

Kapitel 3

Hochdimensionale Integrale

Das Pfadintegral kann nur für sehr einfache Systeme wie den harmonischen Oszillator oder das freie Teilchen explizit berechnet werden. Für kompliziertere Systeme macht man Gebrauch von Störungstheorien (z.B. semiklassische Entwicklung, Störungstheorie in der Wechselwirkung, Hochtemperaturentwicklung) oder numerischen Methoden. Wir haben gesehen, dass die Pfadintegrale für thermodynamische Größen und Korrelationsfunktionen durch endlichdimensionale Integrale approximiert werden. Dabei wird die Zeit diskretisiert, $s \in \{0, \epsilon, \dots, n\epsilon = \tau\}$, und die Wirkung durch eine Riemannsche Summe genähert. Diese hängt von den Werten

$$\mathbf{q} = \{q_0, q_1, \dots, q_n\} = \{q(0), q(\epsilon), \dots, q(n\epsilon)\}$$

des Weges $q(s)$ an den Gitterpunkten $s_k = k\epsilon$ ab. In dieser Gitterapproximation ist jeder Erwartungswert durch ein endlich-dimensionales Integral gegeben,

$$\langle O \rangle = \frac{\int \mathcal{D}\mathbf{q} O(\mathbf{q}) e^{-S(\mathbf{q})}}{\int \mathcal{D}\mathbf{q} e^{-S(\mathbf{q})}}, \quad \text{mit} \quad \int \mathcal{D}\mathbf{q} = \int_{-\infty}^{\infty} \prod_1^n dq_j, \quad (3.1)$$

mit der in (2.53) eingeführten euklidischen Gitterwirkung $S(\mathbf{q}) = S(q_1, \dots, q_n)$ (statt S_E schreiben wir in diesem Abschnitt S).

3.1 Hochdimensionale Integrale

Nicht nur in Quantenstatistik, Festkörperphysik, euklidischen Quantenfeldtheorie, Hochenergiephysik oder anderen Teilgebieten der Physik und Chemie gilt es hochdimensionale Integrale möglichst effizient zu berechnen und dabei den Fluch der Dimension zu vermeiden. Zum Beispiel lässt sich der Erwartungswert von Zinsderivativen als hochdimensionales

Integral schreiben. Bei einer Laufzeit von 30 Jahren zu je 12 Monaten und Verwendung eines eigenen Zinssatzes für jeden Monat handelt es sich hier um 360-dimensionale Integrale. Integrale von noch viel höherer Dimension sind in Physik und Chemie nicht ungewöhnlich. Hier sind effiziente Algorithmen gefragt, die derartige Integrale bis auf einen abschätzbaren Fehler berechnen.

3.1.1 Numerische Algorithmen

Numerische Integrationsmethoden werden seit Jahrhunderten benutzt. Es gibt zwei bekannte Kategorien: Formeln, welche den Integrand an äquidistanten Stützstellen auswerten (Newton-Cotes Integrationsregeln) und Formeln, welche den Integranden an sorgfältig ausgewählten, aber nicht äquidistanten Stützstellen auswerten (Gaußsche Integrationsregeln). Für spezielle Integranden führt die zweite Klasse meistens zu besseren Resultaten.

Die numerischen Algorithmen beruhen auf der Riemannschen Definition von Integralen. Um nachzuprüfen, ob ein Funktion $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar ist, wähle man eine Einteilung des Intervalls,

$$\gamma : a = x_0 < x_1 < x_2 < \dots < x_{n-2} < x_{n-1} < x_n = b \quad (3.2)$$

und definiert die zu dieser Einteilung gehörende Riemannsche Unter- und Obersumme

$$U(f, \gamma) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \cdot \inf\{f(x) | x_i \leq x \leq x_{i+1}\}$$

$$O(f, \gamma) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \cdot \sup\{f(x) | x_i \leq x \leq x_{i+1}\},$$

mit $O(f, \gamma) \geq U(f, \gamma)$. Ist

$$\sup_{\gamma} U(f, \gamma) = \inf_{\gamma} O(f, \gamma),$$

dann heißt f im Riemannschen Sinne integrierbar und

$$\int_a^b f(x) dx \equiv \sup_{\gamma} U(f, \gamma) \quad (3.3)$$

das *Riemannsche Integral* von f . Diese Definition kann leicht auf mehrdimensionale Integrale ausgedehnt werden und wird bei numerischen Rechnungen gebraucht.

Die meisten Algorithmen beruhen darauf, dass jede glatte Funktion durch *Interpolationspolynome* approximiert werden kann. Wir erinnern daran, dass es genau ein Polynom P_m vom Grade $\leq m$ gibt, welches an $(m+1)$ vorgegebenen Stützstellen $x_0, x_1, \dots, x_{m-1}, x_m$

vorgegebene Werte f_0, \dots, f_m annimmt, wobei $f_i = f(x_i)$ ist. Zur expliziten Konstruktion definiert man die $m + 1$ *Lagrangeschen Polynome* vom Grade m :

$$L_p^{(m)}(x) = \prod_{\substack{i=0 \\ i \neq p}}^m \frac{x - x_i}{x_p - x_i}, \quad p = 0, \dots, m, \quad \text{mit} \quad L_p^{(m)}(x_q) = \delta_{pq}. \quad (3.4)$$

Das interpolierende Polynom vom Grade m ist dann

$$P_m(x) = \sum_{p=0}^m f(x_p) L_p^{(m)}(x). \quad (3.5)$$

Es gilt nun der folgende

Satz: *Es sei f eine auf dem Intervall Δ $(m + 1)$ -mal stetig differenzierbare Funktion, und sei P_m das zu den Stützstellen $x_0, \dots, x_m \in \Delta$ gehörige Interpolationspolynom vom Grade $\leq m$. Dann existiert zu jedem $x \in \Delta$ ein Punkt $\xi(x)$ (gelegen im kleinsten Intervall, welches die Punkte (x_0, \dots, x_m, x) enthält) derart, dass*

$$f(x) - P_m(x) = \frac{f^{(m+1)}(\xi(x))}{(m + 1)!} L^{(m)}(x), \quad L^{(m)}(x) = \prod_{i=0}^m (x - x_i). \quad (3.6)$$

Aufgrund des Satzes ergibt sich für das Integral $\int dx f(x)$ von der kleinsten bis zur größten Stützstelle die Formel

$$\int_{x_0}^{x_m} dx f(x) = \sum_{p=0}^m f(x_p) \underbrace{\int dx L_p^{(m)}(x)}_{\gamma_p^{(m)}} + \int dx \frac{f^{(m+1)}(\xi(x))}{(m + 1)!} L^{(m)}(x). \quad (3.7)$$

Die $\gamma_p^{(m)}$ heißen Gewichte und die x_p Knoten der Integrationsformel. Für äquidistante Knoten an den Stellen

$$x_0, x_1 = x_0 + \epsilon, x_2 = x_0 + 2\epsilon, \dots, x_m = x_0 + m\epsilon \quad (3.8)$$

erhalten wir mit der Substitution $x = x_0 + \epsilon t$, $t \in [0, m]$ die Gewichte

$$\gamma_p^{(m)} = \epsilon \int_0^m \prod_{\substack{i=0 \\ i \neq p}}^m \frac{t - i}{p - i} dt := \epsilon w_p^{(m)} = \epsilon w_{m-p}^{(m)}, \quad p = 0, 1, \dots, p. \quad (3.9)$$

