

Jahrbuch der
Heinrich-Heine-Universität
Düsseldorf

Heinrich Heine
HEINRICH HEINE
UNIVERSITÄT
DÜSSELDORF

2007/2008



d|u|p

düsseldorf university press

**Jahrbuch der
Heinrich-Heine-Universität
Düsseldorf
2007/2008**

**Jahrbuch der
Heinrich-Heine-Universität
Düsseldorf
2007/2008**

**Herausgegeben vom Rektor
der Heinrich-Heine-Universität Düsseldorf
Univ.-Prof. Dr. Dr. Alfons Labisch**

**Konzeption und Redaktion:
Univ.-Prof. em. Dr. Hans Süßmuth**

d|u|p

© düsseldorf university press, Düsseldorf 2008
Einbandgestaltung: Wiedemeier & Martin, Düsseldorf
Titelbild: Schloss Mickeln, Tagungszentrum der Universität
Redaktionsassistentz: Georg Stüttgen
Beratung: Friedrich-K. Unterweg
Satz: Friedhelm Sowa, L^AT_EX
Herstellung: Uniprint International BV, Meppel, Niederlande
Gesetzt aus der Adobe Times
ISBN 978-3-940671-10-3

Inhalt

Vorwort des Rektors Alfons Labisch	11
Grußwort des Amtsnachfolgers H. Michael Piper	17
Gedenken	19
Hochschulrat	
ANNE-JOSÉ PAULSEN	
Der Hochschulrat der Heinrich-Heine-Universität Düsseldorf	23
Rektorat	29
ALFONS LABISCH	
Zur Lage und zu den Perspektiven der deutschen Universität in unserer Zeit	31
MATTHIAS HOFER, NATALIE BÖDDICKER und HILDEGARD HAMMER	
Lehren – entweder man kann es, oder man kann es lernen! Hochschuldidaktik an der Heinrich-Heine-Universität Düsseldorf	43
HILDEGARD HAMMER, DORIS HILDESHEIM, VICTORIA MEINSCHÄFER und JUTTA SCHNEIDER	
Die Campus-Messe der Heinrich-Heine-Universität	61
Medizinische Fakultät	
<i>Dekanat</i>	79
<i>Neu berufene Professorinnen und Professoren</i>	81
BERND NÜRNBERG (Dekan)	
Düsseldorfer Hochschulmedizin 2008: Die Zukunft hat längst begonnen	93
INGE BAUER, LEONIE HALVERSCHEID und BENEDIKT PANNEN	
Hepatoprotektive Wirkungen des Hämoxygenase-Stoffwechsels: Der Einfluss von Anästhetika	99
ARNDT BORKHARDT	
Biologische Grundlagen der Immunrestitution nach allogener Stammzelltransplantation bei Kindern und Jugendlichen	117
LARS CHRISTIAN RUMP und OLIVER VONEND	
Pathomechanismen der arteriellen Hypertonie	127
JÖRG SCHIPPER	
Gründung und Aufbau des „Hörzentrums Düsseldorf“	141

ATTILA STEPHAN ANTAL, GABRIELA KUKOVA und BERNHARD HOMEY Juckreiz: Vom Symptom zum Mechanismus	147
WOLFGANG WÖLWER und WOLFGANG GAEBEL Kompetenznetz Schizophrenie: Konzept, Ergebnisse, Perspektiven	153
STEPHAN LUDWIG ROTH und WILFRIED BUDACH Überlebensvorteil durch präoperative Radiochemotherapie beim lokal fortgeschrittenen, nicht-inflammatorischen Brustkrebs	171
GEORG WINTERER Nikotin: Molekulare und physiologische Mechanismen im Zentralen Ner- vensystem – Ein neues nationales Schwerpunktprogramm der Deutschen Forschungsgemeinschaft	191
Mathematisch-Naturwissenschaftliche Fakultät	
<i>Dekanat</i>	201
<i>Neu berufene Professorinnen und Professoren</i>	203
ULRICH RÜTHER (Dekan) Die Mathematisch-Naturwissenschaftliche Fakultät im Jahr 2008	209
MARTIN MÖHLE Nachkommen und Vorfahren im Blickpunkt der Mathematischen Populationsgenetik	213
JÜRGEN KLÜNERS Faktorisierung von Polynomen – Ein wichtiges Problem der Computeralgebra	225
MARTIN LERCHER Wie Bakterien an neue Gene kommen und was sie damit machen	237
MATTHIAS U. KASSACK, ALEXANDRA HAMACHER und NIELS ECKSTEIN Resistenzmechanismen von Tumoren gegen Platinkomplexe: Neue Drug Targets und diagnostische Marker	249
MARGARETE BAIER Sicherheit und Kontrolle im pflanzlichen Kraftwerk – Beiträge zur Regulation des plastidären antioxidativen Schutzsystems	263
SEBASTIAN S. HORN, REBEKAH E. SMITH, and UTE J. BAYEN A Multinomial Model of Event-Based Prospective Memory	275

Philosophische Fakultät

<i>Dekanat</i>	287
<i>Neu berufene Professorinnen und Professoren</i>	289
ULRICH VON ALEMANN (Dekan)	
Wissenschaft. Leben – Die Philosophische Fakultät als tragende Säule von Lehre und Forschung	293
MICHAEL BAURMANN	
Soziologie des Fundamentalismus: Der Ansatz der sozialen Erkenntnistheorie	301
AXEL BÜHLER und PETER TEPE	
Kognitive und aneignende Interpretation in der Hermeneutik.....	315
ROBERT D. VAN VALIN, JR.	
Universal Grammar and Universals of Grammars	329
GERD KRUMEICH	
Nationalsozialismus und Erster Weltkrieg – Ein Forschungsprojekt des Historischen Seminars	339
ANNETTE SCHAD-SEIFERT	
Heiratsverhalten, sinkende Geburtenrate und Beschäftigungswandel in Japan	359
KARL-HEINZ REUBAND	
Rauchverbote in Kneipen und Restaurants. Reaktion der Bürger und der gastronomischen Betriebe – Das Beispiel Düsseldorf	373

Wirtschaftswissenschaftliche Fakultät

<i>Dekanat</i>	383
GUIDO FÖRSTER (Dekan)	
Situation und Perspektiven der Wirtschaftswissenschaftlichen Fakultät	385
WINFRIED HAMEL	
Autonomie des Unternehmens – ein frommes Märchen	395
ULRIKE NEYER	
Die Verzinsung der Mindestreserve und die Flexibilität der Geldpolitik im Eurogebiet	405

