Es ist nicht ganz klar, daß es nicht geht

Das wissenschaftliche Werk von Christoph Zenger und sein Einfluß auf material science, life sciences, financial engineering und data mining

Michael Griebel

Institut für Angewandte Mathematik der Universität Bonn

Heute vormittag hatten wir das Vergnügen, eine Reihe von wissenschaftlichen Vorträgen aus den Bereichen Mathematik, Numerik, Wissenschaftliches Rechnen, Strömungsmechanik und Astrophysik zu hören. Nach mir wird Herr Kollege Bungartz zum philosophischen Werk unseres Jubilars Christoph Zenger sprechen. Mir verbleibt die Ehre, in der nun folgenden Stunde das wissenschaftliche Werk von Christoph Zenger darzustellen und zu würdigen. Hierbei möchte ich zunächst seine Karriere und seine zentralen Beiträge zur Mathematik und Informatik vorstellen und anschließend aufzeigen, welche Möglichkeiten sich mit seinen Ideen und Methoden in einer Reihe moderner interdisziplinärer Wissenschaftsbereiche eröffnen, was für die TU München als berühmte und berüchtigte Reformuniversität des angehenden einundzwanzigsten Jahrhunderts sicher von Interesse ist.

Klar, wir kennen Christoph Zenger alle als umtriebigen Macher im Spannungsbereich zwischen Informatik, Mathematik und Ingenieurwissenschaften, als Gründungssprecher des Bayerischen Forschungsverbundes für technisch-wissenschaftliches Hochleistungsrechnen (FORTWIHR), als aktuellen Sprecher des Sonderforschungsbereichs 438 „Mathematische Modellierung, Simulation und Verifikation in materialorientierten Prozessen und intelligenten Systemen“, als Mitglied des Sonderforschungsbereichs 342 „Methoden und Werkzeuge für die Nutzung paralleler Rechnerarchitekturen“, des Sonderforschungsbereichs 411 „Grundlagen der aeroben biologischen Abwasserreinigung“ und des Graduiertenkollegs „Kooperation und Ressourcenmanagement in verteilten Systemen“. Wir kennen ihn auch als Herausgeber der Zeitschrift Numerische Mathematik, als Direktor der Ferienakademie der Universität Erlangen-Nürnberg und der Technischen Universität München, als Direktoriumsmitglied des LRZ und des KONWIHR, als Sekretär der Kommission für Informatik der Bayerischen Akademie der Wissenschaften, als Mitglied in diversen Ausschüssen und wissenschaftlichen Beiräten (Kommission für Rechenanlagen der DFG, Carl F. Gauss Minerva Center for Scientific Computing des Weizmann Institute of Science, Rehovot/Israel, Universität Konya/Türkei, Lenkungsausschuß Bundes-Höchstleistungsrechenzentrum der Univ. Stuttgart, Direktor NATO Sommerschule Antalya). Und wir kennen ihn als verantwortungsbewußten Hochschullehrer (Vorlesungsreihen „Technik und Ethik“, und „Entwicklung der 3. Welt, Herausforderung für Wissenschaft und Technik“), als jährlichen talarbewehrten Marschierer in der Münchner

Innenstadt an Fronleichnam und als Drahtzieher (neben anderen Kollegen natürlich) bei der Planung und Beschaffung des schnellsten zivilen Hochleistungsrechners der Welt (Stichwort: Kühne Visionen werden Realität) für das Rechenzentrum der Bayerischen Akademie der Wissenschaften gegen die harte Konkurrenz nördlicherer Bundesländer und Stadtstaaten unter wechselnden politischen Randbedingungen.

Aber was wissen wir eigentlich über seine früheren Forschungsergebnisse und Arbeiten in der Zeit vor FORTWIHR und den aktuellen Sonderforschungsbereichen? Was wissen wir über seine vorherigen 60-24 (Jugend und Studium) - 10 (FORTWIHR) = 26 Jahre wissenschaftlichen Wirkens?

Nun, Prof. Bauer, Prof. Stoer und manch anderer kennt das alles gut und wird sicherlich nicht überrascht werden. Aber im Zuge der Vorbereitung zu diesem Vortrag ist es mir gelungen, mich selbst zu überraschen, als ich mir frühere Arbeiten von Christoph Zenger genauer angeschaut habe. Ich hoffe, daß mir das mit meinen Ausführungen auch bei dem einen oder anderen von Ihnen gelingt. Ich bin sicher, daß mir dies auf jeden Fall bei den jüngeren Mitarbeitern des Lehrstuhls Zenger gelingen wird.

Das Gesamtwerk läßt sich zusammenfassend sicherlich als vielseitig, breit, interdisziplinär und aktuell charakterisieren. Es hat eine weite Ausstrahlung und ermöglicht eine Fülle schöner Anwendungen, worüber später noch zu berichten sein wird. Ihm zugrunde liegen feste Prinzipien und einfache klare Konzepte. Dies sind u.a. Lokalität, Hierarchie, Produktbildung und Dimensionsrekursivität. Und natürlich ein unumstößlicher Optimismus, daß eine gestellte Aufgabe auch irgendwie zu bewältigen sein wird, getreu dem Motto: *Es ist nicht ganz klar, daß es nicht geht.*

Lassen Sie mich nun im folgenden versuchen, zunächst den wissenschaftlichen Werdegang Christoph Zengers und eine Auswahl seiner zentralen Ergebnisse darstellen.

1 Werk

1.1 Die frühen Jahre

Von Hause aus, und das ist vielleicht nicht allen hier bekannt, ist Christoph Zenger eigentlich Physiker (*Physiker sind sowieso die besseren Mathematiker*). Ob nun die Zeichen der Zeit (hier sei an das aktuelle Broadway-Theaterstück *Copenhagen* des Briten M. Frayn erinnert, siehe Spiegel 36/2000, Seite 250 ff) damals so standen oder die intellektuelle Neigung dies mit sich brachte, auf jeden Fall studierte Christoph Zenger von 1959-1964 Physik an der Universität in München. Seine Diplomarbeit schrieb er bei Prof. Koppe, einem Heisenberg-Schüler, in Theoretischer Physik über die Wignersche Störungstheorie in der Quantenmechanik. Durch die Berufung von Prof. Koppe nach Kiel waren dabei eine Reihe von Fahrten in den hohen Norden notwendig. Nach eigenem Bekunden beinhaltete die Diplomarbeit ein interessantes neues Ergebnis, nur die räumliche Distanz mag der Grund gewesen sein, warum sich eine gemeinsame Veröffentlichung mit Koppe darüber letztendlich nicht ergeben hat.

Heirat und Anstellung als Wissenschaftlicher Mitarbeiter am Rechenzentrum der TU München gingen Hand in Hand und erfolgten im Jahr 1964. Im November 1965 kam dann der Wechsel an das neugegründete Leibniz Rechenzentrum der Bayerischen Akademie der Wissenschaften, wo Christoph Zenger dort bis 1976 als Wissenschaftlicher Mitarbeiter, ab 1969 dann als Leiter der Abteilung Wissenschaftliche Benutzerbetreuung verblieb.

Das müssen interessante Zeiten gewesen sein: Nachdem die PERM („Programmgesteuerte elektronische Rechenanlage München") schon vorher an der TUM von den Herren Sauer und Piloty gebaut worden war, wurde in der Richard-Wagner-Straße 18, im Rechenzentrum der kurz vorher gegründeten Kommission für elektronisches Rechnen der Bayerischen Akademie der Wissenschaften, der Betrieb auf der Rechenanlage TR4 der Firma Telefunken aufgenommen. Es wurde dort ein ALGOL-Compiler und ein Betriebssystem für die TR4 erstellt sowie Forschung auf dem Gebiet der Informatik („Informationsverarbeitung") betrieben. 1967 wurde der Grundstein des jetzigen LRZ-Gebäudes gelegt, und im Herbst 1968 fiel schließlich die Entscheidung für die Doppelprozessoranlage TR440 der Firma AEG-Telefunken. Die akademische Informatik existierte zwar noch nicht, stand aber als eigener Studiengang Informationsverarbeitung im Jahr 1967 kurz bevor.

In diesen Zeiten waren Fragen der Linearen Algebra und damit verbundene Fragen zu Matrizen, ihrem Spektrum und Wertevorrat dringlich und von praktischem Interesse. Auf diesem Gebiet begann Christoph Zenger inspiriert durch [6] vordringlich zu arbeiten. 1967 erfolgte die Promotion an der Fakultät für Allgemeine Wissenschaften der TUM unter Betreuung von Prof. Bauer und Prof. Stoer mit der Arbeit *Verallgemeinerte Wertevorräte von Matrizen* [93]. Hierbei ist besonders interessant, daß wesentliche Teile im Schwimmbad in der Sonne entstanden sind. (*Warum sind Sie so nervös mit Ihrer Promotionsarbeit? Das ist alles nicht so schlimm, ich habe meine Arbeit im Schwimmbad gemacht. Nur dort hatte ich Ruhe.*)

Im Nachlauf der Promotionsarbeit entstand der Artikel *On convexity properties of the Bauer field of values of a matrix*, der 1968 in der Zeitschrift Numerische Mathematik erschien [94]. Mit dieser Arbeit gelang es Christoph Zenger, frühen Ruhm zu erwerben. Mithin findet sich in der Lecture Note Serie der renommierten London Mathematical Society, Band 10, Numerical Ranges II von F.F. Bonsall, J. Duncan [12], dem frühen Standardwerk zu numerischen Wertebereichen, ab Seite 18 ein eigenes Kapitel mit dem Titel *A Theorem of Zenger*, siehe Abbildung 1.

Insbesondere Theorem 3 wurde zudem von Crabb geringfügig erweitert und wird immer wieder als Zenger-Crabb Theorem gerne verwendet und zitiert. Es wird u.a. gezeigt, daß die konvexe Hülle des Punktspektrums im räumlichen numerischen Wertebereich liegt. Dies hat wichtige Implikationen in Operatortheorie, Funktionalanalysis, multilinearer Algebra und numerischer Analysis, siehe auch `http://www.math.wm.edu/~ckli/wonra.html`.

19. A THEOREM OF ZENGER

The stop press of NRI (§14, Remark (13)) drew attention to Chr. Zenger [78] in which it was proved that the convex hull of the point spectrum is contained in the spatial numerical range. The present section contains the proof of this remarkable result and of its extension due to M. J. Crabb [130].

Theorem 2. (Zenger [78].) *Let* X *be a normed linear space over* $\underset{\sim}{C}$, *let* $a_1, \ldots, a_r$ *be linearly independent elements of* X, *and let* $\alpha_k \geq 0$ $(k = 1, \ldots, r)$, $\sum_{k=1}^{r} \alpha_k = 1$. *Then there exist* $(x, f) \in \Pi(X)$, *and complex numbers* $w_1, \ldots, w_r$, *such that* $x = w_1a_1 + \ldots + w_ra_r$ *and*

$$f(w_ka_k) = \alpha_k \qquad (k = 1, \ldots, r).$$

Notation. Given a normed linear space X over $\underset{\sim}{C}$, and a linear mapping T of X into X we denote by $p\,Sp(T)$ the *point spectrum* of T, i. e. the set of $\lambda \in \underset{\sim}{C}$ such that $Tx = \lambda x$ for some non-zero $x \in X$.

Theorem 3. (Zenger [78].) *Let* X *be a normed linear space over* $\underset{\sim}{C}$ *and let* T *be a linear mapping of* X *into* X. *Then* $co\ p\,Sp(T) \subset V(T)$.

Abbildung 1. Das Theorem von Zenger

Ohne hier weiter ins Detail gehen zu wollen: Auch über die fachliche Leistung und den eigentlichen Inhalt hinaus ist ein eigenes Theorem auf jeden Fall eine gute Sache, umso mehr, da eine durchaus mögliche Frage unter manchen Kollegen der Reinen Mathematik andernorts bei Gesprächen über Kollegen von auswärts diejenige nach dem eigenen Theorem ist.

Zwei Jahre später entstand dann in Zusammenarbeit mit E. Deutsch der Artikel *Inclusion domains for the eigenvalues of stochastic matrices* [28]. Eine stochastische Matrix ist eine nichtnegative Matrix A, in der jede Spaltensumme gleich 1 ist. Sie entspricht direkt einem Markov Prozeß, und ihre Einträge a_{ij} beschreiben die Wahrscheinlichkeiten des Übergangs vom Zustand i in den Zustand j in einer stationären Markovkette. Schaefer, Brauer, Deutsch und Hadeler gaben bereits vorher Abschätzungen für die nichttrivialen Eigenwerte stochastischer Matrizen. In [28] wird nun eine Menge $G(A)$ komplexer Zahlen konstruiert (genauer gesagt, ein verallgemeinerter Bauerscher Wertevorrat), die alle nicht-trivialen Eigenwerte von A beinhaltet. Der entsprechende numerische Radius $r(A) = \max_{z \in G(A)} |z|$ ergibt sich zu

$$r(A) = \frac{1}{2} \max_{i,j} \left(a_{ii} + a_{jj} - a_{ij} - a_{ji} + \sum_{k \neq i, k \neq j} |a_{ik} - a_{jk}| \right), \qquad (1)$$

und es wird gezeigt, daß $r(A)$ die beste aller bisherigen Schranken ist. Auch in der jüngsten Ausgabe von E. Senetas *Non-negative matrices and Markov chains* [81] ist $r(A)$ immer noch die beste bekannte Schranke. Hier bestehen enge Bezüge zu den bekannten Ergodizitätskoeffizienten von Markovprozessen, insbesondere dem Koeffizienten von Dobrushin. Adolf Rhodius führt dies in seinem Artikel „On explicit forms for ergodicity coefficients“ [73] genauer aus und diskutiert in [74] neue Formeln für den Deutsch-Zenger Radius. Schließlich wird der Deutsch-Zenger Radius sogar zur Charakterisierung von „almost scrambling“ Matrizen eingesetzt [75], die bei inhomogenen Markovketten und Wahrscheinlichkeitsautomaten eine wichtige Rolle spielen.

In diese Zeit müssen auch Christoph Zengers Ausflüge in die Privatwirtschaft gefallen sein: Als freier Programmierer mit Cobol Background fuhr er eine Zeit lang mit einem Kollegen aus dem LRZ mit dem Nachtzug in den Norden und in der übernächsten Nacht wieder zurück: *Ich kann im Zug seit dieser Zeit hervorragend schlafen, im Liegewagen müssen Sie sich nur einen Platz suchen, der nicht über den Achsen liegt.*

Als Berater (heute würde man Consultant sagen) für eine große Tiefbaufirma hatte sich Zenger an die Entwicklung und Konstruktion eines Rüttelgeräts, eine Baumaschine zum Verdichten des Bodens, gemacht: *Ich habe die bestehenden Rüttler verbessert. Ja, als Physiker weiß man, wie das gehen muß: Der Trick war, den Massenschwerpunkt in die Drehachse zu legen. Das verbessert den Wirkungsgrad ungeheuer. Der Rüttler ist dann sogar nach meinen Vorschlägen gebaut worden. Und als der Techniker den Rüttler eingeschaltet hat, hat der zu rütteln und zu rütteln begonnen, und der Boden hat* *sich verdichtet und verdichtet. Und dann ist der Rüttler explodiert. Der Techniker hat wohl eigenmächtig Details der Aufhängung verändert. Das Konzept war auf jeden Fall in Ordnung.* Bleibt nachzutragen, daß ein Nachfolgeprodukt es sogar bis zum Patentantrag geschafft hat und erst nach mehreren Instanzen nicht patentiert wurde.

1.2 In der Ruhe liegt die Kraft

Im Jahr 1975 erfolgte schließlich die Habilitation an der Fakultät für allgemeine Wissenschaften der TU München mit der Schrift *Minimale subadditive Einschließungsgebiete für die Eigenwerte von Matrizen* [96]. In ihr untersucht Christoph Zenger, wie man aus Einschließungsgebieten für das Spektrum von zwei Matrizen A und B auf ein Einschließungsgebiet für das Spektrum von $A + B$ folgern kann. Als Hauptergebnis wird gezeigt, daß die konvexe Hülle eines beliebigen Wertevorrats bezüglich einer Norm eine beste konvexe subadditive Eigenwerteinschließung darstellt. Diese Fragestellungen haben einen starken Bezug zur Theorie der dissipativen Operatoren und der monotonen Operatoren und resultieren in Sätzen, die eng verwandt sind mit denen von Varga über reguläre Zerlegungen von Matrizen.

Nach der Habilitation war Christoph Zenger von 1976 bis 1980 als Hochschullehrer an der Technischen Universität München im Fachbereich Informatik tätig, zunächst als Wissenschaftlicher Rat. In dieser Zeit erfolgte eine Hinwendung zur numerischen Behandlung partieller Differentialgleichungen mittels Finiter Differenzen Verfahren. Ein wichtiges Thema hierbei ist die Genauigkeit der erzielten diskreten Lösung. Für konventionelle Diskretisierungen des Poissonproblems mit etwa dem 5-Punkte Stern läßt sich bei glatten Lösungen eine Genauigkeit der Ordnung $O(h^2)$ erzielen, wobei hier h die Maschenweite des verwendeten uniformen Gitters bezeichnet. Ist jedoch der Rand des betrachteten Gebiets nicht glatt und weist etwa einspringende Ecken, Schlitzform oder Schlimmeres auf, dann besitzt die Lösung eine Singularität. In Folge dessen sinkt die Genauigkeitsordnung u.U. dramatisch. Dies gilt sowohl in der Nähe der einspringende Ecke als auch weit davon entfernt (Pollutionseffekt). Gegenmaßnahmen hierzu sind geeignet adaptiv verfeinerte oder gradierte Gitter, spezielle globale singuläre Finite Elemente Ansatzfunktionen oder gewisse Modifikationen einer h-abhängigen Zahl von Finiten Differenzen Sternen in der Nähe der einspringenden Ecke. All diese Ansätze sind jedoch nicht-lokal und erhöhen den Berechnungsaufwand des Gesamtverfahrens.

