

# Zufallsmatrizen, das Zählen von Flächen und Freie Wahrscheinlichkeitstheorie

Roland Speicher

Zufallsmatrizen haben nach einer Blüte in der Physik begonnen, auch in der Mathematik der letzten 15 Jahre eine zentrale Rolle zu spielen. Wir wollen in diesem Artikel einen wesentlichen Grund für die Bedeutung der Zufallsmatrizen aufzeigen: die Wechselwirkung zwischen analytischen und kombinatorischen/geometrischen Aspekten. Grob gesagt beschreiben Zufallsmatrizen Integrationen über (hochdimensionale) Räume. Wir beschränken uns hier auf die einfachsten Klassen von Zufallsmatrizen, dann sind diese Integrale Gaußsche Integrale bezüglich der Normalverteilung. Erstaunlicherweise haben solche Integrale einerseits eine kombinatorische Bedeutung (d. h. sie zählen interessante kombinatorische oder geometrische Objekte), sind aber andererseits auch analytischen Manipulationen zugänglich. Letztere führen dann oft zu neuen, nicht offensichtlichen Beziehungen zwischen den kombinatorischen Größen. Die Intention dieses Artikels ist es, diese Wechselwirkung von kombinatorischen und analytischen Gesichtspunkten an einfachen, aber typischen Beispielen zu veranschaulichen, und dabei auch den Bogen zur Freien Wahrscheinlichkeitstheorie zu schlagen.

Die Freie Wahrscheinlichkeitstheorie wurde von Dan Voiculescu im Zusammenhang mit grundlegenden Fragen über Operatoralgebren vor etwa dreißig Jahren eingeführt und hat sich seitdem zu einem sehr aktiven und erfolgreichen Gebiet der Mathematik mit erstaunlichen Querverbindungen weiterentwickelt. Eine Facette dabei ist die Verbindung zu Zufallsmatrizen, welche wir im letzten Teil des Artikels ansprechen werden.

Wir möchten betonen, dass unsere Betrachtungen sich nur sehr selektiv auf wenige illustrative Aspekte der Zufallsmatrizen und Freien Wahrscheinlichkeitstheorie konzentrieren. Diese Gebiete sind viel reichhaltiger, siehe z. B. [1, 2, 7, 10, 11, 13, 14]. Wir hoffen, dass unser Appetizer eine Anregung gibt, sich näher mit diesen modernen Gebieten der Mathematik auseinanderzusetzen.

## I Eindimensionale Gaußsche Integrale

Als Einstieg betrachten wir eindimensionale Gaußsche Integrale. Wir bezeichnen mit  $\gamma$  das Gaußmaß (oder die Normalverteilung) auf  $\mathbb{R}$ , also

$$d\gamma(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

und betrachten Momente bezüglich dieser Verteilung, also Integrale der Form  $\int_{\mathbb{R}} x^n d\gamma(x)$ . Wir haben unsere

Normierung und Skalierung so gewählt, dass  $\gamma$  ein Wahrscheinlichkeitsmaß mit Varianz 1 ist:

$$\int d\gamma(x) = 1, \quad \int x^2 d\gamma(x) = 1.$$

Weiterhin verschwinden wegen der Symmetrie von  $e^{-x^2/2}$  in  $x$  alle ungeraden Momente. Als nicht-triviale Größen bleiben also die geraden Momente. Wie man relativ einfach nachrechnen kann, sind diese Integrale durch natürliche Zahlen gegeben, nämlich:

$$\int x^{2n} d\gamma(x) = (2n - 1)!!, \quad (I)$$

wobei wir folgende Notation benutzen:

$$(2n - 1)!! := (2n - 1) \cdot (2n - 3) \cdots 5 \cdot 3 \cdot 1.$$

Aus der analytischen Perspektive gibt es keinen offensichtlichen Grund, warum wir als Ergebnisse hier immer natürliche Zahlen erhalten. Diese Zahlen  $(2n - 1)!!$  haben nun aber eine sehr konkrete Bedeutung, sie zählen nämlich folgende kombinatorische Objekte:

$$(2n - 1)!! = \text{Anzahl der Paarungen von } 2n \text{ Elementen.}$$

Dabei ist eine Paarung von  $2n$  Elementen eine Zerlegung der Menge  $\{1, 2, \dots, 2n\}$  in  $n$  disjunkte Teilmengen mit jeweils zwei Elementen.

Unsere Behauptung, dass die Gaußschen Integrale Paarungen zählen, erscheint auf den ersten Blick nicht sehr plausibel, sie ist aber nichts anderes als die Tatsache, dass die Gaußverteilung der Grenzwert im zentralen Grenzwertsatz ist; insbesondere erhalten wir die Gaußverteilung auch als Grenzwert der mit  $\sqrt{k}$  skalierten  $k$ -fachen Faltung der Gaußverteilung mit sich selbst, d.h. unsere Momente können wir auch berechnen als

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^k} \left( \frac{x_1 + \dots + x_k}{\sqrt{k}} \right)^{2n} d\gamma(x_1) \dots d\gamma(x_k).$$

Man überzeugt sich dann leicht, dass nach Ausmultiplizieren der Potenz in diesem Grenzwert nur die Terme überleben (und jeweils den gleichen Beitrag liefern), bei denen jeweils zwei der  $2n$  Faktoren übereinstimmen, also genau die Terme, die Paarungen entsprechen.