Wenden wir das allgemeine Resultat (3.7) auf die konstante Funktion $f = 1$ an, so ergibt

sich die Summenformel $\sum_p \gamma_p^{(m)} = m\epsilon$ oder auch

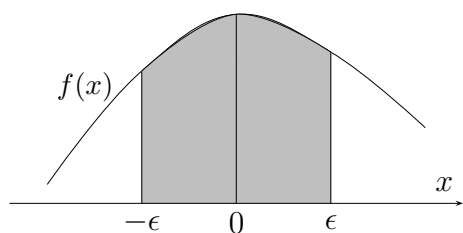
$$w_0^{(m)} + w_1^{(m)} + \dots + w_m^{(m)} = m. \tag{3.10}$$

Die Newton-Cotes-Formeln lauten nun

$$\int_{x_0}^{x_m} dx f(x) \sim \sum_{p=0}^m \epsilon w_p^{(m)} f(x_0 + \epsilon p), \quad x_m = x_0 + m\epsilon. \tag{3.11}$$

Man findet folgende Gewichte

m	Name	$w_p^{(m)} (p = 0, 1, \dots, m)$							
0	Rechteckregel	1							
1	Trapezregel	$\frac{1}{2}$	$\frac{1}{2}$						
2	Simpson-Regel	$\frac{1}{3}$	$\frac{4}{3}$	$\frac{1}{3}$					
3	3/8 – Regel	$\frac{3}{8}$	$\frac{9}{8}$	$\frac{9}{8}$	$\frac{3}{8}$				
4	Milne-Regel	$\frac{14}{45}$	$\frac{64}{45}$	$\frac{24}{45}$	$\frac{64}{45}$	$\frac{14}{45}$	(3.12)		
5		$\frac{95}{288}$	$\frac{375}{288}$	$\frac{250}{288}$	$\frac{250}{288}$	$\frac{375}{288}$			
6	Weddle-Regel	$\frac{41}{140}$	$\frac{216}{140}$	$\frac{27}{140}$	$\frac{272}{140}$	$\frac{27}{140}$		$\frac{216}{140}$	$\frac{41}{140}$



Wir illustrieren die Fehleranalyse für die Simpson-Regel. Dazu betrachten wir die Differenz zwischen dem Integral der Funktion $f(x)$ von $-\epsilon$ bis ϵ (siehe Abbildung) und der Näherung (3.11) für $m = 2$, also den Fehler

$$E_2(\epsilon) = \int_{-\epsilon}^{\epsilon} dx f(x) - \frac{\epsilon}{3} (f(-\epsilon) + 4f(0) + f(\epsilon)).$$

Wir leiten $E_2(\epsilon)$ dreimal nach ϵ ab und erhalten

$$E_2'''(\epsilon) = -\frac{\epsilon}{3} (-f'''(-\epsilon) + f'''(\epsilon)).$$

Dies kann betragsmäßig wie folgt abgeschätzt werden:

$$|E_2'''(\epsilon)| = \frac{\epsilon}{3} |f'''(\epsilon) - f'''(-\epsilon)| \leq \frac{2\epsilon}{3} M_3 \quad \text{mit} \quad M_3 = \sup_{t \in [-\epsilon, \epsilon]} |f'''(t)|.$$

Die Integration ergibt die Fehlerabschätzung

$$|E_2(\epsilon)| \leq M_3 \cdot \frac{\epsilon^4}{36}. \tag{3.13}$$

Falls die Funktion f mindestens viermal stetig differenzierbar ist, kann man auf E_2''' den Mittelwertsatz anwenden,

$$E_2'''(\epsilon) = \frac{2\epsilon}{3} \cdot f^{(4)}(\xi),$$

und es folgt die verbesserte Abschätzung

$$|E_2(\epsilon)| \leq M_4 \cdot \frac{\epsilon^5}{90} \quad \text{mit} \quad M_4 = \sup_{t \in [-\epsilon, \epsilon]} |f^{(4)}(t)|. \tag{3.14}$$

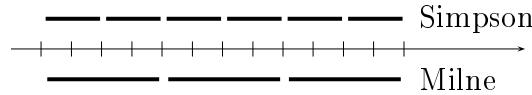
Hieraus ergibt sich die bemerkenswerte Tatsache, dass durch die *Keplersche Fassregel* sogar kubische Polynome exakt integriert werden. Für die anderen Verfahren erhält man analoge Fehlerschranken für das Integral von der kleinsten bis zur größten Stützstelle ($M_m = \sup_{[x_0, x_m]} |f^{(m)}|$):

m	Name	$E_m(\epsilon)$	m	Name	$E_m(\epsilon)$
0	Rechteckregel	$\frac{1}{2} \epsilon^2 M_1$	4	Milne-Regel	$\frac{8}{945} \epsilon^7 M_6$
1	Trapezregel	$\frac{1}{12} \epsilon^3 M_2$	5		$\frac{275}{12096} \epsilon^7 M_6$
2	Simpson-Regel	$\frac{1}{90} \epsilon^5 M_4$	6	Weddle-Regel	$\frac{9}{1400} \epsilon^9 M_8$
3	3/8 – Regel	$\frac{3}{80} \epsilon^5 M_4$			

(3.15)

Allgemein gilt, dass für gerade m sogar Polynome vom Grad $m+1$ exakt integriert werden. Für große m werden die Koeffizienten in den Newton-Cotes Formeln allerdings gross und haben wechselnde Vorzeichen. Dies führt zu Differenzen grosser Zahlen und auch deshalb werden die Newton-Cotes Verfahren höherer Ordnung in der Praxis kaum eingesetzt. Für nicht genügend oft differenzierbare Funktionen können die auf Interpolationspolynomen beruhenden Methoden völlig falsche Resultate liefern!

Zusammengesetzte Integrationsformeln: Indem das Integrationsintervall, über das die Funktion f integriert werden soll, in kleinere, gleich große Teilintervalle zerlegt wird, gelangt man zum Rechteck-, Trapez-, Simpson- oder den höheren Integrationsverfahren. Die Anzahl Intervalle sollte ein Vielfaches von m sein. Zum Beispiel wird beim Simpsonverfahren die Keplersche Fassregel auf Doppelintervalle angewandt.



Wir betrachten die zusammengesetzte Simpsonregel etwas näher. Das Integrationsintervall $[a, b] = [x_0, x_{2n}]$ enthalte $2n$ Teilintervalle der Länge ϵ , $b - a = 2n\epsilon$. Die $2n + 1$ Stützstellen sind $x_j = a + \epsilon j$, $j = 0, 1, \dots, 2n$. Das Integral wird approximiert durch

$$\begin{aligned}
 S_2(f) &\approx \frac{\epsilon}{3} \left(\{f(x_0) + 4f(x_1) + f(x_2)\} + \{f(x_2) + 4f(x_3) + f(x_4)\} + \dots \right. \\
 &\quad \left. \dots + \{f(x_{2n-2} + 4f(x_{2n-1}) + f(x_{2n})\} \right) \\
 &= \frac{\epsilon}{3} \left(f(x_0) + 4 \sum_{j=0}^{n-1} f(x_{2j+1}) + 2 \sum_{j=1}^{n-1} f(x_{2j}) + f(x_{2n}) \right).
 \end{aligned}$$

Der Fehler kann wie folgt abgeschätzt werden

$$\left| \int_a^b f(x) dx - S_2(f) \right| \leq \frac{1}{90} \epsilon^5 \cdot n \underbrace{\sup_{t \in [a,b]} f^{(4)}(t)}_{M_4} = \frac{b-a}{180} \epsilon^4 M_4. \tag{3.16}$$

