

Grundlagen der Monte Carlo Methoden

Christian Theis*
Winfried Kernbichler†

15. Mai 2002

Inhaltsverzeichnis

1	Einführung	2
2	Lösung von Integralen mittels MC	2
2.1	Verallgemeinerung der Methode	4
2.2	Laufzeitverhalten & Konvergenz	7
3	Erzeugung beliebig verteilter Zufallsvariablen	8
3.1	Zufallsgeneratoren, Zufallszahlen & Physik	8
3.2	Zufallsgeneratoren & Gleichverteilung	10
3.3	Beliebig verteilte Zufallszahlen	10
3.3.1	Rejection Method	10
3.3.2	Transformationsmethode	12
3.3.3	Normalverteilung	13
3.4	Beispiel: Mittlerer Teilchenabstand	15
4	Der Metropolis Algorithmus	17
4.1	Beispiel: Punktladungen auf einer Kreisscheibe	19
4.2	Beispiel: Aktienmarkt	21
4.2.1	Black-Sholes Gleichung	23

*ctheis@itp.tu-graz.ac.at

†Inst. für Theoretische Physik, Tel.: +43/3 16/8 73-81 82; winfried.kernbichler@itp.tu-graz.ac.at

1 Einführung

Der Begriff Monte Carlo Methoden kennzeichnet nicht "einen" Algorithmus, sondern eine Gruppe von numerischen Methoden, die Zufallszahlen zur approximativen Lösung oder zur Simulation verschiedener Prozesse einsetzen. Solche stochastische Algorithmen weisen in der Regel folgende Charakteristik auf:

- häufig die einzige Simulationsmethode, die in vernünftiger Rechenzeit brauchbare Resultate liefert
- unter Einsatz von mehr Rechenzeit ist die Lösung systematisch verbesserbar

Die Einsatzgebiete dieser Methoden sind äußerst vielfältig. z.B. bei der Berechnung von Eigenschaften ungeordneter Medien, bei Systemen der Gleichgewichtsstatistik in denen thermische Bewegung eine Rolle spielt, beim Durchgang von Strahlung durch Materie [1], bei Stoß umwandlungen, bei Warteschlangenproblemen, etc. Ebenso eignen sich die Methoden für eigentlich analytische Fragestellungen, wie der Auswertung hochdimensionaler Integrale, oder bestimmter Typen von Differentialgleichungen unter komplexen Randbedingungen (z.B. Poisson-Gleichung mit bewegten Rändern).

In dieser Einführung wird zuerst auf das Grundprinzip der MC Methode eingegangen und danach folgen Beispiele für die Auswertung von Integralen und weiterer Einsatzgebiete in der Physik und Finanzwissenschaft.

2 Lösung von Integralen mittels MC

Wohl das bekannteste Beispiel für die Lösung von Integralen durch MC ist die Berechnung der Zahl π . Dieses Beispiel demonstriert sehr gut die Methode für die Auswertung von Integralen und wird nachher verallgemeinert. Man betrachtet ein Quadrat dem ein Viertelkreis eingeschrieben wird.

Man "schießt" nun blindlings auf dieses Quadrat, indem man zwei voneinander unabhängige, gleichverteilte Zufallszahlen im Intervall $[0,1)$ würfelt, welche wie in Abbildung 2 die Koordinaten des Schusses darstellen. Die Wahrscheinlichkeit, daß der Treffer innerhalb des Kreises liegt verhält sich zur Wahrscheinlichkeit eines Fehlschusses, wie die Kreisfläche zur Quadratfläche. Bei jedem Treffer innerhalb des Kreises erhöht man einen Zähler c , und gleichzeitig wird für jeden Schuß ein Zähler N erhöht.

Aus diesem einfachen Ansatz erhält man nun

$$\pi(N) = 4 \frac{c}{N} \tag{1}$$

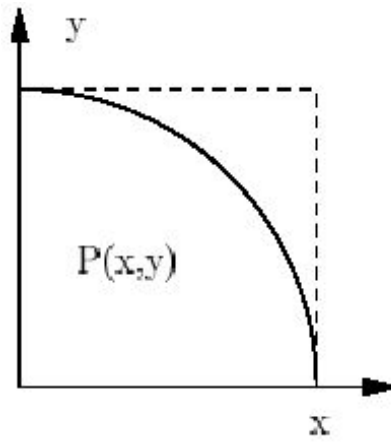


Abbildung 1: Bestimmung von π mittels der MC Methode

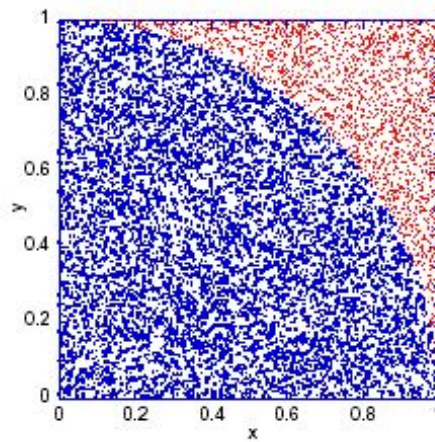


Abbildung 2: 10000 gleichverteilte Punkte $(x,y) \in [0, 1)$. 7874 davon erfüllen die Bedingung $x^2 + y^2 < 1$ und liegen innerhalb des Viertelkreises.

Für den Fehler gilt somit

$$\varepsilon = |\pi - \pi(N)| \quad (2)$$

2.1 Verallgemeinerung der Methode

Bevor wir nun zum allgemeinen Ansatz der MC Integration kommen, müssen zur Erinnerung noch ein paar grundsätzliche Definitionen der Wahrscheinlichkeitstheorie festgehalten werden.

Definition Zufallsvariable

Eine Zufallsvariable ist ein Funktional, das jedem Ergebnis $\omega \in G$ (Gesamtheit der möglichen Ereignisse) eine reelle Zahl $x = X(\omega)$ zuordnet. x heißt Realisierung von X . Die Menge R der möglichen Realisierungen x heißt Wertebereich von X . G wird hierbei auf R abgebildet.

Beispiele:

- Zuweisen von 0 für Kopf und 1 für die Zahl einer Münze beim Münzwurf
- Beim Würfeln weist man dem Ereignis "n Augen" die Zahl n zu

Definition Verteilungsfunktion $\Phi(x)$

Unter der Verteilungsfunktion versteht man die Wahrscheinlichkeit

$$\Phi(x) = P(X < x) \quad (3)$$

dass eine Zufalls-Variable X Werte kleiner als x annimmt. Man nennt die Verteilungsfunktion auch kumulative Wahrscheinlichkeit.

Definition Wahrscheinlichkeitsdichte $\varphi(x)$

(probability density function - PDF)

ist die Ableitung der Verteilungsfunktion $\Phi(x)$ nach der Zufallsvariablen

$$\varphi(x) = \frac{d\Phi(x)}{dx} \quad (4)$$

Die PDF legt also die Wahrscheinlichkeit fest, daß die Zufallsvariable X Werte aus einem infinitesimalen Intervall dx bei x annimmt, dividiert durch die Intervallgröße dx . Beim Würfeln wäre dies z.B. $1/6$ bei einem nicht manipulierten Würfel.