Grafisch wollen wir eine solche Paarung von  $2n$  Elementen durch ein  $2n$ -Eck veranschaulichen, bei dem wir die Seiten paarweise gemäß unserer Paarung identifizieren. In Abbildung 1 sind die drei Paarungen eines Vierecks gezeigt, die der Rechnung  $\int x^4 d\gamma(x) = 3 \cdot 1 = 3$  entsprechen.

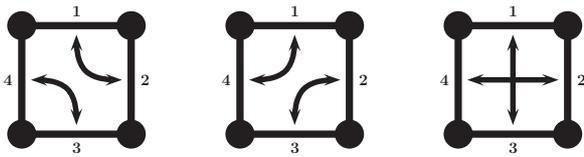


Abbildung 1. Die drei Paarungen  $\{(1, 2), (3, 4)\}$ ,  $\{(1, 4), (2, 3)\}$ ,  $\{(1, 3), (2, 4)\}$  der Menge  $\{1, 2, 3, 4\}$

## 2 Geometrischer Einschub: Das Geschlecht von Paarungen

Wenn man die in Abbildung 1 gezeigten drei Paarungen ernst nimmt und versucht, die der Paarung entsprechenden Identifikationen konkret auszuführen, so kann man dies für die erste und zweite Paarung ohne Probleme tun; die dritte Identifikation ist jedoch nicht ohne Kreuzung der verbleibenden Kanten in der Ebene machbar. Will man ein kreuzungsfreies Bild, so kann man dies allerdings erreichen, indem man das Ganze auf der Oberfläche eines Torus ausführt, wie in Abbildung 2 gezeigt. Die beiden ersten Paarungen sind von Geschlecht  $g = 0$ , ein so gepaartes 4-Eck kann auf der Oberfläche einer Kugel gemalt werden; die dritte Paarung hingegen ist von Geschlecht  $g = 1$ . Allgemein kann man einer Paarung eines  $2n$ -Ecks ein Geschlecht  $g$  zuordnen als das kleinste Geschlecht (also die Anzahl der Henkel) einer Oberfläche, auf der man die nach der Paarung verbleibenden  $n$  Kanten kreuzungsfrei malen kann. Man sollte sich klar machen, dass man dieses Geschlecht gemäß der Euler-Formel sehr einfach kombinatorisch aus Daten des nach der Identifizierung erhaltenen Graphen beschreiben kann. Die Euler-Formel sagt nämlich, dass  $2 - 2g = V - E + F$ ; dabei sind  $V$ ,  $E$ , und  $F$  die Anzahlen der Vertizes, der Kanten und der Flächen nach der Identifizierung. Da wir durch eine Paarung die  $2n$  Kanten jeweils paarweise identifizieren, gilt immer  $E = n$ , und für die Flächen bleibt immer nur die äußere Komponente unseres Polygons übrig, also  $F = 1$ . Nur die Anzahl der Vertizes  $V$  ist je nach Paarung verschieden und nicht ohne Weiteres aus der Paarung ablesbar. Auf jeden Fall haben wir die Beziehung

$$V = n + 1 - 2g. \quad (2)$$

Mit  $\varepsilon_g(n)$  bezeichnen wir die Anzahl der Paarungen eines  $2n$ -Ecks vom Geschlecht  $g$ . Diese Zahlen stellen interessante geometrische Information dar und spielen zum Beispiel eine wichtige Rolle in der Arbeit [5] von Harer und Zagier zur Bestimmung der Euler-Charakteristik des Modulraumes von Kurven. Im Folgenden wollen wir sehen, wie Zufallsmatrizen erlauben, nicht-triviale Information über diese Größen zu erhalten.

Während der Fall von Geschlecht Null relativ einfach zu verstehen ist – als Übungsaufgabe versuche man zu zeigen, dass  $\varepsilon_0(n)$  gleich der  $n$ -ten Catalan-Zahl  $C_n := \frac{1}{n+1} \binom{2n}{n}$  ist – bereiten höhere Geschlechter ungleich

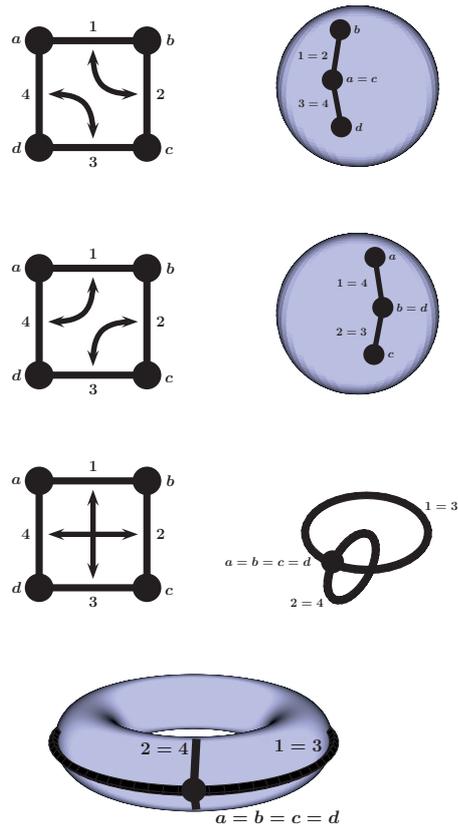


Abbildung 2. Graphen der drei Paarungen von vier Elementen nach Identifikation der Kanten: Die beiden ersten haben Geschlecht  $g = 0$  und können ohne Kreuzungen auf einer Kugeloberfläche gemalt werden; die dritte Paarung kann ohne Kreuzung nur auf einem Torus gemalt werden und hat Geschlecht  $g = 1$ . Man beachte auch, dass die beiden ersten Graphen drei Vertizes haben, während der dritte nur einen Vertex besitzt, im Einklang mit Formel (2).

größere Probleme und es gibt keine offensichtliche Formel für allgemeines  $\varepsilon_g(n)$ .