Juristische Fakultät

<i>Dekanat</i>	421
DIRK LOOSCHELDERS (Dekan)	
Situation und Perspektiven der Juristischen Fakultät	423
NICOLA PREUSS	
Die Reform der Juristenausbildung unter den Rahmenbedingungen des reglementierten Rechtsberatungsmarktes	429
KLAUS-DIETER DRÜEN	
Steuerliche Förderung von Wissenschaft und Forschung	443
CHRISTIAN KERSTING	
Informationshaftung Dritter: Vertrauen auf Verlässlichkeit	457
JAN BUSCHE, ANETTE TRAUDE und JOHANNA BOECK-HEUWINKEL	
Herausforderungen und Chancen bei der Sicherung und Verwertung von „Intellectual Property“ durch die Hochschulen – Der Düsseldorfer Weg	471

**Zentrale wissenschaftliche Einrichtungen
der Heinrich-Heine-Universität Düsseldorf*****Humanwissenschaftlich-Medizinisches Forschungszentrum
Zur Diskussion gestellt: Stammzellforschung***

JOHANNES REITER	
Menschenwürde oder Forschungsfreiheit?	487
DIETER BIRNBACHER	
Ist die Stammzellforschung unmoralisch?	495

**Gesellschaft von Freunden und Förderern der
Heinrich-Heine-Universität Düsseldorf e.V.**

OTHMAR KALTHOFF	
Jahresbericht 2007	503

Private Stiftungen für die Heinrich-Heine-Universität Düsseldorf

CHRISTOPH J. BÖRNER und H. JÖRG THIEME	
Die Schwarz-Schütte-Förderstiftung für die Wirtschaftswissenschaftliche Fakultät	507

Sonderforschungsbereiche der Heinrich-Heine-Universität Düsseldorf

JEAN KRUTMANN und FRITZ BOEGE	
Der Sonderforschungsbereich 728 „Umweltinduzierte Alterungsprozesse“	517
PETER WESTHOFF	
Wie Zellen verschieden werden – Der Sonderforschungsbereich 590.....	531

Graduiertenkollegs der Heinrich-Heine-Universität Düsseldorf

REGINE KAHL

Das Graduiertenkolleg 1427

„Nahrungsinhaltsstoffe als Signalgeber

nukleärer Rezeptoren im Darm“ 545

Graduiertenausbildung an der Heinrich-Heine-Universität Düsseldorf

CHRISTIAN DUMPITAK, LUTZ SCHMITT und DIETER WILLBOLD

Die NRW-Forschungsschule BioStruct – Neue Wege interdisziplinärer

Graduiertenausbildung an der Heinrich-Heine-Universität Düsseldorf 555

Nachwuchsforschergruppen an der Heinrich-Heine-Universität Düsseldorf

DANIEL SCHUBERT

Epigenetische Kontrolle der Pflanzenentwicklung 565

**Kooperation der Heinrich-Heine-Universität Düsseldorf
und des Forschungszentrums Jülich**

KARL ZILLES

Medizin im Forschungszentrum Jülich 579

KARL-ERICH JAEGER und MANFRED KIRCHER

Der Cluster für Industrielle Biotechnologie – CLIB²⁰²¹ 601**Ausgründungen aus der
Heinrich-Heine-Universität Düsseldorf**

JOACHIM JOSE, RUTH M. MAAS und GUNTER FESTEL

Autodisplay Biotech GmbH – Entwicklung von maßgeschneiderten

Ganzzellbiokatalysatoren und *small protein drugs* 611**Zentrale Einrichtungen der
Heinrich-Heine-Universität Düsseldorf*****Zentrale Verwaltung***

SÖNKE BIEL

Hochschulstandortentwicklungsplanung 625

Universitäts- und Landesbibliothek

IRMGARD SIEBERT

Elektronische Medien in der Informationsversorgung der Universitäts- und

Landesbibliothek Düsseldorf 639

Zentrum für Informations- und Medientechnologie

- ELISABETH DREGGER-CAPPEL und STEPHAN OLBRICH
 Erneuerung der Server- und Speicherinfrastruktur am ZIM –
 Basis für zentrale Dienste zur dezentralen IKM-Versorgung 653

Sammlungen in der Universitäts- und Landesbibliothek Düsseldorf

- JUDITH VOLLMER und MAX PLASSMANN
 40 Jahre „1968“ – 30 Jahre Studierendenstreik 1977/1978.
 Studentischer Protest im Spiegel der Plakat- und Flugblattsammlungen des
 Universitätsarchivs Düsseldorf 669

- GISELA MILLER-KIPP
 Die Sammlung „Janusz Korczak“ der Universitäts- und Landesbibliothek
 Düsseldorf und ein Versuch, Janusz Korczak als „Klassiker“ der Pädago-
 gik zu lesen 687

- RUDOLF SCHMITT-FÖLLER
 Die Flechtheim-Sammlung der Universitäts-
 und Landesbibliothek Düsseldorf 697

Geschichte der Heinrich-Heine-Universität Düsseldorf

- ULF PALLME KÖNIG
 Die Gründungsgeschichte der Juristischen Fakultät
 der Heinrich-Heine-Universität 723

- SVENJA WESTER und MAX PLASSMANN
 Univ.-Prof. Dr. Hans-Joachim Jesdinsky und die
 Einführung der Medizinischen Statistik an der Universität Düsseldorf 727

Forum Kunst

- JÜRGEN WIENER
 Architektur, Stadt- und Landschaftsplanung der Heinrich-Heine-Universität:
 Eine Bestandsaufnahme 743

Chronik der Heinrich-Heine-Universität Düsseldorf

- ROLF WILLHARDT
 Chronik 2007/2008 775

- Campus-Orientierungsplan** 787

- Daten und Abbildungen aus dem Zahlenspiegel
 der Heinrich-Heine-Universität Düsseldorf** 793

- Autorinnen und Autoren** 805

MARTIN MÖHLE

Nachkommen und Vorfahren im Blickpunkt der Mathematischen Populationsgenetik

Einführung

Mathematische Populationsgenetik ist ein Teilgebiet der Mathematik mit interdisziplinärem Bezug zu den Lebenswissenschaften, insbesondere im Bereich der Biologie und der Informatik. Ein wesentliches Ziel der Mathematischen Populationsgenetik ist die Entwicklung mathematischer Modelle zur Beschreibung und Analyse der zeitlichen Entwicklung von Populationen. Aus der Mathematik tragen insbesondere die Analysis, die Kombinatorik und die Stochastik zur Entwicklung dieses Gebietes bei. Auf der Seite der Biologie sind vor allem Evolution und Genetik die treibenden Teilbereiche. In der Informatik sind es die Bereiche Datenbanken und Algorithmen, die das Gebiet prägen.

