

Irrflüge auf dem Computer

1 Versuchsplatz

- Würfel / Münze
- Software RDW (Computerexperiment)
- Excel (Auswertung)

1.1 Inhalt des Versuchs

- Ablauf eines Computerexperiments
- Zufallszahlen, Gaußverteilung
- Statistische Auswertung:
Histogramme, Mittelwerte,
Standardabweichung, mittlerer Fehler des Mittelwerts
- Simulation des Diffusionspfades eines Teilchens



2 Allgemeines zum Versuch

Was ist ein Computer-Experiment und wofür ist es gut? Im Computer-Experiment untersucht man das Verhalten eines "Modellsystems". Ein solches Modellsystem bildet reale Systeme idealisiert mittels eines Algorithmus ab. Dabei stellt es beispielsweise die Positionen und Geschwindigkeiten von Molekülen, Bindungswinkel einer Polymerkette oder Dipolmomente einer Vielzahl von kleinen magnetischen Zentren dar. In der Realität wie im Modell ändern sich die betrachteten Größen mit der Zeit. Dafür sind unterschiedliche Einflussgrößen verantwortlich, wie z.B. gegenseitige Wechselwirkungen von Teilchen. So könnten sich z. B. viele kleine magnetische Dipole parallel zueinander ausrichten und einen Ferromagneten bilden. In der Realität bewegen sich mikroskopische Objekte durch thermische Anregung und folgen damit einem zufälligen Bewegungsmuster. Diese zufällige Bewegung ist die Triebkraft der Diffusion. Um diese zufällige Bewegung durch einen Computeralgorithmus darzustellen, wird auf Zufallszahlen-Generatoren (s. Anhang 1) zurückgegriffen.

Bei der Modellbetrachtung ist es wichtig, dass ein betrachtetes Experiment so häufig durchgeführt wird, bis die Mittelwertbildung sinnvoll ist (z. B. einer kinetischen Energie, einer Polymerknäuelgröße, oder einer Magnetisierung). Dies ist der Fall, wenn man bei der Wiederholung eines Computerexperimentes, bei dem Zufallszahlen genutzt werden, zwar unterschiedliche Einzelwerte erhält, diese aber zu nahezu identischen Mittelwerten führen. Beispielsweise würde beim Würfeln mit einem (idealen) 6-seitigen Würfel nach nur 10 Würfeln der Mittelwert stark variieren; bei einer genügend großen Anzahl Würfeln wird der Mittelwert aber dem Wert 3,5 sehr nahe kommen. Diese Zahlenwerte sind nicht durch systematische Fehler

verfälscht, wie es beim realen Experiment der Fall wäre: Bei einem realen (nicht-idealen) Würfel könnte unter Umständen eine Augenzahl statistisch häufiger auftreten als andere, so dass der Mittelwert auch bei vielen Würfeln vom erwarteten Wert 3,5 abweicht. Die Fehlerquellen für Simulationen sind unter anderem abhängig vom genutzten Algorithmus und den Zufallszahlengeneratoren, daher sind sie im Allgemeinen gut bekannt und kontrollierbar. Der statistische Fehler tritt sowohl bei der praktischen Durchführung als auch bei der Simulation auf. Durch eine Erhöhung der Anzahl der Messwerte kann der statistische Fehler verringert werden. Für das betrachtete Computer-Experiment heißt das, dass mit Erhöhung der Rechenzeit der statistische Fehler beliebig klein gerechnet werden kann.

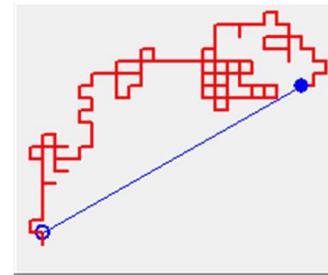


Abb. 1 - „Irrflug“ auf einem Gitter – hier in zwei Dimensionen. Die blaue Linie verbindet Pfadanzfang und -ende.

Das in diesem Versuch durchgeführte Computerexperiment simuliert eine „Irrflug“ (engl. "random walk"). Der Irrflug entspricht im vorliegenden Fall dem Diffusionspfad eines Teilchens: Diesen Diffusionspfad kann man sich dabei so vorstellen, dass das Teilchen nacheinander durch wiederholte Stöße mit Nachbarpartikeln um Strecken jeweils zufälliger Längen in jeweils zufällige Richtungen verschoben wird. Unterschiedliche Schrittlängen kommen dabei durch statistische Verteilung der Energien beim Stoß (Boltzmann-Verteilung) und auch durch die statistisch unterschiedliche Teilchendichte (zumindest in Flüssigkeiten in insbesondere in Gasen) in der unmittelbaren Umgebung des Teilchens zustande. Die zufälligen Richtungen ergeben sich zum einen daraus, dass die Richtung, aus der ein stoßendes Nachbarpartikel auftritt, zufällig ist und dass weiterhin der Stoß nur sehr selten ein zentraler Stoß ist, bei dem der Verbindungsvektor der Teilchen in Richtung des Vektors ihrer Relativgeschwindigkeiten liegt. Das Computermodell des diffusiven Irrflugs ist dagegen wesentlich einfacher: Das Teilchen bewegt sich auf einem d -dimensionalen Gitter, die möglichen Teilchenpositionen haben also feste Abstände und die Achsen des Gitters haben feste Winkel. Im einfachsten Fall haben benachbarte Gitterpunkte entlang jeder Achse gleiche Abstände und die Winkel in einem d -dimensionalen Gitter mit $d > 1$ sind rechte Winkel. Jeder betrachtete Schritt kann in eine zufällige Richtung auf dem Gitter erfolgen. In einem eindimensionalen „Gitter“ gibt es dabei nur zwei Richtungen („vorwärts“ und „rückwärts“), im zweidimensionalen Gitter vier, im dreidimensionalen sechs usw., siehe dazu Abb. 1. Die Schrittlänge λ ist zunächst stets der Abstand zwischen direkt benachbarten Gitterpunkten. Nach einer im Experiment vorgegebenen Anzahl N an Schritten ist der Endpunkt des Irrflugs erreicht. Der Abstand zwischen Start- und Endpunkt wird als Verschiebung R bezeichnet:

$$R = r_{i=N} - r_{i=0} \quad (1)$$

Hier ist $r_{i=0}$ die Startposition und $r_{i=N}$ die Position nach N Schritten.

