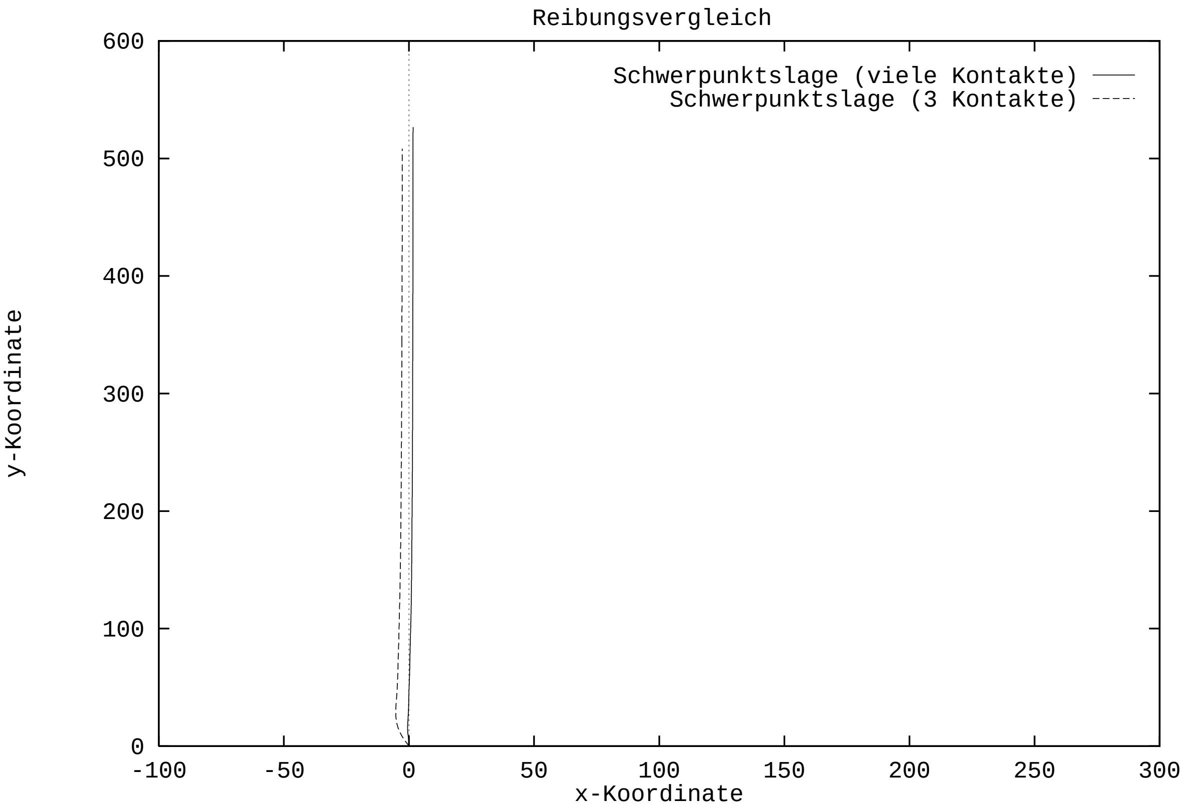


Abbildung 68: Die translatorische Bewegung des Körperschwerpunktes der in Abb. 67 beschriebenen Bewegung.

ben. Um zu einer übersichtlichen Darstellung der Ergebnisse zu kommen, wählen wir als Kenngröße das Maximum des Rotationswinkels, welches nach der ersten Beschleunigungsphase erreicht wird. Diese Kenngröße ist in Abb. 69 in Abhängigkeit von der Exzentrizität des Kraftansatzpunktes dargestellt. Dabei bedeutet der Exzentrizitätswert 1, daß die anschiebende Kraft am äußersten Rand des Körpers angreift, und ein Wert von 0 entspricht zentralem Anschieben. Die Exzentrizitätswerte sind dabei normiert auf die halben Seitenlängen des umschreibenden Rechtecks der polygonalen Kontaktregion. Zusätzlich zu diesen stochastisch gewählten Kontaktpunkttripeln wurden der Fall flächiger Reibung, sowie der Fall dreier randständiger Punktkontakte ebenfalls mit aufgetragen.

Das Resultat dieser experimentellen Untersuchung bestätigt die Ergebnisse der theoretischen Betrachtungen. Da bei den randständigen Punktkontakten die größten Reibradien und somit die größten Reibmomente wirken, wird bei ihnen die Rotationsbewegung am stärksten gehemmt. Dies äußert sich darin, daß die beobachteten maximalen Winkel am kleinsten ausfallen, und zwar unabhängig von der jeweiligen Exzentrizität des Kraftansatzpunktes. Werden jeweils drei Punktkontakte zufällig gewählt, so variiert die auftretende Rotationsbewegung deutlich, entsprechend der Variation der jeweiligen Reibradien. Diese Fälle bilden in Abb. 69 ein Band oberhalb der Winkelwerte randständiger Kontaktpunkte. Der Fall flächiger Reibung, der in diesem Experiment durch jeweils 100 zufällig in der Kontaktregion angeordnete Punktkontakte approximiert wurde, weist stets kleinere

Reibmomente auf als der Fall randständiger Punktkontakte, entsprechend ergeben sich größere Rotationswinkel. Die hierbei beobachteten Werte liegen im mittleren Bereich des Bandes, welches von stochastisch gewählten Kontaktpunkttripeln belegt wird.

Aus Abb. 69 ergibt sich, daß bei kleineren Exzentrizitäten eine stärkere Rotationsbewegung erfolgt. Dieses Phänomen läßt sich wie folgt erklären. Die anschiebende Kraft bewirkt nicht nur die Drehbewegung des betrachteten Körpers, sondern beschleunigt diesen auch translatorisch. Bei kleineren Exzentrizitäten des Kraftansatzpunktes ergibt sich zunächst ein kleineres Drehmoment, so daß der translatorische Anteil der Gesamtbewegung größer ausfällt. Wir betrachten nun die Reibungskräfte innerhalb der Kontaktebene als zweidimensionales Kraftfeld. Bei reiner Rotation ist dieses Kraftfeld konzentrisch, und bewirkt ein hohes Drehmoment. Bei einer Zunahme des translatorischen Bewegungsanteils nimmt die Krümmung der Feldlinien ab, das Feld wird zunehmend homogen, und das resultierende Reibmoment nimmt ab. Genau dies geschieht also bei kleineren Exzentrizitäten des Kraftansatzpunktes, und daher wird in diesen Fällen die Rotationsbewegung weniger abgebremst.