Definition Normierungsbedingung

Das Integral über die Wahrscheinlichkeitsdichte muß 1 (=das sichere Ereignis) ergeben.

$$\int_{-\infty}^{\infty} \varphi(x) dx = 1 \quad (5)$$

Im Falle eines Spielwürfels würde sich diese Summe zu $1/6+1/6+1/6+1/6+1/6+1/6 = 1$ ergeben.

Definition Erwartungswert

einer Zufallsvariable x mit der Wahrscheinlichkeitsdichte $\varphi(x)$

$$\langle x \rangle = \int x \varphi(x) dx \approx \frac{1}{N} \sum_i^N x_i \quad (6)$$

Definition Varianz

einer Zufallsvariablen x ist definiert als

$$Var[x] = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \quad (7)$$

Definition Standardabweichung

$$Std[x] = \sqrt{Var[x]} \quad (8)$$

Wir gehen nun von einer m -dimensionalen Zufallsvariable $x \in R^m$ aus und definieren eine ein-dimensionale Zufallsvariable $O(x)$, die eine eindeutige Funktion von x darstellt. Der Erwartungswert von $O(x)$ lautet nun

$$\langle O \rangle = \int O(x) \varphi(x) d^m x \quad (9)$$

Das Integral entspricht folglich dem Mittelwert der Zufallsvariablen $O(x)$. Der Zentrale Grenzwertsatz besagt nun, daß das arithmetische Mittel

$$\frac{1}{N} \sum_i^N O(x_i) \quad (10)$$

der Zufallszahlen $O(x_i)$ für hinreichend große N einer Normalverteilung mit Erwartungswert $\langle O \rangle$ genügt. Somit kann man hochdimensionale Integrale vom Typ

$$\int O(x)\varphi(x)dx$$

aus einer Stichprobe $\{x_1, \dots, x_i\}$ von Zufallszahlen der Verteilung $\varphi(x)$ durch das arithmetische Mittel und dessen Fehler approximieren.

Theorem Der zentrale Satz der Monte Carlo Rechnung

zur Auswertung von Integralen lautet also

$$\langle f(x) \rangle = \int f(x)\varphi(x)dx \approx \frac{1}{N} \sum_i^N f(x_i) \pm \sqrt{\frac{\langle f(x)^2 \rangle - \langle f(x) \rangle^2}{N}} \quad (11)$$

Beispiel

Wir wählen nun für die Verteilung $\varphi(x)$ eine Gleichverteilung im Intervall $x \in [a, b]$, $\varphi(x) = \frac{1}{b-a}$. Durch Einsetzen in Gl. (11) erkennt man, daß man unter Verwendung der Summenschreibweise das Integral leicht auswerten kann

$$\frac{1}{b-a} \int_a^b g(x)dx \approx \frac{1}{N} \sum_i^N g(x_i) \quad (12)$$

$$\Leftrightarrow \int_a^b g(x)dx \approx \frac{b-a}{N} \sum_i^N g(x_i) \quad (13)$$

Man tastet das Integrationsgebiet somit durch das Würfeln der Zufallszahlen ab. Da dies in unserem Fall mit einer Gleichverteilung erfolgt spricht man vom einfachsten Fall des *simple samplings*. Dieser naive Ansatz einer Gleichverteilung funktioniert jedoch nur bei Integranden die genügend glatt und auf begrenzten Intervallen definiert sind. Bei Integranden die stark variieren ist die Wahrscheinlichkeit für den Integranden passende Zufallszahlen zu würfeln jedoch gering und man müßte eine große Anzahl an Zufallszahlen würfeln. Da dies in der Praxis nicht gewünscht ist, geht man von der Gleichverteilung ab und verwendet andere Verfahren zur Erzeugung von Zufallszahlen die einer gewünschten Verteilung folgen.

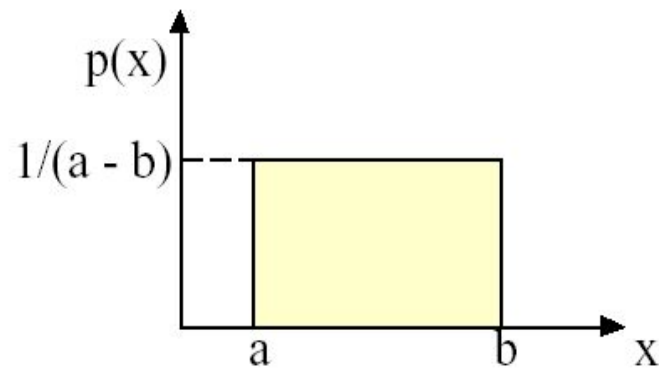


Abbildung 3: Wahrscheinlichkeitsdichte einer Gleichverteilung im Intervall $[a, b]$

2.2 Laufzeitverhalten & Konvergenz

Natürlich könnte man Integrale der Form $\int f(x)\varphi(x)dx$ auch mit herkömmlichen numerischen Integrationsmethoden lösen. Dies wäre im Fall niedriger Dimensionen auch sinnvoll, jedoch betrachten wir nun die Integration einer n -dimensionalen Funktion und dem Integrationsgebiet eines Quaders im n -dimensionalen Hyperraum mit dem Volumen h^n . Bereits bei einer ungenauen Berechnung unter Verwendung von nur 10 Funktionswerten für jede der n Koordinaten, benötigt man bereits 10^n Werte! Für den Fehler der Trapezmethode zur numerischen Integration gilt

$$\varepsilon = \frac{1}{N^{\frac{2}{n}}} \quad (14)$$

N ... Anzahl der notwendigen berechneten Funktionswerte des Integranden

Der Vergleich mit dem Fehler der Monte Carlo Integration

$$\varepsilon = \frac{1}{N^{\frac{1}{2}}} \quad (15)$$

zeigt, daß bei höheren Dimensionen ($n > 4$) die Konvergenz herkömmlicher numerischer Integrationsmethoden der Monte Carlo Integration unterliegt. Bei typischen Integralen der Finanzwirtschaft liegen z.B. Dimensionen $n = 365$ vor und es ist daher beim Vergleich der Rechenzeiten leicht ersichtlich, daß MC der einzige praktikable Weg ist in vernünftiger Zeit Resultate zu erzielen. Wichtig ist, daß bei Erhöhung der Anzahl der verwendeten Samples die Genauigkeit bei Monte Carlo verbessert werden kann, was bei herkömmlichen Integrationsmethoden nicht der Fall ist.