In der Arbeit *Improved Difference Schemes for the Dirichlet Problem of Poisson's Equation in the Neighborhood of Corners* [97] gelingt es Christoph Zenger zusammen mit H. Gietl in eleganter Weise, das oben erwähnte Problem durch Modifikationen einiger weniger Differenzensterne nahe der Singularität zu beheben. Verblüffend hierbei ist, daß dies rein lokal möglich ist, das heißt, die Zahl der zu modifizierenden Sterne ist nicht von h abhängig, sondern eine sehr kleine Konstante. Dieses Vorgehen erlaubt erstmalig, die

Konvergenzordnung des glatten Falls auf kostengünstige und einfache Weise wiederzugewinnen. Zudem läßt sich nun auch Extrapolation zur weiteren Verbesserung der Konvergenzordnung wieder einsetzen. Abbildung 2 zeigt den Effekt der Zengerschen Modifikation für das zweidimensionale Poisson Problem im Fall des L-Gebiets (Problem I: $u(z) = im(z^{2/3})$, Problem II: $u(z) = |z|^2$).

Table 1

h	Usual five-point difference method	Value of the modified method	Value after one h^2-extrapolation step	Estimation for the first coefficient τ_0
Problem I at $z=-\frac{1}{2}$, Exact value: $u(z)=0.54556181$				
$\frac{1}{4}$	0.53116415	0.54386897		0.9999314
$\frac{1}{8}$	0.54042663	0.54515189	0.54557953	1.0000168
$\frac{1}{16}$	0.54362352	0.54546397	0.54556800	1.0000069
$\frac{1}{32}$	0.54481010	0.54553752	0.54556204	1.0000022
Problem II at $z=-\frac{1}{2}$				
$\frac{1}{4}$	0.76340455	0.78243770		1.498002
$\frac{1}{8}$	0.78309117	0.79049182	0.79317653	1.566218
$\frac{1}{16}$	0.78947940	0.79240927	0.79304841	1.5919381
$\frac{1}{32}$	0.79170306	0.79286813	0.79302108	1.601634

Abbildung 2. Kleiner Effekt aber große Wirkung: Die Zenger-Korrektur des Differenzensterns

In einer Reihe von Dissertationen [41,101,34,1,76] wurde dieses Konzept in den Folgejahren sukzessive auf andere dateninduzierte Singularitäten (rechte Seiten, allgemeinere Gebiete und Randdaten, Koeffizientenfunktionen des Operators) und allgemeinere Differentialoperatoren übertragen.

Hier war Christoph Zenger seiner Zeit und auch vielen heutigen Bemühungen voraus. Die Idee ist bestechend: Singularitäten in der Lösung können eigentlich nur durch entsprechende Entartungen in den Daten einer Differentialgleichung, also der Randkurve, der rechten Seite, der Randdaten oder der Koeffizientenfunktionen des Differentialoperators induziert werden. Und diese sind Eingabedaten für das Problem, also a priori bekannt. Jüngste Entwicklungen aus dem Bereich der N-Term Approximation und der adaptiven Verfeinerung ohne Saturationsannahme [29,24] nutzen das gleiche Prinzip in modifizierter Weise. Bei der Fehlerschätzung und adaptiven Netzverfeinerung werden im Prinzip lediglich die Daten adaptiv aufgelöst und damit der Algorithmus getrieben. Erst vor zwei Jahren konnte die Existenz eines solchen Approximationsverfahrens gezeigt werden, und erst letztes Jahr war es meines Wissens möglich, eine lauffähige praktikable Implementierung zu schaffen [4]. Das zeigt, daß eine a posteriori Fehlerschätzung, wie sie in vielen modernen PDE Lösern gerade implementiert wurde, eigentlich gar nicht nötig ist, sondern vielmehr die Datendiskretisierungsfehler beherrscht werden müssen. Der o.e. Waveletalgorithmus [24,4] ist aber immer noch nicht nicht rein lokal, wie dies bei den Zengerkorrekturen der Fall ist.

Im Jahr 1980 erfolgte schließlich der Ruf auf eine C4-Professur an die Hochschule der Bundeswehr in München im Fachbereich Informatik. Die folgenden zwei Jahre waren nach Christoph Zengers eigenem Bekunden seine schönste und ruhigste Zeit: *Herr Bungartz, wenn sie viel spazierengehen wollen und ruhig arbeiten und forschen wollen, dann machen sie es wie ich und gehen an die BWH. So viel Muße hatte ich nie wieder und so früh war ich jeden Tag nie wieder zu Hause.*

1.3 In den besten Jahren

Im Jahr 1982 ereilte Christoph Zenger schließlich der Ruf auf die Nachfolge Samelson zurück an die TUM, und die kleine Münchner Runde war abgeschlossen. Es erfolgte der Einstieg in die Informatik-Lehre. So wurde mit sparsamsten Mitteln der Aufbau des Fortran Praktikums für Ingenieure angegangen. Reinhard Fößmeier schrieb dabei Teile eines Fortran-Compilers selbst, Dietmar Friede entwickelte zusammen mit der Rechnerbetriebsgruppe unter Herrn Gerold den berüchtigten PAX Rechner auf der Basis des Intel 80286 Prozessors unter dem Unix Betriebssystem. Leider war der Name des Rechners nicht immer Programm, auf jeden Fall standen damit erstmalig workstation-artige Arbeitsplatzrechner zur Verfügung. Der Aufbau einer Arbeitsgruppe beginnt. Neben Reinhard Fößmeier umfaßt sie 1985 zunächst Uli Rüde, der aus USA zurückgekehrt war und die Kenntnis der Mehrgitterverfahren zur Lösung partieller Differentialgleichungen einbrachte, sowie Ralph Steinbrüggen und mich selbst. Neben Ursula Hill-Samelson stößt 1989 Hans Bungartz hinzu, und danach wächst die Gruppe über die Zeit kontinuierlich auf fast 20 Personen.

In dieser Zeit erfolgte die Hinwendung zu hierarchischen Verfahren: *Herr Griebel, da gibt's so einen jungen Kerl in Aachen mit einem schwierigen französischen Namen, der hat in seiner Habilitation was ganz Interessantes mit hierarchischen Basen gemacht, schauen Sie sich das mal an.* Gemeint war hier Harry Yserentant, dem es mittels hierarchischer linearer Basen [32] in [91] gelungen war einen, zumindest in zwei Raumdimensionen, fast optimalen Vorkonditionierer für das Lösen der bei der Diskretisierung von elliptischen partiellen Differentialgleichungen entstehenden linearen Gleichungssysteme zu konstruieren.

Darauf aufbauend wurde von Christoph Zenger schließlich 1989 die Idee ausgebrütet, die eindimensionalen hierarchischen Basen in einer Produktkonstruktion auf den zwei- und mehrdimensionalen Fall zu verallgemeinern: Die sogenannten dünnen Gitter waren (wieder)-geboren[1].

Bereits die erste Arbeit *Sparse Grids*, 1990 als Blauer Bericht in der Reihe des SFB 342 gedruckt und dann 1991 als Beitrag in den Notes on Numerical Fluid Mechanics 31 [100] erschienen, beinhaltete alle wesentlichen Komponenten: Theoretische Grundlagen, Fehleranalyse und Komplexitätsdiskussion, sowie, programmiert von Christoph Zenger Junior, die adaptive Behandlung des Poisson-Problems in zwei Dimensionen.

Die Idee ist wiederum einfach aber bestechend und hat weitreichende Folgen: Wir betrachten eine eindimensionale glatte Funktion, beispielsweise die Parabel $u(x) = -(x-1)(x+1)$ über einem Gebiet $\Omega = [-1,1]$. Eine Interpolation auf einem uniformen Gitter Ω_k der Maschenweite $2 \cdot 2^{-k}$ mit stückweise linearen Basisfunktionen $\phi_{k,i}(x), i = 1, ..., 2^k - 1$ führt zur Darstellung $u_k(x) = \sum_{i=1}^{2^k-1} u_{k,i}\phi_{k,i}$ wobei die Koeffizienten $u_{k,i}$ gerade die Werte von u in den Gitterpunkten annehmen. Die Größenordnung der Koeffizienten ist hier $O(1)$. Die Abbildungen 3 (links) und 4 (links) stellen dies anschaulich dar.

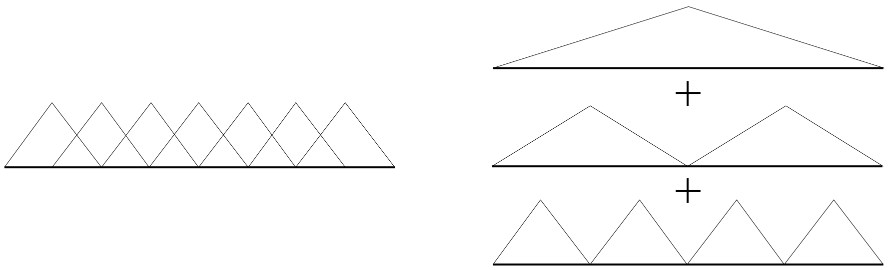

Abbildung 3. Knotenbasis (links) und hierarchische Basis (rechts), $k = 3$

Verwendet man jedoch statt der Knotenbasis die eindimensionale hierarchische Basis

$$\{\phi_{l,i}, l = 1, .., k, i = 1, .., 2^l - 1, i \text{ ungerade}\},$$

[1] Später haben wir durch einen Hinweis von E. Novak festgestellt, daß sich dieses Prinzip bis auf das Jahr 1963 zu einer Arbeit von S.A. Smolyak [82] zurückverfolgen läßt. In der Approximationstheorie ist ein analoges Vorgehen mit Fourierbasen unter dem Namen „hyperbolic crosses“ bekannt und führt zu den sogenannten Korobov-Räumen. Auch ist die Variante von Dünn-Gitter-Verfahren mittels der Kombinationstechnik als „blending scheme“ oder „boolean scheme“ bei Interpolation und Integration bekannt [5,27]. Die Anwendung auf partielle Differentialgleichungen war aber auf jeden Fall neu.

siehe auch Abbildung 3 (rechts), so ergibt sich die Darstellung

$$u_k(x) = \sum_{l=1}^{k} \sum_{\substack{i=1 \\ step\ 2}}^{2^l-1} u_{l,i}\phi_{l,i}(x),$$

und für die zugehörigen hierarchischen Koeffizienten $u_{l,i}$ beobachtet man einen Abfall um genau den Faktor $1/4$ von Level zu Level, siehe Abbildung 4 (rechts). Allgemeiner gilt für Funktionen u mit beschränkter zweiter Ableitung das Abfallsgesetz $|u_{l,i}| \leq C \cdot 2^{-2l}$.

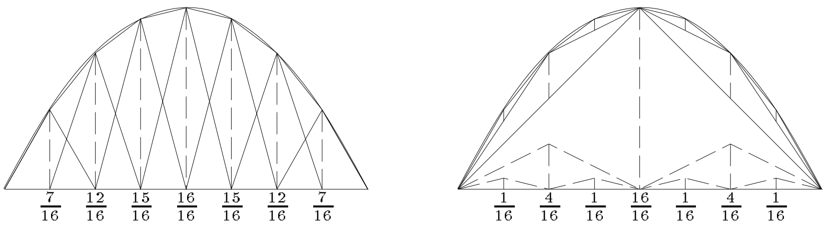

Abbildung 4. Koeffizienten für die Standardbasis und für die hierarchische Basis, $k = 3$

Christoph Zengers Idee war es nun, die hierarchische Basis mit Hilfe einer Produktkonstruktion auf den d-dimensionalen Fall zu verallgemeinern. Über dem Gebiet $\Omega = [-1, 1]^d$ verwendet er die hierarchische Basis

$$\{\phi_{\mathbf{l},\mathbf{i}}, l_j = 1, .., k, i_j = 1, .., 2^{l_j} - 1, i_j \text{ ungerade}, j = 1, .., d\}$$

mit den stückweise d-linearen Hut-Funktionen

$$\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) := \prod_{j=1}^{d} \phi_{l_j,i_j}(x_j).$$

Hier bezeichnet $\mathbf{l} = (l_1, .., l_d) \in \mathbb{N}^d$ den Multi-Index für das Level, $\mathbf{i} = (i_1, .., i_d) \in \mathbb{N}^d$ den Multi-Index für den Ort und $\mathbf{x} = (x_1, .., x_d)$.

Die Träger der hierarchischen Basisfunktionen sind nun im allgemeinen nicht mehr uniform sondern können konstruktionsbedingt verschieden lang in die verschiedenen Koordinatenrichtungen sein. Sie sind im allgemeinen auch nicht mehr so lokal wie konventionelle hierarchische Basen, siehe Abbildung 5 für den zweidimensionalen Fall.

Die Interpolation einer Funktion $u(\mathbf{x})$ auf einem uniformen Gitter Ω_k der Maschenweite $2 \cdot 2^{-k}$ führt nun zur hierarchischen Darstellung

$$u_k(\mathbf{x}) = \sum_{l_1=1}^{k} ... \sum_{l_d=1}^{k} \sum_{\substack{i_1=1 \\ step\ 2}}^{2^k-1} ... \sum_{\substack{i_d=1 \\ step\ 2}}^{2^k-1} u_{\mathbf{l},\mathbf{i}}\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x}). \tag{2}$$

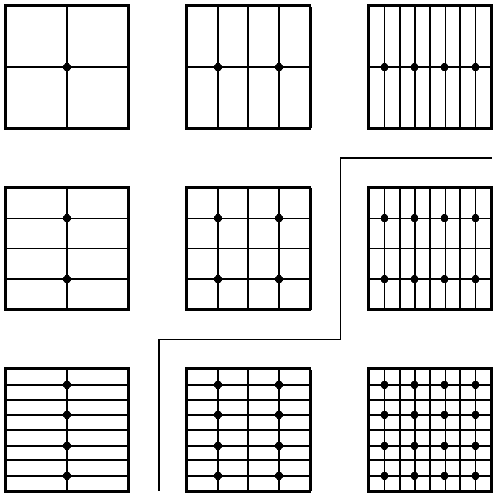

Abbildung 5. Träger der hierarchischen Basisfunktionen in 2D, k=3

Betrachten wir nun exemplarisch das zweidimensionale Paraboloid $u(x_1, x_2) = (1-x_1)(1+x_1)(1-x_2)(1+x_2)$ über $[-1, 1]^2$. Die hierarchischen Koeffizienten $u_{\mathbf{l},\mathbf{i}}$ fallen nun in x_1- und x_2-Richtung wie im eindimensionalen Fall um den Faktor $1/4$ von Level zu Level ab. In der Diagonalrichtung beobachtet man sogar einen Abfall um den Faktor $1/16$, siehe Abbildung 6 (rechts). Allgemeiner gilt für Funktionen u mit beschränkter gemischter $2d$-ter Ableitung[2] das Abfallgesetz $|u_{\mathbf{l},\mathbf{i}}| \leq C \cdot 2^{-2|\mathbf{l}|_1}$ mit $|\mathbf{l}|_1 = l_1 + ... + l_d$. Da nun bei der Interpolation in x_1- und x_2-Richtung ja auf Level k abgeschnitten wird, also nur die Koeffizienten bis zur Größe 2^{-2k} in Betracht gezogen werden, liegt es

[2] Eine genauere Darstellung der Räume mit beschränkter gemischter Ableitung findet man in [15,19,47]

$\frac{49}{256}$	$\frac{21}{64}$	$\frac{105}{256}$	$\frac{7}{16}$	$\frac{105}{256}$	$\frac{21}{64}$	$\frac{49}{256}$
$\frac{21}{64}$	$\frac{9}{16}$	$\frac{45}{64}$	$\frac{3}{4}$	$\frac{45}{64}$	$\frac{9}{16}$	$\frac{21}{64}$
$\frac{105}{256}$	$\frac{45}{64}$	$\frac{225}{256}$	$\frac{15}{16}$	$\frac{225}{256}$	$\frac{45}{64}$	$\frac{105}{256}$
$\frac{7}{16}$	$\frac{3}{4}$	$\frac{15}{16}$	1	$\frac{15}{16}$	$\frac{3}{4}$	$\frac{7}{16}$
$\frac{105}{256}$	$\frac{45}{64}$	$\frac{225}{256}$	$\frac{15}{16}$	$\frac{225}{256}$	$\frac{45}{64}$	$\frac{105}{256}$
$\frac{21}{64}$	$\frac{9}{16}$	$\frac{45}{64}$	$\frac{3}{4}$	$\frac{45}{64}$	$\frac{9}{16}$	$\frac{21}{64}$
$\frac{49}{256}$	$\frac{21}{64}$	$\frac{105}{256}$	$\frac{7}{16}$	$\frac{105}{256}$	$\frac{21}{64}$	$\frac{49}{256}$

$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{4}$	$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	1	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{4}$	$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$

Abbildung 6. Koeffizienten für die Standardbasis und die hierarchische Basis, 2D Fall, k=3

nahe, dies auch in Diagonalrichtung zu tun: Läßt man nun alle Gitterpunkte weg deren hierarchischer Index $(\mathbf{l}, \mathbf{i})$ die Bedingung $|\mathbf{l}|_1 > k + d - 1$ erfüllt, so entsteht ein sogenanntes dünnes Gitter. Beispiele für dünne Gitter sind in Abbildung 7 gegeben.

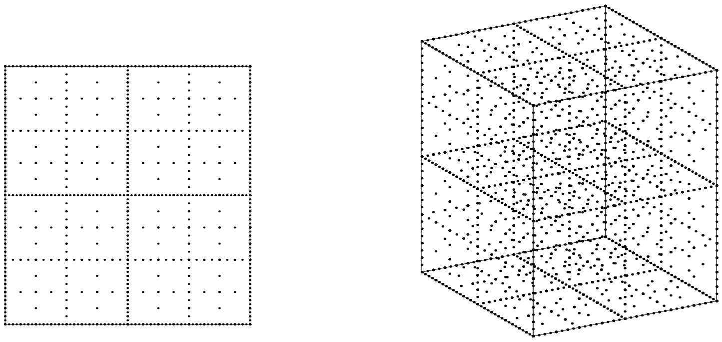

Abbildung 7. Zweidimensionales dünnes Gitter des Levels 6 (links) und dreidimensionales dünnes Gitter des Levels 5 (rechts)

Im Vergleich zur Funktionsdarstellung auf dem uniformen Gitter werden also viele Gitterpunkte weggelassen, die aber alle nur kleine hierarchische Koeffizienten tragen. Die Analyse der Zahl der in einem dünnen Gitter auftretenden Punkte ergibt direkt eine Komplexität der Ordnung $O(2^k k^{d-1})$. Dies ist eine substantielle Einsparung im Vergleich zum uniformen Gitter mit $O(2^{dk})$ Punkten. Es zeigt sich zudem, daß sich für Funktionen mit beschränkter $2d$-gemischter Ableitung der Fehler punktweise und für alle L_p-Normen nur unwesentlich von $O(2^{-2k})$ auf $O(2^{-2k} k^{d-1})$ erhöht. Bezüglich der Energienorm wird sogar die gleiche Komplexität nämlich $O(2^{-k})$ wie auf dem uniformen Gitter erzielt. Der sogenannte Fluch der Dimension, das heißt die exponentielle Abhängigkeit der Kosten, um eine vorgegeben Genauigkeit zu erreichen, ist bei genügend glatten Funktionen also überwunden worden. Die Dimensionsabhängigkeit taucht nurmehr im Exponenten der Levelzahl k auf. Dünne Gitter sind somit sehr gut für höherdimensionale Probleme geeignet.

