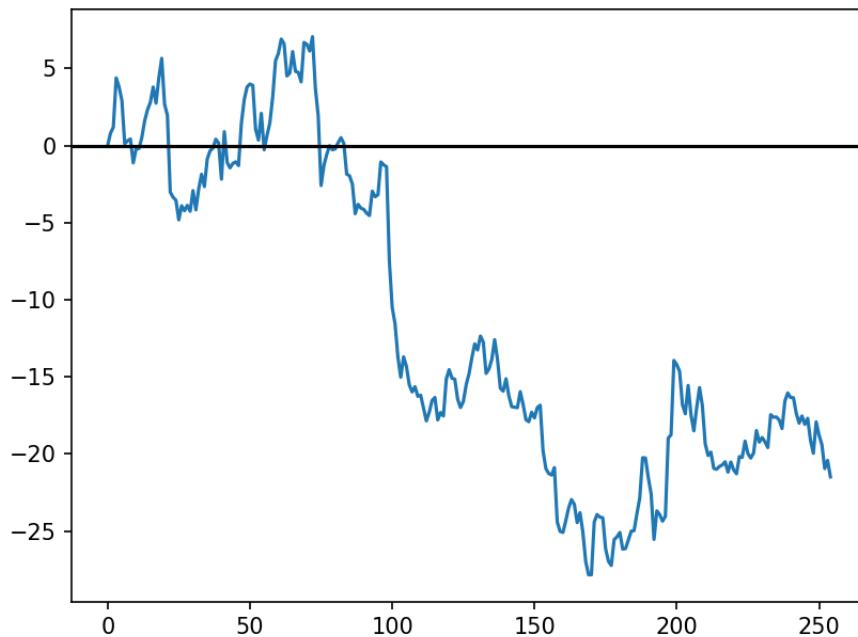


KANTONSSCHULE SURSEE

Wettbewerbsarbeit für Schweizer Jugend forscht

# Modellierung von Aktienkursen durch modifizierte Random Walks



**Abbildung 1** symmetrischer random walk mit t-verteiltern Schrittlängen

**Chiara Willimann**

Tann 8  
6214 Schenkon

**Betreuende Lehrperson:**

Michael Muri  
St. Karlistrasse  
6004 Luzern

**Fachperson von Schweizer Jugend forscht:**

Dr. Xavier Richard

Schenkon, den 31. März 2024

## Abstract

Im Rahmen dieser Arbeit wird das Konzept einfacher, eindimensionaler random walks auf  $\mathbb{Z}$  präsentiert. Daraufhin wird untersucht, inwieweit dieses statistische Modell Eigenschaften mit Aktienkursen teilt. Dazu werden relevante Eigenschaften dieser random walks präsentiert und erläutert. Der Fokus liegt primär auf den Nullstellen und den Extremstellen der random walks. Anschliessend werden diese Eigenschaften der random walks mit den Eigenschaften von Aktienkursen verglichen. Anhand dieser Analyse soll das gewählte Random-Walk-Modell durch eine Modifizierung verbessert werden. Schliesslich wird durch eine ausführliche, computergestützte Datenanalyse das Modell mit echten Kursdaten von an der Schweizer Börse kotierten Unternehmen überprüft. Die Ergebnisse deuten darauf hin, dass durchaus Gemeinsamkeiten zwischen random walks und Aktienkursen existieren. Es ist gelungen, eine sinnvolle Modifikation zu finden, indem die einzelnen Schritte der random walks mit Hilfe der Student-t-Verteilung modelliert werden. In diesem Zusammenhang wird aufgezeigt, dass eine Annahme der Normalverteilung der Schrittgrössen zu einer entscheidenden Unterschätzung der Wahrscheinlichkeit von Extremevents in Form von sehr grossen Kursveränderungen in einem kurzen Zeitraum führen.

Die anschliessende Analyse des Modells im Vergleich zu echten Kursdaten zeigt einige Schwachstellen des Modells, wie beispielsweise das Phänomen des 'Volatility Clusterings'. Dennoch konnte eine signifikante Ähnlichkeit in Bezug auf durchschnittliche Ausprägungen und die Verteilung zentraler Eigenschaften zwischen Aktienkursen und random walks festgestellt werden.

In Anbetracht der ständigen Debatte über die Effizienz der Finanzmärkte und der zugrundeliegenden Modelle der Kursbewegungen, ist eine Untersuchung der Beziehung zwischen random walks und Aktienkursen von Bedeutung. Neben der theoretischen Relevanz bietet dieses Thema auch praktische Anwendungsmöglichkeiten. Ein verbessertes Verständnis der Preisbewegungen an den Finanzmärkten kann zur Verbesserung der Handelsstrategien führen sowie einen Beitrag zur Verhinderung von volkswirtschaftlichen Krisen leisten.

# Inhaltsverzeichnis

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Einleitung</b>   | <b>3</b>  |
| 1.1      | Idee und Ziele . . . . .  | 3         |
| 1.1.1    | Leitfragen . . . . .  | 4         |
| 1.2      | Aufbau . . . . .  | 4         |
| 1.3      | Aktien . . . . .  | 4         |
| 1.3.1    | Was sind Aktien? . . . . .  | 4         |
| 1.3.2    | Effiziente Märkte . . . . .   | 5         |
| 1.3.3    | Die Random-Walk-Theorie . . . . .                                     | 5         |
| <b>2</b> | <b>Random Walks</b>   | <b>7</b>  |
| 2.0.1    | Random Walk auf $\mathbb{Z}$ . . . . .                                | 7         |
| 2.1      | Grundlegende Beweisidee . . . . .                                     | 8         |
| 2.1.1    | Abzählen der Wege . . . . .   | 9         |
| 2.1.2    | Spiegelungsprinzip . . . . .  | 10        |
| 2.2      | Nullstellen . . . . .   | 10        |
| 2.2.1    | Nullstellen bei Aktienkursen . . . . .                                | 10        |
| 2.2.2    | Nullstellen bei Random Walks auf $\mathbb{Z}$ . . . . .               | 11        |
| 2.2.3    | Das Hauptlemma . . . . .  | 11        |
| 2.2.4    | Anzahl der Nullstellen . . . . .                                      | 12        |
| 2.2.5    | Die erste Nullstelle . . . . .  | 14        |
| 2.2.6    | Die letzte Nullstelle . . . . .                                       | 15        |
| 2.2.7    | Arcus-Sinus-Gesetz . . . . .  | 17        |
| 2.3      | Extremstellen . . . . .   | 19        |
| 2.3.1    | Maximum und Minimum . . . . .   | 19        |
| 2.3.2    | Lage der Extremstellen . . . . .                                      | 23        |
| <b>3</b> | <b>Modifikation der Random Walks</b>                                  | <b>26</b> |
| 3.1      | Länge der einzelnen Schritte . . . . .                                | 26        |
| 3.1.1    | Normalverteilte Schrittlänge . . . . .                                | 27        |
| 3.1.2    | t-verteilte Schrittlänge . . . . .                                    | 29        |
| 3.2      | Relevanz der Modellierung von Extremereignissen . . . . .             | 32        |
| 3.2.1    | Versagen der Finanzmodelle als Teilursache von Finanzkrisen . . . . . | 33        |
| 3.2.2    | Volatility Clustering . . . . .                                       | 33        |
| <b>4</b> | <b>Datenanalyse</b>   | <b>34</b> |
| 4.0.1    | Auswahl der Aktien . . . . .  | 34        |
| 4.0.2    | Normalisierung der Aktienkurse . . . . .                              | 36        |
| 4.1      | Vergleich durch Visualisierung . . . . .                              | 38        |

|          |   |           |
|----------|---|-----------|
| 4.1.1    | Zeitraum: 20 Börsentage . . . . .   | 38        |
| 4.1.2    | Zeitraum: 1 Jahr . . . . .  | 40        |
| 4.1.3    | Zeitraum: 10 Jahre . . . . .  | 42        |
| 4.2      | Vergleich der Eigenschaften bezüglich Nullstellen und Extremstellen . . . . .     | 42        |
| 4.2.1    | Vergleich der Durchschnittswerte . . . . .  | 42        |
| 4.2.2    | Anzahl Nullstellen . . . . .  | 44        |
| 4.2.3    | Die erste Nullstellen . . . . .   | 45        |
| 4.2.4    | Die letzte Nullstelle . . . . .   | 46        |
| 4.2.5    | Maximum . . . . .   | 47        |
| 4.2.6    | Minimum . . . . .   | 50        |
| 4.3      | Vergleich durch eine Abschnittsuche innerhalb eines langen random walks . . . . . | 52        |
| <b>5</b> | <b>Diskussion</b>   | <b>55</b> |
| <b>6</b> | <b>Schlusswort</b>  | <b>58</b> |
|          | <b>Literaturverzeichnis</b>   | <b>60</b> |
|          | <b>Abbildungsverzeichnis</b>  | <b>62</b> |
|          | <b>Tabellenverzeichnis</b>  | <b>62</b> |
| <b>A</b> | <b>Programmcode zur Datenanalyse</b>  | <b>63</b> |
| A.1      | 'main.py' . . . . .   | 63        |
| A.2      | 'random walk generator.py' . . . . .  | 66        |
| A.3      | stocks.py . . . . .   | 67        |
| A.4      | stock data.py . . . . .   | 67        |
| A.5      | histogram.py . . . . .  | 70        |
| A.6      | goodness of fit.py . . . . .  | 71        |
| A.7      | plot.py . . . . .   | 72        |
| A.8      | Liste der Ticker-Symbole . . . . .  | 73        |
| <b>B</b> | <b>Digitale Dateiablage</b>   | <b>74</b> |
|          | <b>Deklaration</b>  | <b>75</b> |

# Kapitel 1

## Einleitung

Aktien haben einen grossen Einfluss auf unsere Gesellschaft und sind prägend für die Wirtschaft. Der Handel mit Wertpapieren, wozu auch Aktien gehören, ist besonders in den Medien allgegenwärtig. Auch viele Privatpersonen zeigen Interesse an risikoreicheren Investments in Einzelaktien oder Aktienfonds, diese stellen vermehrt auch eine Möglichkeit zur privaten Altersvorsorge dar [9]. Im Gegensatz zu langfristigen Investments in Rentenfonds gibt es auch viele Privatanleger, die sich durch hohe Risikobereitschaft hohe Gewinne versprechen [9].

Es liegt in der Natur des Wertpapierhandels, dass dieser immer risikobehaftet ist. Investoren haben ein Interesse darin, Methoden zu finden, um dieses Risiko abzuschätzen und die Finanzmärkte besser zu verstehen. Dies führt zur Entwicklung mathematischer Modelle, um die Entwicklung von Aktienkursen zu analysieren und besser zu verstehen [9]. Heute ist die Finanzmathematik ein grosses Forschungsgebiet innerhalb der Mathematik und dessen Erkenntnisse kommen in der Wirtschaft direkt zur Anwendung [9].

### 1.1 Idee und Ziele

Im Fokus dieser Arbeit steht die Random-Walk-Theorie. Diese Theorie besagt, dass sich Aktienkurse zufällig und unvorhersehbar bewegen. Ein Ziel besteht darin, einfache, symmetrische random walks auf  $\mathbb{Z}$  zu verstehen. Ausgewählte Eigenschaften sollen präsentiert und begründet werden. Es soll aufgezeigt werden, welche Rolle das in Abschnitt 2.2 vorgestellte Hauptlemma spielt, um verschiedene Aussagen zu beweisen. Durch eine Auswertung von ausgewählten Aktienkursen soll untersucht werden, welche Eigenschaften die Aktienkurse mit den random walks teilen. Es soll eine Simulation der random walks erstellt werden und schliesslich anhand der gewonnenen Erkenntnisse modifiziert werden, um den Eigenschaften der Aktienkurse besser zu entsprechen. Die Arbeit soll diverse Aspekte und Themenbereiche betreffen, einerseits im Bereich der Mathematik sowie der Finanzökonomie. Nicht zu vernachlässigen ist aber auch die Herausforderung, neue Fähigkeiten bezüglich Datenauswertung und -visualisierung mit der Programmiersprache 'Python' und im Schreiben eines wissenschaftlichen Texts mit  $\text{\LaTeX}$  zu erlangen.

### 1.1.1 Leitfragen

Folgende Leitfragen sollen durch diese Arbeit führen und dessen Ausrichtung aufzeigen. Schliesslich werden im Kapitel 5 diese Fragen nochmals aufgenommen und diskutiert, inwiefern sie im Rahmen dieser Arbeit beantwortet werden konnten.

- Welche Eigenschaften weisen einfache, symmetrische random walks auf  $\mathbb{Z}$  auf?
- Wie können random walks modifiziert werden, um Aktienkurse besser zu modellieren?
- Worin unterscheiden sich das Random-Walk-Modell und die empirischen Kursdaten?
- Welche Aussage kann über die Korrektheit der Random-Walk-Theorie gemacht werden?

## 1.2 Aufbau

Zuerst werden die nötigen Grundlagen im Bezug auf Aktien und Aktienmärkte aufgezeigt und auf die verwendete Methodik der Auswertung der Aktienkurse eingegangen. Anschliessend soll das grundlegende Konzept der einfachen, symmetrischen random walks auf  $\mathbb{Z}$  vorgestellt werden. Es werden ausgewählte, relevante Eigenschaften von einfachen, symmetrischen random walks auf  $\mathbb{Z}$  aufgezeigt und begründet. Zuerst liegt der Fokus auf den Nullstellen. In diesem Abschnitt 2.2 wird auch die zentrale Rolle des Hauptlemmas deutlich. Anschliessend widmet sich das Kapitel 2 den Extremstellen der random walks. Dieser theoretische Teil der Arbeit bildet die Grundlage und das Verständnis des gewählten Grundmodells, was bei einer entsprechenden Weiterführung durch eine Modifizierung vorausgesetzt wird. Das Konzept der einfachen, symmetrischen random walks auf  $\mathbb{Z}$  wird in einer Simulation implementiert. Die Simulation wird mit Hilfe der Programmiersprache 'Python' umgesetzt und soll einerseits die Wahl einer sinnvollen Modifikation begründen sowie eine ausgiebige Datenanalyse ermöglichen. Zusätzlich soll die Simulation auch die Möglichkeit bieten, einzelne random walks zufällig zu generieren und diese graphisch darzustellen. Im Kapitel 3 werden mögliche Modifizierungen vorgestellt und diskutiert. Für sinnvoll erklärte Modifikationen werden schliesslich im Programmcode ergänzt. Abschliessend werden die Werte der Modifikation mit den Ausgangswerten verglichen und die Arbeit wird mit einer Diskussion und Reflexion der Ergebnisse abgerundet. Im Anhang befinden sich die angesprochenen Programmcodes.

## 1.3 Aktien

In diesem Abschnitt werden zuerst die grundlegenden Begriffe im Zusammenhang mit Aktien und dem Aktienmarkt vorgestellt. Schliesslich wird das methodische Vorgehen bezüglich der Auswahl und Normalisierung der Aktienkurse erläutert. Die ausgewählten Aktienkurse werden im Verlauf dieser Arbeit auf ihre Eigenschaften untersucht und mit den random walks verglichen.

### 1.3.1 Was sind Aktien?

Aktien sind meist kleine Anteile eines Unternehmens, einer sog. Aktiengesellschaft. Es wird unterschieden zwischen Quotenaktien, wo der Aktionär einen festen Anteil des Unternehmens besitzt und Aktien mit konstantem Nennwert, der jedoch nicht mit dem Kurswert einer Aktie verwechselt werden darf [9]. Aktien werden an Börsen gehandelt und der Aktienkurs ergibt sich aus Angebot und Nachfrage, was wesentlich durch die Höhe der Dividendenausschüttung und das zukünftige Potential der Aktie bestimmt ist. Aktien dienen den Unternehmen als eine Möglichkeit zur Kapitalbeschaffung [9]. Für die Aktionäre sind Aktien eine riskante Investitionsmöglichkeit, wobei sie sich eine hohe Dividendenauszahlungen und einen steigenden Aktienkurs erhoffen [9].

### 1.3.2 Effiziente Märkte

Die Markteffizienzhypothese besagt, dass aktuelle und vergangene Wirtschaftsbewegungen nicht verwendet werden können, um aktuelle oder zukünftige Kursbewegungen vorherzusagen [11]. Die Theorie wurde wesentlich geprägt von Eugene Fama, der im Jahr 1970 eine Untersuchung zu diesem Thema veröffentlichte [3]. Begründet wird diese Theorie damit, dass der Markt alle verfügbaren Information enthält und diese sich bereits in den Kurspreisen zeigen. Daher ist es nach dieser Theorie unmöglich, den Markt konsistent und langfristig zu schlagen und extreme Gewinne zu erzielen [11]. Marktpreise reagieren also ausschliesslich auf neue Informationen, was zu unvorhersehbaren, zufälligen Kursbewegungen führt, die unabhängig von den vergangenen Kursbewegungen sind. Es müssen einige Bedingungen erfüllt werden, damit ein effizienter Markt vorliegen kann. Es müssen alle Informationen allen Marktteilnehmern kostenfrei zu Verfügung stehen und alle Marktteilnehmer handeln rational und stehen in direkter Konkurrenz zueinander. Weiter dürfen keine Transaktionskosten anfallen [2]. Es wird zwischen drei Arten der Markteffizienz unterschieden:

- Die **schwache Form** der Markteffizienz basiert auf der Annahme, dass historische Information bereits vollständig im Preis reflektiert werden und ausschliesslich neue Informationen sich auf zukünftige Kursbewegungen auswirken [11]. Sich am Trend eines Kurses zu orientieren, um zukünftige Bewegungen vorherzusagen ist laut dieser Theorie nicht möglich und es gibt keine regelmässig auftretenden Muster in den Kursbewegungen. Einzelne Preisbewegungen sind nicht voneinander abhängig und von zufälliger Natur [11].
- Weiter gibt es eine **halbstarke Form**, die einerseits die Annahmen der schwachen Form einschliesst und zusätzlich noch die Behauptung aufstellt, dass zusätzlich alle öffentlich verfügbaren Informationen sich sofort im Markt widerspiegeln. Bei diesen öffentlichen Informationen handelt es sich beispielsweise um allgemeine Wirtschaftsdaten, öffentlich verfügbare Analysen und politische Nachrichten. Der Schluss daraus ist, dass sich einem Investor keine öffentlich verfügbare Informationsquelle bietet, um konsistente Überrenditen zu generieren [11].
- Die dritte Form der Markteffizienz ist die **starke Form**. Es wird davon ausgegangen, dass zusätzlich zu allen Informationen, die in den ersten beiden Formen widerspiegelt werden, auch alle privaten Informationen oder allgemein alles, was durch irgendeine Methoden erkennbar ist, in den aktuellen Marktpreis bereits einfließt. Somit sind selbst Insiderinformationen nutzlos, um den Preis eines Wertpapiers vorherzusagen. Übermässige Gewinne generieren ist langfristig generell unmöglich [12].

