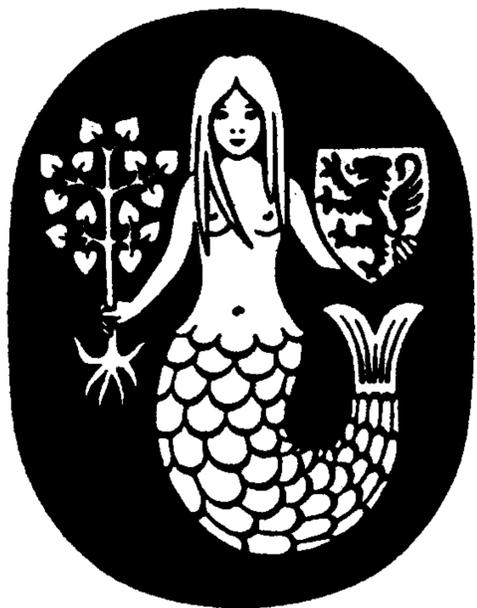

Genealogica & Heraldica

*Under the High Patronage of
Her Majesty Queen Ingrid*



Edited by
Sven Tito Achen

Report of
The 14th International
Congress of
Genealogical and
Heraldic Sciences
in Copenhagen
25.-29. August 1980

Published by
Samfundet for Dansk
Genealogi og Personhistorie
and Heraldisk Selskab
(Societas Heraldica
Scandinavica)

G · E · C GAD · Copenhagen
1982

Theoretische Genealogie

HERMANN ATHEN (†)

1. Was ist theoretische Genealogie, und was kann sie nützen?

Genealogie ist im Laufe der Zeit von verschiedenen Aspekten her definiert und praktiziert worden, ursprünglich mehr vom empirischen Standpunkt her, heute stärker von wissenschaftlichen Aspekten, in erster Linie historisch, soziologisch und zunehmend auch biologisch-genetisch. Sie ist aber nur dann als Wissenschaft zu bezeichnen, wenn neben den angegebenen Komponenten auch noch die Aspekte der Quantifizierbarkeit und der Mathematisierbarkeit (kurz der Computerisierbarkeit) zu realisieren sind. Zwar ist in den vergangenen Jahrzehnten manches zum quantitativen Aspekt publiziert worden, beispielsweise von *S. Rösch* neben einer Reihe von Zeitschriftenaufsätzen ein umfangreiches Heft (1) über quantitative Genealogie. Der Aspekt der Mathematisierbarkeit geht, obwohl er natürlich teilweise auch schon bei der Quantifizierung wirksam wird, noch einen Schritt weiter. Er soll dazu führen, gewisse allgemeingültige, d.h. theoretische Gesetzmäßigkeiten aufzustellen, die sich in der praktischen Arbeit des Genealogen darstellen und bewähren. Solche mathematischen Gesetzmäßigkeiten sind selbstverständlich auch jetzt schon in der Genealogie aufzufinden, z.B. in Ahnentafeln die Kekulé-Benummerung mittels Dualziffern, die Sie sicher alle kennen. Aus der Ahnennummer läßt sich dann sofort die Filiationslinie rekonstruieren (Bild 1).

Auf den (sehr einfachen) mathematischen Beweis gehe ich hier nicht ein; er dürfte allgemein bekannt sein.

Mit mathematischen Hilfsmitteln lassen sich aber auch strittige Fragen oft überzeugend klären.

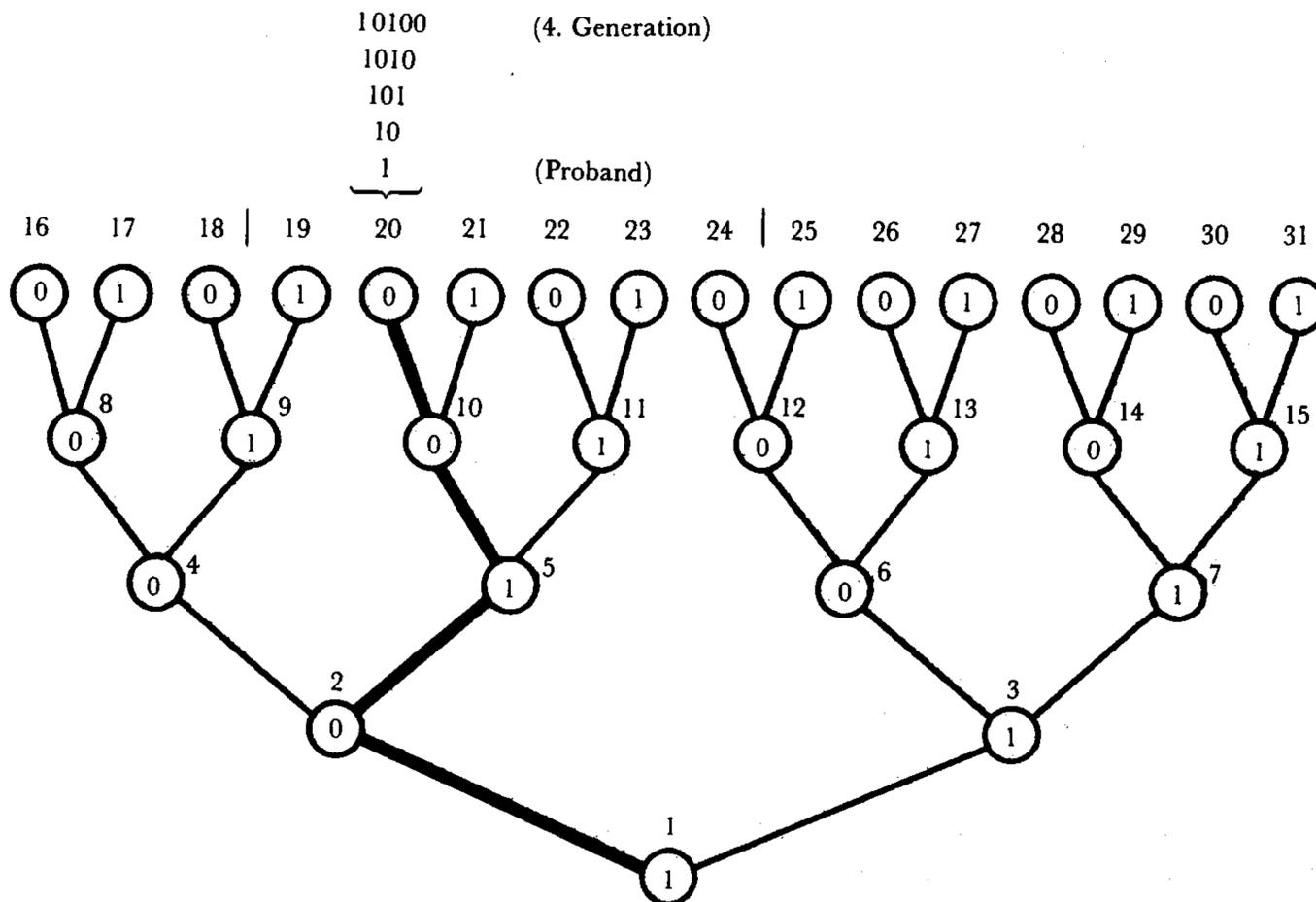
Als Beispiel diene einleitend der nächste Abschnitt, dem dann einige weitere Beispiele folgen sollen.

