



Forschungszentrum Karlsruhe
Technik und Umwelt

Wissenschaftliche Berichte
FZKA 6040

Hochleistungsrechnen im Virtuellen Rechenzentrum Karlsruhe

**Beiträge zu einem Kolloquium
aus Anlaß der Einweihung des
Vektor-Parallel-Computers
SNI VPP300-16
im Forschungszentrum Karlsruhe
am 4. Juni 1997**

K.-P. Mickel

**Hauptabteilung
Informations- und Kommunikationstechnik**

Dezember 1997

Forschungszentrum Karlsruhe

Technik und Umwelt

Wissenschaftliche Berichte

FZKA 6040

**Hochleistungsrechnen im
Virtuellen Rechenzentrum Karlsruhe**

Beiträge zu einem Kolloquium aus Anlaß der Einweihung des
Vektor-Parallel-Computers SNI VPP300-16
im Forschungszentrum Karlsruhe
am 4. Juni 1997

Herausgeber:

Klaus-Peter Mickel

Hauptabteilung Informations-
und Kommunikationstechnik

Forschungszentrum Karlsruhe GmbH, Karlsruhe

1997

**Als Manuskript gedruckt
Für diesen Bericht behalten wir uns alle Rechte vor**

**Forschungszentrum Karlsruhe GmbH
Postfach 3640, 76021 Karlsruhe**

**Mitglied der Hermann von Helmholtz-Gemeinschaft
Deutscher Forschungszentren (HGF)**

ISSN 0947-8620

Inhaltsverzeichnis

K.-P. Mickel (Forschungszentrum Karlsruhe) Begrüßung und Eröffnung des Kolloquiums	1
H. Borgwaldt (Forschungszentrum Karlsruhe) Wie das Hochleistungsrechnen im Forschungszentrum Karlsruhe begann	3
A. Schreiner, W. Schönauer (Universität Karlsruhe) Warum ein gemeinsames Virtuelles Rechenzentrum von Universität und Forschungszentrum ?	11
K.-P. Bohnen (Forschungszentrum Karlsruhe) Simulation komplexer Strukturen auf Hochleistungsrechnern: Molekülkristalle aus Fullerenen	17
L. Eigenmann et al. (Universität Karlsruhe) Strömungs- und Wärmetechnische Optimierung einer KFZ-Kupplung mittels numerischer Simulation	23
M. Krauß (Forschungszentrum Karlsruhe) Einsatz von Hochleistungsrechnern zur Simulation physikalischer Phänomene	34
W. Rodi et al. (Universität Karlsruhe) Simulation turbulenter Strömungen auf Hochleistungsrechnern	49

Begrüßung und Eröffnung des Kolloquiums

Klaus-Peter Mickel

Hauptabteilung Informations- und Kommunikationstechnik
Forschungszentrum Karlsruhe

Sehr verehrte Gäste,

im Namen des Vorstandes des Forschungszentrums Karlsruhe begrüße ich Sie sehr herzlich zum heutigen Kolloquium.

Eine ganze Reihe von Gästen hat heute den Weg zu uns gefunden; stellvertretend begrüße ich Herrn Prof. Klose, der 20 Jahre Mitglied des Vorstandes des FZK gewesen ist, und der während seiner Tätigkeit in unserem Vorstand zeitweise auch für unsere Hauptabteilung zuständig gewesen ist. Außerdem begrüße ich die Leiter der Rechenzentren der Universitäten Darmstadt und Würzburg, Herrn Dr. Ohrnberger und Herrn Schliffer, den Leiter der Zentralen Datenverarbeitung des Deutschen Krebsforschungszentrums in Heidelberg, Herrn Dr. Böhm, sowie die Repräsentanten des Unternehmens Siemens-Nixdorf AG, denen ich an dieser Stelle für die Unterstützung dieser Veranstaltung danke.

Ich freue mich sehr, daß dieses Kolloquium auch in der Universität Karlsruhe und in unserem Forschungszentrum selbst eine so große Resonanz gefunden hat. Und so kann ich zahlreiche Angehörige der Universität Karlsruhe begrüßen, und hier ganz besonders die Mitarbeiterinnen und Mitarbeiter des Rechenzentrums der Universität. Wir betreiben die VPP300 ja in einem gemeinsamen Virtuellen Rechenzentrum. - Wir werden später in einem Referat hören, was sich hinter diesem geheimnisvollen Begriff verbirgt.

Weiterhin begrüße ich aus den Instituten des Forschungszentrums zahlreiche Kolleginnen und Kollegen, sowie besonders herzlich die Mitarbeiterinnen und Mitarbeiter unserer Hauptabteilung HIK.

Last but not least heiße ich die Referenten des heutigen Tages willkommen, die sehr gerne bereit waren, heute über ihre vielfältigen Erfahrungen mit Hochleistungsrechnern zu berichten.

Wir haben das Datum der heutigen Veranstaltung mit Bedacht gewählt: Gestern vor genau 10 Jahren nämlich haben wir bei uns den ersten von Siemens gelieferten Hochleistungsrechner installiert. Es war eine VP50 mit der - für damalige Verhältnisse -

gewaltigen Leistung von 150 Mega Flops und einem Hauptspeicher von - damals - fast unglaublichen 64 Mega Byte.

In einer Veröffentlichung über diese Installation stand damals wörtlich zu lesen: „Mit der Installation des Vektorrechners VP50 wurde eine neue bedeutende Stufe erklommen.“

Heute, nach nur 10 Jahren, ist von dieser bedeutenden Stufe nicht mehr viel übrig geblieben: Die VP50 hatte nicht einmal ein halbes Prozent der Rechenleistung unserer heutigen VPP300 und ihr Hauptspeicher entspricht nur zwei Promille des Speichers, den wir heute mit der VPP300 in Betrieb nehmen.

Ich erlaube mir nun ein Gedankenexperiment: Wenn wir die Entwicklung der vergangenen 10 Jahre in die Zukunft extrapolieren, so werden wir am 5. Juni 2007 einen Hochleistungsrechner einweihen können, der 8,6 Tera Flops leisten wird, und der 16 Tera Byte Hauptspeicher haben wird - aus heutiger Sicht unglaubliche Werte.

Und so lade ich Sie schon mal prophylaktisch zum 5. Juni 2007 hierher ein.

Doch zurück in die Gegenwart:

Es ist uns gelungen, für das heutige Kolloquium einige Referenten zu gewinnen, die das Hochleistungsrechnen aus ganz unterschiedlichen Blickwinkeln beschreiben werden. Wir werden die Vergangenheit des Hochleistungsrechnens im Forschungszentrum Karlsruhe betrachten, wir werden einiges über aktuelle Forschungsergebnisse hören, die mit Hilfe unserer Hochleistungsrechner erarbeitet werden konnten, und wir werden an manchen Stellen Ausblicke auf die Zukunft des Hochleistungsrechnens bekommen. So wünsche ich Ihnen heute ein fruchtbares Kolloquium mit interessanten Vorträgen und anregenden Diskussionen.

Ich eröffne hiermit das Kolloquium aus Anlaß der Einweihung des Vektor-Parallel-Computers VPP300.

Wie das Hochleistungsrechnen im Forschungszentrum Karlsruhe begann

Horst Borgwaldt

Bis 1992: Institut für Neutronenphysik und Reaktortechnik
Forschungszentrum Karlsruhe

Heute, im Jahr 1997, umfaßt der Begriff Hochleistungsrechnen sowohl den Einsatz von Vektorrechnern als auch von Parallelrechnern und sogar von Vektor-Parallelrechnern wie der VPP300/16 von Siemens Nixdorf (SNI). Dieser Beitrag beschränkt sich dagegen auf Vektorrechner, denn mit ihnen hat es im Forschungszentrum Karlsruhe (FZK) begonnen. Die Ansätze zum Parallelrechnen lagen deutlich später.

Die Einführung des Vektorrechnens in unserem Forschungszentrum verlief in zwei Phasen. In der ersten Phase, von Mitte 1983 bis Mitte 1987, hatten wir fast ausschließlich den Zugang zum Vektorrechner Cyber-205 an der Universität Karlsruhe. Ausnahmen waren wenige Mitarbeiter, die im Rahmen internationaler Zusammenarbeit schon Erfahrungen mit Vektorrechnern in Forschungszentren der USA sammeln konnten. Ferner konnten, über unser Rechenzentrum (früher HDI, jetzt HIK), in gewissem Umfang Tests auf Anlagen in Jülich (KFA), Stuttgart, Kaiserslautern u.a. durchgeführt werden. Die zweite Phase begann Mitte 1987 mit der Beschaffung eines eigenen Vektorrechners Siemens VP50 durch unser Forschungszentrum.

Die 'Initiative Vektorrechner'

In den Jahren 1980-1982 etwa verspürten einige Mitarbeiter (darunter der Verfasser dieses Beitrags) aus 3-4 Instituten und dem Rechenzentrum sehr deutlich, daß auch im Forschungszentrum Karlsruhe die Zeit für den Einsatz von Vektorrechnern gekommen war.

Als erster brauchbarer Vektorrechner wurde seit 1976 die CRAY-1 im Forschungszentrum Los Alamos betrieben. Wir hatten schon eine Anlage beim MPI in Garching besichtigen können. Über Architektur, Technologie, Funktionen und Leistung von Vektorrechnern lag damals hinreichende Information vor.

Es gelang, Institutsleiter und Vorstand unseres Zentrums wenigstens zu einem Teil positiv zu beeinflussen. Die darauf anlaufende Diskussion drehte sich im wesentlichen um zwei mögliche Wege:

Der erste Weg sah vor, zunächst etwa drei Jahre nur extern, auf Anlagen verschiedener Hersteller Erfahrungen zu gewinnen. Weitgehend offen blieb dabei die Frage der Kosten.

Der zweite Weg, der dann, u.a. aus finanzpolitischen Gründen, gewählt wurde, war der Eintritt mit einem festen Kapitalbeitrag in eine Kooperation. Der Partner einer solchen Kooperation stand in unmittelbarer Nähe bereit: Die Universität Karlsruhe. So kam es zu einem Vertrag zwischen der Universität und unserem Forschungszentrum.

Die Universität beschaffte einen Vektorrechner Cyber-205 der Firma Control Data (CDC) sowie einen IBM-kompatiblen Frontrechner Siemens 7865, der auch als Verbindungsrechner zu unserem Zentrum diente. Die Cyber-205 hatte anfangs nur eine Vector-Pipe und wurde später auf zwei Pipes und damit auf doppelte Leistung hochgerüstet.

Unser Forschungszentrum leistete, neben anderen Partnern der Universität, einen Beitrag zur Finanzierung. Wir bekamen ab Mitte 1983 Zugang zur Cyber-205, in bestimmtem Umfang und einem gewissen zeitlichen Rahmen. Wir erhielten vom Rechenzentrum der Universität auch die sowohl notwendige als auch sehr kompetente Unterstützung und Beratung.

Die Ausgangssituation

Die Cyber-205 hatte, für einfache Vektor-Operationen (Addition, Multiplikation), nach heutigem Maßstab eine eher bescheidene Nennleistung von 50 Megaflops (= Millionen Gleitkomma-Operationen pro Sekunde). Jedoch brachte es unser damaliger Universalrechner, eine IBM-3033, auch nur auf gemessene 2,2 Megaflops.

Zudem zeigte die Cyber-205 in ihrem Skalar-Mode bereits die 1,8-fache Leistung unseres damaligen Universalrechners [1]. Somit lag für den Übergang zum Vektorrechnen eine günstige Ausgangslage vor, welche jedoch keinen Bestand hatte.

Zur Architektur der Cyber-205 [2, Kap. 2]

Im Gegensatz zu allen sonst bekannten Vektorrechnern hatte die Cyber-205 keine Vektorregister, welche normalerweise den relativ schnellen Start von Vektor-Operationen erlauben. Bei der Cyber-205 lagen Operanden und Resultat-Vektoren im Hauptspeicher. Entsprechend war die Aufsetzzeit für Vektor-Operationen mit etwa 1 μ sec recht erheblich. Sie bedeutete 50 verlorene Maschinen-Zyklen. Bei einer Vektorlänge von 50 erhielt man so gerade 50% der Nennleistung.

Gravierender war eine weitere Einschränkung. Für Vektor-Operationen mußten die Vektoren dicht (lückenlos, engl. contiguous) im Speicher liegen. Diese Bedingung ist beispielsweise für Spalten-Vektoren einer Matrix erfüllt, aber nicht für Zeilen-Vektoren. In diesem allgemeineren Fall mußte der Compiler Umspeicher-Befehle einfügen, um so mit geeigneten (dichten) Hilfs-Vektoren operieren zu können. Dadurch mußte sich natürlich die ausgewiesene Rechenleistung in vielen Anwendungen deutlich verschlechtern.

Es gab somit nur zwei Möglichkeiten, die Cyber-205 mit Gewinn zu nutzen:

(1) Der zu vektorisierende Code hatte schon in der Original-Version 1-dimensionale Datenstrukturen mit langen, dichten Vektoren. Dies erlebte der Verfasser mit zwei aus den USA übernommenen Codes.

(2) Man entschloß sich, Codes vollständig neu zu schreiben, unter Beachtung dieser Bedingungen.

Die dritte, theoretische Lösung, die in übernommenen Codes vorliegenden 2- oder 3-dimensionalen Datenstrukturen (mit meist kurzen Vektoren) aufzubrechen, hätte meist einen unangemessenen Programmieraufwand erfordert.

Es gab zu dieser Zeit im üblichen Austausch von Codes kaum schon für Vektorrechner angepaßte Versionen. Das normale Procedere bestand darin, ein geeignetes, ursprünglich für einen Universalrechner geschriebenes Programm schrittweise umzuformen und so für den Vektorrechner zu optimieren. Bei dieser schon für sich aufwendigen 'Vektorisierung' aber blieben die Datenstrukturen im wesentlichen immer erhalten.

Erfahrungen mit der Cyber-205

Im ersten Vertragsjahr war die Nutzung der Cyber-205 durch unser Zentrum ausgesprochen gering und somit schlecht. In den folgenden zwei Jahren war das Bild eher befriedigend, jedoch nur bei pauschaler Betrachtung. Die Anzahl der Benutzer blieb gering (unter 20) und nur drei (später vier) Institute nahmen die Rechenkapazität der Cyber-205 wirklich in Anspruch [3].

Daß die von uns angestrebte Breitenwirkung ausblieb, kann man auf eine Reihe von Gründen zurückführen:

(1) Zeitgleich mit dem Zugang zur Cyber-205 beschaffte das Forschungszentrum Karlsruhe einen neuen Universalrechner, eine Siemens 7890 mit einer mehr als dreifachen Leistung des Vorgängers. Zudem war diese Leistung auch bis zu 80% höher als die der Cyber-205 im Skalar-Mode [1]. Damit entfiel für die Benutzer das 'natürliche Gefälle' hin zum Vektorrechner.

(2) Die Entfernung zwischen beiden Rechenzentren war mit den damaligen Kommunikationsmitteln zu vertretbaren Kosten nicht voll zu überbrücken.

Anfangs konnte man nur im Rechenzentrum der Universität arbeiten. Nach kurzer Zeit gab es in unserem Rechenzentrum einige TSS-Terminals mit Zugriff auf den Frontrechner in der Universität sowie einen dedizierten CDC-Drucker für Cyber-Ausgabe. Schließlich gab es auch TSS-Terminals in den Instituten sowie Druckausgabe über die normalen Drucker und weitere Verbesserungen. Die Bemühungen der Elektroniker und Systemprogrammierer auf beiden Seiten sorgten für immer besseren Service. Wir konnten Dateien von 5-10 Tausend Zeilen problemlos übertragen und bekamen Druckausgaben in normalem Umfang. Nicht möglich war es jedoch, Rohdaten im Umfang einiger Megabits zur Nachbearbeitung für Graphik zu übertragen.

- (3) Die Cyber-205 hatte ein eigenes Betriebssystem und so auch eine eigene Steuersprache. Dies war besonders für reine Code-Anwender, im Gegensatz zu den Entwicklern, eine kleine zusätzliche Hemmschwelle.
- (4) Die interne Zahlen-Darstellung der Cyber-205 unterschied sich von dem gewohnten IBM-Format. Für die Übertragung von unformatierten Daten in beiden Richtungen wurden im Cyber-Code Konversions-Routinen benötigt.
- (5) Auf die architektur-bedingte Bevorzugung langer, dichter Vektoren im Gegensatz zu den oft kurzen Vektoren aus mehrdimensionalen Datenstrukturen wurde oben hingewiesen.
- (6) Die Cyber-205 hatte für die Vektorisierung von Schleifen eine sehr leistungsfähige Hardware. Es gab alles, um die Befehlsausführung unter Bedingungen zu steuern (IF- und BLOCK IF-Konstrukte). Es gab ferner Vektor-Operationen zum Umspeichern von Daten nach Adreßlisten (GATHER und SCATTER) oder unter Kontrolle von Masken-Bitvektoren (COMPRESS und EXPAND), für die Maximum-Suche, Berechnung von Summen, Skalarprodukten etc.

Demgegenüber war die Software eher schwach. Statt den Compiler so zu ertüchtigen, daß er die vorhandene Hardware voll einsetzte, gab man dem Programmierer zahlreiche FORTRAN-Erweiterungen an die Hand, darunter die 'Q8-Calls'. Sie erlaubten den direkten Einsatz beliebiger Vektorbefehle der Hardware.

Nach Einsatz solcher FORTRAN-Erweiterungen war der Code nicht mehr portabel. Darauf aber legten wir aus mehreren Gründen großen Wert (Code-Entwicklung und Einsatz auf verschiedenen Rechnern, Austausch mit anderen Zentren). Gruppenarbeit wurde durch die schlechte Lesbarkeit des modifizierten Codes erschwert (Beispiele finden sich in [4]).

Notwendig war oft die parallele Führung einer portablen und einer Cyber-optimierten Code-Version, mit Hilfe von Quellcode-Verwaltungsprogrammen wie CDC-UPDATE oder HISTORIAN.

Der als Ausweg angebotene Preprocessor VAST konnte ebenfalls nicht voll befriedigen.

Für den Code-Entwickler gab es auch zwei sehr nützliche Aspekte, in denen sich das einfache Betriebssystem der Cyber-205 von Vorteil zeigte:

- (1) Man konnte die Code-Optimierung durch reproduzierbare, präzise Zeitmessungen, auf μsec -Ebene, verfolgen.
- (2) Die SPY-Option lieferte ohne erheblichen Aufwand die Verteilung der Rechenzeiten über den Code.

Es gab auch durchaus ermutigende Ergebnisse von Code-Entwicklungen auf der Cyber-205. Einige Produktions-Codes, aus den Bereichen Thermofluidodynamik (CFD), rechnergestützter Entwurf (CAD), Festkörper-Physik u.a., ergaben als Leistungsgewinne Faktoren 4, 6 und sogar 10 im Vergleich zum Universalrechner Siemens 7890(1, 3). Dabei zeigten sich in der Regel zusätzliche Gewinne, weil die Codes durch die

Vektorisierung und die somit verbesserte Struktur auch auf dem Universalrechner etwa 50% schneller wurden.

Es wurden eine Reihe von Spezial- und Test-Codes entwickelt, um Algorithmen, Iterationsverfahren sowie allgemein Vektorisierungs-Techniken zu erproben. Später dienten sie oft zum Leistungsvergleich verschiedener Vektorrechner-Modelle.

Insgesamt konnten die beteiligten Benutzer in dieser Zeit auf der Cyber-205 wertvolle Erfahrungen im Einsatz von Vektorrechen-Techniken erarbeiten.

Der Arbeitskreis Hochleistungsrechner [3]

Im Januar 1985 setzte der Vorstand des Forschungszentrums einen Arbeitskreis Hochleistungsrechner ein, unter der Leitung von Herrn Höbel (INR). Bei etwa 20 Mitgliedern waren in ihm vertreten alle Institute und Abteilungen mit wesentlichem Bedarf an Rechenkapazität, eine Projektleitung (PSB-PL), unser Rechenzentrum sowie aus dem Vorstandsbereich die Ressorts Planung und Erfolgskontrolle (PEK), Technischer Einkauf (EKM-TAB) und Finanzen (HF).

Eine der Aufgaben des Arbeitskreises war die Analyse der in zwei Jahren gewonnenen Erfahrungen, eine zweite die Ausarbeitung einer Empfehlung für das weitere Vorgehen, etwa ab 1987.

Wir verschafften uns umfassende Information aus der Fachliteratur, durch den Besuch von Fachtagungen und Seminaren (speziell auch an der Universität Karlsruhe) und durch Firmenkontakte.

Bereits 1985 stellten 4 Hersteller uns ihre Produkte vor und gaben zusätzlich einen Ausblick auf die abzusehende Entwicklung der folgenden 1-2 Jahre. Es waren dies die Firmen CDC, Cray, Siemens und IBM (mit der Vektorfunktion VF).

Der japanische Hersteller Fujitsu, in Deutschland vertreten durch die Firma Siemens, war damals ein Newcomer. Die VP-Serie [2, Kap. 13] hatte eine leistungsfähige Hardware. Neuartig und auffallend war das Konzept eines großen, variabel konfigurierbaren Vektorregister-Felds, geeignet zum Operieren mit langen und kurzen Vektoren.

Etwa 8-10 Codes wurden gebündelt für Leistungsvergleiche. Dieses Benchmarking wurde im Herbst 1985 von den Firmen Cray und IBM in den USA durchgeführt. An den Tests bei Siemens in München konnten einige Mitarbeiter unseres Zentrums mitwirken.

Aufgrund der bisherigen Erfahrungen interessierte uns neben der reinen Rechenleistung insbesondere die Qualität der Compiler und anderer Software. Es ist bekannt, daß der Fujitsu-Compiler damals einen neuen Standard setzte:

- (1) Er akzeptierte portables FORTRAN 77 ohne irgendwelche Erweiterungen. Wie schon bei anderen Herstellern wurden Compiler-Direktiven eingesetzt, wo der Compiler zusätzliche Information zum Optimieren brauchte.
- (2) Die Vektorisierung von Code-Schleifen erfolgte sehr weitgehend, fast wie 'von Hand' optimiert.

- (3) Er lieferte eine sehr gute Diagnostik. Sie umfaßte die Auflistung der Vektorisierungs-Hindernisse mit Vorschlägen zu Verbesserungen im Quellcode. Die Optimierungs-Strategie wurde aufgezeigt und eventuell geeignete Compiler-Direktiven wurden vorgeschlagen.