Ist die Anzahl N der Schritte pro Irrflug klein und ist die Anzahl N_f gemittelter Irrflüge ebenfalls klein, so werden – aufgrund der zufälligen Natur der Einzelschritte – unterschiedliche Experimente unterschiedliche Ergebnisse liefern. Wird die Anzahl der Schritte N pro Irrflug jedoch sehr groß und wird weiterhin über eine sehr große Anzahl N_f an Irrflügen gemittelt, so nähert sich der Mittelwert $\langle R \rangle$ immer mehr dem Wert Null – unabhängig von der gewählten Schrittweite und der tatsächlichen Anzahl der Schritte. Betrachtet man dagegen das „mittlere Verschiebungsquadrat“ $\langle R^2 \rangle$, findet man einen allgemeingültigen Zusammenhang zwischen dem mittleren Verschiebungsquadrat, der Schrittweite und der Anzahl der Schritte:

$$\langle R^2 \rangle = \langle (r_{i=N} - r_{i=0})^2 \rangle = N \cdot \lambda^2 \quad (2)$$

Für die Beschreibung der Diffusionseigenschaften von Teilchen auf molekularer Skala ist die Annahme einer sehr großen Schrittzahl und die Betrachtung des Mittelwerts über die Irrflüge sehr vieler Teilchen sinnvoll und passend: So erfährt beispielsweise ein einzelnes Teilchen in einem idealen Gas unter Standardbedingungen ca. 10^{10} Stöße pro Sekunde (es führt also ebenso viele Schritte eines Irrflugs aus) und in einem Kubikzentimeter befinden sich ca. $2,4 \cdot 10^{20}$ Teilchen, über deren Irrflüge man mitteln könnte, um die allgemeinen Diffusionseigenschaften dieser Gasteilchen zu beschreiben. In Flüssigkeiten sind die Teilchendichte und die Häufigkeit der Stöße i. d. R. noch um einige Größenordnungen höher.

Für diese statistische Betrachtung vieler, sehr langer diffusiver Irrflüge kann man die mikroskopischen Eigenschaften, also die genaue Schrittfolge, ihre Längen und Richtungen, ignorieren. Man interessiert sich nur noch für das Verschiebungsquadrat, gemittelt über viele Irrflüge. Es zeigt sich, dass man zu jedem Irrflug mit gegebener Schrittlänge und gegebener Schrittzahl weitere Irrflüge finden kann, die bei längerer Schrittweite und geringerer Schrittzahl (bzw. kleinerer Schrittweite und größerer Schrittzahl) die gleichen statistischen Eigenschaften aufweisen, sprich ein ebenso großes mittleres Verschiebungsquadrat ergeben. Diese Eigenschaft des Diffusionsprozesses bezeichnet man als „Skaleninvarianz“. Solange die Skaleninvarianz gilt, folgt das mittlere Verschiebungsquadrat einem Gesetz der Form:

$$\langle R^2 \rangle = N^{2\nu} \cdot \lambda^2 \quad (3)$$

Weiterhin bezeichnet N die Anzahl der Schritte und λ die Schrittweite. Aus der Dimensionsbetrachtung wird klar, dass die Wurzel des mittleren Verschiebungsquadrats, die „quadratisch gemittelte Verschiebung“ $R_{\text{rms}} = \sqrt{\langle R^2 \rangle}$ proportional zur Schrittweite λ sein muss. Der Index „rms“ steht dabei für "root (of the) mean square". Da es sich bei beiden Größen, R_{rms} und λ , um Strecken (in z. B. Metern) handelt, wäre ein anderer Zusammenhang (mit einem anderen Exponenten für λ) physikalisch nicht sinnvoll. Die Abhängigkeit der quadratisch gemittelten Verschiebung R_{rms} von der Schrittzahl N ist jedoch komplizierter. Der Zusammenhang zwischen der quadratisch gemittelten Verschiebung und der Schrittzahl wird durch den

Skalenexponenten ν (griechisches "nü") beschrieben. Dieser nimmt für einen klassischen Diffusionsprozess einen Wert von $1/2$ an (es folgt wieder Gleichung 1).

Die statistische Betrachtung von Mittelwerten über viel Irrflüge mit jeweils großer Schrittzahl führt außerdem dazu, dass auch die zur Vereinfachung angenommene Bewegung der Teilchen auf einem orthogonalen Gitter statistisch äquivalent wird mit einer Bewegung, bei der unterschiedliche Schrittweiten und unterschiedliche Winkel zwischen zwei Teilstrecken des Irrflugs auftreten, solange man die Schrittzahl entsprechend passend wählt.

Unabhängig von der genauen Geometrie und Schrittweite ergeben sich für die Verschiebungsquadrate vieler Irrflüge *Normalverteilungen*. Diese Tatsache folgt aus dem *zentralen Grenzwertsatz* der Wahrscheinlichkeitsrechnung. Dieser besagt, dass die Verteilungsfunktion, die sich durch das Aufsummieren von unabhängigen und identisch verteilten Zufallszahlen $x_1, x_2, x_3, \dots, x_N$ ergibt, mit einer wachsenden Anzahl von diesen Zufallszahlen (einer wachsenden Anzahl von N) gegen eine Normalverteilung (auch: Gaußverteilung) strebt. So strebt auch die Binomialverteilung bei ausreichend großem N gegen eine Normalverteilung. Dieser Grenzwertsatz erklärt auch, warum so viele empirische Verteilungen, mit vielen unabhängigen Einflussgrößen, einer Normalverteilung ähneln. Siehe dazu auch Abschnitt 1.3 der Versuchsdurchführung.

An Stelle der festen Schrittweite λ des Irrflugs auf dem Gitter in der Computersimulation tritt dann für die Diffusion realer Teilchen eine (statistische) *mittlere* Schrittweite, die im Fall des idealen Gases der mittleren freien Weglänge entspricht. Die Gültigkeit von Gleichung 1 und 2 bleibt erhalten, allerdings immer noch nur unter der Voraussetzung, dass die Schrittzahl N ausreichend groß ist und die Anzahl der gemittelten Irrflüge gegen unendlich geht ($N_f \rightarrow \infty$).

Dass sich eine Normalverteilung ergibt, folgt aus der Wahrscheinlichkeit, mit der ein Teilchen nach N Schritten einer zufälligen, eindimensionalen diffusiven Bewegung eine Verschiebung von $R = n \cdot \lambda$ gegenüber seiner Ausgangsposition bei $r = 0$ erfährt. Dabei ist $n = (N_+ - N_-)$ mit N_+ und N_- der jeweiligen Anzahl an Schritten in positive bzw. negative Richtung. Für die Wahrscheinlichkeit P erhält man als Näherung für große N (vgl. auch Atkins, 4. Aufl., Kap. 21.3.4)

$$P = \sqrt{\frac{2}{\pi N}} \cdot e^{-\frac{n^2}{2N}} \quad (4)$$

Diese Gleichung hat bereits die Form einer Normalverteilung („Gaußkurve“, s. Anhang 2).

