Schwerpunkt: Analyse von Algorithmen

Tätigkeitsbericht für das Jahr 2001

---

1 Mitarbeiter

1. Michael *Drmota* (Inst. f. Geometrie, Koordinator)
2. Michael *Fuchs* (FWF-Projekt S-8302 MAT, Doktorand)
3. Bernhard *Gittenberger* (Inst. f. Geometrie, bis 30.9.2001 Zivildienst)
4. Thomas *Klausner* (FWF-Projekt S-8302 MAT, Doktorand 20 Std, Werkverträge)
5. Alois *Panholzer* (Inst. f. Algebra u. Computermathematik)
6. Wolfgang *Steiner* (FWF-Projekt S-8302 MAT, Doktorand)

2 Wissenschaftliche Publikationen

Im ersten Jahr des Forschungsschwerpunkts (2000) sind bereits 8 Arbeiten erschienen, die von Mitarbeitern des Forschungsschwerpunkts geschrieben wurden, siehe den Tätigkeitsbericht 2000.
Im zweiten Jahr (2001) sind die folgenden Arbeiten erschienen:

[1] M. Drmota, An analytic approach to the height of binary search trees, *Algorithmica* **29** (2001), 89–119.

[2] M. Drmota, The asymptotic number of leftist trees, *Algorithmica* **31** (2001), 304–317.

[3] M. Drmota, The joint distribution of q-additive functions, *Acta Arith.* **100** (2001), 17–39.

[4] M. Drmota, D. Gardy, and B. Gittenberger, A unified presentation of some urn models, *Algorithmica.* **29** (2001), 120–147.

[5] M. Drmota and G. Larcher, The sum-of-digits-function and uniform distribution modulo 1, *J. Number Theory* **89** (2001), 65–96.

[6] C. Martinez, A. Panholzer and H. Prodinger, *Partial match queries in relaxed multidimensional data structures*, Algorithmica **29** (2001), 181–204.

[7] A. Panholzer and H. Prodinger, Kirkman's hypothesis revisited. *Integers*, **1**: paper A05, 2001.

Die folgenden Arbeiten sind bereits zur Publikation angenommen worden.

[8] M. Drmota, The variance of the height of binary search trees, *Theoret. Computer Science* **270** (2002), 913–919.

[9] M. Drmota, The variance of the height of digital search trees, *Acta Inf.***38** (2002), 261–276.

[10] B. Chauvin, M. Drmota, and J. Jabbour-Hattab The profile of binary search trees, *Ann. Appl. Prob.*, im Erscheinen.

[11] M. Drmota, An analytic approach to the height of binary search trees. II., *JACM*, im Erscheinen.

[12] M. Drmota and W. Szpankowski, Generalized Shannon Code Minimizes the Maximal Redundancy, *Proceedings LATIN 2002*, im Erscheinen.

[13] M. Drmota and W. Szpankowski, The Precise Minimax Redundancy, 2002 International Symposium on Information Theory, Lausanne 2002, im Erscheinen.

[14] M. Drmota, H.-K. Hwang and W. Szpankowski, Precise Average Redundancy of an Idealized Arithmetic Coding, Data Compression Conference, Snowbirds, 2002, im Erscheinen.

[15] M. Drmota and D. Panario, A Rigorous Proof of the Waterloo Algorithm for the Discrete Logarithm Problem, *Design, Codes and Cryptography*, 2002, im Erscheinen.

[16] M. Drmota and J. Schoissengeier, Digital Expansions with Respect to Different Bases, *Monatsh. Math.*, im Erscheinen.

[17] M. Drmota and W. Steiner, The Zeckendorf Expansion of Polynomial Sequences, *J. Theorie Nombres Bodx.*, im Erscheinen

[18] M. Fuchs, Digital Expansion of Exponential Sequences, *Journal de Theorie des Nombres de Bordeaux*, im Erscheinen.

[19] M. Fuchs, On Metric Diophantine Approximation in the Field of Formal Laurent Series, *Finite Fields and Their Applications*, im Erscheinen.

[20] A. Panholzer and H. Prodinger, Moments of level numbers of leaves in binary trees, *Journal of Statistical Planning and Inference* **101** (2002), 267–279.

[21] J.-F. Marckert and A. Panholzer, Noncrossing trees are almost conditioned Galton-Watson trees, *Random Structures & Algorithms* **20** (2002), 115–125.

[22] A. Panholzer, On generalized Fibonacci Permutations. *Journal of Information & Optimization Sciences*, im Erscheinen.

[23] A. Panholzer, The height distribution of non-crossing trees, *Ars Combinatoria*, im Erscheinen.

[24] W. Steiner, Generalized de Bruijn digraphs and the Distribution of Patterns in α-Expansions, *Discrete Mathematics*, im Erscheinen.

Des weiteren wurden folgende Manuskripte fertiggestellt und zur Publikation in wissenschaftlichen Zeitschriften eingereicht.

[25] M. Drmota and J.-F. Marckert, Reinforced weak convergence of stochastic processes, Statist. Probab. Letters, zur Publikation eingereicht.

[26] M. Drmota and H. Prodinger, The Height of q-Binary Search Trees, *DMTCS*, zur Publikation eingereicht.

[27] M. Drmota, D. Gardy and B. Gittenberger, General urn models with several types of balls and Gaussian limiting fields, *Random Struct. Algorithms*, zur Publikation eingereicht.

[28] M. Drmota, M. Fuchs and E. Mastavicius, Functional Limit Theorems for Digital Expansions, *Acta Math. Hung.*, zur Publikation eingereicht.

[29] M. Drmota and B. Gittenberger, The width of Galton-Watson trees, *J. Combinat. Theory*, zur Publikation eingereicht.

[30] M. Drmota and Th. Klausner, The Distribution of Patterns in Random Trees, zur Publikation eingereicht.

[31] M. Fuchs, On a Problem of W. J. LeVeque concerning Metric Diophantine Approximation, *Trans. Amer. Math. Soc.*, zur Publikation eingereicht.

[32] M. Fuchs, Invariance Principles in Metric Diophantine Approximation, *Monatsh. Math.*, zur Publikation eingereicht.