Hier ist zu bemerken, daß andere Hersteller, wie Cray und IBM, sehr schnell nachzogen. Es ergab sich so ein Qualitäts-Standard für fortschrittliche, moderne Compiler. Wir beobachteten damals eine sehr schnelle Entwicklung in diesem Bereich. Sowohl bei Cray als auch bei Siemens (Fujitsu) erfolgten einige Tests mit einem Standard-Compiler und mit einer neuen, noch nicht allgemein verfügbaren Compiler-Version. Dabei zeigten sich teilweise Leistungs-Gewinne von 50-80%.

Die vom Arbeitskreis erarbeiteten Kriterien für einen zu beschaffenden Vektorrechner waren im wesentlichen folgende:

- (1) Der Vektorrechner sollte im Skalar-Mode nicht langsamer sein als ein verfügbarer Universalrechner.
- (2) Der Rechner sollte im Vektor-Mode auch für kürzere Vektoren aus mehrdimensionalen Datenstrukturen ausreichende Performance zeigen.
- (3) Die Einbettung in das damalige reine IBM-Umfeld des Rechenzentrums sollte problemlos sein. - Dies bezog sich auf Betriebssystem und Steuersprache.
- (4) Der Daten-Transfer zwischen Vektorrechner und Umfeld sollte in beiden Richtungen problemlos sein. - Dies bezog sich auf die interne Zahlen-Darstellung.
- (5) Der Compiler sollte fortschrittlich sein, wie oben dargestellt.

Erste Erfahrungen mit dem Vektorrechner Siemens VP50

Das Forschungszentrum Karlsruhe entschied sich zur Beschaffung eines Vektorrechners Siemens VP50 im Sommer 1987. Es war dies das Einstiegs-Modell der VP-Serie mit einer Nennleistung von 143 Megaflops. Skalar glich die VP50 der Siemens 7890. Benchmark-Tests hatten gezeigt, daß bei Produktions-Codes, die meist nicht in hohem Grade vektorisierbar waren, die Rechenleistung der VP50 im Vergleich zur VP200 recht beachtlich war. Dies, obwohl die VP200 die 4-fache Nennleistung hatte.

Unsere positiven Erwartungen bestätigten sich durchweg. Ein Einzelbeispiel sei hier angeführt [5]. Ein größerer, in den USA auf einem Universalrechner entwickelter Thermofluidodynamik-Code (ca. 25000 Zeilen) konnte in etwa 2 Jahren soweit vektorisiert werden (anfangs auf der Cyber-205, danach auf der VP50), daß die relative Leistung verglichen mit dem Original-Code auf dem Universalrechner Siemens 7890 ein Faktor 20 war. Für typische Anwendungen bedeutete dies eine Reduzierung der Rechenzeit von 10 Stunden auf 30 Minuten! Ein Faktor 10 kam hierbei von der Hardware, während ein weiterer Faktor 2 durch die verbesserte Struktur des vektorisierten Codes zu erklären ist.

Die bessere Anbindung des Vektorrechners erleichterte auch die für Code-Entwicklung erforderliche Zusammenarbeit. Neben der notwendigen Unterstützung durch unser

Rechenzentrum gab es jetzt bei der Optimierung und Code-Validierung auch die Hilfe von Kollegen und, in dem angeführten Fall, eines beratenden Ingenieurs, eines Experten für Stoffdaten.

Ein weiterer Beleg dafür, daß die Beschaffung der eigenen VP50 auf einer richtigen Entscheidung beruhte, zeigte sich darin, wie sie von den Benutzern angenommen wurde. In relativ kurzer Zeit war der Rechner gut ausgelastet.

Die weitere Entwicklung

Die Universität Karlsruhe betrieb inzwischen die Beschaffung eines Vektorrechners SNI S600/20. Als Etappen auf diesem Weg wurde zunächst im Januar 1989 eine Siemens VP400-EX installiert und im zweiten Schritt eine SNI S400/10. Die VP400-EX mit einer Nennleistung von 1700 Megaflops ging Anfang 1991 ans Forschungszentrum Karlsruhe und ersetzte die VP50.

Durch neue vertragliche Vereinbarungen zwischen unserem Forschungszentrum und der Universität Karlsruhe hatten wir immer Zugang zu den an der Universität installierten Vektorrechnern höchster Leistung. Somit wurde die 1983 sehr erfolgreich begonnene Kooperation Forschungszentrum-Universität auf dem Gebiet Hochleistungsrechnen kontinuierlich weiter verfolgt und findet jetzt, 1997, seine Fortsetzung.

Referenzen:

[1] H. Borgwaldt, D. Sanitz, E.G. Schlechtendahl:

KfK Nachrichten, Jahrgang 16 2/84, S. 109-116

[2] W. Schönauer:

Scientific Computing on Vector Computers, North-Holland, Amsterdam, New York, 1987

[3] W. Höbel (Herausgeber):

Abschlußbericht des Arbeitskreises Hochleistungsrechner, Dez. 1986

(unveröffentlicht)

[4] H. Borgwaldt:

Proc. Third International Topical Meeting on Reactor Thermal Hydraulics, Newport (Rhode Island, USA), Oct. 15-18, 1985, Paper 2.1

[5] H. Borgwaldt, H.-H. Frey:

ODIN-Symposium, 13.-14. April 1989, Rechenzentrum der Universität Karlsruhe, S. 59-72

Warum ein gemeinsames Virtuelles Rechenzentrum von Universität und Forschungszentrum?

Adolf Schreiner, Willi Schönauer

Rechenzentrum der Universität Karlsruhe

Wenn man im WWW (World Wide Web) unter der URL

<http://www.uni-karlsruhe.de/Uni/RZ/VRZ>

nachschaut, findet man die folgende Definition:

Virtuelles Rechenzentrum:

Seit dem 06. März 1996 bilden die Rechenzentren der Universität Karlsruhe und des Forschungszentrums ein virtuelles Rechenzentrum. Unter einem virtuellen Rechenzentrum versteht man - in Analogie zu einem virtuellen Unternehmen - die enge Kooperation zweier Einrichtungen zur Erreichung eines gemeinsamen Ziels. Beide Einrichtungen bleiben dabei aber rechtlich selbständig, sie bilden jedoch gemeinsame Entscheidungsgremien und operative Einheiten wie ein virtuelles Direktorium sowie Teams und Arbeitsgruppen. Auch über die gegenseitige Nutzung von Maschinen und sonstigen Ressourcen werden Absprachen getroffen. Im Falle der Gründung des virtuellen Rechenzentrums zwischen dem Forschungszentrum Karlsruhe und der Universität Karlsruhe ist eine Hochgeschwindigkeitsdatenleitung mit 155 Mbit/s die technische Grundlage.

Wie kam es dazu?

Niemand, weder am Forschungszentrum noch an der Universität Karlsruhe, hatte die spontane Idee, ein virtuelles Rechenzentrum Karlsruhe zu gründen. Die Gründung dieses Zentrums ist vielmehr ein wichtiger Meilenstein einer langen Entwicklung - wir sagen bewußt Meilenstein und nicht Abschluß -, denn in einem so dynamischen Bereich wie der Informationstechnik sind Prognosen (insbesondere wenn sie die Zukunft betreffen) höchst problematisch. Das virtuelle Rechenzentrum ist eine adäquate - und man darf es ruhig sagen, eine pionierhafte - Antwort auf die technische Entwicklung im Bereich der Informations- und Kommunikationstechnik. Von einem kann man ausgehen, wenn man aus der Vergangenheit auf die Zukunft schließt: daß die beiden Partner, das Forschungszentrum Karlsruhe und die Universität, auch weiterhin als Vorreiter im Konsens innovative Antworten auf die technische Entwicklung geben werden.

Erlauben Sie uns daher einen kurzen Rückblick auf die Geschichte der Zusammenarbeit beider Einrichtungen auf dem Gebiet der Informationstechnik. Die Tatsache, daß das Forschungszentrum und die Universität in ihrer Forschungsarbeit eng kooperieren, ja gewisse Institute sogar in Personalunion betreiben, war wohl der Anstoß, auch auf dem Gebiet der Informationstechnik nach möglichen Kooperationsformen zu suchen. Bereits in den 60er Jahren, also zu Zeiten der auf sehr teure monolithische Mainframes konzentrierten Datenverarbeitung, bemühte man sich, die DV-Anlagen beider Einrichtungen in einer GmbH zusammenzufassen. Diese in das enge Korsett einer gesellschaftsrechtlich genau definierten Einrichtung gepreßte Kooperation - praktisch im Sinne eines Outsourcing der Datenverarbeitung - erwies sich als nicht praktikabel. Zur Gründung einer solchen GmbH hätte es eines Staatsvertrags zwischen dem Bund und dem Land bedurft. Die GWDG (Gesellschaft für Wissenschaftliche Datenverarbeitung) in Göttingen ist ein Beispiel für eine solche Organisation. Sie versorgt die Universität Göttingen und die dortigen Max-Planck-Institute mit Datenverarbeitungskapazität. Als pragmatischen Ausweg wählte man daraufhin die loseste Art der Kooperation: Es wurde eine 40,8 Kbit-Standleitung zwischen dem Forschungszentrum und der Universität eingerichtet und in der Universität eine Remote-Lese- und Ausgabestation installiert. Die Universität bezog auf diese Weise bis zu 20 % der von ihr benötigten Rechenkapazität von den IBM-Anlagen des Forschungszentrums, und das Forschungszentrum nutzte seinerseits die UNIVAC-Anlage der Universität -allerdings nur in sehr geringem Umfang-, wobei der Datentransfer per PKW erfolgte. 1983 beschaffte die Universität Karlsruhe den ersten Hochleistungs-Vektorrechner im Land, eine CYBER 205, das Forschungszentrum beteiligte sich an dieser Investition mit etwa 20 % und rief seinen Kapazitätsanteil an dieser Anlage ebenfalls über die bestehende 40,8 Kbit-Leitung ab. Ohne die Beteiligung des Forschungszentrums hätten wir die CYBER 205 wohl nie bekommen. Wir sind dem Forschungszentrum dafür sehr dankbar. Die Nutzer beider Anlagen standen in engem Kontakt, bei keinem der von beiden Seiten veranstalteten Seminare und Symposien im Bereich des Supercomputing fehlten Beiträge der anderen Seite. Wir denken da besonders an die Herren Borgwaldt und Sanitz. Auch in Grundsatzfragen wie beispielsweise der Rechnerkonfigurierung holte das Rechenzentrum der Universität stets den Rat der Mitarbeiter des Rechenzentrums und der führenden Benutzer des Forschungszentrums ein. Dieses bewährte sich besonders bei der Beschaffung der Nachfolgeanlage für die CYBER 205, des Siemens Vektorrechners VP 400 und seiner Aufrüstung zum VP 600/20. Auch bei dieser Investition beteiligte sich das Forschungszentrum etwa in gleicher Weise. Als der VP 400 durch den VP 600/20 ersetzt wurde, übernahm das Forschungszentrum den VP 400 und nutzte aber weiterhin den größeren Rechner VP 600/20 für extrem große Programme. Als an der Universität der VP 600/20 zur Ablösung anstand, beschäftigte man sich im Forschungszentrum ebenso wie an der Universität mit einem neuen Konzept des Supercomputing. Die Marktsituation war insofern besonders diffizil, als eine Aufspaltung des Supercomputing stattfand in Produkte, die die bisherige Vektorrechner-technik um Parallelisierungsmöglichkeiten erweiterten und hierbei die Lauffähigkeit bisheriger Programme auch auf der neuen Architektur gewährleisteten und einer anderen Produktreihe, die von der

Workstationentechnik her kommend nur noch auf massive Parellelisierung setzte. Als Entscheidungshilfe kam allerdings hinzu, daß inzwischen die Datenkommunikation im Breitbandbereich bis hin zu 155 Mbit/sec wirtschaftlich tragbar wurde. In Fortsetzung des bereits bisher gepflegten Erfahrungsaustausches zwischen beiden Einrichtungen, aber auch besonders stimuliert durch Gespräche, die mit der Kontaktaufnahme zu auf beiden Seiten neu eingetretenen Führungskräften erfolgten, wurde innerhalb weniger Wochen die Idee geboren, wie man in beiden technisch äußerst interessanten Entwicklungslinien auch künftig präsent sein kann. Das Forschungszentrum und die Universität kamen überein, daß das Forschungszentrum die Vektor-Parallelrechner-Linie betreibt, einen SNI/Fujitsu VPP300, und die Universität ein massiv-paralleles System, eine IBM RS/6000 SP.

Ursprünglich hätte die Notwendigkeit bestanden, jeweils zwei kleinere, den verschiedenen Welten angehörende Systeme zu betreiben. Denn schließlich wollte keine von beiden Seiten auf die weitere Nutzung der bisherigen Vektorrechnerprogramme verzichten und die Weiterentwicklung auf diesem Gebiet versäumen, und andererseits wollte auch keiner in der Workstation-kompatiblen Parallelrechnerwelt die Entwicklung versäumen. Die Übereinkunft beider Einrichtungen ermöglichte es nun, an einem Ort die Vektor-Parallel- und am anderen die Massiv-Parallel-Rechner zu betreiben. Jeder der Partner beteiligte sich an dem vom anderen betriebenen System mit etwa 25 %. Die schnelle 155 Mbit-Datenverbindung zwischen beiden Einrichtungen würde den Nutzern die unterschiedlichen Aufstellungsorte nicht in Erscheinung treten lassen. So hat die Universität zwar nur 4 Prozessoren des VPP300, kann aber die Funktionalität von 16 Prozessoren für Extremjobs nutzen, wobei insbesondere der Hauptspeicher von 32 GB eine wichtige Rolle spielt. Auf ihren eigenen 4 Prozessoren hätte die Universität nur 8 GB nutzen können. Umgekehrt kann das Forschungszentrum die Funktionalität von 100 Prozessoren nutzen, obwohl es nur 28 Prozessoren sein Eigen nennt - in Zukunft wird es sogar die Kapazität und Funktionalität von 256 Prozessoren der IBM SP nutzen können.

Ebenso wichtig ist die Entlastung auf dem Personalsektor. Am Forschungszentrum wird die Systembetreuung für den VPP, an der Universität die für die SP durchgeführt. Ohne das Virtuelle Rechenzentrum hätten beide Rechenzentren jeweils zwei Systeme zu betreuen gehabt.

Zu einer derartig engen Kooperation gehört natürlich ein passender organisatorischer Rahmen. Dieser sollte den einzelnen Partnern größtmögliche Flexibilität erhalten und doch alle Synergieeffekte nutzen, die bei einem gemeinsamen Agieren auftreten können. Hier kann man die heutigen Kommunikationsmöglichkeiten, die Steuerung von Rechnern über beliebige Entfernungen hinweg, die Kommunikation zwischen Menschen über computergestütztes Videoconferencing, wechselseitigen Zugriff auf die Arbeitsplatzrechner des anderen etc. nutzen, also die neuen Möglichkeiten der Telekooperation, des Teleworking. Es liegt daher nahe, nicht nur die Computer des Forschungszentrums und der Universität wechselseitig zu nutzen, sondern auch bei der Betreuung der Rechner, der Benutzerberatung und sonstiger Dienstleistungen in Form von virtuellen Arbeitsgruppen zu kooperieren, virtuell in dem Sinne, daß die Mitglieder einer solchen

Gruppe disziplinarisch weiterhin ihren Einrichtungen angehören, sich die Arbeit jedoch teilen. Auf informelle Weise geschieht eine derartige gegenseitige Unterstützung schon lange über das Internet. Es ist also keine Utopie, Arbeitsteilung auf diesem Weg zu institutionalisieren.

In Anlehnung an den in der Wirtschaft bereits üblichen Begriff des virtuellen Unternehmens nannten wir diesen Verbund das "Virtuelle Rechenzentrum Karlsruhe".

Was sind die wichtigsten Charakteristika eines solchen virtuellen Rechenzentrums?

Das virtuelle Rechenzentrum ist eine Kooperationsform rechtlich weiterhin eigenständiger Rechenzentren. Sie sind verbunden über Hochgeschwindigkeitsleitungen, so daß sowohl Computerleistung als auch Beratung und andere Dienstleistungen ohne wesentliche Einschränkungen von der Vereinigungsmenge der Nutzer der beteiligten Rechenzentren abgerufen werden können.

Arbeitsgruppen der beteiligten Rechenzentren mit bisher gleichen Aufgabenstellungen werden als eine virtuelle Einheit betrachtet. Innerhalb dieser virtuellen Einheit wird Doppelarbeit vermieden. Ein virtuelles Direktorium tritt zu regelmäßigen Sitzungen zusammen. Die Planungen beider Seiten werden in langfristiger Vorschau gegenseitig offengelegt, so daß eine Konvergenz ermöglicht wird.

Die Auflösungsklauseln dieser Kooperation müssen so gestaltet sein, daß kurzfristig für keinen der beteiligten Partner Überraschungen auftreten können.

Welches sind die Vorteile eines solchen virtuellen Rechenzentrums?

Gegenüber einem Alleingang hat man erhebliche Einsparungen an Ressourcen sowie die Freisetzung von Personal für zusätzliche Aufgaben. Der mit der technischen Entwicklung einhergehende zunehmende Betreuungsaufwand für Geräte, Software, Nutzer etc. könnte ohne eine derartige Rationalisierung nur mit wesentlicher Aufstockung des Personals oder überhaupt nicht erbracht werden. Die Dezentralität des virtuellen Rechenzentrums bei gleichzeitiger Überbrückung der Distanz durch die neuen Kommunikationsmedien erlaubt größte Flexibilität. Nur wer die Chancen des Teleworking nutzt, wird künftig im Wettbewerb bestehen können. Gegenüber der GmbH ist der Verzicht auf Overhead sowie der größere Freiraum jedes einzelnen Partners von Vorteil.

Was ist das Hauptproblem des virtuellen Rechenzentrums?

Der Erfolg hängt sehr von der Kooperationsbereitschaft aller Mitglieder des virtuellen Rechenzentrums, vom Direktorium bis zum hochgradig spezialisierten Mitarbeiter, ab. Diese Kooperationsbereitschaft bedarf einer permanenten, intensiven Motivationsarbeit durch die Führungskräfte.

Ganz allgemein:

Das größte Risiko besteht darin, nicht mit der Zeit zu gehen. Eines der Auslesekriterien für das wirtschaftliche Überleben jedweder Einrichtung wird sein: Zählst Du unter denjenigen, die das Potential der Telekommunikation nutzen, zur Spitzengruppe, zum Mittelfeld oder zu den Nachzüglern? Den Letzten beißen bekanntlich die Hunde, aber glücklicherweise sind das Forschungszentrum und die Universität Karlsruhe mit ihrem Virtuellen Rechenzentrum mindestens vom Start weg unter den Ersten.

Zum Schluß noch ein paar Worte zur praktischen Arbeit. Am Anfang hatten wir alle 4 Wochen eine Koordinierungssitzung, auf der die aktuellen Probleme diskutiert wurden. Das waren sehr viele. Wir sind auch im Wechsel in das Forschungszentrum Karlsruhe und in die Universität gefahren. Heute koordinieren wir noch alle 6 Wochen, aber mittels einer Workstation mit Video-Ausrüstung, und die Probleme sind sehr viel weniger geworden. Jetzt gehen wir daran, weitere Arbeitsteilung zu realisieren, in Software-Pflege und Beratung.

Das schwierigste Problem war die Zusammenführung der Benutzernummern der zwei Institutionen in einer einzigen DCE/DFS-Zelle, DCE.VRZ.DE. Aber wir hatten Glück: Das Forschungszentrum Karlsruhe hatte Benutzernummern meist unter, die Universität meist über 10 000. Unsere Benutzer haben eine einheitliche Sicht auf Dateien, Maschinen und Dienste. Auch auf diesem Gebiet haben wir Pionierarbeit geleistet.

Die Kooperation zweier Rechenzentren in einem virtuellen Rechenzentrum bedarf natürlicherweise auch einer vertraglichen Basis. In einer frühen Entwurfsphase dieses Kooperationsvertrags war über den Ausgleich der wechselseitig benutzten Ressourcen zu lesen:

Beide Einrichtungen stimmen jedoch darin überein, daß die zum Einsatz kommenden Verrechnungstechniken die Funktionalität und die Innovationsfähigkeit des virtuellen Rechenzentrums nicht beeinträchtigen dürfen. Sie sollen grobe Ungleichgewichte verhindern, aber nicht in lähmende "Pfennigfuchserie" ausarten.

Das kennzeichnet den Geist der Zusammenarbeit. Die Fachleute haben das dann in besseres "Juristendeutsch" umformuliert, ohne aber den Geist dabei zu ändern.

Nur auf dem Boden eines solchen gegenseitigen Vertrauens kann das virtuelle Rechenzentrum wachsen und gedeihen. Wir wünschen uns alle, daß das so bleibt.

Simulation komplexer Strukturen auf Hochleistungsrechnern: Molekülkristalle aus Fullerenen

Dr. K.-P. Bohnen

Neben Diamant und Graphit ist seit 1985 eine weitere Form des Kohlenstoffs bekannt, die Fullerene. Es handelt sich dabei um hochsymmetrische Moleküle die aus fünf- und sechseckigen Kohlenstoffringen gebildet werden. Der Prototyp ist das C_{60} -Molekül, das in seiner Form einen Fußball widerspiegelt. Die physikalischen und chemischen Eigenschaften dieser Moleküle haben in den vergangenen Jahren großes Interesse gefunden, da man sich von ihnen vielfältige industrielle Anwendungsmöglichkeiten verspricht. Dotierte Fullerene (z.B. K_3C_{60}) zeigen Supraleitung mit Sprungtemperaturen, die für organische Supraleiter sehr hoch sind. Für Informationen über Fullerene wird auf Ref. [1] verwiesen. Trotz des großen Interesses sind noch viele grundlegende Fragen ungelöst. Wie kommt es zur Bildung der Moleküle und wie ist ihre Stabilität? Wie sind die interatomaren Potentiale, die zu Festkörpern aus Fullerenen führen? Welche Eigenschaften haben diese Festkörper etc.? Viele dieser Fragen können heute mittels Simulation auf leistungsfähigen Rechnern beantwortet werden. Da es sich jedoch bei den Fullerenkristallen um sehr große Systeme handelt (die stabile Tieftemperaturstruktur enthält 4 C_{60} -Moleküle in der Einheitszelle), stößt man schnell bei Rechnungen, die ausgehend von der atomaren Information die Eigenschaften des Kristalls vorhersagen wollen, an die Grenzen des verfügbaren Speicherplatzes. Ein Verfahren diese Grenze hinauszuschieben, ist die Benutzung eines Mehrprozessorsystems. Effiziente Nutzung dieser Systeme erfordert Parallelisierung der Computercodes. Am Beispiel der Berechnung von Grundzustandseigenschaften der Tieftemperaturstruktur von C_{60} werden im folgenden erste Erfahrungen mit der Nutzung der VPP-300/16 geschildert.