Die Anwendung auf Differentialgleichungen kann nun dadurch geschehen, daß die hierarchischen Basisfunktionen des dünnen Gitters bei der Galerkindiskretisierung verwendet werden. Mittels des Ceá-Lemmas und der Energienorm-Abschätzung des Interpolationsfehlers lassen sich dann direkt auch die Fehler der Dünngitter-Galerkindiskretisierung einer elliptischen Differentialgleichung zweiter Ordnung zu $O(2^{-k})$ abschätzen. Um unnötigen Aufwand zu vermeiden, sollte dabei die Steifigkeitsmatrix nicht explizit aufgestellt werden. Es ist möglich, die Anwendung der Steifigkeitsmatrix auf einen Dünn-Gitter-Vektor in einer Zahl von Operationen zu programmieren, die proportional zur Zahl der Dünn-Gitterpunkte ist. Bereits 1989 wurde in [42] auch schon das erste Mehrgitterverfahren für Dünn-Gitter-Diskretisierungen vorgeschlagen und untersucht.

Die Verallgemeinerung des Dünn-Gitter Konstruktionsprinzips auf den nicht-glatten Fall ist in natürlicher Weise möglich. Das Weglassen der Gitterpunkte mit hierarchischen Koeffizienten kleiner als eine vorgegebene Schranke ε ergibt nun ein adaptives dünnes Gitter. Der hierarchische Koeffizient kann somit gegebenenfalls mit geeigneten Gewichten versehen als Fehlerindikator in einem adaptiven Gesamtverfahren verwendet werden. Beispiele für adaptive dünne Gitter sind in Abbildung 8 gegeben.

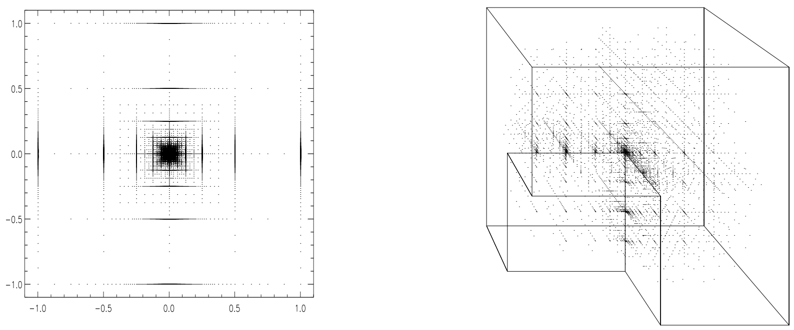

Abbildung 8. Adaptive dünne Gitter in zwei und drei Dimensionen

In den folgenden Jahren erfolgte die Ausarbeitung der Dünn-Gitter Galerkin-Diskretisierung in einer Reihe von Dissertationen: In [15] wurden dreidimensionale Poisson-Probleme und adaptive Verfeinerung studiert, in [3] geschah die Erweiterung auf den Helmholtz-Operator, in [30] gelang die Erweiterung auf Probleme zweiter Ordnung mit allgemeineren ortsabhängigen Koeffizientenfunktionen, in [84] wurde die Anwendung auf die biharmonische Gleichung studiert, weitere Konvergenzuntersuchungen und Mehrgitterverfahren erfolgten in [71], in [70] wurde die Parallelisierung von Dünn-Gitter-Algorithmen vorangetrieben und in der Habilitation [14] wurden schließlich dünne Gitter mit hierarchischer Basis höherer Ordnung entwickelt. Weitere Arbeiten zu additiven [49] und multiplikativen [53] Mehrgittermethoden rundeten das Bild ab. Bis heute wird am Lehrstuhl Zenger intensiv an der Weiterentwicklung der dünnen Gittern gearbeitet [79].

Neben der Diskretisierung mittels des Galerkin-Verfahrens wurde in der Arbeit *A combination technique for the solution of sparse grid problems* [51] im Jahr 1991 ein weiteres Verfahren zur Lösung von PDEs über dünnen Gittern gefunden, die sogenannte Kombinationstechnik. Diese Methode vermeidet gewisse Schwierigkeiten bei der Galerkin-Diskretisierung und erlaubt es, konventionelle Löser auf uniformen Gittern zum Teil wiederzuverwenden.

Bei der Dünn-Gitter-Kombinationstechnik gehen wir wie folgt vor: Wir diskretisieren und lösen das Problem auf einer bestimmten Folge von Gittern $\Omega_{\mathbf{l}}$ mit uniformen Maschenweiten $h_j = 2^{-l_j}$ für die j-te Koordinatenrichtung. Diese Gitter können verschiedene Maschenweiten für verschiedene Koordinatenrichtungen besitzen. Hierzu betrachten wir $\{\Omega_{\mathbf{l}}\}_{\mathbf{l}}$ mit

$$l_1 + ... + l_d = k + (d-1) - q, \quad q = 0,..,d-1, \quad l_j > 0. \tag{3}$$

Der Finite Elemente-Ansatz mit stückweisen d-linearen Ansatz- und Testfunktionen $\phi_{\mathbf{l},\mathbf{j}}(\boldsymbol{x})$ auf dem Gitter $\Omega_{\mathbf{l}}$ führt nun zur Diskretisierung

$$L_{\mathbf{l}} u_{\mathbf{l}} = f_{\mathbf{l}}$$

des kontinuierlichen Problems $Lu = f$ mit zugehörigen Randbedingungen.

Diese Probleme lösen wir dann mit einer geeigneten Methode, beispielsweise einem Mehrgitterlöser mit teilweiser Semi-Vergröberung. Aber auch jeder andere Löser läßt sich hier als Black-Box gegebenenfalls auf Kosten der Effizienz einsetzen. Die diskreten Lösungen $u_{\mathbf{l}}$ sind in den Räumen $V_{\mathbf{l}}$ der stückweise d-linearen Funktionen auf dem Gitter $\Omega_{\mathbf{l}}$ enthalten. Man beachte, daß all diese Probleme im Vergleich zum Problem auf dem uniformen Gitter Ω_k nun substantiell in ihre Größe reduziert sind. An Stelle nur eines Problems der Größe $O(2^{kd})$ müssen wir nun $O(dk^{d-1})$ Probleme der Größe $O(2^k)$ behandeln. Weiterhin können all diese Probleme unabhängig voneinander gelöst werden, was eine einfache Parallelisierung ermöglicht, siehe [45]. Hierzu existiert auch eine simple aber effektive statische Lastbalanzierungsstrategie [46].

Schließlich kombinieren wir die Ergebnisse $u_{\mathbf{l}}(\boldsymbol{x}) = \sum_{\mathbf{j}} u_{\mathbf{l},\mathbf{j}} \phi_{\mathbf{l},\mathbf{j}}(\boldsymbol{x}) \in V_{\mathbf{l}}$ der verschiedenen Gitter $\Omega_{\mathbf{l}}$ wie folgt:

$$u_k^{(c)}(\boldsymbol{x}) := \sum_{q=0}^{d-1} (-1)^q \binom{d-1}{q} \sum_{l_1+...+l_d=k+(d-1)-q} u_{\mathbf{l}}(\boldsymbol{x}). \tag{4}$$

Die resultierende Funktion $u_k^{(c)}$ lebt im Raum der Dünngitterfunktionen. Abbildung 9 zeigt die Gitter, die in der Kombinationsformel des Levels 4 im zweidimensionalen Fall benötigt werden.

Die Kombinationstechnik läßt sich als eine bestimmte multivariate Extrapolationsmethode interpretieren, die auf dem dünnen Gitter arbeitet, für Details siehe [21,51,72]. Die Kombinationslösung $u_k^{(c)}$ ist im allgemeinen nicht gleich der Galerkin-Lösung $u_k^{(s)}$, jedoch ist ihre Genauigkeit üblicherweise von der gleichen Ordnung, siehe [51]. Hierzu ist eine Reihenentwicklung des Fehlers notwendig, ihre Existenz wurde für PDE-Modellprobleme in [20] gezeigt. In [59] wurde die Kombinationstechnik auf turbulente Strömungsprobleme angewandt. Auch die Kombinationstechnik wurde am Lehrstuhl Zenger intensiv weiterentwickelt und in einer Reihe von FORTWIHR Projekten eingesetzt.

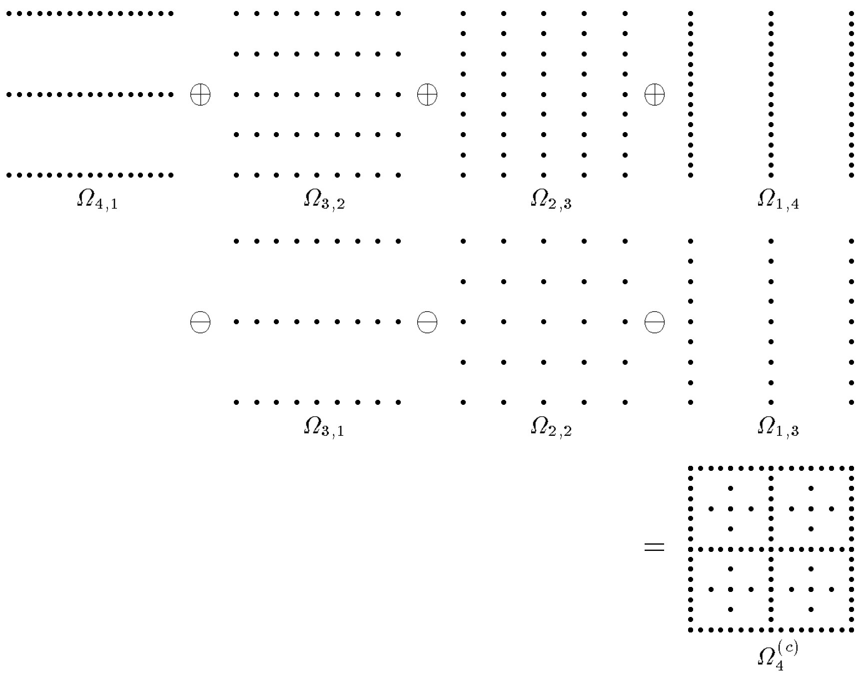

Abbildung 9. Kombinationstechnik mit $k = 4$, $d = 2$

2 Wirken

Schon durch die Beschränkung meiner Redezeit ist es mir nicht möglich, alle Aspekte und Auswirkungen des wissenschaftlichen Werks von Christoph Zenger im Detail vorzustellen. Zudem kann ich das auch nicht in allen Bereichen, da ich mich selbst bei bester Vorbereitung schlicht nicht genügend auskenne, um etwa die Verwendung des Theorems von Zenger, die Deutsch-Zenger Schranke oder die Zenger-Korrekturen in modernen interdisziplinären Aufgabenstellungen oder sogar Killer-Applikationen, wie dies ja Neudeutsch heute heißen muß, genügend würdigen zu können.

Auf jeden Fall konnte ich folgendes herausfinden: Numerische Wertevorräte werden immer noch intensiv studiert. Mittlerweile gibt es eine neue Serie von Workshops (1991 Minneapolis USA, 1992 Williamsburg USA, 1994 Coimbra Portugal, 1996 Sapporo Japan, 1998 Madison USA, 2000 Griechenland) zum Thema „numerical ranges and numerical radii“ mit jeweils bis zu 40 Teilnehmern. Jüngere Ergebnisse fanden ihren Niederschlag 1997 im Buch von Gustafson und Rao [54].

Insbesondere stochastische Matrizen, ihre Radii und Ergodenkoeffizienten und verwandte Fragestellungen spielen heute in vielen praktischen Anwendungsbereichen eine zentrale Rolle. Ein Beispiel ist die Analyse von Netzwerken (Kabel oder Telefonnetzwerke) im Zusammenhang mit Störungen. Der Ergodizitätskoeffizient und der zugehörige Eigenvektor liefern hier Aussagen über die Schwachstellen des Graphen. Auch bei der Partitionierung von Finite Elemente Netzen für das Parallelrechnen spielt der Ergodizitätskoeffizient eine Rolle. Ist L die Laplacematrix/Adjazenzmatrix des Graphen der diskretisierten Gleichung, dann geben der Ergodizitätskoeffizient von I-L und der zugehörige Eigenvektor eine sehr gute Approximation für die optimale Aufteilung des Problems auf verschiedene Rechnerknoten (Stichwort: spektrale Bisektion). Zudem bewerten moderne Internet-Suchmaschinen wie etwa Google Webseiten gemäß ihrer Konnektivität. Mit anderen Worten, alle indizierten Webseiten spannen einen Graphen auf, wobei die Seitenbezeicher die Knoten sind und die Links zwischen Webseiten die Kanten sind. Indem man die Zahl der Links aufsummiert, entsteht eine stochastische Matrix. Es gilt nun, die wichtigsten Seiten zu identifizieren. Auch hier können der Ergodizitätskoeffizient und der numerische Radius sowie Schranken dafür eine große Hilfe sein. Schließlich sei noch aus dem Bereich Bioinformatik, in dem eine Vielzahl von stochastischen Matrizen vorkommen, auf die jüngsten Anstrengungen in Berlin am ZIB bei P. Deuflhard hingewiesen. Auch hier treten bei der Identifizierung metastabiler Konformationen von Biomolekülen bestimmte Markovketten mit stochastischen Transitionsmatrizen auf [80]. Deren Clusteranalyse erlaubt eine substantielle Reduktion der Dimensionalität des Problems.

Ich möchte mich aber im folgenden auf die Technik der dünnen Gitter konzentrieren, deren Entwicklung und Anwendung auf partielle Differentialgleichungen ich in meiner Münchner Zeit mit begleiten durfte.

Mittlerweile ist andernorts die Dünn-Gitter-Idee aufgegriffen worden und wird in verschiedenen Projekten verfolgt. Hier ist zuerst das CWI in Amsterdam zu nennen. Von P. Hemker und B. Koren werden dabei Konvektions-Diffusionsprobleme und die Eulergleichungen mittels einer Variante der Kombinationstechnik behandelt. Bei der GMD wurde in der Gruppe Trottenberg die Kombinationstechnik zur parallelen Lösung von Problemen aus der Meteorologie eingesetzt und in Nürnberg-Erlangen gibt es in der Gruppe Durst diverse Ansätze, um dünne Gitter bei der numerischen Strömungssimulation (Turbulenz, CVD) einzusetzen. Zunächst auch in Erlangen, nun in Stuttgart, werden in der Gruppe Ertl dünne Gitter bei der schnellen Visualisierung und graphischen Darstellung großer Datenmengen erfolgreich eingesetzt. Weiterhin werden in Tübingen in der Nachwuchsgruppe des dortigen SFBs von R. Hiptmair Dünn-Gitter-Techniken auf ihren Einsatz bei Magnetfeldproblemen untersucht. In Bonn wurden Finite Differenzen [43,78] über dünnen Gittern entwickelt, es gelang eine Erweiterung auf allgemeinere Dünn-Gitter-Räume und Sobolevklassen [62,47], auf Integralgleichungen [50] und erste Integrodifferentialgleichungen [57] sowie Wavelets und Interpolets [48,64]. In der Community der „information-based complexity“ wird das Dünn-Gitter-Prinzip/die Smolyak-Konstruktion intensiv untersucht und verwendet, um obere Schranken für die Komplexität diverser Probleme herleiten zu können [35,68,69,90].

In Bonn haben wir eine Dünn-Gitter Bibliothek aufgebaut, auf die auch von auswärts zugegriffen werden kann[3]. Sie enthält mittlerweile mehr als 240 Einträge. Interessant ist dabei das Histogramm der Arbeiten und Artikel über die Jahre. Dies ist in Abbildung 10 dargestellt. Wir sehen deutlich, daß sich die dünnen Gitter auf exponentiellem Wachstumskurs befinden und noch keine Wachstumssättigung abzusehen ist. Sicherlich werden in naher Zukunft noch weitere Arbeitsgruppen infiziert werden.

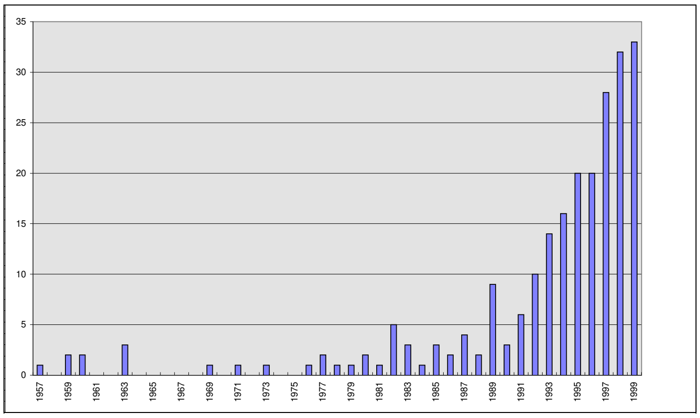

Abbildung 10. Zeitliche Entwicklung von Dünn-Gitter Publikationen

Ob der Ausbruch der dünnen Gitter aus dem akademischen Elfenbeinturm gelingt und der Quantensprung (sic, es ist wirklich keine Entfernung) in den vielbesungenen Mittelstand und die Industrie in naher Zukunft bevorsteht, ist noch unklar. Zwar ist die Gründung der Sparsytec AG immer noch nicht erfolgt (warum eigentlich nicht, Hans?) und aus diesem Grund werden dünne Gitter noch nicht am Neuen Markt oder der NASDAQ gehandelt, aber erste Patentanmeldungen lassen hoffen.