Typische einfache Beispiele für Modelle der Mathematischen Populationsgenetik sind das exponentielle Wachstum einer Population (beschrieben durch eine deterministische Differentialgleichung der Form $N'(t) = cN(t)$) oder das Hardy-Weinberg-Gesetz für die Stabilisierung von Genotyp-Frequenzen, aber auch einfache stochastische Gesetze, etwa die Mendelschen Vererbungsregeln. Einen guten Einblick in die Mathematische Populationsgenetik gibt das Buch von Ewens.¹ Es schlägt in beeindruckender Weise die Brücke zwischen Biologie und Mathematik.

Ein wichtiges Teilgebiet der Mathematischen Populationsgenetik beschäftigt sich mit der Beschreibung und Analyse von Populationen hinsichtlich ihrer Nachkommens- und Vorfahrensstruktur. Von grundlegender Bedeutung ist in diesem Zusammenhang das Wright-Fisher-Modell, auf das nun näher eingegangen wird.

Das Wright-Fisher-Modell

Das nach den Genetikern Ronald Aylmer Fisher² und Sewall Wright³ benannte Wright-Fisher-Modell ist eines der klassischen Populationsmodelle der Mathematischen Populationsgenetik. Es wird oft als Referenzmodell (Nullmodell, Vergleichsmodell) verwendet und modelliert die zufällige Reproduktion ohne Mutation und Selektion einer eingeschlechtlichen (haploiden) endlichen Population von $N \in \mathbb{N} := \{1, 2, \dots\}$ Individuen. Es wird davon ausgegangen, dass sich die Population in diskreten Generationen fortpflanzt. Die Individuen einer jeden Generation werden zufällig von 1 bis N nummeriert. Jedes Individuum

¹ Vgl. Ewens (2004).

² Vgl. Wright (1969).

³ Vgl. Fisher (1958).

$k \in \{1, \dots, N\}$, das in der r -ten Generation lebt, bekommt eine zufällige Anzahl $v_k^{(r)}$ von Nachkommen. Die Gesamtheit all dieser Nachkommen bildet die nächste $(r+1)$ -te Generation. Man spricht daher auch von einem Modell mit nicht-überlappenden Generationen. Da die Populationsgröße als konstant ($= N$) vorausgesetzt wird, muss also in jeder Generation r die Bedingung $v_1^{(r)} + \dots + v_N^{(r)} = N$ erfüllt sein. Eine typische Verteilung, die dieser Randbedingung genügt, ist die symmetrische Multinomialverteilung, gegeben durch

$$P(v_1^{(r)} = j_1, \dots, v_N^{(r)} = j_N) = \frac{N! N^{-N}}{j_1! \cdots j_N!}$$

für $j_1, \dots, j_N \in \mathbb{N}_0 := \{0, 1, 2, \dots\}$ mit $j_1 + \dots + j_N = N$. Im Wright-Fisher-Modell wird davon ausgegangen, dass die Kindervariablen $v_1^{(r)}, \dots, v_N^{(r)}$ in jeder Generation r genau dieser Verteilung folgen. Setzt man zusätzlich noch die Unabhängigkeit dieser Zufallsvariablen von Generation zu Generation voraus, so ist dadurch das Modell präzise beschrieben.

Das Wright-Fisher-Modell lässt sich noch einfacher rückwärts in der Zeit formulieren. Es stellt sich nämlich heraus, dass die symmetrische Multinomialverteilung rückwärts in der Zeit gerade der Strategie entspricht, dass sich jedes Kind zufällig einen Elter wählt. Dies klingt zunächst etwas unrealistisch (Kinder wählen sich doch keine Eltern), aber dieses Modell stellt eine Art zufällige Abstammung sicher, das heißt, kein Individuum nimmt bevorzugt an der Reproduktion teil, was einer gewissen Homogenität der Population entspricht. Jedes Individuum hat gleichberechtigt die Chance, Nachkommen zu zeugen. Es wird kein Individuum besonders bevorzugt. Man spricht auch von einem neutralen beziehungsweise selektionsfreien Modell.

Es gibt viele weitere Verteilungen, die so genannten austauschbaren Verteilungen, die ebenfalls diese Homogenitätseigenschaft besitzen. Die symmetrische Multinomialverteilung bietet sich jedoch als Referenzverteilung an, vor allem, weil der rechnerische Umgang mit ihr relativ einfach ist. Ein haploides Populationsmodell mit konstanter Populationsgröße N heißt austauschbar, wenn die Verteilung der Kindervariablen invariant gegenüber Permutationen der Individuen ist, das heißt, wenn $(v_{\pi 1}^{(r)}, \dots, v_{\pi N}^{(r)})$ dieselbe Verteilung besitzt wie $(v_1^{(r)}, \dots, v_N^{(r)})$ für jede Generation r und für jede Permutation π der Zahlen von 1 bis N . Das neben dem Wright-Fisher-Modell bekannteste austauschbare Populationsmodell ist das Moran-Modell. Bei diesem Modell geht man davon aus, dass ein zufällig ausgewähltes Individuum zwei Nachkommen zeugt, ein weiteres zufällig ausgewähltes Individuum gar keine Nachkommen hat und alle $N-2$ übrigen Individuen genau einen Nachkommen haben. Mathematisch lässt sich dieses Modell sehr einfach dadurch definieren, dass in jeder Generation r der Kindervektor $(v_1^{(r)}, \dots, v_N^{(r)})$ eine zufällige Permutation von $(2, 1, \dots, 1, 0)$ ist.

Nachkommen

Ein typisches Vorgehen ist, dass man eine feste Anzahl $i \in \{1, \dots, N\}$ von Individuen aus der gegenwärtigen Generation 0 auswählt und deren Nachkommen in Generation r , das heißt r Generationen in der Zukunft, betrachtet. Im einfachsten Fall ist man schlicht an der Anzahl $X_N(r)$ der Nachkommen dieser i Individuen in der r -ten Generation interessiert.