Eine in der Festkörperphysik wohl etablierte Methode zur Berechnung von Grundzustandseigenschaften, beruht auf der Energieberechnung mit Hilfe der Dichtefunktionaltheorie. Der numerisch aufwendigste Schritt besteht in der Lösung einer inhomogenen Differentialgleichung bei der die Inhomogenität eine Funktion der Lösung ist. Entwicklung der Lösungsfunktionen nach einem geeigneten Satz von Basisfunktionen führt auf ein verallgemeinertes Eigenwertproblem

$$H\psi = \epsilon S\psi$$

In unserem Fall handelt es sich um ein reelles Eigenwertproblem. Die Größe der Matrizen hängt von der Zahl der benutzten Basisfunktionen ab. Für 50 Basisfunktionen pro Atom erhält man bei 240 C-Atomen (4 C_{60} -Moleküle) in der Einheitszelle einen Speicherbedarf > 3.5 GByte, bei 100 Funktionen/Atom liegt er schon über 14 GByte (doppelt genaue Arithmetik wurde vorausgesetzt). Die Genauigkeit der erzielten Ergebnisse hängt von der Zahl der Basisfunktionen ab. Wie die obigen Zahlenbeispiele

zeigen, ist Parallelisierung unumgänglich. Bei der von uns verwendeten Parallelisierungsstrategie wurde großer Wert auf Übertragbarkeit der Programme (im Rahmen des virtuellen Rechenzentrums steht auch eine SP2 mit 256 Knoten zur Verfügung), geringer Aufwand für den Benutzer, effiziente Nutzung der Maschine, Flexibilität und Skalierbarkeit gelegt. Es wurde daher nach verfügbarer Software Ausschau gehalten, die auf MPI, PVM oder HPF aufbaut. Als Erweiterung von LAPACK (Software für Probleme der linearen Algebra) ist seit November 1996 ScaLAPACK verfügbar. Diese wurde daher zur Matrixdiagonalisierung verwendet. Die spezielle Routine PDSYGVX wurde verwendet, mußte aber aus Speicherplatzgründen modifiziert werden. Im folgenden werden Ergebnisse als Funktion der Parameter N (Dimension der Matrix), N_B (Blockgröße bei der Verteilung der Matrix, siehe Abb. 1), N_P Anzahl der verwendeten Prozessoren), N_R , N_C (Anordnung der Prozessoren) und $N_P = N_R \times N_C$ dargestellt.

Aufteilung einer $N \times N$ Matrix in $N_B \times N_B$ BLöcke

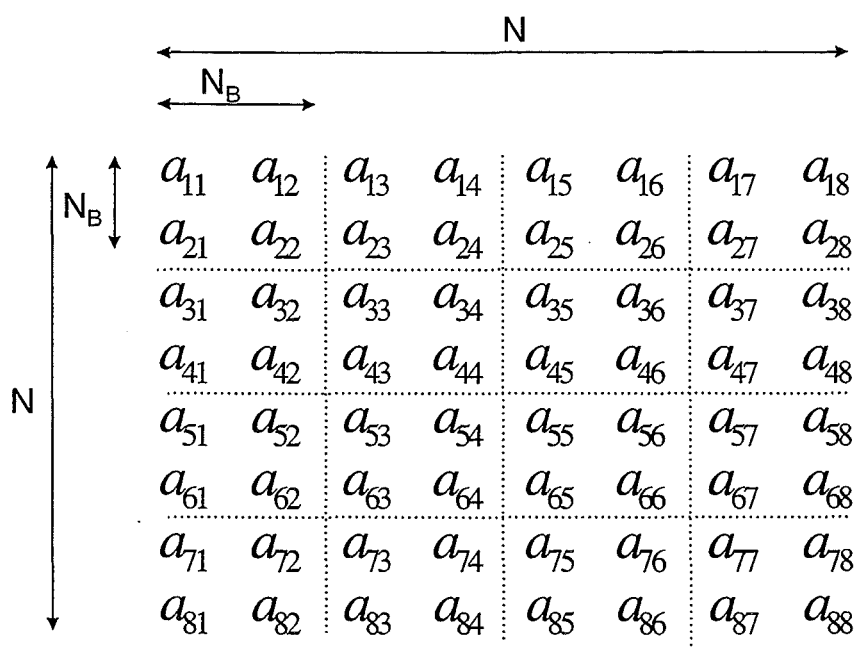


Abb. 1

Lastverteilung einer symmetrischen
8x8 Matrix auf 4 Prozessoren
bei einer Blockgröße 4

$N_R:$	0	1						
$N_C:$	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}
0		a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	a_{28}
			a_{33}	a_{34}	a_{35}	a_{36}	a_{37}	a_{38}
				a_{44}	a_{45}	a_{46}	a_{47}	a_{48}
				a_{55}	a_{56}	a_{57}	a_{58}	
1					a_{66}	a_{67}	a_{68}	
						a_{77}	a_{78}	
							a_{88}	

Prozessor	Elemente
(0,0)	10
(0,1)	16
(1,0)	0
(1,1)	10

Abb. 2

Lastverteilung einer symmetrischen
8x8 Matrix auf 4 Prozessoren
bei einer Blockgröße 1

$N_R:$	0	1	0	1	0	1	0	1										
$N_C:$	0	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}									
	1		a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	a_{28}									
	0			a_{33}	a_{34}	a_{35}	a_{36}	a_{37}	a_{38}									
	1				a_{44}	a_{45}	a_{46}	a_{47}	a_{48}									
	0					a_{55}	a_{56}	a_{57}	a_{58}									
	1						a_{66}	a_{67}	a_{68}									
	0							a_{77}	a_{78}									
	1								a_{88}									

Prozessor	Elemente
(0,0)	10
(0,1)	10
(1,0)	6
(1,1)	10

Abb. 3

Tabelle 1

Blockgröße $N_B = 260$

N	t-Real	CPU	GRID	MFLOPS
8531	1:40	1:38	1x1	1239
14855	2:30	7:24	2x2	310
16233	2:56	8:36	2x2	350
17628	2:11	16:02	3x3	270
19291	2:31	18:29	3x3	310
20299	2:17	16:23	3x3	400
21146	2:05	14:41	3x3	580
23345	2:22	16:41	3x3	590

Stand Juni 1997

Tabelle 2

Matrixdimension = 23345

N_B	t-Real	CPU	GRID	MFLOPS
260	2:22	16:41	3x3	590
325	2:18	16:11	3x3	620

Stand Juni 1997

Es werden nur Ergebnisse für quadratische Anordnung der Prozessoren $N_R = N_C$ präsentiert, da für Diagonalisierungsprobleme eine eindimensionale Topologie deutlich schlechtere Antwortzeiten liefert wie von den Entwicklern der ScaLAPACK erwähnt und durch Testrechnungen bestätigt wurde. Ein sehr kritischer Parameter ist N_B , mit dem sich Lastausgleich und Performance deutlich beeinflussen lassen.

Große Werte von N_B führen zu guter Vektorisierung, haben aber den Nachteil, daß sehr große Puffer für die Nachrichtenkommunikation benötigt werden, was den für Anwendungen verfügbaren Speicher einschränkt. Außerdem führt großes N_B zu sehr schlechter Lastverteilung insbesondere bei der Behandlung reeller symmetrischer Eigenwertprobleme (Abb. 2,3). Das optimale N_B ist schwierig zu bestimmen. Es sollte nach unseren Erfahrungen für die VPP-300/16 bei einer Blockgröße $N_B \sim 200-300$ liegen, bei der die Matrix-Matrix Operationen in BLAS mit 70 % - 80 % der Maximalleistung der Maschine laufen. In Tab. 1 sind die Ergebnisse für festes $N_B = 260$ als Funktion der Matrixdimension N und der verwendeten Anzahl von Prozessoren zusammengefaßt. Das Ergebnis für 1 Prozessor wurde mit der Routine DSYGV aus LAPACK erzielt, während alle Ergebnisse für die Mehrprozessoranwendungen mit PDSVGVX aus ScaLAPACK erzielt wurden, wobei diese Routine aus Speicherplatzgründen leicht modifiziert wurde. Das Skalierungsverhalten mit der Dimension der Matrix und der Zahl der Prozessoren ist sehr zufriedenstellend, die absolute Leistung in MFLOPS/Prozessor sollte für die größten Anwendungen wünschenswerterweise auch bei ~ 1000 MFLOPS liegen. Dies wird in der Zukunft noch weitere detaillierte Analysen der einzelnen involvierten Routinen erfordern. Tabelle 2 zeigt Ergebnisse für die größte Matrix für verschiedene Blockgrößen.

Zusammenfassend läßt sich feststellen, daß die VPP-300/16 eine sehr leistungsfähige Maschine ist, die dem Benutzer neue Möglichkeiten eröffnet. Nutzung als Parallelrechner erfordert entsprechende Investition vom Anwender, da Software z. Zt. weitgehend nicht verfügbar ist. Eine der wenigen Ausnahmen bildet die ScaLAPACK. Auch hier muß der Nutzer natürlich die INPUT-Information in geeigneter Weise bereitstellen. In der Anwendung wurden Leistungen von $9 \times 620 = 5580$ MFLOPS mit relativ geringem Aufwand erreicht. Dies läßt hoffen, daß das Potential der Maschine noch längst nicht ausgereizt ist und daher auch ab-initio Rechnungen für große Fullerenkristalle mit der benötigten Genauigkeit durchgeführt werden können.

Den Herren Schmitz, Sanitz, Geers und Howar vom virtuellen Rechenzentrum der FZK und der Uni Karlsruhe möchte ich für die großzügige Unterstützung bei der Parallelisierung danken.

[1]M.S. Dresselhaus, G. Dresselhaus, P.C. Eklund, Science of Fullerenes and Carbon Nanotubes, Academic Press, San Diego 1996.

Strömungs- und Wärmetechnische Optimierung einer KFZ-Kupplung mittels numerischer Simulation

Lars Eigenmann, Ingo Weißert, Rainer Koch, Sigmar Wittig

Institut für Thermische Strömungsmaschinen

Universität Karlsruhe

Einleitung

Die Kühlung thermisch hochbelasteter Komponenten ist eine Problemstellung, die im Maschinenbau von enormer Bedeutung ist. Die Optimierung des Kühlsystemes erfordert eine exakte Kenntnis der Strömungsvorgänge in den verschiedenen Komponenten des Systems. Eine besondere Herausforderung stellt die Charakterisierung der dreidimensionalen Strömung durch rotierende Bauteile [1,2,6,7] dar.

Zur numerischen Simulation solcher Probleme werden heute hauptsächlich Finiten-Volumen-Methoden eingesetzt. Im Gegensatz zum Experiment lassen sich so zahlreiche Parameter, wie Geometrie und Randbedingungen variieren. Die numerische Simulation liefert detaillierte Informationen im gesamten Strömungsfeld, wodurch sich schon in einem frühen Stadium Bauteiloptimierungen vornehmen lassen.

Als typisches Beispiel wird die numerische Untersuchung einer KFZ-Kupplung (siehe auch [8]) diskutiert. Neuste Untersuchungen haben gezeigt, daß die aerodynamische Auslegung der Kupplungskühlung stark verbessert werden kann. Zielsetzung ist durch optimierte Gestaltung der Bauteile die innere Wärmeabfuhr zu verbessern.

Aufbau der Kupplung

Die Kupplung besteht aus einem Kupplungskorb, der die Kupplungsscheibe umschließt (siehe Abbildung 1). Zur Kühlung sind mehrere Öffnungen in den Kupplungskorb eingebracht. Durch die Rotation des Kupplungskorbes innerhalb des Kupplungsgehäuses ergibt sich eine Durchströmung von Kupplungskorb und -scheibe. Die Durchströmung sorgt für die Wärmeabfuhr von den heißen Teilen. Wärmeleitungsvorgänge können in diesen Untersuchungen vernachlässigt werden.

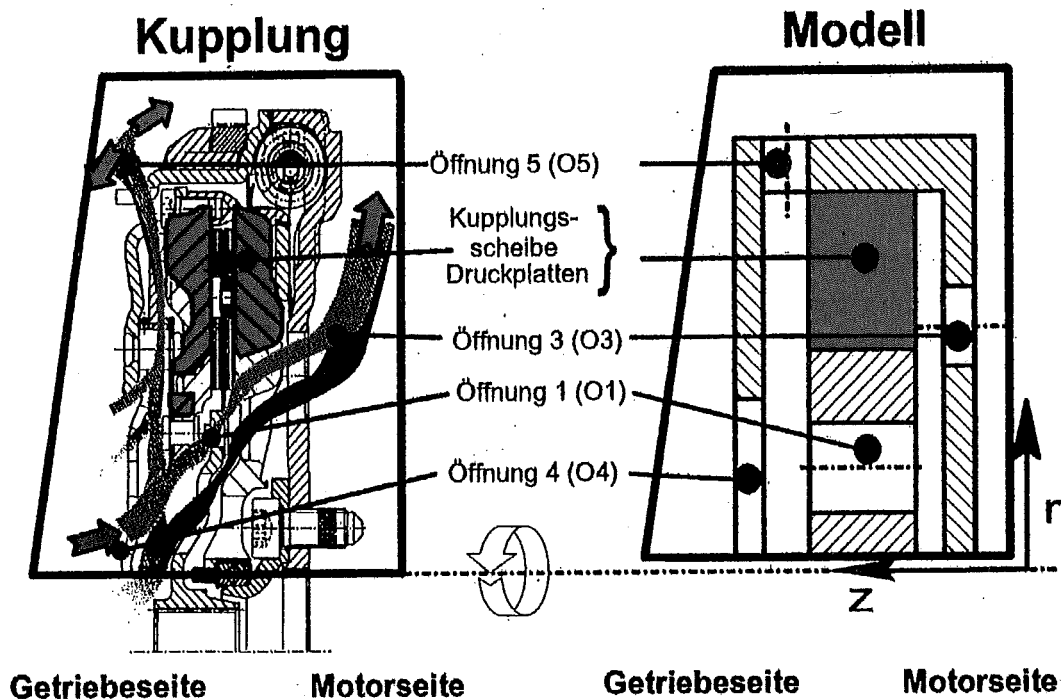


Abbildung 1: Aufbau und Modell der Kupplung

Abbildung 1: Aufbau und Modell der Kupplung

Die Formgebung der KFZ-Kupplung zeichnet sich durch außerordentlich hohe Komplexität aus, so daß eine exakte Diskretisierung der Originalkonfiguration und die anschließende Berechnung nur unter enormem Aufwand möglich wäre. Die vorgestellten Untersuchungen sind daher an stark vereinfachten Modellen durchgeführt worden. Die Hauptkomponenten sind:

Rotierende Kupplungsscheibe:

Sie besteht aus Druckplatte und Scheibe. Beide müssen durch die innere Strömung gekühlt werden.

Rotierender Kupplungskorb:

Dieser umgibt die Kupplungsscheibe und beinhaltet mehrere Öffnungen, die eine interne Durchströmung induziert.

Stillstehendes Kupplungsgehäuse

Hiermit wird das Getriebe mit dem Motor verbunden. Das Gehäuse ist im Gegensatz zur Kupplungsscheibe kühl.

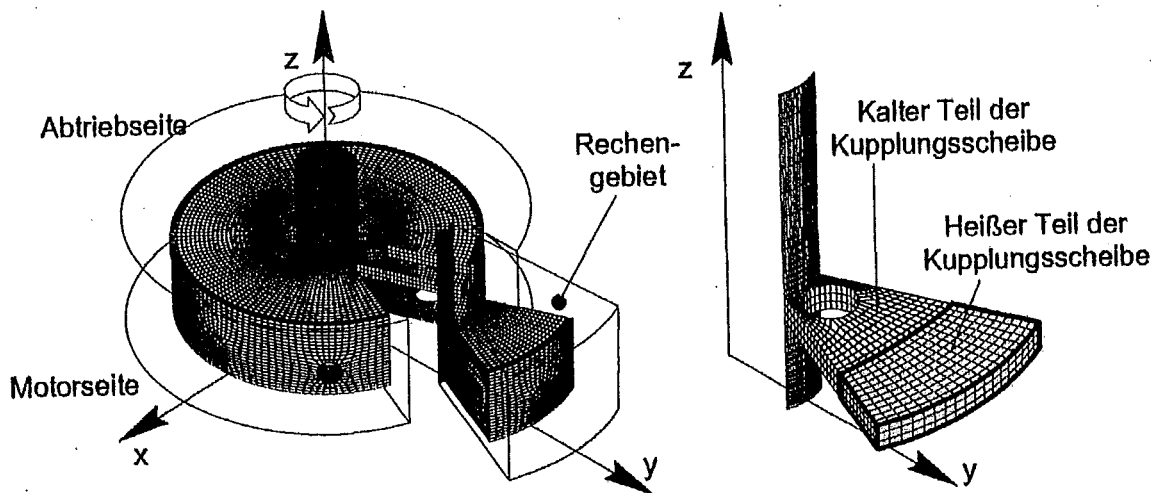


Abbildung 2: Rechengitter der Kupplung

Numerische Beschreibung

Die mathematische Beschreibung von Strömungen durch komplexe, 3-dimensionale Geometrien erfolgt anhand der Reynoldsgemittelten Navier Stokes Gleichungen in Kombination mit Erhaltungsgleichungen für Masse und Energie. Die partiellen Differentialgleichungen für einen unabhängigen Skalar Φ lauten:

$$\frac{\partial(\rho u_i \phi)}{\partial x_i} = \frac{\partial}{\partial x_i} \left[\Gamma_\phi \left(\frac{\partial \phi}{\partial x_i} \right) - \overline{\phi u_i' \phi'} \right] + S_\phi \quad (1)$$

Die Schließung des Gleichungssystems geschieht mittels eines Turbulenzmodelles, das die Verbindung zwischen dem turbulenten Transport von Impuls und Energie und den Strömungsgrößen herstellt. Mit dem Ansatz nach Boussinesq [9] in Kombination mit

dem Standard k-ε Modell wurde auch schon in früheren Arbeiten [6,7] sowohl bei stationären Strömungen als auch bei Strömungen durch rotierende Öffnungen gute Ergebnisse erzielt. Der wandnahe Bereich der turbulenten Strömung wird mittels des logarithmischen Wandgesetzes überbrückt. Für die Diskretisierung der komplexen, dreidimensionalen Geometrie werden die Transportgleichungen auf einem konturangepaßten, strukturierten Gitter gelöst.

Die Diskretisierung der Grundgleichungen erfolgt mit Ansätzen 2. Ordnung aus einer Kombination des von Schneider und Raw [12] vorgeschlagenen MWS-Ansatzes (Mass Weighted Scheme) sowie des von Raithby [11] entwickelte LPS-Ansatzes (Linear Profile Scheme).

Rotierendes Koordinatensystem

Die Betrachtung der Strömung in rotierenden Bauteilen ergibt für das feste Inertialsystem ein instationäres, mit der Drehzahl sich periodisch wiederholendes Stromfeld. Durch den Wechsel des Bezugssystemes in das mitbewegte Koordinatensystem, wird der Strömungstyp zu einem quasistationären Strömungsfall. Hierbei treten in den Navier-Stokes Gleichungen zusätzliche Quellterme auf, die die Coriolis- und Zentripetalkraft beinhalten. Die Transformation zwischen den zwei Bezugssystemen erfolgt mittels

$$\left[\frac{D\vec{w}}{Dt} \right]_B = \left[\frac{D\vec{c}}{Dt} \right]_I - 2\vec{\Omega} \times \vec{w} - \vec{\Omega} \times (\vec{\Omega} \times \vec{r}) - \frac{d\vec{\Omega}}{dt} \times \vec{x} - \vec{a}$$

Bezugs-
=
Inertial-
Coriolis-
Zentrifugal-
Instationärer
Führungs-
Beschleunigung
(2)

Durch die Annahme eines gleichmäßig rotierenden Systems mit stationärem Strömungsfeld reduziert sich Gleichung (2) zu:

$$\vec{S}_u = -2\dot{\rho}\vec{\Omega} \times \vec{w} - \rho\vec{\Omega} \times (\vec{\Omega} \times \vec{r}) \quad (3)$$

Bei der Diskretisierung des Gleichungssystems mit den obengenannten Ergänzungen und den erwähnten Diskretisierungsschemata werden jedoch zusätzliche Terme vernachlässigt.

Diskretisierungsfehler

Im folgenden soll nun der durch die zusätzlichen Terme entstehende Diskretisierungsfehler für eine einfache UPWIND-Diskretisierung in Zylinderkoordinaten abgeleitet werden. Die Rotationsachse ist hierbei die z-Achse.