### 1.3.3 Die Random-Walk-Theorie

Das stochastische Modell der random walks soll nun mit Aktienkursbewegungen in Verbindung gebracht werden. Die Random-Walk-Theorie bietet die Grundlage für diese Verbindung. Im Rahmen dieser Arbeit sollen die Ähnlichkeiten zwischen random walks und Aktienkursen untersucht werden und daraus eine kritische Stellungnahme zur Random-Walk-Theorie ermöglichen.

Die von Bachelier im Jahre 1900 entwickelte Random-Walk-Theorie entspricht der Theorie der schwachen Form der Markteffizienz, die im Abschnitt (1.3.2) vorgestellt wurde [11]. Sie besagt, dass jegliche Preisbewegungen der Vergangenheit bereits im aktuellen Preis widerspiegelt sind und es daher von keinem Nutzen für Investoren ist, jüngste Preistrends als Auswahlgrundlage zu wählen [12]. Es handelt sich laut dieser Theorie um reine Zeitverschwendung, Analysen der vergangenen Kursbewegungen zu machen, da zum Beispiel die Preisbewegungen der letzten Tage keine Information über die Preisbewegung der Gegenwart und der Zukunft liefern. Demnach existieren auch keine Muster und Regelmässigkeiten die sich periodisch wiederholen. Die Random-Walk-Theorie besagt, dass jede Preisbewegung unabhängig von vorhergehenden Bewegungen ist und daher betrachtet sie Kurs-

schwankungen als zufällig [11].

Die Kursbewegungen auf dem Aktienmarkt zu verstehen, ist von grossem Interesse für Investoren und das Entwickeln von Modellen zur Modellierung und Vorhersage der Kurse ist ein zentrales Anliegen der Finanzmathematik [15]. Ein solches Modell ist die Random-Walk-Theorie, wonach Kursbewegungen einem zufälligen Prozess entsprechen und somit nicht vorhergesagt werden können.

Unterstützer der Random-Walk-Theorie wenden keine Zeit für Analysen vergangener Preise auf und richten ihre Handelsstrategie nach dieser Theorie. Nach dieser Theorie sind primär passive Anlagestrategien sinnvoll und es wird ein möglichst diverses Portfolio an Wertpapier zusammengestellt. Eine Möglichkeit dazu bieten sogenannte Indexfonds oder 'Exchange Traded Funds', die auch als 'ETF' bezeichnet werden. Durch ein diverses Portfolio können gerade auch Einzelrisiken eingedämmt werden. Von einem Versuch, den idealen Ein- und Ausstiegszeitpunkt zu finden, wird abgesehen, da es nach der Random-Walk-Theorie nicht möglich ist einen idealen Zeitpunkt vorherzusagen [12].

Es soll auch angemerkt werden, dass die Random-Walk-Theorie umstritten ist und es auch alternative Theorien gibt, die eine Vorhersage der Preisbewegung in gewissem Masse in Betracht ziehen [2].



## Kapitel 2

# Random Walks

Der Begriff 'random walk', der im Deutschen als 'Irrfahrt' bezeichnet wird, ist eine mathematische Beschreibung eines zufälligen Prozesses. Der Begriff stammt von Karl Pearson, der im Jahr 1905 den Aufsatz 'The Problem of the Random Walk' im Magazin 'Nature' veröffentlichte [7]. Heute wird das Konzept der random walks in verschiedenen wissenschaftlichen Disziplinen verwendet um zufällige Bewegungen und Veränderungen zu modellieren. Besonders im Bereich der Finanzmathematik, der Physik, der Biologie und der Informatik kommt das Konzept der random walks zur Anwendung [7].

### 2.0.1 Random Walk auf $\mathbb{Z}$

Als einfaches Beispiel für einen eindimensionalen random walk wird eine Person betrachtet, die sich auf einem Zahlenstrahl der ganzen Zahlen  $\mathbb{Z}$  bewegt und den nächsten Schritt mittels einem Münzenwurf bestimmt. Landet die Münze auf Kopf, bewegt sich die Person um einen Schritt nach rechts, auf die nächstgrössere ganze Zahl. Landet die Münze auf Zahl, bewegt sich die Person einen Schritt nach links auf die nächstkleinere ganze Zahl. Den Weg, den die Person durch mehrmaliges Wiederholen dieses Vorgangs zurücklegt, nennt man *random walk*. An diesem Beispiel wird erkennbar, dass jeder neue Schritt zufällig nach rechts oder nach links abläuft und unabhängig von den vorherigen Schritten ist [7].

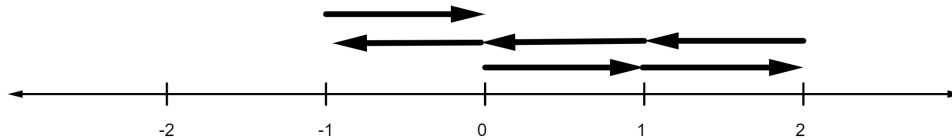
Zuerst sollen einfache, symmetrische random walks und die dazugehörige Notation definiert werden. Es wird definiert, dass der random walk in 0 seinen Anfangspunkt hat und sich auf der Zahlenmenge  $\mathbb{Z}$  der ganzen Zahlen bewegt. Die Zufallsfolge  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  besteht aus einer Abfolge unabhängiger Zufallsvariablen. Die Unabhängigkeit der Zufallsvariablen in  $\mathbf{X}$ , die auch als Gedächtnislosigkeit bezeichnet wird, ist ein zentrales Merkmal der random walks. Jede Zufallsvariable  $X_n$  hat die Eigenschaft

$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 1/2. \quad (2.1)$$

Es sind die beiden Ergebnisse  $\{X_n = 1\}$  beziehungsweise  $\{X_n = -1\}$  denkbar, wobei ersteres Ereignis einem Schritt nach rechts und letzteres Ereignis einem Schritt nach links entspricht. Die Gleichwahrscheinlichkeit dieser Ereignisse führt zur Bezeichnung *symmetrisch*. Von einem *einfachen* random walk spricht man, weil nur die beiden Ereignisse möglich sind und sich diese nur im Vorzeichen unterscheiden, nicht aber in der Schrittlänge. Die Variable  $n$  gibt an, dass es sich um die  $n$ -te Zufallsvariable handelt. Man spricht auch vom  $n$ -ten Schritt des random walks [7].

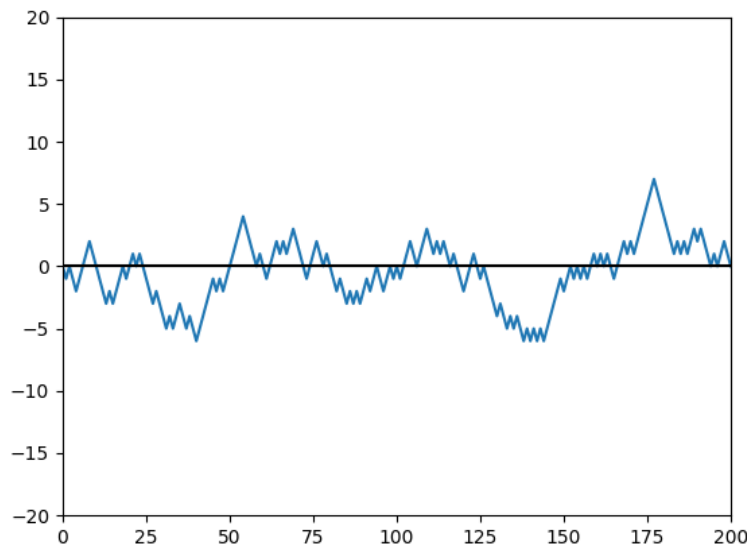
Die Position des random walks auf den ganzen Zahlen zum Zeitpunkt  $n$  wird durch die Zufallsvariable  $S_n$  angegeben. Es wird festgelegt, dass der Ausgangspunkt des random walks  $S_0 := 0$  ist, wobei  $n \geq 1$  sein muss. Die Zufallsvariable  $S_n$  wird gesetzt als [7]

$$S_n := \sum_{k=1}^n X_k. \tag{2.2}$$



**Abbildung 2.1** Ein random walk auf  $\mathbb{Z}$  der Länge  $n = 6$ , dargestellt auf einem Zahlenstrahl

Zur Darstellung des Verlaufs des random walks ist ein einfacher Zahlenstrahl (Abb. 2.1) ungeeignet, weil sich die einzelnen Schritte bei häufigen Überlappungen um den Ursprung schlecht darstellen lassen. Geeigneter ist ein Koordinatensystem, wobei die y-Achse den Zahlenstrahl, auf dem sich der random walk bewegt, darstellt. Die x-Achse zeigt an, um den wievielten Schritt des random walks es sich handelt. Der zeitliche Verlauf wird sichtbar. Da der random Walk in 0 startet und  $n \geq 1$  definiert wurde, werden zur Darstellung des random walks nur der I. und der IV. Quadrant des Koordinatensystems benötigt.



**Abbildung 2.2** Ein random walk auf  $\mathbb{Z}$  der Länge  $n = 200$

## 2.1 Grundlegende Beweisidee

Das Modell der random walk, das auf den ersten Blick einfach und intuitiv verständlich scheint, bietet eine wahre Herausforderung, wenn dessen Eigenschaften bezüglich Nullstellen und Extremstellen erklärt und bewiesen werden sollen. Da random walk als statistisches Modell eine grosse Rolle innerhalb dieser Arbeit einnehmen und in den folgenden Kapiteln nicht jede Eigenschaft abschliessend bewiesen wird, soll an dieser Stelle auf die grundlegende Beweisidee eingegangen werden. Um diese Beweise überhaupt anstellen zu können, ist es wichtig zu spezifizieren, dass diese nur für einfache,

symmetrische random walks auf  $\mathbb{Z}$  gelten. Denn ausschlaggebend sind die einheitliche Schrittgrösse sowie die Gleichwahrscheinlichkeit der beiden möglichen Richtungen.

Jeder Schritt eines einfachen, symmetrischen random walks auf  $\mathbb{Z}$  kann als Bernoulli-Experiment aufgefasst werden. Es handelt sich dabei um ein Zufallsexperiment, das nur zwei Ergebnisse aufweist. Es ist das Ergebnis 'Erfolg' oder das Gegenergebnis 'Misserfolg' möglich. Wir definieren einen Schritt nach rechts auf dem Zahlenstrahl als 'Erfolg' und einen Schritt nach links als 'Misserfolg'.

Bei einem random walk hängen wir  $n$  Bernoulli-Experimente aneinander und erhalten eine Bernoulli-Kette mit Länge  $n$ . Zentral bei einer Bernoulli-Kette ist die gleichbleibende Erfolgswahrscheinlichkeit  $p$ . Diese ist bei unserem symmetrischen random walk gegeben und beträgt immer  $p = 1/2$ .

### 2.1.1 Abzählen der Wege

Das Zählen von möglichen Wegen, die ein random walk nehmen kann, ist eine wichtige Methode, um seine statistischen Eigenschaften zu beschreiben. Soll die Wahrscheinlichkeit bestimmt werden, dass ein random walk einer gewissen Länge eine gewisse Eigenschaft hat, wird die Menge aller möglichen Wege mit dieser Eigenschaft bestimmt und durch die gesamthaft mögliche Anzahl verschiedener Wege dieser Länge dividiert.

Um die allgemeine Anzahl möglicher Wege anzugeben, wird jeweils die Notation  $|W|$  verwendet. Da alle Schritte des random walks voneinander unabhängig sind und jeweils den Wert 1 oder -1 annehmen können, gibt es für einen random walk der Länge  $k$  genau

$$|W_n| = 2^n \quad (2.3)$$

verschiedene Wege, welche gleichwahrscheinlich auftreten [7].

Für die Wahrscheinlichkeit der verschiedenen Wege gilt also [7]

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \left(\frac{1}{2}\right)^n, (x_1, \dots, x_n) \in \{-1, 1\}^n. \quad (2.4)$$

Es wird ein symmetrischer random walk mit  $h$  Aufwärtsschritten und  $v$  Abwärtsschritten sowie mit Anfangspunkt  $a$  und Endpunkt  $b$  betrachtet. Da die Wahrscheinlichkeit für Auf- und Abwärtsschritte identisch ist, gilt  $p = q = 1/2$  (2.1). Aufgrund dieser Gleichwahrscheinlichkeit treten alle möglichen Wege mit gleicher Wahrscheinlichkeit auf. Soll die Anzahl möglicher Wege von einem Punkt  $a$  zum Punkt  $b$  gefunden werden, betrachtet man zuerst die Differenz der Positionen. Es lässt sich daraus bestimmen, wie viele Auf- und Abwärtsschritte nötig sind. Es stellt sich die Frage, auf wie viele Arten sich  $h$  Aufwärtsschritte aus  $n$  Gesamtschritten ausgewählt werden können. Im Allgemeinen wird dieser Sachverhalt als **Kombination ohne Wiederholung** genannt. Es soll 'n über h' Möglichkeiten geben, um die Anzahl möglichen Wege zu finden. Dafür wird der Binomialkoeffizient definiert als

$$\binom{n}{h} = \frac{n!}{h!(n-h)!} \quad (2.5)$$

An dieser Stelle ist wichtig zu bemerken, dass es keine Rolle spielt, ob die Auf- oder Abwärtsschritte betrachtet werden, denn es gilt

$$\binom{n}{h} = \binom{n}{n-h} = \binom{n}{v} = \binom{n}{n-v}. \quad (2.6)$$

Für die Wahrscheinlichkeit  $w_n(a, b)$ , dass der 'Wanderer' in  $n$  Schritten von  $a$  nach  $b$  gelangt gilt [1]

$$w_n(a, b) = \binom{n}{h} p^h q^v = \binom{n}{h} \left(\frac{1}{2}\right)^n, \quad (2.7)$$

wobei

$$n = h + v, \quad h - v = b - a, \quad h = \frac{n+b-a}{2} \quad \text{und} \quad v = \frac{n+a-b}{2} \quad \text{gilt. [1]}$$

Das bedeutet also, dass die Gesamtanzahl Schritte die Summe der Auf- und Abwärtsschritte ist. Möchte man statt der Wahrscheinlichkeit eines Weges die absolute Anzahl Wege beschreiben, gilt [1]

$$|W_n(a, b)| = \binom{n}{h}. \quad (2.8)$$

### 2.1.2 Spiegelungsprinzip

In folgendem Abschnitt steht die Symmetrie als Eigenschaft des gewählten Modells der random walks im Mittelpunkt. Es soll gezeigt werden, dass Spiegelungen als bijektive Abbildungen möglich sind und den Beweisen der Eigenschaften der random walks meist als Basis dienen.

Beweise im Zusammenhang mit random walks bedienen sich oft dem Spiegelungsprinzip, wesshalb dieses hier präsentiert werden soll. Abzählungsprobleme lassen sich auf diese Weise simplifizieren. Mathematisch gesehen steht das 'Spiegelungsprinzip' für eine bijektive Abbildung zwischen zwei Mengen von Wegen, welche als  $W_1$  und  $W_2$  bezeichnet werden. Die Abbildungsvorschrift ist dabei jeweils eine problemspezifische Achsenspiegelung. Dabei muss nicht zwingend der ganze Weg gespiegelt werden, oft werden bestimmte Teilwege an einer Spiegelungsachse gespiegelt. Ein einfaches Beispiel für eine Anwendung des Spiegelungsprinzips ist die Spiegelung an der x-Achse. Die Abbildung ordnet also einem Tupel  $(a_1, a_2, \dots, a_n)$  das Tupel  $(-a_1, -a_2, \dots, -a_n)$  zu.

## 2.2 Nullstellen

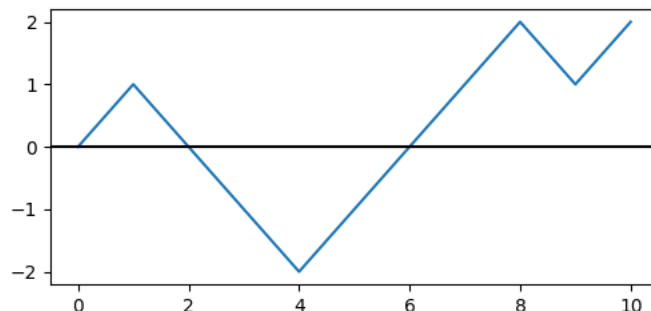
Dieses Kapitel ist den Nullstellen der random walks gewidmet. Nullstellen sind im Zusammenhang mit random walks, aber auch für die Analyse und Interpretation von Aktienkursen von grosser Bedeutung. Um relevante Eigenschaften bezüglich Nullstellen zu ergründen, steht das Hauptlemma als zentrales Werkzeug zur Verfügung. Andererseits begegnet man dem Spiegelungsprinzip, das als Hilfsmittel in den Beweisen dient. Folgend auf eine allgemeine Definition von Nullstellen bei Aktien und random walks werden die nötigen Werkzeuge ergründet, welche schliesslich für die Herleitung der Eigenschaften wichtig sind. Im letzten Teil werden theoretische Werte, Werte aus der Simulation sowie Werte der Aktienkurse gegenübergestellt und in Diagrammen präsentiert.