2. Kann man die Verwandtschaft quantifizieren?

Um es vorwegzunehmen: Man kann! Dies ist nämlich überhaupt kein Problem, wenn man einfach wie im Römischen oder Kanonischen Recht nur von dem Generationenabstand ausgeht und die Verwandtschaft durch einen geeigneten mittleren "verwandtschaftlichen Generationenabstand" definiert. Problematisch wird die Sache, wenn genetische Gesichtspunkte ins Spiel kommen, wie es bisher (leider) fast immer geschehen ist. In der deutschen genealogischen Literatur stehen sich zwei Standpunkte diametral gegenüber: Aus genetischen und (zutreffenden) statistischen Erwägungen folgert *S. Rösch* (2), daß die Anzahl der Gene in der Filiationslinie vom Urahn zum Probanden von Generation zu Generation jeweils um die Hälfte abnimmt. Führen mehrere Filiationslinien vom Probanden zum Urahn, so ergibt sich ein mittlerer verwandtschaftlicher Generationenabstand d aus der Summe

$$2^{-d} = 2^{-d_1} + 2^{-d_2} + 2^{-d_3} + \dots + 2^{-d_n}$$

Ähnlich verfährt in einer neueren Arbeit im Schweizerischen Genealogischen Jahrbuch auch *V. Theus* (3). Demgegenüber schreibt *H.F. Friederichs* im Handbuch der Genealogie (4): "Die weit verbreitete biometrische Theorie, daß die Erbwirksamkeit mit der stellenmäßigen Entfernung eines Merkmalsträgers vom Probanden abnehme, erfordert einen ausdrücklichen Hinweis auf ihre Unhaltbarkeit". Wie sieht es nun wirklich aus?



Ahn Nr. 20: 20 (dezimal) = 10100 (dual);
 10100 \cong 4. Generation

Filiationlinie: 1 = Proband (Nr. 1)
 10 = Vater (Nr. 2 = 1 · 2 + 0)
 101 = Großmutter (Nr. 5 = 2 · 2 + 1)
 1010 = Urgroßvater (Nr. 10) 3
 10100 = Ururgroßvater (Nr. 20)

Bild 1. Kekulé – Benummerung

Die Vererbung ist ein Zufallsvorgang. Konkrete Zufallsprobleme sind meist leichter zu mathematisieren, wenn man ein Urnenmodell o.dgl. angeben kann. Das ist hier möglich. In der Wirklichkeit wird ja eine zufällig "ausgewählte" Hälfte der Gene mit den Chromosomen vom Elter an das Kind weitergegeben, während die andere (ebenfalls zufällige) Hälfte vom andern Elter stammt. Wir stellen demgemäß den Chromosomen-Vorrat eines Individuums durch Urnen mit insgesamt k Paaren aus schwarzen und weißen Kugeln dar. Beim Menschen ist $k = 23$. Von diesen Chromosomen sind

aber 2, nämlich das X- und das Y-Chromosom, geschlechtsgebunden und unterliegen einem besonderen Vererbungsmechanismus, wie A. Richter im einzelnen (5) nachgewiesen hat. Das hat zur Folge, daß die Vererbung der 44 Autosomen in der Urne mit $k = 22$ gesondert von derjenigen der 2 Geschlechtschromosomen betrachtet werden muß, aber auch die Folge, daß im Individualfall für jede Filiationlinie ein differenzierter Erbgang eintritt.

In dem folgenden Urnenmodell seien die schwarzen Kugeln die vom Stammelter überkommenen Autosomen, die weißen Kugeln die durch Zeugung

nach und nach ergänzten fremden Autosomen. Die den Stamm-Elter repräsentierende Urne enthält also anstelle der Autosomen anfänglich 22 Paare von je 2 schwarzen Kugeln. Der Urne wird zunächst das erste Paar entnommen, aus ihm blindlings eine Kugel gewählt und durch eine weiße Kugel wieder zu einem Paar ergänzt; dieses schwarz-weiße Kugel-paar kommt in die erste Folgeurne. Dasselbe geschieht auch mit allen anderen Kugelpaaren der Anfangsurne. Aus der 1. Folgeurne ziehen wir nun nach und nach alle schwarz-weißen Paare, wählen aus jedem Paar blindlings eine Kugel aus und ergänzen die verbleibende Kugel mit einer weißen Kugel wieder zu einem Paar. Weiß-weiße Paare scheiden aus, da sie ja kein Erbgut des Stammelters

mehr enthalten. Die schwarz-weißen Paare kommen in die 2. Folgeurne. Wenn die 1. Folgeurne auf diese Weise geleert ist, tritt die 2. Folgeurne an die stelle der 1. Folgeurne usw. Als einführendes Beispiel betrachten wir zunächst das einfache, allerdings numerisch unrealistische Urnenmodell mit anfänglich 2 Paaren zu je 2 schwarzen Kugeln (Bild 2).

Man sieht, daß die Möglichkeit einer leeren Folgeurne von Generation zu Generation zunimmt, d.h. daß die genetische Verwandtschaft allmählich abstirbt. In keiner Urne können sich mehr als halb so viele schwarze Kugeln wie in der Ausgangsurne befinden. An das Schema sind sowohl die Überganganteile als auch die resultierenden Anteile an-