Mathematisch lässt sich zeigen, dass der Prozess $X_N := (X_N(r))_{r \in \mathbb{N}_0}$ eine Markoffsche Kette und ein Martingal ist. Falls sich der Prozess zur Zeit r im Zustand i befindet, so bewegt er sich im nächsten Schritt in den Zustand j mit Übergangswahrscheinlichkeit

$$P(X_N(r+1) = j | X_N(r) = i) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}.$$

Der Prozess X_N wird Nachkommensprozess genannt. In Anwendungen (Biologie, Medizin) werden oft große Populationen, etwa Zellkulturen, betrachtet, das heißt, man ist daran interessiert, wie der Prozess X_N für große Gesamtpopulationsgröße N aussieht. Dazu geht man wie folgt vor: Da der Prozess X_N mit wachsendem N immer größere Werte annehmen kann, betrachtet man zunächst relative Anzahlen $X_N(r)/N$ von Nachkommen, um die Anzahlen auf Werte zwischen 0 und 1 zu normieren. Die zweite und wesentlichere Idee ist, N Generationen in einem Zeitschritt zusammenzufassen; das heißt, man misst die Zeit in Einheiten von $r = [tN]$ Generationen, wobei $[tN]$ der ganzzahlige Teil der reellen Zahl tN und t eine beliebige reelle nichtnegative Zahl ist. Mathematisch lässt sich für das Wright-Fisher-Modell folgendes Konvergenzresultat herleiten. Falls $X_N(0)/N$ mit wachsender Gesamtpopulationsgröße N in Verteilung gegen eine Grenzvariable X_0 konvergiert, so konvergiert der normierte und zeitskalierte Nachkommensprozess $(X_N([Nt])/N)_{t \geq 0}$ mit wachsendem N in Verteilung gegen einen zeitkontinuierlichen Prozess $X = (X_t)_{t \geq 0}$, die so genannte Wright-Fisher-Diffusion.

Konvergenzresultate dieser Art laufen in der Stochastik unter dem Begriff Diffusionsapproximation. Für die Anwendung hat dies die Konsequenz, dass sich die Nachkommenschaft hinreichend gut mit Hilfe der Wright-Fisher-Diffusion approximieren lässt, sofern die Gesamtpopulationsgröße genügend groß ist. Die Wright-Fisher-Diffusion kann funktionalanalytisch durch ihren Generator $Af(x) = \frac{1}{2}x(1-x)f''(x)$, $f \in C^2([0, 1])$, beschrieben werden oder durch die stochastische Differentialgleichung $dX_t = \sqrt{X_t(1-X_t)}dB_t$, wobei $B = (B_t)_{t \geq 0}$ eine Standard Brownsche Bewegung ist. Das Konvergenzresultat lässt sich auf austauschbare Populationsmodelle verallgemeinern. Die Zeit ist dann in Einheiten der effektiven Populationsgröße $N_e = [1/c_N]$ zu skalieren,⁴ wobei

$$c_N := \frac{1}{N(N-1)} \sum_{k=1}^N E(v_k^{(r)}(v_k^{(r)} - 1)) = \frac{E(v_1^{(r)}(v_1^{(r)} - 1))}{N-1} = \frac{\text{Var}(v_1^{(r)})}{N-1}$$

die so genannte Verschmelzungswahrscheinlichkeit ist, das heißt die Wahrscheinlichkeit, dass zwei zufällig aus einer Generation ausgewählte Individuen einen gemeinsamen Vorfahr haben. Im Wright-Fisher-Modell stimmt die effektive Populationsgröße N_e mit der tatsächlichen Populationsgröße N überein – ein weiterer Grund, weshalb das Wright-Fisher-Modell oft als Referenzmodell herangezogen wird. Im Moran-Modell ist die effektive Populationsgröße $N_e = [N(N-1)/2]$ deutlich größer als die tatsächliche Populationsgröße. Dies ist auch qualitativ klar, denn pro Generation tut sich aus der Sicht der Evolution im Moran-Modell relativ wenig im Vergleich zum Wright-Fisher-Modell.

⁴ Vgl. Möhle (2001).

Sampling-Verteilungen

In vielen Anwendungen ist eine bestimmte Eigenschaft (Typ) der Individuen von Interesse. Das Wright-Fisher-Modell wird deshalb oft dadurch erweitert, dass jedes Individuum zusätzlich von einem bestimmten Typ ist und sich der Typ eines jeden Nachkommen vom Typ des Elter-Individuums mit einer gewissen Mutationswahrscheinlichkeit p unterscheidet. Im einfachsten Fall wird davon ausgegangen, dass bei Mutation ein neuer Typ entsteht, der zuvor noch nie in der Population aufgetaucht ist (*infinitely-many-alleles*-Modell). Betrachte nun eine große Population ($N \rightarrow \infty$) nach sehr langer Zeit ($r \rightarrow \infty$). Wählt man eine Stichprobe von n Individuen aus, so kann man sich fragen, wie viele Typen in dieser Stichprobe vorkommen. Man spricht auch von der Typverteilung und von hoher genetischer Variabilität, falls viele Typen (nahe bei n) in der Stichprobe vorhanden sind. Unter der Annahme, dass die Mutationswahrscheinlichkeit p klein ist, genauer gesagt, unter der Voraussetzung, dass der Grenzwert $\theta := \lim_{N \rightarrow \infty} 2Np$ existiert, lässt sich die Verteilung der Typen genau angeben. Man nennt θ die Mutationsrate. Die Ewens-Sampling-Formel⁵ besagt, dass die Typ-Verteilung durch

$$p(a_1, \dots, a_n) = \frac{n!}{\theta(\theta+1)\cdots(\theta+n-1)} \prod_{j=1}^n \frac{\theta^{a_j}}{j^{a_j} a_j!}$$

gegeben ist. Dabei bezeichnet a_j die Anzahl der Typen, die genau j -mal in der Stichprobe vorkommen, $1 \leq j \leq n$, so dass also $a_1 + 2a_2 + \cdots + na_n = n$ gilt. Eine typische Folgerung daraus ist, dass die erwartete Anzahl von Typen asymptotisch gleich $\theta \log n$ ist.

Moderne Beweise der Ewens-Sampling-Formel basieren meist auf dem Kingman-Coalescent, auf den wir später noch im Detail zu sprechen kommen. Ewens selbst stand in den 1970er Jahren natürlich die Coalescent-Theorie noch nicht zur Verfügung.