Der Geschwindigkeitsvektor des ruhenden Bezugssystems lautet

$$\vec{c} = \begin{pmatrix} u_z \\ v_z \\ w_z \end{pmatrix}_{\varphi, r, z} \quad (4)$$

und der des rotierenden Bezugssystemes

$$\vec{c} = \begin{pmatrix} u_z + \omega \cdot r \\ v_z \\ w_z \end{pmatrix}_{\varphi, r, z} \quad (5)$$

Die Diskretisierung eines finiten Volumens (siehe Abb. 3) nach dem Gauß'schen Satz ist für den

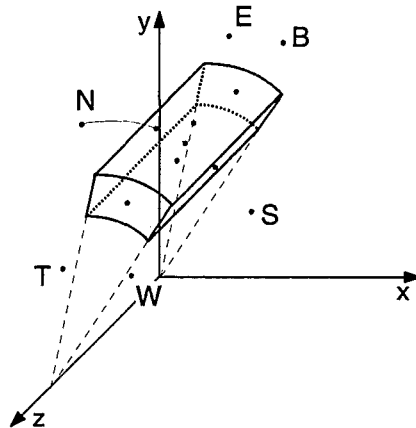


Abbildung 3: Kontrollvolumen in Zylinderkoordinaten

konvektiven Anteil durch

$$\begin{aligned} \iiint_V \frac{\partial}{\partial x_i} (\rho u_i \Phi) dV &= \iint_A (\rho u_i \Phi) dA \\ &\approx \rho_e u_e A_e \Phi_e - \rho_w u_w A_w \Phi_w + \rho_n u_n A_n \Phi_n - \rho_s u_s A_s \Phi_s \\ &\quad + \rho_t u_t A_t \Phi_t - \rho_b u_b A_b \Phi_b \end{aligned} \quad (6)$$

und dem diffusiven Anteil ($\Gamma_\Phi = \mu_t$)

$$\begin{aligned} \iiint_V \frac{\partial}{\partial x_i} \left(\Gamma_\Phi \frac{\partial \Phi}{\partial x_i} \right) dV &= \iint_A \left(\Gamma_\Phi \frac{\partial \Phi}{\partial x_i} \right) dA \\ &\approx \frac{\mu_{te} A_e}{r_e \Delta \varphi_e} \Delta \Phi_e - \frac{\mu_{nw} A_w}{r_w \Delta \varphi_w} \Delta \Phi_w + \frac{\mu_{tn} A_n}{r_n} \Delta \Phi_n - \frac{\mu_{ts} A_s}{r_s} \Delta \Phi_s \\ &\quad + \frac{\mu_{tt} A_t}{r_t} \Delta \Phi_t - \frac{\mu_{tb} A_b}{r_b} \Delta \Phi_b \end{aligned} \quad (7)$$

gegeben. Die Koeffizienten Φ_i und $\Delta \Phi_i$ in (6) und (7) müssen durch eine im Inneren der Rechenzellen bekannten Wert ersetzt werden. Die Terme unter Verwendung des UPWIND-Ansatzes für den konvektiven Anteil ergeben sich nach

$$\Phi_e = \Phi_p \quad (8)$$

sowie mit Hilfe des Zentral Differenzen Ansatzes für den diffusiven Anteil nach

$$\Delta\Phi_e = (\Phi_E - \Phi_P) \cdot (9)$$

Beim Einsetzen von Gl. (8) und (9) und Substitution des lokalen Geschwindigkeitsvektors (5) in die linke Seite der Gleichung (3) und des stationären Geschwindigkeitsvektors (4) in die rechte Seite ergeben sich bei Vergleich der beiden Seiten zusätzliche Terme aus Konvektion und Diffusion in Abhängigkeit der Kreisfrequenz ω der Form

$$S_{uz} = k_{uz} \cdot \omega \cdot \Delta r \quad , (10)$$

wobei Δr wiederum von der Richtung der Radialgeschwindigkeit abhängig ist.

Somit stellt sich bei Verwendung des rotierenden Bezugssystemes und dem UPWIND-Schemata ein Fehler erster Ordnung ein. Dieser Fehler ist gleichbedeutend mit dem wohlbekanntem Diskretisierungsfehler der numerischen Diffusion.

Wie in [8] gezeigt wurde, kann dieser Fehler durch die Verwendung von hochauflösenden Diskretisierungsansätzen verringert werden. Eine gänzliche Vermeidung ist bei rotierendem Bezugssystem nicht erreichbar. Bei Einsatz feiner Rechengitter bzw. höherwertiger Diskretisierungsansätze kann dieser Einfluß auf ein vernachlässigbares Maß reduziert werden. Dies bedingt aber im Gegenzug höhere Rechenzeiten als auch ein Ansteigen der Speicherbedarfs, wodurch für solch komplexe Aufgaben Supercomputer benötigt werden.

Randbedingungen der numerischen Berechnungen

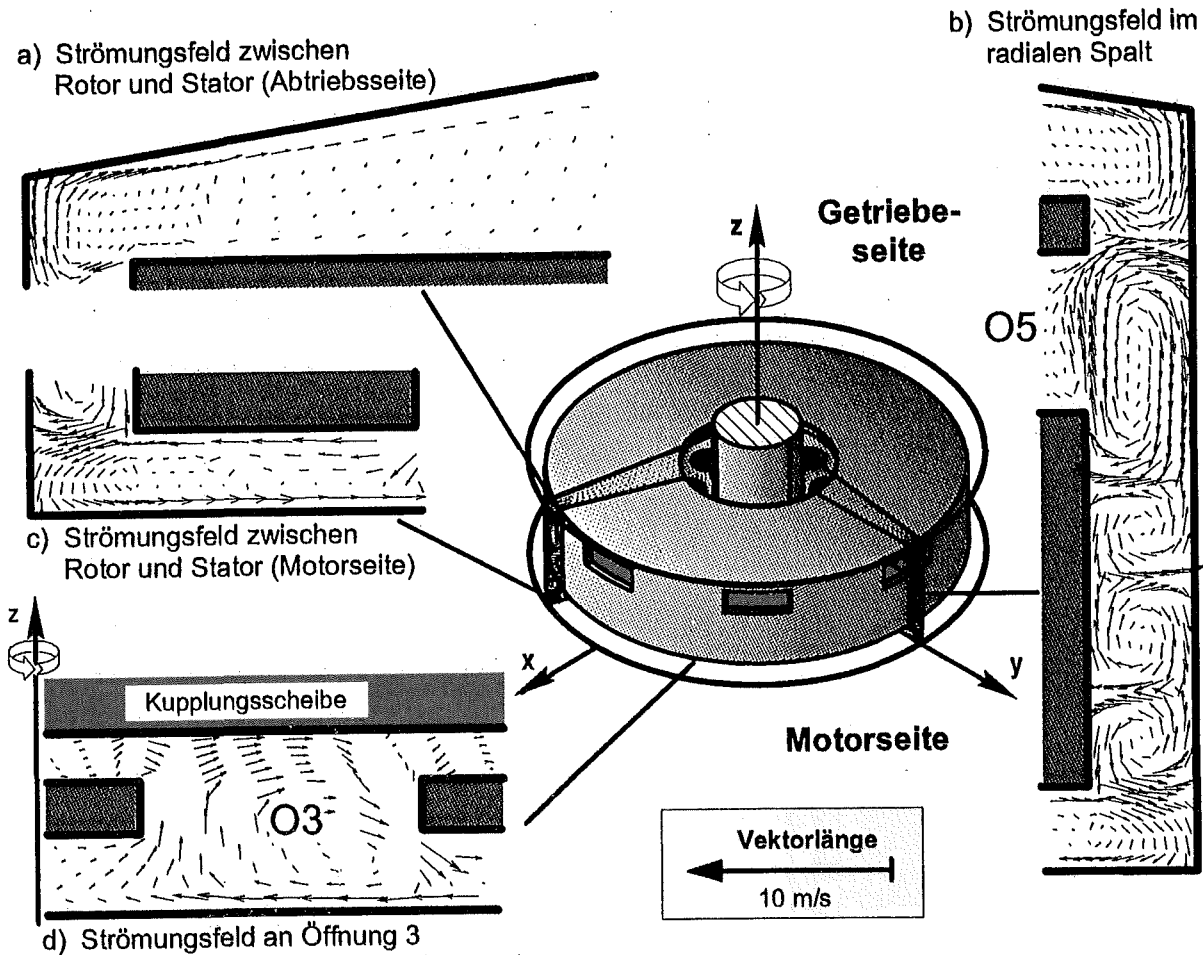
Der heiße Bereich der Kupplungsscheibe ist in Abbildung 2 dargestellt. Er wird in den vorgestellten Untersuchungen mit einer Wandtemperatur von $T_w = 200^\circ\text{C}$ angenommen. Die restlichen Komponenten besitzen eine mittlere Wandtemperatur von $T_w = 100^\circ\text{C}$. Konduktion in den einzelnen Bauteilen wird bei diesen numerischen Untersuchungen nicht berücksichtigt.

Da die Kupplungsgeometrie in Umfangsrichtung periodisch ist, kann die zur Berechnung herangezogene Geometrie auf 1/8 der Gesamtgeometrie reduziert werden. Die Anzahl der Knotenpunkte betrug bei allen Konfigurationen 40 in radialer, 45 in axialer sowie 19 in Umfangsrichtung, so daß sich eine Gesamtzahl von 34200 Gitterpunkten ergab.

Ergebnisse

Bisherige Bauform

In einem ersten Schritt wurde die Kupplung, so wie sie bisher gestaltet war, untersucht. Die Strömung zwischen dem Kupplungsgehäuse und dem Kupplungskorb ist in Abbildung 4 dargestellt.



Die motorseitige Kupplungsscheibe wird aufgrund der Öffnung O3 recht gut gekühlt. (siehe Abb. 4d). Durch die Positionierung der Öffnung O3 direkt vor der heißen Fläche der Kupplungsscheibe wird der Wärmeübergang aufgrund der sich in der Öffnung einstellenden Wirbelströmung verbessert. Dieses Prinzip zur Optimierung des Wärmeüberganges an rotierenden Scheiben [4] durch deren Belochung wird unter anderem bei der Kühlung von Scheibenbremsen angewandt.

Die getriebeseitige Kühlung durch konvektiven Wärmetransport ist im Gegensatz hierzu sehr gering. Ein gerichteter Stofftransport vorbei an der heißen Kupplungsscheibe findet nicht statt (siehe Abb. 4a). Dies verdeutlicht der geringe Massenstrom, der aus dem Kupplungskorb an der Öffnung O5 austritt. Dadurch können sich in dem davorliegenden Spalt zwischen Kupplungskorb und -gehäuse Taylorwirbel ausbilden, deren Geschwindigkeitsgradienten einen hohen Stoffaustausch induzieren, aber zur Kühlung der innenliegenden, heißen Bauteile keinen Beitrag leisten. Insgesamt zeigt die bisherige Bauform der Kupplung keine ausreichende Kühlung der Kupplungsscheibe, wodurch ein ausfallsicherer Betrieb nicht gewährleistet werden kann.

Optimierte Version

Zur Verbesserung der Kühlung der heißen Kupplungsscheibe wurden daher am Institut für Thermische Strömungsmaschinen zahlreiche numerischen Untersuchungen durchgeführt. Die Ergebnisse der optimierten Kupplungsgeometrie sollen nun im folgenden vorgestellt werden.

Wesentliche Verbesserungen konnten mit radialen Rippen erreicht werden, die zu einer Erhöhung des konvektiven Transportes durch die Bohrungen führen. Die berippte Ausführung der Kupplungsscheibe ist in Abb. 5 dargestellt.

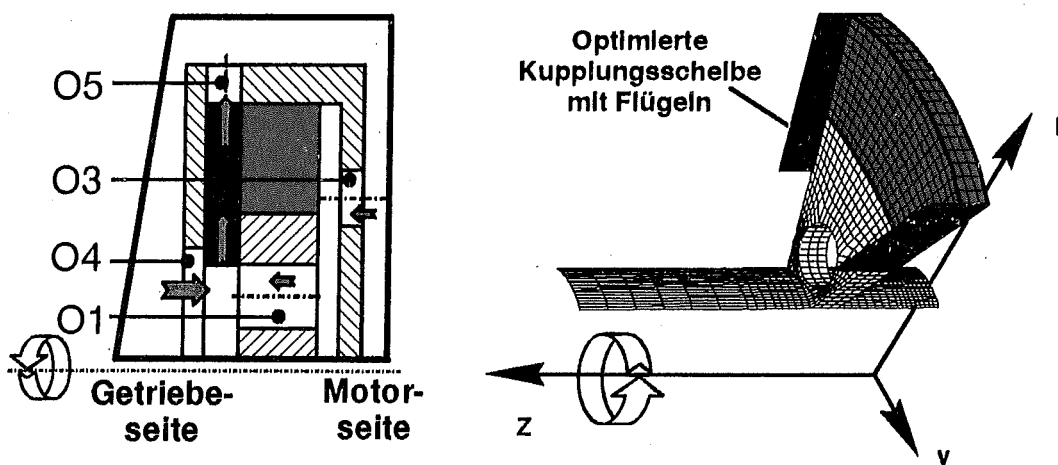


Abbildung 5: Berippte Kupplungsscheibe

Durch die Rotation der Kupplungsscheibe wird die Strömung in tangentielle Richtung beschleunigt (siehe Abb. 6), wodurch sich ein radialer Druckgradient in dem Kupplungskorb einstellt. Die Berippung der Kupplungsscheibe vergrößert damit die Strömungsgeschwindigkeit der Luft in der Kupplung. Das hier angewandte Prinzip entspricht dem eines Radialverdichters. Zusätzlich trägt die Berippung zu einer Erhöhung der Turbulenz bei, welche sich positiv auf den Wärmeübergang auswirkt. Hingegen zeigt die Berippung wenig Einfluß auf die Sekundärströmung an der Öffnung O3, an welcher aber schon bei der Basiskupplung eine ausreichende Kühlung vorhanden ist (s.o.).

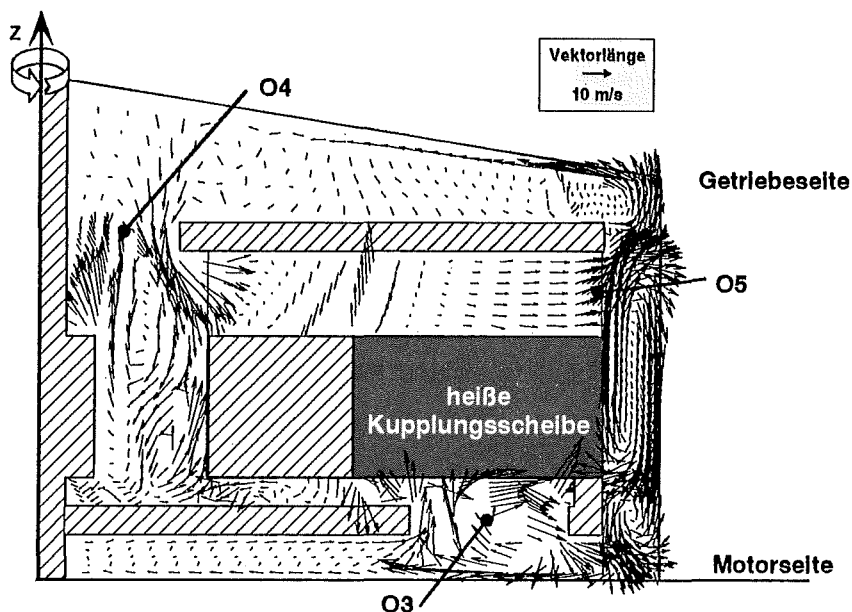


Abbildung 6: Strömungsfeld der berippten Kupplung

Ein quantitativer Vergleich zwischen bisheriger Ausführung und optimierter Kupplung ist in Abbildung 7 anhand der Massenströme durch die einzelnen Bohrungen dargestellt. Deutlich zeigt sich eine Verbesserung des Massenstromes in der Bohrung O5 um den Faktor 4. Auch durch die übrigen Bohrungen konnte der Luftmassenstrom deutlich gesteigert werden. Die damit verbundene Kühlwirkung gestattet somit einen ausfallsicheren Betrieb der Kupplung auch unter extremen Belastungen.

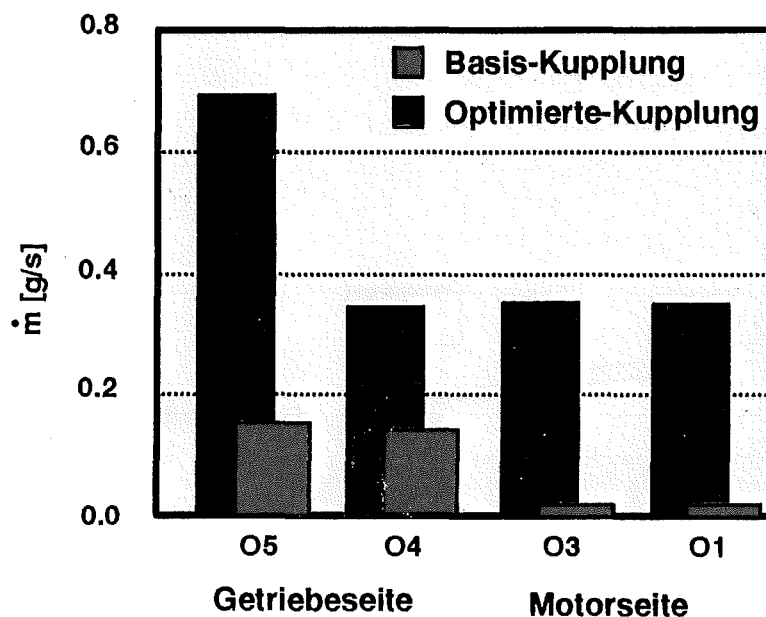


Abbildung 7: Vergleich der Massenströme

Zusammenfassung

Zur Optimierung der konvektiven Kühlung einer rotierenden Kupplung wurden mehrere Konfigurationen untersucht (siehe auch [8]), wobei Strömung und Wärmeübergang mit einem numerischen Verfahren an vereinfachten Modellen berechnet wurde. Im Zuge dieser Untersuchungen konnten zwei maßgebliche Effekte identifiziert werden, durch deren Anwendung der Wärmeübergang signifikant erhöht werden kann. Diese Effekte lassen sich durch die Bezeichnungen 'Radialverdichterprinzip' für die getriebeseitige und 'Scheibenbremsenprinzip' für die motorseitige, heiße Kupplungsscheibe beschreiben.

Die numerischen Ergebnisse, die anhand einer vereinfachten Geometrie der realen Kupplung erhalten wurden, gaben deutliche Hinweise zur Verbesserung des Wärmeüberganges. Trotzdem ist die Übertragung der quantitativen Resultate auf die realen Kupplungen wegen der getroffenen Vereinfachungen aber nicht ohne weiteres möglich, jedoch konnten die wesentlichen Strömungs- und Wärmeübergangseffekte für die unterschiedlichen Konfigurationen herausgearbeitet werden. Dies betrifft vor allem die Belochung der Scheiben in unmittelbarer Umgebung der heißen Segmente und die Einbringung von Nuten bzw. Stegen, die als 'Turbulenzgeneratoren' wirken.

Danksagung

Die vorliegenden Arbeiten wurden von der LuK GmbH & Co, Germany gefördert. Für die Zusammenarbeit und die Unterstützung sei Herrn Dr. Albers, Herrn K. Kasper und Herrn M. Lipps der LuK GmbH gedankt.

Literatur

- [1] W. F. McGreehan and M.J. Schotsch, Flow characteristics of long orifices with rotation and corner radiusing, ASME 87-GT-162, 1987
- [2] N. Hay and A. Spencer, Discharge coefficients of cooling holes with radiused and chamfered inlets, ASME 91-GT-269, 1991
- [3] A. Lichtarowicz, R.K. Duggins, and E. Markland, Discharge coefficients for incompressible non-cavitating flow through long orifices, Journal Mechanical Engineering Science, Vol. 7, No. 2, 1965
- [4] J. M. Owen, Fluid flow and heat transfer in rotating disc systems, Heat and Mass Transfer in Rotating Machinery, Hemisphere Publishing, New York, 1984
- [5] W. Waschka, T. Scherer, S. Kim, and S. Wittig, Study of heat transfer and leakage in high rotating stepped labyrinth seals, ISROMAC-4, 1992
- [6] S. Wittig, S. Kim, R. Jakoby, and I. Weissert, Experimental and numerical study of orifice discharge coefficients in high speed rotating disks, ASME 94-GT-142, 1994
- [7] S. Wittig, S. Kim, T. Scherer, R. Jakoby, and I. Weissert, Durchfluß an rotierenden Wellen- und Scheibenbohrungen und Wärmeübergang an rotierenden Wellen, Abschlußbericht, Forschungsvereinigung Verbrennungskraftmaschinen (FVV), Vorhaben Nr. 465 und 536, Heft R 574, 1995
- [8] S. Wittig, S. Kim, T. Scherer, I. Weissert, Numerical Study for Optimizing Heat Transfer in High Speed Rotating Components, International Symposium on Transport Phenomena and Dynamics of Rotating Machinery (ISROMAC-6), Honolulu, Hawaii, Vol. II, 1996
- [9] M. Jischa, Konvektiver Impuls-, Wärme,- und Stoffaustausch, F. Vieweg & Sohn, Braunschweig/Wiesbaden, 1982
- [10] B.E. Launder and D.E. Spalding, The numerical computations of turbulent flows, Computer Methods in Applied Mechanics and Engineering, Vol. 3, pp. 269-289, 1974
- [11] G.D. Raithby, Skew upstream differencing schemes for problems involving fluid flow, Computer Methods in Applied Mechanics and Engineering, Vol. 9, pp. 153-164, 1976
- [12] G. Schneider, M. Raw, A skewed, positive influence Coefficient Upwinding Procedure for Control-Volume-Based Finite Element Convection-Diffusion Computation, Numerical Heat Transfer, Vol. 8, pp. 1-26, 1987

Einsatz von Hochleistungsrechnern zur Simulation physikalischer Phänomene

Parallelisierung des Karlsruher Hochstrom-Dioden-Simulationscodes KADI2D

Matthias Krauß

Hauptabteilung Informations- und Kommunikationstechnik

Forschungszentrum Karlsruhe

Einleitung

Im Forschungszentrum Karlsruhe werden seit einigen Jahren Hochleistungs-Ionenquellen an einer eigens dafür entwickelten experimentellen Anlage, dem Hochspannungs-Impulsgenerator KALIF, untersucht. Die physikalischen Vorgänge, die sich in Hochleistungs-Ionendioden abspielen, sind extrem komplex. Sie sind bis heute noch nicht vollständig verstanden. Ein Grund hierfür liegt in einer Reihe von sog. Instabilitäten, von denen einige experimentell nachgewiesen sind, andere jedoch nur vermutet werden. Es gibt aber gewichtige experimentelle Anhaltspunkte für ihre Existenz in Hochleistungs-Dioden.

Die experimentellen Forschungen alleine reichen jedoch nicht aus, um zu umfassenden Antworten zu kommen, auf deren Basis dann leistungsstärkere Dioden entwickelt werden können, deren Einsatz von der Grundlagenforschung hochverdichteter Materie bis hin zur Kernfusion reicht.

Daher wurde, ebenfalls im Forschungszentrum, mit der Entwicklung eines Dioden-Simulationscodes, KADI2D, begonnen. Numerische Simulationen bieten gegenüber den recht starren experimentellen Geräten eine hohe Flexibilität bei der Realisierung verschiedener Lösungsstrategien. Diese reichen von der Veränderung der komplexen Dioden-Geometrie bis hin zur Implementierung verschiedener Emissionsmodelle, welche auch in anderen Zusammenhängen von großem Interesse sind und deren Untersuchung ein aktuelles Forschungsgebiet darstellt.

Kernstück des Simulationscodes ist ein Maxwell-Lorentz-Löser für das hoch nichtlineare hyperbolische Differentialgleichungssystem, bestehend aus den elektromagnetischen Feldgleichungen, den Maxwell-Gleichungen, und den als Lorentz-Gleichungen bekannten Bewegungsgleichungen für geladene Teilchen.