Den Zusammenhang mit der *Diffusion* erhält man durch Vergleich mit der Lösung des 2. Fickschen Gesetzes für die entsprechenden Nebenbedingungen, nämlich einer Stoffmenge n_0 von Teilchen, welche aus einer ebenen Fläche A bei $r = 0$ (eindimensional senkrecht zu dieser Fläche) in ein unendliches Medium diffundieren:

$$c(r, t) = \frac{n_0}{A\sqrt{\pi Dt}} \cdot e^{-\frac{r^2}{4Dt}} \quad (5)$$

Für den Vergleich muss man sich nun klarmachen, dass $r = R = n \cdot \lambda$ ist und dass die Zeit t , die für $N = N_+ + N_-$ Schritte benötigt wird, $t = N \cdot \tau$ ist, wenn die (mittlere) Zeit pro Schritt als τ bezeichnet wird. Dann kann man den Exponenten in Gleichung 4 auch schreiben als

$$\frac{n^2}{2N} = \frac{r^2 \tau}{2\lambda^2 t} \quad (6)$$

Der Vergleich der rechten Seite mit dem Exponenten in Gleichung 5 ergibt die *Einstein-Smoluchowski-Gleichung*:

$$D = \frac{\lambda^2}{2\tau} \quad (7)$$

Das eigentliche „Experiment“ besteht nun im Erzeugen von Irrflügen, die einen Diffusionsprozess simulieren: Als Ortskoordinate des Startpunkts wird der Koordinatenursprung definiert, für einen Diffusionsprozess in 3 Dimensionen also $(0, 0, 0)$. Durch den Zufallsgenerator werden (ganzzahlige) Zufallszahlen im Bereich von 1 bis 6 erzeugt. Jeder dieser Zufallszahlen wird ein positiver oder negativer Wert in eine der drei Raumrichtungen zugewiesen (z. B. 1: $r_x \rightarrow r_x + 1$; 2: $r_x \rightarrow r_x - 1$; 3: $r_y \rightarrow r_y + 1$ usw.), in die der Irrflug im folgenden Schritt fortgeführt wird. Für den gesamten Irrflug wird dieser Prozess N mal durchgeführt. Für den einzelnen Irrflug wird nach Abschluss die Verschiebung nach $R = (r_x^2 + r_y^2 + r_z^2)^{1/2}$ bestimmt.

Als Ergebnis der Simulation erhalten wir einen Diffusionspfad; dieser kann (in zwei Dimensionen) z. B. so aussehen wie in Abb. 1 rot dargestellt. Anfangs- (leerer blauer Kreis) und Endpunkt (gefüllter blauer Kreis) des Diffusionspfades sind durch eine blaue Linie verbunden, deren Länge die Verschiebung R darstellt. Für das in Abb. 1 gegebene Beispiel ist für Schritte mit Richtungswechsel ein Winkel von 90° , 180° oder 270° möglich. Würde dem Irrflug stattdessen ein hexagonales Gitter zu Grunde liegen, mit Winkeln von 60° , 120° , 180° , 240° oder 300° , würden sich bei gleicher Schrittzahl und Schrittweite andere Verschiebungen ergeben. Dies wäre sowohl im zwei- wie im dreidimensionalen Raum gültig. Betrachtet man aber den Grenzfall eines unendlich langen Irrfluges ($N \rightarrow \infty$), ist die Geometrie des Gitters wegen des zentralen Grenzwertsatzes (s. o.) nicht mehr ausschlaggebend.

Da die Gleichungen 2 und 3 exakt nur für $N_f \rightarrow \infty$ gelten, wird man im Experiment bei begrenzter Anzahl gemittelter Verschiebungen Abweichungen erwarten. Um die Gültigkeit von Gleichung 2 für unterschiedliche (vorgegebene) N zu überprüfen, trägt man die aus N_f Irrflügen gleicher Schrittzahl N erhaltenen Mittelwerte $\langle R^2 \rangle$ gegen diese Schrittzahl auf. Es sollte

sich für verschiedene N (näherungsweise) eine Gerade ergeben, wobei der lineare Zusammenhang umso besser erfüllt sein sollte, je größer N_f ist. Mit Gleichung 3 lässt sich neben der Schrittweite λ auch der Skalenexponent ν bestimmen. Dies geht am einfachsten, in dem man die Gleichung durch Logarithmieren linearisiert:

$$\ln\left(\sqrt{\langle R^2 \rangle}\right) = \ln R_{\text{rms}} = \ln(\lambda) + \nu \cdot \ln(N) \quad (8)$$

Es ergibt sich wieder eine Geradengleichung mit der Steigung ν und dem Achsenabschnitt $\ln(\lambda)$. Es spielt für die Bestimmung des Exponenten ν keine Rolle, zu welcher Basis man den Logarithmus wählt. Ebenso gut kann man den dekadischen Logarithmus \lg bilden:

$$\lg\left(\sqrt{\langle R^2 \rangle}\right) = \lg R_{\text{rms}} = \lg(\lambda) + \nu \cdot \lg(N) \quad (9)$$

Auch von dieser Gerade werden die logarithmierten Mittelwerte wegen der begrenzten Anzahl gemittelter Irrflüge pro Schrittzahl abweichen. Trotzdem findet man den richtigen Exponenten mit guter Genauigkeit, wenn man durch die erhaltene „Punktwolke“ eine Ausgleichs- oder Regressionsgerade legt. Im Fall eines einfachen linearen Zusammenhangs kann die „einfache lineare Regression“ verwendet werden, mit der die Parameter ν und $\log(\lambda)$ direkt berechnet werden können. Ein linearer Zusammenhang ist daher nicht nur für die Beurteilung mit dem Auge optimal, sondern erleichtert auch die quantitative Bestimmung der gesuchten Parameter.

Bei nicht-linearen Zusammenhängen lässt sich das Verfahren nicht anwenden. Das Bestimmen der Parameter einer nicht-linearen Gleichung, die einen durch statistische Fehler „verrauschten“ Verlauf experimenteller Datenpunkte optimal beschreibt, bezeichnet man oft als „fitten“, die mit den erhaltenen „Fitparametern“ berechnete Kurve als „Fit“: Die Summe der Fehlerquadrate $\Sigma(y-y_{\text{fit}})^2$ oder der Güteparameter χ^2 („chi-quadrat“)

$$\chi^2 = \frac{\Sigma(y-y_{\text{fit}})^2}{y_{\text{fit}}} \quad (10)$$

soll durch systematische Variation (Anpassung, engl. „fit“) der Parameter minimiert werden. Die Algorithmen, mit denen diese Fehlerquadrat-Summe möglichst effizient und in kurzer Zeit minimiert werden, sind dabei unterschiedlich.