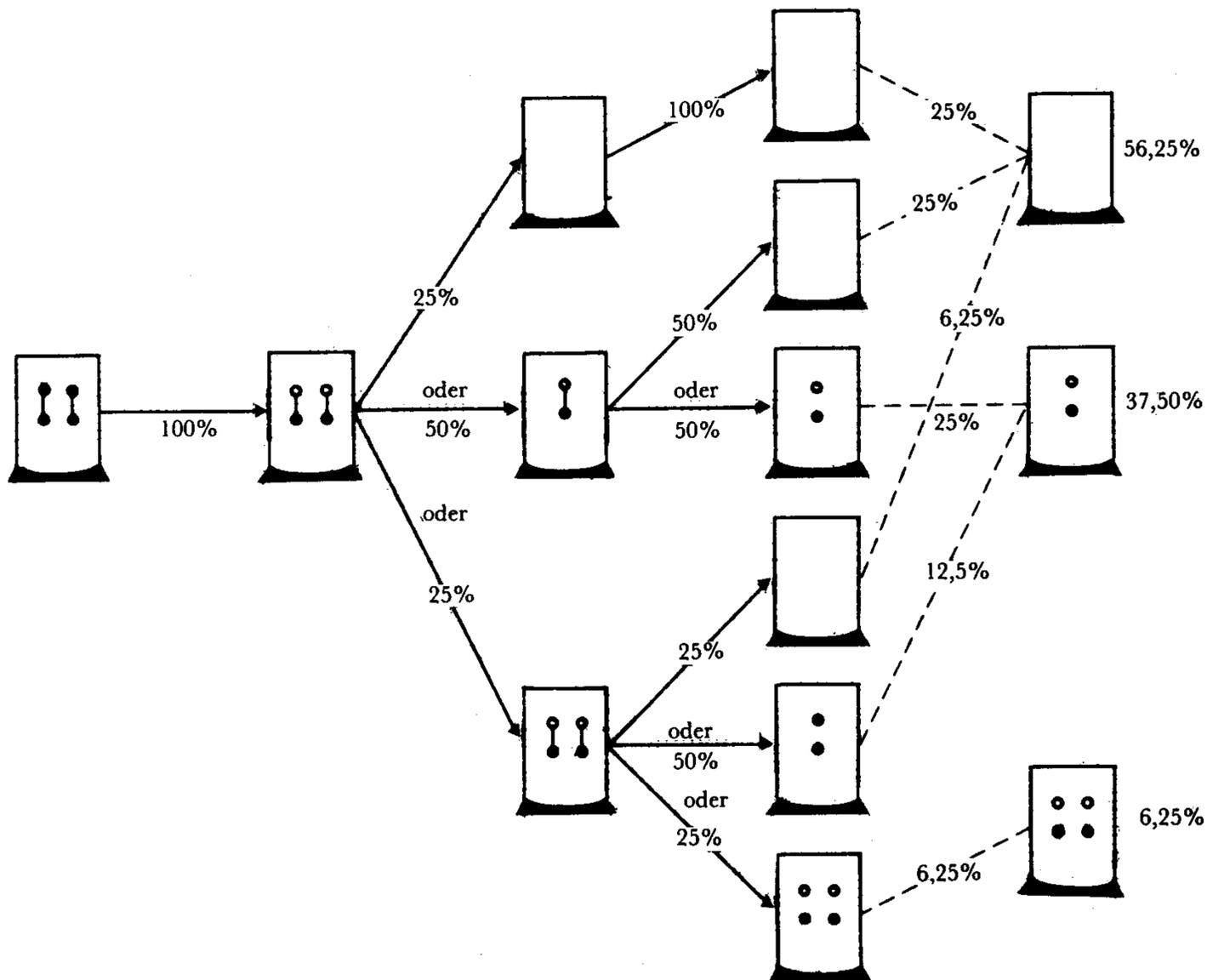


Bild 2. Vererbungsmodell

Hermann Athen

geschrieben. In diesem Beispiel führt der Autosomenaustausch nach drei Generationen in 56,25% der Fälle zu genetischer Nicht-Verwandtschaft, in 37,50% der Fälle zur Viertel- und in 6,25% der Fälle zur Halb-Verwandtschaft. In Kurzform schreiben wir dies als einen *Zustandsvektor* auf (56,25% = 0,5625):

$$z = (0,5625; 0,3750; 0,0625; 0; 0)$$

Im allgemeinen Falle einer Anfangsurne mit $2k$ schwarzen Kugeln ergeben sich die Überganganteile von einer Urne mit i schwarz-weißen Kugelpaaren in eine Folgeurne mit j schwarz-weißen Kugelpaaren als Glieder einer Matrix:

$$M = \{p_{ij}\}; \quad p_{ij} = \binom{i}{j} \cdot 2^{-i}, \text{ wo } i, j \in [0; 22].$$

Somit finden wir die Zustandsvektoren der Reihe nach für beliebige Generationen n durch Matrix-Multiplikation:

$$\begin{aligned} z_0 &= (0; 0; 0; \dots; 1) \\ z_1 &= z_0 \cdot M = (0; 0; \dots; 1; \dots; 0; 0) \\ z_2 &= z_1 \cdot M \quad \text{usw.} \end{aligned}$$

Im Modellfall mit $2k = 4$ folgt auf diese Weise:

$$\begin{aligned} z_0 &= (0 \quad ; 0 \quad ; 0 \quad ; 0; 1) \\ z_1 &= (0 \quad ; 0 \quad ; 1 \quad ; 0; 0) \\ z_2 &= (0,25 \quad ; 0,50 \quad ; 0,25 \quad ; 0; 0) \\ z_3 &= (0,5625; 0,3750; 0,0625; 0; 0) \\ z_4 &= (0,7656; 0,2188; 0,0156; 0; 0) \\ z_5 &= (0,8789; 0,1172; 0,0039; 0; 0) \quad \text{usw.} \end{aligned}$$

Der Prozeß verschiebt sich demnach immer mehr zum Zustand der Nicht-Verwandtschaft.

Interessant sind nun die Mittelwerte der jeweils verbleibenden Autosomen: In z_0 sind es $1 \cdot 4 = 4$ Chromosomen; in z_1 sind es $1 \cdot 2 = 2 = 4 \cdot 2^{-1}$ Chromosomen; in z_2 sind es $0,5 \cdot 1 + 0,25 \cdot 2 = 1 = 4 \cdot 2^{-2}$ Chromosomen; in z_3 sind es $0,375 \cdot 1 + 0,0625 \cdot 2 = 0,5 = 4 \cdot 2^{-3}$ usw. Das Halbierungsgesetz $1/2; 1/4; 1/8; \dots$ gilt also sicher für die Mittelwerte der verbliebenen Autosomen, was auch mathematisch allgemein bewiesen werden kann.

Generell gilt demnach: Von den $2k = 44$ freien Autosomen des Stamm-Elters verbleiben in der n^{ten} Folgegeneration im Mittel $2k \cdot 2^{-n}$ Autosomen. Für die Autosomen-Anzahl $2k = 44$ findet man mit Hilfe eines Klein-Computers die mittlere Anzahl m_i der in der Generation i verbliebenen Autosomen und den prozentualen Anteil p_i der abgestorbenen Verwandtschaft, bei der keine Autosomen des Stamm-Elters mehr feststellbar sind:

i	0	1	2	3	4	5	6	7	8	9	10
m_i	44	22	11	5,5	2,75	1,375	0,6875	0,3478	0,1719	0,0859	0,04
p_i (%)	0	0	0	0,18	5,30	24,18	49,73	70,72	84,15	91,75	95,7
$3s_i$	0	0	7,04	6,09	4,65	3,41	2,45	1,75	1,24	0,88	0,62
$h_i^{(m)}$	-	100	75	50	37,5	28,1	21,9	17,2	13,7	10,9	8,8
$h_i^{(w)}$	-	100	75	62,5	50	40,6	32,8	26,6	21,5	17,4	14,1

Die "Autosomen-Verwandtschaft" ist also nach der 6. Generation eher unwahrscheinlich als wahrscheinlich. Der Streubereich $3s_i$ (dreifache Standardabweichung) zeigt an, daß die Mittelwerte m_i fast generell auf 0, d.h. auf den verwandtschaftlichen Generationenabstand ∞ redu-

ziert werden. Die Zeilen $h_i^{(m)}$ und $h_i^{(w)}$ für männliche bzw. weibliche Probanden zeigen an, durch wieviel % der Ahnen (ohne Implex) noch eine Erhöhung des Mittelwertes m_i infolge X- oder Y-Vererbung erfolgen kann; teilweise kommt man dabei auf ein Mehrfaches von m_i .