Im folgenden möchte ich nun erste neue Anwendungen des Dünn-Gitter-Prinzips in modernen Disziplinen wie der Finanzmathematik, dem Data-Mining, den Materialwissenschaften und den Biowissenschaften zeigen.

[3] http://bibliothek.iam.uni-bonn.de/duennbib.html

2.1 Option-Pricing

Die Bewertung von Finanzderivaten hat im letzten Jahrzehnt einen starken Aufschwung erfahren und ist zu einem wichtigen Instrument in der Finanzwirtschaft geworden. Man kann nun spätestens seit der Beinahe-Pleite des Hedge Fonds Long-Term Capital Management (LTCM), an dem mithin die Nobel-Laureaten Myron Scholes und Robert Merton stark mitgewirkt hatten, solchen Entwicklungen kritisch gegenüberstehen. Einer Schätzung zufolge hatte LTCM Kontrakte im Gesamtwert von einer Billion (1.000 Milliarden!) Dollar abgeschlossen, bei denen es unter anderem um Derivate, Swaps und andere komplexe Finanzierungsinstrumente ging. Durch dessen Zahlungsunfähgkeit wurde das amerikanische Bankensystem fast zum Kollaps gebracht und erst durch die vereinten Anstrengungen staatlicher und internationaler Banken konnte der Konkurs abgewendet werden. Trotzdem ist die Bewertung von Finanzderivaten sinnvoll. Einerseits ist die Möglichkeit einer Bewertung eine notwendige Voraussetzung, damit ein Finanzinstitut überhaupt ein neues Derivat anbieten kann, andererseits ist die Bewertung nötig, um ein Portfolio geeignet zu strukturieren, abzusichern und zu optimieren.

Es gibt nun eine Reihe verschiedener Arten von Finanzderivaten, die beispielsweise Zins- oder Aktien-basiert sind. Hierzu zählen u.a. Bonds, Swaps, Futures, CMOs und Optionen. Ihre Bewertung erfolgt, unter der Annahme, daß keine Arbitragemöglichkeit besteht, mittels partieller Differentialgleichungen oder über den Martingalansatz. Letzterer ist allgemeiner. Dabei spezifiziert man zunächst einen stochastischen Prozeß für den zugrundeliegenden Wert (zum Beispiel den Zinssatz oder den Aktienkurs). Dann bestimmt man das äquivalente Martingalmaß, das den zugrundeliegenden Prozeß in ein Martingal umwandelt und berechnet den Wert des Derivats als den Erwartungswert seiner discounteten Auszahlungsfunktion unter diesem risikoneutralen Maß. Unter bestimmten Voraussetzungen ist der Zugang über eine partielle Differentialgleichung und das Martingal gleichwertig. Der Zusammenhang wird dabei über verallgemeinerte Feynman-Kacz Formeln hergestellt [55]. Der Martingalansatz ist jedoch universeller einsetzbar und leichter auf neue Situationen übertragbar.

Der Preis eines Finanzderivats läßt sich somit als Erwartungswert ausdrücken. In kontiunierlicher Zeit beinhaltet der Integrand dabei wiederum ein Pfadintegral, das etwa mit dem Euler-Verfahren/Trapezregel diskretisiert werden kann oder es wird gleich bedingt durch die Anwendung ein zeitdiskretes Modell verwendet. In beiden Fällen führt dies zu einem hochdimensionalen Integrationsproblem. Für die einfache Europäische Call Option ist es möglich, eine geschlossene Lösung anzugeben (Black-Scholes Formel), kompliziertere Optionstypen benötigen jedoch ein numerisches Lösungsverfahren. Dies gilt analog auch für die anderen Arten von Finanzderivaten.

Welche Integrationsverfahren für hochdimensionale Integranden haben wir nun zur Verfügung? Sicherlich ist die klassische multivariate Quadratur nicht geeignet. Hier begegnet man dem Fluch der Dimensionalität: Der

Aufwand skaliert hier exponentiell mit der Dimension. Man hat die Komplexität der Ordnung $O(N^{-r/d})$ wobei r die Glattheit des Integranden und d seine Dimension bezeichnen. Hingegen ist das Monte Carlo Verfahren von der Dimension unabhängig. Dabei wird der Integrand u an einer zufälligen Folge von N Punkten $\mathbf{x}_i$ ausgewertet, und man erhält die Quadraturformel

$$I_N u = \frac{1}{N} \sum_{i=1}^{N} u(\mathbf{x}_i).$$

Das Monte Carlo Verfahren konvergiert jedoch nur sehr langsam (und auch nur im statistischen Sinne). So ist die Genauigkeit, die mit N Funktionsauswertungen erzielt werden kann, von der Ordnung $O(1/\sqrt{N})$. Von besonderem Interesse sind deswegen die im letzten Jahrzehnt entwickelten Quasi Monte Carlo Methoden. Hier wird der Integrand für eine deterministisch bestimmte Folge von Punkten $\mathbf{x}_i$ ausgewertet. Es gibt hierbei verschiedene Konstruktionen wie die Halton-, Sobol- oder Faure-Folgen, die sich in ihrer Präasymptotik unterscheiden, aber alle in einer Konvergenzordnung des Typs $O((\log N)^d/N)$ resultieren. Diese ist also fast eine Ordnung besser als für die Monte Carlo Methode. Zudem ist der Fehler deterministisch. Prototyp ist hierbei das Programm FinDer von J. Traub [87], das mittlerweile in vielen Banken verwendet wird. In der Finanzsoftware von IBM soll es ein analoges Produkt geben.

Alternativ kann man aber auch dünne Gitter einsetzen. Dünn-Gitter Integrationsformeln erhält man direkt, indem man den Integranden u mittels eines Dünn-Gitter-Interpolanden ((2) mit $|\mathbf{l}|_1 \leq k + d - 1$) approximiert, Summation und Integration vertauscht und die Dünn-Gitter-Basisfunktionen einzeln integriert. Man erhält Quadratur-Formeln des Typs

$$I_k^{(s)} u = \sum_{|\mathbf{l}|_1 \leq k+d-1} \sum_{\mathbf{i}} \gamma_{\mathbf{l},\mathbf{i}} u_{\mathbf{l},\mathbf{i}},$$

wobei die Gewichte $\gamma_{\mathbf{l},\mathbf{i}}$ nun gerade die Integrale der hierarchischen Basisfunktionen und $u_{\mathbf{l},\mathbf{i}}$ die hierarchischen Koeffizienten sind. Interessanterweise ist dies genau die Technik von Smolyak [82]. Hier wird aus einer eindimensionalen Quadraturformel mittels Differenzbildung, Tensorproduktbildung sowie Abschneidens wie beim dünnen Gitter eine multivariate Quadraturformel erzeugt. Die linearen hierarchischen Basen und die Dünn-Gitter-Konstruktion übersetzen sich in die Smolyak-Konstruktion, welche auf der eindimensionalen Trapezregel basiert. Alternativ lassen sich eine Reihe anderer Regeln verwenden. Ein Beispiel sind die Formeln von Clenshaw-Curtis [67]. In [10] wurde eine adaptive Dünn-Gitter-Quadraturformel genauer untersucht und auf Verfahren höherer Ordnung verallgemeinert.

Eine zur Dünn-Gitter-Quadratur analoge Konstruktion ist mittels der Kombinationstechnik möglich. Man integriere dazu Formel (4) und vertausche Summation und Integration. Die Funktionen $u_{\mathbf{l}}$ sind dabei der Interpoland von u auf dem Gitter $\Omega_{\mathbf{l}}$ mit zugehörigem Raum $V_{\mathbf{l}}$.

In [38] haben wir eine noch bessere Technik entwickelt: Wir verwenden in der Smolyak-Konstruktion als 1D Formel die Gauß–Kronrod Formel sowie ihre hierarchischen Erweiterungen nach Patterson. Die Kronrod Formel ergänzt dabei die Gauß Formel hierarchisch auf dem nächstfeineren Gitter so, daß der Exaktheitsgrad maximal wird. Patterson verwendete diese Erweiterungsidee sukzessive auf eine ganze Folge von feiner werdenden geschachtelten Gittern. Für glatte Integraden ergibt sich hier exponentielle Konvergenz und asymptotisch ist dieses Verfahren optimal. Abbildung 11 gibt Beispiele verschiedener dünner Gitter im 2D Fall.

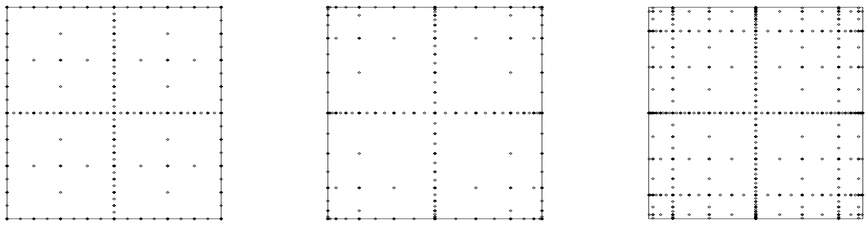

Abbildung 11. Dünne Gitter für die Trapez-, die Clenshaw–Curtis- und die Gauß–Patterson Regeln für $d = 2, l = 6$

Es zeigt sich, daß dieses neue Verfahren in einer Vielzahl von hochdimensionalen Anwendungen den Quasi Monte Carlo Methoden überlegen ist. Diese können höhere Glattheit des Integranden nicht ausnutzen, die neue Gauß-Patterson Dünn-Gitter-Quadratur kann dies hingegen in optimaler Weise. In [38] wird diese Technik genauer mit den anderen verglichen und auf eine Reihe von Problemen (Pfadintegral, Absorptionsproblem, Integralgleichung) angewandt. Weiterhin lösen wir dort mit der Braunschen Brücke ein 360-dimensionales Integrationsproblem für eine collaterated mortgage obligation (CMO). Diese Methoden wurde mittlerweile in [83] auch mit Erfolg bei der Bewertung einer Reihe von Zinsderivaten (Zerobonds, Swaps, MBS für die Modelle von Hall-White und Black-Karasinski) mit glatten Integranden eingesetzt[4].

Ziel ist es nun, das Gauß–Patterson Dünn-Gitter Verfahren auch bei der Bewertung von Optionen einzusetzen. Eine Option ist ein Vertrag zwischen zwei Parteien, der dem Halter das Recht aber nicht die Pflicht gibt, eine festgesetzte Menge von z.B. Aktien zu einem bestimmten Zeitpunkt zu einem vorher festgelegten Preis zu kaufen/verkaufen. Eine Option stellt damit

[4] Da man sich oft im präasymptotischen Bereich befindet, können suboptimale Verfahren hier noch eine Weile mithalten. Insbesondere für hohe Dimensionen und niedrige Genauigkeitsanforderungen kann so der klare Vorsprung unseres Verfahrens geringer werden.

einen intrinsischen Wert dar, da der Halter die Option nicht ausüben muß. Die Frage ist nun, wie der Preis für eine Option fair bewertet werden kann. Konkret haben wir hier mit dem Martingalansatz die Bewertung

$$V = e^{-rT} E(P(\{S_i^j\})).$$

Dabei bezeichnet r die konstante Zinsrate, $t = T$ den Zeitpunkt der Optionsausübung, P die Auszahlungsfunktion, S_i^j den Aktienwert der i-ten Aktie des Portfolios zur Zeit $j\Delta t$, mit $\Delta t = T/M$. Hier ist M die Zahl der diskreten Zeitpunkte $j = 1, \ldots, M$, und N bezeichnet die Zahl der betrachteten Aktien $i = 1, \ldots, N$. e^{-rT} ist der Discountfaktor, und $t = 0$ ist der Zeitpunkt der Optionsbewertung, also jetzt.

Wir nehmen nun an, daß die Aktien des Portfolios einem Modell folgen, das durch die stochastische Differentialgleichung

$$dS_i = \mu_i S_i dt + \sum_{k=1}^{N} \sigma_{ik} S_k dW_k \tag{5}$$

beschrieben ist, wobei (σ_{ik}) die zeitkonstante Volatilitätsmatrix bezeichnet, μ_i die zeitkonstante Drift für die Aktie i bezeichnet[5] und dW_i ein geometrischer Brownscher Prozeß für die Aktie i ist, d.h. W_i ist das Wiener Maß. Integration und Itô-Formel ergibt nach Diskretisierung in der Zeit mit den Zeitpunkten $j\delta t$ die Werte S_i^j

$$S_i^{j+1} = S_i^j \cdot e^{(\mu_i - \frac{1}{2}(\sum_{k=1}^N \sigma_{ik}^2))\Delta t + \sum_{k=1}^N \sigma_{ik}\sqrt{\Delta t} W_k^j},$$

und somit gilt[6]

$$S_i^j = S_i^0 e^{\sum_{l=1}^j \left[(\mu_i - \frac{1}{2}(\sum_{k=1}^N \sigma_{ik}^2))\Delta t + \sum_{k=1}^N \sigma_{ik}\sqrt{\Delta t} W_k^l\right]}.$$

Hier sind W_k^l Zufallsvariable, die $N(0,1)$-verteilt sind.

Die Definition des Erwartungswerts $E(.)$ ist nun für eine allgemeine Funktion f

$$E(f(\{x_i^j\})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\{x_i^j\})\ dg(x_1^1) \ldots dg(x_N^M)$$

mit der Standard-Normalverteilung

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Mit der Forderung, daß keine Arbitrage existiert, muß der Erwartungswert bezüglich des zum unterliegenden Prozeß äquivalenten Martingalmaßes

[5] Es gibt auch Ansätze mit zeitabhängigen Volatilitätsmatrizen und zeitabhängiger Drift. Dann folgen σ_{ik} und μ_i eigenen SDEs, die ihr Verhalten modellieren.

[6] Im Computerprogramm wird die erste rekursive Formel genutzt. Die jetzt folgende nicht-rekursive aufsummierte Form ist aber notwendig, um den späteren Integranden definieren zu können.

genommen werden. Es werden dabei in unserem Fall im Integral des Erwartungswerts die μ_i durch r ersetzt. Damit findet auch der Wechsel der SDE (5) in die risikoneutrale Form statt.

Was nun noch fehlt sind die konkreten Auszahlungsfunktionen. Diese sind abhängig von jeweiligen Typ der Option. Generell haben sie die Struktur

$$P(\{S_i^j\}) = \max(0, H(\{S_i^j\}) - K) \tag{6}$$

für Call Optionen und

$$P(\{S_i^j\}) = \max(0, K - H(\{S_i^j\}))$$

für Put Optionen. Dabei ist K der Ausübungspreis. Beispiele sind pfadabhängige Optionen ($N = 1, M > 1$) und performanceabhängige Optionen ($N > 1, M = 1$).

Populärstes Beispiel für pfadabhängige Optionen sind sogenannte Asiatische Optionen. Hierbei wird für die Auszahlungsfunktion der Mittelwert über die Aktienkurse zu allen Zeitpunkten zwischen dem momentanen Zeitpunkt und dem Ausübungszeitpunkt genommen.

Wir betrachten im folgenden den Fall des geometrischen Mittels, d.h.

$$H(\{S_i^j\}) = (\prod_{k=1}^{M} S_i^k)^{1/M}. \tag{7}$$

Hierzu existiert eine geschlossene Lösung in Form einer verallgemeinerte Black-Scholes Formel. Kleine Änderungen und Variationen (z.B. schon das arithmetische Mittel) sind nicht mehr ohne weiteres analytisch lösbar und benötigen ein numerisches Integrationsverfahren.

Ein Beispiel für performanceabhängige Optionen sind sogenannte „Tailored Options“ [63]. Hier hängt der Preis der Option von der relativen Performance eines Aktienkurses in bezug auf die anderen Aktienkurse einer betrachteten Menge von Aktien ab. Ein Beispiel für eine Call Option ist

$$P(\{S_i^j\}) = \begin{cases} \max(S_1^1 - K, 0) & \text{falls } S_1^1/S_1^0 > S_i^1/S_i^0, 1 < i \leq N \\ a \cdot \max(S_1^1 - K, 0) & \text{sonst} \end{cases} \tag{8}$$

wobei a ein Parameter aus $[0, 1)$ ist. Eine Mischung aus beiden Optionstypen ist selbstverständlich auch möglich. Mit $M = N = 1$ ergibt sich als Spezialfall die Europäische Call Option mit der Black Scholes Formel als analytischer Lösung.

Nun ist also zur Optionsbewertung ein im allgemeinen hochdimensionales Integrationsproblem zu lösen. Die Dimensionen rühren dabei von den Zeitpunkten her, über die bei pfadabhängigen Call Optionen gemittelt wird, oder/und von der Zahl der betrachteten Aktien bei performanceabhängigen Optionen.

Auf Grund der Definition des Erwartungswerts ist das Integrationsgebiet bisher noch $(-\infty, \infty)^{N \cdot M}$. Um unsere Dünn-Gitter Integrationsmethode anwenden zu können, transformieren wir das Integral des Erwartungswerts mit

Hilfe der kumulativen Normalverteilung $G(y) := \int_{-\infty}^{y} g(x)dx$ auf $[0,1]^{N\cdot M}$ und erhalten damit

$$E(f(\{x_i^j\})) = \int_0^1 \dots \int_0^1 f(\{G(x_i^j)\})\; dx_1^1 \dots dx_M^N. \tag{9}$$

Ziel ist, die Gauß-Patterson Dünn-Gitter-Technik nun für die Integration von (9) mit (6) und (7) bzw. (8) gewinnbringend einzusetzen. Für Optionen ergibt sich nun jedoch folgendes Problem: Die Auszahlungsfunktion ist durch die Natur einer Option nicht mehr glatt. Dies drückt letztendlich aus, daß man eine Option ja nicht ausüben wird, wenn der Kauf oder Verkauf der abgesicherten Aktie einen Verlust ergeben würde. Der Integrand weist bzgl. einer $(M\cdot N-1)-$dimensionalen Mannigfaltigkeit einen Knick (pfadabhängige Optionen) oder sogar Sprünge (performanceabhängige Optionen) auf[7].

Beispiele für den Integranden im zweidimensionalen Fall sind in Abbildung 12 gegeben. Deutlich sieht man den Knick bzw. den Sprung des Integranden. Die Glattheitsvoraussetzungen für die Anwendung unseres Dünn-Gitter-Verfahrens sind damit nicht mehr gegeben.