Neben einem diffusiven Prozess kann auch die Gestalt eines Polymerknäuels als Resultat eines Irrflugs beschrieben werden. Während bei der Betrachtung der Diffusion das diffundierende Teilchen auf seinem Pfad auch an einen Platz zurückkehren kann, an dem es bereits war, kann ein Segment der Polymerkette nicht dort sein, wo bereits ein anderes ist. Deswegen muss man bei der Simulation eines Polymerknäuels in der Regel einen Irrflug **mit** Selbstvermeidung

simulieren. Bei der Diffusion wird ein Irrflug **ohne** Selbstvermeidung (also *mit* Selbstüberschneidung des Pfads) betrachtet. Die Unterschiede zwischen einem Irrflug mit und ohne Selbstvermeidung sind in Abb. 2 schematisch dargestellt. Das mittlere Verschiebungsquadrat des Diffusionsprozesses entspricht dem quadratisch gemittelten End-zu-End-Abstand in der Betrachtung der Polymerknäule. Während der Skalensexponent unabhängig von der Dimension für die Diffusion ohne Selbstvermeidung $\nu = 1/2$ ist, weicht er für die Diffusion **mit** Selbstvermeidung für $d \leq 3$ von diesem Wert ab. Erst für vierdimensionale Diffusion erhalte man wieder $\nu = 1/2$.

3 Orientieren Sie sich über

- Transportphänomene
- Grundlagen der Diffusion, Ficksche Gesetze
- statistische Betrachtung der Diffusion

4 Versuchsdurchführung

Hinweis: Die Aufgabennummer entspricht dem zugehörigen Button im Programm.

4.1 Selbstüberschneidende Irrflüge in einer Dimension

Aufgaben:

Erzeugen Sie mit „Zufallszahlen“, die Sie durch einen Münzwurf erhalten, einen eindimensionalen Irrflug mit 10 Schritten der Schrittweite 1. Die eine Seite der Münze wird einem Schritt +1 zugeordnet, die andere einem Schritt -1. Berechnen Sie die Verschiebung R und das Verschiebungsquadrat R^2 . Da das Erzeugen dieser Zufallszahlen und das Berechnen der Verschiebungen mühselig und zeitaufwendig ist, überlassen wir die Erzeugung der Irrflüge im Weiteren dem Computer.

1.0 Erzeugung von Irrflügen mit dem Computer – 1D-Irrflug

Zunächst beschränken wir uns weiterhin auf eine Dimension. Wir erzeugen Zufallszahlen mit dem Computer, wobei wir nur die beiden Werte 1 und -1 („rechts“ und „links“) zulassen. Wir bilden einige Male die Summen aus jeweils 10 solcher Zufallszahlen (wie beim Münzwurf). Jede dieser Summen entspricht der Verschiebung nach einem „Irrflug“ mit „10 Schritten“. Danach lassen wir 1000 Zufallszahlen generieren („1000 Schritte“). Dieser Teil dient nur der Illustration und dem Vergleich mit dem

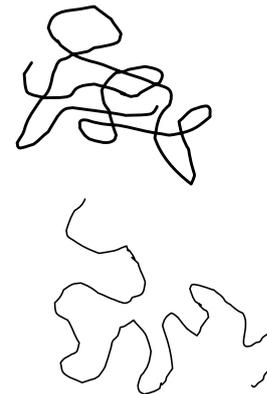


Abb. 2 - Pfade ohne Selbstvermeidung (oben) und mit Selbstvermeidung (unten) in zwei Dimensionen

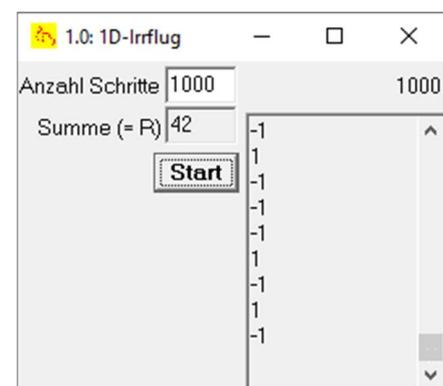


Abb. 3 - Dialogfenster 1.0 1D-Irrflug

Münzwurf, es ist keine Auswertung erforderlich (s. Abb. 3). Es wird deutlich, dass dadurch eine große Anzahl von „Münzwürfen“ in kurzer Zeit simuliert werden kann.

1.1 Gaußverteilungen, Abhängigkeit des mittleren Verschiebungsquadrates von der Anzahl der Schritte

1.1.1 Erzeugen Sie zunächst 100, dann 1000, und schließlich 10000 solcher Flüge (Anzahl der Schritte $N = 1000$, s. Abb. 4) und stellen Sie die Verteilungen der erhaltenen Summen jeweils in einem Histogramm dar. Überzeugen Sie sich, dass die erhaltenen Verteilungen für große Anzahlen der Irrflüge einer Gaußverteilung immer ähnlicher werden (zur Gaußverteilung s. Anhang 3). Berechnen Sie dazu aus Häufigkeiten jeweils die Wahrscheinlichkeitsdichten.

Hinweis: Zur Erstellung des Histogramms folgen Sie ggf. der Online-Hilfe von Excel oder dem Merkblatt „Arbeiten mit Excel“. Sie müssen in Excel in einer (von Hand erzeugten) Spalte angeben, welche Daten jeweils zu einer „Klasse“ zusammengefasst werden sollen (z. B.] -105; -95],] -95; -85], ...,]95; 105]).

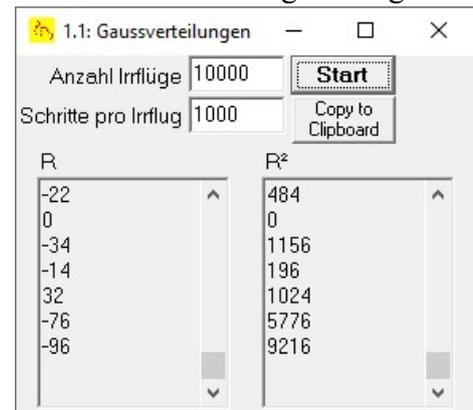


Abb. 4 - Dialogfenster 1.1 Gaußverteilungen

1.1.2 Erzeugen Sie außerdem je 1000 Flüge zu 5 verschiedenen Schrittzahlen: 10, 100, 1000, 10000 und 100 000 Schritte. Überzeugen Sie sich, dass die mittleren Verschiebungsquadrate $\langle R^2 \rangle$ proportional zur Anzahl der Schritte sind.