Die Tabelle zeigt also letztendlich, daß mit genetischen Prinzipien eine Definition des Verwandtschaftsgrades mit letzter Schärfe nicht zu fassen ist. Dabei wirkt das Crossing-Over, welches mathematisch zu einer fiktiven Vermehrung der Chromosomenzahl führt, zusätzlich im diffusen Sinne ein. Auf jeden Fall nimmt aber in jeder Filialionslinie auf dem Wege vom Stammelter zum Probanden die Wahrscheinlichkeit gemeinsamer Erbanlagen in der Ahnenfolge ab. Insbesondere zeigt die Betrachtung der "Autosomen-Verwandtschaft", daß die statistische Definition des Verwandtschaftsgrades noch am zweckmäßigsten ist, weil sie sich mathematisch als leicht und exakt quantifizierbar erweist. Damit ist also die Definition des Verwandtschaftsgrades von *S. Rösch* durchaus berechtigt und die Einlassung von *H. F. Friedrichs* wissenschaftlich nicht begründet, jedenfalls nicht im Hinblick auf die oft recht mißverständlichen genetischen Behauptungen.

3. Welche Probleme der Genealogie kann man mathematisieren?

Prinzipiell gibt es in dieser Hinsicht kaum eine Grenze. Ich möchte zwar einige charakteristische Probleme aufzählen, ohne dabei aber Vollständigkeit zu beanspruchen und ohne sie im einzelnen zu behandeln. Einige Probleme habe ich selber gelöst, aber noch nicht veröffentlicht, andere sind seit längerem oder kürzerem in der Literatur bekannt. Zu den ganz oder teilweise gelösten Problemen gehören:

- (1) Der Implex in der Aszendenz.
- (2) Die Generationenverschiebung in der Aszendenz.
- (3) Die verwandtschaftliche Distanz und ihr Grenzwert in genetischstatistischer Hinsicht.
- (4) Das Filialionsproblem in genetischer Behandlung.
- (5) Das Deszendenzproblem mit Implex.
- (6) Das Aussterben von Familiennamen.
- (7) Das Aussterben von Filialionslinien.
- (8) Ahnen- und Nachkommenstatistiken.

- (9) Zusammenhänge zwischen Aszendenz und Deszendenz.
- (10) Simulationen genealogischer Kollektive mit Computern.
- (11) Genealogische Schlußfolgerungen aus Namenshäufigkeiten. – Usw.

Die vorstehenden Probleme erfordern z.T. erheblichen mathematischen Aufwand. Es sollen nur noch einige theoretische Aspekte von (1) und (2) dargestellt werden.

4. Der Implex und die Generationenverschiebung als Beispiele der Theoretischen Genealogie

Im Idealfall wächst die Ahnenzahl eines Probanden von Generation zu Generation auf das Doppelte an, d.h. für das Verhältnis der Ahnenzahl A_{n+1} in der $(n + 1)$ ten Generation zu der Ahnenzahl A_n in der vorangehenden n ten Generation gilt mit dem Vervielfachungsfaktor 2 die sog. *Differenzgleichung*

$$\frac{A_{n+1}}{A_n} = 2 \quad \text{oder} \quad A_{n+1} = 2A_n$$

Nach k Generationen beträgt also die Ahnenzahl

$$A_k = 2^k \quad (2^k\text{-Gesetz}),$$

und die Gesamtzahl aller Ahnen von der 0ten bis zur k ten Generation

$$\bar{A}_k = 2^{k+1} - 2 = 2(2^k - 1).$$

Jeder weiß, zu welchen astronomischen Zahlen dieser Ansatz führt: Bis zum Jahre 0 hätte ein heute lebender Proband mehr als 150 Trillionen ($= 1,5 \cdot 10^{20}$) Ahnen zählen müssen, eine Anzahl, die mehr als 35 Milliarden mal so groß ist wie die heutige Weltbevölkerung, und die als 6000 Lichtjahre lange Warteschlange fast bis zum Rand der Milchstraße reichen würde. In Wirklichkeit ist die Ahnenzahl pro Generation natürlich viel kleiner als die obige Formel angibt, und sie reicht in den wenigsten Fällen über einige Tausend hinaus. Das bedeutet: Der Vervielfachungsfaktor ist i.a. kleiner

Hermann Athen

als 2 und strebt mit wachsender Generationennummer immer mehr gegen 1.

Wir formulieren diesen Sonderfall eines sog. *logistischen Prozesses* zunächst allgemein mit einer Differenzgleichung

$$\frac{A_{n+1}}{A_n} = a + b \cdot f(A_n; A_{n-1}; \dots; A_0),$$

wo a und b noch unbekannte Konstanten und f eine noch zu bestimmende Funktion bedeuten. Hierin ist mit a = 2 und b = 0 bzw. f ≡ 0 auch der obige Idealfall enthalten. Zur Vereinfachung nehmen wir an, daß f nur von A_n abhängt, d.h. daß stets

$$f(A_n; A_{n-1}; \dots; A_0) = f(A_n).$$

Nach dem logistischen Postulat nimmt A_n mit immer größer werdender Generationennummer n nicht mehr zu, d.h. A_n hat einen Grenzwert N, und es kann A_{n+1} = A_n gesetzt werden, wenn n genügend groß ist. Es gilt also in diesem Falle

$$1 = a + b \cdot f(N) \quad \text{oder} \quad b = \frac{1-a}{f(N)}$$

und daher

$$\frac{A_{n+1}}{A_n} = a + (1-a) \cdot \frac{f(A_n)}{f(N)}.$$

Aus diesem Ansatz lassen sich viele Näherungslösungen des Aszendenz-Problems entwickeln, von denen hier einige angedeutet werden sollen.

(1) Im Idealfall ist a = 2 und f ≡ 0. Wir behalten davon nur die Forderung a = 2 bei und erhalten

$$\frac{A_{n+1}}{A_n} = 2 - \frac{f(A_n)}{f(N)}$$

Die weitere Vereinfachung f(A_n) = A_n ergibt

$$(*) \quad \boxed{\frac{A_{n+1}}{A_n} = 2 - \frac{A_n}{N}}$$

Um jetzt diesen vereinfachten logistischen Prozeß in Gang zu setzen, erinnern wir uns daran, daß die Ahnenanzahl in der 1. Generation immer 2 betragen muß (jeder Proband hat Vater und Mutter). Mit A₁ = 2 kommt

$$\frac{A_2}{A_1} = 2 - \frac{A_1}{N} \quad \text{oder} \quad A_2 = 4\left(1 - \frac{1}{N}\right).$$

Daraus kann man A₂, dann A₃ usw., also schließlich A_n berechnen:

$$\boxed{A_n = N \cdot \left[1 - \left(1 - \frac{2}{N}\right)^{2^{n-1}}\right]}$$

Dies ist der einfachste denkbare Ansatz und m.W. auch die einzige in der Literatur bisher bekannt gewordene ernsthafte Lösung des Implex-Problems. Auf diese – auf ganz anderem Wege gewonnene – Lösung von H. v. Schelling (6) kommen wir weiter unten nochmals zurück.