### 2.2.1 Nullstellen bei Aktienkursen

Die Art von Nullstellen, die in dieser Arbeit betrachtet wird, bezieht sich nicht darauf, dass der Aktienkurs den Wert 0 annimmt, sondern setzt nach der Normalisierung an, wo die Nullachse auf Höhe des Anfangswerts liegt. Es wird also der Zeitpunkt gesucht, wo die Aktie gleich bewertet ist wie im gewählten Anfangspunkt. Um diese Art von Nullstellen der Aktienkurse zu bestimmen, muss zusätzlich zum Fall, dass die Aktie am Tag  $n$  mit  $n \geq 1$  den Anfangswert der Aktie hat, auch der Fall, dass der normalisierte Aktienkurs vom Tag  $n$  zum Tag  $n+1$  das Vorzeichen wechselt, betrachtet werden. Im Gegensatz zum random walk auf der Menge der ganzen Zahlen, betragen die Sprünge der Aktienkurse von einem auf den nächsten Tag nicht ausschliesslich 1 oder  $-1$ , sondern können unterschiedlich gross sein. Daher liegt auch eine Nullstelle vor, wenn ein normalisierter Aktienkurs von einem auf den nächsten Tag das Vorzeichen wechselt. Im Gegensatz zum gewählten Modell des random walks ist nicht nur nach einer geraden Anzahl Schritte eine Nullstelle möglich, sondern nach jedem Tag  $n$  mit  $n \geq 1$ .

### 2.2.2 Nullstellen bei Random Walks auf $\mathbb{Z}$

Es handelt sich bei einem Zeitpunkt  $2k$  eines random walks um ein Nullstelle eines Weges der Länge  $2n$ , wenn  $1 \leq k \leq n$  und  $S_{2k} = 0$  gelten. Dies ist nur möglich, wenn die Anzahl Schritte bis zur Nullstelle gerade ist, weil eine Nullstelle nur entstehen kann, wenn ein random walk bis zur Nullstelle die Länge  $2k$  hat und aus genau  $k$  Auf- und  $k$  Abwärtsschritten besteht [7].



**Abbildung 2.3** random walk mit Nullstelle bei  $2k = 2$  und  $2k = 6$

Der Fall, dass an der Stelle  $2k$  eine Nullstelle auftritt, wird als  $Z_{2k}$  beschrieben. Die Wahrscheinlichkeit  $z_{2k}$ , dass ein random walk an der Stelle  $2k$  eine Nullstelle hat, ist gegeben durch [7]

$$z_{2k} = \mathbb{P}(F_{2k}). \quad (2.9)$$

### 2.2.3 Das Hauptlemma

In folgendem Abschnitt soll das Hauptlemma präsentiert werden, welches später als Grundlage für weitere Beweise dienen wird. Es wird das Prinzip des Abzählens der Wege verwendet, um die Gültigkeit des Hauptlemmas zu zeigen. Die Länge eines Weges wird oft mit  $2n$  statt mit  $n$  beschrieben, da Nullstellen nur nach einer geraden Anzahl Schritten möglich sind und daher Brückenwege nur nach  $2n$  Schritten möglich sind. Um das Hauptlemma verständlich präsentieren zu können, werden zuerst sechs Kategorien vorgestellt, welchen Wege der Länge  $2n$  mit entsprechenden Eigenschaften zugeordnet werden. Es wird unterschieden zwischen [7]

- **positiven Wege**  $W_{2n}^{>0}$ : Der 'Wanderer' berührt die x-Achse am Startpunkt, verlässt diese in die positive Richtung und kehrt niemals zur x-Achse zurück. Somit ist ein Weg positiv, wenn  $s_1 > 0, s_2 > 0, \dots, s_n > 0$  gilt.
- **negativer Weg**  $W_{2n}^{<0}$ : Der 'Wanderer' befindet sich ständig unterhalb der x-Achse und berührt diese, abgesehen vom Startpunkt, nie. Es gilt also  $s_1 < 0, s_2 < 0, \dots, s_n < 0$ .
- **nullstellenfreie Wege**  $W_{2n}^{\neq 0}$ : Ein nullstellenfreier Weg zeichnet sich dadurch aus, dass  $s_1 \neq 0, s_2 \neq 0, \dots, s_n \neq 0$  gilt. Ein nullstellenfreier Weg ist immer auch ein positiver oder ein negativer Weg.
- **nichtnegativen Weg**  $W_{2n}^{\geq 0}$ : Der 'Wanderer' befindet sich immer auf der x-Achse oder darüber. Es gilt  $s_1 \geq 0, s_2 \geq 0, \dots, s_n \geq 0$ .
- **nichtpositiven Weg**  $W_{2n}^{\leq 0}$ : Der 'Wanderer' befindet sich immer auf der x-Achse oder darunter, daher muss  $s_1 \leq 0, s_2 \leq 0, \dots, s_n \leq 0$  gelten
- **Brückenwege**  $W_{2n}^0$ : Diese Art von Wegen, zeichnet sich dadurch aus, dass sie sich am Ende

auf der Position 0 befindet, genau wie am Anfang. Beschrieben wird dies mit  $s_n = 0$ . Da eine solche Nullstelle nur nach einer geraden Anzahl Schritten möglich ist, wird im Zusammenhang mit Brückenwegen meist mit  $2n$  statt  $n$  gearbeitet.

Ein zentraler Aspekt in vielen Herleitungen der Eigenschaften von random walks ist das folgende Hauptlemma. Zentral ist dessen Rolle als Hilfssatz. Abgesehen davon lohnt es sich aber durchaus, die eigentliche Aussage dieses Lemmas zu betrachten. Diese Zusammenhänge sind eventuell nicht intuitiv auf den ersten Blick erklärbar.

**Hauptlemma 2.1.** *Für die Anzahl der Wege gilt, [5]*

$$|W_{2n}^0| = |W_{2n}^{\geq 0}| = |W_{2n}^{\neq 0}| = \binom{2n}{n}, \quad n \geq 1. \quad (2.10)$$

Das Hauptlemma besagt also, dass es gleich viele Wege der Länge  $2n$  gibt, die Brückenwege, nichtnegativ oder nullstellenfrei sind. An dieser Stelle werden wird diese interessante Tatsache so zu Kenntnis genommen. Interessierte Leser können sich gerne dem abwechslungsreichen Beweis in [7] auf Seiten 13 bis 17 widmen.

Unter Annahme des Hauptlemmas, reicht die Bestimmung der Wahrscheinlichkeit für einen Brückenweg der Länge  $n$  aus, um daraus die identischen Wahrscheinlichkeiten der anderen, im Hauptlemma erwähnten, Wege zu kennen.

**Lemma 2.1.** [5] Für die Wahrscheinlichkeit  $z_{2n}$ , dass es sich bei einem Weg der Länge  $2n$  um einen Brückenweg, nullstellenfreien Weg respektive nichtnegativen Weg handelt, gilt

$$z_{2n} = \frac{1}{2^{2n}} \binom{2n}{n} \quad (2.11)$$

**Beweis** Ein Brückenweg startet in 0 und endet nach  $2n$  Schritten ebenfalls in 0. Wir verwenden  $2n$  statt  $n$  für die Längenangabe, da ein Brückenweg nur nach einer geraden Schrittzahl auftreten kann. Nach der Gleichung 2.8 ist bekannt, dass es sich bei

$$|W_{2n}(0, 0)| = |W_{2n}^0| = \binom{2n}{n} \quad (2.12)$$

Wegen um Brückenwege handelt.

Es wird die die mögliche Anzahl Brückenwege, nichtnegative oder nullstellenfreie Wege (2.10) durch die Gesamtanzahl der möglichen Wege mit Länge  $2n$  geteilt (2.3). Der dadurch erhaltene Anteil entspricht der Wahrscheinlichkeit  $z_{2n}$ , weil alle Wege mit gleicher Wahrscheinlichkeit auftreten.

□

## 2.2.4 Anzahl der Nullstellen

Betrachtet man einen random walk mit unendlich vielen Schritten, kehrt dieser mit 100%-iger Wahrscheinlichkeit zu seinem Ursprung zurück und dies passiert unendlich oft. Ein random walk mit unendlicher Länge hat also unendlich viele Nullstellen. Gleichzeitig braucht ein solcher random walk aber auch unendlich lange, um zum Ursprung zurückzukehren. Nun wird die Anzahl der Nullstellen eines endlich langen random walks der Länge  $2n$  betrachtet. Für die Wahrscheinlichkeit  $z_{k,2n}$ , dass ein random walk der Länge  $2n$  genau  $k$  Nullstellen besitzt, gilt [4]

$$\mathbb{P}(N_{2n} = k) = \frac{2^k}{2^{2n}} \binom{2n-k}{n}. \quad (2.13)$$

Damit abgeschätzt werden kann, wie viel Nullstellen ein einfacher, symmetrischer random walk im Durchschnitt hat, wird die Definition des Erwartungswert  $\mathbb{E}(X)$  benötigt. Dieser Wert gibt den Durchschnittswert oder der mittlere Wert einer Zufallsvariable an. Für diskrete Wahrscheinlichkeitsverteilungen, wie sie durch die Gleichung 2.13 gegeben ist, wird der Erwartungswert als

$$\mathbb{E}(X) = \sum_x x \cdot \mathbb{P}(X = x) \quad (2.14)$$

definiert. [10]

Für den Erwartungswert der Anzahl Nullstellen gilt [7]

$$\mathbb{E}(N_{2n}) = (2n + 1) \frac{\binom{2n}{n}}{2^{2n}} - 1. \quad (2.15)$$

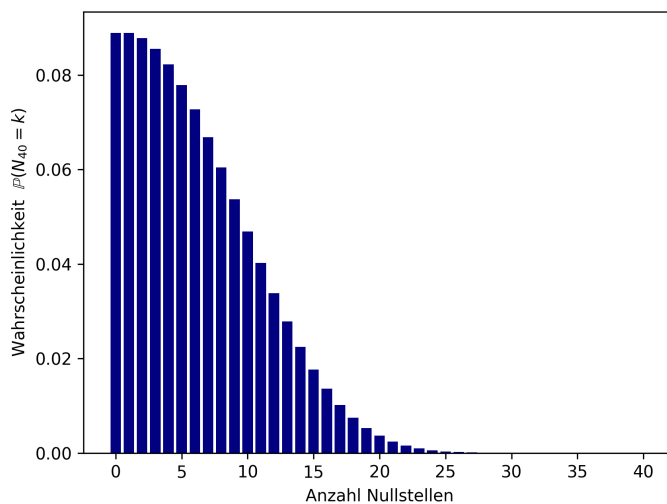
Die Gültigkeit dieser Gleichung kann für das Beispiel eines random walks der Länge  $2n = 254$  gezeigt werden, indem jeweils die mögliche Anzahl Nullstellen mit dessen Wahrscheinlichkeit multipliziert wird und diese 127 Werte schliesslich addiert werden [7]. Es wird also

$$\mathbb{E}(N_{2n}) = \sum_{k=1}^n k z_{k,2n} \quad (2.16)$$

berechnet und mit dem erhaltenen Wert für der Gleichung 2.15 mit  $2n = 254$  verglichen. Es wird deutlich, dass man für beide Berechnungen den gleichen Wert für den Erwartungswert  $\mathbb{E}(N_{2n}) = 11.754$  erhält.

### Verteilung der Anzahl Nullstellen

Durch die Gleichung (2.13) ist die Wahrscheinlichkeit, dass ein random walk der Länge  $2n$  genau  $k$  Nullstellen hat, gegeben. Nun soll die Verteilung der Anzahl Nullstellen genauer betrachtet werden.



**Abbildung 2.4** Wahrscheinlichkeitsverteilung der Anzahl Nullstellen für einen random walk der Länge  $2n = 254$

Das Diagramm 2.4 zeigt die Wahrscheinlichkeitsverteilung der Anzahl Nullstellen eines random walks mit 40 Schritten. Es wird eine starke visuelle Ähnlichkeit mit der positiven Hälfte der Gaussischen Glockenkurve, welche die Dichtefunktion der Normalverteilung darstellt. [10].

### Nullstellenfreie Random Walks

Ein nullstellenfreier random walk bedeutet, dass der 'Wanderer' niemals zu seinem Ausgangswert zurückkehrt. Eine Rückkehr zum Ausgangswert liegt vor, wenn

$$S_n = 0. \quad (2.17)$$

In der gewählten Darstellungsweise bedeutet dies, dass der Polygonzug die x-Achse nie schneidet oder berührt. Die nullstellenfreien random walks können zusätzlich in positive und negative Wege eingeteilt werden, wobei es sich um einen positiven Weg handelt, wenn  $s_1 > 0, s_2 > 0, \dots, s_n > 0$  gilt und um einen negativen Weg, wenn  $s_1 < 0, s_2 < 0, \dots, s_n < 0$  gilt [5].

Die Gleichung für die Berechnung der Wahrscheinlichkeit (2.13), dass ein random walk der Länge  $2n$   $k$  Nullstellen hat, wurde bereits präsentiert. Um nun die Wahrscheinlichkeit  $z_{0,2n}$  eines nullstellenfreien Weges der Länge  $2n$  zu erhalten, wird  $k = 0$  gesetzt und es gilt [7]

$$z_{0,2n} = \frac{1}{2^{2n}} \binom{2n}{n}. \quad (2.18)$$

Durch das Hauptlemma 2.10 ist die Anzahl der nullstellenfreien Wege der Länge  $2n$  bekannt. Um die Wahrscheinlichkeit eines nullstellenfreien Weges zu erhalten, wird durch die Gesamtanzahl der möglichen Wege der Länge  $2n$  dividiert (2.3).

□

#### 2.2.5 Die erste Nullstelle

Als erste Nullstelle wird der Zeitpunkt  $F_{2n}$  beschrieben, an dem der 'Wanderer' erstmals zum Ursprung zurückkehrt und somit in der gewählten Darstellungsweise sich auf der x-Achse befindet. Werden random walks mit unendlich vielen Schritten betrachtet, brauchen diese im Durchschnitt unendlich lange, um zum Ursprung zurückzukehren und somit die erste Nullstelle auftritt [8]. Um die durchschnittliche Position der ersten Nullstelle eines random walks mit endlicher Länge zu finden, müssen zuerst die Eigenschaften eines random walks betrachtet werden. Es gibt genau  $2^{2k}$  Wege der Länge  $2k$  (2.3). Nun werden die Wege in zwei Teile aufgeteilt, ein Teil bis zur ersten Nullstelle und ein zweiter Teil ab der ersten Nullstelle bis zum Ende. Durch diese Unterteilung fällt das Abzählen der möglichen Wege einfacher. Die Wahrscheinlichkeiten der Teilwege können einzeln bestimmt werden und werden dann multipliziert. Die Nullstelle kann jeweils nur nach einer geraden Anzahl Schritten auftreten, somit sind Nullstellen an den Stellen  $2, 4, 6, \dots, 2k$  möglich. Da es bis zu einer Nullstelle genau gleich viele Auf- wie Abwärtsschritte geben muss, wird zuerst bedenkt, auf wie viele Arten man aus  $2k$  Schritten  $k$  Schritte auswählen kann. Dies kann man mit Hilfe des Binomialkoeffizienten tun. Da es sich an der Stelle  $2k$  um die erste Nullstelle handeln soll, müssen noch die Fälle ausgeschlossen werden, wo an der Stelle  $2k$  zwar eine Nullstelle vorkommt, es sich aber nicht um die erste Nullstelle des random walks handelt. Es gibt genau  $2^{2k}$  random walks mit  $2k$  Schritten. Für die Anzahl der random walks  $|W_{2k}|$  mit der ersten Nullstelle bei  $2k$  gilt [5]

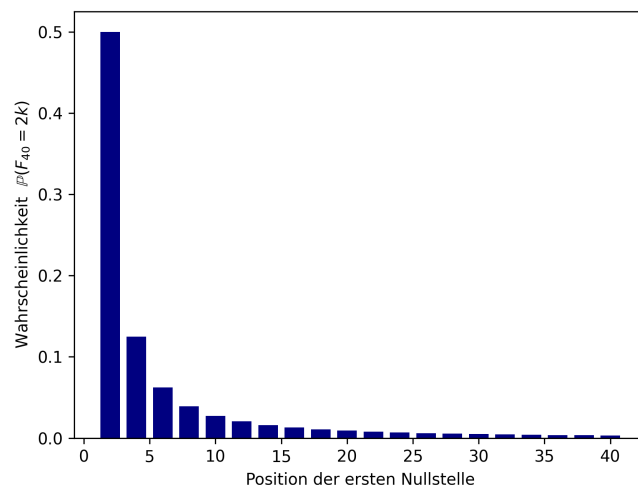
$$|W_{2k}| = \binom{2k}{k} \frac{1}{2k-1}. \quad (2.19)$$

Die Wahrscheinlichkeit, dass die erste Nullstelle nach  $2k$  Schritten auftritt, wird berechnet durch

$$\mathbb{P}(F_{2n} = 2k) = \binom{2k}{k} \left(\frac{1}{2}\right)^{2k} \frac{1}{2k-1}. \quad (2.20)$$



Aus der obigen Gleichung, welche die Wahrscheinlichkeitsverteilung der Position ersten Nullstelle definiert, kann diese Verteilung visualisiert werden. Dazu wird ein Balkendiagramm erstellt.