**Beispiel 1.** Für  $g = 0$  erhalten wir also die Catalan-Zahlen

$$\varepsilon_0(1) = 1, \quad \varepsilon_0(2) = 2, \quad \varepsilon_0(3) = 5, \quad \varepsilon_0(4) = 14, \quad \dots$$

Für höhere Geschlechter kann man sich zumindest für kleine  $n$  durch direkte Inspektion aller Paarungen von  $2n$  Elementen folgende Anzahlen verschaffen:

$$\varepsilon_1(2) = 1, \quad \varepsilon_1(3) = 10, \quad \varepsilon_1(4) = 70, \quad \varepsilon_2(4) = 21$$

Wie wir gesehen haben, „zählen“ unsere eindimensionalen Gaußschen Integrale Paarungen, allerdings können sie nicht zwischen verschiedenen Geschlechtern unterscheiden; wir erhalten hier die eher triviale Aussage

$$\sum_{g \geq 0} \varepsilon_g(n) = \int x^{2n} d\gamma(x).$$

Gehen wir allerdings zu höheren Dimensionen über, so erhalten wir mehr Spielraum, der uns letztendlich erlauben wird, zwischen verschiedenen  $g$  zu unterscheiden.

### 3 Mehrdimensionale Gaußsche Integrale

Wir betrachten nun die Verallgemeinerung unserer Gaußmaße von  $\mathbb{R}$  auf  $\mathbb{R}^k$  für eine beliebige natürliche Zahl  $k$ . Wiederum zentrieren wir unsere Verteilung (d. h., die ersten Momente verschwinden), allerdings wollen wir die zweiten Momente offenlassen, deren Information wird in einer quadratischen Form gespeichert. Sei also  $B = (b_{ij})_{i,j=1}^k$  eine strikt positive  $k \times k$ -Matrix, dann betrachten wir auf  $\mathbb{R}^k$  das Maß mit Dichte

$$d\mu(x) = \frac{1}{(2\pi)^{k/2}} \sqrt{\det B} \cdot e^{-\frac{1}{2}\langle Bx, x \rangle} dx_1 \cdots dx_k.$$

Dabei ist  $x = (x_1, \dots, x_k) \in \mathbb{R}^k$  und  $\langle Bx, x \rangle = \sum_{i,j=1}^k b_{ij}x_i x_j$ .

Die Normierung ist so gewählt, dass  $\mu$  ein Wahrscheinlichkeitsmaß ist; die ersten Momente verschwinden und die zweiten Momente sind gegeben durch die Kovarianzmatrix  $(c_{ij})_{i,j=1}^k = C := B^{-1}$  gemäß

$$\int_{\mathbb{R}^k} x_i x_j d\mu(x) = c_{ij}.$$

Wiederum sind aus Symmetriegründen alle ungeraden Momente gleich Null, und für die geraden Momente gilt die sogenannte *Wick-Formel*, die beliebige Momente mit Hilfe der zweiten Momente ausdrückt:

$$\int_{\mathbb{R}^k} x_{i(1)} \cdots x_{i(2n)} d\mu(x) = \sum_{\pi \in \mathcal{P}_2(2n)} \prod_{(s,t) \in \pi} c_{i(s)i(t)}. \quad (3)$$

Hier bezeichnen wir mit  $\mathcal{P}_2(2n)$  die Menge aller Paarungen von  $2n$  Elementen. In dem Produkt multiplizieren wir  $n$  Faktoren, die den  $n$  verschiedenen Paaren der Paarung entsprechen. Diese Formel ist eine kanonische Verallgemeinerung des eindimensionalen Falles. Um ein  $2n$ -tes Moment zu berechnen, müssen wir wiederum alle Paarungen von  $2n$  Elementen betrachten, allerdings gibt nun nicht jede Paarung den gleichen Beitrag (wir zählen also nicht nur alle Paarungen), sondern trägt mit einem Faktor bei, welcher ein Produkt der  $n$  zweiten Momente ist, die den gepaarten Variablen entsprechen. Das folgende Beispiel für den Fall  $n = 2$  mit unseren drei wohlbekanntesten Paarungen sollte dies klar machen.

$$\begin{aligned} \int x_{i(1)} x_{i(2)} x_{i(3)} x_{i(4)} d\mu(x) &= c_{i(1)i(2)} c_{i(3)i(4)} \\ &+ c_{i(1)i(4)} c_{i(2)i(3)} + c_{i(1)i(3)} c_{i(2)i(4)}. \end{aligned}$$

Die Formel (3) ist nach dem Physiker Gian-Carlo Wick benannt, der sie im Jahre 1950 in der Physik einführte; dort ist sie ein wesentliches Handwerkszeug in der Quantenfeldtheorie und zur Berechnung von Feynman-Integralen. Allerdings findet sich diese Formel zur Berechnung von Momenten von Gauß-Integralen auch schon in Arbeiten des Stochastikers Leon Isserlis aus dem Jahre 1918. Wiederum ist diese Isserlis-Wick-Formel

recht einfach zu beweisen, wenn man beachtet, dass unsere Gaußverteilung die Grenzwertverteilung in einem (multivariaten) zentralen Grenzwertsatz ist.