Zur numerischen Lösung der Feldgleichungen werden explizite Finite-Differenzen Verfahren (staggered-grid-Leapfrog Verfahren), sowie alternativ dazu explizite Finite-Volumen Verfahren (vom Typ Godunov) verwendet.

Die ersten Verifikationsrechnungen wurden auf einem kartesischen Gitter durchgeführt. Zur Zeit werden strukturierte randangepaßte Gitter verwendet. Um zu verlässlichen, d.h. mit relativ kleinen statistischen Fehlern behafteten Ergebnissen zu kommen muß einerseits mit einem sehr feinen Gitter, andererseits mit ca. 10^6 Teilchen gerechnet werden. Dies macht sowohl den Einsatz von Hochleistungsrechnern mit ihren modernen DV-Techniken wie das parallele Rechnen, als auch die Verwendung effizienter numerischer Verfahren wie Gitterverfeinerung auf der Basis adaptiver Methoden unbedingt notwendig. In einem ersten Schritt wurde der reine Lorentz-Löser ohne den Maxwell-Löser mit überzeugendem Ergebnis parallelisiert. Hierzu wurde ein Message-passing basiertes Parallelisierungskonzept verwendet. Die konkrete Programmierung wurde im MPI –Standard realisiert.

Ziel der numerischen Entwicklung ist die Bereitstellung eines 3-dimensionalen zeitabhängigen Algorithmus zur Simulation der wesentlichen physikalischen Prozesse in einer Hochleistungs-Ionendiode als effizientes Werkzeug zur Unterstützung des Experimentators bei der Konstruktion und dem Design zukünftiger leistungsfähigerer Anlagen.

Wie bereits erwähnt, sind die Eigenschaften einer Hochleistungs-Ionendiode wegen den zugrundeliegenden hochgradig nicht linearen Gleichungen sehr komplex. Dies bedingt höchste Sorgfalt sowohl bei der Auswahl der physikalischen und numerischen Modelle, als auch bei der Evaluierung der numerischen Ergebnisse, d.h. bei der Auswahl geeigneter, aussagekräftiger Testbeispiele.

Es würde den Rahmen dieses Vortrages sprengen, auf alle Details der zugrunde liegenden Theorie, sowie deren numerischen Feinheiten einzugehen. Vielmehr sollen im folgenden nur jene Schritte, die zur Entwicklung eines aussagekräftigen Dioden-Simulationscodes notwendig sind, veranschaulicht werden.

Hierzu wird in Kapitel 2 zunächst die KALIF-Anlage in ihrem aktuellen Ausbau erläutert.

Außerdem wird kurz auf Zielsetzung und Aufgabenstellung der Hochleistungs-Ionendiodenforschung, sowie den daraus resultierenden Leistungsausbau der Anlage, eingegangen.

In Kapitel 3 wird die Methode der Simulation, ihre Zielsetzung, sowie ihre Vorteile gegenüber dem reinen Experiment am Beispiel der Ionendiode zunächst allgemein dargestellt.

In Kapitel 4 werden dann die konkreten, d.h. auf die Diode bezogenen Probleme, die bei der Erstellung eines Simulationscodes zu lösen sind, behandelt. Hierzu gehören neben der Festlegung des Simulationsgebietes die eigentliche Modellbildung, sowie die numerischen Approximationen.

In Kapitel 5 wird dann der Bogen zur Datenverarbeitung geschlagen. Es wird gezeigt, warum der Einsatz von Hochleistungsrechnern notwendig und hilfreich ist. Es werden verschiedene Alternativen aufgezeigt.

Schließlich werden die verschiedenen Parallelisierungsstrategien, ihre Vor- und Nachteile erläutert. Die mit diesem Verfahren auf dem Vektor-Parallelrechner VPP300 erzielten Ergebnisse werden in einem abschließenden Kapitel vorgestellt.

Die experimentelle Anlage KALIF

Aufbau der Anlage

Bevor wir uns der Simulation einer Hochleistungs-Ionendiode zuwenden, wollen wir uns ein Bild von der gesamten Anlage machen, denn wir werden in einem späteren Kapitel

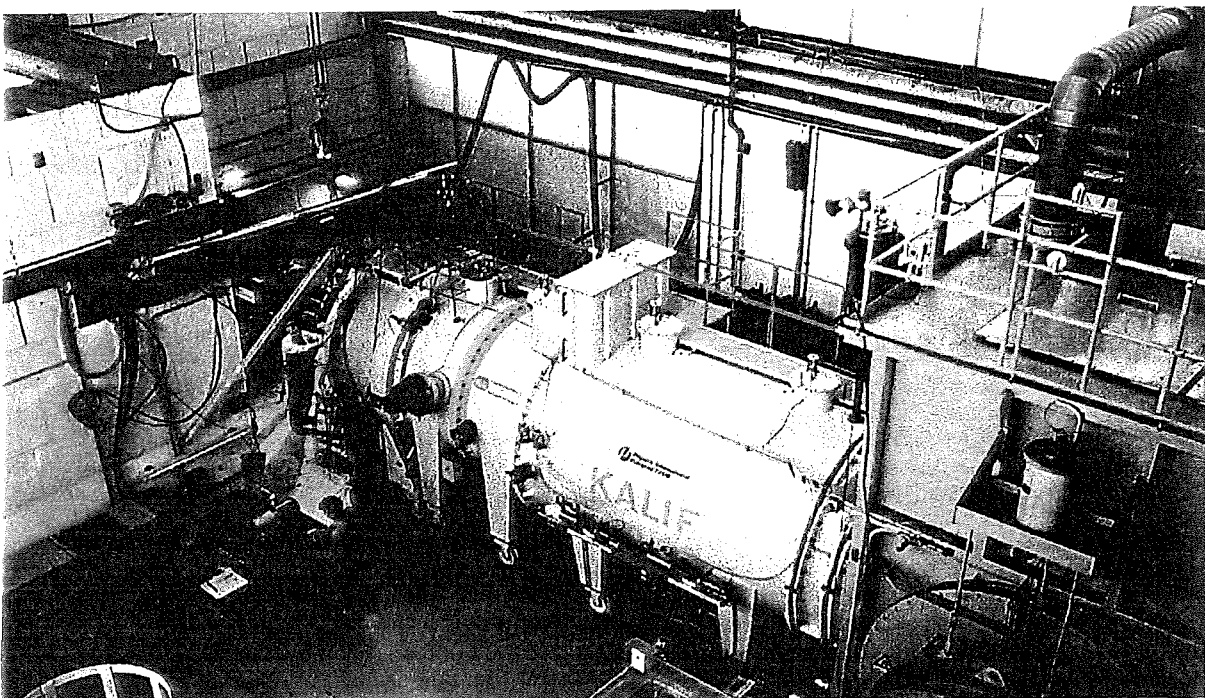


Abb.1: Gesamtbild der KALIF-Anlage (die Ionen-Quelle ist verdeckt)

eben diese reale Geometrie der Anlage in ein abstraktes, für formale Rechnungen geeignetes Rechengebiet übertragen, was ohne detaillierte Kenntnis der Anlagen-Geometrie nicht möglich ist. Die Anlage (s. Abb. 1) besteht im wesentlichen aus drei Teilen, dem Marx-Generator, dem Wellenleiter mit Zwischenspeicher und der Ionendiode (s. Abb.2), die mit diversen Meß- und Beobachtungsinstrumenten ausgestattet ist [1, 2].

Aber erst das Zusammenspiel aller drei Teile ermöglicht die Erzeugung von Ionenströmen hoher Leistung und Qualität. Vereinfacht kann man sich das folgendermaßen vorstellen:

Der elektromagnetische Impuls - der Träger der Energie, die notwendig ist um Ionen in der Diode frei zusetzen - wird zunächst im Marx-Generator erzeugt. Von dort wird er in einem koaxialen Pulsleiter zur Diode transportiert, welche unmittelbar am Ende des Wellenleiters sitzt (s. Abb. 2) und wo er schließlich radial eingespeist wird. Dieser

elektromagnetische Impuls, hat eine Pulsbreite von 50-100 ns und besitzt eine Maximalamplitude zwischen 0.5 und 2 MV .

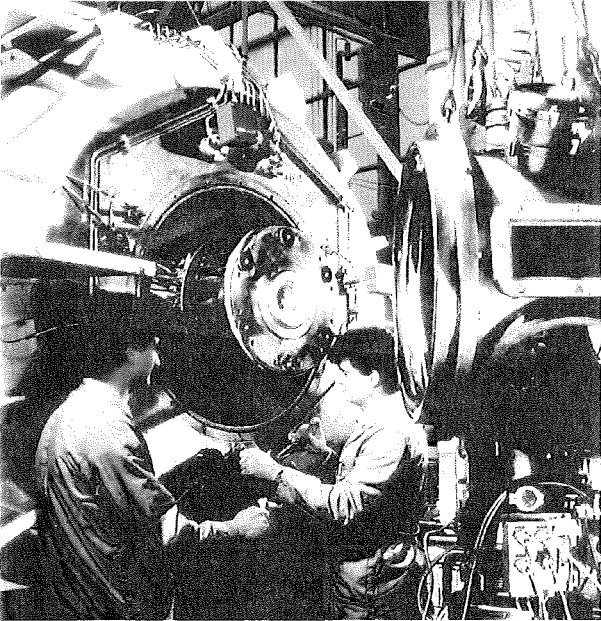


Abb. 2: Eingebaute Ionen-Quelle

Aufgrund der enorm hohen Spannung zwischen den beiden Elektroden übersteigt die elektrische Feldstärke im Vakuum innerhalb kürzester Zeit den Schwellenwert von 20 MV/m im überwiegenden Teil der Diode. Dies führt an der Kathode zur spontanen Emission von Elektronen und im speziell präparierten Gebiet der Anode zur Emission von Ionen. Im Verlaufe der nächsten Nanosekunden spielen sich in der Diode

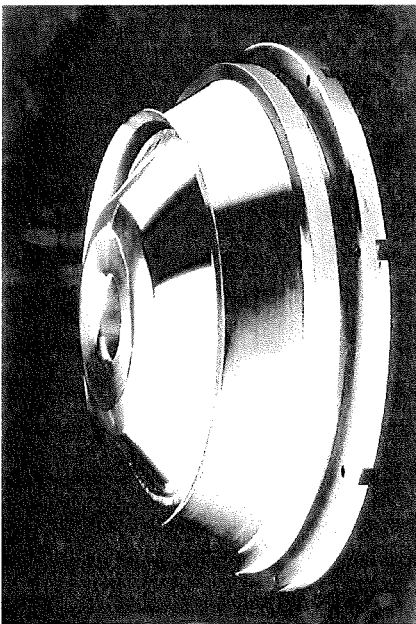


Abb. 3: Kathode

verschiedene Vorgänge ab, die untereinander über komplexe Wechselwirkungen

elektromagnetischer Natur zusammenhängen und die letztendlich zum Entstehen eines mehr oder weniger gebündelten Ionenstrahls führen. So werden die Elektronen durch externe, aber auch durch selbsterzeugte Magnetfelder abgelenkt.

Ferner driften sie wegen der Rotationssymmetrie der Diode in Θ -Richtung. Die Rotationssymmetrie der Diode ist in Abb. 3 deutlich zu erkennen. Die Dioden-Anode ist mit einer speziellen Folie präpariert. Hier kommt es zur Emission von Ionen. Aufgrund äußerer und innerer elektromagnetischer Felder werden die Ionen in Richtung Kathode beschleunigt. Sie verlassen schließlich den Innenbereich der Diode durch die ionendurchlässige Kathode (Mylarfolie) und gelangen in einen gasgefüllten Driftbereich, wo sie neutralisiert werden. Ziel ist es, diesen Ionenstrahl möglichst gut zu fokussieren. Dies gelingt nur, wenn die elektromagnetischen Felder in der Diode genau bekannt sind. Hierzu müssen aber sowohl die Ionen- als auch die Elektronenbahnen berechnet werden. Diese Rückkopplung der Ionen- und Elektronenbahnen an die elektromagnetischen Felder machen aus den zunächst linearen Maxwell-Gleichungen ein hochgradig nichtlineares Maxwell-Lorentz-System.

Einsatzgebiete intensiver Ionenstrahlen

Bei der KALIF-Anlage (LI steht für Leicht-Ionen) besteht der Ionenstrahl aus Protonen, die eine kinetische Energie von 1.7 MeV erreichen. Dadurch ist es möglich, im Fokus des Ionenstrahls eine Leistungsdichte von 1 TW/cm² zu erzeugen. Mit solchen Strahlen können in Targets Temperaturen von ca. 1 Mega-Kelvin, sowie Drucke von ungefähr 1 Mega-bar erzeugt werden. Diese hohen Temperaturen und Drucke führen in der Materie zu extremen Zuständen. Die Erforschung der Eigenschaften solcher hochverdichteter Materie ist einerseits für die Grundlagenforschung, andererseits auch für die Kernfusions-Technologie von allgemeinem Interesse.

Ziel künftiger Forschungsarbeiten

Wie bereits erwähnt, wird der Ionenstrahl in dem an den Diodenbereich sich anschließenden gasgefüllten Driftbereich fokussiert. Je besser die Fokussierung, desto höher sind die erreichbaren Leistungsdichten. Mit der zur Zeit betriebenen Anlage können Ionenstrahlen mit einer Strahldivergenz von ca. 17 mrad erzeugt werden. Mittelfristig möchte man die Strahldivergenz auf 10 mrad reduzieren. Hierzu ist es jedoch notwendig, jene Mechanismen zu verstehen, welche die Fokussierungseigenschaften der Ionendiode negativ beeinflussen. Dabei spielen, so vermutet man, eine Reihe von Instabilitäten eine entscheidende Rolle. Beispielhaft sei hier die Diokotroninstabilität herausgegriffen. Die Diokotroninstabilität wurde von O. Buneman während des zweiten Weltkrieges entdeckt. Buneman war zu dieser Zeit mit Forschungsarbeiten zur Radartechnologie beschäftigt. Die Ursache für die Diokotroninstabilität liegt, vereinfacht ausgedrückt, in einer Scherung im Geschwindigkeitsprofil der Elektronen. Ein Analogon in der Hydrodynamik ist die Kelvin-Helmholtz-Instabilität: in laminar

strömenden parallelen Schichten, die einen Geschwindigkeitsgradienten aufweisen, führt eine kleine Störung zu Turbulenzen, bzw. übertragen auf den elektromagnetischen Fall, zu der Diokotroninstabilität.

Die Auswirkungen der Diokotron-Instabilität auf einen Strahl geladener Teilchen ist in den Abbildungen 4a– 4c deutlich zu erkennen: der ursprünglich räumlich eng lokalisierte Elektronenstrahl verwischt allmählich.

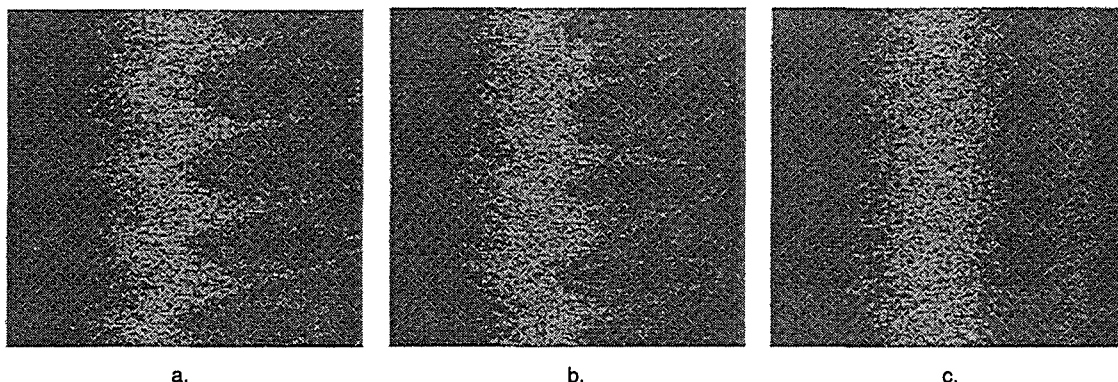


Abb. 4: Diokotron-Instabilität am Beispiel eines Elektronenstrahls

Die Simulation, das Experiment des Theoretikers

Modelle und Simulationen jeder Art sind Hilfsmittel zum Umgang mit der Realität. Sie sind so alt wie die Menschheit selber. Mit dem Aufkommen von Computern aber hat der Begriff Simulation eine neue bis dahin nicht bekannte Dimension bekommen, da es nun möglich war, jede mathematische und logische Formulierung in beliebiger Kombination schnell und genau abarbeiten zu können. Aber erst mit der Entwicklung von Hochleistungsrechnern hat diese Art der Simulation, die Computer-Simulation, ihren unbestrittenen Platz in der Wissenschaft eingenommen. Sie ist heute ein unverzichtbares Hilfsmittel für jeden Wissenschaftler. Die Darstellung des Verhaltens dynamischer Systeme durch Computer-Simulationsmodelle hat inzwischen alle anderen Darstellungsmöglichkeiten, die früher einmal eine Rolle spielten, wie hydraulische, elektrische und mechanische Analogien, abgelöst. Es gibt dafür natürlich gewichtige Gründe:

Man kann völlig unabhängig von der Art des betrachteten Systems mit einheitlichen Methoden und den darauf basierenden vielseitig verwendbaren Software-Programmen arbeiten.

Die Kosten der Modellerstellung und Simulation sind im allgemeinen nur ein Bruchteil dessen, was bei ähnlich umfassender Untersuchung mit realen oder analogen physikalischen Modellen aufzuwenden wäre.

Der zeitliche Ablauf des Experimentes kann erheblich gerafft und damit verkürzt (besonders wichtig bei langsamen Prozessen), aber auch beliebig gedehnt werden (der Simulationscode für die KALIF-Experimente muß eine Zeitspanne von ungefähr 100 ns auflösen), so daß genaue Beobachtungen möglich werden.

Eine Dynamik, die zur Systemzerstörung führen würde (z.B. Simulation von Reaktorunfällen, Crash-Test bei Fahrzeugen, u.v.m.) hinterläßt im Computer keinerlei negativen Spuren: Das Simulationsprogramm ist weiterhin funktionsfähig. Damit können auch gefährliche Systementwicklungen genau und umfangreich untersucht werden.

Das reale System wird keinerlei Risiko unterzogen. Messungen oder Eingriffe am realen System sind gar nicht, oder nur in beschränktem Umfange nötig.

Es ist aber nicht alleine die Erhöhung der Effektivität aufgrund einer beliebig erzielbaren Rechengenauigkeit bei gleichzeitiger Reduktion der Kosten und Vermeidung jeglichen Risikos realer Experimente, welche die Methode der Simulation den oben genannten Darstellungsmöglichkeiten überlegen macht, entscheidend ist auch, daß aus Computersimulationen auf elegante Weise neue Erkenntnisse gewonnen werden können, die aus der ursprünglichen Systemkenntnis nicht direkt folgen. Dies geschieht gewöhnlich durch geeignete Parameterstudien, mit denen man die verschiedenen Aspekte der Systemdynamik gezielt betrachten kann.

Der Leser könnte nun dem Gedanken verfallen, der Naturwissenschaftler könne gänzlich auf alle realen Experimente verzichten, indem er geeignete Simulationsprogramme entwickelt. Dem ist allerdings nicht so:

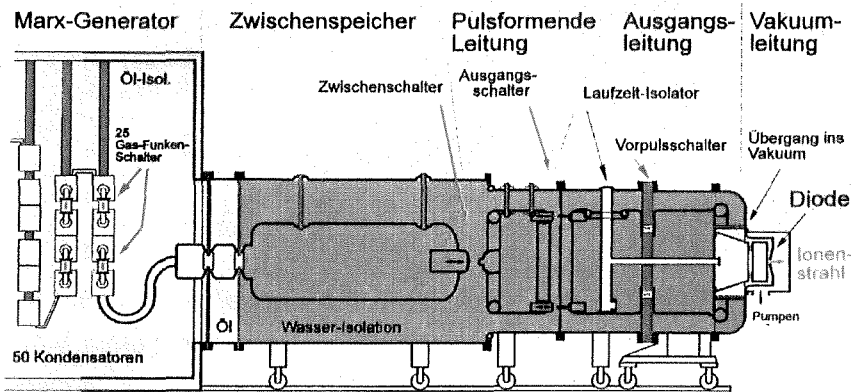
Das Kernstück jeder Simulation ist die Modellbildung. Modellbildung bedeutet aber immer Vereinfachung, Zusammenfassung, Weglassen unwichtiger Teilaspekte, d.h. Abstraktion und ist damit ohne Auswahl und Entscheidungsprozesse prinzipiell nicht möglich. Dieser Selektionsvorgang kann in der Regel nicht vollständig formalisiert werden. Die Intuition des Wissenschaftlers spielt dabei eine wesentliche Rolle. Dies bedeutet: Die in der Modellbildung getroffenen Vereinfachungen müssen stets einer umfassenden Gültigkeitsprüfung unterzogen werden. Es sind aber gerade die mit realen Experimenten gewonnenen Meßdaten, anhand derer der Wahrheitsgehalt einer Simulation nachgewiesen werden kann und muß. So erweist sich letztlich die Computersimulation mit all ihren oben aufgeführten Vorteilen als eine ideale und notwendige Ergänzung zum realen Experiment.

Die Simulation der Hochstrom-Diode

Reduktion des Simulationsgebietes

Versuchen wir nun, die in Kapitel 3 mehr allgemein erläuterte Simulationstechnik auf die Hochleistungs-Ionendiode zu übertragen. Zunächst brauchen wir ein geometrisches Modell für das Simulationsgebiet. Dieses soll so einfach wie möglich sein. Betrachten wir uns dazu die Anlage im Querschnitt (s. Abb. 5). Der elektromagnetische Impuls, im Marx-Generator erzeugt, über einen koaxialen Pulsleiter zur Ionendiode transportiert, wird in die Ionendiode radial eingespeist. Der größte Anteil seiner Energie wird dazu verwendet, Ionen im Diodenbereich freizusetzen und zu beschleunigen. Ein kleiner aber dennoch nicht zu vernachlässigender Teil des elektromagnetischen Impulses wird in der Diode reflektiert und gelangt somit wieder in den Bereich des Wellenleiters, wo er sich

mit dem einlaufenden Impuls überlagert. Jede Simulation muß diesem wichtigen Teilaspekt Rechnung tragen. Dies bedeutet: Das Simulationsgebiet sollte nicht nur den Diodenbereich, sondern zusätzlich große Teile des Wellenleiters umfassen. Der Wellenleiter ist aber im Vergleich zur Diode um ein Vielfaches größer. Ein Simulationscode, der beide Bereiche berücksichtigt, würde die Kapazität an Speicherplatz und Rechenleistung eines jeden Hochleistungsrechners sprengen.