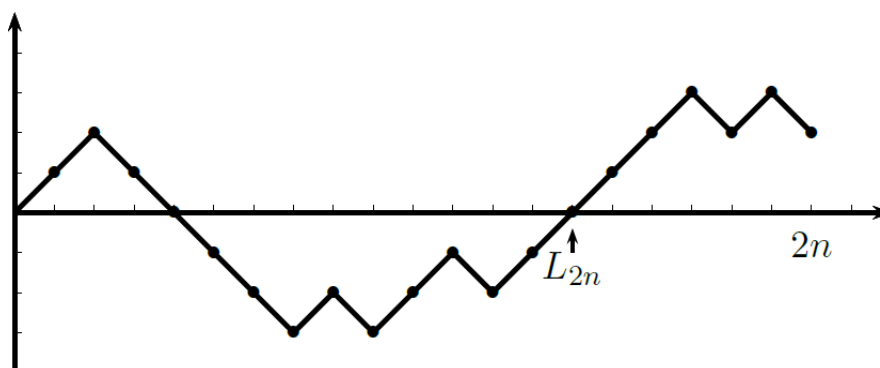


**Abbildung 2.5** Wahrscheinlichkeitsverteilung der ersten Nullstelle  $F_{40}$

Die Wahrscheinlichkeitsverteilung der ersten Nullstelle zeigt einerseits auf, dass der random walk mit hoher Wahrscheinlichkeit bereits nach zwei Schritten sich wieder im Ursprung befindet und somit seine erste Nullstelle aufweist, dies ist mit 50%-iger Wahrscheinlichkeit zu erwarten. Zusätzlich fällt vor allem auf, dass die Wahrscheinlichkeit mit steigender Position der ersten Nullstelle zwar stark abnimmt, trotzdem aufsummiert eine beträchtliche Wahrscheinlichkeit besteht, dass die erste Nullstelle erst relativ spät auftritt.

### 2.2.6 Die letzte Nullstelle

Betrachtet man einen random walk als Münzenwurfspiel zwischen zwei Spielern, so würde man erwarten, dass beide etwa zu einer Hälfte der Zeit das Spiel anführen und es zu relativ häufigen Führungswechseln kommt. Daraus folgend käme man zum Schluss, dass die letzte Nullstelle  $k$ , was im Kontext eines Münzenwurfspiels die gleiche Anzahl an Kopf wie Zahl bedeutet, mit grösserer Wahrscheinlichkeit gegen Ende des Spielverlaufs auftreten müsste. In diesem Fall wird man jedoch von der Intuition in die Irre geführt, denn wie in diesem Abschnitt gezeigt wird, ist die Verteilung der letzten Nullstelle symmetrisch. Die letzte Nullstelle tritt unabhängig von der Spiellänge mit Wahrscheinlichkeit  $1/2$  in der zweiten Spielhälfte auf. Am wahrscheinlichsten tritt die letzte Nullstelle ganz am Anfang oder ganz am Ende des random walks auf [7]. Das folgende Theorem zeigt auf, dass jeweils die letzte Nullstelle mit identischer Wahrscheinlichkeit an der Stelle  $k$  und an der Stelle  $n - k$  auftritt [7].



**Abbildung 2.6** Letzte Nullstelle eines random walks zum Zeitpunkt  $L_{2n}$  [7]

**Theorem 2.1.** [4] Für die Wahrscheinlichkeit, dass die letzte Nullstelle zum Zeitpunkt  $L_{2n}$  auftritt, gilt

$$\mathbb{P}(L_{2n} = 2k) = z_{2k, 2n} = z_{2k}^0 z_{2n-2k}^{\neq 0} = \frac{1}{2^{2n}} \binom{2k}{k} \binom{2n-2k}{n-k}. \quad (2.21)$$

**Beweis** Der Beweis stammt aus [7] auf Seite 18 und 19.

Ein random walk mit der letzten Nullstelle nach  $2k$  Schritten kann in zwei Teile aufgeteilt werden. Der Teilweg der Länge  $2k$  bis zur letzten Nullstelle muss ein Brückenweg sein, damit an der Stelle  $2k$  sich eine Nullstelle befindet. Bei dem Teilweg der Länge  $2n - 2k$  ab der letzten Nullstelle muss es sich um einen nullstellenfreien Weg handeln [7]. Die Wahrscheinlichkeit, dass der Teilweg bis  $2k$  ein Brückenweg ist, ist bekannt durch das 2.1. Die Wahrscheinlichkeit eines Brückenweges der Länge  $2k$  wird mit  $z_{2k}^0$  bezeichnet und beträgt

$$z_{2k}^0 = \frac{1}{2^{2k}} \binom{2k}{k}. \quad (2.22)$$

Für den zweiten, nullstellenfreien Teilweg kann die Wahrscheinlichkeit auf ebenfalls nach 2.1 berechnet werden, da das Hauptlemma besagt, dass es gleich viele Brückenwege und nullstellenfreie Wege gibt. Die Länge des Teilweges ändert sich jedoch und beträgt  $2n - 2k$ . Für die Wahrscheinlichkeit  $z_{2n-2k}^{\neq 0}$ , dass es sich bei einem Weg der Länge  $2n - 2k$  um einen nullstellenfreien Weg handelt, gilt [7]

$$z_{2n-2k}^{\neq 0} = \frac{1}{2^{2n-2k}} \binom{2n-2k}{n-k}. \quad (2.23)$$

Um die Wahrscheinlichkeit  $\mathbb{P}(L_{2n} = 2k)$  zu erhalten, werden die beiden Wahrscheinlichkeiten für die zwei Teilwege multipliziert. Dadurch erhält man das Theorem 2.1.

□

**Theorem 2.2.** [7] Für den Erwartungswert der Zufallsvariable  $L_{2n}$  gilt

$$\mathbb{E}(L_{2n}) = n. \quad (2.24)$$

**Beweis** [7] Bis zur letzten Nullstelle handelt es sich um einen Brückenweg, von da an um einen nullstellenfreien Weg. Es ist bekannt durch das 2.1, dass diese beiden Arten von Wegen bei gleicher Länge mit gleicher Wahrscheinlichkeit auftreten. Um die Wahrscheinlichkeit von  $L_{2n}$  zu erhalten, werden die beiden Wahrscheinlichkeiten für die Teilwege multipliziert. Wenn die Nullstelle an der Stelle  $2n - 2k$  statt an der Stelle  $2k$  auftritt, hat der Brückenweg die Länge  $2n - 2k$  und der nullstellenfreie Weg die Länge  $2k$ , also genau entgegengesetzt zu den Längen in den Gleichungen 2.22 und 2.23. Es gilt also [7]

$$z_{2k}^0 z_{2n-2k}^{\neq 0} = z_{2n-2k}^0 z_{2k}^{\neq 0}. \tag{2.25}$$

Daher ist die Verteilung der Position der letzten Nullstelle symmetrisch zu  $n$  und der Erwartungswert von  $L_{2n}$  liegt bei  $n$  [7].

□

**Theorem 2.3.** *Nach [7] gilt für die Varianz*

$$\mathbb{V}(L_{2n}) = \binom{n+1}{2}. \tag{2.26}$$

Der Beweis dieses Theorems kann in [7] auf Seite 19 nachgelesen werden.

Wird die Verteilung der letzten Nullstelle eines random walks als Säulendiagramm betrachtet, fällt auf, dass es sich um eine U-förmige Verteilung handelt. Der Erwartungswert liegt in der Mitte und besitzt die kleinste Wahrscheinlichkeit. Diese Eigenschaften kann durch das Arcus-Sinus-Gesetz beschrieben werden.

### 2.2.7 Arcus-Sinus-Gesetz

Die Arcus-Sinus-Verteilung tritt im Zusammenhang mit Wahrscheinlichkeitsverteilungen von Eigenschaften der random walks mehrmals auf. An dieser Stelle im Zusammenhang mit der Position der letzten Nullstelle. Wird der Zeitpunkt  $L_{2n}$  durch die Anzahl der Schritte  $2n$  geteilt, erhält man eine Zufallsvariable  $X$ , die auf dem Einheitsintervall  $[1, 0]$  liegt. Diese Zufallsvariable gibt den Bruchteil der gesamten Anzahl vergangener Schritte bis zur letzten Nullstelle an. Die Arcus-Sinus-Verteilung von  $X$  soll nun verdeutlicht werden. [7]

**Theorem 2.4.** *Es gilt für alle  $x$  mit  $0 \leq x \leq 1$  [7]*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{L_{2n}}{2n} \leq x\right) = \int_0^x \frac{1}{\pi \sqrt{(1-t)t}} dt = \frac{2}{\pi} \arcsin \sqrt{x} \tag{2.27}$$

Grosse Fakultäten können mit Hilfe der Stirling-Formel angenähert werden [5], es gilt also

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \tag{2.28}$$

Mit Hilfe der Stirling-Formel kann gezeigt werden, dass

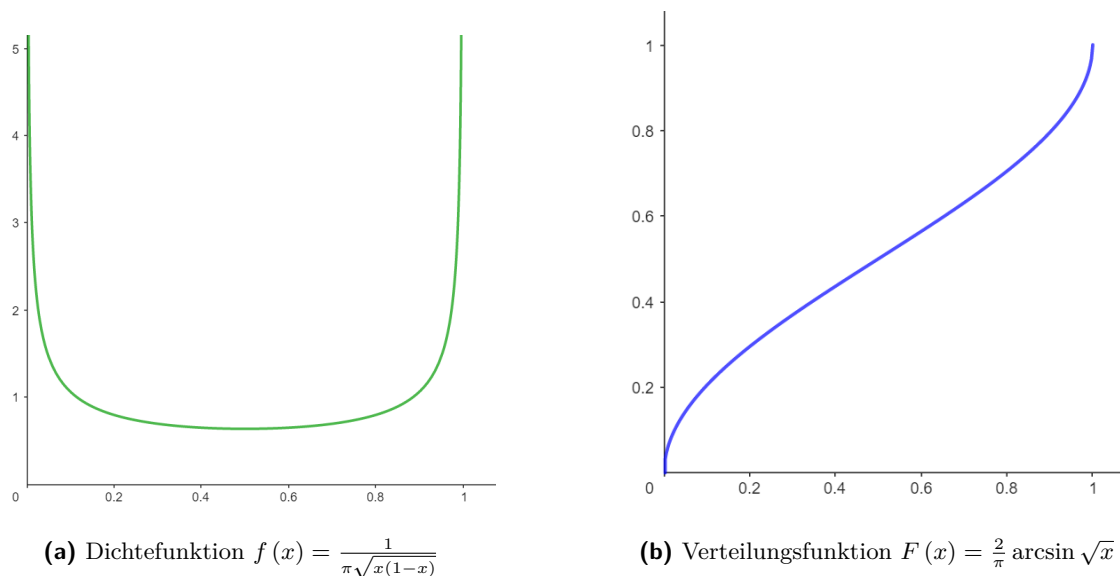
$$z_{2n} \sim \frac{1}{\sqrt{\pi n}} \tag{2.29}$$

gilt, wobei die Notation bedeutet, dass es sich um eine Annäherung handelt, die umso genauer ist, wenn  $n$  sehr gross ist [5]. Gerade dann ist es aufgrund des stark verminderten Rechenaufwands im Vergleich zur Fakultät sinnvoll, die Stirling-Formel zu verwenden.

Um die Verteilung der Wahrscheinlichkeit der Position der letzten Nullstelle anzunähern, wird der Graph der folgenden Funktion verwendet: [7]

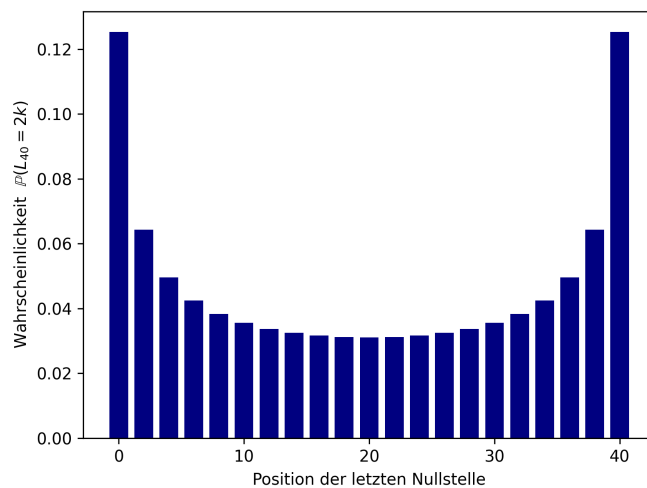
$$z_{2k,2n} \sim \frac{1}{n} f(x) \sim \frac{1}{n \pi \sqrt{x(1-x)}}, \tag{2.30}$$

wobei  $n$  die die halbe Länge eines  $2n$  Schritte langen random walks ist.



**Abbildung 2.7** Dichte- und Verteilungsfunktion der Arcus-Sinus-Verteilung

Die konkrete Wahrscheinlichkeitsverteilung der Position der letzten Nullstelle eines random walks mit 40 Schritten soll nun aufgezeigt werden, indem die einzelnen Wahrscheinlichkeiten der möglichen Positionen der letzten Nullstelle berechnet werden. Dafür wird die Formel 2.1 verwendet. Der Zusammenhang mit der Arcus-Sinus-Verteilung wird visuell eindeutig sichtbar.



**Abbildung 2.8** Wahrscheinlichkeitsverteilung der letzten Nullstelle  $L_{40}$

## 2.3 Extremstellen

In folgendem Kapitel soll die maximale Entfernung der Irrfahrt von ihrem Ursprungsort betrachtet werden. Es soll diese maximale Entfernung sowie die Lage der Extremstellen bei einem symmetrischen random walk auf  $\mathbb{Z}$  statistisch betrachtet werden. Im Zusammenhang mit Aktienkurse sind Extremstellen ein relevantes Merkmal, weil sie einerseits das Ausmass von Kursbewegungen abstecken und so andererseits das Risiko eines Kursverlustes aber auch die Gewinnchancen durch Kursanstiege dargestellt wird.

Das Kapitel beginnt mit einer Definition des Maximums sowie des Minimums bei einfachen, symmetrischen random walks auf  $\mathbb{Z}$ . Anschliessend wird die Wahrscheinlichkeit, dass das Maximum eines solchen random walks einen gewissen Wert annimmt, präsentiert. Im Zusammenhang mit random walks mit fester Schrittgrösse wäre es durchaus sinnvoll, zu betrachten, wie oft im Verlauf eines random walks dieses Maximum erreicht wird. Da Aktienkurse keine feste Schrittlänge haben und ein mehrmaliges Auftreten des Maximum nur sehr unwahrscheinlich ist, wird dieses Merkmal im Rahmen dieser Arbeit vernachlässigt. Das selbe gilt natürlich in gleicher Weise für das Minimum. Durchaus von Interesse ist hingegen die Lage der Extremstellen, daher wird im Abschnitt (2.3.2) die Position der ersten Maximalstelle diskutiert. Im letzten Abschnitt des Kapitels wird, wie bereits im vorhergehenden Kapitel zu den Nullstellen, einen Vergleich zwischen den aus der Theorie dieses Kapitels erhaltenen Erkenntnissen, der Auswertung der Simulation sowie der Auswertung der Aktienkurse bezüglich der Extremstellen gezogen.

### 2.3.1 Maximum und Minimum

Es wird unterschieden zwischen Minimum und Maximum. Das Maximum wird definiert durch

$$M_n := \max\{S_0, S_1, \dots, S_n\} \quad (2.31)$$

und für das Minimum gilt

$$m_n := \min\{S_0, S_1, \dots, S_n\}. \quad (2.32)$$

Im folgenden Abschnitt sowie in den anderen Abschnitten dieses Kapitels wird nur die Verteilung des Maximums betrachtet, da sich bei einem symmetrischen random walk die Verteilung des Maximums von jener des Minimums nur bezüglich des Vorzeichens unterscheidet. Daher gilt für die Zufallsvariable des Minimums  $m_n$  [7]

$$m_n = -M_n. \quad (2.33)$$

Es wird definiert, dass

$$\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\} \quad (2.34)$$

gelten soll, wobei es sich dabei um die Gauss-Klammern handelt, welche einen Wert auf die nächstkleinere Ganzzahl abrundet.

**Theorem 2.5.** *Für das Maximum  $M_n$  eines symmetrischen random walks gilt [7]*

$$\mathbb{P}(M_n = k) = \frac{1}{2^n} \binom{n}{\lfloor \frac{n+k+1}{2} \rfloor}, \quad n \leq 1; \quad k = 0, 1, \dots, n. \quad (2.35)$$

**Beweis** Es wird der Herleitung in [7] auf Seite 49 gefolgt.

Das Maximum liegt immer im Bereich  $0 \leq M_n \leq n$ . Um die Verteilung des Maximums herzuleiten, werden zuerst die beiden Extremfälle betrachtet, dessen Wahrscheinlichkeit schon durch vorhergehende Überlegungen im Kapitel zu den Nullstellen bekannt ist. Einerseits gibt es den Fall, dass ein Weg ausschliesslich aus Aufwärtsschritten besteht. Aus  $n$  Schritten sollen also  $n$  Aufwärtsschritte

ausgewählt werden, dazu gibt es genau eine Möglichkeit. Es ist gegeben durch Gleichung 2.3, dass es genau  $2^n$  Wege der Länge  $n$  gibt. Daraus folgt für die Wahrscheinlichkeit eines Weges mit  $n$  Aufwärtsschritten und somit einem Maximum von  $M_n = n$  [7]

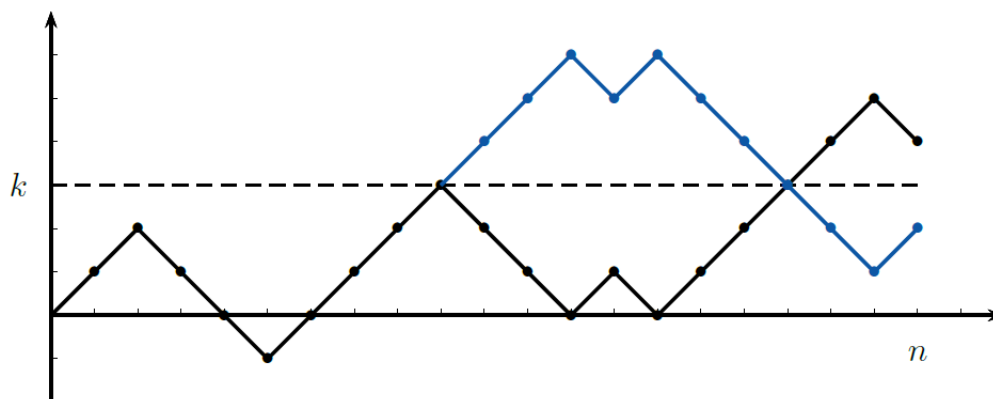
$$\mathbb{P}(M_n = n) = \frac{1}{2^n}. \tag{2.36}$$

Das kleinstmögliche Maximum liegt bei  $M_n = 0$ , wobei es sich in diesem Fall um einen nichtpositiven random walk handelt, dessen Wahrscheinlichkeit durch das Hauptlemma (2.11) gegeben ist [7].