Wählen wir unsere Kovarianzmatrix  $C$  als Diagonalmatrix, so treten in dieser Wick-Formel nur Paarungen auf, die gleiche  $i$ -Indizes miteinander paaren. Durch geeignete Wahlen der Indizes  $i(1), \dots, i(n)$  können wir nun also verschiedene Paarungen voneinander unterscheiden. Allerdings ist nicht offensichtlich, ob wir so einen Zugriff auf das Geschlecht unserer Paarungen bekommen. Wenn wir die Vektoren im  $\mathbb{R}^k$  (für  $k = N^2$ ) nun aber zu Matrizen im  $\mathbb{R}^{N \times N}$  umsordieren, so ermöglicht die Matrizenmultiplikation gemeinsam mit der Wick-Formel genau eine solche Differenzierung.

### 4 Gaußsche Zufallsmatrizen

Daher wollen wir nun statt über Vektoren über Matrizen integrieren, insbesondere über den Raum

$$\mathcal{H}_N := \{\text{selbstadjungierte komplexe } N \times N\text{-Matrizen}\}.$$

Den Raum dieser Matrizen können wir kanonischerweise als Vektorraum mit dem  $\mathbb{R}^{N^2}$  identifizieren, gemäß

$$A = (a_{ij})_{i,j=1}^N \cong (a_{ii} (i = 1, \dots, N), \Re a_{ij} (i < j), \Im a_{ij} (i < j)),$$

wenn man beachtet, dass die Bedingung  $A = A^*$  bedeutet:  $\bar{a}_{ij} = a_{ji}$  ( $i \geq j$ ). Insbesondere sind also alle  $a_{ij}$  reell. Vermöge dieser Identifizierung können wir also ein Gaußsches Maß auf dem  $\mathbb{R}^{N^2}$  auf den Raum der selbstadjungierten  $N \times N$ -Matrizen transportieren. Unsere Konzession an die Tatsache, dass wir den  $\mathbb{R}^{N^2}$  nicht einfach als Vektorraum ansehen, sondern als Raum von  $N \times N$ -Matrizen, liegt nun darin, dass wir eine quadratische Form  $B$  zur Definition unseres Gaußmaßes benutzen, welche die Matrizenstruktur berücksichtigt, nämlich

$$\langle BA_1, A_2 \rangle := \text{Tr}(A_1 \cdot A_2), \quad (4)$$

wobei  $\text{Tr}$  die übliche Spur von Matrizen bezeichnet. Somit betrachten wir auf  $\mathcal{H}_N \cong \mathbb{R}^{N^2}$  ein Gaußmaß mit Dichte

$$d\mu(A) = \frac{1}{(2\pi)^{N^2/2}} 2^{(N^2-N)/2} e^{-\frac{1}{2}\text{Tr}(A^2)} dA. \quad (5)$$

Der Raum  $\mathcal{H}_N$  versehen mit diesem Wahrscheinlichkeitsmaß heißt *Ensemble von Gaußschen Zufallsmatrizen*.

Wenn wir nun Momente bezüglich dieses Maßes berechnen, so sind wir wiederum nur an solchen Momenten interessiert, die die Matrizenstruktur berücksichtigen, d. h. wir wollen nicht über beliebige Funktionen der Einträge  $a_{ij}$  unserer Matrix  $A$  integrieren (ansonsten würden wir ja einfach Gaußsche Integrale im  $\mathbb{R}^{N^2}$  betrachten), sondern nur über spezielle Funktionen der Form  $\text{Tr}(A^m)$ .

Aus Symmetriegründen verschwinden wieder Integrale über Spuren von ungeraden Potenzen von  $A$ . Somit bleiben als die wesentlichen nicht-trivialen Größen Integrale der Form  $\int \text{Tr}(A^{2n}) d\mu(A)$ . Durch Auswerten der Spur und der Matrixmultiplikation erhält man

$$\sum_{i(1), \dots, i(2n)=1}^N \int_{\mathcal{H}_N} a_{i(1)i(2)} \cdots a_{i(2n)i(1)} d\mu(A). \quad (6)$$

Das Integral über die Produkte der Einträge ist nach der Wick-Formel durch eine Summe über Paarungen gegeben; dabei zählen wir, gemäß unserer Wahl (4) der quadratischen Form  $B$ , nur die Paarungen, die  $a_{ij}$  mit  $a_{ji}$  identifizieren. Indem wir die Summe über die Indizes mit der Summe über die Paarungen vertauschen, erhalten wir in (6) für jede Paarung einen Beitrag der Form  $N^V$ , wobei  $V$  die Anzahl der verschiedenen Vertizes in unserem  $2n$ -Eck nach der Identifizierung von Seiten gemäß der Paarung bezeichnet. Diese können wir nach der Euler-Formel (2) mit dem Geschlecht der zugehörigen Paarung in Zusammenhang setzen und erhalten folgende fundamentale Aussage:

**Satz 2 (Entwicklung nach dem Geschlecht).** Für unsere Gaußschen Zufallsmatrizen gilt für alle natürlichen  $N$  und  $n$

$$\int_{\mathcal{H}_N} \text{Tr}(A^{2n}) d\mu(A) = \sum_{g \geq 0} \varepsilon_g(n) N^{n+1-2g}. \quad (7)$$

In der mathematischen Literatur tauchte diese Entwicklung zum ersten Mal wohl in der Arbeit von Harer und Zagier [5] aus dem Jahre 1986 auf; in der physikalischen Literatur waren solche Entwicklungen seit langem eher Folklore und lagen insbesondere Feynman-Diagramm-Rechnungen in der Quantenfeldtheorie zugrunde. Als wegweisend zur Etablierung solcher topologischen Entwicklungen in der Physik kann hier insbesondere die Arbeit von t'Hooft [6] aus dem Jahre 1974 genannt werden sowie [3]. In der mathematischen Gemeinschaft waren diese physikalischen Ideen in den 80er Jahren größtenteils unbekannt und wurden seitdem erst langsam wiederentdeckt (und auf eine rigorose Grundlage gestellt). Der Artikel [15] von Zvonkin (siehe auch [9]) war ein wichtiger Zwischenschritt, um den Mathematikern die physikalischen Ideen nahezubringen und lieferte eine wesentliche Anregung für den vorliegenden Beitrag.

**Beispiel 3.** Unter Benutzung der Zahlen  $\varepsilon_g(n)$  aus Beispiel 1 haben wir somit:

$$\begin{aligned} \int_{\mathcal{H}_N} \text{Tr}(A^2) &= N^2 \\ \int_{\mathcal{H}_N} \text{Tr}(A^4) &= 2N^3 + N \\ \int_{\mathcal{H}_N} \text{Tr}(A^6) &= 5N^4 + 10N^2 \\ \int_{\mathcal{H}_N} \text{Tr}(A^8) &= 14N^5 + 70N^3 + 21N \end{aligned}$$

Wir haben also gesehen: Unsere Gaußintegrale über die hermiteschen Zufallsmatrizen zählen Paarungen, gewichtet mit ihrem Geschlecht. Dies ist zwar recht erstaunlich, allerdings ist vermutlich nicht offensichtlich, ob diese Verknüpfung uns erlaubt, neue Erkenntnisse über die Zahlen  $\varepsilon_g(n)$  zu gewinnen.

## 5 Basiswechsel: Gaußsche Integrale in Eigenwert-Koordinaten

Bis jetzt haben wir unsere Gaußintegrale als Integrale bezüglich den Einträgen der Matrizen gesehen (und die Berechnung in diesen Koordinaten hat uns letztlich die Entwicklung nach dem Geschlecht gegeben). Für Matrizen drängt sich aber eine andere Darstellung geradezu auf: die nach Eigenwerten und Eigenvektoren. Der folgende klassische Satz sagt uns, wie wir unsere Integrale in diesen Koordinaten auswerten können.

**Satz 4 (Weyls Integrationsformel, 1924).** Sei  $F$  eine unitär invariante Funktion auf  $\mathcal{H}_N$ , also  $F(UAU^*) = F(A)$  für alle  $A \in \mathcal{H}_N$  und alle unitären  $N \times N$ -Matrizen  $U$ . (Somit hängt  $F$  also nur von den Eigenwerten  $\lambda_1, \dots, \lambda_N$  von  $A$  ab, und wir können es in der Form  $F(A) = F(\lambda_1, \dots, \lambda_N)$  schreiben, wobei  $F$  symmetrisch in seinen Argumenten ist.) Dann können wir unser Gaußintegral über die Funktion  $F$  auch in den Eigenwerten folgendermaßen durchführen:

$$\int_{\mathcal{H}_N} F(A) d\mu(A) = c_N \int_{\mathbb{R}^N} F(\lambda_1, \dots, \lambda_N) \cdot \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j)^2 d\gamma(\lambda_1) \cdots d\gamma(\lambda_N), \quad (8)$$

wobei die Konstante  $c_N$  nur von  $N$  abhängt.

Dies ist nichts anderes als die Anwendung der Transformationsformel für mehrdimensionale Integrale, wobei man von den Einträgen der Matrix auf Eigenwerte und Eigenvektoren transformiert. Da unsere Funktion nur von den Eigenwerten abhängt, kann man die Integration über die Eigenvektoren trivialerweise ausführen, dies ergibt nur einen konstanten Beitrag, der in der Konstante  $c_N$  steckt. (Die technischen Einzelheiten hier sind etwas verwickelter, da die Transformation von Einträgen auf Eigenwerte/Eigenvektoren keine Bijektion ist.) Die Jakobi-Determinante aus der Transformationsformel ist in unserem Fall explizit durch die in (8) auftauchende Vandermonde-Determinante  $\prod (\lambda_i - \lambda_j)^2$  gegeben. Im Gegensatz zu den Einträgen sind die Eigenwerte also nicht unabhängig voneinander, sondern durch diesen Vandermonde-Term gekoppelt. Da dieser klein wird, wenn zwei der Eigenwerte geringen Abstand haben, ist die Wahrscheinlichkeit einer solchen Konfiguration also klein; diese Tatsache wird normalerweise als *Abstoßung der Eigenwerte von Zufallsmatrizen* angesprochen.

