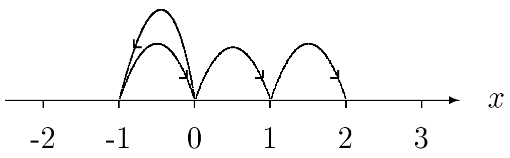


Abbildung 5.7: Ein-dimensionaler Zufallsweg: Start bei $x = 0$ in der Folge: l-r-r-r

```
    x1ave[i]=0;
    x2ave[i]=0;
  }

  for( int iw=1; iw<=nwalk; iw++ )  // perform nwalk random walks
  {
   rkps0( 234567+iw );              // initialize KPS RNG anew
   xn=0;                            // initial position is at zero
   for(int i=1; i<=nstep; i++ ){    // perform nstep steps
    if( rkps() ) xn++;              // step right
    else         xn--;              // step left
    x1ave[i] += xn;                 // averages
    x2ave[i] += xn*xn;
   }
  }
  for(int i=1; i<=nstep; i++ )      // output loop
   cout << i << ' ' << x1ave[i] << ' ' << x2ave[i] << '\n';

  return 0;
}
```

In Abbildung 5.8 (links) sind zwei verschiedene random walks dargestellt. Der Mittelwert bzw. das erste Moment der Position zur Zeit t ist:

$$\langle x_t \rangle = \frac{1}{N} \sum_{i=1}^{N} x_t(i) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{t} \gamma_{ij} \ , \tag{5.16}$$

wobei γ_{ij} die j-te Zufallsvariable im i-ten Lauf, mit $\gamma_{ij} = \pm 1$ ist. N ist die Anzahl der RW und $x_t(i)$ die Position des RW im i-ten Lauf zur Zeit t. Bildet man den Mittelwert aller Zufallswege, so ist $\langle x_t \rangle = 0$, da sich die Summe aller γ_{ij}, die ja mit gleicher Wahrscheinlichkeit -1 oder +1 sind, aufheben.

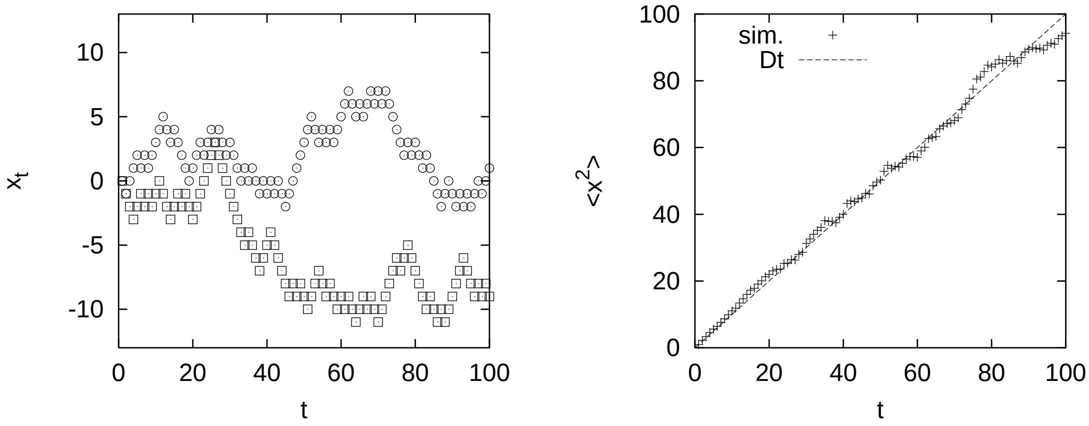


Abbildung 5.8: Position x_t von zwei verschiedenden random walks (links) und mittleres Verschiebungsquadrat $\langle x^2 \rangle$ im Mittel über 300 RW (rechts) als Funktion der Zeit.

Für die Bewegung im Raum ist der mittlere quadratische Abstand

$$\langle x_t^2 \rangle = \frac{1}{N} \sum_{n=1}^{N} x_t^2 \tag{5.17}$$

wichtig und vor allem aussagekräftiger.