In Abbildung (4) ist das Laufzeitverhalten bei der Integration zur Volumsberechnung

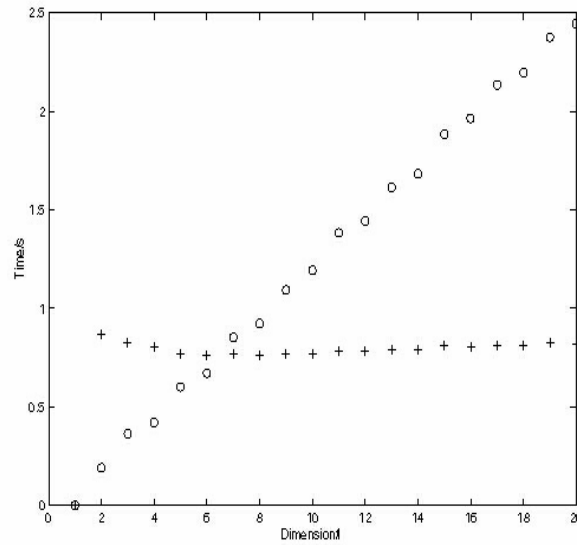


Abbildung 4: Vergleich der Laufzeit zur Volumsberechnung einer n-dimensionalen Einheitskugel. (+ Monte Carlo, o Trapezmethode)

einer n-dimensionalen Einheitskugel dargestellt. Hier ist zu beachten, daß nicht der quantitative Vergleich der Rechenzeit im Vordergrund steht, da dies vom verwendeten Rechner, der Implementierung, Programmiersprache usw. abhängt und zudem für beide Methoden der gleiche Fehlerbereich gefordert werden müßte. Abbildung (4) soll den qualitativen Zusammenhang zwischen Rechenzeit und Dimension in Bezug auf das verwendete Verfahren illustrieren.

Das Volumen einer n-dimensionalen Einheitskugel ist gegeben durch

$$V^n = \int \cdots \int R^2 - \sum_{i=1}^n x_i^2 dx_1 dx_2 \dots dx_n. \quad (16)$$

Für die Einheitskugel mit Radius 1 läßt sich Gl.(16) als Rekursionsformel angeben

$$V^{n+1} = V^n \int_{-1}^1 (1 - s^2)^{n/2} ds. \quad (17)$$

Als Startwert für Gl.(16) setzt man für die Einheitskugel $V^1 = 2$.

Der Vergleich der Rechenzeit in Abbildung (4) von Monte Carlo Integration und Trapezmethode illustriert den Vorteil der MC Methode besonders bei hohen Dimensionen.

3 Erzeugung beliebig verteilter Zufallsvariablen

3.1 Zufallsgeneratoren, Zufallszahlen & Physik

Der Kern von Monte Carlo Simulationen ist das Erwürfeln irgendwelcher Zufallsgrößen. Man kann nun in die Erzeugung dieser Zufallszahlen die zu untersuchende Physik stecken, indem man sie in Form einer Verteilung modelliert. Aus der Sicht der Gibbschen Formulierung der statistischen Mechanik betrachtet man nicht ein einzelnes Teilchensystem, sondern eine Anzahl an ungenauen Kopien dieses Systems. Über dieses Ensemble an Kopien wird nun der Mittelwert gebildet. Die Erzeugung der ungenauen Kopien erfolgt durch zufällige Veränderung der gerade betrachteten Konfiguration. (siehe Bsp. mittlerer Teilchenabstand). Die Abweichung wird jedes mal erwürfelt. Man spricht von einem random walk im Gibbschen Phasenraum. Bei diesen Systemen ist es erforderlich einen "gezinkten" Würfel zu benutzen, der eine Folge von Zufallszahlen mit vorgegebenen statistischen Eigenschaften generiert [1].

In diesem Abschnitt werden Verfahren vorgestellt, wie man Zufallszahlen mit definierten Verteilungsdichten $\varphi(x)$ erstellt. Zuerst möchte ich aber noch kurz allgemein auf Zufallszahlengeneratoren eingehen.

3.2 Zufallsgeneratoren & Gleichverteilung

Wirkliche Zufallszahlen zu erzeugen ist mit deterministischen Maschinen wie einem Computer nicht möglich. Daher wurden Verfahren entwickelt um sogenannte "Pseudozufallszahlen" zu errechnen. Streng genommen dürfte man mit diesen errechneten Zufallszahlen keine Statistik betreiben, da die Generatoren nach einer bestimmten Zeit die Sequenz der erzeugten Zahlen wiederholen. Es ist nun ein Zeichen der Qualität eines Generators wenn die Periode sehr lang ist. Die Diskussion über die Qualität und Tests von Zufallszahlengeneratoren ist jedoch ein sehr komplexes Thema und würde hier zu weit führen. Für nähere Information sei auf das Standardwerk [9] und [8] verwiesen.

Das bekannteste Verfahren zur Erzeugung einer Gleichverteilung ist der **Multiplicative Linear Congruential Generator**. Die Methode zur Berechnung der Zufallszahlen lautet

$$x_{n+1} = (ax_n + c)(\text{mod } m)$$

Häufig wird die Konstante $c = 0$ gewählt und es muß gelten $0 < a < m$. Der Modul m gibt den Bereich an, in dem die Zahlen liegen. Die Wahl der Konstanten ist jedoch die Periodenlänge und weitere Güteermere wie Korrelation ausschlaggebend!

Ein weiteres Verfahren ist das **Fibonacci** Verfahren:

$$x_{n+1} = x_n + x_{n-1}(\text{mod } m)$$

Im allgemeinen ist zu bemerken, daß man in der Regel auf bewährte verfügbare Generatoren zurückgreifen sollte, da durch ungünstige Wahl der Konstanten sehr leicht störende Einflüsse wie Korrelation auftreten! Bevor man der Versuchung erliegt selbst einen Zufallszahlengenerator zu schreiben ist es ratsam sich eingehend mit [9], Statistischer Analyse und aktuellen Veröffentlichungen zu diesem Thema beschäftigen. Trotz der angeratenen Verwendung von verfügbaren Generatoren sollte man diese jedoch nicht "blind" verwenden, da es eine Unmenge an für ernsthafte Anwendungen unbrauchbaren Zufallszahlengeneratoren gibt!

3.3 Beliebige verteilte Zufallszahlen

3.3.1 Rejection Method

Die Voraussetzung ist daß die Verteilungsdichte beschränkt ist $0 < \varphi(x) < M$. Man würfelt nun eine gleichverteilte Zufallszahl r_1 im Intervall $[0,1)$ welche auf den Bereich $[a,b]$ abgebildet wird:

$$x = a + (b - a)r_1 \quad (18)$$

Dieser Wert wird mit der Wahrscheinlichkeit $\frac{\varphi(x)}{M}$, wie in Abb.(5) dargestellt, akzeptiert. Dazu muß man noch eine zweite gleichverteilte Zufallszahl r_2 aus $[0,1)$ ermitteln und man akzeptiert x aus Gl.(3) nur unter der Voraussetzung

$$r_2 \leq \frac{\varphi(x)}{M}$$

Ist diese Bedingung nicht erfüllt werden x , r_1, r_2 verworfen und man beginnt erneut. Der Algorithmus liefert also nicht bei jedem Schritt eine Zufallszahl x , da es oft zur Ablehnung des "Vorschlags" kommt. Zufallszahlen x werden häufiger akzeptiert, wenn die Verteilungsdichte für diesen Wert größer ist und umgekehrt. Die Effizienz des Verfahrens ist gegeben durch $E = \frac{1}{M(b-a)}$.

Die beschränkende Funktion ist hier eine Konstante M . Manchmal jedoch hat man den Fall, daß die zu erzeugende Verteilung eine schmale und hohe Spitze aufweist und die Funktionswerte im restlichen Bereich viel tiefer liegen. Hier würde man nun viele Vorschläge verwerfen müssen. Eine Lösung ist nun die beschränkende Funktion nicht als Konstante M anzunehmen sondern als Funktion $G(x)$. *Man würde x nun aus dieser Verteilung $G(x)$ und nicht einer Gleichverteilung ziehen!!!*

Algorithmus Rejection Method

- Ziehe x aus einer Gleichverteilung im Intervall $[a,b]$. oder $G(x)$.