Neuere Forschungen auf diesem Gebiet betrachten allgemeinere Sampling-Verteilungen⁶ für Coalescent-Prozesse mit simultanen multiplen Verschmelzungen von Vorfahrenlinien. Bis heute sind, abgesehen vom Kingman-Coalescent und dem so genannten sternförmigen Coalescent, lediglich Rekursionen, das heißt keine expliziten Lösungen für diese Sampling-Verteilungen bekannt. Solche und ähnliche Rekursionen können zum statistischen Schätzen von Parametern eingesetzt werden.⁷

Vorfahren

In Analogie zur Betrachtung von Nachkommen vorwärts in der Zeit ist man in der Mathematischen Populationsgenetik in natürlicher Weise auch an einer Betrachtung der Population rückwärts in der Zeit interessiert. Beispielsweise wählt man n Individuen der Gegenwart aus und betrachtet die Anzahl $D_N(r)$ der Vorfahren dieser n Individuen r Generationen in der Vergangenheit. Zu beachten ist, dass die Generationen nun rückwärts in der Zeit gezählt werden, das heißt, $r = 0$ ist die gegenwärtige Generation der Kinder, $r = 1$ die Generation der Eltern und so weiter. Auch hier lässt sich für das Wright-Fisher-Modell eine Grenzaussage über den Vorfahrenprozess $(D_N(r))_{r \in \mathbb{N}_0}$ herleiten. Der zeit-skalierte Prozess $(D_N(\lfloor Nt \rfloor))_{t \geq 0}$ konvergiert mit wachsender Gesamtpopulationsgröße N

⁵ Vgl. Ewens (1972).

⁶ Vgl. Möhle (2006).

⁷ Vgl. Birkner und Blath (2008).

gegen einen zeitkontinuierlichen Markoffschen Todesprozess $(D_t)_{t \geq 0}$ mit Zustandsraum $\{1, \dots, n\}$ und infinitesimalen Raten $d_k = k(k-1)/2$. Diese Vorfahrensprozesse haben 1 als absorbierenden Zustand, das heißt, nach hinreichend langer Zeit in die Vergangenheit ist man beim gemeinsamen Vorfahr angelangt – ein Phänomen, das typisch für eingeschlechtliche Populationen ist.

Auch dieses Konvergenzresultat lässt sich auf allgemeinere austauschbare Populationsmodelle erweitern, wobei dann als Zeitskalierung wieder die effektive Populationsgröße ins Spiel kommt.⁸ Dass diese Zeitskalierung identisch mit derjenigen ist, die schon im Zusammenhang mit den Nachkommen auftaucht, ist kein Zufall. Es besteht ein enger Zusammenhang zwischen Nachkommen und Vorfahren, auf den nun genauer eingegangen wird.

Dualität

Das Prinzip der Dualität taucht in vielen Bereichen der Lebenswissenschaften auf und stellt eine Verallgemeinerung des Begriffes der (zeitlichen) Reversibilität dar. Es wird unter anderem in der Physik der interagierenden Teilchensysteme⁹ als Analysewerkzeug verwendet. Im Allgemeinen stellt die Dualität einen Zusammenhang zwischen zwei Markoffschen Prozessen dar, im Kontext der Mathematischen Populationsgenetik ist dies ein Zusammenhang zwischen Nachkommen und Vorfahren.¹⁰ Für die zuvor betrachteten zeitkontinuierlichen Prozesse $X = (X_t)_{t \geq 0}$ (Nachkommen) und $D = (D_t)_{t \geq 0}$ (Vorfahren) lässt sich zeigen, dass $E(X_t^n | X_0 = x) = E(x^{D_t} | D_0 = n)$ für beliebige Zeiten $t \geq 0$, $n \in \mathbb{N}$ und $0 \leq x \leq 1$. Mathematisch stellt dies einen Bezug zwischen den Momenten von X_t und der erzeugenden Funktion von D_t dar. Die in der Dualitätsgleichung auftretende Funktion $H(x, n) = x^n$ wird Dualitätsfunktion genannt. Ist die (oder eine) Dualitätsfunktion – wie in dem hier betrachteten Fall – bekannt, so hat dies die Konsequenz, dass die Wahrscheinlichkeitstheoretische Struktur des einen Prozesses durch diejenige des anderen Prozesses bereits eindeutig bestimmt ist. Nachkommen und Vorfahren stehen also in einem engen Bezug zueinander, der durch den Begriff der Dualität mathematisch präzisiert wird. Es gibt wesentlich komplexere Formen der Dualität. So sind etwa die Fleming-Viot-Prozesse dual zu den Coalescent-Prozessen. Wir werden darauf später genauer eingehen.

Kingman-Coalescent

In vielen Fällen ist man nicht nur an Anzahlen von Vorfahren interessiert, sondern man möchte vielmehr die Genealogie (Stammbaum) von n Individuen verstehen. Der Kingman-Coalescent (lateinisch: *coalescere* = verschmelzen, vereinigen, sich verbinden) ist ein stochastischer Prozess, der diese Verwandtschaftsstruktur in robuster Weise beschreibt. Er ist zu einem wichtigen Werkzeug der Mathematischen Populationsgenetik geworden. Wie zuvor entspricht eine Zeiteinheit im Coalescent-Prozess N Generationen im Wright-Fisher-Modell. Mathematisch ist der Kingman-Coalescent ein zeitkontinuierlicher Markoffscher Prozess $\Pi = (\Pi_t)_{t \geq 0}$ mit binären Verschmelzungen von Vorfahrenlinien. Der Zustands-

⁸ Vgl. Möhle (2001).

⁹ Vgl. Liggett (1985: 84, Definition 3.1).

¹⁰ Vgl. Möhle (1999).

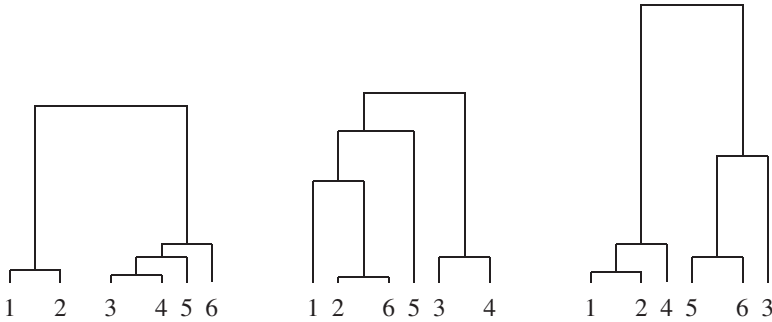


Abb. 1: Grafische Darstellung von drei Realisationen des Kingman-6-Coalescent

raum ist die Menge \mathcal{P} der Partitionen von \mathbb{N} . In Anwendungen wird – ausgehend von einer Stichprobe von n Individuen – meist nur der eingeschränkte Prozess $(\rho_n \Pi_t)_{t \geq 0}$ benötigt, wobei $\rho_n : \mathcal{P} \rightarrow \mathcal{P}_n$ die natürliche Restriktion von \mathcal{P} auf die Menge \mathcal{P}_n der Partitionen von $\{1, \dots, n\}$ bezeichnet. Die Grundidee ist die, dass (per Definition) je zwei Individuen i und j der Stichprobe zum selben Block der Partition $\rho_n \Pi_t$ gehören, wenn sie einen gemeinsamen Vorfahr zur Zeit t in der Vergangenheit haben. Bei jedem Übergang verschmelzen genau zwei Blöcke der gerade vorliegenden Partition, und zwar mit Rate 1.