In Abbildung 5.8 (rechts) ist dieser Abstand über die Zeit aufgetragen. Es ergibt sich eine Ursprungsgerade, es gilt also:

$$\langle x_t^2 \rangle = Dt \ , \tag{5.18}$$

mit der Diffusionskonstante D. Aus der Grafik ist ersichtlich, daß sie in diesem Fall $D \approx 1$ ist. Im Gegensatz zu ballistischen Teilchen (geradlinige, gleichförmige Bewegung), deren zurückgelegter Weg proportional zur Zeit ist ($r \sim t$), besitzen die RW-Teilchen ein diffusives Verhalten und ihr im Mittel zurückgelegter Weg ist proportional zur Wurzel aus der Zeit:

$$r \mathrel{\hat{=}} \sqrt{\langle x^2 \rangle} \ \sim \sqrt{t} \tag{5.19}$$

Auf den Tintentropfen bezogen, wäre die benötigte Zeit, um sich im Glas zu verteilen, bei doppelt so großem Glas viermal so hoch.

Zur analytischen Berechnung von D benutzen wir die Summendarstellung von linearen und quadratischem Abstand bei einem Lauf $N = 1$:

$$x_t = \sum_{i=1}^{t} \gamma_i \tag{5.20}$$

$$\Rightarrow x_t^2 = \left(\sum_{i=1}^{n} \gamma_i \right)^2 = \sum_{i,j=1}^{n} \gamma_i \gamma_j \tag{5.21}$$

Für $i \neq j$, ist $\gamma_i\gamma_j \pm 1$ und zwar wieder jeweils mit gleicher Wahrscheinlichkeit $p = 0.5$. Bei Addition löschen sich alle diese Terme gegenseitig aus, und es bleiben nur noch die Summanden mit $i = j$ übrig:

$$x_t^2 = \sum_{i=1}^{t} \gamma_i^2 = t \text{ , da } \gamma_i^2 = 1 \tag{5.22}$$

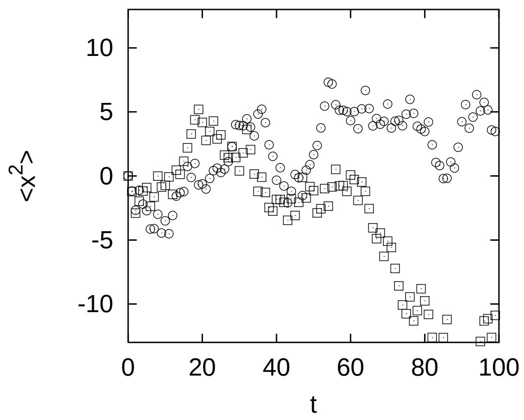
$$\Rightarrow D = \frac{x_t^2}{t} = 1 \text{ .} \tag{5.23}$$

Die ganze Bewegung kann auch dreidimensional realisiert werden. Dafür werden alle Berechnungen für x nochmals unabhängig für die beiden anderen Richtungen y und z durchgeführt. Die Summe der Quadrate wird zur quadratischen Entfernung vom Ursprung

$$r^2 = x^2 + y^2 + z^2 \tag{5.24}$$

Diese Entfernung ist immer noch proportional zur Wurzel der Zeit.

Eine andere Variationsmöglichkeit ist die Implementierung einer zufälligen Weglänge mit maximaler Sprungweite a. Anstatt einer dualen Zufallszahl für den random walk benutzen wir eine reelle Zufallszahl, beispielsweise zwischen $-a$ und a. Abbildung 5.9 (links) zeigt diesen random walk mit zufälliger Weglänge und $a = 2$. In der rechten Grafik ist das mittlere Quadrat der Entfernung aufgetragen und es ist zu erkennen, daß sie wieder proportional zur Zeit ist, aber diesmal mit kleinerem Faktor.



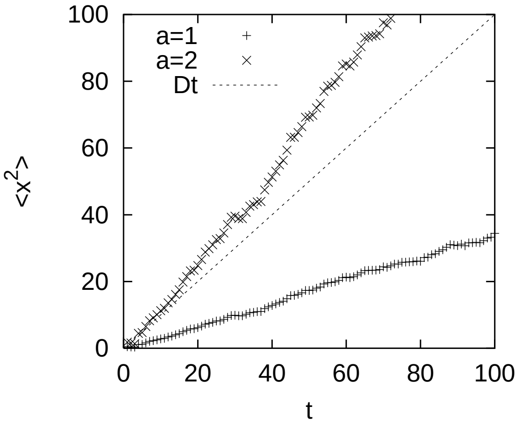


Abbildung 5.9: RW mit zufälliger Sprungweite und -richtung. Zwei Realisierungen (links) und $\langle x^2 \rangle$ im Mittel über 300 Läufe (rechts).