[33] M. Fuchs, A Theorem of Szüsz for Formal Laurent Series over a Finite Base Field, *J. Number Theory*, zur Publikation eingereicht.

[34] B. Gittenberger, On the profile of random forests, *Mathematics and Computer Science*, Proceedings, Versailles, September 2002, zur Publikation eingereicht.

[35] M. Riedel, A. Panholzer and H. Prodinger, Measuring post-quickselect disorder, zur Publikation eingereicht.

[36] A. Panholzer, Gröbner bases and the defining polynomial of a context-free grammar generating function. zur Publikation eingereicht.

[37] A. Panholzer and H. Prodinger, A generating functions proof of a curious identity, zur Publikation eingereicht.

[38] A. Panholzer, Analysis of Multiple Quickselect variants, zur Publikation eingereicht.

[39] W. Steiner, Parry Expansions of Polynomial Sequences, *manuscr. math.*, zur Publikation eingereicht.

3 Vorträge

3.1 Michael Drmota

1. *The Height of Binary and Digital Search Trees*, Finite and Infinite Combinatorics, Budapest, Ungarn, 8. 1. 2001.

2. *Digital Expansions of Polynomial Sequences*, Workshop: Dynamic Odyssee, Luminy, Marseille, Frankreich, 15. 2. 2001.

3. *The Height of Binary and Digital Search Trees*, Purdue University, Indiana, USA, 28. 2. 2001.

4. *The Height of Digital Search Trees*, Seventh Seminar on Analysis of Algorithms, Tatihou, France. 1.–7. 7. 2001

5. *The Moment Method for the Sum-of-Digits Function*, Université de Provence, Marseille, 12. 7. 2001.

6. *Travelling Waves and the Height of Binary Search Trees*, INRIA, Rocquencourt, Frankreich, 24. 9. 2001.

7. *The Shape and Height of Binary Search Trees*, TU Delft, Niederlande, 17. 10. 2002.

3.2 Michael Fuchs

8. *Digital Expansion of Exponential Sequences*, Finite and Infinite Combinatorics, Budapest, Ungarn, 8. 1. 2001.

9. *Digital Expansion of Exponential Sequences*, L'odyssée Dynamique, Marseille, Frankreich, 15. 2. 2001.

10. *Grenzwertsätze in der metrischen diophantischen Approximation* ,ÖMG-Kongress, Wien, Österreich, 17. 9. 2001.

11. *Functional Limit Theorems and Digital Expansion*, Analytic and Probabilistic Number Theory, Palanga, Litauen, 25. 9. 2001.

12. *Metrische Kettenbruchtheorie und Diophantische Approximation*, Wissenswertes aus der Mathematik, eingeladener Vortrag, Wien, Österreich, 29. 10. 2001.

3.3 Thomas Klausner

13. *Grenzverteilung von Mustern in Bäumen*, ÖMG-Kongress, Wien, 20. 9. 2001.

3.4 Alois Panholzer

14. *About two combinatorial problems: Height distribution of non-crossing trees and Generalized Fibonacci permutations*, Universität Versailles, 22. März 2001.

15. *Kombinatorische Analyse von Algorithmen und Datenstrukturen*, Technischen Universität Wien, Mai 2001.

3.5 Wolfgang Steiner

16. *Generalized de Bruijn Graphs and Digital Expansions*, Finite and Infinite Combinatorics, Budapest, Ungarn, 8. 1. 2001.

17. *Markov Partitions of Toral Automorphisms and Digital Expansions*, L'odyssée Dynamique, Marseille, Frankreich, 15. 2. 2001.

18. *Digital Expansions and Rauzy Fractals*, Fractals, Graz, 7. 6. 2001.

19. *Die Verteilung der Ziffernsummenfunktion auf polynomiellen Folgen für G-adische Ziffernsysteme*, 15. ÖMG-Kongress, Wien, 17. 9. 2001.

20. *Central limit theorems for G-additive functions*, Analytic and Probabilistic Number Theory, Palanga, Litauen, 25. 9. 2001.

4 Tagungenbesuche

1. *Finite and Infinite Combinatorics*, Budapest, Ungarn, 5.-10. 1. 2001
Teilnehmer: *M. Drmota, M. Fuchs, W. Steiner*

2. L'odyssee Dynamique, Luminy, Marseille, Frankreich, 4.-17. 2. 2001
Teilnehmer: *M. Drmota, M. Fuchs, W. Steiner*

3. *Journees ALEA de Luminy 2001*, Marseille-Luminy, Frankreich, 26.-30. 3. 2001
Teilnehmer: *Th. Klausner*

4. *Fractals in Graz 2001*, Graz, 6. 2001
Teilnehmer: *M. Fuchs, W. Steiner*

5. *Seventh Seminar on Analysis of Algorithms*, Tatihou, Frankreich, 1.-7. 7. 2001
Teilnehmer: *M. Drmota, Th. Klausner*

6. *15. MG-Kongress*, Wien, 16.-22. 9. 2001
Teilnehmer: *M. Drmota, M. Fuchs, B. Gittenberger, Th. Klausner, W. Steiner*

7. *Analytic and Probabilistic Number Theory*, Palanga, Litauen, 24.-28. 9. 2001
Teilnehmer: *M. Fuchs, W. Steiner*

5 Forschungsaufenthalte

1. Thomas *Klausner*: INRIA Rocquencourt, 15. 1.–15. 4. 2001 (unterstẗzt durch ein Stipendium des ÖAD fuer kurzfristige wissenschaftliche Arbeiten im Ausland).

2. Michael *Drmota*: Purdue University, 22.-29.2.2001 (Einladung von W. Szpankowski).

3. Alois *Panholzer*: Université de Versailles, 12.-24. 3. 2001 (im Rahmen des Amadee-Projekts).

4. Michael *Drmota*: Université de Provence, Marseille, 8.–21. 7. 2001 (Einladung von P. Liardet).

5. Michael *Drmota*: Université de Versailles, 23.–29. 9. 2001 (im Rahmen des Amadee-Projekts).

6. Michael *Drmota*: TU Delft, 16.–21. 10. 2001 (Einladung von C. Kraaikamp).

7. Bernhard *Gittenberger*: Université de Versailles, 25. 10.–3. 11. 2001 (im Rahmen des Amadee-Projekts).

6 Internationale Kooperationen im Jahr 2001

1. Helmut *Prodinger* (University of Witwatersrand, Südafrika): intensive Kooperation mit A. Panholzer (Arbeiten [6. 7. 20, 35, 37]) und mit M. Drmota (Arbeit [26]).