KALIF-1 ist eine koaxiale Pulsleitung; 60 ns FWHM,
mit $V = 1.7 \text{ MV}$, $P = 1.5 \text{ TW}$, $Z = 2.25 \Omega$

Abb. 5: Querschnitt der KALIF-Anlage

Es ist daher notwendig das Simulationsgebiet einzuschränken. Nun ist es in der Tat möglich, die Simulationsrechnungen auf das Gebiet der Ionendiode selbst zu reduzieren. Hierzu braucht man aufgrund der erwähnten Reflexionen eine adäquate Beschreibung der Kopplung von Diode und Wellenleiter. Auf diese Details kann im Rahmen des Vortrages allerdings nicht näher eingegangen werden. Wenden wir uns nun der Diode zu. In Abb. 6a ist die Ionendiode im Querschnitt dargestellt. Sie ist relativ komplex und besteht aus vielen Details, von den die meisten für die Simulationsrechnung unwesentlich sind. Was wir brauchen ist ein vereinfachtes Modell der Ionendiode. Dabei muß die Geometrie des Diodeninnenbereiches sowie dessen Berandung, aber auch die Größe und Form von Kathode und Anode berücksichtigt werden. Die Rotationssymmetrie der Diode macht die Verwendung eines 2-dimensionalen Modells möglich. Das so gewonnene 2-dimensionale Gebiet ist in Abb. 6b dargestellt.

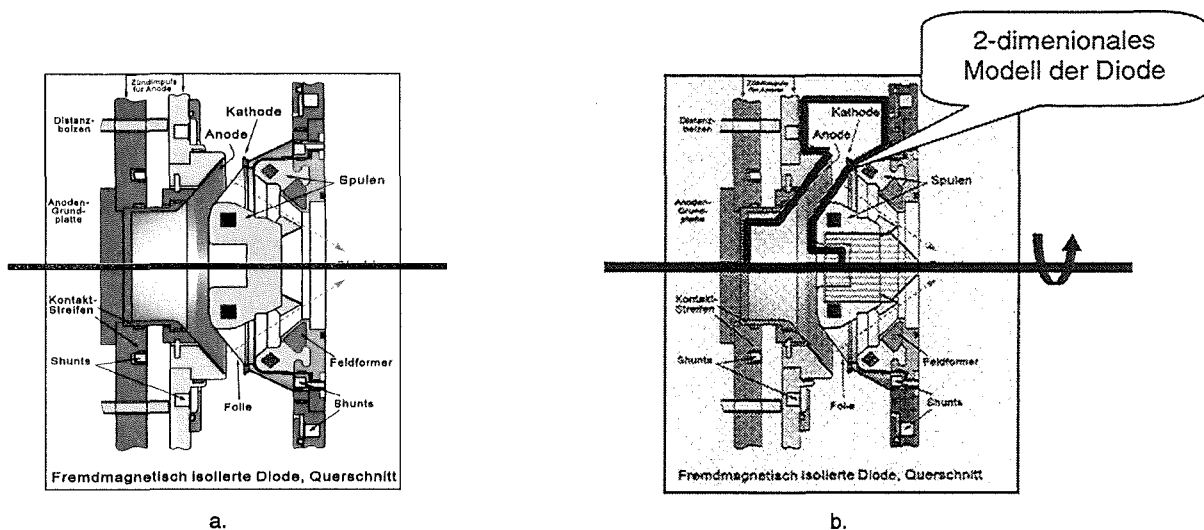


Abb. 6: Querschnitt der Ionen-Quelle

Man muß sich jedoch darüber im Klaren sein, daß durch die Verwendung eines rein 2-dimensionalen Modells, selbst wenn eine Rotationssymmetrie zugrunde gelegt wird, die Beschreibung von 3-dimensionalen Phänomenen, wie z.B. die Diokotron-Instabilität, nicht möglich ist. Es gibt jedoch auch bei der 2-dimensionalen Betrachtung bereits sehr komplexe Vorgänge, deren numerische Simulation für die Experimente von Bedeutung ist. So muß der 2-dimensionale Simulationscode KADI2D trotz seiner Bedeutung für Experimente erst als Zwischenschritt auf einen 3-dimensionalen zeitabhängigen Simulationscode für Hochleistungs-Ionendioden hin gesehen werden. An der Entwicklung eines 3-dimensionalen Diodencodes wird zur Zeit gearbeitet. Im weiteren Verlauf dieses Vortrages werden wir uns auf zwei Dimensionen beschränken.

Das physikalische Modell

Die elektromagnetischen Vorgänge innerhalb der Ionendiode werden einerseits durch die Maxwell-Gleichungen, andererseits durch die Lorentz-Gleichungen, welche die Bewegung, d.h. die räumliche Lage und die Geschwindigkeit der Elektronen und Ionen beschreiben, bestimmt. Die Maxwell-Gleichungen (s. M1 – M4) sind ein lineares System partieller Differentialgleichungen für das elektromagnetische Feld \vec{E}, \vec{B} . Die Lorentz-Gleichungen sind ein System von gewöhnlichen Differentialgleichungen erster Ordnung. Diese beiden grundlegenden Differentialgleichungssysteme können jedoch nicht unabhängig voneinander gelöst werden. Denn durch die Bewegung der Teilchen (Elektronen und Ionen) ändert sich die Ladungs- und Stromdichte in jedem Raumpunkt des Diodenbereiches. Ladungs- und Stromdichte gehen aber als Quellen in die Maxwell-Gleichungen ein (rechte Seite der Maxwell-Gleichungen) und beeinflussen so die zeitliche Entwicklung des elektromagnetischen Feldes und damit letztendlich wieder die Dynamik der Teilchen. Die Beeinflussung der Teilchen durch ihre Eigenfelder ist als Selbstkonsistenz bekannt. Die Selbstkonsistenz führt also dazu, daß das Gesamtsystem der Felder und Teilchen hochgradig nichtlinear ist.

Maxwell-Gleichungen

$$\text{M1. } \partial_t \vec{E} - c^2 \nabla \times \vec{B} = -\frac{1}{\epsilon_0} \vec{j}$$

$$\text{M2. } \partial_t \vec{B} + \nabla \times \vec{E} = \vec{0}$$

$$\text{M3. } \nabla \cdot \vec{E} = \frac{1}{\epsilon_0} \rho$$

$$\text{M4. } \nabla \cdot \vec{B} = \vec{0}$$

ρ, \vec{j} sind die Ladungs- und Stromdichte

Lorentz-Gleichungen

$$\text{L1. } \frac{d\vec{x}_k(t)}{dt} = \vec{v}_k(t)$$

$$\text{L2. } \frac{d\vec{p}_k(t)}{dt} = \vec{F}_L(t)$$

$\vec{F}_L = q_k (\vec{E} + \vec{v}_k \times \vec{B})$ ist die Lorentzkraft, die auf jedes geladene Teilchen wirkt.

Numerische Approximation

Die im vorhergehenden Abschnitt dargelegten physikalischen Gleichungen müssen nun durch adäquate numerische Approximationen ersetzt werden. Hierfür werden ein geeignetes Gittermodell, genaue Diskretisierungsschemata für die beiden Differentialgleichungssysteme sowie einen Gesamtlösungsansatz, der die Selbstkonsistenz der Wechselwirkung von geladenen Teilchen und Feldern widerspiegelt, benötigt. Betrachten wir nun diese drei Punkte im einzelnen.

a. Die Diskretisierungs-Methode

Eine weitverbreitete Methode zur Lösung eines hyperbolischen Differentialgleichungssystems, wie es die Maxwell-Gleichungen darstellen, bilden die Finiten-Differenzen-Verfahren. Dabei werden die in den Differentialgleichungen auftretenden partiellen Ableitungen durch Differenzenquotienten ersetzt. Bei der Wahl der Differenzenquotienten hat man gewisse Freiheitsgrade, die zu expliziten oder impliziten Verfahren führen. Zur Lösung der Maxwell-Gleichungen hat sich ein explizites Verfahren mit zentralen Differenzen, welches auch als Leapfrog-Verfahren bekannt ist, bewährt. Es wird auf einem versetzten Gitter verwendet, was bedeutet, daß die Komponenten der Felder \vec{E} und \vec{B} an verschiedenen Punkten des Gitters berechnet werden. Das Leapfrog-Verfahren ist ein Verfahren zweiter Ordnung. Es wurde erstmals von Yee vorgeschlagen und später von Holland für ko- und kontravariante Feldkomponenten erweitert [3, 4, 5, 6]. Das Leapfrog-Verfahren ist auf verhältnismäßig einfache Weise algorithmisch zu realisieren und liefert zufriedenstellende Ergebnisse. Ein entscheidender Nachteil besteht jedoch in der Verwendung von versetzten Gittern. Dies führt bei der Lösung der Lorentz-Gleichungen zu zusätzlichen Interpolationen, was i.a.

zu einer Erhöhung des Verfahrensfehlers führt. Außerdem können Gitterverfeinerungstechniken nur schwer in den Algorithmus integriert werden.

Diese Probleme können durch die Verwendung von Finite-Volumen-Verfahren auf nicht versetzten Gittern vermieden werden. Hierzu muß man zur integralen Darstellung der Maxwell-Gleichungen übergehen. Die so gewonnenen Integralgleichungen werden dann ihrerseits wieder diskretisiert. Für eine detaillierte Betrachtung der Finiten-Volumen-Verfahren zur Lösung der Maxwell-Gleichungen wird auf die Literatur verwiesen. Es sei an dieser Stelle nur erwähnt, daß im Simulationscode KADI2D Verfahren vom Typ Godunov verwendet werden [7, 8].

B. Das Gittermodell

Um technisch relevante Geometrien, wie z.B. die der Ionendiode, numerisch gut

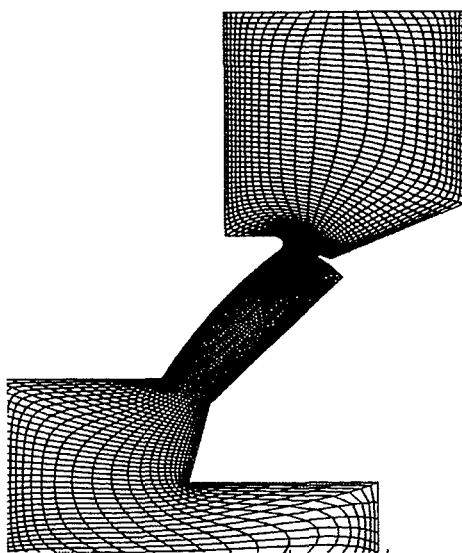


Abb. 7: randangepaßtes Gitter für den Innenbereich der Ionodiode

handhaben zu können, hat sich die Verwendung von randangepaßten Gittern bewährt. Der wesentliche Vorteil gegenüber den nicht randangepaßten Gittern besteht darin, daß der Rand des Simulationsgebietes mit Gitterlinien zusammenfällt. Dadurch wird die Formulierung der Randbedingungen wesentlich vereinfacht. Allerdings müssen auch Nachteile in Kauf genommen werden. So ist das dem randangepaßten Gitter zugeordnete Koordinatensystem i.a. nicht orthogonal. Die zu lösenden Differentialgleichungen, in der Regel in kartesischen Koordinaten formuliert nehmen durch die Transformation auf die krummlinigen nichtorthogonalen Koordinaten eine im Vergleich zu den Ausgangsgleichungen etwas kompliziertere Form an. Dies bereitet jedoch für die numerische Behandlung

keine grundsätzlichen Schwierigkeiten.

Bei der zur Zeit am weitesten verbreiteten numerische Methode zur Erzeugung von randangepaßten Gittern wird die Laplace- bzw. die Poisson-Gleichung gelöst [9]. Dies führt bei Kanten mit kleinen Winkeln zu entarteten Gitterzellen (entartete Zellen sind z.B. Zellen, die im Vergleich zu anderen Zellen sehr kleine Flächen haben). Numerische Algorithmen liefern im Bereich von entarteten Zellen schlechte oder sogar vollkommen falsche Ergebnisse. Die Qualität eines numerischen Algorithmus hängt daherentscheidend von der Qualität des verwendeten Gitters ab. Das Erkennen und Beseitigen von Gitterentartungen geschieht mangels geeigneter automatisierbarer Techniken üblicherweise manuell, was jedoch sehr zeitaufwendig ist.

Neben den oben genannten Verfahren zur Gittererzeugung gibt es jedoch noch andere Verfahren, die für extreme Geometrien geeigneter sind . So wurde bereits gezeigt, daß

durch die Verwendung von biharmonischen Gleichungen zur Gittererzeugung selbst bei komplexen Geometrien bessere Gitter erzeugt werden können.

c. Die Particle-In-Cell-Methode

Wie bereits erwähnt, muß das Maxwell-Lorentz-System selbstkonsistent gelöst werden. Das heißt, der durch die Eigenfelder der Elektronen und Ionen bewirkte Rückkopplungseffekt muß in der numerischen Simulation berücksichtigt werden. Ein geeignetes Verfahren hierfür stellt die Particle-In-Cell-Methode (PIC-Methode) dar. Sie wird im folgenden für einen Zeitschritt erläutert:

Zunächst werden die partiellen Differentialgleichungen (Maxwell-Gleichungen) für die elektromagnetischen Felder \vec{E} und \vec{B} auf einem randangepaßten Gitter gelöst. Dabei werden die für diesen Zeitschritt aktuellen Werte für Ladungsdichte und Stromdichte (ρ, \vec{j}) verwendet. Anschließend werden die berechneten Felder auf die Orte der Elektronen und Ionen interpoliert (die Teilchen befinden sich i.a. nicht an den Gitterpunkten, sondern irgendwo in den Gitterzellen). Nun werden für jedes geladene Teilchen die Lorentz-Gleichungen gelöst. Aus den damit erhaltenen neuen Orten und Geschwindigkeiten der Teilchen lassen sich für jeden Gitterpunkt die geänderten Ladungs- und Stromdichten ermitteln, so daß im nächsten Zeitschritt die daraus resultierenden geänderten elektromagnetischen Felder \vec{E} und \vec{B} berechnet werden können.

Parallelisierung des KADI2D-Codes

Warum High-Performance-Computing

Um zu verlässlichen numerischen Ergebnissen zu kommen, die dem Vergleich mit den experimentellen Meßergebnissen standhalten muß einerseits mit einem sehr feinen Gitter (ca. 10^5 Gitterpunkte), andererseits mit ca. 10^6 Teilchen gerechnet werden. Außerdem werden, um eine genügend hohe zeitliche Auflösung zu erhalten, $10^3 - 10^4$ Iterationszyklen benötigt. Dies führt zu Rechenzeiten, die für 2-dimensionale Simulationsrechnungen deutlich über 10 CPU-Stunden liegen. Die Rechenzeiten für die angestrebte 3-dimensionale Simulation werden sogar auf mehr als 50 CPU-Stunden geschätzt. Durch Rechner spezifische Optimierung, wie z.B. durch die Anpassung des Algorithmus an die spezielle Cache-Größe der verwendeten Hardware u.ä., lassen sich die Rechenzeiten nur in geringem Maße reduzieren. Hingegen kann durch paralleles Abarbeiten geeigneter Teile der Simulationsrechnung auf verschiedenen Prozessoren die Verweilzeit (Elapsed-Time) des Programms deutlich verringert werden. Außerdem wird der pro Prozessor benötigte Speicherplatz ebenfalls reduziert. Daher können mit einem effizient parallelisierten Programm wesentlich größere Probleme mit mehr Gitterpunkten und einer größeren Anzahl von Elektronen und Ionen gerechnet werden,

wodurch genauere Vorhersagen möglich sind. Wir werden nun im folgenden zwei verschiedene Parallelisierungs-Strategien vorstellen.

Die Parallelisierungs-Strategien

Bevor man mit der Parallelisierung eines Programmsystems beginnt, müssen zunächst solche Programmteile lokalisiert werden, bei denen sich eine Parallelisierung überhaupt lohnt. Die wichtigsten Kriterien hierbei sind der Verbrauch von CPU-Zeit und der Speicherplatzbedarf.

Das heißt: Mögliche Kandidaten für eine erfolgreiche Parallelisierung sind jene Programmteile, die den größten Anteil der CPU-Zeit bzw. des Speicherplatzes benötigen. Betrachten wir die schematische Darstellung des Programmsystems KADI2D (Abb. 8):

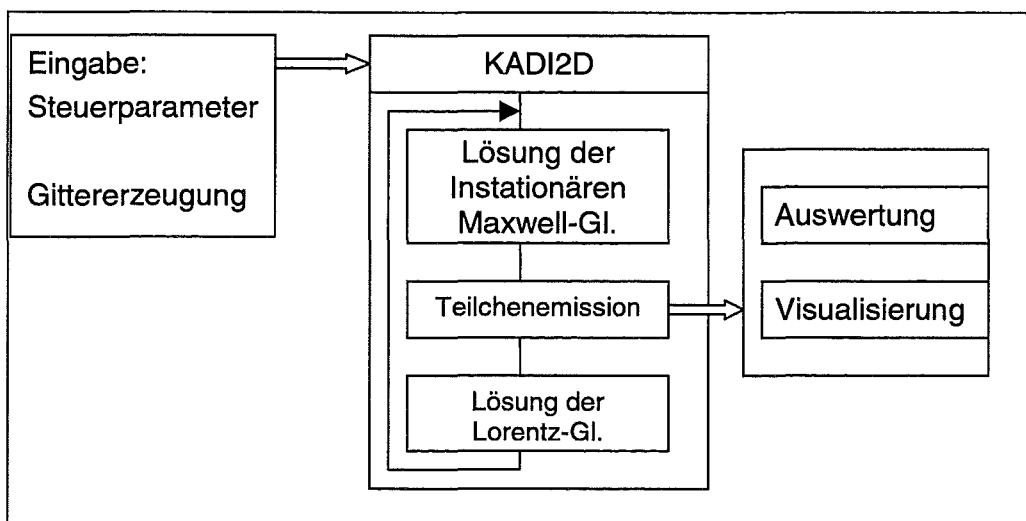


Abb. 8: Schematische Darstellung des Gesamtsystems zur Simulation von Hochleistungs-Ionendioden an der KALIF-Anlage

Eine CPU-Zeit Analyse für das Programmsystem KADI2D zeigt, daß die meiste Rechenzeit zwar im Lorentz-Löser verbraucht wird, der Rechenaufwand des Maxwell-Lösers aber nicht vernachlässigt werden kann. Für 3-dimensionale Simulationsrechnungen dürfte der Maxwell-Löser sogar die dominierende Rolle spielen. In KADI2D soll zunächst die Teilchenbehandlung parallelisiert werden. Dabei machen wir von der speziellen Eigenschaft der Particle-In-Cell-Methode Gebrauch, daß in jedem Zeitschritt jedes Teilchen der Gesamtheit sich unabhängig von allen anderen Teilchen bewegt. Die Lorentz-Gleichungen können daher parallel auf verschiedenen Prozessoren gelöst werden. Dafür gibt es zwei grundsätzlich verschiedene Parallelisierungs-Strategien:

Jedem Prozessor wird eine feste Anzahl von Teilchen zugewiesen, unabhängig von ihrer Lage im Berechnungsgebiet. Um ein Load-Balancing zu gewährleisten, muß jeder Prozessor ungefähr die gleiche Anzahl von Teilchen bearbeiten. Es ist zwar grundsätzlich möglich, dies während der gesamten Simulationsrechnung zu gewährleisten, aber aufgrund der Tatsache, daß Teilchen aus dem Berechnungsgebiet herausfliegen können, ist das Load-Balancing nicht automatisch gegeben. Ein

entscheidender Nachteil dieser Methode ist der hohe Kommunikationsaufwand für den Austausch von Feld- und Teilcheninformationen.

Das Rechengebiet wird in n Teilbereiche unterteilt. Jeder der n Prozessoren muß genau eines dieser Teilgebiete bearbeiten unabhängig von der Anzahl der in den Teilgebieten vorhandenen Teilchen. Ohne spezielle Maßnahmen, wie z.B. lastabhängige Gebietsgrößen, ist ein Load-Balancing grundsätzlich nicht möglich. Ein wesentlicher Vorteil dieser Methode liegt aber in dem relativ geringen Kommunikationsaufwand.

Wie bereits erwähnt, ist der Rechenaufwand zur Lösung der Feldgleichungen nicht zu vernachlässigen. Die Parallelisierung des Maxwell-Lösers wäre daher wünschenswert. Der Maxwell-Löser kann aber sinnvoll nur über Gebietszerlegung parallelisiert werden. Um ein harmonisches Zusammenspiel von Lorentz- und Feld-Löser zu gewährleisten, haben wir uns dafür entschieden, die Parallelisierung der Teilchenbehandlung nach Methode 2 zu realisieren.

Skalierungseigenschaften von KADI2D auf der VPP300

Die in Abb. 9 zusammengefaßten Resultate wurden auf dem Vektor-Parallelrechner VPP300 gewonnen. Hierzu wurde die Teilchenbehandlung parallel auf bis zu 8 Prozessoren durchgeführt. Das elektrische Feld wurde als konstant angenommen. Aus der Abbildung geht klar hervor, daß die benötigte Rechenzeit (Elapsed-Time) zunächst deutlich abnimmt, dann aber aufgrund des steigenden Kommunikationsaufwandes wieder zunimmt. Ein Optimum liegt, bezogen auf das Testbeispiel, bei 4 - 5 Prozessoren. Durch das Parallelisieren des Maxwell-Lösers sind weitere Verbesserungen im Zeitverhalten des Programms zu erwarten.