Trägt man zur Zeit t die Anzahl $N_t(x)$ der Teilchen (von verschiedenen RW) an einem Ort x auf, so entsteht eine Gaußsche Glockenkurve, siehe Abbildung 5.10. Das bedeutet daß $N_t(x)$ normalverteilt ist:

$$n_t(x) = \frac{N_t(x)}{\sum_x N_t(x)} = \frac{1}{\sqrt{\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{\sigma^2}} \text{ .} \tag{5.25}$$

Hierbei ist der Mittelwert $\mu = \langle x_t \rangle = 0$. Für die Standardabweichung σ, also die Breite der Normalverteilung, gilt:

$$\sigma^2 = 2Dt = 2\langle x_t^2 \rangle \tag{5.26}$$

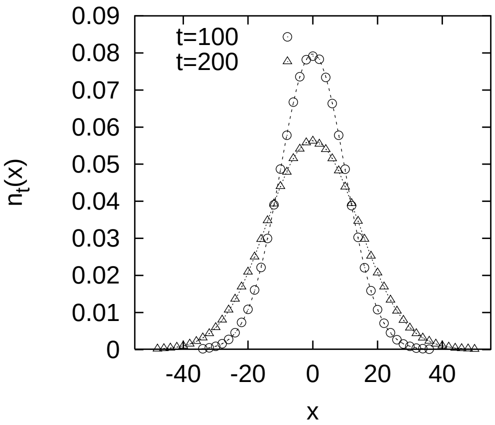


Abbildung 5.10: Gaußsche Glockenkurve der Verteilungen $2n_t(x)$. Symbole sind Daten der Simulation in Abb. 5.8, und die Linien stellen die entsprechenden Funktionen in Gleichung 5.25 dar.

Da die diskrete Simulation zur Gewinnung der Daten verwendet wurde, können bei geraden Zeiten t auch nur gerade Orte x besetzt sein. Zur Zeit $t = 0$ hat die Funktion $n_t(x)$ nur bei $x = 0$ den Wert $n_t(0) = 1$, überall sonst ist verschwindet die Funktion. Im Laufe der Zeit läuft die Funktion auseinander und wird flacher, die Fläche unter $n_t(x)$ beleibt durch die Definition erhalten.

Das mittlere Abstandquadrat eines RW hat also die Bedeutung der Breite der Verteilungsfunktion, die ebenso proportional zur Zeit ist:

$$\sigma \sim t \ . \tag{5.27}$$

Ein RW entfernt sich also mit der Zeit immer weiter von seinem Ausgangspunkt, kehrt jedoch immer wieder zurück, so daß die mittlere Position immer der Ausgangspunkt ist. Der RW unternimmt also immer längere Reisen zur Erkundung des Raumes.

5.5 Random Walk und Diffusion

Wie schon erwähnt besteht ein enger Zusammenhang zwischen einem RW und der Diffusion. Die Verbindung zwischen beiden Prozessen soll kurz genauer erläutert werden.

Ausgehend von einem random walk kann die zeitliche Änderung der Aufenthaltswahrscheinlichkeit $p(x,t)$, den RW zur Zeit t am Ort x zu finden, durch die sog. *Master-Gleichung*

$$\Delta p(x,t+\Delta t) = p(x+\Delta x,t)W_- + p(x-\Delta x,t)W_+ - p(x+\Delta x,t)(W_+ + W_-) \quad (5.28)$$

ausgedrückt werden. Dabei sind W_+ und W_- jeweils die Wahrscheinlichkeiten, daß ein Teilchen in der Zeit Δt um die Entfernung $\pm\Delta x$ springt. Im Falle eines einfachen RW sind $W_+ = W_- = 1/2$.