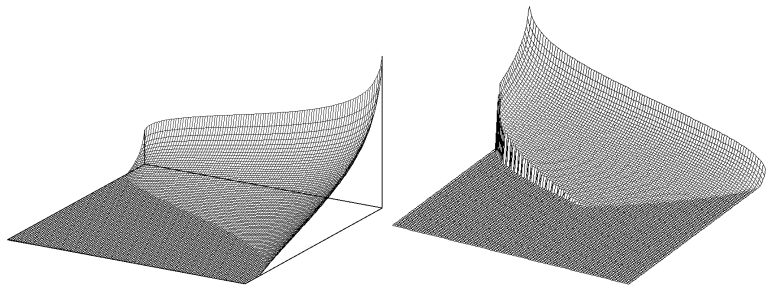

Abbildung 12. Integrand im 2D Fall: Links pfadabhängige Option mit zwei Zeitpunkten, $M = 2, N = 1, S_1^0 = 100, K = 100, T = 1, \sigma_{11} = 0.2, r = 0.05$, rechts performanceabhängige Option mit zwei Aktien, $M = 1, N = 2, S_1^0 = S_2^0 = 100, K = 100, T = 1, \sigma_{11} = 0.2, \sigma_{12} = 0.25, \sigma_{21} = -0.1, \sigma_{22} = 0.25, r = 0.05, a = 0$

Die Idee ist nun, nur über den Träger des Integranden zu integrieren. Hier ist der Integrand eine glatte Funktion, die Knicke und Sprünge befinden sich gerade am Rand des Trägers. Um dieses Gebiet zu bestimmen genügt es, die Nullstellen des Integranden zu berechnen. Wenn man das Integral iteriert berechnet, reduziert sich die Nullstellensuche auf eine einzige (die letzte) Dimension. Wir rechnen deswegen in der letzten Dimension die Nullstelle $\hat{x}$ aus (Newtonverfahren für den Knick oder Bisektionverfahren für den Sprung)

[7] Neben den Singularitäten am Rand des Integrationsgebiets, die durch die Transformation auf $[0,1]^d$ entstanden sind.

und transformieren den Integranden bezüglich der letzten Dimension mit der linearen Abbildung

$$t(x) = x \cdot (1 - \hat{x}) + \hat{x}$$

auf [0,1]. Nun können wir unser Gauß-Patterson Verfahren einsetzen. Implizit ergeben sich dünne Gitter wie in Abbildung 13.

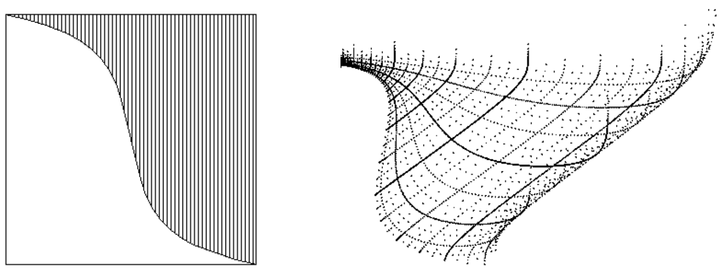

Abbildung 13. Dünnes Gitter auf dem Träger des Integranden

Abbildung 14 gibt einen Vergleich der verschiedenen Integrationsverfahren bei der Bewertung einer pfadabhängigen Option mit 6 Zeitpunkten, $M = 6, N = 1$, und einer performanceabhängigen Option mit $M = 1, N = 2$. Die überlegene Konvergenzordnung unseres Dünn-Gitter Verfahrens mit Transformation und Gauß-Patterson Formel über die anderen Methoden (Monte Carlo (MC), Klassischer Produktansatz ohne (PR) und mit Transformation auf den Träger (PRTR) sowie Gauß-Patterson-Dünne-Gitter ohne (SG) und mit Transformation auf den Träger (SGTR)) ist deutlich zu sehen. Man beachte, daß der Fehler hier logarithmisch aufgetragen ist.

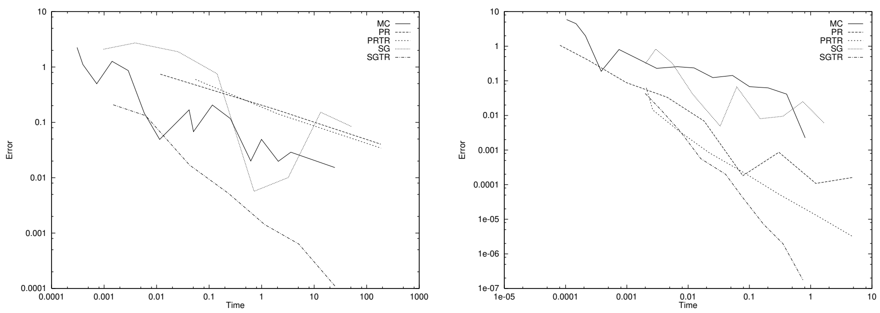

Abbildung 14. Vergleich verschiedener Integrationsverfahren für eine pfadabhängige Option in 6 Dimensionen (links) und eine performanceabhängige Option in zwei Dimensionen

2.2 Data-Mining

Data Mining ist ein Prozeß, um verborgene Muster, Zusammenhänge und Trends in großen Datenmengen aufzuspüren. Dies spielt eine immer wichtigere Rolle in der Wirtschaft und in den Wissenschaften. Typische wissenschaftliche Anwendungen sind das Nachbereiten von Daten in der Medizin (z.B. CT Daten), die Auswertung von Daten in der Astrophysik (z.B. Teleskop und Observatoriumsdaten), das Gruppieren seismischer Daten oder die Auswertung von Satellitenbildern (z.B. beim NASA Erdbeobachtungssystem). Aber vielleicht sogar wichtiger sind finanzielle und kommerzielle Anwendungen. Mit der Entwicklung des Internets und des e-commerce werden nun riesige Datenmengen gesammelt, und zwar mehr oder weniger automatisch, die für Geschäftsentscheidungen und weitere strategische Planungen benutzt werden können. Die Anwendungen reichen von der Kündigung und Stornierung von Verträgen, der Beurteilung von Kreditrisiken, der Segmentierung von Kunden für Raten, der Tarifplanung, der Aufdeckung von Betrugsversuchen und der Aktienanalyse bis zur Umsatzvorhersage.

Üblicherweise läßt sich der Prozeß des Data Mining trennen in einen Planungsschritt, die Vorbereitungsphase, die eigentliche Data Mining Phase, d.h. das maschinelle Lernen, und die Auswertung. Dabei werden u.a. Assoziationsanalyseklassifikation, Clustering und Prognose eingesetzt. Für einen Überblick über die verschiedenen Aufgaben, die im Data Mining Prozeß auftreten, siehe [8,23].

Das Klassifikationsproblem und seine Diskretisierung. Im folgenden betrachten wir detailliert das Klassifikationsproblem. Dabei ist eine Menge von Datenpunkten im d-dimensionalen Merkmalsraum zusammen mit einem Klassenlabel etwa aus $\{-1, 1\}$ gegeben. Aus diesen Daten muß ein Klassifikator konstruiert werden, der es erlaubt, für irgendeinen neu gegebenen Datenpunkt die Klasse zur weiteren Entscheidung vorherzusagen. Dies läßt sich als traditionelles Scattered Data Approximationsproblem mit bestimmten zusätzlichen Regularisierungstermen in hochdimensionalen Räumen interpretieren. Hier gibt uns der Zugang über Regularisierungsnetzwerke einen guten Rahmen [40]. Dieser Zugang erlaubt eine direkte Beschreibung der meisten wichtigen neuronalen Netzwerke, und er erlaubt auch eine äquivalente Beschreibung von Support-Vektor-Maschinen und n-term Approximationsschemata [39].

Konkret folgen wir [37] und betrachten eine gegebene Menge bereits klassifizierter Daten (die Trainingsmenge)

$$S = \{(\boldsymbol{x}_i, y_i) \in \mathbb{R}^d \times \mathbb{R}\}_{i=1}^M.$$

Wir nehmen nun an, daß diese Daten durch Abtasten einer unbekannten Funktion f erhalten wurden, die zu einem bestimmten über $\mathbb{R}^d$ definierten Funktionenraum V gehört. Der Samplingprozeß war dabei durch Rauschen gestört. Das Ziel ist nun, die Funktion f aus den gegebenen Daten so gut wie möglich zu rekonstruieren. Dies ist sicherlich ein schlecht gestelltes Problem, denn es gibt unendlich viele mögliche Lösungen. Um ein gut gestelltes eindeutig lösbares Problem zu erhalten, müssen wir mehr Kenntnisse über f haben. Die Regularisierungstheorie [86,89] fordert dabei zusätzliche Glattheitsbedingungen an die Lösung des Approximationsproblems. Der Zugang über ein Regularisierungsnetzwerk betrachtet das Variationsproblem

$$\min_{f \in V} R(f)$$

mit

$$R(f) = \frac{1}{M} \sum_{i=1}^{M} C(f(\boldsymbol{x}_i), y_i) + \lambda \Phi(f). \tag{10}$$

Dabei beschreibt $C(.,.)$ eine Fehlerkostenfunktion, die den Interpolationsfehler mißt, und $\Phi(f)$ ist ein Glattheitsfunktional, das für $f \in V$ wohldefiniert sein muß. Der erste Term erzwingt die Nähe von f zu den Daten, der zweite Term erzwingt die Glattheit von f und der Regularisierungsparameter λ balanciert diese zwei Terme. Wir betrachten dabei konkret den Fall

$$C(f_N(\boldsymbol{x}_i), y_i) = (f_N(\boldsymbol{x}_i) - y_i)^2$$

und

$$\Phi(f_N) = ||P f_N||_{L_2}^2 \tag{11}$$

für einen gegebenen linearen Operator$P = \nabla$.

Im folgenden beschränken wir das Problem nun explizit auf einen endlichdimensionalen Teilraum $V_N \in V$. Die Funktion f wird dann ersetzt durch

$$f_N = \sum_{j=1}^{N} \alpha_j \varphi_j(\boldsymbol{x}). \tag{12}$$

Dabei sollte $\{\varphi_j\}_{j=1}^N$ den Raum V_N aufspannen und möglichst auch eine Basis dafür sein. Die Koeffizienten $\{\alpha_j\}_{j=1}^N$ entsprechen den Freiheitsgraden. Man beachte, daß die Einschränkung auf einen geeignet gewählten endlichdimensionalen Teilraum bestimmte zusätzliche Regularisierung mit sich bringt (Regularisierung durch Diskretisierung), die von der Wahl von V_N abhängt.

Damit erhalten wir aus dem Minimierungsproblem ein lösbares lineares Gleichungssystem. Zu minimieren ist also

$$R(f_N) = \frac{1}{M} \sum_{i=1}^{M} (f_N(\boldsymbol{x}_i) - y_i)^2 + \lambda ||P f_N||_{L_2}^2, \quad f_N \in V_N \tag{13}$$

im endlich dimensionalen Raum V_N. Wir setzen (12) in (13) ein und erhalten

$$R(f_N) = \frac{1}{M}\sum_{i=1}^{M}\left(\sum_{j=1}^{N}\alpha_j\varphi_j(\boldsymbol{x}_i) - y_i\right)^2 + \lambda\|P\sum_{j=1}^{N}\alpha_j\varphi_j\|_{L_2}^2 \tag{14}$$

$$= \frac{1}{M}\sum_{i=1}^{M}\left(\sum_{j=1}^{N}\alpha_j\varphi_j(\boldsymbol{x}_i) - y_i\right)^2 + \lambda\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j(P\varphi_i, P\varphi_j)_{L_2}. \tag{15}$$

Differentiation nach α_k, $k = 1, \ldots, N$, ergibt

$$0 = \frac{\partial R(f_N)}{\partial \alpha_k} = \frac{2}{M}\sum_{i=1}^{M}\left(\sum_{j=1}^{N}\alpha_j\varphi_j(\boldsymbol{x}_i) - y_i\right)\cdot\varphi_k(\boldsymbol{x}_i) + 2\lambda\sum_{j=1}^{N}\alpha_j(P\varphi_j, P\varphi_k)_{L_2}. \tag{16}$$

Dies ist äquivalent zu $(k = 1, \ldots, N)$

$$\lambda\sum_{j=1}^{N}\alpha_j(P\varphi_j, P\varphi_k)_{L_2} + \frac{1}{M}\sum_{j=1}^{N}\alpha_j\sum_{i=1}^{M}\varphi_j(\boldsymbol{x}_i)\cdot\varphi_k(\boldsymbol{x}_i) = \frac{1}{M}\sum_{i=1}^{M}y_i\varphi_k(\boldsymbol{x}_i) \tag{17}$$

und wir erhalten $(k = 1, \ldots, N)$

$$\sum_{j=1}^{N}\alpha_j\left[M\lambda(P\varphi_j, P\varphi_k)_{L_2} + \sum_{i=1}^{M}\varphi_j(\boldsymbol{x}_i)\cdot\varphi_k(\boldsymbol{x}_i)\right] = \sum_{i=1}^{M}y_i\varphi_k(\boldsymbol{x}_i). \tag{18}$$

In Matrix-Notation erhalten wir schließlich das lineare System

$$(\lambda C + B \cdot B^T)\alpha = By. \tag{19}$$

Hier ist C eine quadratische $N \times N$ Matrix mit Einträgen $C_{j,k} = M \cdot (P\varphi_j, P\varphi_k)_{L_2}$, $j,k = 1, \ldots N$, und B ist eine rechteckige $N \times M$ Matrix mit Einträgen $B_{j,i} = \varphi_j(\boldsymbol{x}_i), i = 1, \ldots M, j = 1, \ldots N$ Der Vektor y enthält die Daten y_i und hat die Länge M. Der unbekannte Vektor α enthält die Freiheitsgrade α_j und hat die Länge N. Mit dem Gradienten $P = \nabla$ im Regularisierungsausdruck in (10) erhalten wir ein Poisson Problem mit einem zusätzlichen Term, der dem Interpolationsproblem entspricht. Die natürlichen Randbedingungen für solch eine Differentialgleichung in beispielsweise $\Omega = [0,1]^d$ sind Neumann-Bedingungen. Die Diskretisierung (12) gibt uns nun das lineare Gleichungssystem (19), wobei C einer diskreten Laplace-Matrix entspricht. Um den Klassifikator f_N zu erhalten, müssen wir dieses System nun lösen.

Wiederum verbietet der Fluch der Dimensionalität, für V_N konventionelle Finite Elemente Räume einzusetzen, die etwa über einem uniformen Gitter leben. Die Komplexität würde exponentiell mit der Zahl der Dimensionen wachsen. Statt dessen setzen wir im folgenden dünne Gitter ein. Konkret verwenden wir die Kombinationstechnik. Wie bereits in (3) - (4) beschrieben, betrachten wir das Minimierungsproblem auf einer Folge von Gittern, lösen die jeweils entstehenden linearen Probleme (19) und kombinieren die Lösungen gemäß (4). Die Komplexität dieses Verfahrens ist dabei bezüglich k wie üblich von der Ordnung $O(k^{d-1}2^k)$. In Bezug auf die Zahl M der Trainingsdaten skaliert der Aufwand bei geschickter Implementierung linear in M. Dies ist ein substantieller Vorteil gegenüber den meisten neuronalen Netzen und den Supportvektormaschinen, die im allgemeinen quadratisch in M skalieren und deswegen nicht für Probleme mit hohen Datenmengen geeignet sind.

Nun wenden wir unseren Ansatz auf verschiedene Test-Datenmengen an. Hier verwenden wir sowohl synthetische Daten, die mit dem Programm DatGen [66] generiert worden sind, als auch Daten, die aus praktischen Data Mining Anwendungen stammen. All diese Datenmengen wurden auf $[0,1]^d$ reskaliert. Um unsere Methode zu validieren, geben wir die Korrektheitsraten auf Test-Datenmengen, falls verfügbar, und die zehnfachen Kreuzvalidierungsergebnisse auf Training und Testdatenmengen an. Bei der zehnfachen Kreuzvalidierung gehen wir wie folgt vor: Wir unterteilen die Trainingsdaten in 10 gleichgroße Teilmengen. Für $i = 1$ bis 10 bestimmen wir die i-te dieser Teilmengen als weitere Testmenge und bauen mit den Daten der verbleibenden neun Teilmengen den Dünngitter-Kombinationsklassifikator auf. Wir werten dann diese Korrektheitsraten auf der aktuellen Training- und Testmenge aus. Auf diese Weise erhalten wir 10 verschiedene Trainings- und Testkorrektheitsraten. Das zehnfache Kreuzvalidierungsergebnis ist dann gerade der Durchschnitt dieser 10 Korrektheitsraten. Für weitere Details siehe [85]. Für eine kritische Diskussion der Auswertung der Qualität von Klassifikations-Algorithmen, siehe [77].

Zweidimensionale Probleme. Wir betrachten zunächst zweidimensionale Probleme mit kleinen Datenmengen, die bestimmte Strukturen aufweisen. Dann behandeln wir Probleme mit großen Mengen synthetischer Daten mit bis zu 5 Millionen Punkten.

Das erste Beispiel ist aus [56,61] übernommen. Hier sind 1000 Trainings-Datenpunkte gegeben, die mehr oder weniger gleichförmig in $\Omega = [0,1]^2$ verteilt sind. Die zugehörigen Datenwerte sind plus eins oder minus eins abhängig von der Position in Ω, wobei eine 4×4 Schachbrettstruktur erscheint, siehe Abb. 15 (links). Wir berechnen die zehnfache kreuzvalidierte Trainings- und Testkorrektheit mit der Dünngitterkombinationsmethode für verschiedene Werte des Regularisierungsparameters λ und für verschiedene Level k. Die Ergebnisse werden in Abb. 15 (rechts) gezeigt.

Wir sehen, daß die zehnfache Testgenauigkeit für Werte von λ zwischen $3 \cdot 10^{-5}$ und $5 \cdot 10^{-3}$ um 95 % liegt. Unsere beste zehnfache Testkorrektheit war 96.20% auf Level 10 bei $\lambda = 4.54e-05$. Die Schachbrettstruktur wird somit mit weniger als 4 % Fehler rekonstruiert.