2. Philippe *Flajolet*, Bruno *Salvy* (INRIA, Rocquencourt): Forschungsaufenthalt von Thomas Klausner (Automatische Bestimmung von Grenzverteilungen mit Maple); Tagungsorganisationen gem. mit Michael Drmota.

3. Brigitte *Chauvin*, Danielle *Gardy*, Jean-Francois *Marckert* (Université de Versailles): Amadee-Projekt des ÖAD, Arbeiten [4, 10, 21, 25, 27].

4. Wojciech *Szpankowski* (Purdue University): Forschungsaufenthalt von Michael Drmota, Tagungsorganisation, Arbeiten [13, 14, 15].

5. Pierre *Liardet* (Univesité de Provence, Marseille): Forschungsaufenthalt von Michael Drmota.

6. Hsien-Kuei *Hwang* (Academica Sinica, Taipei): Gast an der TU Wien, Arbeit [14].

7. Eugenijus *Manstavicius* (Universität Vilnius, Litauen): Gast in Wien im Jahr 2000, Arbeit [28].

7 Gäste an der TU-Wien

1. Jean-François Marckert, Université de Versailles, 7.–21. 4. 2001 (im Rahmen des Amadee-Projects)

2. Philippe *Duchon*, Université de Bordeaux, 17.–23. 6. 2001 (im Rahmen des FWF-Projekts S8302).

3. Danièle *Gardy*, Université de Versailles, vom 13.–23. 5. 2001 (im Rahmen des Amadee-Projects).

4. Hsien-Kuei *Hwang*, Academia Sinica, Taipeh, Taiwan, vom 10.–16. 8. 2001 (im Rahmen des FWF-Projekts S8302).

5. Brigitte *Chauvin*, Université de Versailles, vom 8.–14. 12. 2001 (im Rahmen des Amadee-Projects).

8 Tagungsorganisation

8.1 15. ÖMG-Kongress, Wien, 16.-22. 9. 2001.

Michael Drmota war Mitglied des Organisationskomitees des 15. ÖMG-Kongresses, an dem etwa 500 Mathematiker teilnahmen. Auch Wolfgang Steiner war bei der Tagungsorganisation beteiligt.
Weiters wurde im Rahmen des Kongresses ein Minikolloquium mit dem Titel *Zahlentheoretische Algorithmen* von Gerhard Larcher (Linz) und Robert Tichy (Graz) organisiert, um den FWF-Forschungsschwerpunkt (FSP) *Numbertheoretic Algorithms and Applications*, dem auch das von Michael Drmota geleitete FWF-Projekt *Statistical Properties of Number Systems* (S 8302) angehört, zu präsentieren.

8.2 Analysis of Algorithms, Strobl, 23.-29. 6. 2002

Michael Drmota ist Vorsitzender des Programm- und Organisationskomitees der internationalen Tagung *Eighth Seminar on Analysis of Algorithms*, die vom 23.-29. 6. 2002 in Strobl stattfinden wird. Diese Tagung wird die achte einer Reihe von Tagungen sein, die 1993 in Dagstuhl (BRD) begonnen hat.
Die Vorarbeiten laufen bereits seit Mitte 2000. Es werden etwa 80 Teilnehmer erwartet.
Das Thema der Tagung ist entspricht im wesentlichen den Zielsetzungen dieses Schwerpunkts und ist daher ein wichtiger Bestandteil.

8.3 Mathematics and Computer Science, Versailles, Juni 2002

Viele Themenstellungen dieser Tagung (z.B. Trees, Algorithms) finden sich auch im Schwerpunkt wieder.
Michael Drmota ist Mitglied des Programmkomitees, auch hier haben die Vorarbeiten bereits im Jahr 2001 begonnen.

9 Forschungsprojekte

9.1 FWF-Projekt S-8302, 2000-2005: Statistical Properties of Number Systems

Das von Michael Drmota geleitete FWF-Projekt S-8302 ist ein Teilprojekt des fünfjährigen FWF-Forschungschwerpunktes *Numbertheoretic Algorithms and Applications*, der im Mai 2000 begonnen wurde.

Im Rahmen dieses Projekts waren im Jahr 2001 2 1/2 wissenschaftliche Mitarbeiter beschäftigt: Michael Fuchs, Thomas Klausner (50 %), Wolfgang Steiner.

9.2 Amadee-Projekt des ÖAD 2000–2001: Probabilistic Analysis of Data Structures

Im November 2000 wurde das (im Mai 2000 eingereichte) Projekt *Probabilistic Analysis of Algorithms* im Rahmen des Programms *Amadee 2001-2002* des Wissenschaftlich-Technischen Abkommens mit Frankreich genehmigt (Partneruniversität: Universität Versailles-St. Quentin).
Dieses Kooperationsabkommen ermöglicht den österreichischen Teilnehmern Michael Drmota, Bernhard Gittenberger und Alois Panholzer in den Jahren 2001 und 2002 jeweils 2 Wochen die Universität Versailles-St. Quentin zu besuchen. Entsprechend werden die französischen Teilnehmer Brigitte Chauvin, Daniele Gardy und Jean-Francois Marckert die TU-Wien besuchen.
Diese Kooperation ist schon sehr erfolgreich, so sind etwa die gemeinsamen Arbeiten [4, 10, 21, 27] daraus hervorgegangen.
Es ist geplant, eine Verlängerung dieses Projekts um weitere zwei Jahre zu beantragen.

10 Diplomarbeiten, Dissertationen, Habilitationen

10.1 Diplomarbeiten

1. Andreas *Traxler*: *Höhe und interne Pfadlänge in verschiedenen Baumklassen*, begonnen im Oktober 2001, betreut durch Michael Drmota.

10.2 Dissertationen

1. Christian *Brandner*: *"Uber periodische Kettenbrüche erzeugende Folgen*, abgeschlossen im Mai 2001, betreut durch Michael Drmota.