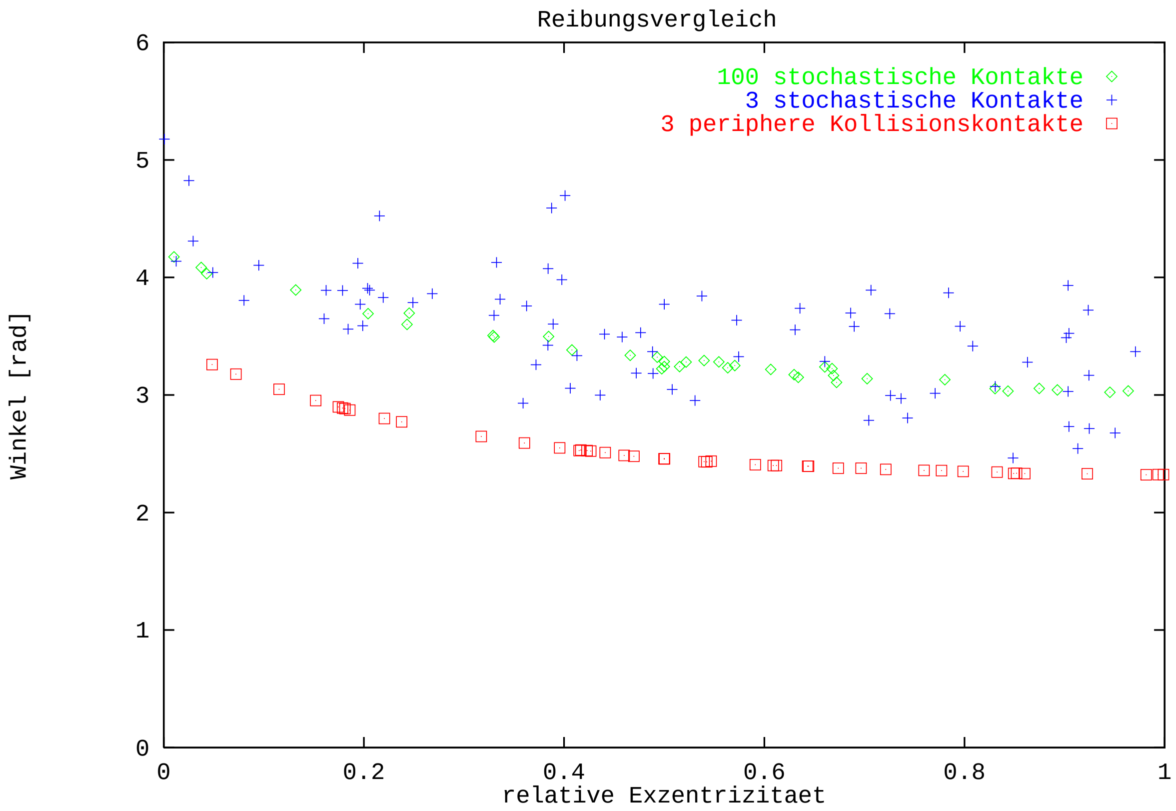


Abbildung 69: Darstellung des maximalen Rotationswinkels der Bewegung von exzentrisch angeschobenen Körpern als Kenngröße zum Vergleich verschiedener Kontaktkonfigurationen (zur Diskussion siehe Text).

9 Beschreibung des Simulationssystems

Die Implementierung der in den vorangegangenen Kapiteln beschriebenen Verfahren zur Bewegungssimulation mit Hinderniskontakten bildet eines der Module eines Systems zur Realisierung von Anwendungen in simulierten Umgebungen. In den folgenden Abschnitten soll nun sowohl dieses Modul zur Bewegungssimulation selbst, als auch seine Einbindung in das Gesamtsystem dargestellt werden.

9.1 Die VR-Plattform DBView

Am Forschungsinstitut der Daimler-Benz AG in Ulm wurde im Verlauf mehrerer Jahre *DBView*, eine universelle Plattform für Anwendungen virtueller Realität (VR) entwickelt. Es handelt sich bei *DBView* um ein modulares Softwaresystem, welches eine Vielzahl verschiedenster Funktionalitäten umfaßt, die zur Realisierung konkreter Anwendungen bedarfsgerecht konfiguriert werden können. Zu den verfügbaren Bestandteilen gehören

- Verfahren zur Gestaltung virtueller Umgebungen
- Schnittstellen zur Interaktion des Benutzers mit virtuellen Komponenten
- Verfahren zur 3D-Meßtechnik
- die Ansteuerung verschiedenster Visualisierungseinrichtungen
- Methoden zur Animation virtueller Objekte mit automatischen Bewegungsabläufen
- Verfahren zur Physiksimulation

Die jüngst entwickelten Verfahren zur Physiksimulation, zu denen neben der in der vorliegenden Arbeit behandelten Bewegungssimulation auch Module zur schnellen Kollisionserkennung (Eckstein [17]) und zur Dynamiksimulation (Sauer [39]) gehören, tragen wesentlich dazu bei, simulierte Umgebungen mit realistischen Verhaltensweisen zu unterlegen. Um die Einbindung solcher Module in das Gesamtsystem zu veranschaulichen, soll nun kurz auf die Architektur von *DBView* eingegangen werden.

9.1.1 Die Architektur von *DBView*

Der Kern von *DBView* setzt auf einem hierarchischen Szenengraphen auf, welcher die Bestandteile einer virtuellen Umgebung repräsentiert. Hierzu zählen zunächst die Beschreibungen der Geometrie aller Szenenobjekte, sowie deren Farben und Oberflächeneigenschaften. Ein weiterer wesentlicher Bestandteil sind Objekttransformationen, welche die Anordnung einzelner Objekte innerhalb der Szene und die Objektbewegung ermöglichen. Auf alle Bestandteile des Szenengraphen kann programmtechnisch zugegriffen werden. Auf diese Weise kann auch die Bewegungssimulation Einfluß auf die Transformation interaktiv kontrollierter Szenenobjekte nehmen. Die Verwaltung des Szenengraphen beruht auf der Software-Bibliothek OpenInventor [47], welche wiederum auf der standartisierten Graphikbibliothek OpenGL [37] aufsetzt. Diese Bibliotheken stellen auch Basismechanismen zur Visualisierung virtueller Umgebungen bereit.

Der Ablauf von VR-Simulationen in *DBView* wird durch Auslösen von Software-Events gesteuert. Insbesondere löst auch jede Benutzereingabe solche Events aus, welche in einer Eventqueue gesammelt werden. *DBView* arbeitet diese Events schließlich in einer zentralen Hauptschleife ab.

Eingehende Events lösen in der Regel Aktionen aus, welche Einfluß auf den Szenengraphen nehmen. Dies kann ein Wechsel des Beobachterstandortes innerhalb der Szene sein, eine Objektbewegung oder eine sonstige Manipulation von Eigenschaften eines Szenenobjektes. Weitere Auslöser von Änderungen des Szenengraphen erzeugt die Animationssteuerung, welche vordefinierte Abläufe, insbesondere Bewegungsabläufe, realisiert. Sind alle anstehenden Events berücksichtigt worden, wird die Szene aus der Perspektive simulierter Kameras abgebildet und schließlich dem Anwender auf dem Bildschirm oder einer sonstigen Visualisierungseinrichtung dargestellt.

Über verschiedene Eingabegeräte ist es einem Benutzer nun möglich, Bewegungsvorgaben für virtuelle Objekte zu spezifizieren. In der einfachsten Variante werden diese im Rahmen der erwähnten Eventabarbeitung direkt umgesetzt. Dabei folgen die manipulierten Objekte exakt den vorgegebenen Bewegungsbahnen, indem die Lagetransformationen dieser Objekte im Szenengraphen entsprechend verändert werden.

9.1.2 Die Einbindung der Bewegungssimulation

Diese einfache Vorgehensweise berücksichtigt jedoch nicht Randbedingungen, welche aus den Merkmalen und Verhaltensweisen realer Objekte ableitbar sind. Hierzu gehört insbesondere die Eigenschaft massiver Objekte, sich nicht gegenseitig zu durchdringen, sondern im Falle einer Berührung eine Ausweichbewegung zu vollziehen, oder die Bewegung ganz zu beenden. Solche Verhaltensweisen erfordern also die Bestimmung einer neuen Objektbewegung, wenn die ursprünglich vorgesehene Bewegung mit den physikalischen Randbedingungen nicht vereinbar ist. Innerhalb von *DBView* stellt die Bewegungssimulation daher ein Filter dar, welches unzulässige Bewegungsvorgaben korrigiert.

Die Bewegungssimulation wirkt jedoch nicht nur auf Szenenobjekte, welche unmittelbar vom Benutzer gesteuert werden. Häufig tritt der Fall auf, daß ein vom Benutzer kontrolliertes Szenenobjekt auf weitere Szenenobjekte einwirkt, und deren Bewegung ebenfalls beeinflußt. Diese Funktionalität wurde im Rahmen der gemeinsamen Simulation von Kontaktclustern in Abschnitt 3.5 behandelt.

Eine weitere Funktionalität stellt schließlich die physikalisch korrekte Simulation der Eigendynamik bewegter Szenenobjekte dar. Mit diesem Themenfeld wollen wir uns jedoch hier nicht weiter befassen, sondern gehen nun auf die Struktur des Modules zur Bewegungssimulation mit Hinderniskontakten näher ein.

9.2 Das Modul zur Bewegungssimulation mit Hinderniskontakten

Die Bewegungssimulation unter Berücksichtigung von Hinderniskontakten und anderen Bewegungseinschränkungen läßt sich als Filter auffassen, welches die aktuelle Ist-Bewegung von simulierten Objekten sowie eine extern vorgegebene Sollbewegung als Eingangsgrößen akzeptiert, und daraus eine neue Ist-Bewegung bestimmt (Abb. 70).