(2) Noch etwas näher an die Wirklichkeit kommen wir durch die Überlegung, daß "jeder sein eigener Ahn" ist, d.h. daß A₀ = 1 und zugleich A₁ = 2 sein muß. Dann gilt

$$\frac{A_{n+1}}{A_n} = \frac{2}{1} = a + (1-a) \cdot \frac{f(1)}{f(N)}$$

$$a = \frac{2f(N) - f(1)}{f(N) - f(1)} \Rightarrow \text{und}$$

$$\frac{A_{n+1}}{A_n} = 2 - \frac{f(A_n) - f(1)}{f(N) - f(1)}$$

Zur Vereinfachung setzen wir wieder f(A_n) = A_n und erhalten als verbesserte Annäherung die Differenzgleichung

$$(**) \quad \boxed{\frac{A_{n+1}}{A_n} = 2 - \frac{A_n - 1}{N - 1}}$$

(3) Eine dritte, noch bessere Näherung ergibt

sich aus der Forderung, daß nach deutschem und nach anderem Recht Geschwister und Halbgeschwister nicht heiraten dürfen. Daraus folgt, daß auch noch $A_2 = 4$ als bekannte Voraussetzung in den Ansatz eingeht. Ohne Kommentar führen wir das Ergebnis an:

$$(***) \quad \frac{A_{n+1}}{A_n} = 2 - \left[\frac{(A_n - 1)(A_n - 2)}{(N - 1)(N - 2)} \right]^t$$

wo t noch eine weitere Anpassungszahl ist und von einer bestimmten Größe ab eine Oszillation von A_n um N bewirkt. Es soll darauf hier aber nicht eingegangen werden.

Die drei Annäherungen mögen an einer realen Ahnentafel (Prinz von Isenburg) mit $N = 400$ "gemessen" werden.

n	P.v.I	(*)	(**)	(***) t=0,5	(***) t=0,4	(***) t=0,3	(*)' mit variabl. N
0	1	1	1	1	1	1	N
1	2	2	2	2	2	2	70
2	4	4	4	4	4	4	70
3	8	8	8	8	8	8	70
4	16	16	16	16	15	15	70
5	28	31	31	31	30	28	100
6	48	59	60	60	56	50	160
7	86	109	110	111	101	87	275
8	159	189	191	191	169	139	380
9	245	289	291	292	253	204	510
10	365	369	370	371	332	273	600
11	500	-	-	-	-	-	↓
12	767	-	-	-	-	-	1000

Der Trend der realen Ahnentafel wird im ganzen gut erkennbar wiedergegeben. Man erkennt aber, daß mit einem konstanten N für alle Generationen eine völlige Anpassung nicht möglich zu sein scheint. Dafür gibt es auch einleuchtende Gründe:

a) Die vorstehenden Formeln sind deterministisch, d.h. sie lassen für Variationen bei verschiedenen Ahnentafeln kaum Möglichkeiten offen. In Wirklichkeit ist aber die Entstehung einer Ahnentafel ein Zufallsprozeß. Es kann sich also bei den durch die obigen Formeln errechneten Zahlen nur um Mittelwerte handeln, ähnlich wie es oben beim Chromosomen-Modell war. In der Tat hat H. v. Schelling (6) ein Urnenmodell (s.u.) entwickelt,

welches, wie schon erwähnt, auf die erste Näherung (*) führt. Berechnet man die dazugehörigen Streuungen, so ergibt sich eine ausreichend hohe Variabilität. Auf die Vertiefung dieses Problems soll hier nicht eingegangen werden.

b) Alle Näherungsformeln setzen voraus, daß die Heiratskreise in den verschiedenen Generationen den gleichen Umfang haben, d.h. daß die Bevölkerung mit einer Art von Zwei-Kindersystem konstant bleibt. Das ist i.a. nicht der Fall. Zudem heiraten immer wieder Heiratsfähige aus dem Heiratskreis heraus. Das bedeutet praktisch eine Vergrößerung der Heiratskreise von N_1 auf N_2 , dann auf N_3 , N_4 usw. (Bild 3). Es ist nicht besonders

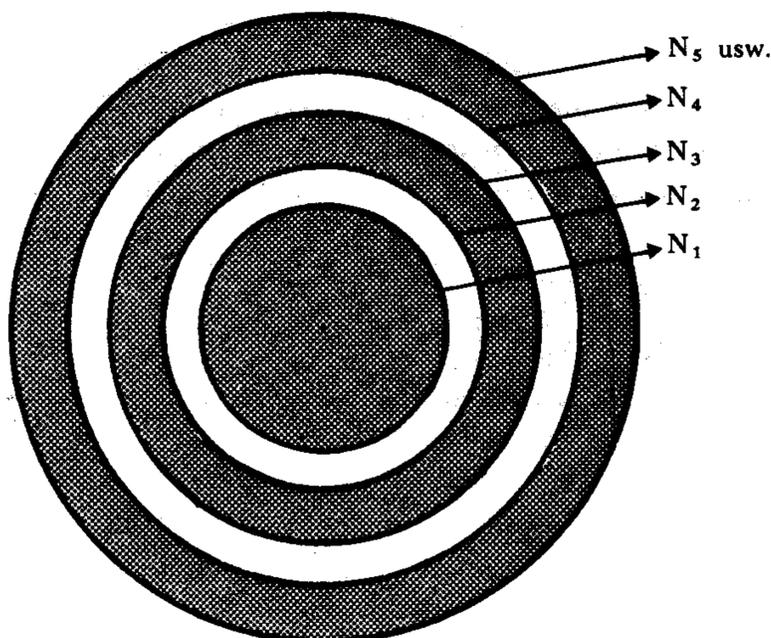


Bild 3. *Expandierende Heiratskreise*

schwierig, diesen Prozeß in die obigen Formeln (*), (**), (***) einzubauen. Man erreicht damit eine ausgezeichnete Anpassung an reale Situationen (vgl. die letzte Spalte der Tabelle). Doch soll hier darauf im einzelnen nicht eingegangen werden. Eine Formel für k Kreise und $100 p$ % "Abwanderer" lautet:

$$A_u = 2 N_0 \cdot \left[(1 + k) - \left(1 - \frac{1}{N_0} \right)^{1 + (2^{n-1}) \cdot (1-p)} - k \left(1 - \frac{1}{k \cdot N_0} \right)^{(2^{n-1}) \cdot p} \right]$$

Eine weitere interessante, mathematisch erfaßbare Erscheinung ist die *Generationenverschiebung*. Die bisherigen Formeln zeigen Ahnengleichheit jeweils in derselben Generation. In der Ahnentafel des Prz. v. Isenburg z.B. entfallen auf die 1024 Ahnenplätze nur 365 verschiedene Ahnen; im Mittel kommt also jeder Ahn etwa 3mal vor. Es ist aber bekannt, daß derselbe Ahn auch in verschiedenen Generationen vorkommen kann (sog. *Generationenverschiebung*). Wie ist das zu erklären und mathematisch zu erfassen?