1.2 Mittlerer Fehler des Mittelwerts

Erzeugen Sie 100 Mittelwerte von jeweils $N_f = 100$ Irrflügen mit einer Schrittzahl von 1000 und stellen Sie die Mittelwerte der Summen $\langle R \rangle$ als Histogramm dar. Erzeugen Sie danach 100 Mittelwerte von jeweils 500 Irrflügen und 100 Mittelwerte von 1000 Irrflügen mit der gleichen Schrittzahl (Anzahl der Irrflüge / Mittelwert $N_f = 500$ bzw. 1000, s. Abb. 5). Stellen Sie wieder die Verteilungen der je 100 Mittelwerte als Histogramme dar.

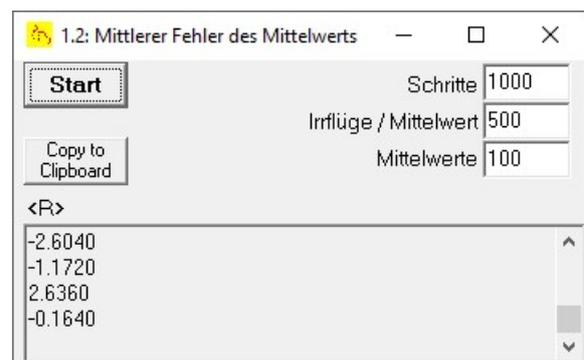


Abb. 5 - Dialogfenster 1.2
Mittlerer Fehler des Mittelwerts

Überzeugen Sie sich, dass der mittlere Fehler des Mittelwertes (die Breite der drei so erhaltenen Verteilungen) mit $N_f^{-1/2}$ skaliert.

Hinweis: Die „Breite der Verteilungen“ erhalten Sie, indem Sie die Standard-Abweichung der betreffenden Spalte berechnen. Folgen Sie der Excel on-line Hilfe.

1.3 Variable Schrittweiten

1.3.1 Überprüfen Sie das Skalengesetz für unterschiedliche (feste) Schrittweiten. Erzeugen Sie dazu zunächst 1000 Flüge mit der Schrittzahl 10000 und der Schrittweite 1. Berechnen Sie nach Gl. 3, wie viele Schritte nötig sind, um das gleiche mittlere Verschiebungsquadrat mit einer (festen) Schrittweite von 5 zu erhalten. Erzeugen Sie so 1000 Flüge mit dieser Schrittweite und vergleichen Sie die erhaltenen mittleren Verschiebungsquadrate.

1.3.2 Lassen Sie nun *variable* (ganzzahlige) Schrittweiten zu. Wählen Sie eine *maximale* Schrittweite von z.B. 100 und erzeugen Sie 10000 Irrflüge mit jeweils 1000 Schritten *zufälliger* Schrittweite zwischen 1 und 100 (Abb. 6). Berechnen Sie die *mittlere* Schrittweite und überzeugen Sie sich, dass auch für variable Schrittweiten die Summen gaußverteilt sind.

Dies ist der Inhalt des *zentralen Grenzwertsatzes*. Weil dieser gilt, dürfen wir uns bei unserer Betrachtung auf Irrflüge beschränken, die auf einem Gitter liegen (Schrittweiten ± 1). Solange die Computersimulation eine Gaußverteilung erzeugt, ist der Diffusionspfad aus statistischer Sicht ein Irrflug. *Feste* Schrittweiten in nur 3 oder 4 Raumrichtungen sind natürlich einfacher zu handhaben. Wie Sie gerade eben gezeigt haben, ist diese Vorgehensweise statistisch einem Irrflug mit beliebigen Schrittweiten und Winkeln äquivalent. Diese Äquivalenz setzt große Schrittzahlen ($N \rightarrow \infty$) voraus. Andernfalls spielt das Gitter durchaus eine Rolle.

Der zentrale Grenzwertsatz ist von allgemeinerer Bedeutung für unser Praktikum. Jedes Instrument hat Fehlerquellen, die die Messung verfälschen. Die einzelnen Fehler sind nicht gaußverteilt, ebenso wenig wie die Schrittweiten unserer Irrflüge, die vielmehr gleichverteilt sind (wie die Würfelergebnisse eines idealen Würfels). Wenn sich jedoch viele Fehler „unkorreliert“ aufaddieren, resultieren Gaußverteilungen der Messwerte (oder auch der mittleren Verschiebungsquadrate). Die einzelnen Schrittweiten sind zufällig und daher unkorreliert.

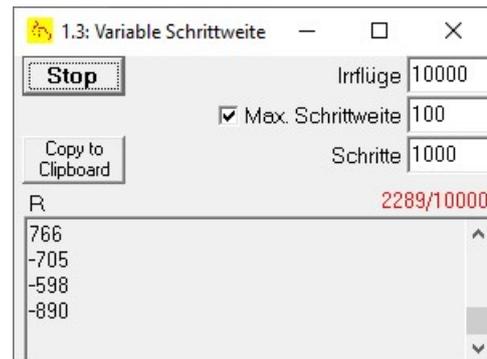


Abb. 6 - Dialogfenster 1.3
Variable Schrittweite

4.2 Irrflüge in 2 und 3 Dimensionen

Aufgaben:

2.1 Irrflüge in 2D (Illustration)

Wir begeben uns nun in die Ebene (wir lassen eine zweite Dimension zu). Wählen Sie drei Schrittanzahlen (≤ 10000) aus und erzeugen Sie Irrflüge. Dieser Teil des Versuches dient nur der Illustration

(s. Abb. 7), es ist keine Auswertung erforderlich. Mit dem Mausrad können Sie die Grafik vergrößern

oder verkleinern. Wenn Sie sie stark verkleinern, wirkt die Bewegung des Teilchens

„gleichmäßig“ - es ist kaum noch zu erkennen, dass die Bewegung auf einem Gitter stattfindet. Einer kleiner Wert für die Verzögerung lässt das Teilchen zwar schnellstmöglich

„diffundieren“, dabei wird aber leider auch die Darstellung durch starkes Flimmern gestört. Suchen Sie einen akzeptablen Kompromiss zwischen Geschwindigkeit und Qualität der

Darstellung.