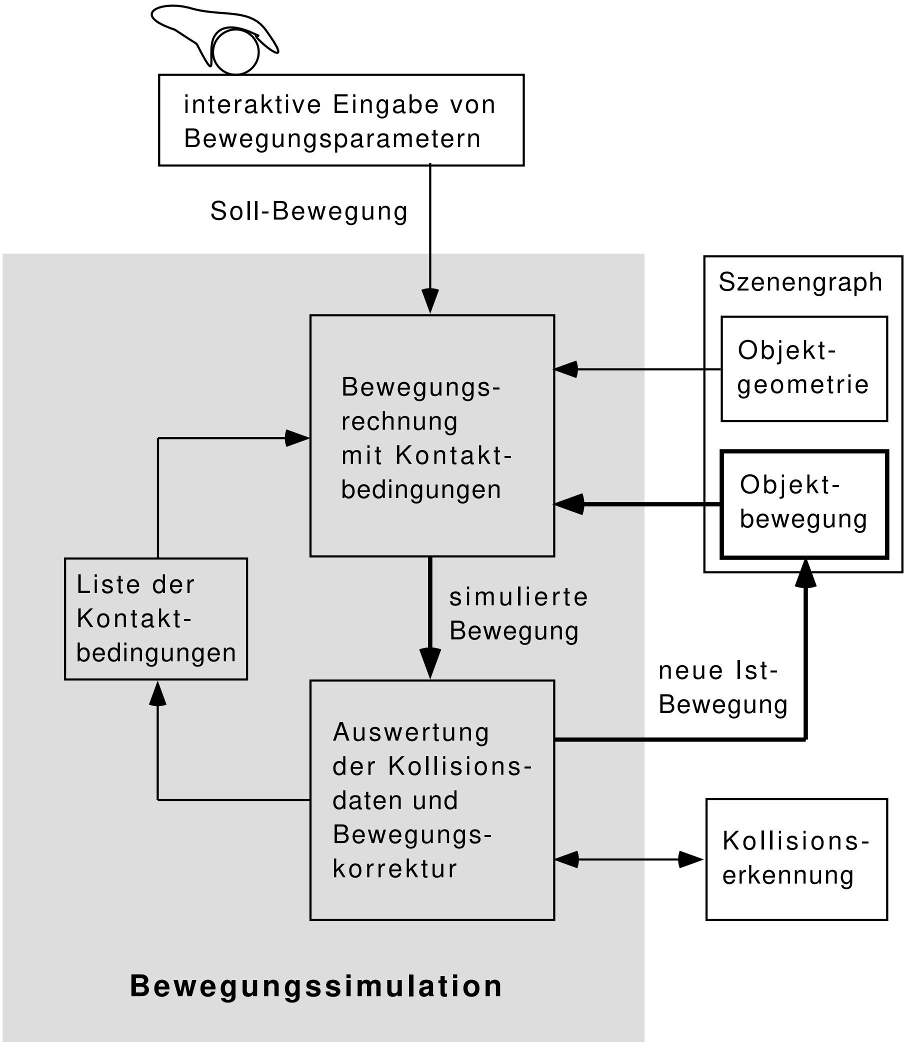


Abbildung 70: Struktur der Bewegungssimulation mit Hinderniskontakten und deren Schnittstellen zum umgebenden System.

Dabei wird zunächst aus der alten Ist-Bewegung, der vorgebenen Soll-Bewegung und den aktuell registrierten Kontaktbedingungen eine neue simulierte Bewegung bestimmt. In einem zweiten Schritt wird diese Bewegung nun durch Aufruf eines externen Modules daraufhin untersucht, ob diese Bewegung im betrachteten Zeitintervall zu Kollisionen führt. Ist dies nicht der Fall, so kann die simulierte Bewegung direkt als neue Ist-Bewegung übernommen werden. Tritt dagegen eine neue Kollision zum Zeitpunkt $t_c \in [t, t + \Delta t]$ auf, so wird die neue simulierte Bewegung nur bis zu diesem Zeitpunkt t_c durchgeführt. Die hierbei eingenommene Objektlage wird nun als neue Ist-Position übernommen. Schließlich wird die Menge der Kontaktbedingungen den aktuellen Verhältnissen angepaßt.

Der erste dieser beiden Schritte, die eigentliche Bewegungssimulation, ist in Abb. 71 noch einmal detaillierter dargestellt.

Zunächst werden, unter Berücksichtigung der Massenträgheit des betrachteten Objektes, als Hilfsgrößen die erforderlichen externen Kräfte F_{ext} und Momente M_{ext} bestimmt, welche die aktuelle Ist-Bewegung in die extern vorgegebene Sollbewegung überführen. Aus

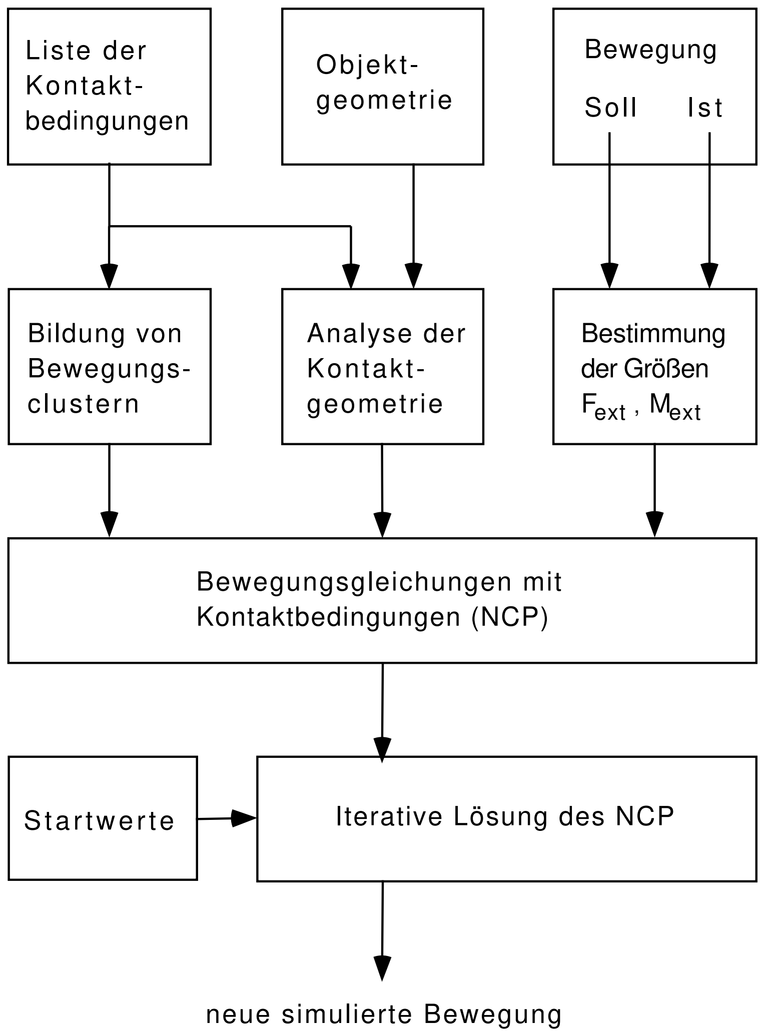


Abbildung 71: Struktur des Teilmodules zur Bewegungsrechnung mit Kontaktbedingungen.

der Liste der bestehenden Kontaktbedingungen werden mittels einer Zusammenhangsanalyse des Kontaktgraphen die bestehenden Kontaktcluster gebildet. Ebenfalls aus der Liste der bestehenden Kontaktbedingungen werden unter Berücksichtigung der Objektgeometrien und der aktuellen Objekttransformationen die aktuellen Kontaktgeometrien berechnet. Hierbei werden gegebenenfalls auch die zur Reibungssimulation erforderlichen Modifikationen der Kraftrichtungen an den Kontaktpunkten vorgenommen. Basierend auf der Analyse der Kontaktcluster und der Kontaktgeometrien wird nun für jedes der Kontaktcluster das Gleichungssystem zur Bestimmung der unbekannten Kontaktkräfte aufgestellt. Die zur iterativen Lösung dieses Gleichungssystems erforderlichen Startwerte werden gemäß einer der in Abschnitt 6.3 diskutierten Vorgehensweisen bestimmt. Damit sind nun alle Voraussetzungen erfüllt, um mit diesem Gleichungssystem, welches die Form eines nichtlinearen Komplementaritätsproblems (NCP) annimmt, die neuen Bewegungsgrößen für die Objekte des betrachteten Kontaktclusters zu bestimmen.