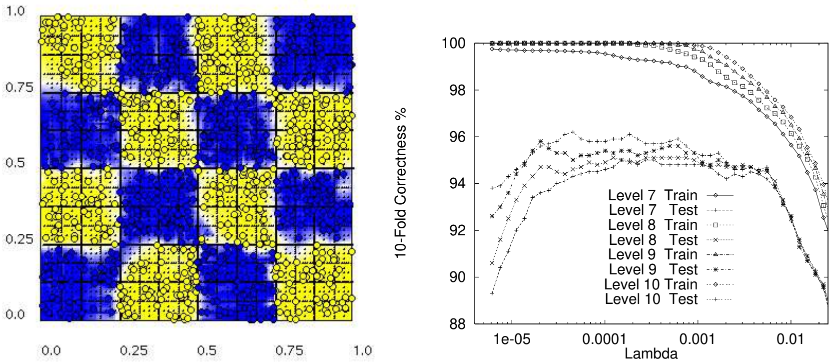

Abbildung 15. Links: Checkerboard-Datensatz [61], Kombinationstechnik mit Level 10, $\lambda = 4.53999\text{e-}05$, Rechts: Plot der Abhängigkeit von λ (in logarithmischer Skalierung) und Level k

Ein anderes zweidimensionales Beispiel mit Struktur ist die Spiralen-Datenmenge, die erstmalig von Alexis Wieland von der MITRE Corp vorgeschlagen wurde, siehe auch [33]. Hier beschreiben 194 Datenpunkte zwei ineinander verwobene Spiralen, siehe Abbildung 16. Dies ist sicherlich ein künstliches Problem, das nicht in praktischen Anwendungen auftaucht. Es dient jedoch als harter Testfall für neue Data Mining Algorithmen. Es ist bekannt, daß neuronale Netzwerke große Probleme mit dieser Datenmenge haben können und manche neuronalen Netzwerke die zwei Spiralen sogar überhaupt nicht trennen können. In Abbildung 16 zeigen wir die Ergebnisse, die mit unserer Dünn-Gitter-Kombinationsmethode für $\lambda = 0,001$ für $k = 6$ und $k = 8$ erzielt wurden. Bereits auf Level 6 werden die zwei Spiralen klar erkannt und aufgelöst. Man beachte, daß hier nur 577 Gitterpunkte im dünnen Gitter enthalten sind. Für Level 8 (2817 Dünn-Gitterpunkte) wird die Form der zwei rekonstruierten Spiralen glatter und die Rekonstruktion genauer.

Sechsdimensionale Probleme. Die BUPA Liver Disorders Datenmenge der Irvine Machine Learning Datenbank [9] besteht aus 345 Datenpunkten mit 6 Merkmalen sowie einem Selektorfeld, das benutzt wird, um die Daten in 2 Mengen mit 145 Punkten und 200 Punkten zu trennen. Hier liegen uns nur Trainingsdaten vor und wir können auch deswegen nur über unsere zehnfachen Kreuzvalidierungsergebnisse berichten. Ein Vergleich mit Testdaten ist nicht möglich.

Wir vergleichen die zwei besten Ergebnisse aus [65], die dort eingeführte geglättete Support Vektor Maschine (SSVM) und den (FSV) Algorithmus [13]. Die Ergebnisse sind in Tabelle 1 zu sehen.

Unser Dünn-Gitter-Kombinationsansatz ergibt auf Level 3 mit $\lambda = 0.0625$ eine 69.23 prozentige zehnfache Testkorrektheit. Aber auch unsere anderen

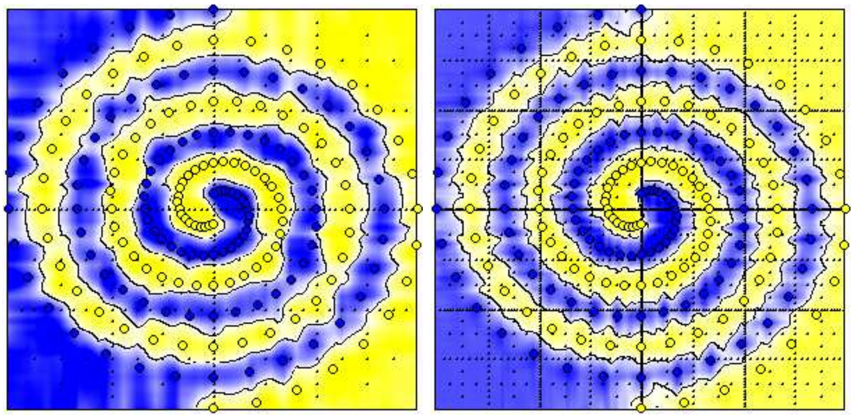

Abbildung 16. Spiraldatensatz, Kombinationstechnik mit Level 6 (links) und 8 (rechts), $\lambda = 0.001$

	SSVM [65]	FSV [65]	Dünn-Gitter-Kombinationsmethode Level 1 $\lambda = 0.0001$	Level 2 $\lambda = 0.1$	Level 3 $\lambda = 0.0625$	Level 4 $\lambda = 0.625$
10-f. Train.-genau., %	70.37	68.18	83.28	79.54	90.20	88.66
10-f. Test.-genau., %	70.33	65.20	66.89	66.38	69.23	68.74

Tabelle 1. Ergebnisse für den BUPA Liver Disorders Datensatz

Ergebnisse sind in etwa in diesem Bereich. Unsere Methode verhält sich nur geringfügig schlechter als die SSVM aber deutlich besser als FSV. Man beachte, daß die Ergebnisse für den robusten linearen Programm-Algorithmus (RLP) [7], die Support-Vektor-Maschine basierend auf der 1-Norm ($SVM_{||.||_1}$) und die klassische Support-Vektor-Maschine ($SVM_{||.||_2^2}$) [13,22,88] schlechter sind [65].

Nun erzeugen wir mit DatGen [66] eine 6-dimensionale Datenmenge mit 5 Millionen Trainingspunkten und 20000 Testpunkten. Wir verwenden den Aufruf

```
datgen -r1 -X0/100,R,O :0/100,R,O:0/100,R,O:0/100,R,O:0/200,
R,O:0/200,R,O -R2 -C2/4 -D2/5 -T10/60 -O502 0000 -p -e0.15
```

Die Ergebnisse sind in Tabelle 2 aufgeführt. Bereits auf Level 1 wird eine Testgenauigkeit von 88 % erzielt, was für diese Daten sehr zufriedenstellend ist. Wir sehen, daß wirklich große Datenmengen von 5 Millionen Punkten gehandhabt werden können. An den CPU-Zeiten, die für die Berechnung der Matrizen $G_{\mathbf{l}} = B_{\mathbf{l}} \cdot B_{\mathbf{l}}^T$ benötigt wurden, sieht man, daß mehr als 96 % der Rechenzeit für das Aufstellen der Matrix benötigt wird. Die Ausführungszeiten skalieren wiederum linear in der Zahl der Datenpunkte.

	# of points	Trainings- Genauig.	Test- Genauig.	gesamte Zeit (sek)	Datenmatrix Zeit (sek)	# der Iterationen
Level 1	50 000	87.9	88.1	158	152	41
	500 000	88.0	88.1	1570	1528	44
	5 Million	88.0	88.0	15933	15514	46
Level 2	50 000	89.2	89.3	1155	1126	438
	500 000	89.4	89.2	11219	11022	466
	5 Million	89.3	89.2	112656	110772	490

Tabelle 2. Ergebnisse für einen sehr großen synthetisches 6D Datensatz, $\lambda = 0.01$

In analoger Weise lassen sich zur Zeit bis zu 15-dimensionale Probleme behandeln, dann stößt das Verfahren auf unseren Rechnern an seine Speichergrenzen. Dies ist jedoch immer noch für die Praxis interessant. In sehr hochdimensionalen Daten sind im allgemeinen starke gegenseitige Abhängigkeiten unter den Dimensionen enthalten. In einem Vorverarbeitungsschritt lassen sich, etwa mit Hilfe der Principal Component Analyse (faktisch einer Singulärwertzerlegung), die effektiven Dimensionen bestimmen und in vielen Fällen damit die Dimensionalität des Problems auf Zahlen kleiner 15 reduzieren.

2.3 Material science und life science

Die Moleküldynamik ist neben dem Monte Carlo Verfahren die zentrale Simulationsmethode bei der Untersuchung von Materialien und Biomolekülen auf der klassischen Ebene. Hier wird die Dynamik eines Systems, etwa die Atome eines Kristalls oder eines Proteins, unter äußeren Kräften durch die Newtonschen Bewegungsgleichungen

$$\begin{aligned} \dot{x}_i &= u_i, \\ m_i \dot{u}_i &= -\nabla_{x_i} V(\{x_j\}_j), \;\; i = 1, \ldots, N. \end{aligned} \tag{20}$$

beschrieben. Hier bezeichnet x_i den Ort des Teilchens i, u_i seine Geschwindigkeit, m_i seine Masse und V die Potentialfunktion. Die Kraft F_i auf ein Partikel i ist definiert als

$$F_i = -\nabla_{x_i} V(\{x_j\}_j)$$

und hängt im Falle konservativer Potentiale nur von den Koordinaten $\{x_i\}_i$ der Partikel ab. Systeme, die allein durch paarweise Wechselwirkungen zwischen Teilchen beschrieben werden können, welche nur vom Abstand $r_{ij} := |x_i - x_j|$ jeweils zweier Teilchen abhängen, werden dann durch ein Potential der Art

$$V(x_1, \ldots, x_N) = \sum_{k,j;k<j} U_{kj}(r_{kj})$$

modelliert, wobei U_{kj} das zwischen den Partikeln k und j herrschende Potential bezeichne.

Die gewöhnliche Differentialgleichung (20) wird nun in der Zeit diskretisiert. Hier ist es günstig, aus Gründen der Energieerhaltung ein sogenanntes symplektisches Integrationsverfahren einzusetzen. Das gebräuchlichste ist das Verlet-Verfahren (eine Variante des Diffenzenschemas zweiter Ordnung). In jedem Zeitschritt ist dabei nun die Kraft auf die Partikel zu bestimmen. Da jedes Teilchen mit jedem wechselwirkt, sind die Kosten hier $O(N^2)$, was für große Teilchenzahlen nicht mehr praktikabel ist.

Ist nun das Potential schnellabfallend, dann kann man es in einem genügend großen Abstand von einem Teilchen einfach abschneiden, also zu Null setzen, ohne einen großen Fehler zu machen. Die Teilcheninteraktion ist damit nur noch auf die nächsten Nachbarn beschränkt und man erhält ein Verfahren der Komplexität $O(N)$, wobei die Proportionalitätskonstante durch die mittlere Zahl der Nachbarn gegeben ist. Mit einer geeigneten Linked-List Datenstruktur, die es erlaubt die Nachbarn eines Partikels in konstanter Zeit zu finden, läßt sich der resultierende Algorithmus ohne Schwierigkeiten implementieren [44].

Anders sieht es jedoch aus, wenn das Potential langreichweitig ist (genauer: in drei Dimensionen langsamer als $1/r^3$ mit dem Abstand r abfällt). Ein Beispiel ist hier das Coulombpotential

$$U(r_{ij}) = -\frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}}$$

das insbesondere in Biosimulationen eine Rolle spielt. Hier darf das Potential nicht ohne weiteres abgeschnitten werden. Es gibt nun einige approximative Verfahren zur Kraftberechnung bei langreichweitigen Potentialen, die eine Komplexität von $O(N \log N)$ oder sogar $O(N)$ aufweisen und es erlauben, die Kraft auf eine vorgegebene Genauigkeit genau auszurechnen. Dies sind diverse Baumverfahren, Multipolmethoden sowie das P^3M-Verfahren. Diese Verfahren für langreichweitige Kräfte haben als grundlegendes Prinzip, daß es nicht notwendig ist, diejenigen Anteile zur Kraft auf ein Partikel exakt zu berechnen, die von weit entfernten Partikeln herrühren. Wir betrachten im folgenden sogenannte gitterbasierte Verfahren wie die Particle-Particle–Particle-Mesh (P^3M) Methode genauer.

Das elektrostatische Potential

$$\Phi(x) = \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{N} q_i \frac{1}{|x - x_i|} \tag{21}$$

einer Partikelverteilung aus N Partikeln an den Orten $\{x_i\}_i$ mit Ladungen $\{q_i\}_i$ erhält man auch als Lösung der Potentialgleichung [8]

$$-\Delta\Phi = \frac{1}{\epsilon_0}\rho \tag{22}$$

mit ρ der durch die Partikel gegebene Ladungsverteilung

$$\rho(x) = \sum_{i=1}^{N} q_i \delta(x - x_i). \tag{23}$$

Die zugehörige Kraft auf ein Partikel mit der Ladung q_i am Ort x_i und die elektrostatische Energie des Gesamtsystems aus N Partikeln erhält man aus der Lösung Φ des Potentialproblems (22) durch

$$F_i = -q_i \nabla\Phi|_{x_i}. \tag{24}$$

Eine Methode zur Berechnung der Kräfte auf die Partikel bestünde also darin, die Kraft auf ein einzelnes Partikel nicht als Summe über die Beiträge aller anderen Partikel zu berechnen, sondern direkt via (24) durch Auswertung der Lösung der Potentialgleichung (22). Die Genauigkeit der dadurch erzielten Approximation an die exakte Kraft auf das Partikel wird durch die Genauigkeit der approximativen Lösung von (22) bestimmt. Standard-Diskretisierungsverfahren wie Finite Elemente konvergieren jedoch aufgrund der Singularitäten der Lösung nur sehr schlecht und führen zu großen Fehlern, insbesondere in der Nähe der Positionen der Partikel. Das Problem besteht darin, daß eine rechte Seite mit geringer Glattheit zu einer Lösung führt, die ebenfalls nur von geringer Glattheit ist. Im Fall von δ-Distributionen als rechter Seite ergeben sich sogar Singularitäten in der Lösung, vgl. (21).

Trotzdem läßt sich die Idee der approximativen Lösung der Potentialgleichung gewinnbringend einsetzen, indem man folgende Ideen verwendet:

- Das Potential wird in einen kurzreichweitigen schnell abfallenden und einen nur langsam abfallenden aber sehr glatten Anteil aufgespalten.
- Der schnell abfallende Anteil des Potentials wird abgeschnitten und mit Hilfe direkter Summation zusammen mit der Linked-List Methode ausgewertet.

[8] Das Gebiet ist eigentlich unendlich groß und die Lösungen fallen im Unendlichen zu Null ab. In der Praxis findet man zwei Zugänge: Das Gebiet wird als endlich aber groß gewählt, z.B. $\Omega = [-a, a]^3$ mit a genügend groß, und es werden homogene Dirichlet Randbedingungen angesetzt. Daneben werden auch gewisse Abschirmrandbedingungen verwendet. Alternativ dazu wird das Gebiet periodisch ins Unendliche fortgesetzt und deswegen periodische Randbedingungen benutzt. Letzterer Zugang ist insbesondere bei regelmäßigen Strukturen wie etwa Kristallen u.a. weit verbreitet. Da die Lösung der Potentialgleichung mit periodischen Randbedingungen nur bis auf eine Konstante bestimmt ist, sind zusätzliche Bedingungen zu stellen.

– Der langsam abfallende glatte Anteil des Potentials wird aus der Lösung eines Potentialproblems mit veränderter rechter Seite mit Hilfe geeigneter Diskretisierungs- und Lösungsverfahren für die Potentialgleichung berechnet.

In analoger Weise können auch die zugehörigen Kräfte in zwei Anteile zerlegt und entsprechend behandelt werden.

Die Zerlegung des Potentials. Das bekannte Superpositionsprinzip läßt sich nun gewinnbringend ausnutzen, um die Lösung von (22) zu approximieren. Addiert und subtrahiert man nämlich zur rechten Seite von (22) eine Summe der Art $1/\epsilon_0 \sum_{i=1}^N \rho_i$, so läßt sich (22) schreiben als

$$\Delta\Phi = \frac{1}{\epsilon_0}\sum_{i=1}^{N}(\rho_i - q_i\delta_i) - \frac{1}{\epsilon_0}\sum_{i=1}^{N}\rho_i. \tag{25}$$

Dann läßt sich die Lösung aus den Lösungen der zwei Probleme

$$\Delta\Phi^{kr} = \frac{1}{\epsilon_0}\sum_{i=1}^{N}(\rho_i - q_i\delta_i) \tag{26}$$

und

$$\Delta\Phi^{lr} = -\frac{1}{\epsilon_0}\sum_{i=1}^{N}\rho_i \tag{27}$$

additiv aus den beiden Teilen Φ^{kr} und Φ^{lr} zu $\Phi = \Phi^{kr} + \Phi^{lr}$ zusammensetzen. Analog ergeben sich dann die Kräfte und die Energie aus diesen zwei Anteilen, z.B. die Kraft zu

$$F_i = -\sum_{\substack{j=1\\ j\neq i}}^{N} q_j \nabla\Phi|_{x_j} = -\sum_{\substack{j=1\\ j\neq i}}^{N} q_j \nabla\Phi^{kr}|_{x_j} - \sum_{\substack{j=1\\ j\neq i}}^{N} q_j \nabla\Phi^{lr}|_{x_j}. \tag{28}$$

Der Vorteil dieser Formulierung des Potentialproblems als zwei separate Probleme besteht darin, daß für eine geeignete Wahl der Funktionen ρ_i die Lösung Φ^{kr} von (26) durch direkte Summation mit geringem Aufwand erlangt werden kann. Da das zugehörige Potential nur kurzreichweitig ist, kann der entsprechende Kraftanteil mit dem Linked-List Verfahren in $O(N)$ Operationen behandelt werden. Die Lösung Φ^{lr} von (27) kann mit Standarddiskretisierungs- und Lösungsverfahren für das Potentialproblem schnell und genau approximiert werden. Dazu muß die von ρ_i induzierte Ladung gleich der Ladung des Partikels q_i sein, d.h. es muß gelten $\int \rho_i(x)dx = q_i$. Weiterhin muß ρ_i punktsymmetrisch um die Partikelposition x_i sein, also nur vom Abstand von x_i abhängen, und kompakten Träger haben oder schnell abfallen (im Sinne obiger Definition langreichweitiger Potentiale). Schließlich sollen die ρ_i glatte Funktionen sein. Die ersten drei Forderungen sorgen dafür, daß das von der Ladungsverteilung $q_i\delta_i - \rho_i$ erzeugte Potential außerhalb des (numerischen) Trägers von ρ_i gleich Null (bzw. sehr klein) ist. Die Glattheitsforderung an die ρ_i sorgt dafür, daß auch die Lösung Φ^{lr} von (27) eine glatte Funktion ist, die sich mittels Standard-Approximationsverfahren gut approximieren läßt[9].