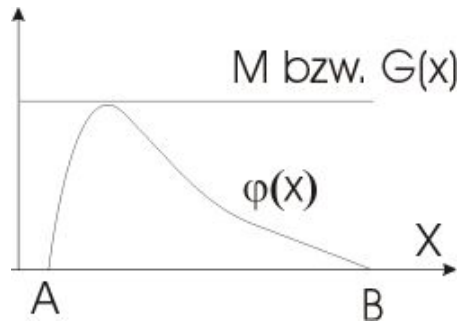


Abbildung 5: Rejection Methode zur Erzeugung einer beliebigen Verteilung

- Ziehe eine gleichverteilte Zufallszahl r aus $[0,1)$.
- Akzeptiere x wenn $r \leq \frac{\varphi(x)}{G(x)}$ bzw $r \leq \frac{\varphi(x)}{M}$

Man darf nicht vergessen bei einer Ablehnung jeweils neue Zufallszahlen zu verwenden, da ein "Recycling" zur Verfälschung des Verfahrens führt!

3.3.2 Transformationsmethode

Man betrachtet die Verteilungsdichte $\varphi(x)$ und die Funktion $y = f(x)$, welche die Variablentransformation durchführt und somit die Abbildung auf die Verteilungsdichte $\varphi(y)$ darstellt (siehe Abb.6). $f(x)$ muß eine bijektive Abbildung sein mit $x = f(y)^{-1}$. Aufgrund der Erhaltung der Wahrscheinlichkeit bei der Transformation muß gelten:

$$|d\Phi(y)| = |d\Phi(x)|. \quad (19)$$

Daraus folgt nun mit Gl.(4)

$$|\varphi(y)dy| = |\varphi(x)dx| \Rightarrow \varphi(y) = \varphi(x) \left| \frac{dx}{dy} \right| \Rightarrow \varphi(y) = \varphi(f(y)^{-1}) \left| \frac{df(y)^{-1}}{dy} \right|$$

Man kann nun Zufallszahlen mit einer vorgegebenen Verteilungsdichte $\varphi(x)$ erzeugen und muß eine Funktion $y = f(x)$ finden, sodaß die Verteilung der y eine Gleichverteilung $\varphi(y) = c$ ist. Es folgt dann

$$\varphi(x) = \varphi(y) \left| \frac{dy}{dx} \right| = c \left| \frac{df(x)}{dx} \right| \quad (20)$$

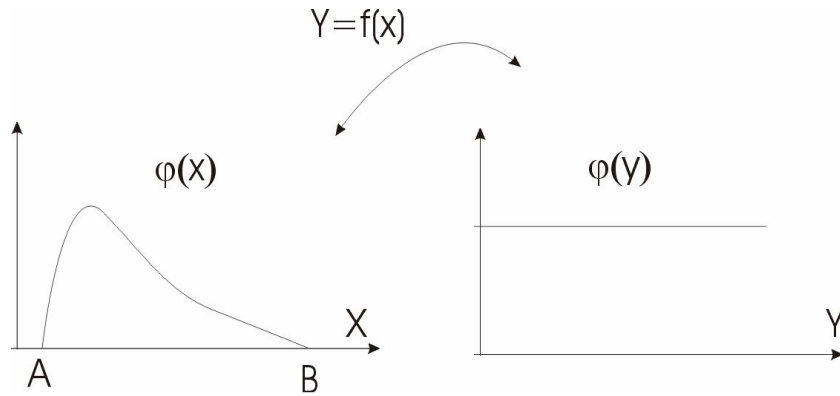


Abbildung 6: Transformation zwischen zwei Verteilungen

Unter der Annahme $c = 1$ erkennt man, daß $f(x) = \Phi(x)$ die Differentialgleichung (20) erfüllt. Daraus folgt

$$x = \Phi^{-1}(y) \quad (21)$$

Bei mehreren Variablen bezeichnet $\left| \frac{dy}{dx} \right|$ die Jakobideterminante!

Algorithmus Transformationsmethode

- Ziehe y aus einer Gleichverteilung $[0,1]$
- Ermittle x durch die Umkehrfunktion $x = \Phi^{-1}(y)$

Beispiel

Wir haben die Gleichverteilung

$$\varphi(\xi) = \begin{cases} 0 & \xi < 0 \\ \xi & 0 \leq \xi \leq 1 \\ 1 & \xi \geq 1 \end{cases} \implies \Phi(\xi) = \xi$$

Die zu erzeugende Verteilung $\varphi(y)$ ist gegeben durch

$$\varphi(y) = \lambda e^{-\lambda y} \implies \Phi(y) = \int_0^y \lambda e^{-\lambda u} du = 1 - e^{-\lambda y}$$

Unter Beachtung von Gl. (19) erhalten wir

$$1 - e^{-\lambda y} = \xi \implies y = -\frac{1}{\lambda} \ln(1 - \xi)$$

ξ wird hierbei aus einer Gleichverteilung $[0,1)$ gezogen.

3.3.3 Normalverteilung

Hier sei noch das Box-Müller Verfahren zur Erzeugung normalverteilter Zufallsvariablen angegeben. Die gemeinsame Verteilungsfunktion für zwei unkorrelierte und normalverteilte Zufallsvariablen lautet

$$\varphi(x_1, x_2) = \varphi(x_1)\varphi(x_2) = \frac{1}{2\pi} e^{-\frac{(x_1^2 + x_2^2)}{2}} \quad (22)$$

Wir führen nun Polarkoordinaten (r, ϑ) anstatt von x_1 und x_2 und es ergibt sich dadurch

$$x_1 = r \cos \vartheta \quad (23)$$

$$x_2 = r \sin \vartheta \quad (24)$$

$$\varphi(r, \vartheta) dr d\vartheta = \varphi(r) dr \varphi(\vartheta) d\vartheta = \left(r e^{-\frac{r^2}{2}} dr \right) \left(\frac{1}{2\pi} d\vartheta \right) \quad (25)$$

Wie man aus Gl.(25) erkennen kann ist die Verteilung des Winkelanteils durch $\varphi(\vartheta) = \frac{1}{2\pi}$ bereits gleichverteilt in $[0, 2\pi]$. Unter Verwendung der Transformationsmethode folgt, daß der Winkel ϑ mit Hilfe der im Intervall $[0,1]$ gleichverteilten Zufallsvariable y_1 ermittelt werden kann. Die Berechnung erfolgt durch

$$\vartheta = 2\pi y_1 \quad (26)$$

Nun bleibt noch der Radialanteil zu behandeln. Wir berechnen nun mit Gl.(4) die Wahrscheinlichkeit

$$y_2 = \Phi(r) = \int_0^r \varphi(\xi) d\xi = 1 - e^{-\frac{r^2}{2}} \quad (27)$$

Aus diesem Ergebnis sieht man, daß y_2 auch gleichverteilt ist im Intervall $[0,1]$. Durch neuerliche Anwendung der Transformationsmethode erhält man

$$r = \sqrt{-2 \ln(1 - y_2)} \quad (28)$$

Durch Einsetzen von Gl.(28) und Gl.(26) in Gl.(23) folgt

$$x_1 = \sqrt{-2 \ln y_1} \cos(2\pi y_2) \quad (29)$$

$$x_2 = \sqrt{-2 \ln y_1} \sin(2\pi y_2) \quad (30)$$

Algorithmus Normalverteilung

- Ziehe y_1 und y_2 aus einer Gleichverteilung $[0,1]$
- Bilde $x_1 = \sqrt{-2 \ln y_1} \cos(2\pi y_2)$
 $x_2 = \sqrt{-2 \ln y_1} \sin(2\pi y_2)$

Die so generierten Zufallsvariablen x_1 und x_2 sind normalverteilt und statistisch unabhängig. Gaußverteilte Zufallszahlen mit gegebener Varianz σ^2 erhält man, indem man x_1 und x_2 mit σ multipliziert. Es gibt noch weitere optimierte & fortgeschrittene Verfahren wie Importance Sampling, Probability Mixing welche man in [1],[8], [6] findet.