2. Michael *Fuchs*: *Limit Theorems in Metric Diophantine Approximation*, abgeschlossen im Dezember 2001, Rigorosum im März 2002, betreut durch Michael Drmota.

3. Wolfgang *Steiner*: *The Distribution of Digital Expansions on Polynomial Sequences*, abgeschlossen im Dezember 2001, Rigorosum im März 2002, betreut durch Michael Drmota.

4. Thomas *Klausner*: *Automatic Gaussian Limiting Distributions with Maple*, begonnen im Mai 2000, betreut durch Michael Drmota.

10.3 Habilitationen

1. Bernhard *Gittenberger*: Diskrete Mathemati, TU Wien, Institut für Geometrie, Juni 2001.

11 Vorsorge für die Verwertung

Alle Ergebnisse dieses Forschungsbereiches werden in entsprechenden wissenschaftlichen Zeitschriften publiziert (siehe oben). Weiters werden laufend facheinschlägige Tagungen besucht und bei Vorträgen über die wissenschaftlichen Ergebnisse berichtet.

Ein wesentlicher Punkt ist auch die Organisation von Tagungen, da hier auf das Programm entsprechend Einfluß genommen werden kann. Vor allem trifft dies auf die für den Juni 2002 geplante Tagung *Analysis of Algorithms* (in Strobl) zu, bei der Michael Drmota Vorsitzender des Programm- und Organisationskomitees ist.

12 Impact für praxisorientierte und technikrelevante Anwendungen

Die Zielsetzung dieses Schwerpunkts ist es, Algorithmen aus den verschiedensten Anwendungen (z.B. Algorithmen für Datenstrukturen, zahlentheoretische Algorithmen) zu analysieren, und zwar im Hinblick auf das durchschnittliche Verhalten.
Auf diese Weise können verschiedene Algorithmen (etwa Sortieralgorithmen) und deren Varianten direkt miteinander verglichen werden. Dies ermöglicht einem Anwender, aus den verschiedenen Möglichkeiten einer Problemlösung die für ihn optimale Variante zu wählen und gleichzeitig den zu erwartenden Aufwand abschätzen zu können.
Ein wichtiger Punkt des Schwerpunktes ist es mit Hilfe von Maple-Prozeduren, die in der Literatur beschriebenen theoretischen Ergebnisse praktisch anwendbar zu machen. Die Aufgabe von Thomas Klausner (der auch beim FWF-Projekt S8302 MAT zu 50 % angestellt ist) ist es, diese Maple-Prozeduren zu programmieren.

13 Wissenschaftlicher Fortschritt

Der Schwerpunkt **Analyse von Algorithmen** kann grob in zwei Teilbereiche eingeteilt werden.

1. Probabilistische und kombinatorischen Analyse von Algorithmen
2. Zahlentheoretische Algorithmen: Ziffernentwicklungen

Im folgenden werden die Forschungaktivitäten des Jahres 2000 in diesen beiden Bereichen beschrieben und dokumentiert.

13.1 Probabilistische und kombinatorischen Analyse von Algorithmen

Dieses Gebiet beschäftigt sich mit der mathematischen Analyse von Algorithmen, die mit Datenstrukturen in Zusammenhang stehen, z.B. mit binären Suchbäumen, mit relationalen Datenbanken etc.
Diesem Teil sind die Arbeiten [1, 2, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 20, 21, 22, 23, 25, 26, 27, 29, 30, 34, 35, 36, 37, 38] zuzuordnen.

Eine der wichtigsten Datenstrukturen zum Sortieren und Suchen sind die **binären Suchbäume**. Sie treten z.B. (implizit) beim Sortieralgorithmen **Quicksort** auf. (Die Analyse von Quicksort ist daher untrennbar mit der Analyse von binären Suchbäume verbunden.) Die probabilistische Analyse von binären Suchbäumen ist auch ein wesentilicher Teil dieses Schwerpunkts.
So wird etwa in der Arbeit [10] ein (fast sicherer) Grenzwersatz für das *Profil* von binären Suchbäumen bewiesen. (Das Profil eines Baumes ist die Folge L_k der Anzahlen der Knoten im Niveau k und beschreibt daher die Gestalt des Baumes.) Das Hauptresultat besagt, daß sich das Profil fast sicher wie das durchschnittliche Profil verhält, das ganz genau angegeben werden kann (etwa mit Hilfe der Stirlinzahlen

1. Art). Es stellt sich heraus, daß der Quotient zwischem dem Profil und dem erwarteten Profil asymptotisch durch den (fast sicheren) Grenzwert eines Martingals beschrieben werden kann. Eine Folgerung daraus ist ein (fast sicherer) Grenzwertsatz für die *Weite* $\max_k L_k$.
Ein anderer äußerst interessanter Parameter ist die Höhe von binären Suchbäumen. Die Höhe entspricht genau der Anzahl der Teilungsschritte in Quicksort. Es ist seit langem bekannt, daß die durchschnittliche Höhe ungefähr $4.3 \cdot \log n$ ist (wobei n die Anzahl der internen Knoten bezeichnet). Es wurde auch vermutet, daß die Varianz der Höhe beschränkt bleibt. Diese Eigenschaft ist ein wenig unerwartet, da sie ja besagt, daß die Verteilung der Höhe extrem konzentriert ist. Diese Vermutung konnte in der Arbeit [8] bewiesen werden, Vorarbeiten dazu wurden bereits in [1] geleistet. (Eine davon unaghängige Lösung stammt von Bruce Reed.) Entsprechende Erweiterungen für die minimale Höhe, die Höhe von m-ären Suchbäume und auch Darstellungen für die Verteilung der Höhe findet man in [11].
Interessanterweise konnte auch gezeigt werden, daß die Varianz von digitalen Suchbäumen beschränkt bleibt ([9]). Der Unterschied zwischen binären und digitalen Suchbäumen ist, daß bei binären Suchbäumen nur die (lineare) Ordnung der Schüssel (der Daten) bei der Links-Rechts-Einteilung im Binärbaum verwendet wird. Bei digitalen Suchbäumen hat jeder Datensatz einen binären Schlüssen, der die Linkt-Rechts-Einteilung entscheidet. Analytisch äußert sich das so, daß bei binären Suchbäumen gewönliche erzeugende Funktionen und bei digitalen Suchbäumen exponentielle erzeugende Funktionen zum Einsatz kommen, wodurch bei der Analyse unterschiedliche Methoden angewandt werden müssen. Im ersteren Fall ist man mit einem endlichen Konvergenzradius konfrontiert und im zweiteren mit ganzen Funktionen.
Es gibt zahlreiche Varianten von Quicksort, wie z.B. *Quickselect*, bei dem nur das größte oder kleinste Element gesucht werden, bzw. das p-t größte oder p-t kleinste Element. Zur Beschleunigung wird auch oft die *median of 3-* oder allgemeiner *median of* $2t+1$-Variante verwendent, um den dazugehörigen binären Suchbaum stärker zu balanzieren. In den Arbeiten [35] und [38] werden verschiedene Parameter solcher Quickselect-Varianten untersucht.
Ein ganz anderern Aspekt wird in der Arbeit [26] beleuchtet. Hier wird anstelle des üblichen probabilisten Modells für binäre Suchbäume ein q-Variante betrachtet, die im Grenzfall $q \to 0$ den klassischen Fall ergibt. Im Gegensatz zum klassischen Fall ist die Höhe bei positivem q normalverteilt mit Mittelwert qn und Varianz $q(1-q)n$.