## 6 Verknüpfung zwischen Geometrie und Analysis: Die Formel von Harer und Zagier

Die alternative Beschreibung unseres Gaußschen Integrals gemäß Satz 4 erlaubt nun, dieses Integral nicht nur in Abhängigkeit von  $N$ , sondern auch von  $n$  zu verstehen. Nach der Entwicklung (7) ist das Integral  $\int_{\mathcal{H}_N} \text{Tr}(A^{2n}) d\mu(A)$  ja ein Polynom in  $N$  (vom Grad  $n+1$ ). Unser neuer Blickwinkel liefert nun die überraschende Erkenntnis, dass wir im Wesentlichen auch polynomiales Verhalten in  $n$  haben.

**Lemma 5.** *Der Ausdruck*

$$\frac{1}{(2n-1)!!} \int_{\mathcal{H}_N} \text{Tr}(A^{2n}) d\mu(A) \quad (9)$$

ist ein Polynom vom Grad  $N-1$  in  $n$ .

Nur für  $N=1$  ist diese Aussage offensichtlich, da wir dann nach (1) für das eindimensionale Gaußintegral das Ergebnis  $(2n-1)!!$  erhalten, der Ausdruck (9) also konstant gleich 1 und somit ein Polynom vom Grad 0 in  $n$  ist.

**Beweis.** Wegen  $\text{Tr}(A^{2n}) = \lambda_1^{2n} + \dots + \lambda_N^{2n}$  folgt aus der Weylschen Formel (8)

$$\begin{aligned} & \int_{\mathcal{H}_N} \text{Tr}(A^{2n}) d\mu(A) \\ &= c_N \int_{\mathbb{R}^N} (\lambda_1^{2n} + \dots + \lambda_N^{2n}) \prod_{i < j} (\lambda_i - \lambda_j)^2 d\gamma(\lambda_1) \dots d\gamma(\lambda_N) \\ &= N \cdot c_N \int_{\mathbb{R}^N} \lambda_1^{2n} \prod_{i < j} (\lambda_i - \lambda_j)^2 d\gamma(\lambda_1) \dots d\gamma(\lambda_N) \\ &= N \cdot c_N \int_{\mathbb{R}} \lambda_1^{2n} p_N(\lambda_1) d\gamma(\lambda_1), \end{aligned}$$

wobei  $p_N(\lambda_1)$ , das Ergebnis der Integrationen über  $\lambda_2, \dots, \lambda_N$ , ein gerades Polynom in  $\lambda_1$  vom Grad  $2(N-1)$  ist, dessen Koeffizienten nur von  $N$  abhängen. Wegen

$$\int \lambda_1^{2n+2k} d\gamma(\lambda_1) = (2n+2k-1)!!$$

und der Tatsache, dass  $\frac{(2n+2k-1)!!}{(2n-1)!!}$  ein Polynom in  $n$  vom Grade  $k$  ist, folgt dann die Behauptung.  $\square$

Aus diesem neuen Verständnis der Integrale folgt dann relativ rasch folgende Aussage.

**Satz 6 (Harer-Zagier, 1986).** *Die Zahlen  $\varepsilon_g(n)$  genügen folgenden Rekursionen:*

$$(n+2)\varepsilon_g(n+1) = (4n+2)\varepsilon_g(n) + (4n^3 - n)\varepsilon_{g-1}(n-1).$$

Diese Rekursionen sind kombinatorisch nicht offensichtlich; sie wurden von Harer und Zagier in der Tat mittels der oben beschriebenen Zufallsmatrizenintegrale bewiesen (um daraus die Euler-Charakteristik der Modulräume

$\mathcal{M}_{g;1}$  abzuleiten); und zwar zu einer Zeit, als Zufallsmatrizen in der Physik zwar Folklore, in der Mathematik aber noch recht exotische Objekte waren. Es dauerte ziemlich lange bevor rein kombinatorische Beweise für die Formel von Harer und Zagier gefunden wurden, siehe [4, 8].

## 7 Verallgemeinerung auf mehrere Zufallsmatrizen und Freie Wahrscheinlichkeitstheorie

Eine neue Dimension eröffnet sich, wenn wir die vorhergehenden Betrachtungen von dem Fall einer Zufallsmatrix auf die gleichzeitige Betrachtung mehrerer Zufallsmatrizen verallgemeinern. Wiederum wollen wir hier nur den allereinfachsten Fall erwähnen, bei dem die verschiedenen Zufallsmatrizen voneinander unabhängig sind; d.h. wir betrachten nun das  $r$ -fache Produktmaß auf dem Raum  $\mathcal{H}_N^{\times r}$  von  $r$ -Tupeln selbstadjungierter  $N \times N$ -Matrizen. Dann kann man Integrale über Spuren der Produkte solcher Matrizen ebenso wie vorher mit Hilfe der Wick-Formel gemäß dem Geschlecht von Paarungen entwickeln. Der einzige Unterschied besteht nun darin, dass nur solche Paarungen beitragen, die jeweils die gleichen Matrizen untereinander paaren; wir erhalten somit