Die ersten fundierten Arbeiten zum Coalescent stammen von Kingman¹¹ aus dem Jahr 1982, jedoch gehen die Ursprünge etwas weiter zurück.¹² In den 1990er Jahren wurden wesentliche Beiträge zur Coalescent-Theorie von Donnelly und Tavaré¹³ geleistet. Für weitere Literatur sei auf die Überblicksartikel von Hudson,¹⁴ Fu und Li¹⁵ und Möhle¹⁶ sowie auf die Bücher von Hein, Schierup und Wiuf,¹⁷ Bertoin¹⁸ und Wakeley¹⁹ verwiesen.

Die Bedeutung des Kingman-Coalescent liegt darin begründet, dass er relativ robust gegenüber Änderungen des zugrunde liegenden Populationsmodelles ist. Nicht nur im Wright-Fisher-Modell, sondern für eine große Klasse von Populationsmodellen kann die Genealogie einer Stichprobe von n Individuen approximativ durch den Kingman-Coalescent beschrieben werden, sofern die Gesamtpopulationsgröße N hinreichend groß ist.²⁰ Man spricht von der Robustheit²¹ des Coalescent. Dieses Invarianzprinzip begründet seine fundamentale Bedeutung für die Mathematische Populationsgenetik. Wie die Normalverteilung im zentralen Grenzwertsatz taucht der Coalescent mit wachsender Ge-

¹¹ Vgl. Kingman (1982a), Kingman (1982b) sowie Kingman (1982c).

¹² Vgl. Kingman (2000).

¹³ Vgl. Donnelly und Tavaré (1995).

¹⁴ Vgl. Hudson (1991).

¹⁵ Vgl. Fu und Li (1999).

¹⁶ Vgl. Möhle (2000).

¹⁷ Vgl. Hein *et al.* (2005).

¹⁸ Vgl. Bertoin (2006).

¹⁹ Vgl. Wakeley (im Druck).

²⁰ Vgl. Möhle und Sagitov (2001).

²¹ Vgl. Möhle (1998).

samtpopulationsgröße in natürlicher Weise als universaler Grenzprozess auf. Darüber hinaus wurde und wird die Theorie zum Coalescent kontinuierlich auf komplexere und somit realitätsnähere Populationsmodelle ausgeweitet. Dazu gehören beispielsweise strukturierte Populationen,²² Coalescent-Prozesse mit Rekombination²³ und Selektion²⁴ sowie Coalescent-Prozesse für diploide und zweigeschlechtliche Populationsmodelle.²⁵

Coalescent-Prozesse mit simultanen multiplen Kollisionen

Neuere Forschungen auf dem Gebiet der austauschbaren haploiden Populationsmodelle haben ergeben, dass neben dem Kingman-Coalescent, der nur binäre Verschmelzungen von Vorfahrenlinien erlaubt, eine ganze Klasse weiterer Coalescent-Prozesse von zentraler Bedeutung ist, die multiple und sogar simultan mehrere multiple Verschmelzungen von Vorfahrenlinien zulassen. Diese Prozesse treten mit wachsender Gesamtpopulationsgröße als Grenzprozesse in austauschbaren Populationsmodellen mit großen Familiengrößen auf, das heißt in Modellen, in denen einzelne Individuen ab und zu sehr viele (in der Größenordnung der Gesamtpopulation) Nachkommen zeugen können. Coalescent-Prozesse mit multiplen Kollisionen wurden unabhängig von Pitman²⁶ und Sagitov²⁷ eingeführt. Diese Prozesse können durch ein endliches Maß Λ auf dem Einheitsintervall charakterisiert werden. Sie werden daher auch Λ -Coalecents genannt. Befindet sich der Prozess in einer Partition mit b Blöcken, so tritt eine Kollision (Verschmelzung) von k Vorfahrenlinien mit Rate

$$\lambda_{b,k} = \int_{[0,1]} x^{k-2}(1-x)^{b-k} \Lambda(dx)$$

auf. Für $\Lambda = \delta_0$ (Dirac-Maß im Punkt 0) ergibt sich der Kingman-Coalescent. Im Fall $\Lambda = \delta_1$ (Dirac-Maß im Punkt 1) spricht man vom sternförmigen Coalescent. Ist Λ die Gleichverteilung (Lebesgue-Maß) auf dem Einheitsintervall, so spricht man vom Bolthausen-Sznitman-Coalescent.²⁸

Kurz darauf wurden alle möglichen Coalescent-Prozesse von Möhle und Sagitov²⁹ und Schweinsberg³⁰ klassifiziert. Diese erlauben zur gleichen Zeit gleich mehrere multiple Verschmelzungen von Vorfahrenlinien. Die Beschreibung dieser Prozesse erfolgt wieder mit Hilfe der infinitesimalen Raten, das heißt durch Angabe des Generators (Erzeugers). Während die Klassifikation nach Möhle und Sagitov mit Hilfe einer konsistenten Folge endlicher symmetrischer Maße $\Lambda_1, \Lambda_2, \dots$ erfolgt, wobei Λ_j ein Maß auf dem endlichen Simplex $\Delta_j := \{(u_1, \dots, u_j) \in [0, 1]^j : u_1 + \dots + u_j \leq 1\}$ ist, charakterisiert Schweinsberg diese Prozesse alternativ mit Hilfe eines einzigen Maßes Ξ auf dem unendlichen Simplex $\Delta := \{(u_1, u_2, \dots) : u_1 \geq u_2 \geq \dots \geq 0, \sum_i u_i \leq 1\}$. Man spricht daher auch vom Ξ -

²² Vgl. Wilkinson-Herbots (1998).