Ein ganz anderer Datentyp sind **relationale Datenbanken**. In den Arbeiten [4, 27] werden diese durch **Urnenmodelle** modelliert und das dynamische Verhalten beschrieben. Es stellt sich heraus, daß das Verhalten um den Mittelwert herum (entsprechend normiert) durch einen Gaußschen Prozess beschrieben werden kann, im multivariaten Fall durch ein Gaußsches Feld.
Das einfachste Modell ist es, auf m (identischen) Urnen, n Bälle zufällig zu verteilen und man betrachtet z.B. die Anzahl der leeren Urnen in Abhängigkeit von m und n. Die Problemstellung wird zunächst mit Hilfe erzeugender Funktionen auf eine *analytische Ebene* gebracht und diese daraufhin mit Hilfe der (multivariaten) Sattelpunktsmethode weiterbehandelt.

In der Arbeit [6] werden **Mustererkennungen** in mehrdimensionalen Datenbanken untersucht, die typischerweise in großen Datenbanken auftreten und der Schlüssel natürlich auch mehrdimensional ist. Es soll auch erlaubt sein, daß nicht alle Komponenten des zu erkennenden Schlssels spezifiziert werden müssen (*partial match query*). Insbesondere werden K-dimensionale baumartige Datenstrukturen (search trees und digital tries), die Verallgemeinerungen der bekannten Kd-tries darstellen, behandelt, wobei der durchschnittliche Aufwand und die Varianz untersucht werden.

Neben den stark anwendungsorientierten Untersuchungen über baumartige Datenstrukturen, wurden auch einige Arbeiten über verschiedene Eigenschaften von **Zufallsbäumen** geschrieben.
Diese Modelle sind meist von den Anwendungen motiviert, wie z.B. die üblichen Binärbäume (die etwa im Zusammenhang mit *Dyck-paths* oder *Dyck-languages* auftreten, die als Spezialfall von **Galton-Watson-Verzweigungsprozessen** gesehen werden können, wobei man alle Realisierungen so eines Verzweigungsprozesses betrachtetn, wo die Gesamtalzahl der Teilchen n ist.
So wird etwa in der Arbeit [29] die *Weite* von Galton-Watson-Bäumen untersucht und gezeigt, daß die Momente der (entspechend normierten) Weite (wie schon lange vermutet) gegen die Momente des Maximums der totalen lokalen Zeit einer Brownschen Exkursion konvergieren. Grundlage dazu ist einerseits ein schwacher Grenzwertsatz (der in einer früheren Arbeit von Drmota und Gittenberger gezeigt wurde, nämlich daß die Weite schwach gegen das Maximum der totalen lokalen Zeit einer Brownschen Exkursion konvergiert) und entprend quantifizierte Straffheitsaussagen, die es erlauben aus der schwachen Konvergenz auch die Konvergenz der Momente abzuleiten.
Dieses Konzept wurde in der Arbeit [25] allmein für stochastische Prozesse entwickelt (und der Begriff der *polynomiellen Konvergenz* eingeführt). Anwendungen dafür sind nicht nur die Weite von Bäumen, sondern auch die Höhe, die Pfadlänge und alle möglichen anderen Parameter.
Die Arbeiten [21, 23] beschäftigen sich mit der Klasse der *non-crossing trees*, das sind planare Wurzelbäume deren Knoten auf einem Kreis liegen und deren Kanten sich nicht überschneiden. In der Arbeit [23] wird die Höhe solcher Bäume untersucht und festgestellt, daß sich die Kontur (entsprechend normiert) asymptotisch wie eine Brownsche Exkursion verhält (in ähnlicher Weise wie die Galton-Watson-Bäume). Diese Analogie wird dann in der Arbeit [21] tatsächlich begründet. Es stellt sich heraus, daß die non-crossing trees durch Galton-Watson-Verzweigungsprozesse beschrieben werden können, deren erste Generation ein etwas anderes Verteilungsgesetz hat wie die darauffolgenden.
Die Verteilung der Blätter in Binärbäumen (und allgemeiner in t-ären Bäumen) ist das Thema der Arbeit [20], in der geschlossene Formeln für die Momente des j-ten Blattes abgeleitet werden.
In der Arbeit [2] stehen ebenfalls spezielle Binärbaume (und auch allgemeiner spezielle Galton-Watson-Bäume) im Vordergrund, nämlich die *leftist trees*, die im Binärfall einer speziellen Datenstruktur entsprechen. Es wird für den allgemeinen Fall eine asymptotische Anzahlformel abgeleitet und gezeigt, daß gewisse Parameter (wie z.B. die Anzahl der Bätter) einen zentralen Grenzwertsatz erfüllen. Methodisch werden wieder erzeugende Funktionen in mehreren Verändelichen und die genaue Analyse der Singularitätenstruktur dieser Funktionen eingesetzt.
Eine ganz andere Frage wird in der Arbeit [30] behandelt, nämlich die Verteilung des Auftretens von vorgegebenen *Mustern* in Zufallsbäumen. Es wird gezeigt, daß die erwartete Anzahl eines Musters proportional zur Knotenanzahl ist und die Verteilung um diesen Mittelwert durch eine Gaußsche Verteilung approximiert werden kann. Wieder kommen erzeugende Funktionen (in zwei Veränderlichen zum Einsatz), von denen gezeigt wird, daß sie ein System von Funktionalgleichungen erfüllen. Eine Analyse der Singularitätenstruktur führ zu diesem Verteilungsresultat.
Schließlich werden in der Arbeit [34] *Zufallswälder* betrachtet, deren (Baum-)Komponenten Galton-Watson-Bäume sind. Es wird gezeigt, daß sich das Profil solcher Wälder (d.h. die Folge L_k der Anzahlen der Knoten im Niveau k) entsprechend normiert durch eine Brownsche Bewegung beschreiben läßt.