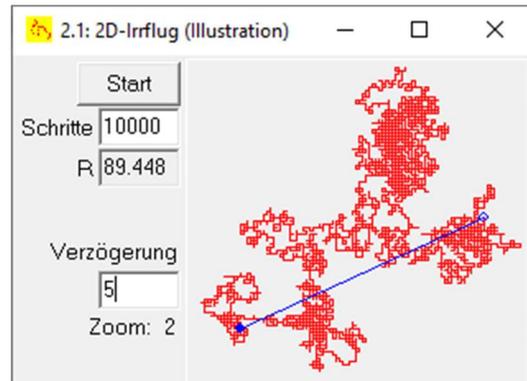


Abb. 7 - Dialogfenster 2.1 Illustration

2.2 Einstein-Smoluchowski-Beziehung und Skalengesetz

Wir kommen jetzt zu Bestimmung des Skalensexponenten und der statistischen (mittleren) Schrittweite anhand von Gleichung 2. Das Programm erzeugt dazu kontinuierlich Irrflüge, wobei die Anzahl der Schritte N und die quadratisch gemittelte Verschiebung R_{rms} in einer Tabelle gespeichert werden. Außerdem führt das Programm eine Regressionsrechnung zur Bestimmung der Parameter der (Geraden-) Gleichung 9 durch und zeigt das Ergebnis an.

2.2.1 Bestimmung des Skalensexponenten und der statistischen Schrittweite

Wählen Sie einen dreidimensionalen Irrflug mit zunächst einer Schrittweite von 1 und berechnen Sie Mittelwerte aus jeweils 500 Irrflügen. Damit sich in der doppelt logarithmischen Auftragung eine Gerade ergibt, müssen Mittelwerte für viele verschiedene Schrittzahlen N erzeugt werden. Wählen Sie daher 1000000 als die maximale Anzahl von Schritten (s. Abb. 8). Falls Sie von Ihren Vorgängern noch Daten auf dem Bildschirm hatten, klicken Sie bitte „Reset“. Nun starten Sie die Erzeugung der Irrflüge. Das Programm erzeugt dann kontinuierlich Datenpunkte mit zufällig gewählten Schrittzahlen zwischen 1 und 100000. Sobald drei Datenpunkte erzeugt sind, wird eine Regressionsgerade angepasst und die Regressionsparameter angezeigt (Schrittweite und Exponent unter „Fit“). Erzeugen Sie so mindestens 200 Datenpunkte. Sie können die erzeugten Daten über die Zwischenablage z.B. nach Excel kopieren oder als Textdatei speichern (TAB-getrennte Spalten, Datei-Endung .dat,) und später einlesen.

Bilden Sie in Excel die Logarithmen von den Anzahlen der Schritte und den quadratisch gemittelten Verschiebungen. Stellen Sie diese Werte dar und fügen Sie eine Regressionsgerade hinzu. Vergleichen Sie die von Excel erhaltenen Werte für den Skalensexponenten und die statistische Schrittweite mit denen vom Programm bestimmten.

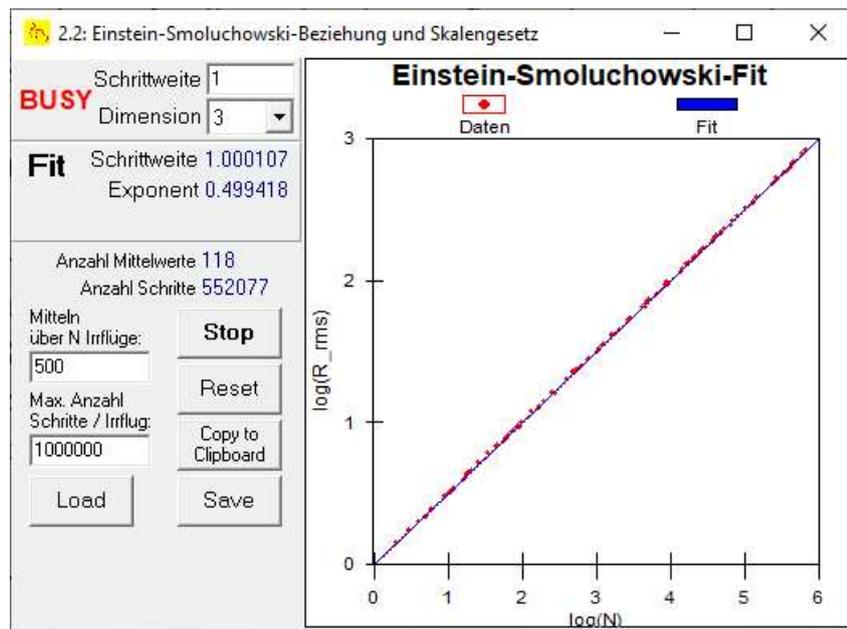


Abb. 8 - Dialogfenster 2.2 Einstein-Smoluchowski-Beziehung, Skalengesetz

2.2.2 Mittelwerte

An dieser Stelle möchten wir Sie auf ein wichtiges Detail aufmerksam machen. Löschen Sie die erzeugten Daten mit „Reset“. Wählen Sie für „Irrflüge pro Mittelwert“ den Wert 1 (es wird also *kein* Mittelwert gebildet) und lassen Sie die Simulation ansonsten mit denselben Parametern wie in Aufgabe 2.2.1 so lange laufen, bis Sie mindestens 500 Datenpunkte erzeugt haben. Sie werden eine Schrittweite < 1 und einen Exponenten $< 0,5$ finden. Es ist also wesentlich, dass *Mittelwerte* über eine hinreichend große Anzahl an Irrflügen gebildet werden. Anderenfalls finden Sie nicht die richtigen Werte.

2.2.3 Datenauswahl

Belassen Sie jetzt die Mittelung bei 1 (*keine* Mittelung) und erzeugen Sie mindestens 500 Irrflüge mit einer maximalen Länge von 10 Schritten. Vergleichen Sie die Ergebnisse des Programms RDW für die Ausgleichsgerade nach Gleichung 9 wieder mit dem Ergebnis, das Sie in Excel erhalten. Sie werden verschiedene Resultate finden. Das soll Sie lehren, keiner Software blind zu vertrauen.

Den Grund möchte man als tückisch bezeichnen. Bei kleinen Pfadlängen endet der Pfad manchmal am Ursprung. Die Verschiebung ist dann 0. Den Logarithmus von 0 darf man aber nicht bilden ($\lim(\ln x) = -\infty$ für $x \rightarrow 0$). Der Weg der Tugend wäre in dieser Situation, nicht den Logarithmus zu bilden, sondern direkt die Funktion $\lambda \cdot N^v$ an die rms-Verschiebungen anzufitten. Dies ist möglich und behebt das Problem. Nachdem dies eine etwas kompliziertere

Funktion ist, kommen wir aber bei Excel nicht mehr einfach mit der Regressionslinie aus. Ganz allgemein kann man sagen, dass „Linearisierungen“ auf den Rohdaten (wie das Logarithmieren in diesem Fall) gefährlich sind. Darauf wird auch im Anhang zum Skript *Adsorption* eingegangen.