3.4 Beispiel: Mittlerer Teilchenabstand

Man betrachte ein zweidimensionales System von n klassischen Teilchen in einem Kasten der Kantenlänge L . Der Teilchendurchmesser betrage s und es wirke ein Hard Core Potential. Der mittlere Teilchenabstand ergibt sich zu

$$\langle r \rangle = \frac{1}{Z} \int \sum_{i < j} \sqrt{[(x_i - x_j)^2 + (y_i - y_j)^2]} d\vec{X} \quad (31)$$

Z ... Phasenraumvolumen.

Als Nebenbedingung legen wir fest, daß die Teilchen sich nicht durchdringen dürfen

$$\sqrt{[(x_i - x_j)^2 + (y_i - y_j)^2]} > s \quad (32)$$

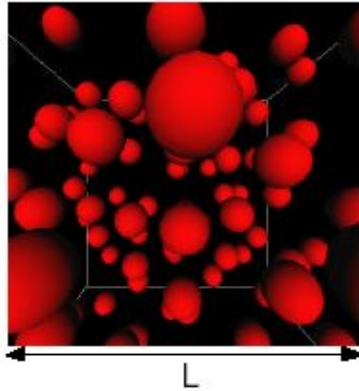


Abbildung 7: OpenGL-Visualisierung der Partikelsimulation zur Ermittlung des mittleren Teilchenabstandes. Man sieht hier aufgrund der Nebenbedingung keine Durchdringung.

Der Phasenraum ist $2n$ dimensional, wodurch also auch das Integral in Gl. (31) die Dimension $2n$ besitzt. Da man zudem auch noch eine Nebenbedingung in Gl. (32) erfüllen muß, ist die MC Methode der einzige praktikable Lösungsweg.

Zur Lösung des Problems wählt man folgende Vorgangsweise:

- Ermittlung eines zufälligen Teilchens i .
- Bestimmung zweier Zufallszahlen $x_i, y_i \in [0, L]$. Diese Zahlen repräsentieren die mögliche nächste Position des Teilchens i . Erfüllt die somit erwürfelte Position die Nebenbedingung in Gl. (32) für alle j , so erhält das Teilchen die neue Position. Ansonsten wird die alte Position beibehalten und man beginnt mit der neuerlichen zufälligen Ermittlung eines Teilchens i .
- In der neu gewonnenen Konfiguration berechnet man

$$\langle r \rangle = \sum_{i < j} \sqrt{[(x_i - x_j)^2 + (y_i - y_j)^2]} \quad (33)$$

Man führt diese Berechnung N mal aus und mittelt dann über alle erhaltenen Werte. Somit erhält man den mittleren Teilchenabstand mit einem Fehler $\Delta \propto \frac{1}{\sqrt{N}}$. Durch Variation der Iterationsanzahl erkennt man, daß bei größerer Zahl an Iterationen der mittlere Teilchenabstand minimiert wird. Es ist in Abb.(8) eindeutig zu erkennen, daß

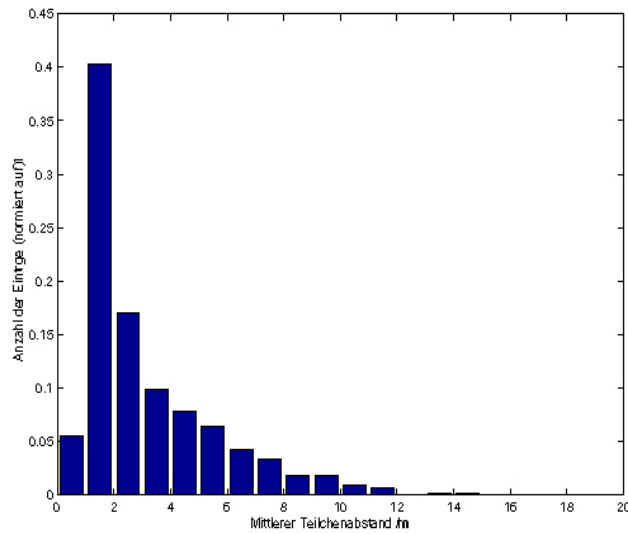


Abbildung 8: Histogramm über simulierte mittlere Teilchenabstände (Teilchenradius 1nm)

die Streuung geringer wird, was mit einem eindeutigen Maximum beim minimalen Teilchenabstand im Histogramm über die ermittelten Abstände sichtbar ist.

Möchte man ein realistischeres Potential berücksichtigen, wird der Schritt der Akzeptanz etwas verändert. Man verwendet nun den sogenannten **Metropolis Algorithmus** um die Akzeptanz zu ermitteln.

4 Der Metropolis Algorithmus

Man betrachtet nun ein System von Teilchen (z.B. Gas). Die Wahrscheinlichkeit, daß sich das System im Zustand C mit der Energie E_C befindet, ergibt sich aufgrund der Boltzmann Verteilung zu

$$p(C) = \frac{1}{\underbrace{\sum_C \exp\left(-\frac{E_C}{kT}\right)}_{Z_T}} \exp\left(-\frac{E_C}{kT}\right) \quad (34)$$

Der Normierungsfaktor Z_T bezeichnet somit die Summe über alle möglichen Zustände. Er ergibt sich aus der Bedingung, daß die Summe aller Wahrscheinlichkeiten gleich 1

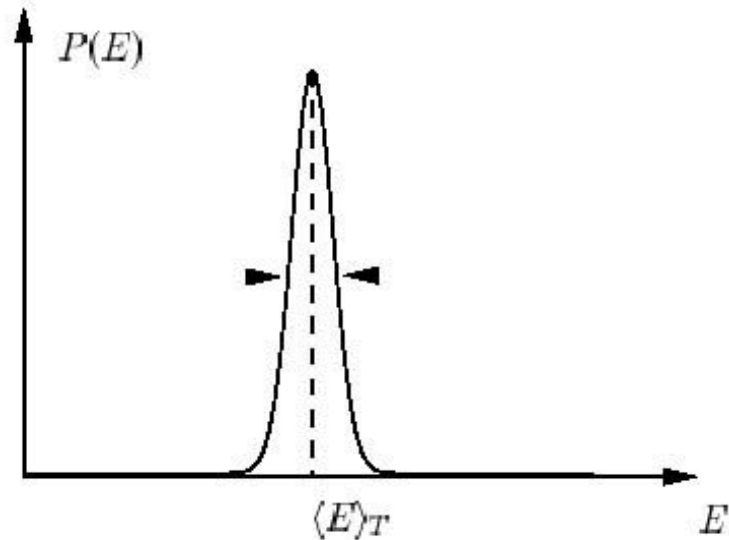


Abbildung 9: Wahrscheinlichkeitsverteilung eines Zustandes in Abhängigkeit von seiner Energie

sein muß! Die Wahrscheinlichkeitsverteilung eines Zustandes in Abhängigkeit von seiner Energie ist bei großen Systemen stark um den Erwartungswert $\langle E_T \rangle$ konzentriert. Dies zeigt sich in Abb.(9) in einer Gausskurve mit geringer Varianz.