Im Grenzfall $\Delta t \to 0$ und $\Delta x \to 0$ sind

$$\partial p/\partial t = \Delta p(x,t+\Delta t)/\Delta t \quad (5.29)$$

und

$$\partial^2 p/\partial x^2 = [p(x+\Delta x,t) - 2p(x+\Delta x,t) + p(x-\Delta x,t)]/\Delta x^2 \ , \quad (5.30)$$

woraus direkt die Diffusionsgleichung

$$\frac{\partial p}{\partial t} = \mathcal{D}\frac{\partial^2 p}{\partial x^2} \quad (5.31)$$

mit der Diffusionskonstante

$$\mathcal{D} = \frac{\Delta x^2}{2\Delta t} \quad (5.32)$$

folgt. Durch Einsetzen verifiziert man leicht die Lösung

$$p(x,t) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \ , \quad (5.33)$$

mit

$$\sigma = \sqrt{2\mathcal{D}t} \ . \quad (5.34)$$

Der Unterschied zwischen Gleichung 5.33 und Gleichung 5.25 liegt nur in der unterschiedlichen Definition der Diffusionskonstanten $D = 2\mathcal{D}$, beide Gleichungen sind normiert, d.h. $\int p(x)dx = 1$. Die Definition in Glg. 5.33 ist allerdings leichter zu rekapitulieren, da sie genau so auf dem 10 DM Schein abgedruckt ist.

Anmerkungen

Das statistische Verhalten eines RW mit diskreter Zeit und diskretem Ort ist (bei hinreichend großer Schrittanzahl) gleich dem Verhalten eines RW mit kontinuierlicher Schrittweite, kontinuierlicher Wartezeitverteilung, oder beidem – solange die Diffusionskonstanten gleich sind.

Verbietet man die Orte, die ein RW bereits besucht hat, so erhält man einen self-avoiding RW (SAW), ein einfaches Modell für eine Polymerkette.

5.6 Fehler, Varianz und lineare Regression

Um den Fehlerbereich einer Meßreihe abzuschätzen, wird die Varianz (=Schwankung) der Meßreihe benutzt, die die Abweichung vom Mittelwert

$$\langle x \rangle_N = \frac{1}{N} \sum_{n=1}^{N} x_n \tag{5.35}$$

beschreibt. Für die Varianz σ gilt:

$$\begin{aligned} \sigma_N^2 &= \langle (x - \langle x \rangle_N)^2 \rangle_N \\ &= \langle x^2 - 2x\langle x \rangle_N + \langle x \rangle_N^2 \rangle_N \\ &= \langle x^2 \rangle_N - 2\langle x \rangle_N \langle x \rangle_N + \langle x \rangle_N^2 \\ &= \langle x^2 \rangle_N - \langle x \rangle_N^2 \end{aligned} \tag{5.36}$$

5.6.1 Fehler

Der statistische Fehler des Mittelwerts wird typischerweise definiert als $\Delta_N = \sigma_N / \sqrt{N-1}$. Diese Wahl soll im Folgenden begründet werden. Geschickt ist eine Wahl, die unabhängig von der Meßprozedur ist. Wird dieselbe Meßreihe auf mehrere kleine Meßreihen aufgeteilt, so sollte der Fehler bei gleichen Werten auch konstant bleiben. Wir unterteilen unsere Meßreihe x_n mit N Punkten in k Einheiten der Länge m. Es gilt also $N = km$ und

$$x_n = x_{ij} \text{ mit } n = (i-1)m + j; \;\; i = 1, \cdots, k; \;\; j = 1, \cdots, m \tag{5.37}$$

Desweiteren haben die Untermeßreihen die Mittelwerte:

$$M_i = \frac{1}{m} \sum_{j=1}^{m} x_{ij} \tag{5.38}$$

Der gesamte Mittelwert ist dann:

$$\langle x \rangle_N = \langle M_i \rangle_k = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{m} \sum_{j=1}^{m} x_{ij} \tag{5.39}$$

Würde man die Varianz der M_i als Fehler nehmen, so bekäme man:

$$\begin{aligned}
\sigma_k^2 &= \langle (M_i - \langle M_i \rangle_k)^2 \rangle_k \\
&= \tfrac{1}{k} \textstyle\sum_{i=1}^{k} (M_i - \langle x \rangle_N)^2 \\
&= \tfrac{1}{k} \textstyle\sum_{i=1}^{k} \tfrac{1}{m^2} \textstyle\sum_{j,n=1}^{m} (x_{ij} - \langle x \rangle_N)(x_{in} - \langle x \rangle_N) \\
&= \tfrac{1}{k} \textstyle\sum_{i=1}^{k} \tfrac{1}{m^2} \textstyle\sum_{j=1}^{m} (x_{ij} - \langle x \rangle_N)^2 \\
&= \tfrac{1}{m} \tfrac{1}{N} \textstyle\sum_{i=1}^{k} \textstyle\sum_{i=1}^{m} (x_{ij} - \langle x \rangle_N)^2 \\
&= \tfrac{1}{m} \sigma_N^2
\end{aligned} \tag{5.40}$$