Allgemeiner gilt für eine äquidistante Einteilung des Integrationsintervalls in $m \cdot n$ Teilintervalle, so dass $b - a = (mn)\epsilon$ ist, die Abschätzung

$$\left| \int_a^b f(x) dx - S_m(f) \right| \leq \frac{b-a}{m\epsilon} E_m(\epsilon) \tag{3.17}$$

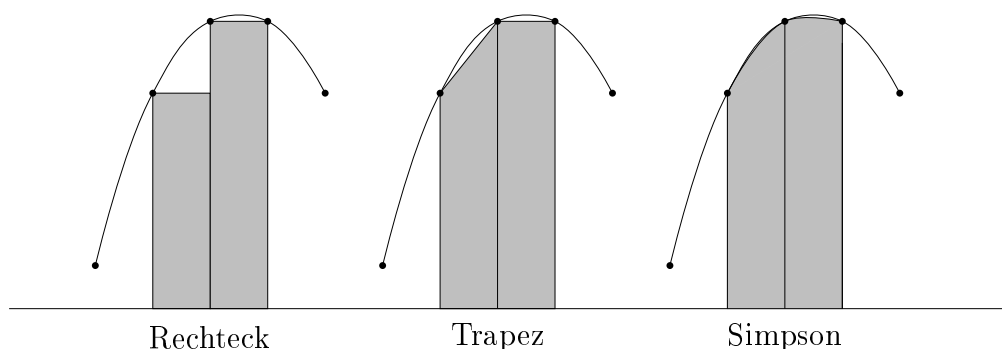
mit $E_m(\epsilon)$ aus (3.15). Dabei ist natürlich $M_m = \sup_{[a,b]} f^{(m)}$.

Mit Hilfe eines C-Programms berechnen wir das Integral einer Funktion über das Intervall $[a, b]$ und zwar auf vier Arten: mit dem Rechteck-, Trapez- und Simpsonverfahren sowie mit Hilfe der Monte-Carlo Methode. Das letzte Verfahren wird weiter unten besprochen. Nochmals zur Erinnerung:

$$\begin{aligned}
 \text{Rechteckregel :} & \quad \sum_{i=0,1,2}^{n-1} \epsilon f(x_i) \\
 \text{Trapezregel :} & \quad \sum_{i=0,1,2}^{n-1} \frac{\epsilon}{2} (f(x_i) + f(x_{i+1}))
 \end{aligned} \tag{3.18}$$

$$\text{Simpson-Methode : } \sum_{i=0,2,4}^{n-2} \frac{\epsilon}{3} (f(x_i) + 4f(x_{i+1}) + f(x_{i+2})).$$

In der letzten Formel soll n eine gerade Zahl sein. Die Näherungen sind in der folgenden Figur skizziert.



Das Programm `1dintegral.c` auf Seite 47 berechnet das Integral $\int_0^1 dx e^x$ für

$$\epsilon \in \{10^{-n} | n = 1, 2, \dots, 6\}.$$

Die Werte für die stückweise konstante, lineare oder quadratische Näherung sind in der folgenden Tabelle enthalten. Für das Simpsonverfahren konvergiert wie erwartet die Näherung sehr schnell gegen den exakten Wert 1.7182818.

$n, \log M$	einfach	Trapez	Simpson	MC
1	1.633799	1.719713	1.718283	1.853195
2	1.709705	1.718296	1.718282	1.793378
3	1.717423	1.718282	1.718282	1.720990
4	1.718196	1.718282	1.718282	1.711849
5	1.718273	1.718282	1.718282	1.719329
6	1.718281	1.718282	1.718282	1.718257

3.1.2 Monte-Carlo Integration

Die Monte-Carlo Methode stammt wahrscheinlich von STANISLAW ULAM. Er fand die Methode 1946, als er sich Gedanken über die Gewinnwahrscheinlichkeiten beim Solitaire-Spiel machte. In seinen Worten:

The first thoughts and attempts I made to practice [the Monte Carlo Method] were suggested by a question which occurred to me in 1946 as I was convalescing from an illness and playing solitaires. The question was what are the chances that a Canfield solitaire laid out with 52 cards will come out successfully? After spending a lot of time trying to estimate them by pure combinatorial calculations, I wondered whether a more practical method than „abstract thinking“ might not be to lay it out say one hundred times and simply observe and count the number of successful plays. . . .

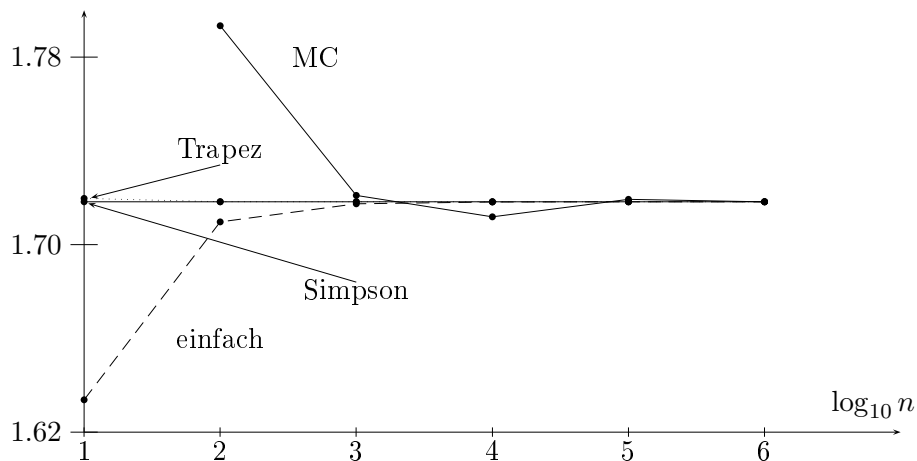
Einige Jahre später wurde die Methode auf das Neutronendiffusionsproblem angewandt, das mit anderen Methoden nicht lösbar schien [25]. Eine wichtige Anwendung ist die Berechnung hochdimensionaler Integrale. Ein sehr einfacher Algorithmus wäre:

- erzeuge M gleichverteilte Punkte $\{x_1, \dots, x_M\}$ im Integrationsgebiet \mathcal{G} ,
- berechne für jeden Punkt den Funktionswert $f(x_i)$, $i = 1, \dots, M$,
- berechne den Mittelwert

$$I(M) = \frac{\text{Vol}(\mathcal{G})}{M} \sum_{i=1}^M f(x_i). \tag{3.19}$$

Für eine Riemann-integrierbare Funktion konvergiert $I(M)$ für große M gegen das Integral $\int_{\mathcal{G}} f$. Die Werte in der letzten Spalte der obigen Tabelle enthalten $I(M = 10, 100, \dots)$ für das Integral der Exponentialfunktion.

Die folgende Abbildung illustriert das Konvergenzverhalten der drei Integrationsmethoden mit äquidistanten Stützstellen und der einfachen Monte-Carlo Integration. Für die Exponentialfunktion liefert die Methode von Simpson schon für zehn Intervalle das richtige Resultat e bis auf die 6. Stelle hinter dem Komma.