Bei der (kombinatorischen) Analyse von Algrithmen entstehen praktisch immer **Rekursionen** für die betrachteten Kenngrößen. Das (asmyptotische) Lösen von Rekursionen ist daher ein wesentlicher Bestandteil der Analyse von Algorithmen, und es ist nicht verwunderlich, daß die dafür entwickelten Methoden auch für Rekursionen angewandt werden können, die in anderen Zusammenhängen auftauchen.
Beispielsweise wird in der Arbeit [37] ein Beweis der Identität

$$(x+m+1)\sum_{i=0}^{m}(-1)^i\binom{x+y+i}{m-i}\binom{y+2i}{i}-\sum_{i=0}^{m}\binom{x+i}{m-i}(-4)^i=(x-m)\binom{x}{m}$$

mit Hilfe einer Doppelrekursion und erzeugenden Funktionen gegeben.
In der Arbeit [22] werden Permutationen σ mit der Eigenschaft $|\sigma(i)-i|\le N$ (für alle i und vorgegebenem $N\ge 1$) betrachtet. Die Anzal $F_{N,n}$ dieser Permutationen (der Länge n) erfüllen eine lineare Rekursion der Ordnung $\binom{2N}{N}$, woraus sich die asymptotische Beziehung $F_{n,N}\sim c_N\,\rho_N^n$ (mit $1/2\le\frac{\rho_N}{N+1}\le 1$) ableiten läßt.
Schließlich wird wird in [7] ein (neuer) kurzer Beweis der Kirkmanschen Vermutung mit Hilfe der Lagrange-Inversion gegeben.

Die Arbeiten [12, 13, 14] beschäftigen sich mit der **Redundanz von Codes**. Hier geht es darum, Codes zu konstruieren, die bezüglich einer gegebenen Klasse von Wahrscheinlichkeitsverteilungen für die Codewörter gleichmäßig kleinsmögliche Redundanz besitzen. (Die Redundanz ist die maximale (oder durchschnittliche) Differenz zwischen den tatsächligen Codelängen und den *optimalen* Codelängen, dem negativen Logarithmus der Wahrscheinlichkeiten.) Gleichzeitig wird diese bestmögliche Redundanz bis auf einen gegen 0 konvergierenden Fehlerterm genau angegeben.

In der Arbeit [36] werden **Kontextfreie Grammatiken** behandelt, und es wird mit Hilfe von **Gröbnerbasen** ein neuer Beweis angegeben, daß solche Grammatiken algebraisch sind. Gleichzeitig wird (eben mit Hilfe von Gröbnerbasen, die mit Hilfe des Buchbergeschen Algorithmus bestimmt werden) ein Algorithmus angegeben, die entsprechende algebraische Gleichung der erzeugenden Funktion eines Nonterminal-Symbols explizit zu ermitteln. Dieser Algorithmus stellt eine wirkliche Alternative zum Kuich-Salomaa-Eliminationsalgorithmus dar, der auf Resultanten beruht.

Ein wesentlicher Bestandteil dieses Schwerpunkts ist es auch **Maple-Prozeduren** zur automatischen Bestimmung von asymptotischen Grenzverteilungen und multivariaten asymptotischen Entwicklungen zu schreiben. Diese Aufgabe wird von *Thomas Klausner* (unter Anleitung von Michael Drmota) durchgeführt. (Thomas Klausner war im Jahr 2001 beim FWF-Projekt S-8302 halbbeschäftigt und bekam auch in den Jahren 2000 und 2001 einen Werkvertrag aus den Schwerpunktsmitteln.)
Für univarite asymptotische Entwicklungen existiert bereits ein umfangreiches Programm (equivalent), das von Bruno Salvy (INRIA, Rocquencourt) entwickelt wurde. Thomas Klausner hat 2001 einen dreimonatigen Forschungsaufenthalt bei Bruno Salvy (INRIA) absolviert, um auf dem Stand der Forschung aufsetzen zu können.
Bisher können asymptotische Entwicklungen bzw. Verteilungsresultate (zentrale Grenzwertsätze) für die Koeffizienten von erzeugenden Funktionen mit endlichem Konvergenzradius behaldelt werden. Derzeit wird an ganzen erzeugenden Funktionen gearbeitet. Dafür ist es notwendig, das Haymansche Konzept der *admissible functions* auf den multivariaten Fall zu erweitern.

13.2 Zahlentheoretische Algorithmen: Ziffernentwicklungen

Dieser Teil des Schwerpunkts wird durch das FWF-Projekt S-8302 MAT *Statistical Properties of Number Systems*, ein Teilprojekt des FWF-Schwerpunktes *Numbertheoretic Algorithms and Applications*, der im Mai 2000 begonnen wurde, bestritten. Im Rahmen dieses Projekts sind 2 1/2 wissenschaftliche Mitarbeiter beschäftigt: Michael Fuchs, Thomas Klausner (50 %), Wolfgang Steiner.
Diesem Teil sind die Arbeiten [3, 5, 15, 16, 17, 18, 19, 24, 28, 31, 32, 33, 39] zuzuordnen.