$$\begin{aligned} & \int_{\mathcal{H}_N \times \dots \times \mathcal{H}_N} \text{Tr}(A_{p(1)} \dots A_{p(2n)}) d\mu(A_1) \dots d\mu(A_r) \\ &= \sum_{g \geq 0} \varepsilon_g(p(1), \dots, p(2n)) N^{n+1-2g}, \quad (10) \end{aligned}$$

wobei wir mit  $\varepsilon_g(p(1), \dots, p(2n))$  die Anzahl der Paarungen  $\pi$  vom Geschlecht  $g$  bezeichnen, für die zusätzlich noch für jedes Paar  $(s, t) \in \pi$  gilt, dass  $p(s) = p(t)$ .

Während für  $r=1$  der Fall  $g=0$  recht trivial war und wir unser Hauptaugenmerk darauf gelegt haben, durch analytische Hilfsmittel die Kombinatorik für allgemeines  $g$  besser zu verstehen, stellt nun für beliebiges  $r$  selbst der Fall von Geschlecht  $g=0$  eine Herausforderung dar, und wir werden im Folgenden nur diesen betrachten.

Man sollte sich zunächst klarmachen, dass wir unsere analytische Herangehensweise aus dem letzten Kapitel zwar imitieren können, also unter Benutzung der Transformationsformel von Einträgen der Matrizen auf ihre Eigenwerte und Eigenvektoren umrechnen können. Allerdings bringt dies nun keinen Vorteil, da wegen der Nicht-Kommutativität unserer  $r \geq 2$  Matrizen die betrachtete Spur im Allgemeinen nicht nur eine Funktion der Eigenwerte ist, sondern auch die Eigenvektoren in einer nicht kontrollierbaren Art und Weise enthält.

Einen gewissen Ersatz liefert aber nun folgende Einsicht, mit Hilfe derer wir wiederum einen mehr analytischen Zugang zu unseren kombinatorischen Objekten bekommen.

**Satz 7.** Betrachte Operatoren  $l_i$  ( $i = 1, \dots, r$ ) auf einem Hilbertraum  $\mathcal{H}$  und einen ausgezeichneten Einheitsvektor  $\Omega \in \mathcal{H}$ , die für alle  $i, j = 1, \dots, r$  die folgenden Relationen erfüllen:

$$l_i^* l_j = \delta_{ij} \text{id}, \quad l_i^* \Omega = 0.$$

Dann gilt mit  $x_i := l_i + l_i^*$

$$\varepsilon_0(p(1), \dots, p(2n)) = \langle x_{p(1)} \cdots x_{p(2n)} \Omega, \Omega \rangle \quad (11)$$

Wir sollten hier anmerken, dass es solche Operatoren auch wirklich gibt (diese hängen mit der sogenannten *Cuntz-Algebra* zusammen) und diese auch sehr einfach (auf dem sogenannten *vollen Fockraum*) explizit konstruiert werden können.

**Beispiel 8.** Insbesondere gilt obiger Satz auch für den Fall  $r = 1$ , d. h. wir erhalten unsere Catalan-Zahlen also auch folgendermaßen:

$$C_n = \varepsilon_0(n) = \langle (l + l^*)^{2n} \Omega, \Omega \rangle, \quad \text{falls } l^* l = \text{id}, l^* \Omega = 0.$$

Der Operator  $l$  ist einer der wichtigsten Operatoren auf Hilberträumen, der sogenannte *einseitige Shift*. Es ist instruktiv, sich klarzumachen, dass wir auf der rechten Seite die Wörter in  $l$  und  $l^*$  zählen müssen, die gemäß den Relationen den Beitrag id liefern; z. B.

$$\begin{aligned} n = 1 &: l^* l & n = 2 &: l^* l^* l l, l^* l l^* l \\ n = 3 &: l^* l^* l^* l l l, l^* l^* l l l^* l, l^* l^* l l^* l l, l^* l l^* l^* l l, l^* l l^* l l^* l \end{aligned}$$

Wir haben oben die geometrischen/kombinatorischen Zahlen  $\varepsilon_0(p(1), \dots, p(2n))$  als Momente von Operatoren geschrieben. In Analogie zur klassischen Wahrscheinlichkeitstheorie sprechen wir diese nun als Momente einer *nicht-kommutativen Verteilung* an. Die *Freie Wahrscheinlichkeitstheorie* – die von Dan Voiculescu um 1983 initiiert und seitdem stetig und sehr erfolgreich vorangetrieben wurde – hat es sich zum Ziel gesetzt, Konzepte und Werkzeuge zur Beschreibung solcher nicht-kommutativen Verteilungen zu entwickeln.

**Beispiel 9.** Man beachte, dass wir im Fall einer Variablen die Momente mit einem Wahrscheinlichkeitsmaß identifizieren können. So sind die Catalan-Zahlen  $C_n$  die Momente der sogenannten *Halbkreisverteilung*, d. h. es gilt

$$\frac{1}{2\pi} \int_{-2}^{+2} t^{2n} \sqrt{4 - t^2} dt = C_n.$$

Somit ist die Verteilung des Realteils  $l + l^*$  des einseitigen Shifts durch diese Halbkreisverteilung gegeben. Auf der Seite der Zufallsmatrizen bedeutet dies dann, dass wir nach geeigneter Skalierung diese Verteilung als asymptotische Eigenwertverteilung unserer Gaußschen Zufallsmatrizen erhalten. Dieses *Wignersche Halbkreisgesetz* (welches von Eugene Wigner 1955 bewiesen wurde), ist in Abbildung 3 veranschaulicht.