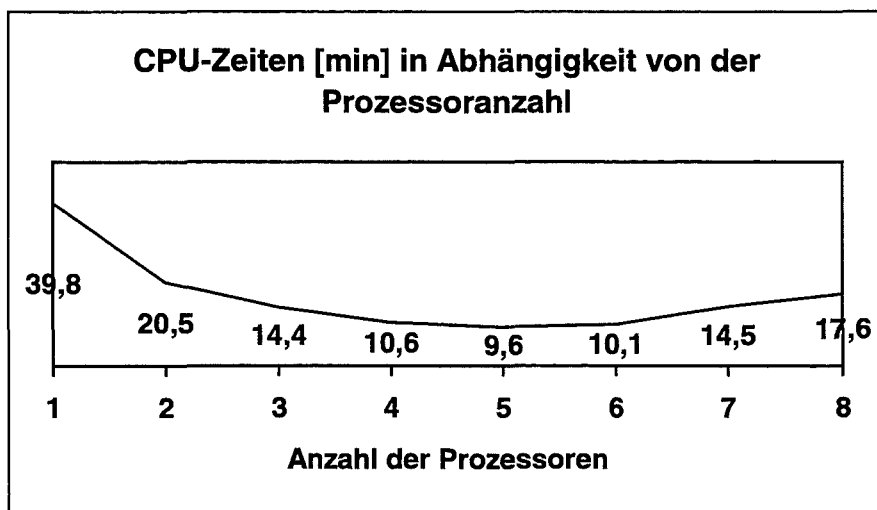


Abb. 9: An der VPP300 gemessene CPU-Zeiten in Abhängigkeit von der Prozessoranzahl

Literaturverzeichnis

- [1] Bauer, W., Bluhm, H.-j., Hoppe, Karow, H., Laqua, H.; KfK-Nachrichten, Jahrg. 24, 1 (1992), pp. 2-13
- [2] Rusch, D. Ratajczak, W.; KfK-Nachrichten, Jahrg. 24, 1 (1992), pp. 13-18
- [3] Holland, R.; 'Finite-Difference Solution of Maxwell's Equations in generalized nonorthogonal Coordinates', IEEE Trans. Nucl. Sci. 30, 4589(1983)
- [4] Yee, K.S.; 'Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media', IEEE Trans. Ant.& Prop. 14, 302(1966)
- [5] Boris, J.P.; 'Relativistic plasma simulation – optimization of a hybrid code', Proc. 4 th. Conf. On Num. Sim. Plasmas, NRL Washington D.C., 1970
- [6] Taflove, A.; 'Re-inventing electromagnetics: Supercomputing solution of Maxwell's equations via direct time integration on space grids', AIAA 92-0333, American Institute of Aeronautics and Astronautics, Washington, D.C., 1992
- [7] Godunov, S.K.; 'Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics', Mat. Sbornik, 47 (1957), pp. 271 –306
- [8] E. Halter, M. Krauß, C.-D. Munz, R. Schneider, E. Stein, U. Voß, T. Westermann; Forschungszentrum Karlsruhe – Umwelt und Technik, FZKA 5654, (1995)
- [9] Halter, E., Kündahl, J., Krauß, M., Munz, C.-D., Schneider, R., Stein, E., Voß, U., Westermann, T.; 'Interaktive Gittererzeugung im Rahmen eines offenen CAD-Systems, 9.DGLR/AG STAB-Symposium 1994, Erlangen, 1994

Simulation turbulenter Strömungen auf Hochleistungsrechnern

W. Rodi

Institut für Hydromechanik

Universität Karlsruhe

EINLEITUNG

Turbulente Strömungen spielen eine wichtige Rolle bei vielen Problemen der Luft- und Raumfahrttechnik, des Maschinenbaus, Chemieingenieur- sowie Bauingenieurwesens. Als Beispiele seien hier genannt: die Strömung um Flug- oder Schiffskörper, Autos, Züge und Gebäude sowie durch Leitungen, Kanäle, Kammern und Wärmetauscher, die Vermischungsprozesse in Reaktoren aller Art, die Wandkühlung von Turbinenschaufeln und Brennkammern, die Ausbreitung eingeleiteter Abwärme oder von Schadstoffen in der Atmosphäre, in Flüssen oder Seen. Die turbulenten Schwankungsbewegungen tragen wesentlich zum Transport von Impuls, Wärme und Stoff bei und haben deshalb einen bestimmenden Einfluß auf die Verteilung von Geschwindigkeit, Druck, Temperatur oder Konzentration und damit auf Strömungskräfte, Vermischung und Verdünnung von Schadstoffen sowie Wärme- und Stoffübergang. Um die Auswirkungen turbulenter Strömungen bei der Optimierung von Anlagen voraussagen zu können, besteht ein großer Bedarf an Verfahren zur Berechnung dieser Strömungen, wobei in jedem Fall auch der Einfluß der Turbulenz simuliert werden muß.

Die Turbulenzbewegungen sind äußerst komplex, da sie immer dreidimensional, instationär und sehr unregelmäßig sind. Die praktisch wichtigste Eigenschaft der Turbulenz ist die durch die turbulenten Schwankungen verursachte starke Vermischung. Turbulenz besteht aus Wirbelelementen, die in Wechselwirkung miteinander stehen. Eine wichtige Eigenschaft der Turbulenz ist, daß diese Elemente über ein weites Spektrum von Elementgrößen verteilt sind (siehe Bild 1), wobei die größten Elemente von der Größe des Strömungsgebietes sind und bei den kleinsten Elementen mit hochfrequenten Schwankungen die viskose Dissipation von kinetischer Energie des Fluids stattfindet.

Die Berechnung praxisrelevanter turbulenter Strömungen ist nur mit Hilfe numerischer Verfahren und leistungsstarker Rechner möglich, wobei es drei prinzipielle Methoden zur Simulation des Turbulenzeinflusses gibt: die direkte numerische Simulation (DNS), die Large-Eddy- Simulation (LES) und die statistische Methode. In diesem Beitrag werden alle drei Methoden zusammen mit Rechenbeispielen kurz vorgestellt und es werden die jeweiligen Anforderungen an die Rechnerkapazität diskutiert.

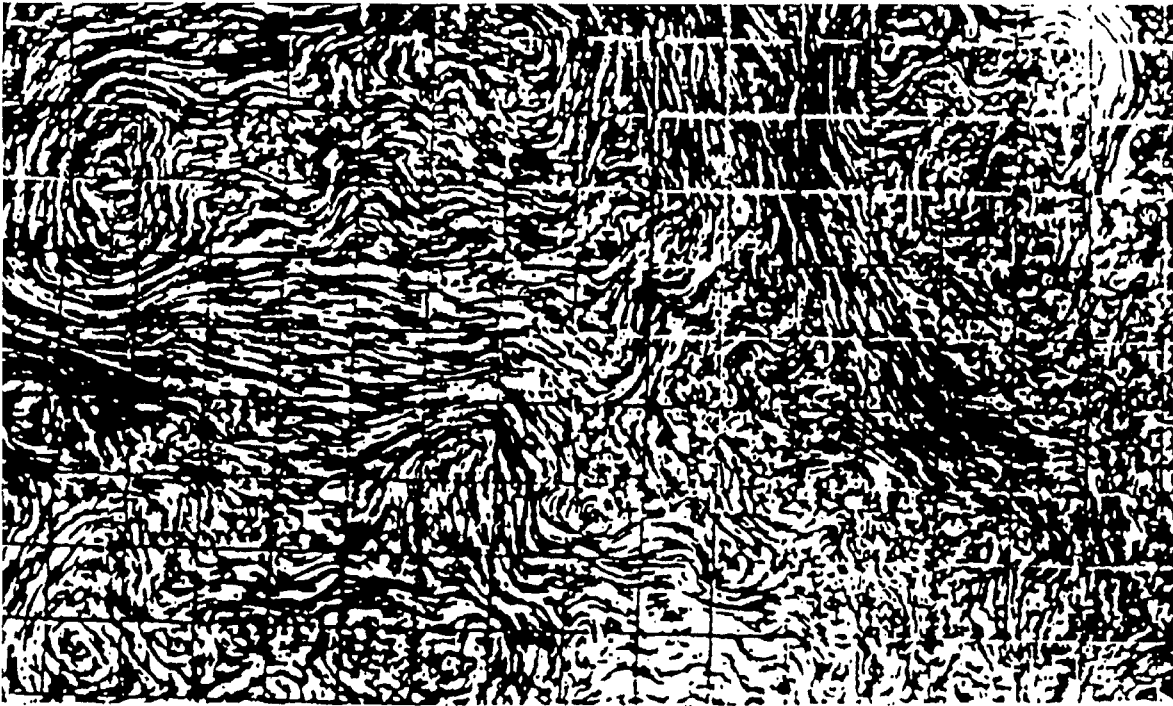


Bild 1: Turbulente Wirbel an der Oberfläche eines gerührten Behälters

2. DIREKTE NUMERISCHE SIMULATION (DNS)

Die exakten Gleichungen zur Beschreibung der unregelmäßigen Turbulenzbewegungen sind bekannt: die dreidimensionalen instationären Navier-Stokes Gleichungen zusammen mit der Kontinuitätsgleichung. Bei der direkten numerischen Simulation werden diese Gleichungen zur Bestimmung aller Details der Turbulenzbewegung gelöst. Das Problem dabei ist, daß mit einem gewählten Rechennetz nur die Turbulenzwirbel aufgelöst werden können, die größer als die Maschenweite sind (siehe Bild 1), und das Rechennetz nicht beliebig fein gewählt werden kann. Die Größe der kleinsten Elemente relativ zu denen der großen Elemente und damit des Rechengebiets, η/D , ist etwa umgekehrt proportional zur Reynoldszahl. Zur Auflösung der kleinsten Wirbel müssen also mindestens $N = D/\Delta x > D/\eta$ Gitterpunkte in jede Raumrichtung gelegt werden und damit N^3 Punkte insgesamt. Eine Abschätzung ist für ebene Kanalströmung in Tabelle 1 für verschiedene Reynoldszahlen $Re = DU/v$ gegeben, wobei U die mittlere Kanalgeschwindigkeit, D die Kanaltiefe und v die kinematische Viskosität ist. Für die niedrigste Reynoldszahl in Tabelle 1 wurden von Kim et al. [1] DNS Rechnungen mit 2 Millionen Gitterpunkten durchgeführt. Da die benötigte Gitterpunktzahl mit Re^3 ansteigt, werden sehr schnell extrem große Gitterpunktzahlen erforderlich, um auch die kleinen Wirbel auflösen zu können. Da die Zahl der Rechenoperationen ebenfalls ungefähr proportional der Zahl der Gitterpunkte ist, steigt auch die benötigte Rechenzeit ungefähr mit Re^3 an. Um aus den instationären Rechnungen verlässliche statistische Durchschnittswerte zu erhalten, sind sehr viele Zeitschritte notwendig und eine Abschätzung der benötigten Rechenzeit auf einer Maschine mit 150 MFlops (Millionen Gleitkommaoperationen pro Sekunde) ist in Tabelle 1 ebenfalls gegeben. Diese zeigt,

daß extrem große Rechenzeiten erforderlich wären und selbst auf wesentlich schnelleren Maschinen viele Stunden notwendig sind.

Re	6600	20000	100000	10^6
N^3	$2 \cdot 10^6$	$40 \cdot 10^6$	$3 \cdot 10^9$	$1.5 \cdot 10^{12}$
CPU	37h	740h	6.5a	3000a

Tabelle 1: Benötigte Gitterpunktzahl und Rechenzeit (auf 150 MFlops Maschine) für DNS-Rechnung einer Kanalströmung

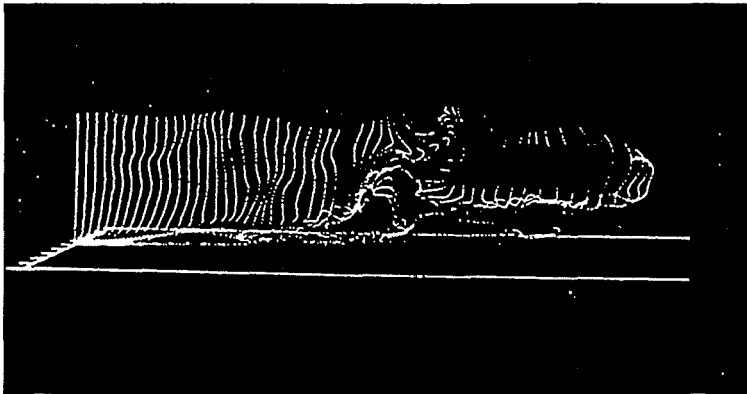
Dies bedeutet, daß DNS zur Zeit und auch in der näheren Zukunft nur für Strömungen mit relativ kleinen Reynoldszahlen möglich ist und deshalb nicht für die Lösung praktischer Probleme in Frage kommt. Bei niedrigen Reynoldszahlen erlaubt die Methode jedoch eine Untersuchung aller Details der turbulenten Strömung inklusive der Struktur und ist deshalb ein wichtiges Werkzeug für Grundlagenuntersuchungen und für das Entwickeln und Austesten der nachfolgend besprochenen statistischen Modelle geworden.

Bild 2 zeigt Ergebnisse einer DNS Rechnung für eine Kanalströmung, wobei durch Einbringen eines Tracers entlang einer vertikalen Linie die sich ausbildende Turbulenzstruktur sichtbar gemacht wurde. Die Rechnung ist gegenübergestellt einer entsprechenden Sichtbarmachungsuntersuchung im Labor, wobei Wasserstoffbläschen eingebracht wurden. Die Gegenüberstellung zeigt, daß die Rechnung in der Tat realistische Turbulenzstrukturen liefert. Inzwischen konnten Kanalströmungen bis zu einer Reynoldszahl von 15.000 berechnet werden. Weiter konnte auch die Strömung über eine rückspringende Stufe realistisch simuliert werden [4] (10^7 Gitterpunkte, mehr als 1000 CPU Stunden auf einer Cray YMP) und es konnte auch der laminar-turbulente Umschlag in der Kanalströmung nachvollzogen werden [5].

3. LARGE-EDDY SIMULATION (LES)

Wegen der Auflösungsprobleme bei großen Reynoldszahlen wurde die Large-Eddy Simulationstechnik entwickelt, bei welcher nur die Wirbel mit Skalen größer als die Maschenweite aufgelöst werden. Der Einfluß der kleinskaligen Bewegung muß dabei durch ein Modell berücksichtigt werden. Dieser Einfluß ist hauptsächlich dissipativ, d.h. dem auflösbaren Teil des Turbulenzspektrums wird Energie entzogen; der Einfluß kann auf zweierlei Weise simuliert werden: Die übliche und physikalisch korrektere Methode ist die Verwendung eines Feinstrukturmodells zur Bestimmung von turbulenten Spannungen, welche durch die nicht aufgelösten Feinstrukturschwankungen eingeführt werden; die andere Möglichkeit besteht im Entzug von kinetischer Energie durch das dissipative Verhalten des numerischen Verfahrens.

DNS Rechnung von Kim (aus [2])



Experiment [3]

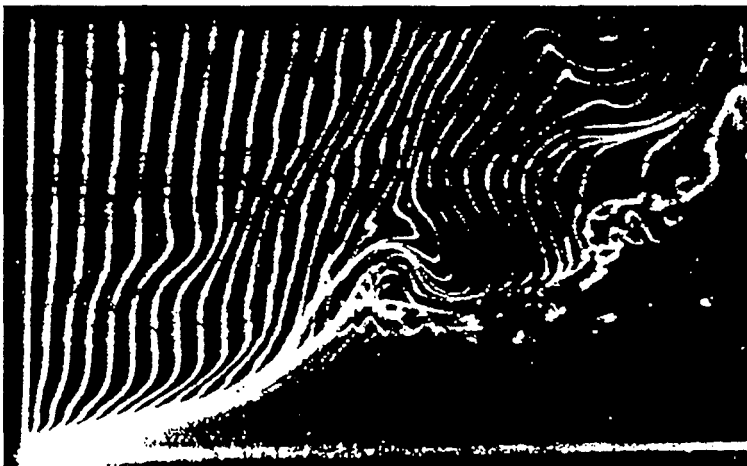


Bild 2: Strömungssichtbarmachung durch Tracer in einer ebenen Kanalströmung

Bei diesem von Kuwahara und seiner Schule [6] verfolgten Weg werden die Navier-Stokes- und Kontinuitätsgleichungen wie bei DNS auf einem möglichst feinen Gitter gelöst und die notwendige Dämpfung wird durch Verwendung eines Aufwindverfahrens 3. Ordnung für die konvektiven Glieder eingeführt.

Bei Verwendung eines Feinstrukturmodells wird der Einfluß der nicht auflösbaren Turbulenzbewegungen explizit simuliert; es wird dabei für diesen Teil des Turbulenzspektrums ein Modell eingeführt - im Gegensatz dazu wird bei den nachfolgend besprochenen statistischen Modellen für das gesamte Turbulenzspektrum ein Modell eingeführt. Oft wird das auch dort verwendete Wirbelviskositätsprinzip eingesetzt und die Wirbelviskosität zu dem Gradienten der aufgelösten Geschwindigkeiten und der Maschenweite in Beziehung gesetzt (Smagorinsky Modell). Dabei wird eine empirische Konstante eingeführt, deren Wert leider von Strömung zu Strömung verschieden ist. Um diese unerwünschte Eigenschaft zu eliminieren, haben Germano et al. [7] das sogenannte dynamische Feinstrukturmodell entwickelt, bei dem praktisch aus den kleinsten, gerade noch aufgelösten Turbulenzbewegungen die zeitliche und räumliche Variation der Smagorinsky Konstante bestimmt wird.

Bei hohen Reynoldszahlen kann das wandnahe, viskositätsbeeinflusste Gebiet nicht ausreichend aufgelöst werden und muß durch sogenannte Wandfunktionen überbrückt werden. Diese nehmen eine Beziehung zwischen der Geschwindigkeit am ersten Gitterpunkt und der Wandschubspannung an oder direkt eine Verteilung der momentanen Geschwindigkeit in dem nicht aufgelösten Bereich [8, 9]. Die Vorgabe von Randbedingungen am Ein- und Ausströmrand ist bei LES Rechnungen problematisch. Ein realistisches momentanes Turbulenzfeld muß am Einströmrand vorgegeben werden, es sei denn, dieser Rand kann weit genug vom hauptsächlich interessierenden Gebiet gelegt werden oder die Ränder sind periodisch.

Im allgemeinen ist die für LES Rechnungen benötigte Computerkapazität ebenfalls sehr groß; dabei ist nicht so sehr die Anforderung an den Speicher ein Problem, sondern die Rechenzeit, da zur Erzielung verlässlicher statistischer Werte eine sehr große Zahl von Zeitschritten gerechnet werden muß, oft mehr als 100.000. Dies führt auch auf Supercomputern zu großen Rechenzeiten. In Bild 3 ist eine LES Rechnung der Strömung um einen Würfel dargestellt, der in einem Kanal plaziert wurde. Die Strömung über und um den Würfel ist stark dreidimensional und es treten mehrere Ablösegebiete auf. Gezeigt werden im Vergleich mit dem Experiment berechnete Stromlinien in der vertikalen Symmetrieebene sowie Stromlinien in Bodennähe. Die Übereinstimmung ist für diese recht komplexe Strömung insgesamt sehr zufriedenstellend und wesentlich besser als mit der ebenfalls gezeigten Rechnung mit einem nachfolgend beschriebenen, statistischen $k - \epsilon$ Turbulenzmodell. Der Preis hierfür ist allerdings recht hoch: die auf einem Vektorrechner Siemens SNI 600/20 benötigte Rechenzeit betrug 160 Stunden. Andererseits benötigt die LES Methode wesentlich weniger Rechenzeit als die DNS Methode: Akselvol und Moin [12] haben mit 700.000 Gitterpunkten in 22 CPU Stunden auf einer Cray YMP dieselbe Stufenströmung mit guter Genauigkeit berechnet, für die Moin et al. [4] bei ihrer DNS Rechnung 10^7 Gitterpunkte und mehr als 1000 CPU Stunden brauchten.

Die LES Methode hat für praktische Rechnungen zur Zeit noch nicht die Einsatzreife erreicht, ist aber sicherlich die Methode der Zukunft. Sie ist vor allem dann geeignet und den nachfolgend besprochenen statistischen Modellen überlegen, wenn große Strukturen den turbulenten Austausch wesentlich bestimmen und die Strömung auch statistisch gesehen instationär ist oder bimodales Verhalten aufzeigt sowie wenn Methoden zur Manipulation der Turbulenz untersucht werden sollen.

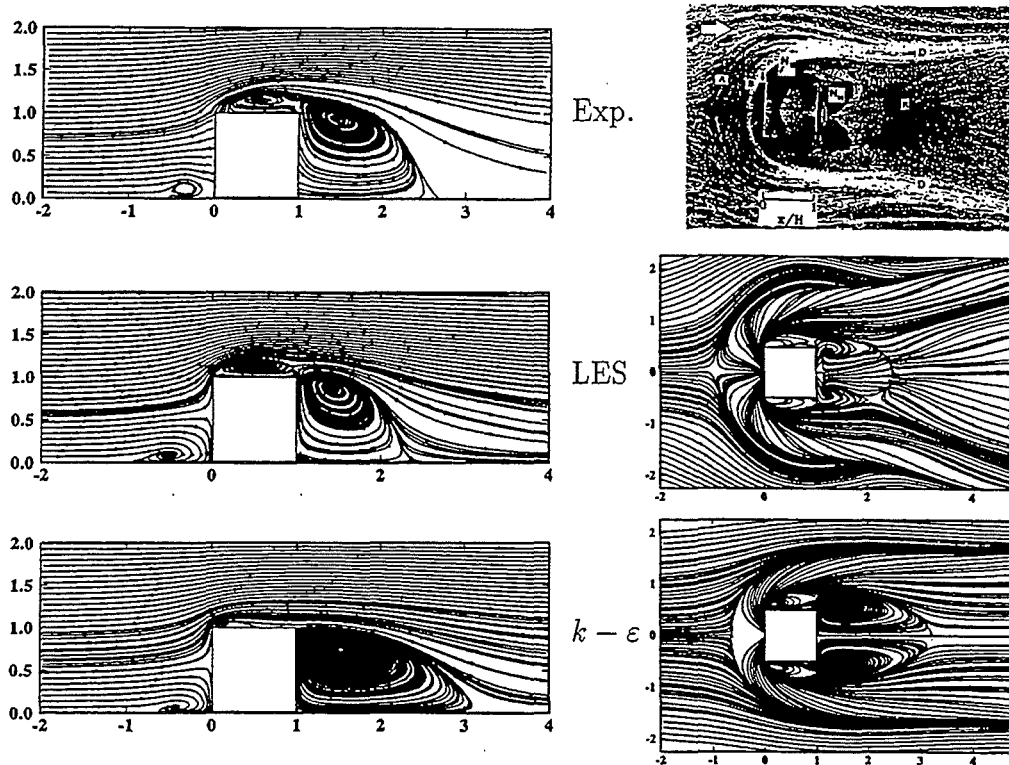


Bild 3: Strömung um einen wandgebundenen Würfel in einem Kanal, LES Rechnung von Breuer und Rodi [10], $k - \varepsilon$ Modellrechnung von Lakehal und Rodi [11], Experiment von Martinuzzi und Tropea [14]

4 STATISTISCHE METHODEN

Die heute in der Praxis eingesetzten Methoden sind statistischer Natur; dabei werden alle stochastischen turbulenten Schwankungen herausgemittelt und nur Gleichungen für die gemittelten Strömungsgrößen gelöst. Bei statistisch zeitabhängigen Strömungen handelt es sich um eine Ensemblemittelung, bei stationären Strömungen muß über einen im Vergleich zum Zeitmaßstab der Turbulenzschwankungen langen Zeitraum gemittelt werden. Die momentanen Strömungsgrößen f werden dann in einen Mittelwert \bar{f} und einen Schwankungswert f' aufgespalten. Die Einführung dieser Aufspaltung in die ursprünglichen zeitabhängigen Navier-Stokes- und Kontinuitäts- gleichungen führt zur gemittelten Form dieser Gleichungen. Die Mittelung führt dabei Korrelationen zwischen den Schwankungsgeschwindigkeit $\overline{u'_i u'_j}$ ein, die wie Spannungen auf das Fluid wirken und deshalb turbulente Spannungen oder Reynolds-Spannungen genannt werden. Diese Spannungen müssen mit einem statistischen Turbulenzmodell bestimmt werden. Die Turbulenzmodellgleichungen schließen das Gleichungssystem und erlauben damit die Berechnung der statistischen Größen. Im Unterschied zu den Feinstrukturmodellen bei LES Rechnungen simulieren die statistischen Turbulenzmodelle das gesamte Spektrum der Turbulenz.