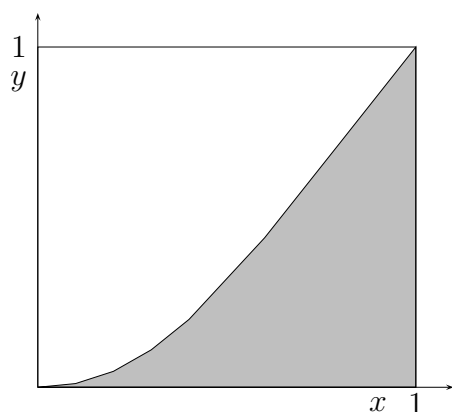
Unpraktikabel werden Standardverfahren wenn die Dimension n des Integrals

$$I = \int dq_1 \dots dq_n f(q_1, \dots, q_n) \equiv \int d^n q f(\mathbf{q}) \quad (3.20)$$

gross wird. Sind zum Beispiel die Integrationsgrenzen in jeder Dimension gleich 0 und 1, und wählt man in jeder Dimension den gleichen Abstand ϵ zwischen den Stützstellen, dann ist deren Anzahl ϵ^{-n} . Der Rechenaufwand ist proportional zur Anzahl Stützstellen. Nehmen wir als Beispiel $\epsilon = 0.1$, was sicherlich eine grobe Einteilung des Intervalls $[0, 1]$ ist, dann ist die Anzahl Stützstellen $\sim 10^n$. Die Auswertung einer Stützstelle auf einem PC dauert etwa 10^{-7} s und die Berechnung eines 12-fachen Integral etwa einen Tag.

Hit-or-miss Monte Carlo Methode und Binomialverteilung

Gesucht sei wieder der Wert des Integrals $I = \int dx f(x)$,



wobei wir ohne Beschränkung der Allgemeinheit annehmen dürfen, dass wir von 0 bis 1 integrieren. Mit einem Zufallszahlengenerator, der zwischen 0 und 1 gleichverteilte Zahlen liefert, werden zwei Zufallszahlen $0 \leq r_1, r_2 \leq 1$ erzeugt,

$$x = r_1 \quad , \quad y = r_2.$$

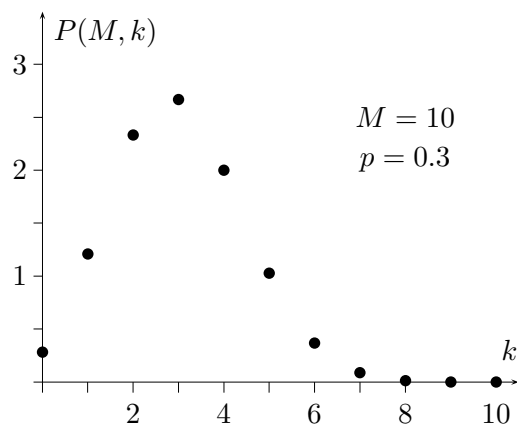
Wir haben getroffen, wenn $y \leq f(x)$ ist. Die Wahrscheinlichkeit für einen Treffer ist

$$p = \frac{\text{Anzahl Treffer}}{\text{Anzahl Versuche}} = \frac{\text{dunkle Fläche}}{\text{Gesamtfläche}} = \frac{I}{1} = I. \quad (3.21)$$

Bei M statistisch unabhängigen Versuchen können wir $k \in \{0, \dots, M\}$ Treffer landen und die Wahrscheinlichkeit dafür ist durch die *Binomialverteilung*

$$P(M, k) = \binom{M}{k} p^k (1-p)^{M-k} \quad \text{mit} \quad \sum_{k=0}^M P(M, k) = 1 \quad (3.22)$$

gegeben. Hier ist zum Beispiel $p^k (1-p)^{M-k}$ die Wahrscheinlichkeit dafür, dass die ersten k Versuche Treffer und die letzten $M - k$ Versuche Nieten ergeben. Der Binomialkoeffizient zählt die Anzahl Möglichkeiten, aus der Menge von M Versuchen k Treffer auszuwählen. Die Binomialverteilung beschreibt eine bei pM lokalisierte Glockenkurve und ist in der folgenden Figur für $M = 10$ und $p = 0.3$ geplotted.



Ihre *erzeugende Funktion* kann leicht berechnet werden,

$$\begin{aligned} Z(t) &= \langle e^{tk} \rangle = \sum_{k=0}^M e^{kt} P(M, k) \\ &= (e^t p + (1 - p))^M. \end{aligned} \quad (3.23)$$

Als Summe von Wahrscheinlichkeiten ist $Z(0) = 1$. Erwartungswerte von beliebigen Potenzen von k können durch ableiten der erzeugenden Funktion gewonnen werden.

Nicht unerwartet ist der *mittlere Anteil Treffer* gleich

$$\left\langle \frac{k}{M} \right\rangle = \frac{1}{M} \sum_{k=0}^M k P(M, k) = \frac{1}{M} \frac{dZ}{dt} \Big|_{t=0} = p. \quad (3.24)$$

Das Quadrat der Streuung um den Ursprung ist

$$\left\langle \frac{k^2}{M^2} \right\rangle = \frac{1}{M^2} \sum_{k=0}^M k^2 P(M, k) = \frac{1}{M^2} \frac{d^2 Z}{dt^2} \Big|_{t=0} = \frac{p}{M} + \left(1 - \frac{1}{M}\right) p^2 \quad (3.25)$$

und für das Quadrat der Streuung um die mittlere Anzahl Treffer findet man

$$\sigma^2 = \frac{1}{M^2} \langle (k - \langle k \rangle)^2 \rangle = \frac{1}{M^2} \frac{d^2 \log Z}{dt^2} \Big|_{t=0} = \frac{p(1-p)}{M}. \quad (3.26)$$

Die Streuung um den Mittelwert vermindert sich relativ langsam mit der Anzahl Versuche, $\sigma \sim M^{-1/2}$. Eine Schätzung von p ist h/M , wobei h die Anzahl Treffer bei M Versuchen ist. Die folgende Tabelle enthält die Schätzwerte p und σ für das Integral

$$I = \int_0^1 f(x), \quad f(x) = \frac{x^2 e^x}{1 - x + x e^x}. \quad (3.27)$$

für verschiedene Anzahl M von Versuchen. Die Streuung um den wahren Wert des Integrals, $I = 0.376370$, nimmt mit M ab. Die Werte in den ersten drei Spalten wurden mit dem Programm `hitmissflaeche.c` auf Seite 48 generiert.

Die grobe Hit-or-Miss Methode kann mit wenig Aufwand verbessert werden. Wenn nämlich p gegen 1 oder 0 strebt so wird σ sehr klein (allerdings wird für $p \rightarrow 0$ der relative Fehler gross). Wir nehmen nun eine Hilfsfunktion $g(x)$, die $f(x)$ möglichst gut

approximiert aber analytisch noch integriert werden kann. Ist das erste Integral in

$$I = \underbrace{\int (f(x) - g(x)) dx}_{p \text{ wird klein}} + \underbrace{\int g(x) dx}_{\text{bekannt}} \quad (3.28)$$

klein und der Integrand zwischen 0 und 1, dann können wir dieses Integral mit dem Hit-or-miss Verfahren mit verminderter Varianz berechnen. Für $f(x)$ in (3.27) wählen wir

$$g(x) = x^2 \quad \text{mit} \quad \int g(x) dx = 1/3.$$

Dann ergeben sich für das Integral die verbesserten Schätzwerte in der drittletzten Spalte und die Varianz in der letzten Spalte der Tabelle. Diese Werte wurden ebenfalls mit `hitmissflaeche.c` berechnet.