²³ Vgl. Hudson und Kaplan (1985).

²⁴ Vgl. Krone und Neuhauser (1997) sowie Neuhauser und Krone (1997).

²⁵ Vgl. Möhle und Sagitov (2003).

²⁶ Vgl. Pitman (1999).

²⁷ Vgl. Sagitov (1999).

²⁸ Vgl. Bolthausen und Sznitman (1998).

²⁹ Vgl. Möhle und Sagitov (2001).

³⁰ Vgl. Schweinsberg (2000).

Coalescent. Die Rate, mit der eine (k_1, \dots, k_j) -Kollision auftritt, $k_1 \geq k_2 \geq \dots \geq k_j$, ist gegeben durch

$$\int_{\Delta} \sum_{l=0}^s \binom{s}{l} (1 - |u|)^{s-l} \sum_{\substack{i_1, \dots, i_{r+j} \in \mathbb{N} \\ \text{alle verschieden}}} u_{i_1}^{k_1} \dots u_{i_r}^{k_r} u_{i_{r+1}} \dots u_{i_{r+j}} \frac{\Xi(du)}{(u, u)},$$

mit stetiger Interpretation für $u \rightarrow 0$. Dabei ist $s := |\{i : k_i = 1\}|$ die Anzahl der Singletons, $r := j - s$, $|u| := \sum_i u_i$ und $(u, u) := \sum_i u_i^2$. Allein die Formel für diese Raten zeigt die Komplexität dieser Prozesse, obwohl sie in natürlicher Weise als Grenzprozesse von Stammbaumprozessen in austauschbaren Populationsmodellen auftauchen.

Funktionale von Coalescent-Stammbäumen

Ein wichtiger Zweig der Coalescent-Theorie beschäftigt sich mit der Analyse von Funktionalen von Stammbäumen. Aus biologischer Sicht ist beispielsweise die Zeit τ_n einer Stichprobe von n Individuen zurück bis zum jüngsten gemeinsamen Vorfahr (englisch: *most recent common ancestor*, kurz MRCA) von besonderer Bedeutung. Diese Zeit entspricht der Länge des Coalescent-Stammbaumes. Weitere wichtige solche Funktionale sind

- (a) die Anzahl X_n der Verschmelzungen (Kollisionen), die stattfinden, bis man beim MRCA angekommen ist,
- (b) die totale Länge L_n des Baumes, das heißt die Summe der Längen aller Zweige des Stammbaumes,
- (c) die Länge Z_n eines zufällig ausgewählten externen Zweiges.

Für Coalescent-Prozesse mit Mutationen sind weitere Funktionale von besonderem Interesse für die Biologie, etwa

- (d) die Anzahl S_n der Mutationen, die im *infinitely-many-alleles*-Modell mit der Anzahl der sich unterscheidenden Stellen übereinstimmt, oder
- (e) die Anzahl der Typen K_n , die bereits im Abschnitt über Sampling-Verteilungen auftauchte.

Von besonderem Interesse ist dabei das Verhalten dieser Funktionale mit wachsendem Stichprobenumfang n . Während diese Funktionale im Fall des Kingman-Coalescent inzwischen weitgehend untersucht sind, steht die Forschung auf diesem Gebiet für Stammbaumprozesse mit simultanen multiplen Kollisionen noch am Anfang. Inzwischen gibt es diverse Resultate zur Anzahl der Kollisionen.³¹ Auch diverse Funktionale im Fall des Bolthausen-Sznitman-Coalescent sind weitgehend verstanden.³² Zurzeit werden vor allem Funktionale von Beta-Coalescent-Prozessen studiert, das heißt von Λ -Coalescents, wobei $\Lambda = \beta(a, b)$ eine Betaverteilung mit Parametern $a, b > 0$ ist.

Fleming-Viot-Prozesse

Der klassische Fleming-Viot-Prozess³³ $F = (F_t)_{t \geq 0}$ kann als Verallgemeinerung der Wright-Fisher-Diffusion angesehen werden. Die Zufallsvariable F_t beschreibt die Typ-

³¹ Vgl. Gnedin und Yakubovich (2007).

³² Vgl. Drmota *et al.* (2007).

³³ Vgl. Fleming und Viot (1979).

verteilung der Population zur Zeit t vorwärts in der Zeit. Der Typraum E darf dabei ein beliebiger kompakter Raum sein. Der Fleming-Viot-Prozess ist ein maßwertiger Prozess, das heißt $F_t : \Omega \rightarrow \mathcal{M}_1(E)$, wobei $\mathcal{M}_1(E)$ die Menge der Wahrscheinlichkeitsmaße auf E bezeichnet. Die Dualität zwischen Wright-Fisher-Diffusion $X = (X_t)_{t \geq 0}$ und dem im Abschnitt über Vorfahren eingeführten Todesprozess $D = (D_t)_{t \geq 0}$ findet hier ihre entsprechende Verallgemeinerung. Es stellt sich nämlich heraus, dass der Fleming-Viot-Prozess $F = (F_t)_{t \geq 0}$ dual zum Kingman-Coalescent $\Pi = (\Pi_t)_{t \geq 0}$ ist.

Allgemeiner lässt sich zu jedem Maß Ξ auf dem unendlichen Simplex Δ ein maßwertiger Prozess $F = (F_t)_{t \geq 0}$ konstruieren, dessen Generator L für geeignete Testfunktionen $G : \mathcal{M}_1(E) \rightarrow \mathbb{R}$ die Gestalt

$$LG(\mu) = \int_{\Delta} \int_{E^{\mathbb{N}}} \left(G((1 - |u|)\mu + \sum_{i=1}^{\infty} u_i \delta_{x_i}) - G(\mu) \right) \mu^{\mathbb{N}}(dx) \frac{\Xi(du)}{(u, u)}$$

hat, mit stetiger Interpretation für $u \rightarrow 0$. Die Klasse der Testfunktionen ist hinreichend groß, so dass dadurch eindeutig ein maßwertiger Prozess, der so genannte Ξ -Fleming-Viot-Prozess, definiert ist. Ist Ξ das Dirac-Maß in 0, so ergibt sich der klassische Fleming-Viot-Prozess. In Verallgemeinerung zu der zuvor erwähnten Dualität zwischen dem klassischen Fleming-Viot-Prozess und dem Kingman-Coalescent ist der Ξ -Fleming-Viot-Prozess $F = (F_t)_{t \geq 0}$ dual zum Ξ -Coalescent $\Pi = (\Pi_t)_{t \geq 0}$ mit simultanen multiplen Kollisionen. Genauer gesagt gilt