Es treten also Nullen auf, die man nicht logarithmieren kann. Das Programm RDW *verwirft* nun diese Werte. Das ist nicht wirklich korrekt, denn es verfälscht die Statistik (Sie kennen die Prozedur aus der Arbeitslosenstatistik). Was Excel macht, ist aber noch „besser“: Zwar wird das Problem, den Logarithmus von 0 zu berechnen, in der Tabelle durch die Meldung „#ZAHL!“ signalisiert, in der Auftragung der Logarithmen werden diese Werte aber als 0 dargestellt und von der „Trendlinie“ auch als solche behandelt. Damit wird das ursprüngliche Verschiebungsquadrat zu 1 ($\log(1) = 0$). Excel *verändert* also die Werte! Und der Fit auf diesen veränderten Werten führt dann auch zu einem veränderten Fitergebnis. Sie erhalten das (annähernd) gleiche Ergebnis wie das Programm RDW, wenn Sie alle als „#ZAHL!“ gekennzeichneten Werte *löschen*.

Bitte merken Sie sich also erstens, dass der Teufel im Detail steckt und zweitens, dass das mittlere Verschiebungsquadrat von Zufallspfaden – ungeachtet aller Teufel im Detail – mit der Wurzel der Schrittzahl skaliert.

5 Anhang

5.1 Anhang 1: Zufallsgeneratoren

Ein Computer kennt keinen echten Zufall. Er verhält sich deterministisch. Trotzdem kann der Computer Zahlenketten erzeugen, die „so gut wie zufällig“ sind, in dem Sinne, dass auch eine detaillierte statistische Analyse den Unterschied zwischen dieser Zufallskette und einer echten Kette von Zufallszahlen nicht (oder nur unter extremem Aufwand) unterscheiden kann. Die Details der Zufallszahlengeneratoren sind oft sogar geheim, da sie im Zusammenhang mit der Kryptographie eine Rolle spielen (ebenso wie die „detaillierte statistische Analyse“ im Zusammenhang mit der Entschlüsselung). Ein Weg, um „Pseudo-Zufallszahlen“ zu erzeugen, beruht auf dem folgenden Verfahren: Man nehme eine hohe Primzahl und schneide die letzten 10 Ziffern ab. Die so erzeugte Zahl multipliziere man mit dieser Primzahl und schneide erneut die letzten 10 Ziffern ab. Diese Prozedur wird iteriert. Die so erhaltene Ziffernfolge sieht aus als wäre sie zufällig erzeugt worden.

Grundsätzlich ist dies ein zyklischer Algorithmus. Irgendwann wird der Lauf des Programms 10 Ziffern erzeugen, die schon einmal vorher erzeugt wurden. Ab dann wiederholen sich die Zahlen und es wird offensichtlich, dass es sich nicht um echte Zufallszahlen handelt. Gute Zufallszahlengeneratoren lösen dieses Problem auf die eine oder andere Weise.

5.2 Anhang 2: Zufallsgrößen und Verteilungen

Zur Analyse der erhaltenen Daten wird die Verteilung dieser Ergebnisse betrachtet. Um eine Verteilung bzw. eine Verteilungsfunktion zu bestimmen, können die erhaltenen Daten in einem Histogramm aufgetragen werden, in dem die Häufigkeit der Messgröße über der Messgröße selbst aufgetragen wird. Die x -Achse, Auftragung der Messgröße, wird dabei in diskrete Intervalle unterteilt. Die Wahrscheinlichkeit, dass einer der zufällig erzeugten Werte in eines der Intervalle der Häufigkeitsverteilung fällt, berechnet sich wie üblich als Quotient aus Häufigkeit der Werte in diesem Intervall und Anzahl aller Werte (Verhältnis „günstiger“ Ereignisse zur Anzahl aller Ereignisse). Mache man die Intervalle unendlich schmal, um eine kontinuierliche Wahrscheinlichkeitsverteilung zu erhalten, ginge auch die Wahrscheinlichkeit überall gegen 0. Man betrachtet daher besser die Wahrscheinlichkeitsdichte, die man erhält, wenn man die Wahrscheinlichkeit für ein Intervall durch die Intervallbreite teilt. An ein so erhaltenes Histogramm von Wahrscheinlichkeitsdichten kann nun eine kontinuierliche Funktion angepasst werden, welche die erhaltene Verteilung möglichst gut beschreibt. Diese Verteilung stellt das Ergebnis des Experimentes dar. Je nach durchgeführtem Experiment können sich unterschiedlichste Verteilungsfunktionen ergeben. Für den Fall, dass eine Summierung sehr vieler Werte x_1, x_2, \dots, x_N ($N \rightarrow \infty$) betrachtet wird, die durch unabhängige,

gleichverteilte Zufallsvariablen beeinflusst werden und die Intervalle der x -Achse sinnvoll gewählt sind, ergibt sich eine als Wahrscheinlichkeitsdichtefunktion (s. Abb. 9) der Form:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{\left(\frac{-(x-x_0)^2}{2\sigma^2}\right)} \quad (\text{A.1})$$

Diese Verteilung wird als Gaußverteilung (oder Normalverteilung) bezeichnet, mit x dem betrachteten Messwert, x_0 dem Mittelwert (oder Erwartungswert, s. u.) und σ der Standardabweichung.

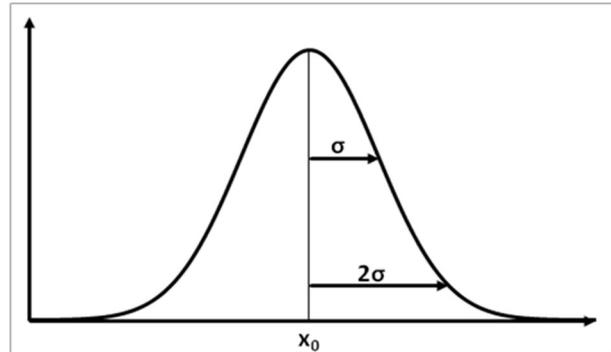


Abb. 9 - Schematische Darstellung einer Gauß-Funktion

Die Standardabweichung einer Stichprobe ist definiert als:

$$\sigma = \sqrt{\frac{\sum_1^N (x - x_0)^2}{N - 1}} \quad (\text{A.2})$$

Sie gibt an, wie stark die Messwerte um den Mittelwert streuen. So sollten etwa 2/3 aller Messwerte im Bereich $x_0 \pm \sigma$ bzw. etwa 95,5 % aller Messwerte im Bereich $x_0 \pm 2\sigma$ liegen. Neben der Standardabweichung für eine Stichprobe gibt es außerdem die Standardabweichung der Gesamtheit, in der der Mittelwert durch einen bekannten Erwartungswert ersetzt wird. Auf deren Details soll hier aber nicht näher eingegangen werden, da in der Datenanalyse, wie sie in der Chemie betrachtet wird, der Erwartungswert nur selten bekannt ist.