$\log_{10} M$	p	$I - p$	σ	p_{verb}	$I - p_{\text{verb}}$	σ_{verb}
1	0.500000	-0.123630	0.158114	0.333333	0.043037	0.000000
2	0.330000	0.046370	0.047021	0.363333	0.013037	0.017059
3	0.399000	-0.022630	0.015485	0.377333	-0.000963	0.006486
4	0.378900	-0.002530	0.004851	0.376833	-0.000463	0.002040
5	0.376570	-0.000200	0.001532	0.377693	-0.001323	0.000651
6	0.374857	0.001513	0.000484	0.376305	0.000065	0.000203
7	0.376273	0.000097	0.000153	0.376303	0.000067	0.000064

Summen von Zufallszahlen, Gaußverteilung und Grenzwertsatz

Das Programm `gaussdistr.c` auf Seite 49 erzeugt die Summe s von n auf dem Intervall $[0, 1]$ gleichverteilten unabhängigen Zufallszahlen x_1, \dots, x_n . Die erzeugende Funktion für die Summe ist

$$Z(t) = \langle e^{ts} \rangle = \int_{I^n} d^n x e^{t(x_1 + \dots + x_n)} = \left(\int_0^1 dx e^{tx} \right)^n = t^{-n} (e^t - 1)^n, \quad (3.29)$$

und für den Mittelwert von s finden wir

$$m = \langle s \rangle = \left. \frac{dZ}{dt} \right|_{t=0} = \int_{I^n} d^n x (x_1 + \dots + x_n) = \frac{n}{2}. \quad (3.30)$$

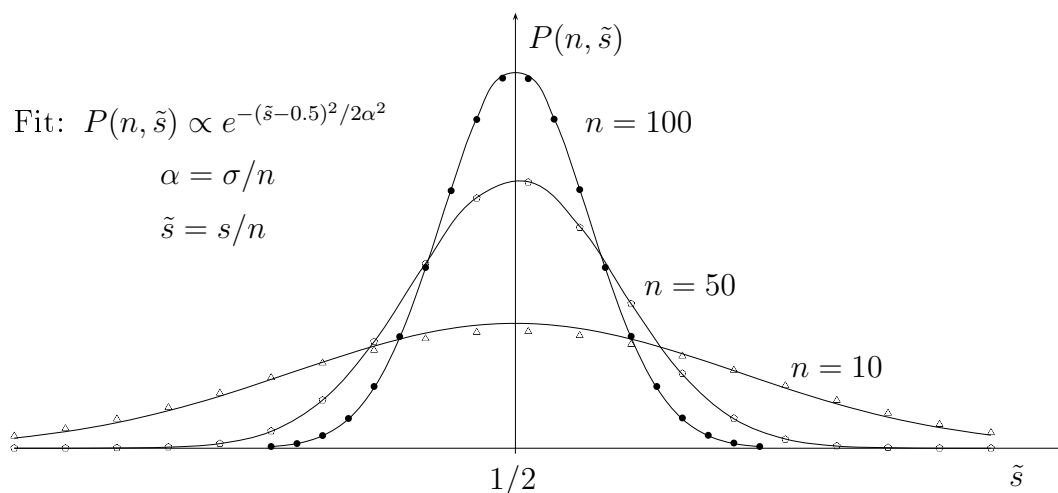
und für dessen Streuungsquadrat

$$\left. \frac{d^2 \log Z}{dt^2} \right|_{t=0} = \sigma^2 = \langle s^2 \rangle - m^2 = \frac{n}{12}. \quad (3.31)$$

Nach dem Gesetz der grossen Zahlen erwarten wir die Gaußverteilung

$$P_s = \frac{1}{\sqrt{2\pi} \sigma} e^{-(s-m)^2/2\sigma^2}. \quad (3.32)$$

Das Programm `gaussdistr.c` auf Seite 49 berechnet die Verteilung von s für die Summe von 10, 50 und 100 Zufallszahlen. Zur Bestimmung der Verteilung werden jeweils eine Million Versuche gemacht. Mit den zufälligen Werten für s wird ein Histogramm erstellt und im array `mean[100]` gespeichert. Wir haben die Zufallsvariable s mit n reskaliert, so dass das Maximum der Verteilung bei $1/2$ liegt. In der folgenden Abbildung werden die Resultate der MC-Simulation (Punkte, Dreiecke, Vierecke) mit den entsprechenden Gaußverteilungen verglichen.



Im Anhang C findet sich ein Beweis für das *Gesetz der grossen Zahlen*. Für gleichverteilte Zahlen in $[0, 1]$ ist der Mittelwert $1/2$ und die Varianz $1/12$. Die Ungleichung (3.56) für die Wahrscheinlichkeit dafür, dass s/n mehr als δ vom Mittelwert $1/2$ abweicht, nimmt folgende Form an

$$\Pr \left[\left| \frac{s}{n} - \frac{1}{2} \right| \geq \delta \right] \leq \frac{1}{12n\delta^2}. \quad (3.33)$$

3.2 Important Sampling

Numerische Integrationsverfahren nähern Integrale durch endliche Summen,

$$\int d^n q f(\mathbf{q}) \sim \sum_{\mu=1}^M f(\mathbf{q}_\mu) \Delta \mathbf{q}_\mu.$$

Für große n und schwach veränderliche Funktionen f kann es vorteilhaft sein, die Stützpunkte \mathbf{q}_μ zufällig zu wählen. In vielen Anwendungen variiert der Integrand allerdings um Größenordnungen für verschiedene Punkte und man vergeudet Rechenzeit wenn man Stützpunkte mit sehr kleinem Integranden auswählt. Beim *important sampling*, zum Beispiel dem Metropolis-Algorithmus, werden bevorzugt Punkte \mathbf{q}_μ mit großem Integranden berücksichtigt. Die Stützstellen liegen überwiegend da, wo der Integrand gross ist und dies verringert die Varianz der einzelnen Schätzung für das Integral.

Dazu nimmt man eine Funktion $g(\mathbf{q})$, deren Integral berechenbar ist und welche $f(\mathbf{q})$ möglichst gut annähert, und schreibt

$$\int_0^1 f(\mathbf{q}) d^n q = \int_0^1 \frac{f(\mathbf{q})}{g(\mathbf{q})} g(\mathbf{q}) d^n q. \quad (3.34)$$

Durch die Erzeugung von Zufallspunkten \mathbf{q}_μ die mit $g(\mathbf{q})d^n q$ verteilt sind, ergibt sich bei M „Messungen“ die Schätzung

$$\int f(\mathbf{q}) d^n q \approx \bar{f} = \frac{1}{M} \sum_{\mu=1}^M \frac{f(\mathbf{q}_\mu)}{g(\mathbf{q}_\mu)}, \quad (3.35)$$

und dabei variieren die Summanden jetzt nicht mehr so stark. Allerdings muß das Integral von g bekannt sein, um aus gleichverteilten Zufallszahlen solche zu erhalten, die mit g verteilt sind.

Bei der Berechnung von Erwartungswerten in Gitterfeldtheorien,

$$\langle O \rangle \approx \frac{1}{Z} \int \mathcal{D}\mathbf{q} O(\mathbf{q}) e^{-S(\mathbf{q})}, \quad \mathcal{D}\mathbf{q} = d^n q, \quad Z = \int e^{-S(\mathbf{q})} \mathcal{D}\mathbf{q}, \quad (3.36)$$

wäre es wünschenswert die Boltzmannverteilung

$$P(\mathbf{q}) = \frac{1}{Z} e^{-S(\mathbf{q})}, \quad (3.37)$$

als Vergleichsfunktion g zu wählen, weil dann nur noch über die im Vergleich zu P in den

meisten Fällen glatten Observablen $O(\mathbf{q})$ gemittelt werden müsste,

$$\langle O \rangle \approx \bar{O} = \frac{1}{M} \sum_{\mu=1}^M O(\mathbf{q}_\mu). \quad (3.38)$$

Hier ist M die Anzahl der erzeugten Punkte \mathbf{q}_μ . Damit wird die Monte Carlo-Schätzung \bar{O} für den Mittelwert von O zu einem arithmetischen Mittel. Passend verteilte $\{\mathbf{q}_1, \mathbf{q}_2, \dots\}$ sind aber nicht ohne Weiteres zu erzeugen.