9.3 Interaktionsschnittstellen

Die interaktive Eingabe der Sollbewegung zur Steuerung simulierter Objekte erfordert ein Eingabegerät, welches die Benutzeraktionen in Bewegungsparameter umsetzt, die dann von der Bewegungssimulation verwertet werden können. Da bei der Manipulation simulierter Objekte 6 Freiheitsgrade zu spezifizieren sind, ist zur intuitiven Bewegungssteuerung ein Interface mit ebenfalls 6 Freiheitsgraden erforderlich. Als solche Eingabegeräte sind sowohl Kraftsensoren (beispielsweise SpaceMouse oder SpaceBall) geeignet, als auch Bewegungssensoren (wie beispielsweise mit Datenhandschuhen). Beide Alternativen unterscheiden sich jedoch wesentlich, sowohl hinsichtlich der Handhabungsweise als auch hinsichtlich der Versorgung der Bewegungssimulation mit Eingabeparametern.

9.3.1 Interaktion durch Kraftsensoren

Die Eingabe von Bewegungsparametern über Kraftsensoren erfolgt dadurch, daß der Benutzer einen etwa handtellergroßen Steuerknopf ergreift und auf diesen Kräfte und Momente ausübt. Der Steuerknopf gibt dabei nur minimal nach (wenige Millimeter und Winkelgrade). Die aufgebrachten Kräfte und Momente werden von Sensoren gemessen und stehen als Parameter für die Simulation zur Verfügung.

Zunächst bietet es sich an, diese Kraftparameter unmittelbar als externe Kräfte in der Bewegungssimulation einzusetzen. Dadurch würde der Zwischenschritt entfallen, eine Sollbewegung explizit zu spezifizieren. Stattdessen würden die externen Kräfte direkt zur Modifikation der jeweiligen Ist-Bewegung verwendet. Diese Vorgehensweise erlaubt somit zwar eine sehr direkte Steuerung der Objektbeschleunigung. Sie ist jedoch weniger geeignet, präzise Objektgeschwindigkeiten und Objektpositionen zu definieren. Mit unmittelbarer Kraftsteuerung ist es beispielsweise schwierig, ein sich bewegendes Objekt ganz zur Ruhe zu bringen. Noch schwieriger wird es, ein gesteuertes Objekt präzise an eine Zielposition zu bewegen und dort zur Ruhe zu bringen.

Eine sehr viel einfacher zu bedienende Interaktionsschnittstelle ergibt sich, wenn die gemessenen Kraftwerte als Geschwindigkeiten, und die gemessenen Momente als Winkelgeschwindigkeiten interpretiert werden. Diese Werte fließen dann direkt als Sollbewegung in die Simulation ein (vgl. Abb 70). Wird der Steuerknopf losgelassen, kommt das kontrollierte Objekt sofort zur Ruhe.

Die präzise Navigation mit diesem Interaktionsparadigma hat sich bereits in vielen Alltagssituationen bewährt, beispielsweise bei der Steuerung von Fahrzeugen im Straßenverkehr. Auch hier werden die auf Gaspedal und Lenkrad ausgeübten Kräfte beziehungsweise Auslenkungen in Geschwindigkeiten umgesetzt.

9.3.2 Interaktion durch Bewegungssensoren

Die Eingabe von Bewegungsparametern über Bewegungssensoren entspricht weitgehend unserer Alltagserfahrung. Hierbei wird durch geeignete Verfahren der 3D-Meßtechnik die räumliche Lage und Orientierung der Hand des Benutzers gemessen, meist in Verbindung mit einem Datenhandschuh. Auf diese Weise kann die Sollbewegung eines interaktiv gesteuerten Objektes unmittelbar und intuitiv vorgegeben werden.

Zunächst erscheint dieses Interaktionsparadigma ideal zur Manipulation simulierter Objekte. Es ist jedoch zu beachten, daß die Bewegung der menschlichen Hand durch einen Regelmechanismus kontrolliert wird, welcher nicht ausschließlich positionsbasiert ist, sondern auch kraftbasiert. Dies sollen zwei Beispiele illustrieren. Das erste Beispiel besteht in der Erfahrung, daß die menschliche Hand kaum vollkommen ruhig gehalten werden kann, ohne sich an einem feststehenden Objekt anzulehnen. Als zweites Beispiel betrachten wir die in Abb. 72 dargestellte Situation. Ein stabförmiges Objekt A (beispielsweise ein Bleistift) werde an einem Ende gehalten, und dort mit einer externen Kraft F_{ext} und einem externen Moment M_{ext} beaufschlagt. Am anderen Ende gleite der Stab entlang einer feststehenden Fläche eines zweiten Objektes B. Die präzise Steuerung des Stabes durch die menschliche Hand wird wesentlich mitbedingt durch die Reaktionskraft F_k, welche von der Fläche auf den Stab zurück wirkt.

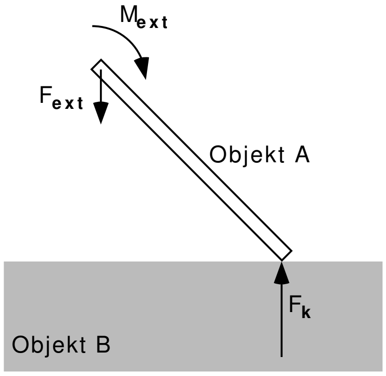


Abbildung 72: Steuerkräfte F_{ext} und M_{ext} und die Reaktionskraft F_k bei der interaktiven Objektmanipulation mit Hinderniskontakten.

Im Falle interaktiver Simulationen stehen praktikable Mechanismen zur Kraftrückwirkung von der simulierten Umgebung auf die steuernde Hand derzeit jedoch nicht zur Verfügung. Berührt der Stab die festehende Fläche im Verlaufe einer Bewegungssimulation, so verhindern die in dieser Arbeit beschriebenen Simulationsmechanismen zwar die Durchdringung von Stab und Fläche. Die steuernde reale Hand erfährt die in der Simulation bestehenden Bewegungsbeschränkungen ohne Kraftrückwirkungsmechanismen jedoch nicht unmittelbar, und kann sich unbeschränkt weiterbewegen. In dieser Situation können daher Diskrepanzen zwischen realer Handbewegung und simulierter Objektbewegung auftreten.

10 Anwendungsbeispiele

Die in der bisherigen Darstellung beschriebenen Verfahren zur Simulation interaktiv bewegter Objekte mit Hinderniskontakten sollen nun noch anhand einiger Beispielanwendungen illustriert werden. Dabei werden auch spezifische Aspekte beleuchtet, die im jeweiligen Beispiel besonders hervortreten.

10.1 Beispiel 1: Flugzeug

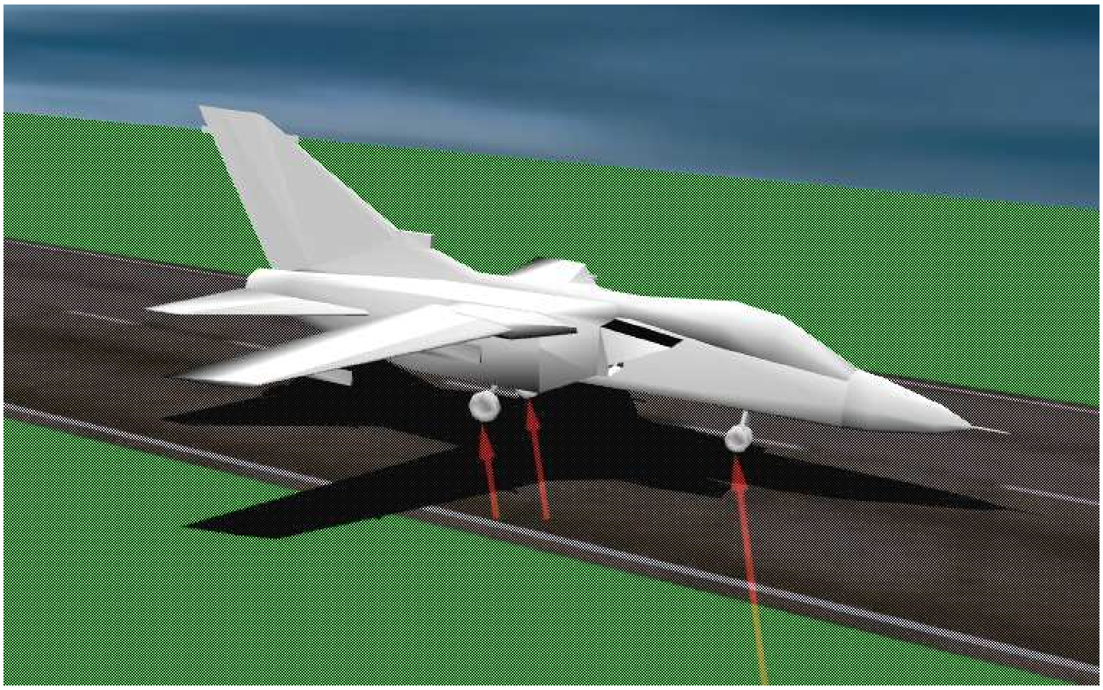


Abbildung 73: Beispiel 1: Flugzeug.