Würde man mit dieser Verteilung einen Erwartungswert unter Verwendung von Gl. (11) errechnen, so würde man viele Summenglieder berechnen, für die $p(E_C) = 0$ ist. Der von Metropolis, Rosenbluth & Teller vorgeschlagene Algorithmus löst dieses Problem auf einfache und elegante Weise. Man erzeugt eine Reihe von Systemzuständen, bei denen jeder Zustand nur von seinem Vorgänger abhängig ist¹. Um von einem aktuellen Zustand X ausgehend zu einem Folgezustand Y zu gelangen, ist folgendes zu tun:

Algorithmus Metropolis

- Man macht einen beliebigen Zustandsvorschlag Y. Dabei ist zu beachten, daß bei der Betrachtung aller möglichen Zugvorschläge jeder Punkt im Phasenraum (sprich jeder mögliche Zustand) erreicht werden kann! Außerdem **muß** die Wahrscheinlichkeit von Y nach X zu gelangen gleich sein, wie umgekehrt.
- Man akzeptiert den Vorschlag Y mit folgender Akzeptanzwahrscheinlichkeit

$$A(X \rightarrow Y) = \min \left(1, \frac{p(Y)}{p(X)} \right) \quad (35)$$

¹Solche Reihen bezeichnet man als Markov Ketten. Ausführlichere Diskussion von Markov Ketten findet man in [1], [2], [6].

Unter Verwendung von Gl. (34) erhält man somit

$$\frac{p(Y)}{p(X)} = \frac{e^{-\frac{E_Y}{kT}} \left(\sum_i e^{-\frac{E_i}{kT}} \right)}{\left(\sum_i e^{-\frac{E_i}{kT}} \right) e^{-\frac{E_X}{kT}}} = e^{-\frac{\Delta E}{kT}} \quad (36)$$

$$A(X \rightarrow Y) = \min \left(1, \exp\left(-\frac{\Delta E}{kT}\right) \right) \quad (37)$$

- Man würfelt nun eine im Intervall [0,1) gleichverteilte Zufallsvariable r und akzeptiert den Vorschlag wenn

$$r < \min \left(1, \exp\left(-\frac{\Delta E}{kT}\right) \right) \quad (38)$$

Ein Zugvorschlag bei dem es zu einer Erniedrigung der Energie kommt, wird somit immer akzeptiert. Ein Vorschlag welcher eine Erhöhung zur Folge hätte, wird mit der Wahrscheinlichkeit $\exp(-\frac{\Delta E}{kT})$ angenommen. Wichtig ist, daß auch ungünstigere Konfigurationen akzeptiert werden, dies jedoch mit einer geringeren Wahrscheinlichkeit. Dies ist notwendig um den gesamten Phasenraum zu untersuchen! Für Details und Beweisführungen bezüglich notwendiger Bedingungen wie Detailed Balance für Gl.(35) sei auf [2] verwiesen.

4.1 Beispiel: Punktladungen auf einer Kreisscheibe

Im folgenden betrachten wir ein 2D Modell eines Ensembles aus N Partikeln, welches durch N Koordinaten innerhalb eines Kreises mit Radius R festgelegt ist

$$\{(Rx_i, Ry_i) | i = 1, \dots, N; x, y \in \mathfrak{R}; x_i^2 + y_i^2 < 1\}$$

Die potentielle Energie ist gegeben durch

$$V = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{q^2}{Rr_{ij}} \quad (39)$$

q...Ladung

R ...Kreisradius, r_{ij} ...Abstand zwischen den Partikeln

Zur Durchführung der Simulation benötigt man gleichverteilte Koordinaten innerhalb des Einheitskreises (Fläche $A = \pi$) mit der Verteilungsdichte $\varphi(x) = \frac{1}{\pi}$. Dafür führt man eine Koordinatentransformation durch

$$x = \sqrt{\xi} \cos(2\pi\zeta), \quad (40)$$

$$y = \sqrt{\xi} \sin(2\pi\zeta) \quad (41)$$

Die neuen Variablen ζ und ξ sind gleichverteilt im Intervall $[0,1)$. Zur Transformation benötigt man die aus der Wahrscheinlichkeitstheorie bekannte Transformationsbeziehung

$$\varphi(x, y)dxdy = \varphi(\xi, \zeta)d\xi d\zeta \frac{dxdy}{\left| \frac{\partial(x,y)}{\partial(\xi,\zeta)} \right| d\xi d\zeta} \quad (42)$$

Die Anzahl der Punkte welche bei (ξ, ζ) innerhalb des Bereiches $d\xi d\zeta$ generiert wurden, wird in den Bereich $\left| \frac{\partial(x,y)}{\partial(\xi,\zeta)} \right| d\xi d\zeta$ um (x,y) transformiert. Daher gilt

$$\begin{aligned} \varphi(x, y) &= \varphi(\xi, \zeta) d\xi d\zeta \frac{1}{\left| \frac{\partial(x,y)}{\partial(\xi,\zeta)} \right|} = \left| \begin{array}{cc} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \zeta} \end{array} \right|^{-1} \varphi(\xi) \varphi(\zeta) \\ &= \left| \begin{array}{cc} \frac{1}{2\sqrt{\xi}} \cos(2\pi\zeta) & -2\pi\sqrt{\xi} \sin(2\pi\zeta) \\ \frac{1}{2\sqrt{\xi}} \sin(2\pi\zeta) & 2\pi\sqrt{\xi} \cos(2\pi\zeta) \end{array} \right|^{-1} \\ &= \frac{1}{\pi(\cos^2(2\pi\zeta) + \sin^2(2\pi\zeta))} = \frac{1}{\pi} \end{aligned} \quad (43)$$

Somit zeigt Gl.(43), daß Gl.(40) eine gültige Transformation für unser Problem darstellt. Da die Temperatur nur in Verbindung mit der potentiellen Energie

$$\frac{V}{k_B T} = \frac{q^2}{Rk_B T} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{1}{r_{ij}},$$

auftritt, entspricht Teilchenzahl N den Modellparametern und die normierte inverse Temperatur ist gleich $\frac{q^2}{Rk_B T}$.