Es soll nun aber eine Herleitung für die Wahrscheinlichkeit aller möglichen Werte, die das Maximum annehmen kann, gezeigt werden. Schliesslich wird deutlich, dass die Verteilung des Maximums direkt im Zusammenhang mit der Verteilung der Endposition  $S_n$  steht [7]. Es soll die Wahrscheinlichkeit  $\mathbb{P}(M_n = k)$  ergründet werden. Dazu wird das Ereignis  $\{M_n \geq k\}$  in drei verschiedene Fälle zerlegt, wobei für  $k$   $0 \leq k \leq n$  gelten muss. Es wird unterschieden, ob  $S_n > k$ ,  $S_n < k$  oder  $S_n = k$ . Es wird also eine Verbindung zwischen Maximum und Endposition hergestellt und man erhält dadurch [7]

$$\mathbb{P}(M_n \geq k) = \mathbb{P}(M_n \geq k, S_n > k) + \mathbb{P}(M_n \geq k, S_n < k) + \mathbb{P}(M_n \geq k, S_n = k). \tag{2.37}$$

Diese Zerlegung in drei verschiedene Fälle bezüglich der Endposition ist insofern sinnvoll, da nun festgestellt werden kann, dass die beiden Wahrscheinlichkeiten  $\mathbb{P}(M_n \geq k, S_n > k)$  und  $\mathbb{P}(M_n \geq k, S_n < k)$  identisch sind und Wege, die die einen Bedingungen erfüllen können injektiv und eindeutig einem Weg zugeordnet werden, der die anderen Bedingungen erfüllt [7]. Wird beispielsweise ein Weg betrachtet, der den ersten Eigenschaften  $M_n \geq k, S_n > k$  entspricht, wird dieser in zwei Teilwege unterteilt. Der Teilweg ab dem erstmaligen Erreichen der Höhe  $k$  wird durch eine Achsenspiegelung an der Geraden  $y = k$  gespiegelt. Somit entsteht ein Weg, der den Eigenschaften  $M_n \geq k, S_n < k$  entspricht. Man stellt fest, dass es sich um eine injektive Abbildung handelt und genau gleich viele Wege mit jeweils diesen Eigenschaften existieren, somit ist bewiesen, dass diese Wahrscheinlichkeiten identisch sind [7].



**Abbildung 2.9** Spiegelung des Teilweges ab erstmaligem Erreichen der Höhe  $k$  an der Achse  $y = k$  [7]

Erfüllt ein Weg die Bedingung  $S_n \geq k$  oder  $S_n = k$ , gilt auch für das Maximum zwingend  $M_n \geq k$  [7]. Auf Basis der gewonnenen Erkenntnisse kann die Gleichung 2.37 umgeformt werden [7]:

$$\mathbb{P}(M_n \geq k) = 2\mathbb{P}(M_n \geq k, S_n > k) + \mathbb{P}(M_n \geq k, S_n = k) \quad (2.38)$$

$$= 2\mathbb{P}(M_n \geq k, S_n \geq k) - \mathbb{P}(M_n \geq k, S_n = k) \quad (2.39)$$

$$= 2\mathbb{P}(S_n \geq k) - \mathbb{P}(S_n = k) \quad (2.40)$$

$$= \mathbb{P}(S_n \geq k) + \mathbb{P}(S_n \geq k + 1) \quad (2.41)$$

Da durch die injektive Abbildung klar wurde, dass die beiden ersten Wahrscheinlichkeiten auf der rechten Seite der Gleichung (2.37) identisch sind, wird im ersten Schritt (2.38) dies zusammengefasst. Im zweiten Schritt (2.39) wird der Fall  $S_n = k$  in die erste Wahrscheinlichkeit einbezogen, da diese doppelt vorkommt, wird genau wiederum die Wahrscheinlichkeit, dass  $S_n = k$  gilt, subtrahiert. Im dritten Schritt (2.40) wird die Gleichung so vereinfacht, dass  $M_n \geq k$  weggelassen wird, weil die  $M_n \geq k$  zwingend gelten muss, wenn  $S_n \geq k$  oder  $S_n = k$  gegeben ist. Schliesslich wird im vierten Schritt (2.41) nochmals so umgeformt, dass die Wahrscheinlichkeit  $\mathbb{P}(S_n \geq k)$  und die Wahrscheinlichkeit  $\mathbb{P}(S_n \geq k + 1)$  addiert werden. Durch die einheitliche Schrittlänge von 1 kann  $k$  nur ganzzahlig sein, schliesslich ergibt sich dadurch, dass  $\mathbb{P}(S_n \geq k) - \mathbb{P}(S_n = k)$  und  $\mathbb{P}(S_n \geq k + 1)$  identisch sind.[7]

Um die Wahrscheinlichkeit  $\mathbb{P}(M_n = k)$  zu erhalten, wird von  $\mathbb{P}(M_n \geq k)$  die Wahrscheinlichkeit  $\mathbb{P}(M_n \geq k + 1)$  subtrahiert wodurch nur der gesuchte Fall übrig bleibt [7]. Man erhält durch Umformung [7]

$$\mathbb{P}(M_n = k) = \mathbb{P}(M_n \geq k) - \mathbb{P}(M_n \geq k + 1) \quad (2.42)$$

$$= (\mathbb{P}(S_n \geq k) + \mathbb{P}(S_n \geq k + 1)) - (\mathbb{P}(S_n \geq k + 1) + \mathbb{P}(S_n \geq k + 2)) \quad (2.43)$$

$$= \mathbb{P}(M_n \geq k) - \mathbb{P}(M_n \geq k + 2) \quad (2.44)$$

Subtrahiert man von der Wahrscheinlichkeit, dass die Endposition  $S_n$  grösser oder gleich  $k$  ist, die Wahrscheinlichkeit  $\mathbb{P}(M_n \geq k + 2)$ , bleiben noch die Wahrscheinlichkeiten für die Fälle  $S_n = k$  und  $S_n = k + 1$  übrig. Es gilt also [7]

$$\mathbb{P}(M_n = k) = \mathbb{P}(S_n = k) + \mathbb{P}(S_n = k + 1). \quad (2.45)$$

Die Endposition eines random walks auf  $\mathbb{Z}$  mit Schrittlänge 1 muss nach einer geraden Anzahl Schritten auch gerade sein und nach einer ungeraden Anzahl Schritte dementsprechend ungerade.  $n$  und  $k$  müssen also die gleiche Parität besitzen und dies führt dazu, dass jeweils eine der summierten Wahrscheinlichkeiten der Gleichung (2.45) gleich 0 sein wird [7]. In Gleichung (2.7) wurde die Wahrscheinlichkeit für einen Weg vom Punkt  $a$  nach  $b$  aufgezeigt. In diesem Fall soll der Weg vom Ursprung nach  $P(n/S_n)$  führen. Diese Wahrscheinlichkeiten sind gegeben durch [5]

$$\mathbb{P}(S_n = k) = \frac{1}{2^n} \binom{n}{\frac{n+k}{2}} \quad (2.46)$$

und

$$\mathbb{P}(S_n = k + 1) = \frac{1}{2^n} \binom{n}{\frac{n+k+1}{2}}. \quad (2.47)$$

Wie bereits oben erläutert, ist immer eine der Wahrscheinlichkeiten gleich 0 und um diese Summe der Wahrscheinlichkeiten etwas kompakter auszudrücken, wird die Notation der Gauss-Klammer verwendet [7]:

$$\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\} \quad (2.48)$$

Es handelt sich dabei um eine Abrundungsfunktion, die eine Zahl auf die nächstkleinere ganze Zahl abrundet. Dies ist für diese Gleichung daher sinnvoll, da die Notation verkürzt werden kann [7] :

$$\mathbb{P}(M_n = k) = \frac{1}{2^n} \binom{n}{\lfloor \frac{n+k+1}{2} \rfloor}. \quad (2.49)$$

Diese Notation deckt beide möglichen Paritäten von  $n$  und  $k$  ab. Im Fall, dass  $n$  und  $k$  ungerade sind, handelt es sich bei dem Term  $\frac{n+k+1}{2}$  um eine Ganzzahl und die Gauss-Klammer verändert nichts. Liegt eine ungerade Parität vor, ist derselbe Term keine ganze Zahl. Durch das Abrunden verschwindet aber dieser Einfluss des  $+1$  und das Resultat dieses Terms entspricht dem Term  $\frac{n+k}{2}$ .

□

Nachdem die Verteilung der Wahrscheinlichkeit für das Maximum eines symmetrischen random walks erläutert wurde, werden hier die Gleichungen für den Erwartungswert und die Varianz des Maximums  $M_{2n}$  präsentiert.

**Theorem 2.6.** *Das Maximum eines symmetrischen random walks ist charakterisiert durch [7]*

- *den Erwartungswert*

$$\mathbb{E}(M_{2n}) = \left(2n + \frac{1}{2}\right) \frac{\binom{2n}{n}}{2^{2n}} - \frac{1}{2}, \quad (2.50)$$

$$\mathbb{E}(M_{2n+1}) = (2n + 1) \frac{\binom{2n}{n}}{2^{2n}} - \frac{1}{2}. \quad (2.51)$$

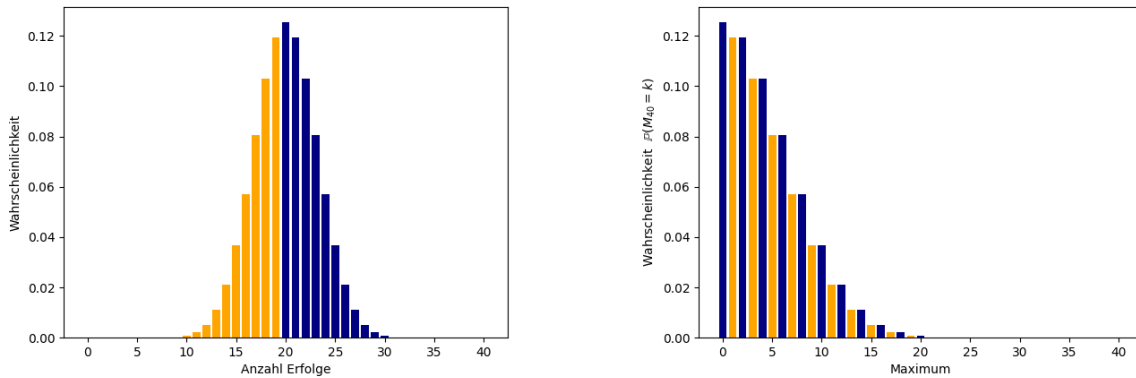
- *die Varianz*

$$\mathbb{V}(M_{2n}) = 2n + \frac{1}{4} - \left( \left(2n + \frac{1}{2}\right) \frac{\binom{2n}{n}}{2^{2n}} \right)^2, \quad (2.52)$$

$$\mathbb{V}(M_{2n+1}) = 2n + \frac{5}{4} - \left( (2n + 1) \frac{\binom{2n}{n}}{2^{2n}} \right)^2. \quad (2.53)$$

Auf einen Beweis dieser Gleichungen für den Erwartungswert und die Varianz wird in dieser Arbeit verzichtet, da dies den Rahmen sprengen würden. Diese Aussagen werden im Buch [7] auf Seite 51 bewiesen und können dort nachgeschlagen werden. Da diese Gleichungen nur im Zusammenhang mit random walks der Länge  $n = 254$  verwendet werden, wurden diese Werte, die man durch die Gleichungen erhält, verglichen mit Mittelwert und Varianz, die mit Hilfe der Programmiersprache 'Python' und der bewiesenen Wahrscheinlichkeitsverteilung für das Maximum berechnet wurden. Es konnte eine Übereinstimmung festgestellt werden. Gleichzeitig wurde auch für die selbe Schrittzahl ein Verteilungsdiagramm erstellt.





(a) Binomialverteilung mit  $n = 254$  und  $p = 0.5$

(b) Diagramm zur Verteilung der Wahrscheinlichkeit des Maximums  $k$

**Abbildung 2.10** Vergleich zwischen der Binomialverteilung (2.10a) und der Verteilung der Wahrscheinlichkeit des Maximums  $M_n$  (2.10b)

Die Abbildung (2.10) zeigt den Zusammenhang zwischen der Binomialverteilung und der Verteilung des Maximums. Werden die Säulen des Diagramm 2.10a nach der Grösse sortiert und beginnend bei der grössten Säule fortlaufend nummeriert, erhält man die Verteilung des Maximums (2.10b). Die Wahrscheinlichkeit (2.45) ist maximal für  $k = 0$  und anschliessend sind jeweils zwei Wahrscheinlichkeiten gleich gross, da das Diagramm 2.10a eine Symmetrie aufweist und abgesehen vom Maximalwert alle Wahrscheinlichkeiten doppelt auftreten. [7]

### 2.3.2 Lage der Extremstellen

Ein einfacher, symmetrischer random walk auf  $\mathbb{Z}$  kann ein Maximum mehrmals annehmen. In diesem Abschnitt liegt der Fokus auf der ersten Maximalstelle und deren Wahrscheinlichkeitsverteilung. Die Beweisidee zur Herleitung dieser Wahrscheinlichkeit stammt aus [7] auf Seite 64 ff..

Das erste Maximum wird definiert als [7]

$$E_{2n} := \{k \in \{0, 1, \dots, 2n\} : S_k = \max_{0 \leq j \leq 2n} S_j\}. \tag{2.54}$$

Zuerst werden die beiden Extremfälle betrachtet. Frühstmöglich kann das Maximum im Ursprung angenommen werden, in diesem Fall gilt  $E_{2n} = 0$ . Mit Bezug zum Hauptlemma kann festgestellt werden, dass dies genau dann der Fall ist, wenn es sich um einen nichtpositiven Weg handelt. Dadurch ergibt sich auch für diesen Extremfall die entsprechende Wahrscheinlichkeit aus dem Hauptlemma.[7] Sie entspricht  $u_{2n} = \binom{2n}{n}/2^{2n}$ . Der entgegengesetzte Extremfall entsteht, wenn  $E_{2n} = 2n$ . Damit dieser Fall auftritt muss  $S_0 < S_{2n}, S_1 < S_{2n}, \dots, S_{2n-1} < S_{2n}$  gelten. Zwischen jedem Punkt des random walks und dem Endpunkt des Weges muss ständig eine positive Differenz zwischen Auf- und Abwärtsschritten herrschen. Mathematisch ausgedrückt bedeutet dies, dass für alle  $j = 1, 2, \dots, 2n - 1$  die Summe  $\sum_{i=j}^{2n-1} X_i > 0$  gelten muss [7]. Per Definition des einfachen, symmetrischen random walk auf  $\mathbb{Z}$  sind die einzelnen Schritte  $X_1, X_2, \dots, X_{2n}$  des random walks unabhängig voneinander und identisch verteilt. Durch diese Erkenntnisse wird klar, dass die Wahrscheinlichkeit dieses Extremfalls identisch mit der Wahrscheinlichkeit des Ereignis  $\{X_1 > 0, X_1 + X_2 > 0, \dots, X_1 + \dots + X_{2n} > 0\}$  ist [7]. Es wurde definiert in (2.2), dass die Summe dieser Schritte der Definition von  $S_1, \dots, S_{2n}$  entspricht. Für die Wahrscheinlichkeit gilt schliesslich

$$\mathbb{P}(E_{2n} = 2n) = \mathbb{P}(S_1 > 0, S_2 > 0, \dots, S_{2n} > 0). \tag{2.55}$$

Diese Wahrscheinlichkeit eines positiven Weges ist bereits bekannt und entspricht  $\mathbb{P}(E_{2n} = 2n) = \frac{1}{2}u_{2n}$ . Nun sollen noch all jene Fälle betrachtet werden, die zwischen den beiden Extremfällen liegen, daher die Fälle  $1 \leq k \leq 2n - 1$ . All jene Wege, die ihr erstes Maximum an der Stelle  $k$  annehmen, werden nun in zwei Teilwege unterteilt. Der erste Teilweg führt vom Ursprung und ist  $k$  Schritte lang. Für diesen Teilweg soll gelten [7]

$$A_k := \{S_0 < S_k, \dots, S_{k-1} < S_k\}, \quad (2.56)$$

damit der Teilweg sein einziges Maximum nach  $k$  Schritten erreicht. Die gleiche Überlegung, wie in (2.55) führt zum Schluss, dass dieser Teilweg einem positiver Weg gleicher Länge entsprechen muss und die Wahrscheinlichkeit gegeben ist durch  $\mathbb{P}(S_1 > 0, \dots, S_k > 0)$  [7].