Zahlreiche Algorithmen basieren auf der (binären) Zifferndarstellung der natürlichen Zahlen. Andererseits stellt sich heraus, daß bei der theoretischen Analyse verschiedenster Algorithmen, insbesondere wenn ein periodisches Verhalten auftritt, die Ziffernentwicklung bzw. auch die Ziffernsumme von n (wobei n z.B. die Anzahl der Daten, die in einem Algorithmus verarbeitet werden sollen, bezeichnet) in das Ergebnis der Analyse eingeht.
Ziel des FWF-Projekts *Statistical Properties of Number Systems*, ist es daher, das statistische Verhalten von verschiedenen Ziffernentwicklungen genauer zu untersuchen, um diverse Algorithmen (bzw. die Resultate der theoretischen Analyse) besser verstehen zu können.
In einer Reihe von Arbeiten [3, 16, 17, 24, 28, 39] wird im wesentlichen die **Ziffernsummenfunktion** (aber auch allgemeinere sogenannte additive Funktionen) untersucht, wobei das Hauptaugenmerk auf Verteilungseigenschaften, also auf zentrale Grenzwertsätze, gelegt wird.
In den Arbeiten [3] und [16] geht es um die Korrelationen von Ziffernentwicklungen bezüglich verschiedener Basen. In [3] wird gezeigt, daß Ziffernentwicklungen zu teilerfremden Basen q_1, q_2 asymptotisch unanbhängig sind, wenn man über die Zahlen $1 \leq n \leq N$ mittelt und den Fall $N \to \infty$ betrachtet. In [16] wird hingegen N festgehalten und über alle Basen $2 \leq q \leq Q$ gemittelt. Die eingesetzten Methoden stammen aus der Wahrscheinlichkeitstheorie und insbesondere aus der klassischen analytischen Zahlentheorie, nämlich um Abschätzungen von Exponentialsummen.
Die Arbeiten [17, 24, 39] behandeln verallgemeinerte Ziffernentwicklungen, z.B. die Zeckendorf-Entwicklung, wo jede natürliche Zahl als Summe von Fibonacci-Zahlen dargestellt wird. Allgemeiner wird als *Basis* anstelle der Potenzen q^k eine Folge G_k verwendet, die eine lineare Rekursion erfüllt (wie eben die Fibonaccizahlen). Entsprechend kann hier die Ziffernsummenfunktion etc. definiert und das statistische Verhalten solcher Funktionen untersucht werden. Es werden wieder zentrale Grenzwertsätze abgeleitet, wobei hier aber wesentlich aufwendigere Methoden als im *klassischen* q-adischen Fall eingesetzt werden müssen.
In der Arbeit [28] werden die oben erwähnten zentralen Grenzwersätze dahingehend verallgemeinert, daß die Ziffernentwicklugen durch Brownsche Bewegungen approximiert werden, also durch eine funktionale Version des zentralen Grenzwertsatzen. Entsprechend werden auch bekannte Gesetze des iterierten Logarithmus zu einem Strassen-Gesetz erweitert. Alle Resultate beziehen sich (im Unterschied zu Vorarbeiten) auf die gemeinsame Verteilung Ziffernentwicklungen zu verschiedenen Basen, außerdem werden nicht alle natürlichen Zahlen zugrundegelegt, sondern nur die Werte eines ganzzahligen Polynoms.
Schließlich werden in der Arbeit [18] statistische Eigenschaften von q-adischen Ziffernentwicklungen von exponentiellen Folgen a^n betrachtet. Mit Hilfe der Bakerschen Methode und Abschätzungen für Exponentialsummen wird gezeigt, daß die relative Häufigkeit mit der ein Ziffernblock in den ersten bzw. letzten Ziffern der ersten N Folgeglieder einer exponentiellen Folge auftritt, dem erwarteten Wert entspricht.

Ausgehend von Untersuchungen von Niederreiter über das Komplexitätsprofil von Folgen (über endlichen Körpern) hat Michael Fuchs in den Arbeiten [19] und [33] eine **metrische Kettenbruchtheorie und diophantische Approximation im Körper der formalen Laurentreihen** entwickelt. In [19] werden klassische Sätze der metrischen Theorie der diophantischen Approximation werden in den Körper der formalen Laurentreihen übertragen. In [33] wird ein zentraler Grenzwertsatz für die Anzahl der Lösungen der Ungleichung $|\alpha - P/Q| < f(|Q|)/|Q|^2$ abgeleitet.

Die vorangegangenen Untersuchungen haben auch das Interesse für klassische Fragestellungen der **klassischen metrischen diophantischen Approximation** in den reellen Zahlen geweckt. In den Arbeiten [31] und [32] werden alle bisherigen Ergebnisse weitgehend verallgemeinert und zu einem entsprechenden Abschluß gebracht. Das Hauptergebnis von [31] ist die Verallgemeinerung und der vollständige Beweis einer Vermutung von LeVeque aus dem Jahre 1959. In [32] werden klassische Sätze der metrischen diophantischen Approximation von LeVeque und Philipp werden verallgemeinert, indem fast sichere Invarianzprinzipien und Invarianzprinzipien in Verteilung abgeleitet werden. Diese Ergebnisse haben funktionale Versionen der Sätze von LeVeque und Philipp zur Folge.

In der Arbeit [5] werden Zusammenhänge der Ziffernsummenfunktion $s_q(n)$ (zur Basis q) und **Zufallszahlen** im mehrdimensionalen Fall aufgestellt. Es wird u.a. bewiesen, daß die Folge $(\alpha_1 s_{q_1}(n), \ldots, \alpha_k s_{q_k}(n))$ genau dann modulo 1 gleichverteilt ist (also eine Folge von Zufallszahlen generiert), wenn alle α_i irrational sind, wobei vorausgesetzt ist, daß die Basen q_i paarweise teilerfremd sind. Es werden auch Diskrepanzabschätzungen angegeben.

Schließ lich behandelt die Arbeit [15] den sogenannten Waterloo-Algorithmus zur Berechnung des **diskreten Logarithmus in endlichen Körpern**. Die bisherigen Aufwandsaufschätzungen dieses Algorithmus waren teilweise noch heuristisch. Diese Heuristiken konnten hier exaktifiziert werden.