Nun benötigt man noch die Änderung der Energie bei einem Positionswechsel eines Teilchens k von $(x_k, y_k) \Rightarrow (x'_k, y'_k)$. Diese ergibt sich zu

$$\frac{\Delta V}{k_B T} = \frac{q^2}{Rk_B T} \sum_{j \neq k} \frac{1}{r'_{kj}} - \frac{1}{r_{kj}} \quad (44)$$

Zur Lösung des Problems wählt man folgende Vorgangsweise:

- Festlegen der Anfangsparameter $\frac{q^2}{Rk_B T}$ und N
- Generierung einer zufälligen Anfangskonfiguration der Teilchenkoordinaten
- Zufällige Wahl eines Teilchens k
- Zufällige Wahl einer neuen Teilchenposition
- Berechnung der Energieänderung unter Verwendung von Gl.(44)
- Falls $\frac{\Delta V}{k_B T} < 0$ wird die neue Position akzeptiert, andernfalls akzeptiert man die Position nur, wenn für eine im Intervall $[0,1]$ gleichverteilte Variable r gilt: $r < \exp\left(-\frac{\Delta V}{k_B T}\right)$
- Wiederhole ab Punkt 3 bis die maximale Iterationsanzahl erreicht ist.

Man erkennt in Abbildung (10) bei Durchführung der Simulation die zu erwartende Kreisverteilung auf der Scheibe. Durch Variation der inversen Temperatur kann man sehr gut die Abhängigkeit der Akzeptanz studieren. Bei einem Wert von 0 wird somit jeder Vorschlag akzeptiert und die Energie wird nicht minimiert. Erst bei größeren Werten erkennt man die eindeutige Tendenz zur Energieminimierung, da die Wahrscheinlichkeit ungünstige Konfigurationen zu akzeptieren kleiner wird.

4.2 Beispiel: Aktienmarkt

Der Aktienmarkt weist auch stochastische Veränderungen auf und ist damit ein weiteres Anwendungsbeispiel für Monte Carlo Methoden. Zuerst einen kurzen Überblick über das notwendige finanztheoretische Wissen.

Der Aktienpreis $S(t)$ kann als Zufallsvariable gesehen werden, welche durch einen Diffusionsprozeß idealisiert werden kann. Es gilt die Gleichung

$$dS = \mu S dt + \sigma S \varepsilon \sqrt{dt} \quad (45)$$

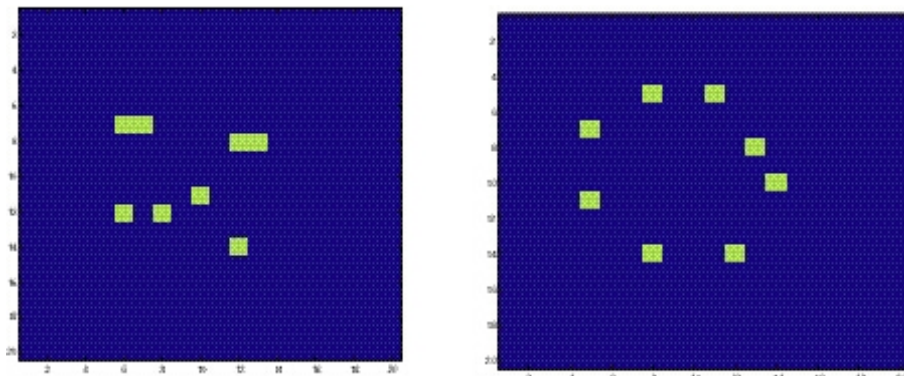


Abbildung 10: Simulation der Energieminimierung von 8 geladenen Teilchen mit 6000 Iterationen und einer inversen Temperatur von $10/kT$

μ bezeichnet den sogenannte "drift term" und σ die "volatility" oder Schwankung. ε ist eine normalverteilte Zufallsvariable.

Unter der Annahme, daß der zweite Term auf der rechten Seite 0 ist folgt die Lösung

$$S(t) = S_0 e^{\mu t} \quad (46)$$

Der erste Term beschreibt somit das Preiswachstum bei einer Rate μ . Nehmen wir nun an, daß der erste Term in Gl.(45) gleich 0 ist und somit keine Preissteigerung auftritt. Weiters definieren wir $U = \ln S$ und mit Gl.(45) folgt

$$U(t) - U(0) = \sigma \sqrt{\Delta t} \sum_{i=1}^N \varepsilon_i \quad (47)$$

Die Summe über N Zufallsvariablen ε_i ergibt nach dem Zentralen Grenzwertsatz eine Normalverteilung.

Die Änderung des Aktienpreises bezogen auf den aktuellen Preis kann nun formuliert werden

$$\frac{dS}{S} = \mu dt + \sigma \xi \sqrt{dt} \quad (48)$$

wobei ξ eine normalverteilte Zufallszahl ist. Diese kann z.B. mit Hilfe des Box Müller Verfahrens (siehe Kapitel 3.3.3) erzeugt werden.

Abbildung 11 zeigt die mögliche Entwicklung des Aktienpreises mit $\mu = 0.14$ und $\sigma = 0.2$ und einem Anfangswert von 20\$ indem Gl.(48) simuliert wurde.

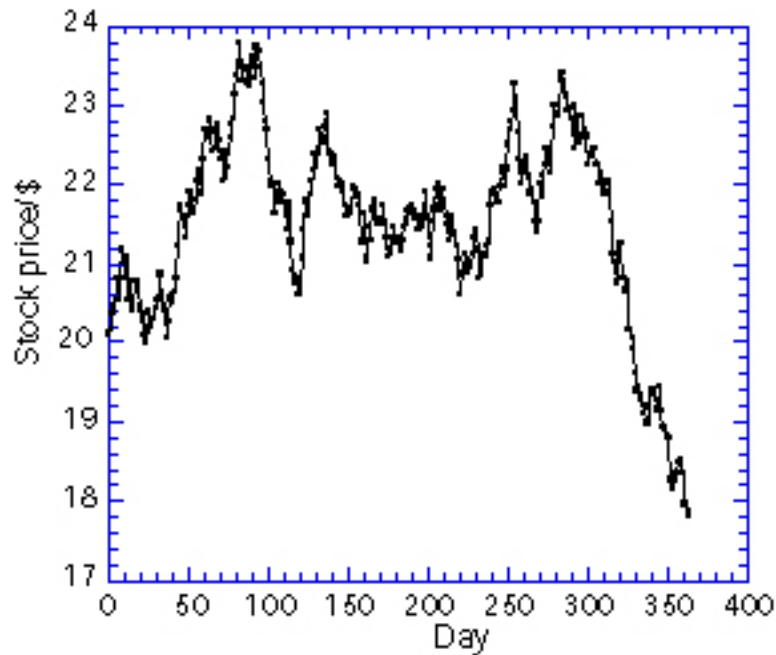


Abbildung 11: Entwicklung eines simulierten Aktienpreises

Wenn man nun über die auftretenden Preise histogrammiert erhält man [Abbildung 12](#).

Es ist aus [Abbildung 12](#) leicht ersichtlich, daß der Gewinn bzw. der Verlust bei größerer Schwankung σ (volatility) relativ groß ist. Ein Analyst kann nun durch Analyse des bisherigen Verlaufs des Aktienkurses eine Schätzung der Parameter μ und σ vornehmen. Dafür gibt es mehrere Methoden wie z.B. Maximum Likelihood oder bessere welche man in [\[11\]](#) nachlesen kann. Wenn man nun eine ganze Reihe von Simulationen von [Gl.\(48\)](#) durchführt und über diese histogrammiert, kann man die Schwankungsbreite des Preises und einen Trend abschätzen. Etwas komplexer wird die Sachlage bei der Betrachtung von Finanztiteln wie Optionen, Futures usw. die auf der Preisentwicklung von Aktien basieren.