Diskretisierung mit dünnen Gittern. Wir müssen also das Problem (27) geeignet diskretisieren und lösen. Etwa kann man ein genügend feines uniformes Gitter Ω_k mit Maschenweite $2a2^{-k}$ verwenden, dort mit stückweise linearen Finiten Elementen diskretisieren und das entstehende lineare System mit einem Mehrgitterverfahren lösen. Das ist bis auf das MG-Verfahren im Prinzip der von Hockney und Eastwood [58] vorgeschlagene Zugang. Der Diskretisierungsfehler ist dann von der Ordnung $O(2^{-2k})$, der Aufwand ist von der Ordnung $O(2^{3k})$ in drei Raumdimensionen.

Statt dessen können wir zur Diskretisierung und Lösung von (27) ein Dünn-Gitter-Verfahren verwenden. Dies ist nun im Gegensatz zum ursprünglichen Problem (22) möglich, da nun die für das dünne Gitter notwendigen Glattheitsvoraussetzungen erfüllt sind. Hier kann im Fall homogen verteilter Partikel sowohl der auf dem Galerkin-Verfahren basierende Dünn-Gitter-Ansatz als auch die Kombinationstechnik verwendet werden. Die Komplexität ist dann substantiell reduziert. Wir erhalten eine Genauigkeit der Ordnung $O(k^2 2^{-2k})$ mit nur $O(k^2 2^k)$ Aufwand. Vernachlässigen wir der Einfachheit halber die k^2-Terme und vergleichen die beiden Verfahren, dann sehen wir, daß sich der Dünn-Gitteransatz implizit wie ein Verfahren höherer, nämlich sechster Ordnung verhält. Nun kann man natürlich statt der linearen Basen auf dem uniformen Gitter auch höhere Polynome als stückweise Ansatzfunktionen verwenden und erhält, da die Lösung des Problems glatt ist, einen Diskretisierungsfehler der Ordnung $2^{-k(p+1)}$, wobei p den Polynomgrad bezeichne. Jedoch gibt es analog dazu auch zugehörige dünne Gitter mit Polynomen höherer Ordnung [14] und der Komplexitätsvorteil der dünnen Gitter bleibt erhalten, ja er wird sogar immer größer mit steigendem p. Etwa

[9] Ein Beispiel für ρ_i sind die Gauß'sche Glockenkurven

$$\rho_i(\boldsymbol{x}) := q_i \left(\frac{G}{\sqrt{\pi}}\right)^3 \exp(-G^2(x - x_i)^2)$$

mit Parameter G. Dann fällt ρ_i exponentiell mit dem Abstand von x_i ab, ist punktsymmetrisch um x_i und ist eine C^∞-Funktion, d.h. sie ist unendlich oft stetig differenzierbar. Die Lösung von (27) mit diesen ρ_i's als rechter Seite ist dann ebenfalls aus C^∞ und läßt sich damit auf einem Gitter gut approximieren. Zudem lassen sich in diesem speziellen Fall die Lösungen der o.e. Potentialprobleme im periodischen Fall sogar analytisch mittels der sogenannten Fehlerfunktion in Form von unendlichen Reihen angeben.

für $p = 2$ verhält sich das Dünn-Gitter-Verfahren implizit (und unter Vernachlässigung der k^2-Terme) wie ein Verfahren der Ordnung 9 und mit $p = 3$ würden wir schon implizit ein Verfahren der Ordnung 12 erreichen. Lediglich ein exponentiell konvergentes Verfahren kann hier noch überlegen sein. Dies gibt es mit der hp-Version der Finiten Elemente oder mit der Fourier-Spektralmethode. Die Fourier-Transformation des Problems (27) und die anschließende spektrale Diskretisierung und Lösung im Fourier Raum mittels FFT wird in der PME-Methode [26] verfolgt. Für homogen verteilte Partikel ist sie somit zumindest asymptotisch überlegen. Jedoch muß auch hier auf die Partikelpositionen zurückinterpoliert werden, was ein gewisser limitierender Faktor ist. Mittels Splines höherer Ordnung wird dies in der „smoothed" PME-Methode [31] weiter verbessert. Erste Vergleiche zwischen der Dünn-Gitter-Kombinationsmethode und dem PME-Verfahren haben ergeben, daß die Dünn-Gitter-Diskretisierung in praxisrelevanten Anwendungen mit vielen hunderttausend Partikeln ebenbürtig ist.

Im Fall inhomogen verteilter Partikel oder Atome reduziert sich die Effizienz der FFT-basierten Techniken substantiell. Auch eine Diskretisierung auf einem uniformen Gitter oder einem regulären dünnen Gitter wird weniger effizient, denn die verwendete Gitterweite muß fein genug sein um auch nun noch die inhomogen verteilten Partikel genügend genau erfassen zu können. Hier bieten sich adaptive Methoden an. Die adaptive Dünn-Gitter-Methode höherer Ordnung ist direkt einsetzbar, während Fouriertechniken i.a. nicht ohne weiteres adaptiv eingesetzt werden können. Ich glaube auch, daß adaptive dünne Gitter höherer Ordnung den in diesem Fall verwendeten Baum-Algorithmen und Multipol-Verfahren ebenbürtig wenn nicht sogar überlegen sind. Genaueres müssen zukünftige Untersuchungen zeigen.

Anwendungen aus den Material- und Biowissenschaften. Im folgenden zeigen wir einige Simulationsergebnisse aus dem Bereich der Material- und der Biowissenschaften. Abbildung 17 gibt einen Eindruck, wie mit der Moleküldynamikmethode das Schmelzen von Salzkristallen, hier Kochsalz, untersucht werden kann. Dabei tritt im Potential durch die Na+ und Cl- Ionen ein nicht zu vernachlässigender langreichweitiger Anteil auf. Durch Simulation läßt sich hier der Phasenübergang Fest-Flüssig, und damit der Schmelzpunkt des Salzes sehr genau bestimmen. Abbildung 18 zeigt einen Schnappschuß einer Rechnung mit Dimethylcarbonat-Molekülen. Nun sind neben den langreichweitigen Potentialanteilen auch innermolekulare Wechselwirkungen im Potential vorhanden, die die Struktur des Moleküls bestimmen. Diese sind kurzreichweitig und lassen sich im Linked-List Teil des Codes mit behandeln. Von Interesse ist hier die Untersuchung des Übergangs von der Flüssigphase in den Glaszustand und Cis-/Trans-Strukturfragen des Moleküls.

Eine weitere Anwendung der Moleküldynamiksimulation ist die Untersuchung der Ausbreitung von Mikrorissen in Kristallen. Abbildung 19 gibt ein Beispiel. Mittels weitergehender Streßanalysen läßt sich die Ausbreitung eines

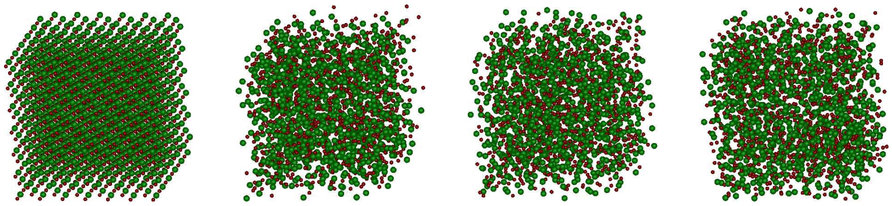

Abbildung 17. NaCl-Schmelzen. Startkonfig.: uniformes Gitter, 1014 Na, 1014 Cl Ionen, Maxwell-Boltzmann Geschwindigkeitsverteilung bei 1800 K, Partikelkonfiguration für NaCl zu den Zeiten 0.00, 0.47, 0.94 and 1.41 (Zeiteinheit: 1.03e-15 sec)

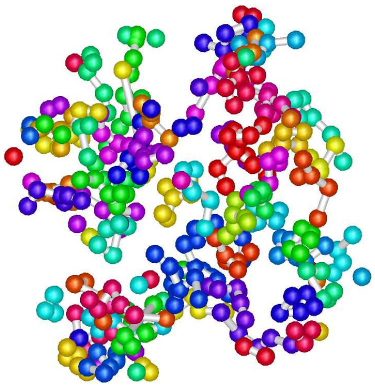

Abbildung 18. Moleküldynamiksimulation von Dimethylcarbonat

Risses auch automatisch verfolgen (das Partikel mit dem maximalen Stress befindet sich an der Rißspitze) und damit Aussagen zur Schädung eines Materials auf Mikroebene machen. Eine Herausforderung für die Zukunft ist, diese Ergebnisse in makroskopische Kontinuumsmodelle einfließen zu lassen und diese dadurch zu verbessern.

Schließlich lassen sich auch komplexe Molekülstrukturen mittels Moleküldynamikmethoden untersuchen. Ein interessantes Beispiel sind die sogenannten Nanotubes. Diese sind eine röhrenförmige Verallgemeinerung der bekannten Buckyballs, der C_{60} Kohlenstoff-Fußballmoleküle. Nanotubes wurden 1991 von S. Ijima [60] entdeckt und lassen sich mittlerweise gezielt fertigen. Sie sind nur wenige Nanometer dick, können aber eine Länge im Millimeterbereich aufweisen. Sie besitzen je nach Typ metallische oder halbleitende Eigenschaften, sind sehr leicht aber zeichnen sich durch ihre extreme Festigkeit aus, die an Diamant heranreicht. Ihr zukünftiges Einsatzspektrum reicht von Nanodrähten in Chips über die Verstärkung von Composit-Material, und chemischen Muskeln bis hin zum Speichermedium für Wassserstoff. Es gibt

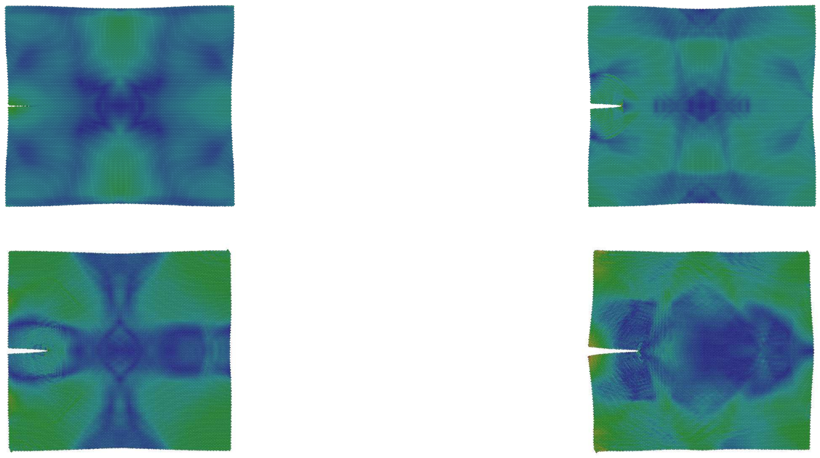

Abbildung 19. Entwicklung eines Mikrorisses in harten Materialien (2D)

bereits erste industrielle Anwendungen in speziellen Flachbildschirmen von Samsung. Abbildung 20 und 21 zeigen simulierte Zug- und Biegeversuche mit einwandigen Nanotubes, Abbildung 22 einen Zugversuch mit mehrwandigen Nanotubes.

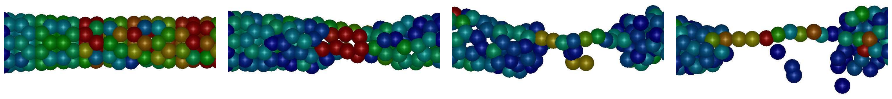

Abbildung 20. Zugexperiment mit Kohlenstoff-Nanotubes

Neben ihrem Einsatz in den Materialwissenschaften ist die Moleküldynamik ein wichtiges Werkzeug zur Untersuchung von Bioproteinen. Die intramolekularen Potentialanteile beinhalten nun Abstände, Winkel und Diederwinkel in Form von Mehrkörperpotentialen. Diese und das Lennard-Jones-Potential lassen sich kurzreichweitig mittels des Linked-List Codeanteils behandeln. Bioproteine sind jedoch grundsätzlich in Wasser gelöst und müssen deswegen für eine realistische Simulation von einer Vielzahl von Wassermolekülen umgeben werden. Weiterhin sind gelöste Salze im Wasser allgegenwärtig. Das Wasser und die Salzionen sind elektrostatisch und benötigen

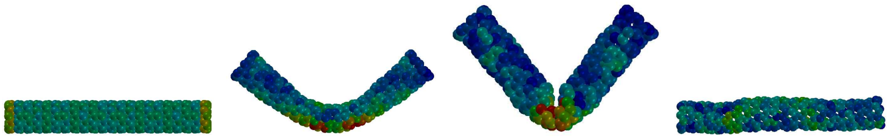

Abbildung 21. Biegen von Nanotubes

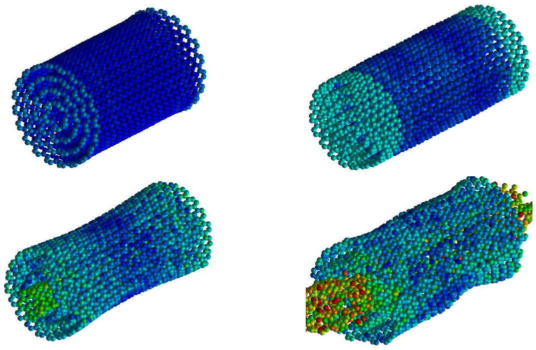

Abbildung 22. Zugexperiment mit mehrwandigen Kohlenstoff-Tubes

das Coulomb-Potential. Eine erste Simulation von Bovine Pankrease Tryptin Inhibitor (BPTI), dem am häufigsten studierten Biomolekül ist in Abbildung 23 zu sehen. Aufgabenstellung ist hier etwa die Gestalt des Proteins in Wasser zu errechnen (die Moleküldaten sind in dehydriertem Zustand mittels Röntgenkristallographie oder NMR (solution nuclear magnetic resonance) vermessen und erlauben deswegen meist keine Aussagen über die tatsächliche Gestalt eines Proteins). Wir haben mittlerweile eine Schnittstelle an die Brookhaven Protein Data Bank implementiert, die es uns nun erlaubt, nahezu sämtliche bekannte Proteine zu simulieren. Beispiele sind in Abbildung 24 gegeben. Zentrale Fragestellungen neben der Form sind Schlüssel-Schloß Mechanismen, um Liganden und Inhibitoren finden zu können. Untersucht wird dabei u.a. wieviel Kraft nötig ist, um einen Wirkstoff aus einer Bindungstasche eines Biomoleküls herauszuziehen. Für kleine Proteine ist mittlerweile auch die Entfaltung aus der Helix-Struktur simulierbar. Das geht gerade noch mit heutigen Rechnern. Zukunftsmusik ist jedoch sicherlich die Proteinfaltung. Die dabei wirkenden Mechanismen sind bis heute weitgehend unverstanden. Sie laufen zudem auf einer für die Simulation (noch) zu großen Zeitskala im Millisekundenbereich ab, die von Moleküldynamik-Rechnungen aufgrund ihrer Zeitschritte im Femtosekundenbereich noch nicht erreicht werden kann.

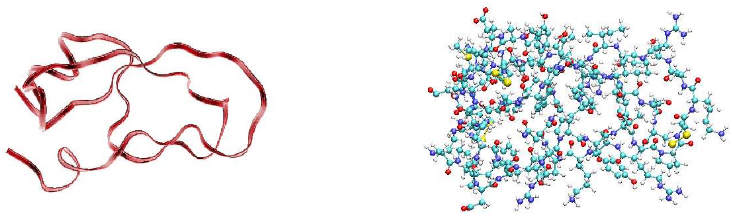

Abbildung 23. BPTI in „ribbon“ und „ball-and-stick“-Darstellung

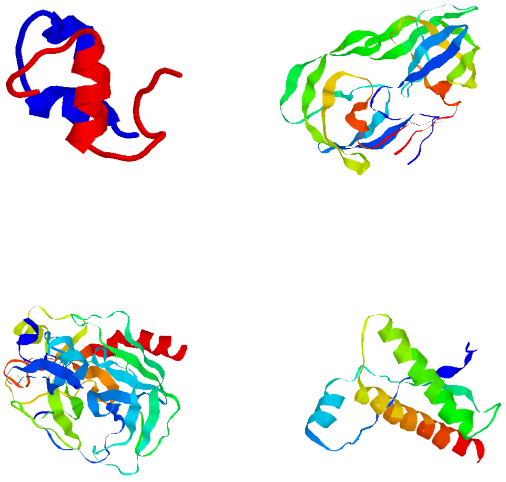

Abbildung 24. Biomoleküle, wie sie von der Protein Data Base Brookhaven zur Verfügung gestellt werden. Reaktiver Teil des menschlichen Insulinhormons (oben links), HIV-1 Protease mit Inhibitor SDZ283-910 (oben rechts), Protease Inhibitordomain des Alzheimer-Amyloid-β-Peptids (unten links), Rinder-Prion (unten rechts)

3 Schlußbemerkungen

Wir haben gesehen, daß dünne Gitter in einer Vielzahl von praktischen Problemen über die Lösung partieller Differentialgleichungen hinaus Anwendung finden. Insbesondere für höherdimensionale Aufgabenstellungen sind sie bestens geeignet und beginnen nun allmählich ihre Vorteile in der Praxis zu zeigen.

Vielleicht gelingt es in naher Zukunft auch, die Methodik der Zenger-Korrekturen auf die in der Dünn-Gitter-Kombinationstechnik auftretenden Teilprobleme anzuwenden und so auch dieses Verfahren vom Pollutionseffekt der Singularitäten zu befreien. Neben der Tatsache, daß sich dabei nun zwei Zengersche Entwicklungen gegenseitig befruchten würden, wäre dies für viele höherdimensionale Aufgabenstellungen ein echter Fortschritt. Eine adaptive Verfeinerung der dünnen Gitter ist in hohen Dimensionen durchaus aufwendig. Zudem würde vielleicht auch der Kampf strukturierte Gitter gegen unstrukturierte Gitter, den Christoph Zenger tapfer gegen viele Widerstände und modische Strömungen aufgenommen hat, bald entschieden, oder zumindest in eine neue Runde gehen.