Wir haben folgendes Problem: Die n -dimensionalen Integrale

$$\langle O \rangle = \int dq_1 \dots \int dq_n O(\mathbf{q}) P(\mathbf{q}), \quad \int \mathcal{D}\mathbf{q} P(\mathbf{q}) = 1, \quad (3.39)$$

sollen für verschiedene Funktionen (Observablen) $O(\mathbf{q})$, aber die gleiche Wahrscheinlichkeitsdichte $P(\mathbf{q})$ berechnet werden. Dazu sollen Algorithmen gefunden werden, die nach P verteilte Punkte generieren. Der folgende Metropolis-Algorithmus [25] (er wird später begründet werden) erzeugt $\{\mathbf{q}_\mu\}$, die gemäß $P(\mathbf{q})$ verteilt sind:

1. Beginne mit $\mu = 0$ und einem beliebigen Startpunkt \mathbf{q}_μ im Integrationsbereich.
2. Wähle einen zweiten zufälligen Punkt \mathbf{q}' und ein Zufallszahl $r \in [0, 1]$.
3. Ist $P(\mathbf{q}')/P(\mathbf{q}_\mu) > r$ dann setze man $\mathbf{q}_{\mu+1} = \mathbf{q}'$, andernfalls $\mathbf{q}_{\mu+1} = \mathbf{q}_\mu$.
4. Erhöhe μ um eins und wiederhole die Schritte 2, 3 und 4.

Die so erzeugten Punkte \mathbf{q}_μ im Integrationsgebiet sind gemäß $P(\mathbf{q})$ verteilt, so dass

$$\bar{O} = \frac{1}{M} \sum_{\mu=1}^M O(\mathbf{q}_\mu) \quad (3.40)$$

ein Schätzwert für $\langle O \rangle$ ist, der für große M gegen $\langle O \rangle$ konvergiert. Jeder Punkt \mathbf{q}_μ der Markovkette heisst *Konfiguration*.

Das Programm `samplingflaeche.c` auf Seite 50 berechnet mit Hilfe des Metropolis-Algorithmus Schätzwerte für das eigentliche Integral

$$I = 128 \cdot \frac{\int_0^1 dx dy dz x^3 y^2 z \exp(-x^2 - y^2 - z^2)}{\int_0^1 dx dy dz \exp(-x^2 - y^2 - z^2)} = 128 \cdot \langle x^3 y^2 z \rangle \approx 2.4313142,$$

wobei für P die Exponentialfunktion gewählt wurde. Die Konvergenz zum exakten Resultat ist langsam, der Fehler ist von der Ordnung $1/\sqrt{M}$. Die folgende Tabelle enthält die berechneten Schätzwerte. Der letzte Eintrag ist das Resultat von $M = 10^6$ MC-Iterationen

und hat einen Fehler von -0.00555 .

M	5 000	10 000	15 000	20 000	25 000	30 000	35 000
\bar{I}	2.33113	2.31536	2.33432	2.38934	2.3568	2.34805	2.35253
E	0.10018	0.11595	0.09699	0.04197	0.07449	0.08327	0.07878
M	40 000	45 000	50 000	55 000	60 000	65 000	70 000
\bar{I}	2.34528	2.34193	2.35193	2.35089	2.35659	2.35952	2.36130
E	0.08603	0.08939	0.07938	0.08043	0.07473	0.07179	0.07001
M	75 000	80 000	85 000	90 000	95 000	100 000	1 000 000
\bar{I}	2.36969	2.37196	2.36937	2.38248	2.38742	2.38448	2.43686
E	0.06162	0.05935	0.06194	0.04884	0.04390	0.04683	-0.00555

3.3 Elemente der Wahrscheinlichkeitstheorie

In den bisher angestellten und auch kommenden Untersuchungen spielen Begriffe und Ergebnisse der Wahrscheinlichkeitstheorie eine nicht unerhebliche Rolle. Es lohnt sich deshalb die für uns wesentlichen Begriffe dieser Theorie zusammenzutragen. Die axiomatische Begründung der Theorie wurde in den 1930er Jahren von KOLMOGOROW entwickelt. Sie handelt von Ereignissen die als Mengen aufgefasst werden und denen Wahrscheinlichkeiten zugeordnet sind. Wahrscheinlichkeiten sind reelle Zahlen zwischen 0 und 1. Die Zuordnung von Wahrscheinlichkeiten zu Ereignissen genügt gewissen Anforderungen.

Wir denken uns einen Vorgang, ein Ereignis oder ein Experiment mit zufälligem Ergebnis. Ω sei die Menge der Elementarereignisse. Ein allgemeines Ereignis ist eine Teilmenge von Ω und entsprechend ist der Ereignisraum Σ eine Menge von Teilmengen von Ω . Ein Maß $Pr : \Sigma \rightarrow [0, 1]$ im Sinne der Maßtheorie heisst Wahrscheinlichkeitsmaß wenn $Pr(\Omega) = 1$ gilt. Das Tripel (Ω, Σ, Pr) wird als Wahrscheinlichkeitsraum bezeichnet.

Ein Wahrscheinlichkeitsmaß sollte folgende Axiome erfüllen:

- Für jedes Ereignis $A \subset \Omega$ ist die Wahrscheinlichkeit $Pr(A) \in [0, 1]$.
- Das sichere Ereignis hat die Wahrscheinlichkeit 1, $Pr(\Omega) = 1$.
- Die Wahrscheinlichkeit einer Vereinigung abzählbar vieler disjunkter Ereignisse ist die Summe der Wahrscheinlichkeiten der Einzelereignisse (σ -Additivität)

$$Pr(A_1 \cup A_2 \cup \dots) = \sum Pr(A_i) \quad \text{falls} \quad A_i \cap A_j = \emptyset.$$

Aus diesen Kolmogorowschen Axiomen folgen zum Beispiel die Eigenschaften

$$Pr(\Omega \setminus A) = 1 - Pr(A) \quad \text{und} \quad Pr(A_1 \cup A_2) = Pr(A_1) + Pr(A_2) - Pr(A_1 \cap A_2).$$

Das zufällige Ergebnis $X : \Omega \rightarrow \mathbb{R}$ heisst *Zufallsvariable* mit Erwartungswert (Mittelwert)

$$\langle X \rangle = \sum_{w \in \Omega} X(w) \cdot Pr[w] = \sum_{x \in X(\Omega)} x \cdot \sum_{\substack{w \in \Omega \\ X(w)=x}} Pr[w] = \sum_{x \in X(\Omega)} x \cdot P_X(x), \quad (3.41)$$

wobei $Pr[w]$ die Wahrscheinlichkeit für das Ereignis w ist und

$$P_X(x) = Pr(X^{-1}(x)) \quad (3.42)$$

die Wahrscheinlichkeit dafür angibt, dass X den Wert x annimmt. Der mittlere Wert bei einer wiederholten Messung von X ist gerade $\langle X \rangle$. Die Wahrscheinlichkeit dafür, ein Ergebnis im Intervall Δ zu finden ist

$$P_X(\Delta) = \sum_{x \in \Delta} P_X(x). \quad (3.43)$$

Ist X eine stetige reellwertige Zufallsvariable, dann wird P_X zu einer Wahrscheinlichkeitsdichte und

$$P_X(\Delta) = \int_{\Delta} P_X(x) dx. \quad (3.44)$$

Für eine stetige Funktion f ist der Erwartungswert der Zufallsvariablen $f(X)$ gleich

$$\langle f(X) \rangle = \sum_{w \in \Omega} f(X(w)) \cdot Pr[w] = \sum_{x \in X(\Omega)} f(x) \cdot P_X(x). \quad (3.45)$$