Bei der in Abb. 73 dargestellten Beispielanwendung handelt es sich um ein Flugzeugmodell, welches mit einem Kraftsensor in allen 6 Bewegungsfreiheitsgraden interaktiv gesteuert werden kann. Dieses Flugzeugmodell kann nun interaktiv auf dem Modell einer Rollbahn aufgesetzt werden, ohne mit einem seiner Teile darin einzutauchen. Dieses Beispiel zeigt die Wirkungsweise der vorgestellten Verfahren sehr deutlich. Insbesondere spiegeln die eingeblendeten Kraftpfeile die jeweils auftretenden Kontaktkräfte und Reibungskräfte wieder.

Dieses Flugzeugbeispiel wurde vielfach von ungeübten Benutzern erfolgreich getestet. Es zeigte sich, daß die interaktive Handhabung simulierter Objekte mit Hinderniskontakten in der realisierten Weise auch ohne Übung sofort gelingt, und von den Testpersonen als einfach handhabbar und sehr intuitiv empfunden wird. Da die durchzuführende Handhabungsaufgabe keine besondere Geschicklichkeit erfordert, kommt auch die fehlende Kraftrückkopplung hierbei nicht zum Tragen.

10.2 Beispiel 2: Autoradio

Abbildung 74: Beispiel 2: Radioeinbau.

In diesem zweiten Beispiel besteht die interaktive Handhabungsaufgabe darin, mit dem Datenhandschuh das virtuelle Modell eines Autoradios zu ergreifen, und in den dafür vorgesehenen Schacht in der Mittelkonsole eines simulierten Fahrzeuginnenraums einzusetzen (siehe Abb. 74). Dazu wird in der virtuellen Szene ein synthetisches Handmodell verwandt, welches die Bewegungen des Datenhandschuhes einschließlich der Fingerstellungen nachvollzieht. Das Ergreifen des Radios erfolgt dabei durch eine einfache Gestenerkennung. Wenn die Spitzen von Daumen und Zeigefinger des Datenhandschuhes sich berühren, während sie die Oberfläche des Radios schneiden, wird die Bewegung des Radios an die Handbewegung gekoppelt. Lösen sich die Spitzen von Daumen und Zeigefinger wieder voneinander, so löst sich auch das gegriffene Objekt wieder von der Hand.

Die Aufgabe, das Radio in den Schacht einzuführen, ist wesentlich anspruchsvoller, als die im ersten Beispiel beschriebene Handhabung des Flugzeuges. Da das Radio nur unwesentlich kleiner ist als die Schachtöffnung, muß es zum Einführen ziemlich präzise positioniert werden. Wird das Radio verkantet, kollidiert eine der hinteren Kanten mit der Schachtöffnung, und das Radio vollzieht als Ausweichbewegung eine seitliche Kippbewegung. Da dem Benutzer keine unmittelbare Kraftrückkopplung zur Verfügung steht, ist er bei der Handhabung ausschließlich auf die visuelle Darstellung der Objektbewegung und eventuell eingeblendete Indikatoren (Kontaktkraftpfeile) angewiesen. Dieses Fehlen der aus der Alltagserfahrung vertrauten haptischen Wahrnehmung erschwert die Durchführung der Einbauaufgabe.

10.3 Beispiel 3: Schraube

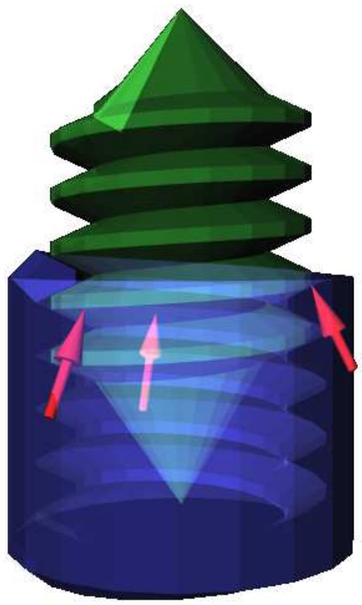


Abbildung 75: Beispiel 3: Schraube und Mutter.

Das in Abb. 75 dargestellte dritte Anwendungsbeispiel besteht aus einer Schraube, welche in eine Gewindebohrung eingeführt wird. Dieses Beispiel demonstriert die Robustheit des Verfahrens zur automatischen Bestimmung einer nichtredundanten Anzahl von Punktkontakten, sowie zur Simulation der verbleibenden Bewegungsfreiheitsgrade auch in komplexen geometrischen Anordnungen. Zwischen den Gewindegängen von Innen- und Außengewinde besteht nur wenig Spielraum, so daß bei der interaktiv gesteuerten Schraubbewegung ständig neue Kontakte entstehen, und andere sich wieder lösen. Durch die Regelmäßigkeit der Gewindegeometrie treten typischerweise an sehr vielen Stellen gleichzeitig Kollisionen auf, unter denen die jeweils früheste Kollision identifiziert werden muß . Die beschriebenen Mechanismen zur Bestimmung und Verwaltung von Kontaktpunkten bewältigen auch solche anspruchsvolle Szenarien stabil. Bei der interaktiven Steuerung der Schraubbewegung verkanten Schraube und Mutter häufig, und es treten sehr viele verschiedene Kontaktkonfigurationen und Kräfteverteilungen auf. Das beschriebene iterative Verfahren zur Lösung des NCP und damit zur Bestimmung der Kontaktkräfte konvergiert bis auf Ausnahmefälle sehr gut. Konvergenzprobleme wurden nur bei extremen Bewegungsvorgaben beobachtet (siehe hierzu Abschnitt 6.4).

Stellt man Schraubenachse und Mutter senkrecht und überläßt die Schraube einer simulierten Schwerkraftwirkung, während die Mutter räumlich fixiert wird, so zentriert sich die Schraube selbständig zur Mutter und gleitet bei einem geeignet gewählten Reibungskoeffizienten in einer Rotationsbewegung langsam nach unten. Dieses Ergebnis unterstreicht das resultierende physikalisch glaubwürdige Verhalten der realisierten Bewegungssimulation mit Hinderniskontakten.

Dieses Schraubenbeispiel stellt ebenfalls einen guten Test für die Effizienz des verwendeten Kollisionserkennungsmoduls dar. Durch die Verzahnung der beteiligten Körper und die geringen Abstände zwischen den Gewindegängen lassen sich mit groben Hüllkörpern praktisch keine Kollisionen ausschließen. Erst in den tieferen (detaillierteren) Hierarchieebenen wirken sich die schnellen Hüllkörpertests positiv aus.