4.2.1 Black-Sholes Gleichung

Hier nur ein kurzer Einblick in die in der Finanztheorie äußerst wichtige Black-Sholes Gleichung² zur Bestimmung der Preisentwicklung von Finanztiteln (z.B. Aktien, Optionen, Futures...). Eine europäische "call option" ermöglicht dem Besitzer einen Finanztitel zu einem bestimmten Zeitpunkt (maturity) zu einem bestimmten Preis (strike price) zu erwerben. Zu beachten ist, daß hier nur die Möglichkeit eingeräumt wird, jedoch muß

²F. Black and M. Sholes, The pricing of options and corporate liabilities, Journal of Political Economy 81, 637 (1973).

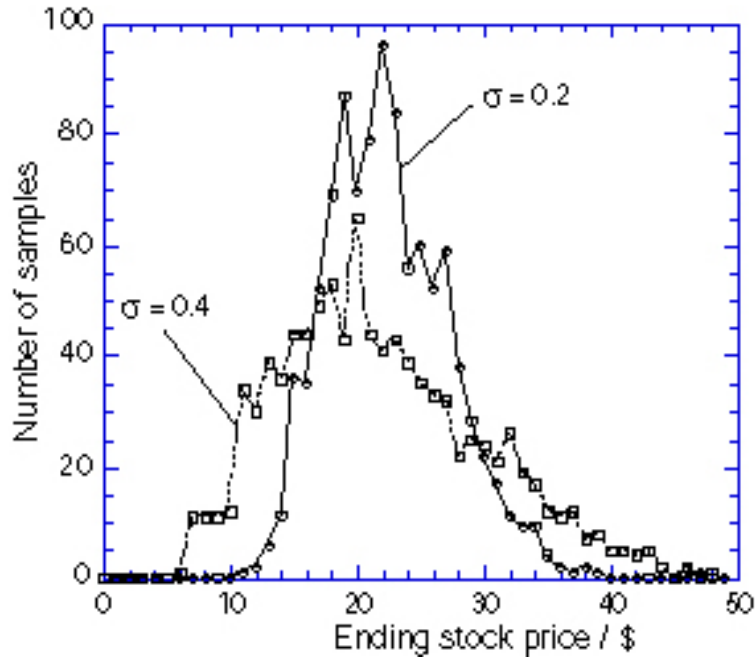


Abbildung 12: Histogramm eines simulierten Aktienpreises.

der Besitzer dieser Option sie nicht einlösen. Betrachten wir nun einen Investor der eine call option für IBM Aktien mit einem strike price von 100\$ kauft. Der aktuelle Aktienpreis liegt bei 98\$, die Laufzeit beträgt 2 Monate und der Preis der call Option ist 5\$. Falls die Aktie nach Ende der Laufzeit weniger als 100\$ wert ist, wird der Investor seine Option sicher nicht einlösen. (Es wäre wohl sinnlos eine Aktie welche weniger als 100\$ wert ist um 100\$ zu kaufen!). In diesem Fall verliert der Investor sein ursprüngliches Investitionskapital von 5\$. Falls der Aktienpreis nun z.B. bei 115\$ liegt, wird er von seinem Kaufrecht um 100\$ jedoch sicherlich Gebrauch machen. Er macht somit $15\$ - 5\$ = 10\$$ Reingewinn (net profit). Die Black Sholes Analyse macht folgende Vorraussetzungen:

- Der zugrunde liegende Aktienpreis folgt einer einfachen Diffusionsgleichung (Gl.(45))
- In einem kompetitiven Markt gibt es keine risikofreien Gewinnmöglichkeiten durch Arbitrage³.
- Die risikofreie Zinsrate am Geldmarkt ist konstant für alle risikofreien Investitionen.

Die Aufgabe eines Investment-Analysten ist nun die Zusammenstellung eines Portfolios aus Finanztiteln, welches einen möglichst guten Kompromiß zwischen Risiko und

³Kauf von Finanztiteln mit sofortigem Verkauf in anderen Märkten und erzielen von Profit durch Unterschiede in Preis bzw. Währungskurs

möglichem Gewinn darstellt. Dafür benötigt er eine Abschätzung des Wertes der Option, welche von der Preisentwicklung der ihr zugrundeliegenden Aktie abhängt. Unter Annahme, daß diese Entwicklung einem Diffusionsprozeß (siehe Gl.48) folgt, stellten F. Black und M. Sholes folgende Gleichung für den Preis f einer Option in einem Portfolio mit der Wachstumsrate r auf

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \quad (49)$$

Man schätzt nun aufgrund der vorhandenen Daten über die bisherige Entwicklung des Aktienkurses die Parameter μ und σ aus Gl.(48) ab und berechnet damit eine ganze Reihe von möglichen Entwicklungen S_T^i (Pfade) für eine Zeitdauer von T Tagen. Der zu erwartende Preis der Option ergibt sich bei einem Strike-Preis K zu

$$\langle f \rangle = \frac{1}{N} \sum_{i=1}^N e^{\mu T} \max\{S_T^i - K, 0\} \quad (50)$$

Dies ist nun ein sehr idealisiertes Modell, welches verbessert werden kann, indem der "drift term" μ während der Berechnung auch Schwankungen unterliegt, die als Diffusionsprozeß über Monte Carlo bzw. Quasi Monte Carlo Methoden modelliert werden können. Dies ginge jedoch in diesem Rahmen zu weit und kann in [10] nachgelesen werden.

Literatur

- [1] Ewald Schachinger, *Physik & Simulation von Strahlungstransport*, TU-Graz (1999) [2](#), [8](#), [14](#), [18](#)
- [2] N. Metropolis, et al., *Journal of Chem. Physics.* *21*, 1087 (1953) [18](#)
- [3] Kurt Binder, Dieter W. Heermann, *Monte Carlo Simulation in Statistical Physics*, Springer Verlag (1992)
- [4] P. K. MacKeown, *Stochastic Simulation in Physics*, Springer Verlag, Singapore (1997)
- [5] W. von der Linden, A. Prüll, *Wahrscheinlichkeitstheorie, Statistik & Datenanalyse*, TU-Graz (2001)
<http://www.itp.tu-graz.ac.at/lectures>
- [6] M. Kalos, P. Whitlock, *Monte Carlo Methods Volume 1*, Wiley & Sons (1986) [14](#), [18](#)

- [7] C. Lang, N. Pucker, *Mathematische Methoden in der Physik*, Spektrum Akademischer Verlag (1998)
- [8] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, New York (1998), ISBN 0521-35465-X [10](#), [14](#)
- [9] D.E. Knuth, *The Art of Computer Programming, Vol. 2*, Addison Wesley Publ. Co. (1981) [10](#)
- [10] Jeremy Staum, *Simulation in financial engineering*, Proceedings of the 2001 Winter Simulation Conference (2001) [24](#)
- [11] J.D. Hamilton, *Time series analysis*, Princeton University Press (1994)

[23](#)