$$E(H_n(F_t, \xi) | F_0 = \mu) = E(H_n(\mu, \rho_n \Pi_t) | \rho_n \Pi_t = \xi)$$

für beliebige Zeiten $t \geq 0$, für jedes Wahrscheinlichkeitsmaß $\mu \in \mathcal{M}_1(E)$, für jedes $n \in \mathbb{N}$ und für jede Partition $\xi \in \mathcal{P}_n$. Die Dualitätsfunktion $H_n : \mathcal{M}_1(E) \times \mathcal{P}_n \rightarrow \mathbb{R}$ ist dabei gegeben durch

$$H_n(\mu, \xi) := \int_{E^b} x_1^{k_1} \cdots x_b^{k_b} \mu(dx_1) \cdots \mu(dx_b),$$

wobei k_1, \dots, k_b die Größen der Blöcke der Partition ξ sind. Nicht zuletzt wegen dieser Dualitätseigenschaft stehen Fleming-Viot-Prozesse und andere verwandte maßwertige Prozesse daher im Interesse der aktuellen Forschung.

Literatur

- BERTOIN, Jean (2006). *Random Fragmentation and Coagulation Processes*. Cambridge.
- BIRKNER, Matthias und Jochen BLATH (2008). „Computing likelihoods for coalescents with multiple collisions in the infinitely many sites model“, *Journal of Mathematical Biology* 57, 435–465.
- BOLTHAUSEN, Erwin und Alain S. SZNITMAN (1998). „On Ruelle’s probability cascades and an abstract cavity method“, *Communications in Mathematical Physics* 197, 247–276.
- DONNELLY, Peter und Simon TAVARÉ (1995). „Coalescents and genealogical structure under neutrality“, *Annual Review of Genetics* 29, 401–421.
- DRMOTA, Michael, Alex IKSANOV, Martin MÖHLE und Uwe RÖSLER (2007). „Asymptotic results concerning the total branch length of the Bolthausen-Sznitman coalescent“, *Stochastic Processes and their Applications* 117, 1404–1421.
- EWENS, Warren J. (1972). „The sampling theory of selectively neutral alleles“, *Theoretical Population Biology* 3, 87–112.

- EWENS, Warren J. (2004). *Mathematical Population Genetics*. New York.
- FISHER, Ronald A. (1958). *The Genetical Theory of Natural Selection*. Oxford und New York.
- FLEMING, Wendell H. und Michel VIOT (1979). „Some measure-valued Markov processes in population genetics theory“, *Indiana University Mathematics Journal* 28, 817–843.
- FU, Yun-Xin und Wen-Hsiung LI (1999). „Coalescing into the 21st century: an overview and prospects of coalescent theory“, *Theoretical Population Biology* 56, 1–10.
- GNEDIN, Alexander und Yuri YAKUBOVICH (2007). „On the number of collisions in Λ -coalescents“, *Electronic Journal of Probability* 12, 1547–1567.
- HEIN, Jotun, Mikkel H. SCHIERUP und Carsten WIUF (2005). *Gene Genealogies, Variation and Evolution – a Primer in Coalescent Theory*. Oxford.
- HUDSON, Richard R. und Norman L. KAPLAN (1985). „Statistical properties of the number of recombination events in the history of a sample of DNA sequences“, *Genetics* 111, 147–164.
- HUDSON, Richard R. (1991). „Gene genealogies and the coalescent process“, *Oxford Surveys in Evolutionary Biology* 7, 1–44.
- KINGMAN, John F. C. (1982a). „On the genealogy of large populations“, *Journal of Applied Probability* 19A, 27–43.
- KINGMAN, John F. C. (1982b). „Exchangeability and the evolution of large populations“, in: Giorgio KOCH und Fabio SPIZZICHINO (Hrsg.). *Exchangeability in Probability and Statistics*. Amsterdam, 97–112.
- KINGMAN, John F. C. (1982c). „The coalescent“, *Stochastic Processes and their Applications* 13, 235–248.
- KINGMAN, John F. C. (2000). „Origins of the coalescent: 1974–1982“, *Genetics* 156, 1461–1463.
- KRONE, Stephen M. und Claudia NEUHAUSER (1997). „Ancestral processes with selection“, *Theoretical Population Biology* 51, 210–237.
- LIGGETT, Thomas M. (1985). *Interacting Particle Systems*. New York.
- MÖHLE, Martin (1998). „Robustness results for the coalescent“, *Journal of Applied Probability* 35, 438–447.
- MÖHLE, Martin (1999). „The concept of duality and applications to Markov processes arising in population genetics models“, *Bernoulli* 5, 761–777.
- MÖHLE, Martin (2000). „Ancestral processes in population genetics – the coalescent“, *Journal of Theoretical Biology* 204, 629–638.
- MÖHLE, Martin (2001). „Forward and backward diffusion approximations for haploid exchangeable population models“, *Stochastic Processes and their Applications* 95, 133–149.
- MÖHLE, Martin und Serik SAGITOV (2001). „A classification of coalescent processes for haploid exchangeable population models“, *Annals of Probability* 29, 1547–1562.
- MÖHLE, Martin und Serik SAGITOV (2003). „Coalescent patterns in diploid exchangeable population models“, *Journal of Mathematical Biology* 47, 337–352.
- MÖHLE, Martin (2006). „On sampling distributions for coalescent processes with simultaneous multiple collisions“, *Bernoulli* 12, 35–53.
- NEUHAUSER, Claudia und Stephen M. KRONE (1997). „The genealogy of samples in models with selection“, *Genetics* 145, 519–534.
- PITMAN, Jim (1999). „Coalescents with multiple collisions“, *The Annals of Probability* 27, 1870–1902.
- SAGITOV, Serik (1999). „The general coalescent with asynchronous mergers of ancestral lines“, *Journal of Applied Probability* 36, 1116–1125.
- SCHWEINSBERG, Jason (2000). „Coalescents with simultaneous multiple collisions“, *Electronic Journal of Probability* 5, 1–50.

- WAKELEY, John (im Druck). *Coalescent Theory: An Introduction*. Greenwood Village.
- WILKINSON-HERBOTS, Hilde M. (1998). „Genealogy and subpopulation differentiation under various models of population structure“, *Journal of Mathematical Biology* 37, 535–585.
- WRIGHT, Sewall (1969). *Evolution and the Genetics of Populations*. Bd. 1. Chicago.