10.4 Beispiel 4: Lampenwechsel

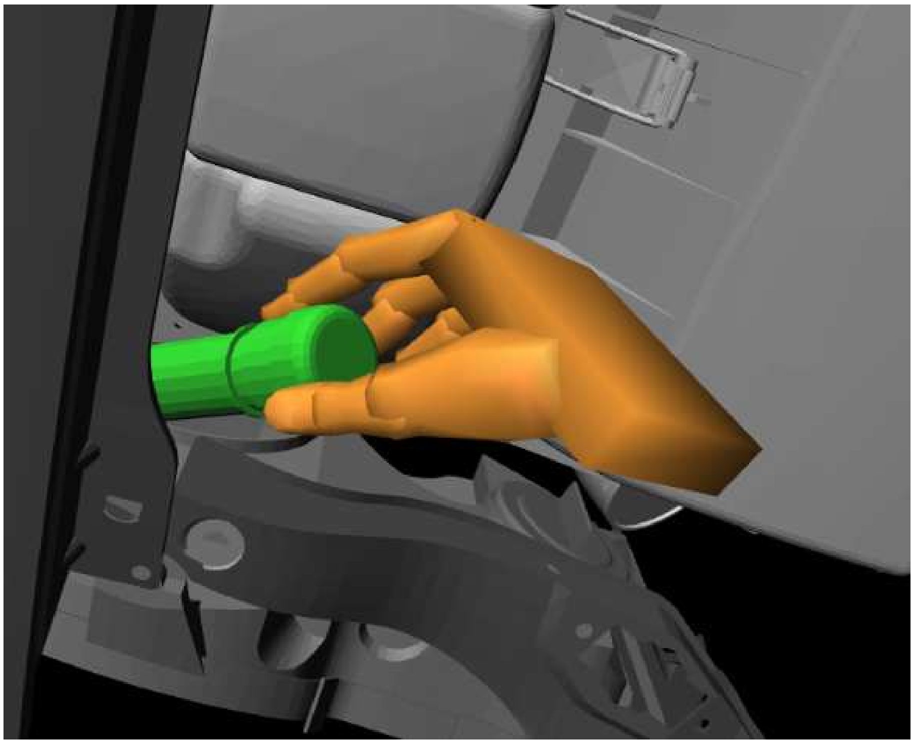


Abbildung 76: Beispiel 4: Simulation des Ausbaus einer Glühbirne aus dem Scheinwerfer eines Pkw mit dem Datenhandschuh.

Dieses letzte Beispiel zeigt einen typischen Anwendungsfall aus der industriellen Praxis. Bei der Konstruktion eines neuen Produktes strebt man an, Arbeitsschritte der Montage und Wartung möglichst frühzeitig durch Montagesimulationen durchzuspielen, um deren Durchführbarkeit und ergonomische Fragestellungen zu evaluieren. Abb. 76 zeigt den vorderen Bereich eines Pkw-Motorraums. Die Handhabungsaufgabe in diesem Beispiel besteht darin, die Glühbirne des Fahrscheinwerfers auszuwechseln. Wie aus der Abbildung ersichtlich ist, besteht zwischen dem Scheinwerfergehäuse und anderen Baugruppen des Motorraumes nur ein geringer Freiraum zur Handhabung und Montage der Glühbirne. Aus diesem Grunde ist es erforderlich, die Durchführbarkeit des Lampenwechsels durch eine Einbausimulation zu verifizieren.

Die Szene zeichnet sich durch eine hohe geometrische Komplexität aus[13], was sich insbesondere auf die Rechenzeiten der Kollisionserkennung auswirkt. Aufgrund der effizienten Techniken des verwendeten Kollisionserkennungsmodules bleiben jedoch die resultierenden Simulationszeiten in einem Bereich, der interaktive Einbauuntersuchungen zuläßt. Die gemessenen Rechenzeiten für ein Simulationsintervall bewegen sich im Bereich von 1ms bis einige 100ms. Die konkreten Werte sind in starkem Maße davon abhängig, welche Sollbewegungen interaktiv vorgegeben werden, und wie nahe sich die Glühbirne an anderen Szenenobjekten entlangbewegt.

Aufgrund der vorliegenden Anordnung der verschiedenen Baugruppen ist die die Zielposition der Glühbirne im Scheinwerfer kaum direkt einzusehen. Um Handhabungsaufgaben in solchen Situationen realistisch simulieren zu können, wären Verfahren zur haptischen Rückkopplung besonders wünschenswert. Solange solche nicht zur Verfügung stehen, bleibt nur der Weg, fehlende Kraftrückkopplung durch zusätzliche visuelle Informationen zu kompensieren, indem beispielsweise verdeckende Objekte halbtransparent dargestellt werden.

[13]Die Szene besteht aus 16 kollisionssensitiven Objekten mit insgesamt etwa 62000 Flächen, 186000 Kanten, und 39000 Eckpunkten.

11 Zusammenfassung und Ausblick

Die beschriebenen Verfahren zur Bewegungssimulation mit Hinderniskontakten und ihre Implementierung im Rahmen der VR-Plattform *DBView* erlauben eine intuitiv richtige und physikalisch glaubwürdige interaktive Handhabung virtueller Objekte in simulierten Umgebungen. Durch die spezielle Formulierung der Kontaktbedingungen werden die Kontaktabstände an den Berührstellen stabil eingehalten. Durch die automatische Bestimmung kanonischer Punktkontaktkonfigurationen zur Nachbildung beliebiger Kontaktsituationen wird eine effiziente Realisierung der Bewegungssimulation errreicht. Der kraftbasierte Ansatz ermöglicht ebenfalls die Nachbildung von Reibungseffekten. Sowohl Gleitreibung als auch Haftreibung wurden im Rahmen der Kontaktsimulation realisiert.

Die Beschreibung aller Kontaktsituationen durch eine minimale Anzahl von Punktkontakten erlaubt die im Rahmen der numerischen Genauigkeit exakte Einhaltung der kinematischen Kontaktbedingungen. Ein Vergleich der Reibungseigenschaften von realen Flächenkontakten mit den Eigenschaften der vorgestellten Substitution durch Punktkontakte wurde sowohl anhand theoretischer Betrachtungen als auch empirisch durchgeführt. Dabei zeigte sich, daß die Reibungssimulation mit wenigen Punktkontakten zwar bei Rotationsbewegungen typischerweise zu große Reibmomente aufweist, und somit der Rotationsanteil von reibungsbehafteten Gleitbewegungen stärker abgebremst wird als diese bei realen Flächenkontakten der Fall wäre. Das resultierende Bewegungsverhalten bleibt jedoch qualitativ stets richtig, und somit auch intuitiv glaubwürdig.

Die Echtzeitfähigkeit des vorgestellten Verfahrens hängt zunächst von der Anzahl der gleichzeitig zu berücksichtigenden Anzahl von Punktkontakten ab. Auf einer kleineren Workstation können Simulationen mit 10 Punktkontakten in unter 10 ms pro Simulationsschritt durchgeführt werden. Somit können sicher mindestens drei sich gegenseitig berührende Objekte in Echtzeit simuliert werden. Eine andere mögliche Einschränkung der Echtzeitfähigkeit ergibt sich aus dem zur Kollisionserkennung erforderlichen numerischen Aufwand, der wesentlich von der Komplexität der beteiligten Objektgeometrien bestimmt wird.

Im Rahmen dieser Arbeit wurden ausschließlich kontaktbedingte Bewegungseinschränkungen in Betracht gezogen. In speziellen Anwendungen kommen auch Situationen vor, welche zwar im engeren Sinne auch kontaktbedingte Bewegungseinschränkungen darstellen, die sich jedoch effizienter durch spezifische Bewegungsbedingungen realisieren lassen. Hierzu gehören beispielsweise der Fall einer Schublade, die sich entlang ihrer Führung bewegt, oder der Fall einer Schraube, welche sich in einer Gewindebohrung bewegt. Ganz allgemein fallen auch alle Arten von Gelenken und Scharnieren in diese Kategorie. Solche spezifischen Bewegungsbeschränkungen stellen eine mögliche Erweiterung der allgemeinen kontaktbedingten Bewegungsbeschränkungen dar. Zu ihrer Realisierung ist die Formulierung entsprechender Randbedingungen für die Lösung der Bewegungsgleichungen erforderlich. Da der bisher realisierte Ansatz bereits ein System individueller Kontaktbedingungen vorsieht, stellt eine solche Erweiterung voraussichtlich kein wesentliches Problem dar.

Eine weitere Art von Bewegungseinschränkungen gehen auf anwendungsspezifische Anforderungen zurück. So könnte beispielsweise die Bewegung eines interaktiv gesteuerten Behälters so eingeschränkt werden, daß er nicht irrtümlich gekippt werden kann. Auf diese Weise können selbst bei interaktiven Handhabungssimulationen anwendungsspezifische Randbedingungen umgesetzt werden. Auch solche Bewegungseinschränkungen lassen sich vermutlich ohne wesentliche Schwierigkeiten integrieren.