Die Vereinfachung der Doppelsumme gelingt durch die Auslöschung bei fluktuierenden Vorzeichen, wenn die Zufallsvariablen x_{ij} symmetrisch verteilt sind. Das Ergebnis zeigt, daß die Varianz von der Aufteilung der Meßreihe abhängt und somit nicht direkt für eine Fehlerberechnung verwendet werden darf. Aufteilungsunabhängig ist dagegen die Definition

$$\tilde{\Delta}_N = \frac{\sigma_N}{\sqrt{N}} \ , \tag{5.41}$$

wie durch die folgende Rechnung gezeigt wird:

$$\tilde{\Delta}_k^2 = \frac{\sigma_k^2}{k} = \frac{\sigma_N^2}{km} = \frac{\sigma_N^2}{N} = \tilde{\Delta}_N^2 \tag{5.42}$$

Die endgültige Definition für den Fehler lautet:

$$\Delta_N = \frac{\sigma}{\sqrt{N-1}} \ , \tag{5.43}$$

da bei $N = 1$ Messungen keine Aussage über den Fehler getroffen werden kann.

Sind die Meßpunkte x_n die Positionen eines RW, so kennt man die Varianz und damit den zu erwartenden Fehler in der Schätzung der mittleren Position des RW.

5.6.2 Lineare Regression

Gegeben seien N Datenpunkte (x_n, y_n) mit $n = 1, \cdots, N$, die in etwa auf einer Geraden liegen. Durch lineare Regression kann man die Steigung m und den Achsenabschnitt b der Geraden $y = mx + b$ bestimmen.

Parallel zur y-Achse ist der Abstand eines Punktes zu dieser Geraden

$$d_n = y_n - (mx_n + b) \ . \tag{5.44}$$

Gesucht ist nun jene Gerade, die die Summe der Abstandsquadrate

$$S = \frac{1}{N} \sum_{n=1}^{N} d_n^2 \tag{5.45}$$

minimiert. Sieht man die d_n als Meßwerte an, so muß die Gerade gefunden werden, die zur kleinsten Varianz der d_i führt. Das Minimum von S findet man indem man nach den beiden Unbekannten m und b ableitet, und die Nullstellen der Ableitungen identifiziert. Die Ableitungen von S sind:

$$\frac{\partial S}{\partial m} = \frac{-2}{N} \sum_{n=1}^{N} x_n (y_n - m x_n - b) = 0 \tag{5.46}$$

und

$$\frac{\partial S}{\partial b} = \frac{-2}{N} \sum_{n=1}^{N} (y_n - m x_n - b) = 0 \ . \tag{5.47}$$

Die Gleichungen 5.46 und 5.47 lassen sich unter Verwendung der Definitionen im vorigen Abschnitt vereinfachen:

$$\langle xy \rangle_N = m \langle x^2 \rangle_N + b \langle x \rangle_N \tag{5.48}$$

und

$$b = \langle y \rangle_N - m \langle x \rangle_N \ . \tag{5.49}$$

Durch Einsetzen von Glg. 5.49 in Glg. 5.48 ergibt sich durch Auflösen nach m schließlich die Steigung (und damit b in Glg. 5.49) als Funktion der verschiedenen Mittelwerte

$$m = \frac{\langle xy \rangle_N - \langle x \rangle_N \langle y \rangle_N}{\langle x^2 \rangle_N - \langle x \rangle_N^2} \ . \tag{5.50}$$

Vor der linearen Regression sollte amn allerdings eine Sichtprüfung der Daten durchführen, um festzustellen, ob die Daten wenigstens annähernd auf einer Geraden liegen. Ist dies nicht der Fall, so macht die "lineare" Regression keinen Sinn und man sollte ein anderes Anpassungsverfahren anwenden. Hier seien lediglich die "Spline-" oder die "Padé-Approximation" als Alternative Verfahren erwähnt.

5.7 Weitere Anwendungsbeispiele

Nach Zellularautomaten und Random Walks sollen nun etwas interessantere Anwendungsmöglichkeiten von Zufallszahlen in der numerischen Physik vorgestellt werden. Perkolation, Epidemie-Modelle und Diffusionsbestimmte Aggregation sind die im Folgenden genauer diskutierten Beispiele.