Für mehrere nicht-kommutierende Variablen ist diese Gleichsetzung zwischen Momenten und Wahrscheinlichkeitsmaßen nicht mehr möglich, und eine der fundamentalen Fragen in der Freien Wahrscheinlichkeitstheorie lautet: Was soll eine sinnvolle analytische Entsprechung davon sein? Die Entwicklung von Theorien zur Beantwortung dieser Frage ist zur Zeit ein sehr aktives und spannendes Gebiet.

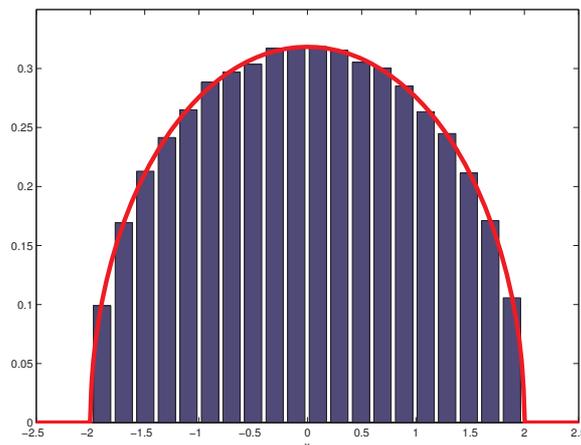


Abbildung 3. Vergleich der Halbkreisverteilung (rote Kurve) mit dem Histogramm (blau) der 4000 Eigenwerte einer gemäß der Verteilung (5) zufällig gewählten Gaußschen  $4000 \times 4000$ -Zufallsmatrix

## Literatur

- [1] G. Akemann, J. Baik, and P. Di Francesco (eds.): *The Oxford Handbook of Random Matrix Theory*, Oxford University Press, 2011.
- [2] G. Anderson, A. Guionnet, and O. Zeitouni: *An Introduction to Random Matrices*, Cambridge University Press, 2010.
- [3] E. Brezin, C. Itzykson, G. Parisi, and J.-B. Zuber: Planar diagrams. *Communication in Mathematical Physics* 59 (1978), 35–51.
- [4] I. Goulden and A. Nica: A direct bijection for the Harer-Zagier formula. *Journal of Combinatorial Theory, Series A* 111 (2005), 224–238.
- [5] J. Harer and D. Zagier: The Euler characteristic of the moduli space of curves. *Inventiones Mathematicae* 85 (1986), 457–485.
- [6] G. t’Hooft: A planar diagram theory for strong interactions. *Nuclear Physics B* 72 (1974), 461–473.
- [7] N. Katz and P. Sarnak (eds.): *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society, 1999.
- [8] B. Lass: Démonstration combinatoire de la formule de Harer-Zagier. *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics* 333 (2001), 155–160.
- [9] S. Lando and A. Zvonkin. *Graphs on surfaces and their applications*. Encyclopaedia of Mathematical Sciences, vol. 141, Springer, 2004.
- [10] U. Haagerup: Random matrices, free probability and the invariant subspace problem relative to a von Neumann algebra, *Proceedings of the International Congress of Mathematicians*, Vol. I (Beijing, 2002), Higher Ed. Press, Beijing, 2002, 273–290.
- [11] A. Nica and R. Speicher: *Lectures on the combinatorics of free probability*. Cambridge University Press, 2006.
- [12] R. Speicher: Free Probability Theory. In [1], 461–481.

- [13] R. Speicher: Free Probability and Random Matrices. To appear in *Proceedings of the ICM 2014, Seoul*.
- [14] D. Voiculescu: Free probability theory: random matrices and von Neumann algebras, *Proceedings of the International Congress of Mathematicians*, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, 227–241.
- [15] A. Zvonkin: Matrix integrals and map enumeration: An accessible introduction. *Mathematical and Computer Modelling* 26 (1997), 281–304.

Prof. Dr. Roland Speicher, Universität des Saarlandes, Fachrichtung Mathematik, Postfach 151150, 66041 Saarbrücken  
 speicher@math.uni-sb.de

Roland Speicher (geb. 1960) studierte Mathematik und Physik an den Universitäten Saarbrücken, Freiburg und Heidelberg. Diplom in Physik 1986, Promotion (1989) und Habilitation (1994) in Mathematik, alles in Heidelberg. 1995–2000 Heisenberg-Stipendiat. Von 2000 bis 2010 Professor an der Queen's University, Kingston, Kanada. Seit 2010 Professor in Saarbrücken. Seine Forschungsinteressen liegen auf dem Gebiet der Freien Wahrscheinlichkeitstheorie und Zufallsmatrizen. 2012 Jeffery-Williams-Prize der Kanadischen Mathematischen Gesellschaft CMS, 2014 Vortrag auf dem ICM in Seoul, 2014 ERC Advanced Grant *Non-commutative distributions in free probability*.