Die vorgestellte Bewegungssimulation bewirkt die Einhaltung vorgegebener Kontaktbedingungen durch Einsatz geeigneter Zwangskräfte, nämlich durch die an den Berührstellen wirkenden Kontaktkräfte. Diese Kräfte werden zunächst nur als interne Hilfsgrößen benötigt, um das resultierende Bewegungsverhalten der betrachteten Körper zu bestimmen. Diese Kräfte sind jedoch durchaus auch als wesentliche Voraussetzung zur Ansteuerung von Kraftrückkopplungseinrichtungen zu betrachten. Die derzeit am häufigsten verwendeten Interaktionsgeräte zur Objektmanipulation in virtuellen Umgebungen, Datenhandschuh und SpaceMouse, bieten keine Möglichkeiten, Reaktionskräfte innerhalb der Simulation dem Anwender erfahrbar zu machen. Eine solche Kraftrückkopplung wäre für Handhabungssimulationen jedoch von großer Bedeutung, da sie die intuitiven Kontrollmöglichkeiten wesentlich verbessern würde. Derzeit stehen allerdings noch kaum akzeptable Kraftrückkopplungssysteme zur Verfügung. Die Schwierigkeiten hierbei liegen vorwiegend bei der Realisierung des mechanischen Aufbaus. Es existieren allerdings bereits prototypische Geräte, die alle 6 Kraftdimensionen (3 Kräfte und 3 Drehmomente) bereitstellen. Das in dieser Arbeit dargestellte Verfahren zur Bewegungssimulation mit Kontaktkräften stellt gerade diese Kräfte und Momente zur Verfügung, welche zur Ansteuerung eines Kraftrückkopplungssystems benötigt werden. Die Anbindung eines solchen Interaktionsinterfaces an das bisherige Simulationssystem stellt daher die konsequente Fortführung des bisher beschrittenen Weges dar.

A Beschreibung von Rotationen mit Quaternionen

Zur Beschreibung von Rotationen werden häufig Rotationsmatrizen $\mathbf{R} \in \mathbb{R}^{3\times3}$ verwendet, da sie sehr einfach angewandt werden können. Die Rotation eines Vektors wird dabei durch eine Matrizenmultiplikation realisiert. Für die Bestimmung einer Rotation sind diese Matrizen eher unhandlich, da sie die drei Bewegungsfreiheitsgrade der Rotation auf neun Zahlenwerte verteilen. Eine sehr viel intuitivere und kompaktere Beschreibungsmöglichkeit für Rotationen bieten die sogenannten Quaternionen $\mathfrak{q} = (q_0, \mathbf{q}) \in \mathbb{R}^4$, welche explizit die Rotationsachse $\mathbf{r} \in \mathbb{R}^3$, $|\mathbf{r}| = 1$ und den Rotationswinkel φ beschreiben:

$$\mathfrak{q}_{\mathbf{r},\varphi} = \left(\cos\frac{\varphi}{2}, \sin\frac{\varphi}{2}\mathbf{r}\right).$$

In diesem Abschnitt stellen wir wesentliche Eigenschaften und Rechenvorschriften für Quaternionen dar, insoweit als sie zum Verständnis dieser Arbeit beitragen.

Die Addition, Multiplikation, und Konjugierung von Quaternionen sind wie folgt definiert:

$$\begin{aligned}
(a_0, \mathbf{a}) + (b_0, \mathbf{b}) &= (a_0 + b_0, \mathbf{a} + \mathbf{b}) \\
(a_0, \mathbf{a}) \cdot (b_0, \mathbf{b}) &= \left(a_0 b_0 - \mathbf{a}^T\mathbf{b}, a_0\mathbf{b} + b_0\mathbf{a} + \mathbf{a} \times \mathbf{b}\right) \\
(a_0, \mathbf{a})^* &= (a_0, -\mathbf{a}).
\end{aligned}$$

Man beachte, daß die Multiplikation nicht kommutativ ist.

Die Norm eines Quaternions ist gegeben durch $|\mathfrak{q}| = \sqrt{\mathfrak{q}\mathfrak{q}^*}$, und die Inverse erhält man durch $\mathfrak{q}^{-1} = \mathfrak{q}^*/|\mathfrak{q}|^2$. Bei Verwendung von Einheitsquaternionen $|\mathfrak{q}| = 1$ vereinfacht sich dies zu der Beziehung $\mathfrak{q}^{-1} = \mathfrak{q}^*$.

Die folgenden Darstellungsformen der Rotation eines Vektors $\hat{\mathbf{x}} \to \mathbf{x}$ sind äquivalent:

$$\begin{aligned}
\mathbf{x} &= \mathbf{R}(\mathfrak{q})\hat{\mathbf{x}} \\
\mathbf{x} &= \mathfrak{q} \cdot (0, \hat{\mathbf{x}}) \cdot \mathfrak{q}^* \\
\mathbf{x} &= \frac{2\mathbf{q}^T\hat{\mathbf{x}}\mathbf{q} + (q_0^2 - \mathbf{q}^2)\hat{\mathbf{x}} + 2q_0(\mathbf{q} \times \hat{\mathbf{x}})}{q_0^2 + \mathbf{q}^2}.
\end{aligned}$$

Die Beziehungen zwischen $\mathbf{R}(\mathfrak{q})$ und $(q_0, \mathbf{q})$ lauten

$$\begin{aligned}
\mathbf{R}(\mathfrak{q}) &= \frac{(q_0^2 - \mathbf{q}^2)\mathbf{E} + 2\mathbf{q}\mathbf{q}^T + 2q_0\mathbf{q}^\times}{q_0^2 + \mathbf{q}^2} && (80) \\
q_0 &= \arccos\left(\frac{1}{2}(\mathbf{R}_{1,1} + \mathbf{R}_{2,2} + \mathbf{R}_{3,3} - 1)\right) \\
\mathbf{q} &= (2\sin\varphi)^{-1}(\mathbf{R}_{3,2} - \mathbf{R}_{2,3}, \mathbf{R}_{1,3} - \mathbf{R}_{3,1}, \mathbf{R}_{2,1} - \mathbf{R}_{1,2})^T.
\end{aligned}$$

Desweiteren gelten die folgenden arithmetischen Regeln:

$$\begin{aligned} (\mathbf{a}\cdot\mathbf{b})^* &= \mathbf{b}^*\cdot\mathbf{a}^* \\ (\mathbf{a}\cdot\mathbf{b})\cdot\mathbf{c} &= \mathbf{a}\cdot(\mathbf{b}\cdot\mathbf{c})\,. \end{aligned}$$

Zur Herleitung der zeitlichen Ableitung einer Rotation in Quaternionenschreibweise, $\dot{\mathbf{q}}$, definieren wir die Rotation als Zeitfunktion über das Quaternionenprodukt $\mathbf{q}(t) = \mathbf{q}_\omega(t)\cdot\mathbf{q}(0)$, wobei $\frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|}$ die Rotationsachse bezeichne, sowie $|\boldsymbol{\omega}|$ die Winkelgeschwindigkeit:

$$\begin{aligned} \mathbf{q}(t) &= \left(\cos\frac{|\boldsymbol{\omega}|}{2}t, \frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|}\sin\frac{|\boldsymbol{\omega}|}{2}t\right)\cdot\mathbf{q}(0), \\ \dot{\mathbf{q}}(t) &= \left(-\frac{|\boldsymbol{\omega}|}{2}\sin\frac{|\boldsymbol{\omega}|}{2}t, \frac{\boldsymbol{\omega}}{2}\cos\frac{|\boldsymbol{\omega}|}{2}t\right)\cdot\mathbf{q}(0). \end{aligned}$$

Daraus ergibt sich die Zeitableitung bei $t = 0$ zu $\dot{\mathbf{q}}(0) = \left(0, \frac{\boldsymbol{\omega}}{2}\right)\cdot\mathbf{q}(0)$, oder kurz

$$\dot{\mathbf{q}} = \frac{1}{2}\boldsymbol{\omega}\cdot\mathbf{q}.$$

Gegeben seien die Orientierung $\mathbf{q}^t$ eines Körpers zu Beginn eines Zeitintervalls, und die Orientierung $\mathbf{q}^{t+\Delta t}$ an dessen Ende. Die entsprechende mittlere Winkelgeschwindigkeit $\boldsymbol{\omega}$ erhalten wir gemäß (24) aus der Beziehung

$$\mathbf{q}^{t+\Delta t} = \mathbf{q}^t + \frac{1}{2}\Delta t\ (0,\boldsymbol{\omega})\cdot\mathbf{q}^{t+\Delta t}.$$

Nach einer Multiplikation von rechts mit $(\mathbf{q}^{t+\Delta t})^*$ erhalten wir daraus

$$(0,\boldsymbol{\omega}) = \frac{1-\mathbf{q}^t\cdot(\mathbf{q}^{t+\Delta t})^*}{\frac{1}{2}\Delta t}.$$

Weitere Details zur Herleitung und Arithmetik von Quaternionen finden sich in [20].