Angenommen, ein Ahn habe vor 400 Jahren gelebt (wir sagen: er hat das Ahnenalter 400); ein

Proband möge mit ihm u.a. in der reinen Tochter- und in der reinen Sohneslinie verwandt sein. Ferner mögen die Töchter ihre jeweiligen Töchter im Alter von 20 Jahren, die Söhne ihre jeweiligen Söhne im Alter von 50 Jahren bekommen. Dann umfaßt die Tochterlinie $400:20 = 20$ Generationen, die Sohneslinie aber $400:50 = 8$ Generationen. Denkbar wären darüber hinaus auch "Misch"-Linien, z.B. 10 Generationen von Tochter zu Tochter und 4 Generationen von Sohn zu Sohn. Während es nun nur eine reine Sohneslinie und nur eine reine Tochterlinie gibt, kann man für diese Mischlinie ca. 1000 Möglichkeiten angeben. In Wirklichkeit kommen aber die Generationenabstände 20, 21, 22, ..., 48, 49, 50 vor. Mit kombinatorischen Hilfsmitteln der Mathematik ist es möglich auszurechnen, wie viele Möglichkeiten es gibt, mit diesen Abständen Filiationslinien von 8, 9, 10, ..., 18, 19, 20 Generationen zu bilden. Damit kann man dann aber auch die Wahrscheinlichkeit berechnen, wie oft ein Ahn jeweils in der möglichen Generationen auftreten kann. Bildlich sind die Dinge schematisch darstellbar. (Bild 4). Rechnet man ohne *Generationenverschiebung*, dann kommen alle (als praktisch gleichaltrig betrachteten) Ahnen etwa gleich oft in der "Schwerpunktgeneration" vor.

Infolge der geschilderten Generationenverschiebung "wandern" aber gewisse Teile nach oben oder nach unten; für jeden Ahn entsteht so ein mehrere Generationen überdeckendes rhombisches Gebiet. In die entstehenden Lücken schieben sich die Rhomben von Ahnen aus anderen Schwerpunktgenerationen, wie angedeutet. Modellmäßig könnte man sich einen "Würfel" mit den Augenzahlen $a + k = 20, 21, \dots, 49, 50$ vorstellen. Man wirft $n = 8, 9, \dots, 19, 20$ Würfel dieser Art zugleich und zählt die Würfe, welche die Augensumme $A = 400$ aufweisen. Dieses Modell ist mathematisch beherrschbar, allerdings nur mit einem erheblichen Aufwand an Theorie. Es soll daher hier nur die Endformel für die Anzahl $f_A^{(n)}$ der Möglichkeiten angegeben werden, mit n Würfeln, die die Augenzahlen

$$f_A^{(n)} = \sum_{j \leq n} (-1)^j \cdot \binom{n}{j} \binom{\frac{1+k}{2} \cdot n - 1 - kj - |A - na - \frac{1-k}{2} \cdot n|}{n-1}$$

In Wirklichkeit sind aber die Generationenabstände nicht gleichwahrscheinlich. Der Altersunterschied zwischen Kind und Eltern liegt häufiger bei 30 als bei 20 oder 40 Jahren. Aus empirischen Erhebungen ergab sich ein Mittelwert von $\mu \approx 30$ mit einer Standardabweichung von $\sigma \approx 8,9$ in einer angenäherten Normalverteilung. Besser wäre zwar eine sog. schiefe Log - Normalverteilung, deren Maximum bei etwa 30 Jahren liegt, d.h. der Wirklichkeit entsprechend sehr viel näher am unteren Randwert 20 als am oberen Randwert 50. Im allgemeinen genügt aber die Normalverteilung

$a + 1, a + 2, \dots, a + k$

$$\varphi(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2}$$

tragen, auf einen Wurf die Augensumme A zu werfen:

mit $\mu \approx 30, \sigma \approx 8, 9$.