Für beliebige Zufallsvariablen X_1, \dots, X_N ist der Erwartungswert linear,

$$\langle X \rangle = \langle X_1 \rangle + \dots + \langle X_N \rangle. \quad (3.46)$$

Sie heissen *unabhängig*, wenn die Wahrscheinlichkeit für die Ereignisse w mit $X_i(w) = x_i$ für beliebige Ergebnisse x_1, \dots, x_N faktorisiert,

$$Pr(\{w | X_1(w) = x_1, \dots, X_N(w) = x_N\}) = P_{X_1}(x_1) \cdots P_{X_N}(x_N). \quad (3.47)$$

Nach dieser Definition kann X nur unabhängig von X sein, wenn $P_X(x) \in \{0, 1\}$ für alle

x gilt. Die erzeugende Funktion von unabhängigen Zufallsvariablen faktorisiert,

$$\begin{aligned} \langle e^{i(t_1 X_1 + \dots + t_N X_N)} \rangle &= \sum_w e^{i(t_1 x_1 + \dots + t_N x_N)} Pr(\{w | X_1(w) = x_1, \dots, X_N(x) = x_N\}) \\ &= \sum_{x_1, \dots, x_N} \prod_{k=1}^N e^{it_k x_k} P_{X_k}(x_k) = \langle e^{it_1 X_1} \rangle \dots \langle e^{it_N X_N} \rangle. \end{aligned} \quad (3.48)$$

Daraus folgt für $t = t_1 = \dots = t_N$ eine nützliche Gleichung für die erzeugende Funktion der verbundenen Korrelationen unabhängiger Zufallsvariablen,

$$\log \langle e^{itX} \rangle = \sum_{i=1}^N \log \langle e^{itX_i} \rangle, \quad X = X_1 + \dots + X_N. \quad (3.49)$$

Leiten wir zweimal nach t ab und setzen danach $t = 0$, so ergibt sich

$$\text{Var}[X] = \sum_{i=1}^N \text{Var}[X_i], \quad \text{Var}[X] \equiv \langle (\Delta X)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2, \quad (3.50)$$

wobei ΔX die Zufallsvariable $X - \langle X \rangle$ bezeichnet. Die Varianz der Summe von unabhängigen Zufallsvariablen X_i ist gleich der Summe der Varianzen der X_i .

Satz [Markov] Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}^+$

$$P[X \geq t] \leq \frac{1}{t} \langle X \rangle. \quad (3.51)$$

Der Beweis ist nicht sehr schwierig,

$$\begin{aligned} \langle X \rangle &= \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq t} x \cdot P_X(x) \geq \sum_{x \geq t} t \cdot P_X(x) \\ &= t \sum_{x \geq t} P_X(x) = t \cdot P[X \geq t]. \end{aligned}$$

Aus der Markovschen Ungleichung ergibt sich die nützliche Chebyshevsche Schranke für die mittlere Abweichung einer reellwertigen Zufallsvariablen von ihrem Mittelwert,

Satz [Chebyshev] Sei X eine Zufallsvariable und $t \in \mathbb{R}^+$. Dann gilt

$$P[|\Delta X| \geq t] \leq \frac{1}{t^2} \text{Var}[X]. \quad (3.52)$$

Der Beweis dieser Ungleichung ist ein Einzeiler,

$$P[|\Delta X| \geq t] = P[(\Delta X)^2 \geq t^2] \stackrel{(3.51)}{\leq} \frac{\text{Var}[X]}{t^2}.$$

Ein sehr wichtiges Theorem ist das

Satz [Gesetz der großen Zahlen] *Gegeben sei eine Zufallsvariable X . Ferner seien $\epsilon, \delta > 0$ beliebig aber fest. Setzt man*

$$K := \frac{\text{Var}[X]}{\epsilon \cdot \delta^2} = \text{const}, \quad (3.53)$$

dann gilt für alle $N \geq K$: Sind X_1, \dots, X_N unabhängige Zufallsvariablen mit derselben Verteilung wie X und setzt man $Z = (X_1 + \dots + X_N)/N$. Dann gilt

$$P[|\Delta Z| \geq \delta] \leq \epsilon. \quad (3.54)$$

Beweis: Offensichtlich ist

$$\langle Z \rangle = \frac{1}{N} \sum \langle X_i \rangle = \langle X \rangle \quad \text{und} \quad \text{Var}[Z] = \frac{1}{N^2} \sum \text{Var}[X_i] = \frac{\text{Var}[X]}{N}.$$

Die letzte Ungleichung ist für sich interessant. Es ist das berühmte Quadratwurzelgesetz für die *relativen* Schwankungen,

$$\frac{\sqrt{\text{Var}[Z]}}{\langle Z \rangle} \leq \frac{1}{N} \frac{\sqrt{\text{Var}[X]}}{\langle X \rangle}. \quad (3.55)$$

Die Größenordnung der relativen Schwankungen von Z ist somit von $O(N^{-1/2})$. Bei einer großen Menge von X_i (Systemen) dürfen wir infolgedessen die Schwankungen vernachlässigen. Mit der Chebyshev-Ungleichung erhalten wir mit Hilfe der vorletzten Gleichung

$$P[|\Delta Z| \geq \delta] \leq \frac{\text{Var}[Z]}{\delta^2} = \frac{\text{Var}[X]}{N \cdot \delta^2} \leq \epsilon. \quad (3.56)$$

Sei nun X_i eine Folge unabhängiger, gleichverteilter Zufallsvariablen mit verschwindenden Mittelwerten und Kovarianzmatrix $\langle X_i X_j \rangle = \delta_{ij} \sigma^2$, und sei

$$Y_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N X_i. \quad (3.57)$$

Die Erzeugende Funktion für Y_N berechnet sich zu

$$\begin{aligned}\langle e^{itY_N} \rangle &= \prod_{i=1}^N \left\langle \exp \left(i \frac{t}{\sqrt{N}} X_i \right) \right\rangle \\ &= \left\langle 1 + i \frac{t}{\sqrt{N}} X_1 - \frac{1}{2} \frac{t^2}{N} X_1^2 + \frac{1}{O(N^{3/2})} \right\rangle^N \xrightarrow{N \rightarrow \infty} \exp \left(-\frac{1}{2} t^2 \sigma^2 \right).\end{aligned}$$

Andererseits ist die erzeugende Funktion einer Gaußschen Zufallsvariable mit Mittel m und Varianz (Streuungsquadrat) σ^2

$$\frac{1}{\sqrt{2\pi}\sigma} \int dx e^{-(x-m)^2/2\sigma^2} e^{itx} = \exp \left(imt - \frac{1}{2} t^2 \sigma^2 \right).$$

Ein Vergleich zeigt, dass für grosse N die Zufallsvariablen Y_N Gaußverteilt sind mit Mittel 0 und Varianz σ^2 .