An dieser Stelle verbleibt mir, meine Glückwünsche zum Jubiläumsgeburtstag auszusprechen. Ich will hier nun mit dem Wunsch enden, daß Christoph Zenger nun nicht wie schon seit Jahren immer wieder halb im Scherz behauptet, die *Landeklappen ausfahren*[10] wird, sondern mindestens noch die nächsten 20 Jahre aktiv und tolerant die Wissenschaft voranbringen wird. Prost, auf das Geburtstagskind.

Literatur

1. V. Apostolescu, *Numerische Behandlung der stationären Umströmung einer Ecke bei zähen inkompressiblen Fluiden*, Dissertation, TU München, 1985.
2. K.I. Babenko, *Approximation by trigonometric polynomials in a certain class of periodic functions of several variables*, Dokl. Akad. Nauk SSSR 132, 672–675, 1960.
3. R. Balder, *Adaptive Verfahren für elliptische und parabolische Differentialgleichungen auf dünnen Gittern*, Dissertation, Institut für Informatik TU München, 1994.
4. A. Barinka, T. Barsch, P. Charton, A. Cohen, S. Dahlke, W. Dahmen, K. Urban, *Adaptive wavelet schemes for elliptic problems - implementation and numerical experiments*, IGPM Report 173, RWTH Aachen, 1999.
5. G. Baszenski, *N–th order polynomial spline blending*, Multivariate Approximation Theory III, ISNM 75, Birkhäuser, Basel, editors W. Schempp and K. Zeller, 35–46, 1985.
6. F.L. Bauer *On the field of values subordinate to a norm*, Numerische Mathematik 4, 103-111, 1962.

[10] Die Zitate C.Z.s sind großen Teils aus meinem schlechten Gedächtnis rekonstruiert, also nicht immer ganz wortgetreu und vielleicht durch den zeitlichen Abstand verklärt. Im Kern stimmen sie aber, glaub' ich.

7. K. P. Bennett and O. L. Mangasarian. *Robust linear programming discrimination of two linearly inseparable sets.* Optimization Methods and Software 1, 23–34, 1992.
8. M. Berry and G. Linoff, *Mastering data mining*, Wiley, 2000.
9. C.L. Blake and C.J. Merz, *UCI Repository of machine learning databases*, 1998, http://www.ics.uci.edu/~mlearn/MLRepository.html
10. T. Bonk, *Ein rekursiver Algorithmus zur adaptiven numerischen Quadratur mehrdimensionaler Funktionen*, LRZ–Bericht Nr. 9501, Leibnitz-Rechenzentrum der Bayerischen Akademie der Wissenschaften, 1995.
11. F.F. Bonsall, J. Duncan, *Numerical ranges of operators on normed spaces and of elements of normed algebras*, London Mathematical Society Lecture Note Series 2, Cambridge University Press, 1971.
12. F.F. Bonsall, J. Duncan, *Numerical ranges II*, London Mathematical Society Lecture Note Series 10, Cambridge University Press, 1973.
13. P. S. Bradley and O. L. Mangasarian. *Feature selection via concave minimization and support vector machines*, in J. Shavlik, editor, Machine Learning Proceedings of the Fifteenth International Conference (ICML '98), 82–90, Morgan Kaufmann, 1998.
14. H.-J. Bungartz, *Finite elements of higher order on sparse grids*, Habilitationsschrift, Institut für Informatik, TU München, 1998.
15. H.-J. Bungartz, *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*, Dissertation, Institut für Informatik, TU München, 1992.
16. H.-J. Bungartz, *An adaptive Poisson solver using hierarchical bases and sparse grids*, in Iterative Methods in Linear Algebra, P. de Groen and R. Beauwens, eds., Elsevier, Amsterdam, 293–310, 1992.
17. H.-J. Bungartz, T. Dornseifer, *Sparse grids: Recent developments for elliptic partial differential equations*, Rep. TUM-I9702, Institut für Informatik, TU München, 1997.
18. H.-J. Bungartz, T. Dornseifer, C. Zenger, *Tensor product approximation spaces for the efficient numerical solution of partial differential equations*, in Proc. Int. Workshop on Scientific Computations, Konya, 1996, Nova Science Publishers, 1997.
19. H.-J. Bungartz, M. Griebel, *A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives*, J. Complexity 15 (1999), 167-199.
20. H.-J. Bungartz, M. Griebel, D. Röschke, C. Zenger, *Pointwise convergence of the combination technique for Laplace's equation*, East-West J. of Numerical Mathematics 1(2) (1994), 21-45.
21. H.-J. Bungartz, M. Griebel, and U. Rüde, *Extrapolation, combination, and sparse grid techniques for elliptic boundary value problems.* Comput. Methods Appl. Mech. Eng., 116:243–252, 1994.
22. V. Cherkassky and F. Mulier, *Learning from data - concepts, theory and methods.* John Wiley & Sons, 1998.
23. K. Cios, W. Pedrycz, and R. Swiniarski, *Data mining methods for knowledge discovery*, Kluwer, 1998.
24. A. Cohen, W. Dahmen, R. DeVore, *Adaptive wavelet methods for elliptic operator equations – convergence rates*, IGPM report 165, RWTH Aachen 1998.

25. M.J. Crabb, The numerical range of an operator, PH.D. Thesis, Edinburgh, 1969.
26. T.A. Darden, D. York, L.G. Pedersen, *Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems*, J. Chem. Phys. 98, 12, 10089-10092, 1993.
27. F.J. Delvos, W. Schempp, *Boolean methods in interpolation and approximation*, Pitman Research Notes in Mathematics Series 230, Longman Scientific & Technical, Harlow, 1989.
28. E. Deutsch, C. Zenger, *Inclusion domains for the eigenvalues of stochastic matrices*, Numerische Mathematik 18, 182-192, 1971.
29. W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal. 33, 1106-1124, 1996.
30. T. Dornseifer, *Diskretisierung allgemeiner elliptischer Differentialgleichungen in krummlinigen Koordinatensystemen auf dünnen Gittern.* Dissertation, Technische Universität München, 1997.
31. U. Essmann, L. Perera, M. Berkowitz, T.A. Darden, H. Lee, L.G. Pedersen, *A smooth particle mesh Ewald method*, Journal of Chemical Physics 103, 8577-8593, 1995.
32. G. Faber, *Über stetige Funktionen*, Mathematische Annalen, 66 (1909), pp. 81–94.
33. S.E. Fahlmann and C. Lebiere. *The cascade-correlation learning architecture*, In Touretzky, editor, Advances in Neural Information Processing Systems 2, Morgan-Kaufmann, 1990.
34. R. Fößmeier, *Differenzenverfahren hoher Ordnung für elliptische Neumann-Probleme auf beliebigen Gebieten*, Dissertation, Institut für Informatik, TU München, 1984.
35. K. Frank, S. Heinrich, S. Pereverzev, *Information complexity of multivariate Fredholm integral equations in Sobolev classes*, J. of Complexity, 12, 17-34, 1996.
36. S. Friedland, C. Zenger, *All spectral dominant norms are stable*, Linear Algebra Appl. 58, 97-107, 1984.
37. J. Garcke, M. Griebel, M. Thess, *Data mining with sparse grids*, SFB-Report 675, Universität Bonn, 2000.
38. T. Gerstner and M. Griebel, *Numerical integration using sparse grids*, Numer. Algorithms, 18:209–232, 1998.
39. F. Girosi, *An equivalence between sparse approximation and support vector machines*, Neural Computation 10, 6, 1455-1480, 1998.
40. F. Girosi, M. Jones, T. Poggio, *Regularization theory and neural networks architectures*, Neural Computation 7, 219-265, 1995.
41. H. Gietl, *Asymptotische Entwicklungen für modifizierte Differentialverfahren beim Dirichlet-Problem der Poisson-Gleichung in Gebieten mit einspringenden Ecken*, Dissertation, TU München, 1980.
42. M. Griebel, *A parallelizable and vectorizable multi-level algorithm on sparse grids*, in Parallel Algorithms for Partial Differential Equations, Notes on Numerical Fluid Mechanics 31, W. Hackbusch, ed., Vieweg, Braunschweig/Wiesbaden, 1991, pp. 94–100.
43. M. Griebel, *Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences*, Computing 61(2), 151–179, 1998.
44. M. Griebel, A. Caglar, S. Knapek, *Computersimulationen in Physik und Chemie*, Springer, to appear 2001.

45. M. Griebel, W. Huber, U. Rüde, T. Störtkuhl, *The combination technique for parallel sparse-grid-preconditioning and -solution of PDEs on multiprocessor machines and workstation networks*, in L. Bouge, M. Cosnard, Y. Robert and D. Trystram (eds.), Lecture Notes in Computer Science 634, Parallel Processing: CONPAR92-VAPP V, pp. 217-228, Springer Verlag, 1992.
46. M. Griebel, W. Huber, T. Störtkuhl, C. Zenger, *On the parallel solution of 3D PDEs on a network of workstations and on vector computers*, in Parallel Computer Architectures: Theory, Hardware, Software, Applications, A. Bode and M. D. Cin, eds., vol. 732 of Lecture Notes in Computer Science, Springer Verlag, 1993, pp. 276–291.
47. M. Griebel, S. Knapek, *Optimized tensor-product approximation spaces*, Constructive Approximation 16, 4, 525-540, 2000, auch als SFB-256 Bericht Nr. 568, Universität Bonn, 1998.
48. M. Griebel, F. Koster, *Adaptive wavelet solvers for the unsteady incompressible Navier Stokes equations*, in J. Malek, J. Necas and M. Rokyta, editors, Advances in Mathematical Fluid Mechanics, Springer Verlag, 2000.
49. M. Griebel, P. Oswald, *On additive Schwarz preconditioners for sparse grid discretizations*, Numer. Math. 66, 449-464, 1994.
50. M. Griebel, P. Oswald, T. Schiekofer, *Sparse grids for boundary integral equations*, Numer. Mathematik, 83(2), 279–312, 1999.
51. M. Griebel, M. Schneider, C. Zenger, *A combination technique for the solution of sparse grid problems*, in Iterative Methods in Linear Algebra, P. de Groen and R. Beauwens (eds.), Elsevier, Amsterdam, 263–281, 1992.
52. M. Griebel, V. Thurner, *The efficient solution of fluid dynamics problems by the combination technique*, Int. J. Num. Meth. for Heat and Fluid Flow 5(3), 251-269, 1995.
53. M. Griebel, C. Zenger, S. Zimmer, *Multilevel Gauß-Seidel-algorithms for full and sparse grid problems*, Computing, 50, 127–148, 1995.
54. K.E. Gustavson, D. Rao, *Numerical Range: The field of values of linear operators and matrices*, Springer Verlag, New York, 1997.
55. D. Heath, M. Schweizer, *Martingales versus PDEs in finance: An equivalence result with examples*, to appear in Journal of Applied Probability, 2000.
56. T. K. Ho and E. M. Kleinberg, Checkerboard dataset, 1996. http://www.cs.wisc.edu/math-prog/mpml.html.
57. R. Hochmuth, S. Knapek, G. Zumbusch, *Tensor products of Sobolev spaces and applications*, Technical Report SFB 256, Univ. Bonn, 2000.
58. R.W. Hockney, J.W. Eastwood, *Computer simulation using particles*, IOP Publising Ltd., London, 1988.
59. W. Huber, *Turbulenzsimulation mit der Kombinationsmethode auf Workstation-Netzen und Parallelrechnern*, Dissertation, Institut für Informatik, TU München, 1996.
60. S. Ijima, *Helical microtubules of graphitic carbon*, Nature 354, 56-58, 1991.
61. L. Kaufman, *Solving the quadratic programming problem arising in support vector classification*, in Advances in Kernel Methods - Support Vector Learning, B. Schölkopf, C. J. C. Burges, and A. J. Smola, eds., MIT Press, 1999, pp. 146–167.
62. S. Knapek, *Approximation und Kompression mit Tensorprodukt–Multiskalen–Approximationsräumen*, Dissertation, Universität Bonn, Institut für Angewandte Mathematik, 2000.

63. R. Korn, *Optimal Portfolios*, World Scientific, 1997.
64. F. Koster, *A proof of the consistency of the finite difference technique on sparse grids*, Computing, 2000, to appear, auch als Report SFB 256, No. 642, Univ. Bonn, 2000.
65. Y.J. Lee, O. L. Mangasarian, *SSVM: A smooth support vector machine for classification*, Technical Report 99-03, Data Mining Institute, University of Wisconsin, Computational Optimization and Applications, 1999.
66. G. Melli, *Datgen: A program that creates structured data*, www.datasetgenerator.com.
67. E. Novak, K. Ritter, *High dimensional integration of smooth functions over cubes*, Numer. Math., 75, 79–98, 1996.
68. E. Novak, I.H. Sloan, H. Woźniakowski, *Tractability of tensor product linear operators*, J. Complexity 13, 4, 387–418, 1997.
69. K. Petras, *Asymptotically minimal Smolyak cubature*, Report Technische Universität Braunschweig, 1999.
70. A. Pfaffinger, *Funktionale Beschreibung und Parallelisierung von Algorithmen auf dünnen Gittern*, Dissertation, Institut für Informatik TU München, 1997.
71. C. Pflaum, *Diskretisierung elliptischer Differentialgleichungen mit dünnen Gittern.* Dissertation, Institut für Informatik, TU München, 1996.
72. C. Pflaum, A. Zhou, *Error analysis of the combination technique*, Numer. Math. 84, 2,327-350, 1999.
73. A. Rhodius, *On the maximum of ergodicity coefficients, the Dobrushin ergodicity coefficient and products of stochastic matrices*, Linear Algebra Appl. 253, 141-154, 1997.
74. A. Rhodius, *On explicit forms for ergodicity coefficients*, Linear Algebra Appl. 194, 71-83, 1993.
75. A. Rhodius, *On almost scrambing stochastic matrices*, Linear Algebra Appl. 126, 79-86, 1989.
76. U. Rüde, *Zur numerischen Behandlung von Singularitäten in elliptischen partiellen Differentialgleichungen*, Dissertation, Institut für Informatik TU München, 1988.
77. S. L. Salzberg, *On comparing classifiers: Pitfalls to avoid and a recommended approach*, Data Mining and Knowledge Discovery 1, 317–327, 1997.
78. T. Schiekofer, *Die Methode der Finite Differenzen auf dünnen Gittern zur Lösung elliptischer und parabolischer partieller Differentialgleichungen*, Dissertation, Inst. für Angewandte Mathematik, Universität Bonn, 1998.
79. S. Schneider, Dissertation, Technische Universität München, 2000.
80. C. Schütte, W. Huisinga, P. Deuflhard, *Transfer operator approach to conformational dynamics in biomolecular systems*, Konrad-Zuse-Zentrum Berlin, Preprint SC 99-36, 1999.
81. E. Seneta, *Non-negative matrices and Markov chains*, Springer Series in Statistics, Springer Verlag, 1981.
82. S.A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR, 148, 1042–1043, 1963, Russian, Engl. Transl.: Soviet Math. Dokl. 4:240–243, 1963.
83. A. Steinbauer, *Quadraturformeln für das Wienermaß*, Dissertation, FAU Erlangen-Nürnberg, 2000.

84. T. Störtkuhl, *Ein numerisches adaptives Verfahren zur Lösung der biharmonischen Gleichung auf dünnen Gittern*, Dissertation, Institut für Informatik TU München, 1995.
85. M. Stone, *Cross-validatory choice and assessment of statistical predictions*, Journal of the Royal Statistical Society 36, 111–147, 1974.
86. A.N. Tikhonov, V.A. Arsenin, *Solutios of ill-posed problems*, W.H. Winston, Washington D.C., 1977.
87. J. Traub, S. Paskov, *Faster evaluation of financial derivatives*, Journal of Portfolio Management 22, 1, 113-120, 1995.
88. V. N. Vapnik, *The nature of statistical learning theory*, Springer, 1995.
89. G. Wahba, *Spline models for observational data*, Series in Applied Mathematics, Vol. 59, SIAM, Philadelphia, 1990.
90. A. G. Werschulz, *The complexity of the Poisson Problem for spaces of bounded mixed derivatives*, The Mathematics of Numerical Analysis, J. Renegar and M. Shub and S. Smale, American Mathematical Society, Lectures in Applied Mathematics 32, 895–914, 1996.
91. H. Yserentant, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412, Vorversion als Bericht 21, Institut für Geometrie und Praktische Mathematik RWTH Aachen, 1983.
92. H. Yserentant, *Hierarchical bases*, in Proc. ICIAM'91, Washington, 1991, R. E. O'Malley, Jr. et al., eds., SIAM, Philadelphia, 1992.
93. C. Zenger, *Verallgemeinerte Wertevorräte von Matrizen*, Dissertation, Fakultät für Allgemeine Wissenschaften der TU München, 1967.
94. C. Zenger, *On convexity properties of the Bauer field of values of a matrix*, Numerische Mathematik 12, 96-105, 1968.
95. C. Zenger, *A comparison of some bounds for the nontrivial eigenvalues of stochastic matrices*, Numerische Mathematik 19, 209-211, 1972.
96. C. Zenger, *Minimale subadditive Einschließungsgebiete für die Eigenwerte von Matrizen*, Habilitationsschrift, Fakultät für allgemeine Wissenschaften der TU München, 1975.
97. C. Zenger, H. Gietl, *Improved Difference Schemes for the Dirichlet Problem of Poissons Equation in the Neighborhood of Corners*, Numerische Mathematik 30, 315-332, 1978.
98. C. Zenger, *Dual operator norms and the spectra of matrices*, Linear Algebra Appl. 58, 453-460, 1984.
99. C. Zenger, *Positivity in complex spaces and its application to Gershgorin discs*, Numerische Mathematik 44, 67-73, 1984.
100. C. Zenger, *Sparse grids*, Parallel Algorithms for Partial Differential Equations, *Notes on Num. Fluid Mech. 31*, W. Hackbusch (ed.), Vieweg, Braunschweig, 1991.
101. L. Zink, *Numerische Behandlung von Eckensingularitäten der biharmonischen Gleichung*, Dissertation, TU München, 1983.