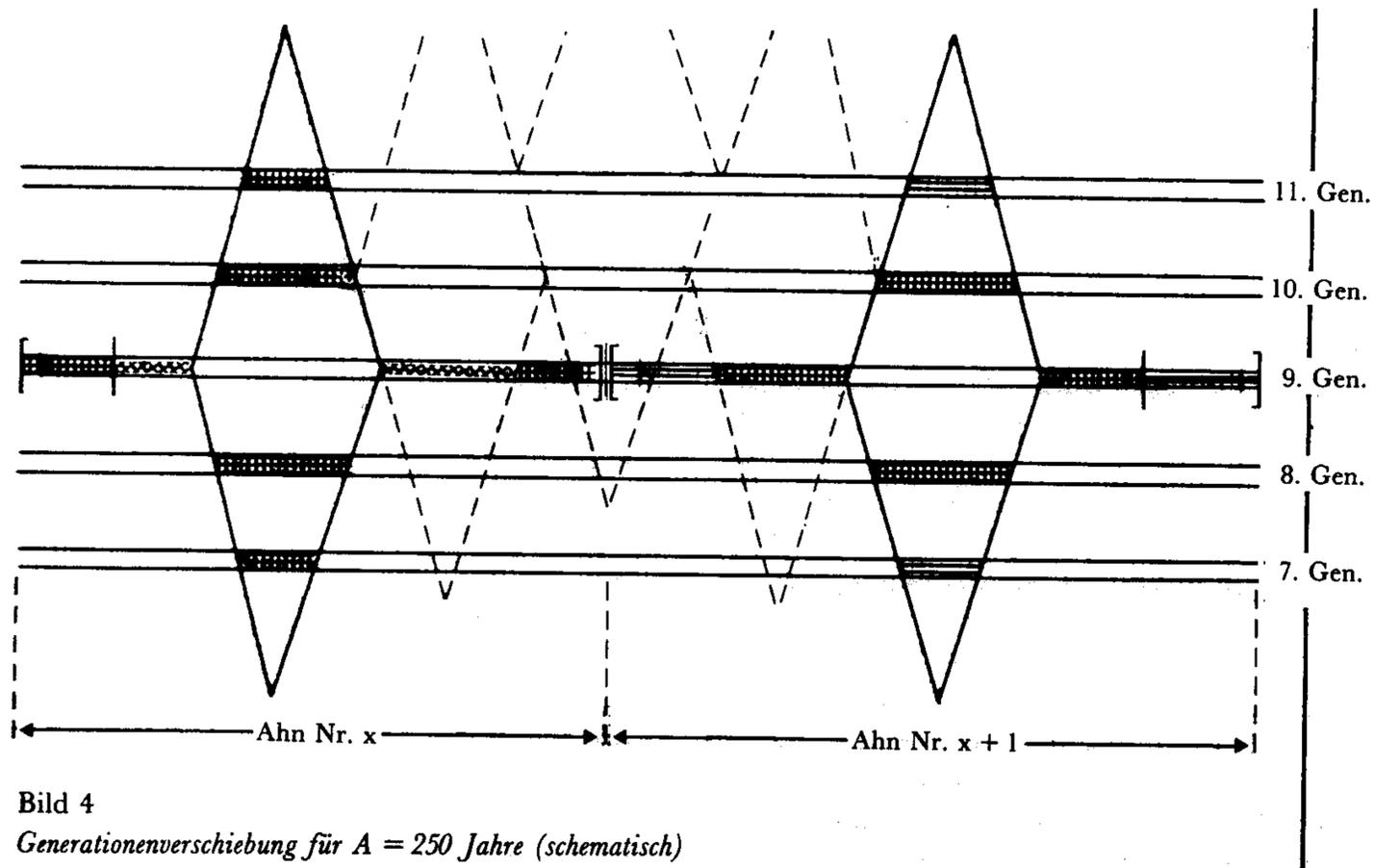


Bild 4
Generationenverschiebung für $A = 250$ Jahre (schematisch)

Hermann Athen

Mit diesen Voraussetzungen geht die obige kombinatorische Formel für $f_A^{(n)}$ über in die exaktere und bequemere Verteilung

$$f_A^{(n)} = \frac{2^n}{\sqrt{2\pi n} \cdot \sigma} \cdot e^{-\frac{1}{2n} \cdot \left(\frac{A-n\mu}{\sigma}\right)^2};$$

$\sigma \approx 8, 9; \mu \approx 30$

Aus der Theorie ergeben sich einige *Faustformeln*:

Schwerpunktgeneration: $s = A/28$

Überdeckte Anzahl a_s von Generationen:

$$a_s = A/50$$

Gesamtzahl der von *einem* Ahn besetzten Plätze:

$$z_s = 2^s/A_s$$

[A_s nach einer der Formeln (*) bis (***)].

Ein durchgerechnetes Beispiel bezieht sich auf Kaiser Maximilian I. (1459–1519), für den die in folgender Tabelle aufgeschriebenen Zahlen z_i angeben, wie oft in der betreffenden Generation n Karl der Große als sein Ahn vorkommt (im Vergleich von Wirklichkeit und Theorie):

n	20	21	22	23	24	25	26	27	28	29	30	31
z_i (emp.)	2	26	147	586	1464	1948	1843	1077	566	188	66	16
z_i (theor.)	2	22	161	623	1405	1975	1822	1157	527	178	46	9

Es soll noch hinzugefügt werden, daß aus den vorstehenden Betrachtungen folgender Grenzwert gefolgert werden kann: Der *verwandtschaftliche Generationsabstand* (s.o.) kann bei keinem Ahn eines beliebigen Probanden den Grenzwert

$$d^* = \frac{\ln N_\infty}{\ln 2}$$

überschreiten. Im Falle Maximilians I. ist $N_\infty \approx 3500$, d.h. $d^* \approx 11,8$, ein Wert, den S. Rösch (7) aus den noch vorhandenen Unterlagen für Maximilian und andere Carolinger bestätigen konnte. Aus der obigen Tabelle geht hervor, daß auch die angegebenen Faustformeln befriedigend erfüllt werden.

6. Zufallsmodelle

Die vorstehenden Betrachtungen lassen sich durch ähnliche Zufallsmodelle wie im § 2 simulieren. In Abwandlung eines Urnenmodells von H. v. Schel-

ling (6) denken wir uns folgenden computerisierten Zufallsprozeß:

Für jede Generation wird die Menge der Männer bzw. die Menge der Frauen, aus der ein Ahn dieser n^{ten} Generation gewählt werden kann, durch die natürlichen Zahlen von $N_1^{(n)}$ bis $N_2^{(n)}$ dargestellt. Aus dieser Menge wählt der Computer mit einem leicht einprogrammierbaren Zufallsgenerator nach dem Zufall abwechselnd 2^{n-1} Zahlen für die Männer und 2^{n-1} Zahlen für die Frauen aus. Alle diese Zahlen werden gesondert für die Männer und für die Frauen gespeichert, und nach Abschluß des Auswahlprozesses zählt der Computer, wie viele männliche bzw. weibliche Ahnen auf diese Weise fixiert worden sind. Die Summe der beiden so ermittelten Anzahlen ist dann die Anzahl der verschiedenen Ahnen in dieser Generation. Jetzt wird derselbe Vorgang für die nächste Generation – jedoch unter Auswahl von 2^n Zahlen – mit einer anderen Zahlenmenge $[N_1^{(n+1)}; N_2^{(n+1)}]$ wiederholt; dabei gilt $N_1^{(n+1)} = N_2^{(n)} + 1$, und außerdem hat die neue Zahlenmenge denselben Umfang wie die der vorangehenden Generation. So werden alle Gene-

rationen durchgespielt. Mit Hilfe der Wahrscheinlichkeitsrechnung läßt sich beweisen, daß dieses Modell im Mittel genau durch Formel (*) wieder-

gegeben wird. Ein entsprechender praktischer Versuch, der 6-mal wiederholt wurde, hatte folgendes Ergebnis:

n	Versuch Nr.						Mittelwert (1)-(6)	Nach (*)	Streubreich
	(1)	(2)	(3)	(4)	(5)	(6)			
1	2	2	2	2	2	2	2,0	2,0	0
2	4	4	4	4	4	4	4,0	4,0	3,4- 4,0
3	8	8	8	6	8	8	7,7	7,9	6,4- 8,0
4	16	16	16	14	14	14	15,0	15,4	12,3- 16,0
5	28	28	30	28	32	30	29,3	29,6	23,7- 32,0
6	52	54	56	48	52	54	52,7	54,6	44,1- 64,0
7	88	88	86	94	104	96	92,7	93,5	77,5-108,5
8	128	138	136	148	134	146	138,3	141,0	122,4-159,6
9	180	186	180	170	168	184	178,0	177,3	164,2-190,0
10	190	190	189	190	190	188	189,5	189,2	185,4-190,0

Die empirischen Werte der einzelnen Versuche liegen nicht nur voll im zugelassenen theoretischen Streubereich, sondern auch die Mittelwerte stimmen gut überein.