Literatur

[1] D. Baraff: *Dynamic simulation of non-penetrating rigid bodies*, Ph.D. Thesis 92-1275, Cornell University, March 1992

[2] D. Baraff, R. Mattikalli: *Impending Motion Directions of Contacting Rigid Bodies*, Technical Report CMU-RI-TR-93-15, Carnegie Mellon University, June 1993

[3] D. Baraff: *Fast contact force computation for nonpenetrating rigid bodies*, SIGGRAPH 94, Orlando, July 1994

[4] D. Baraff: *Linear-time dynamics using Lagrange multipliers*, Technical Report CMU-RI-TR-95-44, Carnegie Mellon University, January 1996

[5] D. Baraff: *An Introduction to Physically Based Modeling: Rigid Body Simulation II - Nonpenetration Constraints*, SIGGRAPH Lecture Notes, SIGGRAPH 1997.

[6] R. Barzel, A.H. Barr: *A modeling system based on dynamic constraints*, Computer Graphics (Proc. SIGGRAPH), vol.22, pp.179-188, 1988

[7] W. Bouma, G. Vaněček: *Collision Detection and Analysis in a Physically Based Simulation*, Proc. Eurographics Workshop on Animation and Simulation Workshop, Wien, pp. 191-203, Sept. 1991

[8] W. Bouma, G. Vaněček: *Contact Analysis in a Physically Based Simulation*, ACM Symposium on solid Modeling and Applications, Montreal, Canada, May 1993

[9] M. Buck, E. Schömer: *Interactive Rigid Body Manipulation with Obstacle Contacts*, Journal of Visualization and Computer Animation 9, pp.243-257, 1998

[10] C. Cai, B. Roth: *On the spatial motion of a rigid body with point contact*, IEEE International Conference of Robotics and Automation, pp.686-695, 1987

[11] K. Carr, R. England (edts): *Simulated and Virtual Realities - Elements of Perception*, Taylor & Francis, 1995

[12] Th. Chadzelek: *Heuristische Bewegungsplanung mit vielen Freiheitsgraden*, Diplomarbeit an der Universität des Saarlandes, Saarbrücken, März 1995

[13] T.H. Cormen, C.E. Leiserson, and R.L. Rivest:: *Introduction to Algorithms*. MIT Press / McGraw-Hill, 1990

[14] R.W. Cottle, G.B. Danzig: *Complementarity pivot theory of mathematical programming*. Linear Allgebra and its Applications, **1**:103-125, 1968

[15] R.W. Cottle, J.S. Pang, and R.E. Stone:: *The Linear Complementarity Problem*. Academic Press, 1992

[16] J.F. Cremer, A.J. Stewart: *The Architecture of Newton, a general-purpose dynamic simulator*, IEEE International Conference of Robotics and Automation, pp.1806-1811, 1989

[17] J. Eckstein: *Echtzeitfähige Kollisionserkennung für VR-Anwendungen*, Dissertation, Universität des Saarlandes, 1998

[18] M.C. Ferris, J.S. Pang: *Engineering and Economic Applications of Complementarity Problems*, Technical Report 95-07, Computer Sciences Department, University of Wisconsin, Madison, May 1995

[19] A. Fischer: *A special Newton-type optimization method*, Optimization 24, pp. 269-284, 1992

[20] P.-G. Maillot: *Using quaternions for coding 3d transformations.* Graphic Gems (A.S. Glassner, edt.) pp. 498-515, Academic Press, Boston, 1990

[21] S. Gottschalk, M.C. Lin, D. Manocha: *OBBTree: A Hierarchical Structure for Rapid Interference Detection*, SIGGRAPH 96, Computer Graphics Proceedings, pp. 171-179, August 1996

[22] J.K. Hahn: *Realistic animation of rigid bodies*, Computer Graphics 22(4):299-308, August 1988

[23] Chr. Kanzow, H. Kleinmichel: *A Class of Newton-Type Methods for Equality and Inequality Constrained Optimization*, Preprint 61, Institut für Angewandte Mathematik, Universität Hamburg, Hamburg, 1992

[24] Chr. Kanzow: *Some Equation-Based Methods for the Nonlinear Complementarity Problem*, Optimization Methods and Software, Vol. 3, pp.327-340, 1994

[25] Chr. Kanzow: *Global Convergence Properties of some Iterative Methods for Linear Complementarity Problems*, SIAM Journal of Optimization, Vol.6, No.2, pp.326-341, May 1996

[26] Chr. Kanzow, H. Kleinmichel: *A New Class of Semismooth Newton-Type Methods for Nonlinear Complementarity Problems*, Manuskript, Institut für angewandte Mathematik, Universität Hamburg, Januar 1997

[27] C.E. Lemke: *Bimatrix equilibrium points and mathematical programming*, Management Science **11**:681-689, 1965

[28] K.E. Lynch: *The Mechanics of Fine Manipulation by Pushing*, Proc. IEEE International Conference on Robotics and Automation, pp.2269-2276, 1992.

[29] K.E. Lynch: *Estimating the Friction Parameters of Pushed Objects*, Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.186-193, 1993.

[30] P. Lötstedt: *Mechanical systems of rigid bodies subject to unilateral constraints.*, SIAM Journal of Applied Mathematics, 42(2):281-296, 1982

[31] O.L. Mangasarian: *Equivalence of the Complementarity Problem to a System of Nonlinear Equations*, SIAM Journal of Applied Mathematics, 11, pp.89-92, 1976 of Robotics Research, Vol.7, No.3, June 1988

[32] M.T. Mason, J.K. Salisbury: *Robot Hands and the Mechanics of Manipulation*, The MIT Press, 1985

[33] D.J. Montana: *The kinematics of contact and grasp*, The International Journal of Robotics Research, Vol.7, No.3, June 1988

[34] B. Mirtich, J. Canny: *Impulse-based dynamic simulation* In K. Goldberg, D. Halperin, J.C. Latombe, and R. Wilson (edts.) The algorithmic foundations of robotics. A.K. Peters, Boston, MA, 1995

[35] B. Mirtich, J. Canny: *Impulse-based simulation of rigid bodies*, Symposium on Interactive 3D Graphics, New York, ACM Press, 1995

[36] B. Mirtich: *Hybrid Simulation: Combining Constaints and Impulses* Proceedings of The First Workshop an Simulation and Interaction in Virtual Environmnts SIVE95, Iowa City, pp.153-158, July 1995

[37] Silicon Graphics Inc.: *The OpenGL Programming Guide - The Official Guide to Learning OpenGL*, Addison-Wesley, Reading, Mass., 1993

[38] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling: *Numerical Recipes*, Cambridge University Press, 1986

[39] J. Sauer, E. Schömer: *Dynamiksimulation starrer Körper für Virtual Reality – Anwendungen*, ASIM98, Zürich 1998

[40] J. Sauer, E. Schömer: *A Hybrid Rigid Body Simulation System for Virtual Reality Applications*, Technischer Bericht, Universität des Saarlandes, 1998

[41] E. Schömer: *Interaktive Montageplanung mit Kollisionserkennung*, Dissertation, Universität des Saarlandes, 1994

[42] Y. Shan, Y. Koren: *Obstacle accomodation motion planning*, IEEE Transactions on Robotics and Automation, Vol.11, No.1, Febr. 1995

[43] J. Krim: *Reibung auf atomarer Ebene*, Spektrum der Wissenschaft, 12/1996, pp.80-85, 1996

[44] D.E. Stewart, J.C. Trinkle: *An Implicit Time-stepping Scheme for Rigid Body Dynamics with Coulomb Friction*, Int. Jounal of Numerical Methods in Engineering, 1996

[45] J. Trinkle, J.-S. Pang, S. Sudarsky, G. Lo: *On Dynamic Multi-Rigid-Body Contact Problems with Coulomb Friction*, Zeitschrift für angewandte Mathematik und Mechanik, 1996

[46] G. Vaněček, Jr., J.F. Cremer: *Project Isaac: Building Simulations for Virtual Environments*, Technical Report, Purdue University, June 1994

[47] J. Wernecke: *The Inventor Mentor - Programming Object-Oriented 3D Graphics with OpenInventor*, Addison-Wesley, Reading, Mass., 1994