Bei der Turbulenzmodellierung gibt es hauptsächlich drei verschiedene Methoden: zwei von ihnen basieren auf dem Wirbelviskositätsprinzip, bei welchem in Analogie zu laminaren Strömungen die turbulenten Spannungen $\overline{u_i' u_j'}$ proportional den Gradienten der mittleren Geschwindigkeiten angenommen werden. Dabei wird eine künstliche Wirbelviskosität ν_t eingeführt. Diese ist keine Stoffgröße sondern hängt von der Struktur der Turbulenz ab, und das Turbulenzmodell muß auch die Verteilung der Wirbelviskosität über das Strömungsgebiet bestimmen. Ein Kernansatz bei der Turbulenzmodellierung ist die Charakterisierung des lokalen Zustands der Turbulenz durch wenige Parameter. Aus Dimensionsgründen ist das Minimum dabei zwei Parameter, nämlich ein Geschwindigkeitsmaßstab V und ein Längenmaßstab L . Ebenfalls aus Dimensionsgründen gilt $\nu_t \propto V \cdot L$, so daß die Wirbelviskosität bestimmt werden kann, wenn man das Geschwindigkeits- und das Längenmaß der Turbulenz abschätzen kann.

Die einfachsten Turbulenzmodelle, die auf Prandtl zurückgehen, beziehen den Geschwindigkeits- und Längenmaßstab der Turbulenz zu Geschwindigkeitsgradienten oder zu anderen charakteristischen Geschwindigkeiten und Längen von Scherschichten mit Hilfe rein algebraischer Beziehungen. Diese Modelle des Mischungswegtyps werden hauptsächlich für anliegende Grenzschichten oder Grenzschichten mit kleinen Ablösezone eingesetzt und eignen sich nicht bei Strömungen mit massiven Ablösegebieten. Sie basieren implizit auf der Annahme lokalen Gleichgewichts der Turbulenz, d.h. sie vernachlässigen den Transport von Turbulenz von einer Stelle zur anderen sowie auch Gedächtniseffekte.

Zur Berücksichtigung der letztgenannten Effekte wurden sogenannte Zweigleichungsmodelle eingeführt, die ebenfalls die Wirbelviskositätshypothese verwenden, aber Geschwindigkeits- und Längenmaßstab der Turbulenz durch Modelltransport-Differentialgleichungen bestimmen. Am häufigsten verwendet wird das k - ϵ Modell, bei dem Gleichungen für die kinetische Energie der Turbulenz k und für die Dissipationsrate ϵ gelöst werden; wegen $\epsilon = k^3/2L$ bestimmt die Lösung der ϵ -Gleichung auch die Längenmaßverteilung. Dieses Modell ist heute praktisch in allen kommerziellen CFD Codes implementiert. Es ist auch auf Strömungen mit massiver Ablösung anwendbar; doch können dabei keine zu großen Genauigkeitsansprüche an die Ergebnisse gestellt werden.

Für komplexe Strömungen und beim Auftreten spezieller Einflüsse wie Auftrieb, Krümmung, Rotation etc. besser geeignet sind die sogenannten Reynolds-Spannungsmodelle. Diese verwenden das Wirbelviskositätsprinzip nicht sondern lösen Modelltransportgleichungen für die verschiedenen Reynoldsspannungen $\overline{u_i' u_j'}$. Auf Grund der größeren Anzahl von Turbulenzmodell-Differentialgleichungen, die miteinander gekoppelt sind, sind diese Modelle numerisch wesentlich aufwendiger und wurden daher auch wesentlich weniger getestet. Wegen der in letzter Zeit gestiegenen Rechnerleistung und der Verbesserung numerischer Methoden haben sie jedoch an Popularität gewonnen und sind auch schon in einigen kommerziellen CFD Codes

implementiert. Ihre Überlegenheit gegenüber dem $k-\varepsilon$ Modell ist allerdings nicht in jedem Fall garantiert.

Auch bei Verwendung von statistischen Turbulenzmodellen, speziell bei $k-\varepsilon$ und Reynoldsspannungsmodellen, ist die numerische Auflösung wandnaher viskositätsbeeinflusster Gebiete bei hohen Reynoldszahlen schwierig bzw. sehr aufwendig. Deshalb werden auch hier oft Wandfunktionen eingesetzt, bei denen die Strömungsgrößen am ersten Gitterpunkt zur Wandschubspannung in Beziehung gesetzt werden. In letzter Zeit wurden aber sogenannten Low-Reynolds-Number Versionen der verschiedenen Modelle entwickelt und auch getestet, bei denen die viskose Unterschicht aufgelöst wird.

Die Lösung der gemittelten Gleichungen zusammen mit einem Turbulenzmodell ist weniger aufwendig als eine LES Rechnung, aber für komplexe Probleme wie die Strömung um ein gesamtes Flugzeug oder durch Stufen von Turbinengittern kann die Rechenzeit doch mehrere Stunden auf heutigen Computern betragen. Holst et al. [13] berichten, daß bei Verwendung eines einfachen algebraischen Turbulenzmodells die Rechenzeit auf einem 150 MFlops Rechner für solche Strömungen in der Größenordnung 10 bis 20 Stunden für stationäre Rechnungen beträgt und bedeutend höher sein kann für instationäre Strömungen oder wenn komplexere Turbulenzmodelle verwendet werden. Für einfachere Geometrien sind die Rechenzeiten natürlich wesentlich geringer. So konnte die in Bild 3 dargestellte Strömung um einen Würfel bei Verwendung des $k-\varepsilon$ Modells mit Wandfunktionen auf dem Siemens SNI 600/20 Vektorrechner in 15 Minuten berechnet werden, und bei Auflösung der wandnahen Schicht in 8 Stunden. Zur Erinnerung sei erwähnt, daß die LES Rechnung für dasselbe Problem auf demselben Rechner 160 Stunden benötigte. Die $k-\varepsilon$ Modellrechnung stimmt allerdings wesentlich weniger gut mit den experimentellen Beobachtungen überein als die LES Rechnung (vor allem das Ablösegebiet hinter dem Würfel wird wesentlich zu lang berechnet).

Viele Probleme der Praxis können jedoch mit dem statistischen $k-\varepsilon$ Modell mit ausreichender Genauigkeit berechnet werden. So zeigt Bild 4 die berechneten Geschwindigkeitsvektoren in einem 3,6 km langen Flußabschnitt des Rheins an der deutsch/niederländischen Grenze [15]. Es handelt sich hierbei um eine zweidimensionale tiefengemittelte Rechnung, bei der nur die horizontalen Geschwindigkeitskomponenten berechnet werden. Der betrachtete Abschnitt ist geometrisch relativ komplex auf Grund der Flußkrümmung und der Vielzahl von Buhnen, zwischen denen sich sogenannte Totwassergebiete mit schwacher, rezirkulierender Strömung ausbilden. Die Übereinstimmung zwischen berechneter und in einem Laborversuch gemessener Geschwindigkeitsverteilung ist insgesamt gut und für praktische Zwecke ausreichend. Die Rechnung benötigte 155 Sekunden auf dem SNI S600/20 Rechner. Als letztes Beispiel werden in Bild 5 Rechnungen für ein rundes Nachklärbecken gezeigt, und zwar für das betrachtete axialsymmetrische Problem die Ergebnisse in einem Radialschnitt. Von oben kommt in der Nähe der Symmetrieachse (links) schlammbeladenes Wasser in das Becken, wird durch ein horizontales Leitblech nach rechts in radiale Richtung umgelenkt und durchströmt das Becken in dieser Rich-

tung. Dabei setzt sich der Schlamm am Beckenboden ab und wird unter dem Leitblech in der Mitte abgezogen, während auf der rechten Seite am äußeren Rand das geklärte Wasser abfließt. Es entsteht eine sehr komplexe Strömung mit mehreren Walzen, wie die im unteren Bildteil dargestellten berechneten Stromlinien zeigen. Außerdem bildet sich am Boden eine Schlammschicht mit klarem Schlammspiegel aus und darüber befindet sich relativ klares Wasser, wie die im oberen Bildteil dargestellte Schlammkonzentrationsverteilung illustriert. Insgesamt liefert die Rechnung realistische und somit für Auslegungszwecke geeignete Ergebnisse; wegen der sich einstellenden starken Dichteschichtung ist die numerische Lösung allerdings nicht sehr stabil und benötigt deshalb mehr Rechenzeit als das letzte Beispiel. Es sei jedoch noch erwähnt, daß die letztgenannten zweidimensionalen Rechnungen auch auf Workstations oder modernen PCs durchgeführt werden können.

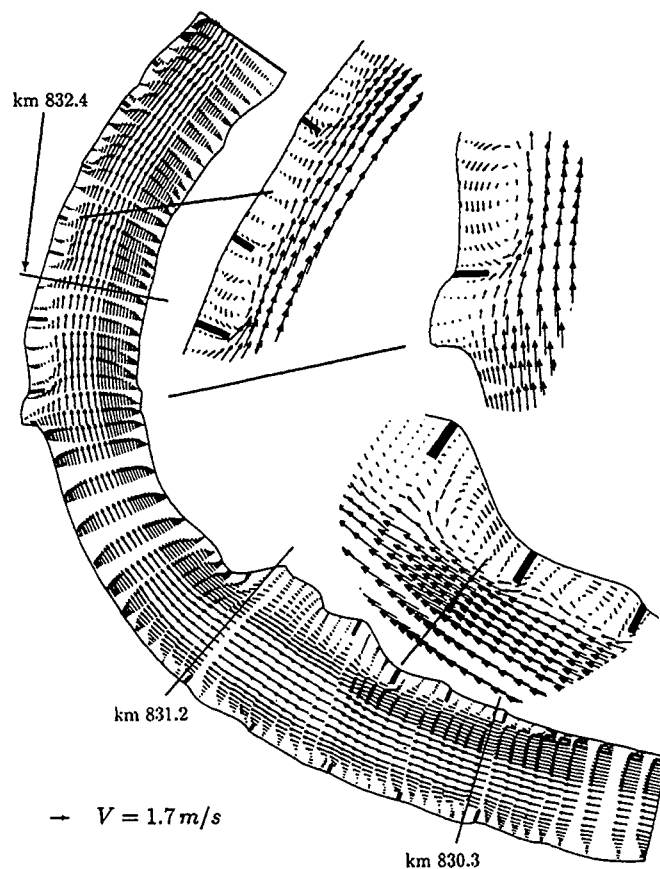


Bild 4: Berechnete Geschwindigkeitsvektoren in einem Abschnitt des Rheins bei Vynen-Rees

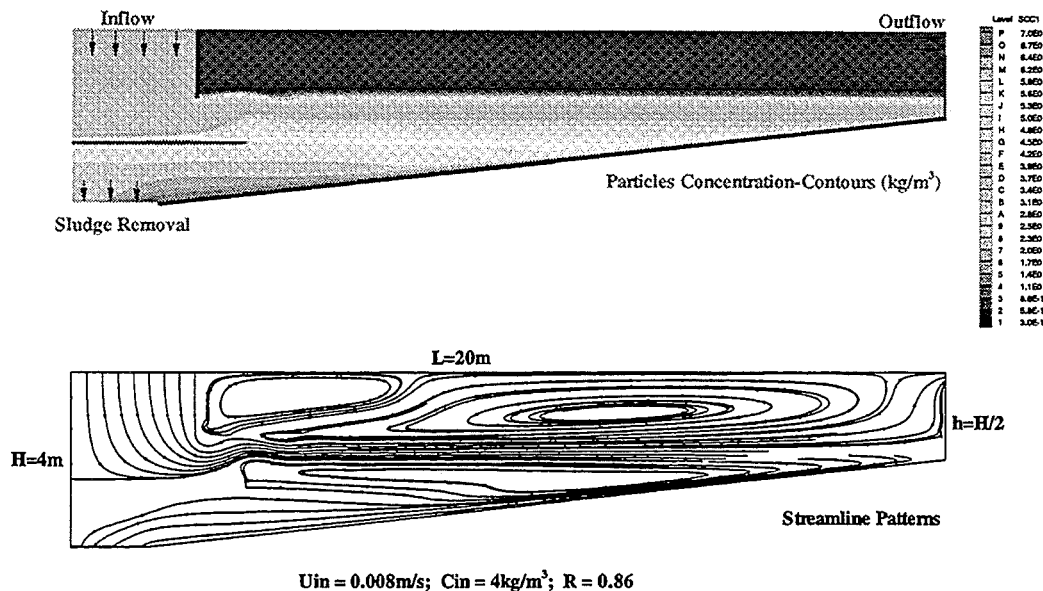


Bild 5: Berechnete Schwebstoffkonzentrationen (oben) und Stromlinien (unten) in einem runden Nachklärbecken

5 AUSBLICK AUF WEITERE ENTWICKLUNGEN

In der näheren Zukunft werden sicherlich hauptsächlich die statistischen Rechenverfahren weiter getestet und bei praktischen Rechnungen eingesetzt. Mit zunehmender Leistungsstärke der Rechner und geringer werdenden Kosten pro Rechenoperation können mehr und mehr Gitterpunkte verwendet und damit immer komplexere Probleme berechnet werden, unter Einsatz immer aufwendigerer Turbulenzmodelle. Die leistungsstärkeren Reynoldsspannungsmodelle werden dabei sicherlich vermehrt eingesetzt. Die Entwicklung der Rechnerleistung ist in Bild 6 dargestellt, in welcher die Rechengeschwindigkeit in MFlops aufgetragen ist gegen das Jahr der Einführung des entsprechenden Computers. Heute gibt es schon Maschinen mit mehreren hundert GFlops und zur Jahrhundertwende werden Teraflops-Maschinen verfügbar sein. Mit solchen äußerst leistungsstarken Maschinen eröffnen sich natürlich völlig neue Dimensionen bei der Berechnung turbulenter Strömungen. Es sollte dabei auch erwähnt werden, daß schon heute mit den relativ preisgünstigen Workstations mit einer Leistung von ca. 100 MFlops dreidimensionale Rechnungen mit statistischen Modellen möglich sind. Die bald verfügbare Teraflops-Leistung eröffnet die Möglichkeit von Large-Eddy-Simulationen auch für praktische Probleme, und da diese Methode vor allem für komplexere Strömungen mit großen Wirbelstrukturen wesentlich geeigneter ist, wird sie für praktische Rechnungen eine immer größere Rolle spielen. Mit den leistungsstärkeren Maschinen können natürlich auch aufwendigere direkte numerische Simulationen bei höheren Reynoldszahlen durchgeführt werden, doch wird der Einsatz dieser Methode in absehbarer Zeit auf Grundlagenuntersuchungen beschränkt bleiben.

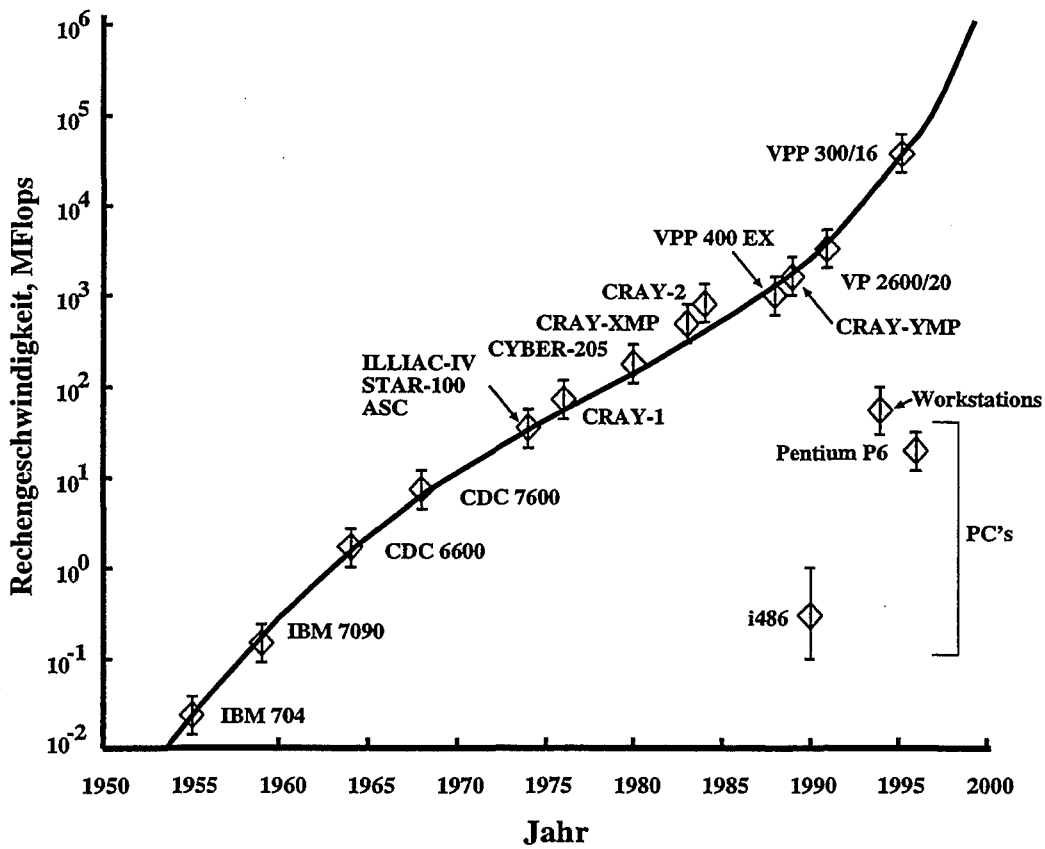


Bild 6: Entwicklung der Rechengeschwindigkeit

6 Literatur

- [1] J. Kim, P. Moin, R.D. Moser, "Turbulence statistics in fully developed channel at low Reynolds number", *J. Fluid Mech.*, 177, pp. 133-166 (1987).
- [2] I. Nezu, H. Nakagawa, "Turbulence in open channel flow", Balkema 1993. [3] S.J. Kline, W.C. Reynolds, F.A. Schraub, P.W. Rundstadler, "The structure of turbulent boundary layers", *J. Fluid Mech.*, 30, pp. 741-773 (1967)
- [4] P. Moin, H. Le, J. Kim, "Direct numerical simulation of turbulent flow over a backward-facing step", *Proc. 9th Symposium on Turbulent Shear Flows*, Kyoto, Japan (1993).
- [5] L. Kleiser, T.A. Zhang, "Numerical simulation of transition in wall-bounded shear flows", *An. Rev. Fluid Mech.*, 23, pp. 495-537 (1991).
- [6] T. Tamura, I. Ohta, K. Kuwahara, "On the reliability of two-dimensional simulation for unsteady flows around a cylinder-type structure", *J. Wind Eng. and Ind. Aerodyn.*, 35, pp. 275-298 (1990).
- [7] M. Germano, U. Piomelli, P. Moin, W.H. Cabot, "A dynamic subgrid-scale eddy-viscosity model", *Physics of Fluids A*, 3, pp. 1760-1765 (1991).
- [8] U. Schumann, "Subgrid-scale model for finite-difference simulations of turbulent flows in plane channel and annuli", *J. Comp. Phys.*, 18, pp. 376-404 (1975).
- [9] A. Werner, H. Wengle, "Large-eddy simulation of turbulent flow over and around a cube in a plane channel", *Proc. 8th Symp. on Turbulent Shear Flows*, Munich, Germany (1991).
- [10] M. Breuer, W. Rodi, "Large-eddy simulation for complex turbulent flows of practical interest", *Notes in Numerical Fluid Mechanics*, Vol. 52, ed. E.H. Hirschel, Vieweg Verlag, pp. 258-274 (1996).
- [11] D. Lakehal, W. Rodi, "Calculation of the flow past a surface-mounted cube with two-layer turbulence models", *Proc. 2nd Int. Symp. on Computational Wind Engineering*, Fort Collins, Colorado, USA (1996).
- [12] A. Akselvol, P. Moin, "Large-eddy simulation of a backward-facing step flow", in *Engineering Turbulence Modelling and Experiments 2*, W. Rodi and F. Martelli eds., Elsevier Science Publishers, pp. 303-313 (1993).
- [13] T.L. Holst, M.D. Salas, R.W. Claus, "The NASA computational aerosciences program - towards Teraflops computing", paper AIAA 92-0558 (1992).
- [14] R. Martinuzzi, C. Tropea, "The flow around a surface-mounted prismatic obstacle placed in a fully developed channel flow", *J. Fluids Eng.*, 115, pp. 85-92, (1993).
- [15] Th. Wenka, "Numerische Berechnung von Strömungsvorgängen in naturnahen Flußläufen mit einem tiefengemittelten Modell", *Dissertation*, Universität Karlsruhe, 1992.
- [16] D. Lakehal, "Numerical prediction of buoyant flows and particle settling in secondary clarifiers", Report No. 738, Institut für Hydromechanik, Universität Karlsruhe, 1997.