14 Finanzen

Dem Schwerpunkt "Analyse von Algorithmen" wurden im Jahr 2001 insgesamt ATS 54.670,– außerordentliche Dotation und ATS 115.260,– ordentliche Dotation zugewiesen.
Von der außerordentlichen Dotation wurden insgesamt ATS 50.467,– ausgebeben, im Detail:

- 1 Netzwerk-Laser-Drucker zu ATS 17.172,–
- 2 PCs (ohne Bildschirm) zu ATS 33.295.–

Von der ordentlichen Dotation wurden insgesamt ATS 70.713,30 ausgebeben, im Detail:

- Reisemittel für Michael Drmota: ATS 6.300,–)
- 1 Bildschirm zu ATS 4.480,02,–
- Druckertoner zu ATS 2.935,–
- Bücher zu ATS 29.347,–
- Sachaufwand zu ATS 3.251,28
- Werkvertrag für Thomas Klauser zu ATS 24.000,–

Die Differenzbeträge werden im folgenden Projektjahr ausgeglichen werden.

Univ.Prof.Dr. Gerd Baron (Institutsvorstand)

Michael Drmota (Koordinator)

15 Nachtrag zum Tätigkeitsbericht 2000

Im Jahr 2000 wurden 8 Arbeiten publiziert, über deren Hauptergebnisse kurz berichtet werden soll. (Dieser Punkt fehlte im Tätigkeitsbericht 2000.)

[1] M. Drmota, *The saturation level in binary search trees,* in: Mathematics and Computer Science, Algorithms, Trees, Combinatorics and Probabilities (D. Gardy and A. Mokkadem eds), Birkhäuser, Basel, 2000, 41–51.

Es wird eine asymptotische Formel für den Erwartungswert der minimalen Höhe in binären Suchbäumen angegeben, wobei erstmals ein expliziter Fehlerterm angegeben werden konnte. Die Methoden basieren auf der genauen asymptotischen Analyse der Lösungen der Differentialgleichung $\Phi'(u) = -1/\alpha^2\Phi(u/\alpha)^2$.

[2] B. Gittenberger, *The number of descendants in simply generated random trees*, in: Mathematics and Computer Science, Algorithms, Trees, Combinatorics and Probabilities (D. Gardy and A. Mokkadem eds), Birkhäuser, Basel, 2000, 65–73.

In dieser Arbeit geht es um die Anzahl der Nachfolger eines Knotens in einem Zufallswurzelbaum. Das Interessante hier ist, daß zwar eine diskrete Grenzverteilung existiert, aber sowohl Erwartungswert (und alle höheren Momente) gegen unendlich divergieren, wobei das asymptotische Verhalten genau angegeben werden kann.

[3] B. Gittenberger and G. Louchard, *On the local time density of the reflecting Browinan bridge*, J. Appl. Math. Stoch. Analysis **13** (2000), 125–136.

Mit Hilfe zweier vollständig verschiedener Zugänge (über die Kacsche Formel und über erzeugende Funktionen) werden explizite Integraldarstellung der mehrdimensionalen Dichten der totalen lokalen Zeit der Brownschen Brücke angegeben.

[4] A. Panholzer and H. Prodinger, *Two proofs of Filipponi's formula for Lucas numbers of odd index*, The Fibonacci Quarterly **38** (2000), 165–166, 2000.

Es werden zwei Beweise von einer Darstellung der Lucas-Zahlen angegeben, einer mit Hilfe erzeugender Funktionen und ein zweiter mit Hilfe von Zeilbergers Algorithmus.

[5] M. Drmota and M. Skalba, *Rarified Sums of the Thue-Morse Sequence,* Trans. Amer. Math. Soc. **352** (2000), 609–642.

Ein bekannter Satz von Newman sagt aus, daß unter den Vielfachen von 3 mehr Zahlen gibt, deren binäre Ziffernsumme gerade ist als ungerade. In dieser Arbeit wird diese Eigenschaft auf alle Vielfachen von Zahlen der Form 2^k+1 verallgemeinert (für $k = 1$ erhält man das Newmansche Resultat wieder). Außerdem wird gezeigt, daß es nur sehr wenige Primzahlen gibt, die diese Eigenschaft haben.

[6] M. Drmota, *The distribution of patterns in digital expansions,* in: Algebraic Number Theory and Diphantine Analysis (F. Halter-Koch and R. F. Tichy eds.), de Gruyter, Berlin, 2000, 103–121.

Es wird gezeigt, daß in Ziffernentwicklungen bezüglich allgemeiner Basen (die eine lineare Rekursion erfüllen) das Auftreten eines vorgegebenen (erlaubten) Ziffernblocks einen zentralen Grenzwersatz erfüllt. Methodisch werden erzeugende Funktionen und Eigenschaften nicht-negativer Matrizen verwendet. Bemerkenswerterweise funktioniert diese Methode auch bei unendlichen linearen Rekursionen.

[7] M. Drmota and M. Skalba, *The parity of the Zeckendorf sum-of-digits-function,* manuscripta math. **101** (2000), 361–383.

In dieser Arbeit wird der soeben erwähnte Newmansche Satz für die Zeckendorfentwicklung bewiesen. Außerdem werden weitere ähnliche Fragestellungen für die Zeckendorfentwicklung untersucht.

[8] B. Gittenberger and J. Thuswaldner, *Asymptotic normality of b-additive functions on polynomial sequences in the Gaussian number field*, J. Number Theory **84** (2000), 317–341.

Interessanterweise gibt es auch in den Gaußschen ganzen Zahlen *natürliche* Ziffernentwicklungen, d.h. jede Gaußsche ganze Zahl z läßt sich in eindeutiger Form $z = \sum a_k b^k$ mit einem endlichen Ziffernvorrat $a_k \in \{0, 1, \ldots, |b|^2 - 1\}$ darstellen. Daher stellt sich auch hier das Problem der statistischen Eigenschaften dieser Ziffenentwicklung, wenn man z.B. alle Gaußschen ganzen Zahlen in einem Kreis $|z| \le N$ betrachtet (bzw. polynomielle Werte davon). Wieder gewinnt man zentrale Grenzwertsätze.