3.4 Programme für Kapitel 3

Hier finden sie die im Kapitel über „Hochdimensionale Integrale“ benutzten Programme `1dintegral.c`, `hitmissflaeche.c`, `gaussdistr.c` und `samplingflaeche.c`

1dintegral.c: Dieses Programm berechnet das Integral $\int \exp(x)$ von 0 bis 1 für

$$\epsilon \in \{10^{-n} | n = 1, 2, \dots, 6\}.$$

mit dem Rechteck-, Trapez- und Simpsonverfahren sowie mit Hilfe der Monte-Carlo-Methode.

```
/* Programm 1dintegral.c
/* numerische Integration von f(x) von |alpha bis beta
/* mit vier verschiedenen Verfahren */
#include <stdio.h> #include <stdlib.h> #include <math.h>
/* Hier die zu integrierende Funktion eingeben */
double f(double x)
{return exp(x);}
/* Zufallszahl zwischen 0 und 1 */
double randa(void)
{return (double)rand()/((double)RAND_MAX);}
int main(void)
{
    double epsi ,Sum, I1 , I2 , I3 , I4 , x0 , x1 , x2; /* Integrationsgrenzen */
```

```

    double alpha=0,beta=1;
    long N,i ,M=0;
    for (N=10;N<1000001;N*=10)
    {
        M=M+1;epsi=(beta-alpha)/N;Sum=0;
        /* einfache Methode */
        for (i=0;i<N;i++)
            {x0=alpha+epsi*i;Sum=Sum+f(x0);}
        I1=Sum*epsi;
        /* Trapezregel */
        Sum=0;
        for (i=0;i<N;i++)
            {x0=alpha+epsi*i;x1=x0+epsi;Sum=Sum+(f(x0)+f(x1))/2.0;}
        I2=Sum*epsi;
        /* Simpson-Methode */
        Sum=0;
        for (i=0;i<N-1;i=i+2)
            {x0=alpha+epsi*i;x1=x0+epsi;x2=x1+epsi;
             Sum=Sum+(f(x0)+4.0*f(x1)+f(x2))/3.0;}
        I3=Sum*epsi;
        /* Monte-Carlo-Methode */
        Sum=0;
        for (i=0;i<N;i++)
            {x0=randa();Sum=Sum+f(x0);}
        I4=epsi*Sum;
        printf("%i_%.6f_%.6f_%.6f_%.6f_\n",M,I1,I2,I3,I4);
    }
    return 0; }

```

hitmissflaeche.c: Das Programm berechnet die Fläche unter einer Funktion mit Hilfe der „Hit-or-Miss“ Monte-Carlo Methode. Naiv und mit Verbesserung.

```

/* Programm hitmissflaeche.c
/* Integration von f(x)
/* mit hit-or-miss-Verfahren */
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>
#define M 10000001 /* Anzahl Versuche */
/* Hier die zu integrierende Funktion eingeben */
double f(double x)
{return x*x*exp(x)/(1-x+x*exp(x));}
/* Funktion fuer verbessertes Verfahren */
double g(double x)
{return x*x*exp(x)/(1-x+x*exp(x))-x*x;}

```

```

int main(void)
{double sum1,sum2,I1,I2,sig1,sig2,x,y;
  long n,m;
  srand48(time(NULL));
  for (m=10;m<M;m*=10)
    {sum1=0;sum2=0;
     for (n=1;n<m+1;n++)
       {x=drand48();y=drand48();
        if (y<f(x)) sum1=sum1+1;
        if (y<g(x)) sum2=sum2+1;
       };
     I1=sum1/m;I2=sum2/m;
     sig1=sqrt(I1*(1-I1)/m);sig2=sqrt(I2*(1-I2)/m);
     n=(int)log10(m);
     /* Ausgabe in TeX-array-Format */
     printf("%i&%.6f&%.6f&%.6f&%.6f&%.6f\\\\\\_\\n",
            n,I1,0.376370-I1,sig1,1/3.0+I2,0.043037-I2,sig2);
    };
  return 0;
}

```

gaussdistr.c: Das Programm berechnet die Verteilung der Summe von 10, 50 und 100 Zufallszahlen. Zur Bestimmung der Verteilung werden jeweils 1 Million Versuche gemacht. Aus den zufälligen Werten für s wird ein Histogramm erstellt und im array mean[100] gespeichert. Wir haben die Zufallsvariable s mit $2m$ reskaliert, so dass das Maximum der Verteilung bei $1/2$ liegt.

```

/* Programm gaussdistr.c
/* Summe von M Zufallszahlen
/* im Intervall [0,1] */
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>
#define PI 3.1415926
#define ANZ 10 /* Wieviele Zufallszahlen addiert werden */
#define M 1000000 /* Anzahl MC-Iterationen */
int main(void)
{ double sum,mean[1000];
  double dM=100.0/(double)M; /* Skalierungsfaktor 100 */
  int i,j,sumi;
  /* Initialisierung */
  for (i=0;i<100;i++)
    mean[i]=0;sum=0;srand48(time(NULL));
  /* M-fache Wiederholung des Experiments */
  for (i=0;i<M;i++)

```

```

    {sum=0;
      /* Summe von ANZ Zufallszahlen in jedem Experiment */
      for (j=0;j<ANZ;j++)
        sum=sum+drand48();
      /* 100 Bins fuer Histogramm */
      sumi=(int)(100.0*sum/ANZ);
      ++mean[sumi];
    };
  for (i=30;i<70;i=i+2)
    printf("(%i,%.2f)",i,mean[i]*dM);
  puts("_");
  for (i=30;i<70;i=i+2)
    {sum=i-49.5;
      printf("(%i,%.2f)",i,sqrt(6*ANZ/PI)*exp(-6*ANZ*sum*sum*dM));
    };
  puts("_");
  return 0;
}

```

samplingflaeche.c: Beim *important sampling* zur Berechnung von Integralen werden bevorzugt Punkte mit großem Integranden berücksichtigt und dies verringert die Varianz der einzelnen Schätzung. Das Programm `samplingflaeche.c` berechnet mit Hilfe des Metropolis-Algorithmus das eigentliche Integral

$$128 \cdot \frac{\int_0^1 dx dy dz x^3 y^2 z \exp(-x^2 - y^2 - z^2)}{\int_0^1 dx dy dz \exp(-x^2 - y^2 - z^2)} \approx 2.4313142.$$

```

/* Pogramm samplingflaeche.c
/* Berechnet 3-d Integral mit important sampling.
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>
#define M 100000 /* Anzahl gemessener MC-Iterationen */
#define MA 1000 /* jede MA'te Konfiguration gemessen */
/* Verteilung */
double P(double *x)
{return exp(-x[0]*x[0]-x[1]*x[1]-x[2]*x[2]);}
/* Funktion */
double f(double *x)
{return 128.0*x[0]*x[0]*x[0]*x[1]*x[1]*x[2];}
int main(void)
{ double I,Sum,x[3],y[3];
  long i,j;

```

```
srand48 (time (NULL));
Sum=0; x[0]=drand48 ();x[1]=drand48 ();x[2]=drand48 ();
for (i=1;i<M+1;i++)
{ for (j=0;j<MA;j++)
  { y[0]=drand48 (); y[1]=drand48 (); y[2]=drand48 ();
    if (P(y)>P(x)*drand48 () )
      { x[0]=y[0];x[1]=y[1];x[2]=y[2];};
  };
Sum=Sum+f (x); I=Sum/i ;
if (i%5000==0)
  printf ("i=%ld , I=%0.5f Fehler=%0.5f\n",i , I,2.4313142-I);
};
return 0;
}
```

3.5 Aufgaben

Aufgabe 7: Numerische Berechnung von Integralen

Berechnen Sie mit Hilfe der Simpson-Regel das Integral

$$\int_0^1 dx e^x$$

und vergleichen Sie mit dem exakten Resultat.