Der zweite Teilweg beginnt ab der Stelle  $k$  und führt bis zum Ende des Weges. In diesem Teil des Weges darf das Maximum, das in  $k$  erreicht wurde, beliebig oft angenommen werden. Es darf jedoch nicht überschritten werden. Dieser Teilweg ist definiert durch die Bedingung [7]

$$B_k := \{S_{k+1} \leq S_k, \dots, S_{2n} \leq S_k\}. \quad (2.57)$$

Ebenfalls durch die Erkenntnisse oben lässt sich schliessen, dass die Wahrscheinlichkeit für einen Teilweg mit diesen Eigenschaften definiert ist durch  $\mathbb{P}(S_1 \leq 0, S_2 \leq 0, \dots, S_{2n-k} \leq 0)$  [7].

Die Wahrscheinlichkeit eines Weges der Länge  $2n$  mit Maximum an Stelle  $k$  ergibt sich also aus den Wahrscheinlichkeiten der beiden Teilwegen [7]:

$$\mathbb{P}(E_{2n} = k) = \mathbb{P}(A_k \cap B_k) \quad (2.58)$$

$$= \mathbb{P}(A_k) \mathbb{P}(B_k) \quad (2.59)$$

$$= \mathbb{P}(S_{k+1} \leq S_k, \dots, S_{2n} \leq S_k) \mathbb{P}(S_1 \leq 0, S_2 \leq 0, \dots, S_{2n-k} \leq 0) \quad (2.60)$$

Die Ereignisse  $A_k$  und  $B_k$  sind stochastisch unabhängig, weil sie unterschiedliche Teilwege mit unabhängigen Schritten betreffen.

Nun sollen die Fälle danach aufgeteilt werden, ob  $k$  gerade oder ungerade ist. Für gerade  $k$  gilt  $k = 2l$  und für ungerade  $k = 2l + 1$ , wobei  $l \in \{1, \dots, n - 1\}$ . Die Wahrscheinlichkeit in 2.60 ergibt sich für gerade  $k$  als [7]

$$\mathbb{P}(S_1 > 0, \dots, S_{2l} > 0) = \frac{1}{2} \mathbb{P}(S_1 \neq 0, \dots, S_{2l} \neq 0) = \frac{1}{2} u_{2l}, \quad (2.61)$$

$$\mathbb{P}(S_1 \leq 0, \dots, S_{2n-2l} \leq 0) = u_{2(n-l)}. \quad (2.62)$$

Die erste Umformung kann begründet werden, dass der erste Teilweg positiv sein muss. Da die Menge der nullstellenfreien Wegen zu gleichen Teil in positive und negative Wege aufgeteilt werden kann, liegt die Wahrscheinlichkeit eines positiven Teilweges bei der halben Wahrscheinlichkeit für einen nullstellenfreien Weg gleicher Länge [5]. Die zweite Umformung begründet sich ebenfalls durch das Hauptlemma. Die Wahrscheinlichkeit eines nichtpositiven Weges der Menge  $2n - 2l$  ist gegeben durch  $u_{2(n-l)}$ .

Nun soll der Fall mit ungeradem  $k$  betrachtet werden. Wiederum ist die Wahrscheinlichkeit, dass ein Weg der Länge  $k$  positiv ist, gesucht. Der Unterschied liegt darin, dass dieser positive Weg eine ungerade Schrittzahl hat. Es ist bekannt, dass eine Nullstelle nur nach einer geraden Anzahl Schritte auftreten kann, wenn der Weg im Ursprung startet. Ist nun ein Weg der geraden Länge  $k - 1$  positiv, muss auch ein positiver Weg der Länge  $k$  dieselbe Wahrscheinlichkeit besitzen, da durch diesen zusätzlichen, ungeraden Schritt die Bedingungen für einen positiven Weg nicht gebrochen werden kann, weil keine Nullstelle entstehen kann [7].

Der zweite Teil ist ebenfalls identisch, mit dem Fall für gerade  $k$  weil ein Weg die Bedingung für

Nichtpositivität nur nach einer ungeraden Anzahl Schritte brechen kann, weil dies nur in einem auf eine Nullstelle folgenden Schritt passieren kann. Daher lässt sich das erste Gleichheitszeichen in 2.64 erklären. Für ungerade  $k$  gilt

$$\mathbb{P}(S_1 > 0, \dots, S_{2l+1} > 0) = \frac{1}{2}u_{2l}, \tag{2.63}$$

$$\mathbb{P}(S_1 \leq 0, \dots, S_{2n-2l-1} \leq 0) = \mathbb{P}(S_1 \leq 0, \dots, S_{2n-2l} \leq 0) = u_{2(n-l)}. \tag{2.64}$$

**Theorem 2.7.** *Es gilt für den Zeitpunkt  $E_{2n}$  der ersten Maximalstelle [7]*

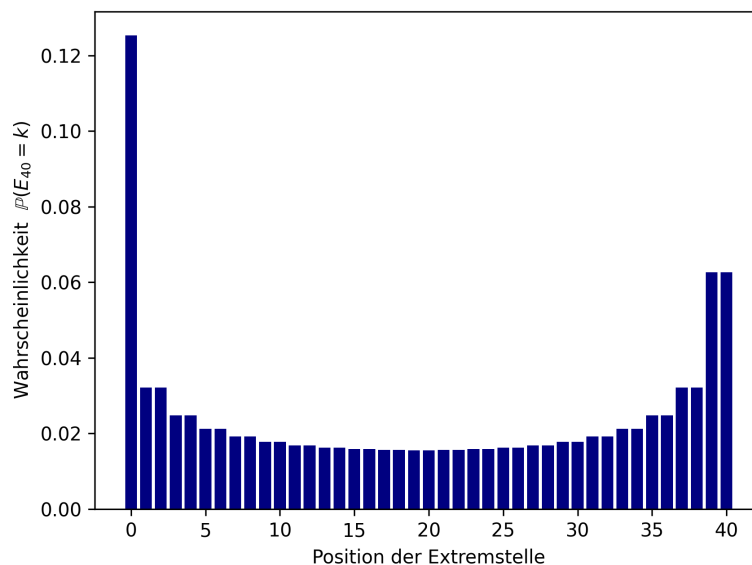
$$\mathbb{P}(E_{2n} = 0) = u_{2n}, \tag{2.65}$$

$$\mathbb{P}(E_{2n} = 2l) = \frac{1}{2}u_{2l}u_{2(n-l)}, \quad l = 1, \dots, n, \tag{2.66}$$

$$\mathbb{P}(E_{2n} = 2l + 1) = \frac{1}{2}u_{2l}u_{2(n-l)}, \quad l = 0, \dots, n - 1. \tag{2.67}$$

### Arcus-Sinus-Gesetz

Bereits im Zusammenhang mit der Verteilung der letzten Nullstelle (2.2.7) wurde das Arcus-Sinus-Gesetz vorgestellt und bewiesen, dass dieses Gesetz auch im Zusammenhang mit  $E_{2n}$  gilt, soll nun gezeigt werden. In der folgenden Abbildung 2.11 wird die Verteilung des ersten Maximums eines random walks dargestellt.



**Abbildung 2.11** Stabdiagramm zur Verteilung von  $E_{2 \cdot 20}$

Der Zusammenhang zwischen der Verteilung der letzten Nullstelle  $L_{2n}$  und dem ersten Maximum  $E_{2n}$  besteht darin, dass die Wahrscheinlichkeit im Nullpunkt identisch ist. Für jedes  $k = 1, 2, \dots, n$  wird die Wahrscheinlichkeit  $\mathbb{P}(L_{2n} = 2k)$  hälftig auf die Stellen  $2k$  und  $2k + 1$  verteilt [5]. Im Zusammenhang mit der letzten Nullstelle wurden nur gerade  $k$  definiert, da Nullstellen nur nach einer geraden Schrittzahl vorkommen können.

**Theorem 2.8.** *Es gilt für alle  $x$  mit  $0 \leq x \leq 1$  [7]*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{E_{2n}}{2n} \leq x\right) = \int_0^x \frac{1}{\pi \sqrt{(1-t)t}} dt = \frac{2}{\pi} \arcsin \sqrt{x} \tag{2.68}$$

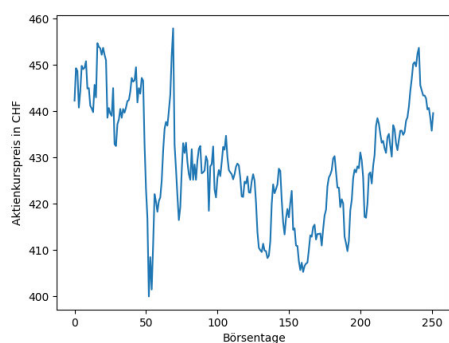
## Kapitel 3

# Modifikation der Random Walks

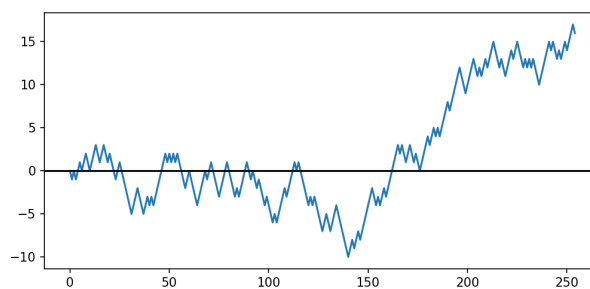
Im folgenden Kapitel soll das Modell der symmetrischen random walks auf  $\mathbb{Z}$  angepasst werden, um Aktienkurse akkurater modellieren zu können. Das Modell der Random Walks mit fixer Schrittlänge eins bietet sich zur Herleitung der theoretischen Eigenschaften der random walks an. Wird jedoch ein Aktienkurs betrachtet, fällt sofort auf, dass dessen Schritte von unterschiedlicher Länge sind. Es liegt daher auf der Hand, dass dies einer Anpassung im bisherigen random-walk-Modell bedarf, die schon rein optisch die random walks den Aktienkursen näher bringt. Es drängt sich aber die Frage auf, welche Eigenschaften eine passende Verteilung der Schritte mitbringen soll. In diesem Kapitel wird einerseits die historische Signifikanz der Wahl einer geeigneten Wahrscheinlichkeitsverteilung in finanziellen Modellen beleuchtet. Weiter wird anhand von echten Daten die passende Wahrscheinlichkeitsverteilung für die Modifizierung des gewählten random-walk-Modells gesucht.

### 3.1 Länge der einzelnen Schritte

Betrachtet man Aktienkurse und random walks auf  $\mathbb{Z}$  mit der selben Länge, fällt auf, dass der random walk mit jedem Schritt genau ein Schritt nach oben oder ein Schritt nach unten macht. Die Aktienkurse hingegen können in der gleichen Zeiteinheit unterschiedlich grosse Schritte nach oben oder unten machen.



(a) Aktienkurs der Zurich Insurance Group Ltd



(b) symmetrischer random walk auf  $\mathbb{Z}$

**Abbildung 3.1** Vergleich zwischen einem Aktienkurs und einem random walk mit gleicher Länge

Statt in der Simulation zufällig aus einem Auf- oder Abwärtsschritt der Länge 1 auswählen zu lassen, soll nun zusätzlich zu zufälligen Vorzeichen auch die Länge des Schrittes zufällig gewählt werden. Für die Zufallsvariable  $X$ , welche die Schrittlänge bestimmt, soll eine geeignete Verteilung gewählt

werden. Dafür gibt es verschiedene Möglichkeiten mit jeweiligen Vor- und Nachteilen. Folgend sollen zwei Möglichkeiten vorgestellt werden.

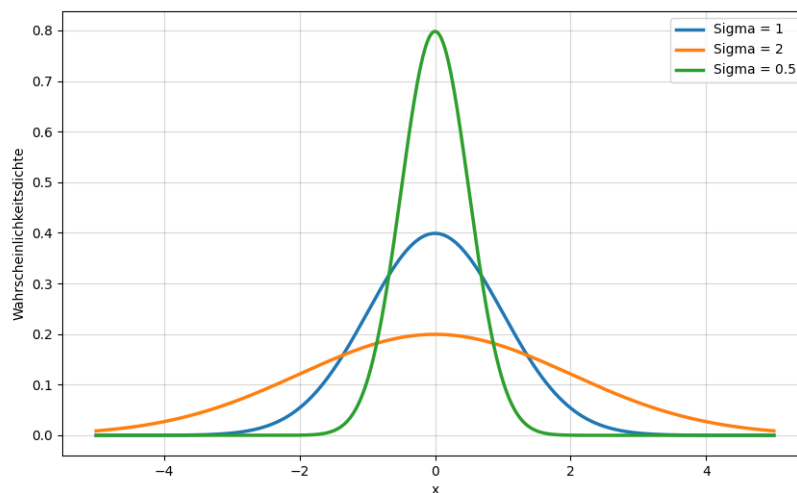
### 3.1.1 Normalverteilte Schrittlänge

Als erste Möglichkeit soll die Schrittlänge durch eine normalverteilte Zufallsvariable modelliert werden. Ein grosser Vorteil der Normalverteilung liegt in ihrer Einfachheit. Sie ist leichter nachvollziehbar, weil sie bekannt und weit verbreitet ist. Die Implementierung in die Simulation gestaltet sich dadurch besonders einfach.

**Normalverteilung** Die Standardverteilung, eine der bekanntesten Wahrscheinlichkeitsverteilungen, wird auch als Gauss-Verteilung bezeichnet und der Graph wird aufgrund der Form oft als 'Glockenkurve' bezeichnet. Die Verteilung weist eine symmetrische Form auf. Es handelt sich um eine kontinuierliche Wahrscheinlichkeitsverteilung, welche durch zwei Parameter gekennzeichnet wird. Der erste Parameter ist der Mittelwert  $\mu$  und der zweite Parameter die Standardabweichung  $\sigma$ . Die Standardabweichung ist ein Mass für die Streuung der Werte um den Mittelwert. Der Graph der Dichteverteilung der Normalverteilung ist durch

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.1)$$

gegeben.



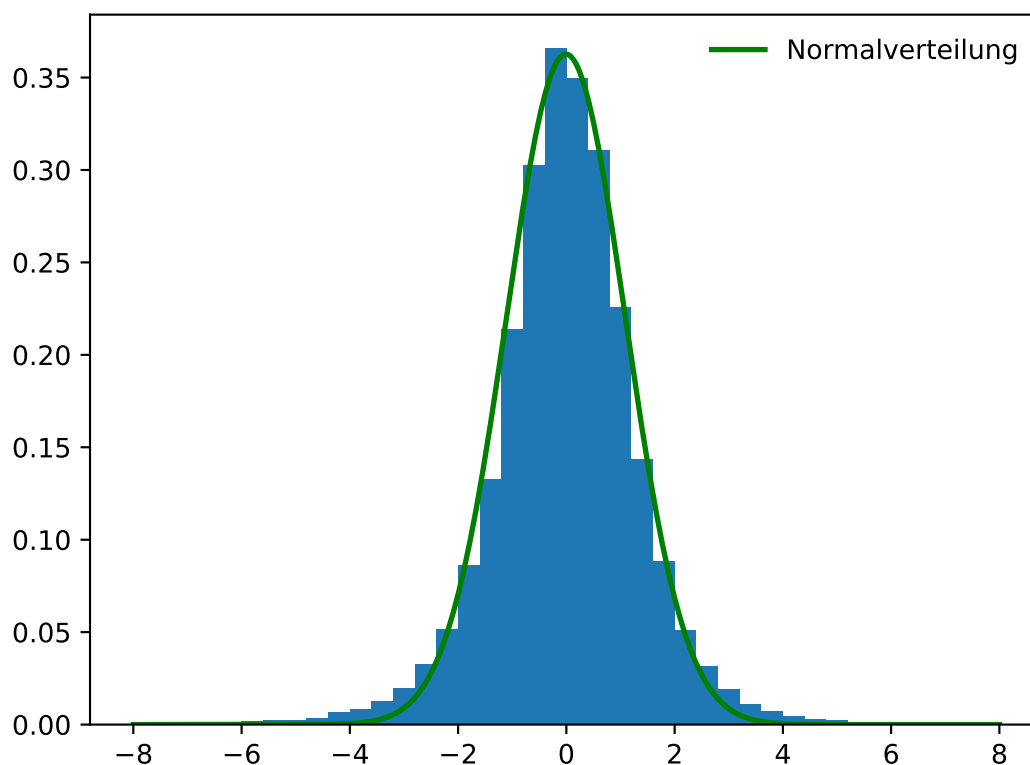
**Abbildung 3.2** Normalverteilung mit  $\mu = 0$  und unterschiedlichen Standardabweichungen  $\sigma$

Es soll nun überprüft werden, ob die Schritte der Aktienkurse normalverteilt sind. Zurückgegriffen wird dafür auf die normalisierten Daten, damit die Schritte jeweils eine Länge von 1 aufweisen und dadurch vergleichbar werden. Für eine Modellierung mit der Normalverteilung müssen zwei Parameter festgelegt werden: Der Mittelwert der Verteilung wird aufgrund der Annahme, dass kein signifikanter Trend vorliegt, als 0 festgelegt. Die Standardabweichung soll festgelegt werden, indem die Standardabweichung aller Schritte berechnet wird. Die Standardabweichung einer Stichprobe wird

folgendermassen mit der Formel

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.2)$$

berechnet, wenn diese auf eine Stichprobe angewendet wird. Nachdem diese beiden Parameter festgelegt wurden, kann ein Histogramm der Schritte erstellt werden und die entsprechende Kurve der Normalverteilung darübergerlegt werden. Die entsprechende Abbildung 3.3 zeigt das erhaltene Diagramm. [10]

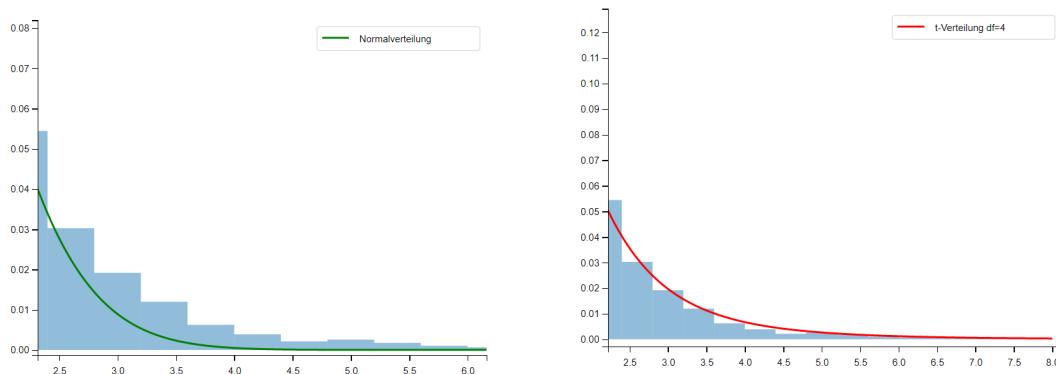


**Abbildung 3.3** Histogramm zur Verteilung der Kursveränderungen der Aktienkurse und Graph der Normalverteilung

Es ist eine gewisse Ähnlichkeit zur Normalverteilung sichtbar, wenn man nur die grobe Form betrachtet. Werden die Unterschiede genauer unter die Lupe genommen, fallen entscheidende Gründe gegen eine Modellierung durch eine Normalverteilung auf. Was die Normalverteilung bei weitem nicht richtig darstellt, sind grosse Schrittlängen. Die Normalverteilung verfügt über schmale Ränder und grosse Schritte sind somit viel unwahrscheinlicher als man in Wirklichkeit bei Schrittlängen von echten Aktienkursen beobachten kann. Diese unpassende Abbildung durch die Normalverteilung im Randbereich fällt aber in dieser Abbildung nicht so stark ins Auge, weil der schmale Randbereich visuell schwer abzuschätzen ist. Trotzdem ist es schon von Auge gut ersichtlich, dass es sich bei der Normalverteilung aus diesem Grund um eine unpassende Verteilungsfunktion handelt. Um diese Eigenschaft von Aktienkursen besser zu modellieren, wäre eine Verteilung mit schweren Rändern passender. Aus diesem Grund soll noch eine weitere Verteilung betrachtet werden, die diese Eigenschaft besser aufgreift, denn gerade diese grossen Sprünge sind im Kontext von Aktienkursen nicht

vernachlässigbar, weil sie eine grosse Kursveränderung darstellen.