5.3 Anhang 3: Skaleninvarianz und Skalenexponent

Vorbemerkung: Die folgenden mathematischen Betrachtungen sprengen den Rahmen eines Grundpraktikums. Betrachten Sie diesen Abschnitt bitte als einen Ausblick auf das weiterführende Studium.

Es sei zunächst behauptet, dass man so tun darf, als sei die Schrittweite variabel, wie bei der Diffusion auch. Es gilt dann für den normalen Diffusionsprozess

$$\langle R^2 \rangle = N \cdot \langle \lambda^2 \rangle \quad (\text{A.3})$$

Wir fassen nun in Gedanken je 4 Schritte zu einem neuen Schritt zusammen und erhalten so einen neuen Irrflug mit der neuen Schrittzahl $N'=N/4$. Aus statistischer Sicht ist der neue Flug dem alten im Grunde recht ähnlich. Allerdings haben sich die Anzahl der Schritte und die mittlere Schrittweite geändert. Wir betrachten den neuen Flug als statistisch äquivalent zum alten Flug, wenn ebenfalls die Relation

$$\langle R^2 \rangle = \langle R'^2 \rangle = N' \cdot \langle \lambda'^2 \rangle \quad (\text{A.4})$$

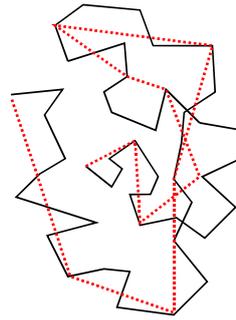


Abb. 10 - Reskalierter Pfad: Statistisch gesehen, ist der Zusammenhang zwischen Verschiebung und Segment-Anzahl für den gepunkteten Pfad derselbe wie für den Pfad mit durchgezogener Linie.

Wenn man an vielen verschiedenen Pfaden je 4 Schritte zu einem neuen Schritt zusammenfasst, gilt für die Mittelwerte die Relation (6)

gilt. Gestrichene Größen bezeichnen dabei den neuen Flug.

Mit $N'=N/4$ und Gleichung A.3 können wir die neue mittlere Schrittweite λ' bestimmen:

$$N \cdot \langle \lambda^2 \rangle = \langle R^2 \rangle = \langle R'^2 \rangle = N' \cdot \langle \lambda'^2 \rangle = \frac{N}{4} \cdot \langle \lambda'^2 \rangle \quad (\text{A.5})$$

und also

$$\langle \lambda'^2 \rangle = 4 \cdot \langle \lambda^2 \rangle$$

Allgemein schreibt man eine solche „Skalentransformation“ in der Form:

$$N \rightarrow N' = \frac{N}{n} \quad (\text{A.6})$$

$$\lambda \rightarrow \lambda' = n^{\nu} \cdot \lambda \quad \nu = \frac{1}{2} \quad (\text{A.7})$$

In dem obigen Beispiel war $n = 4$, aber man hätte auch 5 oder jede andere Anzahl n von alten Schritten zu einem neuen Schritt zusammenfassen können.

Der Exponent ν in Gleichung A.7 ist der „Skalenexponent“. N und λ müssen gleichzeitig transformiert werden, damit der neue Pfad dieselbe Verschiebung hat. Wir haben den Skalenexponenten so gewählt, dass die neue Verschiebung gleich dem alten ist. Diese Bedingung fixiert ν . Für den normalen Irrflug muss der Skalenexponent $1/2$ sein. Dieser

Skalenexponent von $1/2$ gilt jedoch nur für Irrflüge *ohne* Selbstvermeidung. *Mit* Selbstvermeidung, also für Irrflüge, die z. B. die Gestalt eines Polymerknäuels beschreiben, hat man ebenfalls Skaleninvarianz, aber man findet andere Exponenten.

Wann liegt Skaleninvarianz vor? Skaleninvarianz bedeutet, dass das Problem keine natürliche innere Längenskala hat. Eine solche natürliche Längenskala könnte z. B. der Durchmesser einer Pore sein, indem sich ein diffundierendes Teilchen, z. B. Polymerkette, befindet. Es könnte auch die Reichweite einer abstoßenden Wechselwirkung zwischen Kettensegmenten sein. Wenn es eine solche Längenskala *nicht* gibt, hat man Skaleninvarianz. Die Skaleninvarianz ist eine Symmetrie wie die Isotropie auch. Invariant bleiben aber nur die *statistischen* Eigenschaften des Objekts oder einer großen Menge gleichartiger Objekte, nicht die individuelle Eigenschaft (z. B. die Gestalt des Polymerknäuels).

Skaleninvarianz hat stets *Potenzgesetze* zur Folge:

$$R = \lambda \cdot N^\nu \quad (\text{A.8})$$

Betrachten wir zwei Gegenbeispiele. Angenommen die Verschiebung gehorche der Relation $R \sim \exp(N/N_0)$. Dann ist die Größe $(N_0 \cdot \lambda)$ eine charakteristische Länge, die man (mikroskopisch) interpretieren kann. Oder es gelte die Relation $R \sim aN + bN^2$ mit zwei Koeffizienten a und b . Das kann aber auch als $R \sim aN(1 + N/N_0)$ schreiben mit $\lambda \cdot N_0 = \lambda \cdot b/a$ erneut einer charakteristischen Länge. Nur Potenzgesetze enthalten kein solches N_0 .

Wir fassen zusammen: Wenn das Problem keine natürliche innere Längenskala hat, kann man durch Skalentransformation Pfade erzeugen, die dem alten Pfad statistisch äquivalent sind. Skaleninvarianz hat stets Potenzgesetze zur Folge. Die Skalenexponenten sind universell, sie hängen nur von der Raumdimension², nicht aber von den Materialeigenschaften ab.

² Beim Irrflug ohne Selbstvermeidung erhält man für ein-, zwei-, drei und vierdimensionale Diffusion den Skalenexponenten $1/2$.