Eine erste Verallgemeinerung dieses Modells ergibt sich, wenn zwar die Forderung $N_1^{(n+1)} = N_2^{(n)} + 1$ beibehalten wird, aber die Umfänge $[[N_1^{(i)}; N_2^{(i)}]]$ der Zahlenmengen von Generation zu Generation bis zu einem endlichen Höchstwert N_∞ gesteigert werden. Im übrigen wählt und zählt der Computer die Ahnen aus wie im Grundmodell. Dieses letztere Zufallsmodell realisiert expandierende Heiratskreise mit variablem, aber beschränktem Umfang (Tabelle auf S. 263, letzte Spalte; s. auch Bild 3).

Auch die Generationenverschiebung läßt sich durch eine weitere Verallgemeinerung erfassen. Auch hier wachsen die Bereiche $[N_1^{(n)}; N_2^{(n)}]$, $[N_1^{(n+1)}; N_2^{(n+1)}]$, ... wie im vorangehenden Modell bis zu einem endlichen Höchstumfang an. Aber bei der Ziehung der Zahlen stehen nicht gleich sämtliche Zahlen des jeweiligen Bereichs $[N_1^{(n)}; N_2^{(n)}]$ zur Verfügung, sondern immer nur gewisse Teilbereiche. In der 10. Generation könnte man beispielsweise den Bereich [800; 1000] zugrunde legen, jedoch für die ersten 64 Ziehungen von den 512

erforderlichen Ziehungen nur den Teilbereich [800; 860] verwenden, für die nächsten 64 Ziehungen (Nr. 65 bis 128) den Teilbereich [820; 880], ..., für die letzten Ziehungen (Nr. 449 bis 512) den Teilbereich [940; 1000]. Für die 11. Generation mit 1024 Ziehungen beginnt der Gesamtbereich bei 800 und endet bei z.B. 1120. Daher verwendet man jetzt für die ersten 128 Ziehungen den Bereich [820; 910], für die Ziehungen Nr. 129 bis 256 den Bereich [850; 940], für die letzten 128 Ziehungen (Nr. 897 bis 1024) den Bereich [1030; 1120]. Auf diese Weise kann dieselbe Modellzahl, die ja einen bestimmten Ahn repräsentiert, in verschiedenen Generationen vorkommen. Das Bild 5 gibt die sich so übergreifenden Teilbereiche schematisch wieder. Bei der Festlegung der Gesamt- und der Teilbereiche, die im vorangehenden nur schematisiert werden konnte, muß man natürlich darauf achten, daß die von der Theorie geforderten numerischen Zusammenhänge gewahrt werden. (Illustration auf der nächsten Seite).

Alle diese Modelle und andere ähnliche Varianten lassen sich wahrscheinlichkeitstheoretisch beherrschen. Eine Erörterung im einzelnen würde jedoch den vorliegenden Rahmen sprengen.

Hermann Athen

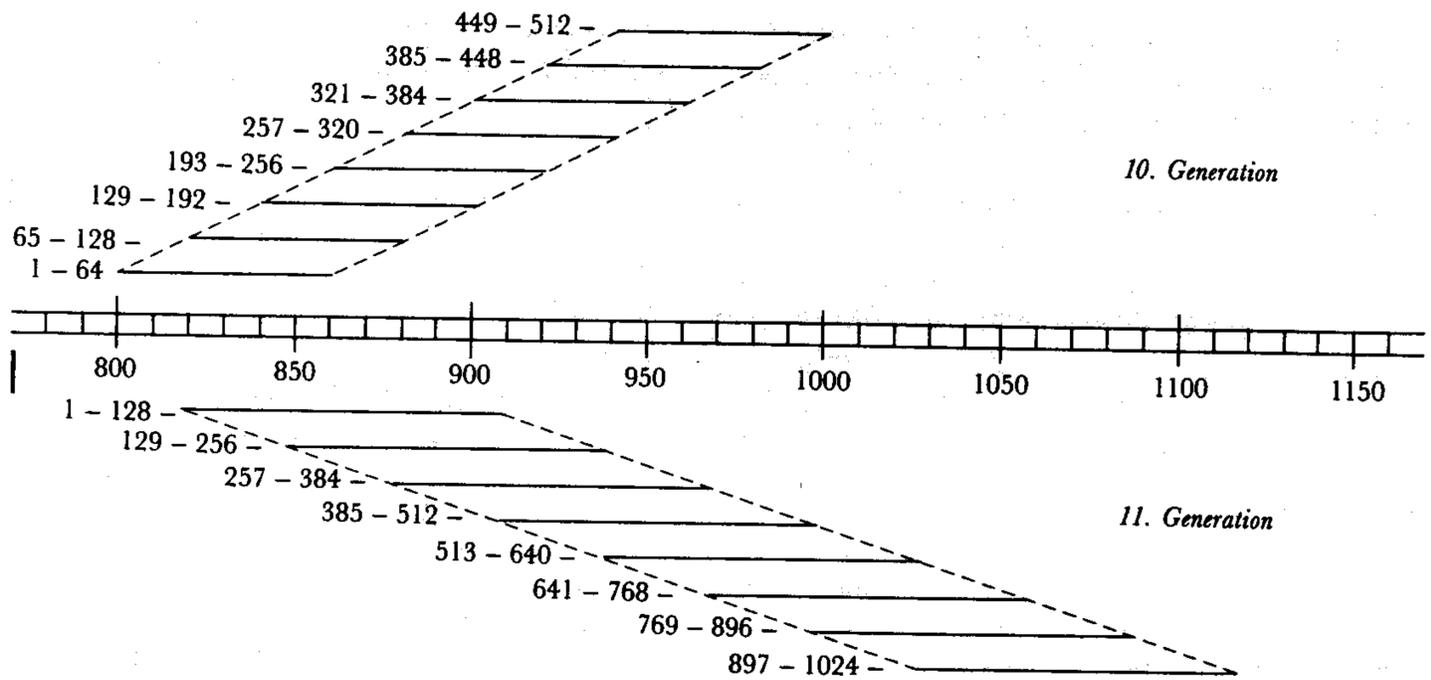


Bild 5 Simulationsmodell

7. Literaturangaben

1. S. Rösch, Grundzüge einer quantitativen Genealogie. Neustadt/Aisch 1955.
2. S. Rösch, Über den Verwandtschaftsgrad. In: Familie und Volk 6, 1957, S. 313/317.
3. V. Theus, Die genetische Verwandtschaft. In: Schweizerische Gesellschaft für Familienforschung. Jahrbuch 1978, S. 1/34.
4. H. F. Friederichs, Humangenetik. In: Handbuch der

Genealogie, Hrsg. Henning und Ribbe, S. 231/235. Neustadt/Aisch 1972.

5. A. Richter, Erbmäßig bevorzugte Vorfahrenlinien bei zweigeschlechtigen Lebewesen. In: Archiv für Sippenforschung 45, 1979, 74, S. 97/109.

6. H. v. Schelling, Studien über die durchschnittliche verwandtschaftliche Verflechtung innerhalb einer Bevölkerung. Fischer Jena 1945.

7. S. Rösch, Caroli Magni Progenies, Pars I. Neustadt/Aisch 1977.