Was die Normalverteilung abbilden kann, sind die kurzen und mittleren Schrittlängen, die auch am häufigsten vorkommen. Bei Aktienkursen treten aber auch unerwartete, extreme Ereignisse auf, die extreme Kursveränderungen mit langer Schrittlänge zur Folge haben. Bei der Normalverteilung fallen die Wahrscheinlichkeiten für extreme Werte schneller ab als dies in der Realität bei Aktienkursen beobachtet wird. Daher soll eine passendere Verteilung gefunden werden, die schwerere Ränder aufweist, also eine sog. *heavy-tailed* Verteilung. [6]



(a) Normalverteilung: Randbereich der Abb. 3.3

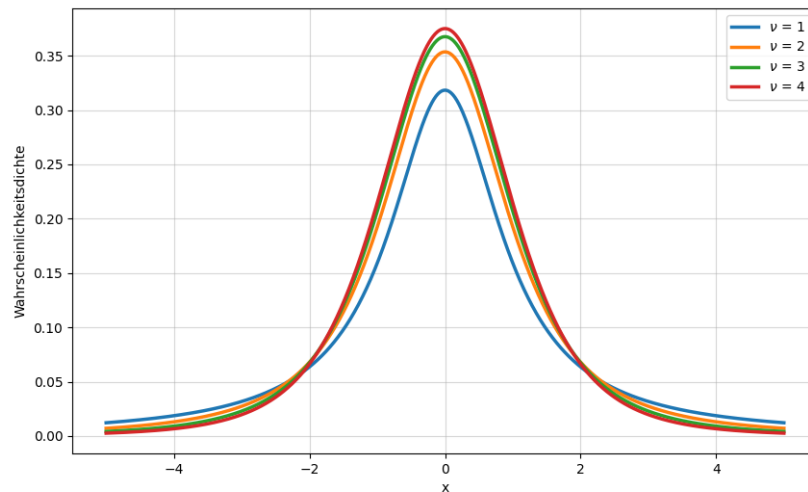
(b) t-Verteilung: Randbereich der Abb. 3.6

**Abbildung 3.4** Vergleich zwischen der Modellierung von grossen Schrittlängen im Randbereich der Normal- und t-Verteilung

### 3.1.2 t-verteilte Schrittlänge

Eine Verteilung, welche die Schrittlängen besser repräsentiert, ist die Studentsche t-Verteilung. Der Vorteil der Studentsche t-Verteilung für die Modellierung von Schrittlängen der Aktienkurse sind die schweren Ränder im Vergleich zur Normalverteilung. Im Unterschied zur Normalverteilung hängt die t-Verteilung nicht von der Standardabweichung der Daten ab, sondern von einem Parameter, der als "Freiheitsgrade  $\nu$ " bezeichnet wird. Wenn  $\nu = 1$  ist die Verteilung am stärksten *heavy-tailed* und nähert sich mit zunehmenden Freiheitsgraden der Normalverteilung an [14].

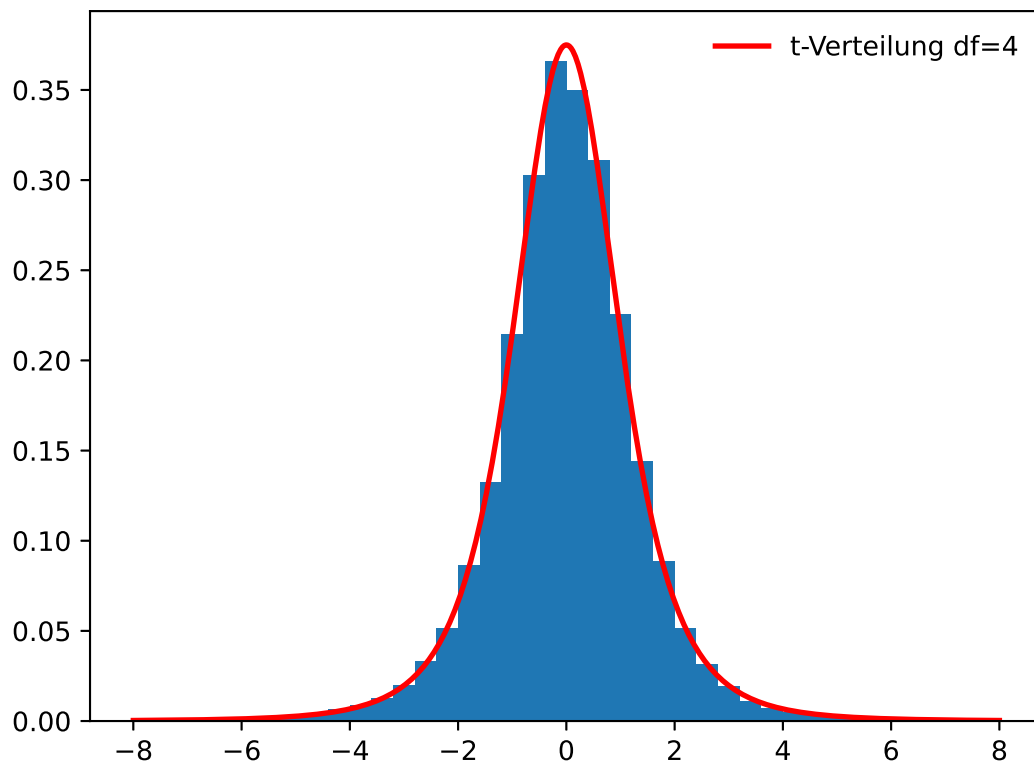
**Studentsche t-Verteilung** Die studentsche t-Verteilung wurde im Jahr 1908 von William Sealy Gosset entwickelt. Gosset hat herausgefunden, dass eine standardisierte Schätzfunktion der Stichproben-Mittelwerte normalverteilter Daten nicht normalverteilt ist. Es muss eine t-Verteilung verwendet werden, wenn die Varianz des Merkmals unbekannt ist und diese geschätzt werden muss. Im Rahmen der Modifizierung des Random-Walk-Modells spielen aber die schweren Ränder der t-Verteilung eine zentrale Rolle und führen zur Wahl dieser Verteilung.



**Abbildung 3.5** Student-t-Verteilung mit verschiedenen  $\nu$

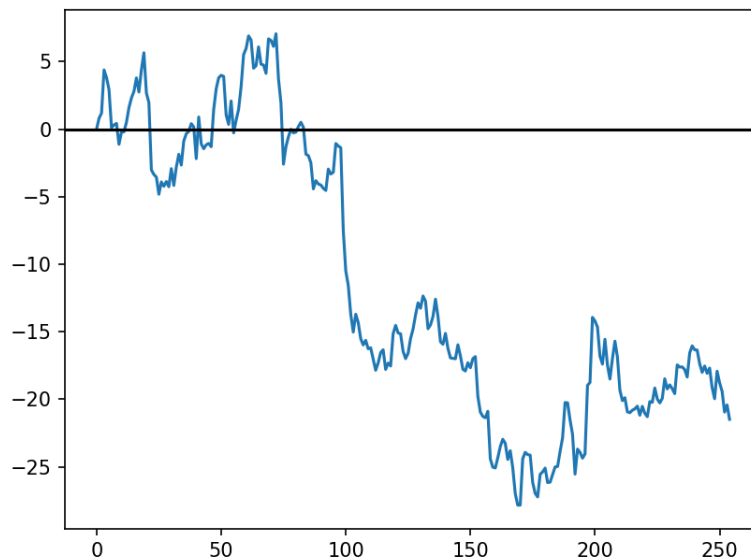
Wie bereits im Abschnitt zur Normalverteilung erwähnt, werden die Daten in einem Histogramm dargestellt und mit dem Graphen der t-Verteilung ergänzt. Es werden verschiedene Werte für die Freiheitsgraden  $\nu$  ausprobiert und der am besten passende Wert ausgewählt. Welcher Wert am besten geeignet ist, kann man besonders im Randbereich erkennen.





**Abbildung 3.6** Histogramm zu den Schrittgrößen der Aktienkurse und Graph der t-Verteilung mit  $\nu = 4$

Durch das Histogramm wird erkennbar, dass die studentsche t-Verteilung die Verteilung der Schrittgrößen viel besser modellieren kann als die Normalverteilung. Daher soll nun diese Verteilung in die Simulation eingefügt werden. Diese Anpassung kann nun mit der Anpassung bezüglich der Endposition kombiniert werden. Um die Nullstellen auszuwerten, muss neu jeder Vorzeichenwechsel als Nullstelle betrachtet werden.



**Abbildung 3.7** symmetrischer random walk mit t-verteilten Schritt­längen

## 3.2 Relevanz der Modellierung von Extremereignissen

Stellt man sich einen Aktienmarkt vor, wo jede Aktie sich analog zu dem im vorherigen Kapitel präsentierten Modell der einfachen, symmetrischen random walks auf  $\mathbb{Z}$  bewegt, hätte dies bezüglich der Risikoabschätzung einen beträchtlichen Unterschied zu Realität. Der entscheidende Punkt ist an dieser Stelle die Schritt­länge, welche bei dem Modell immer exakt 1 pro Zeiteinheit beträgt. Investiert man nun in eine solche imaginäre Aktie, kann das Risiko auf einfachste Weise abgeschätzt werden. Investiert man in eine Aktie während 10 Tagen, ist man sich bewusst, dass diese im Extremfall einen Verlust von 10 Einheiten generiert, analog dazu ist der mögliche Gewinn natürlich auch begrenzt.

Wird nun aber in eine echte Aktie investiert, ist dieser maximale Verlust nicht bekannt. Es können Extremereignisse auftreten, wo die Veränderung der Aktie von einer auf die nächste Zeiteinheit sehr gross ausfällt. Daher ist es von Bedeutung, diese Wahrscheinlichkeitsverteilung der Schritt­größen eines Aktienkurses zu modellieren, um das Risiko solcher Extremereignisse bezüglich der Schritt­grösse abzuschätzen. Ganz allgemein fällt auf, dass ein Grossteil der Schritt­größen um den Durchschnittswert aller Schritt­größen auftreten. Bei den normalisierten Aktienkursen wird diese durchschnittliche Schritt­grösse als 1 definiert. Mit grosser Wahrscheinlichkeit treten Schritt­größen um diesen Durchschnitt auf. Trotzdem besteht auch eine Chance, dass massiv grössere Schritte auftreten, was durch einfaches Betrachten eines Aktienkurses auffällt. Trotz der verhältnismässig geringen Wahrscheinlichkeit für diese grossen Schritte, beeinflussen sie das Risiko beträchtlich. Gerade eben auch weil sie selten auftreten, aber beim Auftreten einen grossen Verlust ausmachen können. Es ist daher zentral, bei der Wahl der passenden Verteilung für die anzunehmende Schritt­grösse, speziell auf die korrekte Darstellung der Extremereignisse zu achten.

Anhand der historischen Daten konnte festgestellt werden, dass die Schritt­grösse durch eine Wahrscheinlichkeitsverteilung mit schweren Rändern modelliert werden sollte. Diese Verteilungen modellieren die zu erwartenden Extremereignisse passender. Auch historisch gesehen hat sich in konkreten Fällen gezeigt, welche verheerenden Auswirkung ein unpassende Wahrscheinlichkeitsverteilung des Risikomodells haben kann.

Oftmals wird schnell die Normalverteilung angenommen, ohne dies kritisch zu hinterfragen und zu überprüfen. Ausschlaggebend dafür ist im Bereich der Finanzmathematik eben auch das ursprüngliche Random-Walk-Modell für die Modellierung von Aktienkursen, welches von dem Französischen Mathematiker Louis Bachelier entwickelt wurde. In seiner Arbeit ging Bachelier davon aus, dass die relative Preisänderung normalverteilt sind. [13]

### 3.2.1 Versagen der Finanzmodelle als Teilursache von Finanzkrisen

Seit Mitte 1990er Jahre gab es mehrere verheerende Finanzkrisen, dessen Folgen international spürbar waren. Beispiele dafür sind die Asienkrise im Jahre 1997, die Dotcom-Blase im Jahr 2000, später speziell auch die Subprime-Hypothekenkrise (2007) und die darausfolgende Weltfinanzkrise (2007-2008). Experten und Beobachter haben als Reaktion auf diese schwerwiegenden Krisen im Finanzsektor kritisiert, dass die verwendeten Modelle im Bereich Portfolio-Management, Risikomanagement aber auch die Modelle der verantwortlichen Behörden die Realität inkorrekt widerspiegelt haben. Solche Krisen verstärken die Bemühungen, bestehende Finanzmodelle und das allgemeine Risikomanagement zu verbessern. Die grundsätzliche Kritik an Modellen, welche sich der Normalverteilung bedienen, ist kein neues Phänomen. Bereits in den 1960er übte der Mathematiker Benoît Mandelbrot Kritik an Bacheliers Modell, welches die Normalverteilung annimmt. Trotzdem wird in Realität auch heute noch standardmässig die Normalverteilung angenommen. Gründe dafür sind einerseits die grosse Bekanntheit der Normalverteilung und der starke Fokus auf normalverteilte Modelle in der universitären Ausbildung im Finanz- und Wirtschaftsbereich. [13]

Anstelle der Normalverteilung eignen sich 'heavy tailed'-Verteilung besser, um die Realität im Falle der Verteilung der Kursveränderungen von Aktienkursen abzubilden. In Realität sind Extremevents, welche sich in diesem Falls als ausserordentlich grosse Kursbewegungen in einem kurzen Zeitraum äussern, viel häufiger zu beobachten, als die Normalverteilung implementieren würde. Oft wird auch in den Medien davon gesprochen, dass ein Aktienkurs einen 'Crash' erlebt hat. Dieses Phänomen entspricht einem solchen Extremevent, wobei es sich um einen grossen Schritt in die negative Richtung handelt. Natürlich sind auch Extremevents in die positive Richtung zu beobachten, finden in der allgemeinen Presse aber weniger Resonanz, weil sie keine weitläufigen, schwerwiegenden Folgen haben.

### 3.2.2 Volatility Clustering

Im Zusammenhang mit Aktien oder andern börsengehandelten Vermögenswerte wird das Ausmass der Schwankungen als Volatilität bezeichnet. Die Volatilität ist definiert als Standardabweichung der Kursänderungen. Im Allgemeinen ist die Volatilität ein wichtiges Mass zur Risikoabschätzung.

Ein Phänomen, welches auf dem realen Aktienmarkt zu beobachten ist, nennt sich '*Volatility Clustering*' und wurde im Jahr 1963 von Mandelbrot erstmals beschrieben. Das Phänomen verweist auf die Häufung von grossen Kursschwankungen in einem kurzen Zeitraum. Im Umgekehrten aber auch auf kleine Kursschwankungen wiederum kleine Schwankungen folgen. Volatility Clustering widerspricht dem einfachen Random-Walk-Modell als vollumfängliches Finanzmodell, weil dieses die Unabhängigkeit der Schritte vorsieht. Volatility Clustering ist in Bezug auf die Risikoabwägung relevant, weil nicht nur mit einzelnen grossen Kursschwankungen gerechnet werden muss, sondern auf solche Extremevents mit erhöhter Wahrscheinlichkeit weitere grosse Schritte folgen. Meist wird dieses Phänomen aber als auf kurzfristige Zeitintervalle begrenztes Phänomen betrachtet, dessen Auswirkungen nur währen einigen Minuten bis wenigen Wochen anhalten.

## Kapitel 4

# Datenanalyse

In den vorhergehenden Kapitel wurden sowohl die random walks auf ihre Eigenschaften untersucht, als auch bereits eine passende Modifikation vorgenommen. Als Ziel wurde festgelegt, dass dieses gefundene Modell, die symmetrischen random walks mit t-verteiltern Schritten, Aktienkurse möglichst gut modellieren können. Anhand echter Daten von Aktienkursen soll das präsentierte Modell nun überprüft werden. Diese Überprüfung des Modells stützt sich sowohl auf optische Vergleiche, als auch auf den Vergleich der Eigenschaften des Modells und der Aktienkurse. Zur Analyse dieser Daten wird die Programmiersprache Python eingesetzt, welche dank vielen verfügbaren libraries den Umgang, die Analyse und die Visualisierung der Daten erleichtert. In einem ersten Schritt müssen die Aktienkursdaten importiert und normalisiert werden. Dieser Prozess wird in den folgenden Abschnitten dargelegt.

### 4.0.1 Auswahl der Aktien

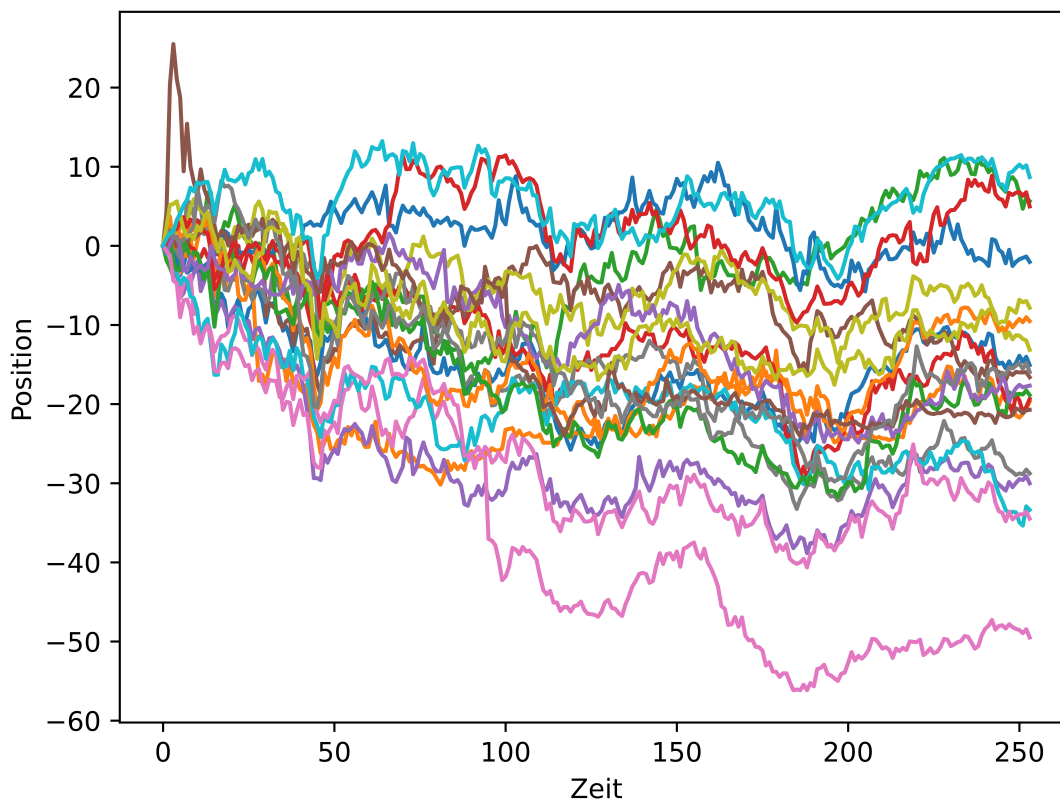
Um einen Vergleich mit Aktienkursen machen zu können, werden historische Börsendaten von unterschiedlichen Aktien benötigt. Es werden Aktien von Schweizer Unternehmen ausgewählt, die an der Schweizer Börse SIX gehandelt werden. An dieser Stelle sollte darauf hingewiesen werden, dass die Anzahl Börsentage in einem Jahr nicht mit der Anzahl Kalendertage übereinstimmt, weil die Börse jeweils am Wochenende und an nationalen Feiertagen geschlossen bleibt.

Der Zugriff auf die historischen Börsendaten erfolgt über zweit Weg. Um die Daten mit Python zu analysieren, wird yfinance, eine Python-Bibliothek, verwendet. Die Bibliothek erlaubt uns, Aktienkurse von Yahoo Finance abzurufen. Die heruntergeladenen Daten werden anschliessend in einem CSV-File gespeichert. In dieser Analyse wird jeweils nur der Wert der Spalte 'close' betrachtet.

Im Allgemeinen wird mit  $x_0$  auf den Wert der betrachteten Aktie am ersten Tag des gewählten Zeitpunkts verwiesen.  $x_n$  meint den letzten Tag des Zeitraums mit  $n + 1$  Tagen.

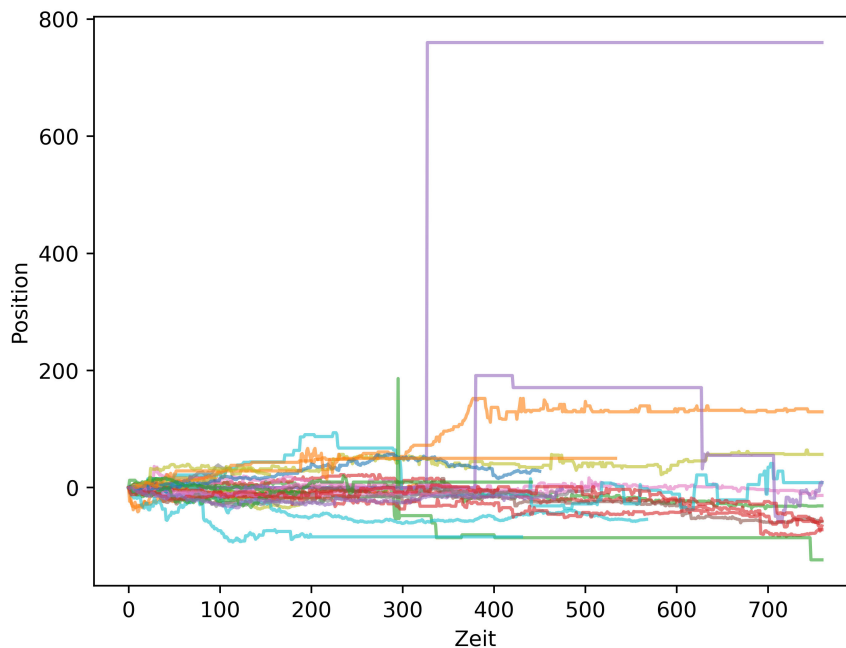
| Date       | Open               | High               | Low                | Close              | Adj Close          | Volume  |
|------------|--------------------|--------------------|--------------------|--------------------|--------------------|---------|
| 2022-12-30 | 28.34000015258789  | 28.459999084472656 | 28.059999465942383 | 28.059999465942383 | 27.268779754638672 | 2387244 |
| 2023-01-03 | 28.3799991607666   | 28.860000610351562 | 28.229999542236328 | 28.739999771118164 | 27.92960548400879  | 3103295 |
| 2023-01-04 | 28.959999084472656 | 29.760000228881836 | 28.799999237060547 | 29.709999084472656 | 28.87225341796875  | 5088804 |
| 2023-01-05 | 29.709999084472656 | 29.829999923706055 | 29.450000762939453 | 29.729999542236328 | 28.891691207885742 | 4388238 |
| 2023-01-06 | 29.809999465942383 | 30.010000228881836 | 29.559999465942383 | 30.010000228881836 | 29.163795471191406 | 2638284 |
| 2023-01-09 | 30.0               | 30.540000915527344 | 30.0               | 30.440000534057617 | 29.5816707611084   | 3819373 |
| ...        | ...                | ...                | ...                | ...                | ...                | ...     |
| 2023-12-22 | 37.20000076293945  | 37.36000061035156  | 37.150001525878906 | 37.2599983215332   | 37.2599983215332   | 1729441 |
| 2023-12-27 | 37.310001373291016 | 37.40999984741211  | 37.11000061035156  | 37.29999923706055  | 37.29999923706055  | 1708447 |
| 2023-12-28 | 37.310001373291016 | 37.369998931884766 | 37.040000915527344 | 37.04999923706055  | 37.04999923706055  | 1593899 |
| 2023-12-29 | 37.2400016784668   | 37.31999969482422  | 37.119998931884766 | 37.29999923706055  | 37.29999923706055  | 1810717 |

Tabelle 4.1 Importierte Kursdaten der Aktie 'ABBN.SW'



**Abbildung 4.1** Kursverlauf von 20 Aktien im Zeitraum von 30.12.2021 bis 30.12.2023

Das Vorgehen gestaltet sich so, dass anhand einer Tickerliste alles an der Schweizer Börse gehandelte Aktien die dazugehörigen Daten eines gewählten Zeitraums importiert werden und in einem CSV-File zwischengespeichert werden. Es fällt jedoch auf, dass gewisse Aktienkursverläufe im gewählten Zeitraum keine Bewegung aufweisen, was aber nicht auf eine zufällige Bewegung zurückzuführen ist. Gerade bei Aktien kleiner Unternehmen kommt es vor, dass eine Aktie in einem Zeitraum nicht aktiv gehandelt wurde oder nicht im ganzen Zeitraum an der Schweizer Börse gehandelt wurde. Bei Aktien von Grossunternehmen, sogenannte *Blue Chips*, tritt dies kaum auf. Es werden alle Aktien aussortiert, die mehr Tage ohne Kursveränderung als eine gewisse Toleranz aufweisen, sowie alle Aktien die, ebenfalls mit einem Toleranzbereich, nicht während dem ganzen Zeitraum aktiv gehandelt wurden.



**Abbildung 4.2** Aussortierte Aktienkurse im Zeitraum vom 31.12.2020 bis zum 31.12.2023 mit einer Toleranz von 400 Tagen ohne Kursbewegung und 100 Tagen Verkürzung

#### 4.0.2 Normalisierung der Aktienkurse

Damit die Aktienkurse mit random walks verglichen werden können, müssen diese zuerst normalisiert werden. Als Zeitspanne wird jeweils für jede Aktie das gleiche Jahr gewählt, wobei die Werte der Aktien mit einem Wert pro Tag in einer Tabelle aufgelistet werden. Da alle ausgewählten Aktien an der Schweizer Börse gehandelt werden, gibt es pro Aktie jeweils die gleiche Anzahl Börsentage. Die gewählte Zeitspanne enthält 255 Börsentage. Der eindimensionale random walk, der als Vergleich genutzt wird, startet jeweils beim Wert 0. Da die verschiedenen Aktienkurse am Anfang des Jahres nicht beim Wert 0 starten, werden diese angepasst, indem von jedem Wert jeweils der Wert des ersten Tages  $x_0$  subtrahiert wird. Im zweiten Schritt der Normalisierung, soll die Änderung einer Aktie von einem zum nächsten Börsentag so angepasst werden, dass sie, über den ganzen betrachteten Zeitraum gesehen, einen Durchschnittswert von 1 annimmt.

Ein Sprung eines Aktienkurses vom einen auf den nächsten Börsentag definieren wir als  $d_n$ . Der erste Sprung  $d_1$  entspricht der Differenz der Tageswerte des Tag 0 und Tag 1. Im allgemeinen kann jeder Sprung durch

$$d_n = x_n - x_{n-1}, \quad n \leq 0 \quad (4.1)$$

definiert werden. Die Sprunggröße  $|d_n|$  entspricht dem Betrag des Sprungs. Im Zusammenhang mit der Normalisierung ist nur der Betrag eines Sprunges, also die Sprunggröße, verändert. Die Richtung der Sprünge sowie die relativen Längenunterschiede bleiben erhalten. Wobei die Richtung dem Vorzeichen des Sprungs entspricht. Zwecks der Normalisierung soll die durchschnittliche Sprunggröße

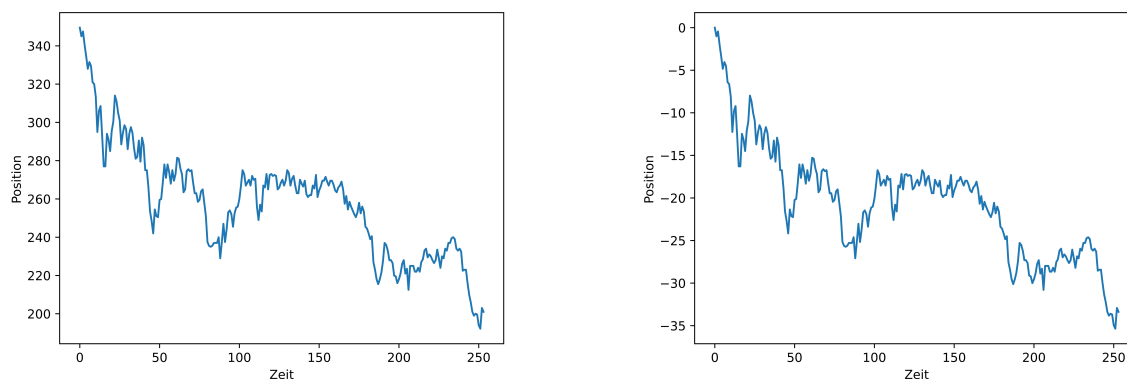
$$\mu = \frac{1}{n} \sum_{k=1}^n |d_k| \quad (4.2)$$

berechnet werden. Dieser Wert der durchschnittlichen Sprunggröße  $\mu$  variiert für die verschiedenen Aktien und nimmt nicht wie bei der Simulation des random walks immer den Wert 1 an. Mit Hilfe

der durchschnittlichen Sprunggrösse  $\mu$  können schliesslich die normalisierten Werte

$$a_n = \frac{x_n - x_1}{\mu} \quad (4.3)$$

bestimmt werden. Durch diese Normalisierung wird die relative statt der absoluten Sprunggrösse dargestellt, wodurch ein Vergleich mit einem random walk sinnvoll wird. Die einzelnen Sprunggrössen können unterschiedlich gross sein, was ein Unterschied zum gewählten Modell des random walks darstellt, da in diesem Modell jeweils nur Sprünge mit der Sprunggrösse 1 möglich sind. Im Durchschnitt hat aber nun jede Aktie eine Sprunggrösse von 1.



(a) Originaler Aktienkurs

(b) Normalisierter Aktienkurs

**Abbildung 4.3** Vergleich zwischen dem originalen und dem normalisierten Aktienkurs der HBM Healthcare Investments AG (HBMN.SW)

Anhand der Abbildung 4.3b kann man erkennen, dass es sich bei der Normalisierung nur um eine Verschiebung in y-Richtung und eine Skalierung handelt. So gesehen wird nur das Referenzsystem um den Kursverlauf verändert, wobei der ursprüngliche Charakter des Kursverlaufes erhalten bleibt. Die ursprüngliche Form der Kursbewegung ist identisch mit der normalisierten.

## 4.1 Vergleich durch Visualisierung

Der erste Vergleich basiert darauf, dass die normalisierten Aktienkurse in einem Diagramm dargestellt werden. Anschliessend werden analog zu den Aktienkursen die gleiche Anzahl random walks mit gleicher Anzahl Schritte generiert und ebenfalls in einem Diagramm dargestellt. Dieser optische Vergleich bietet eine explorative Möglichkeit, einen ersten Eindruck zu erhalten und gewisse Muster zu erkennen. Potentiell erste Gemeinsamkeiten, aber auch Unterschiede werden ersichtlich. Ganz allgemein wird ein visueller Eindruck vermittelt. Ein zusätzlicher Vorteil dieser Methode liegt in der Einfachheit. Natürlich gibt es auch nachteilige Aspekte bei der Anwendung dieser Methode. Neben dem grundsätzlichen Einfluss der Subjektivität bei jeglichen rein visuellen Vergleichen, kann auch eine zufällige Ähnlichkeit nicht ausgeschlossen werden. Aufgrund dieser erwähnten Schwachstellen bietet dieser Vergleich nicht genügend Aussagekraft für eine vollständigen und abschliessenden Analyse der Ähnlichkeiten zwischen dem gewählten Modell der modifizierten random walk und Aktienkursen. In Kombination mit den nachfolgenden Methoden bietet diese Methode durchaus einen Mehrwert.

Aufgebaut ist der Vergleich folgendermassen: Es wird eine Zeitspanne festgelegt und die Aktienkursdaten aus diesem Zeitraum importiert und normalisiert gemäss dem vorherigen Abschnitt 4.3b. Diese Daten werden anschliessend in einem Liniendiagramm visualisiert. Entsprechend der Anzahl visualisierter Aktienkurse werden die gleiche Anzahl random walks generiert, wessen Schrittzahl mit den Anzahl Börsentagen im gewählten Zeitraum übereinstimmt. Die genaue Vorgehensweise ist im Programmcode im Anhang ersichtlich.

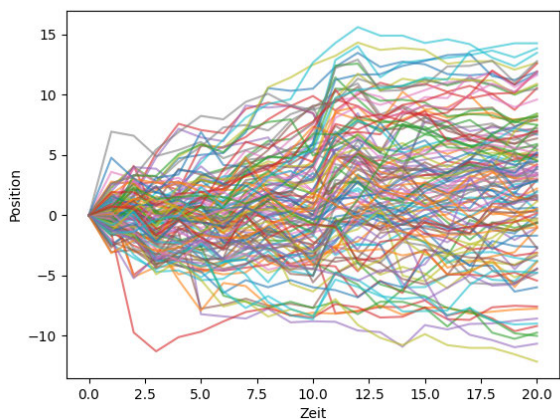
Nachfolgend wird die beschriebene Methode für verschiedene Zeitperioden angewendet. Dies soll auch einen Überblick gewähren, wie sich das Modell kurz- bis langfristig eignet. Mögliche visuelle Merkmale, welche verglichen werden können, sind beispielsweise die Spannweite aller random walks oder Aktienkurse, die allgemeine Form sowie mögliche Symmetrien.

### 4.1.1 Zeitraum: 20 Börsentage

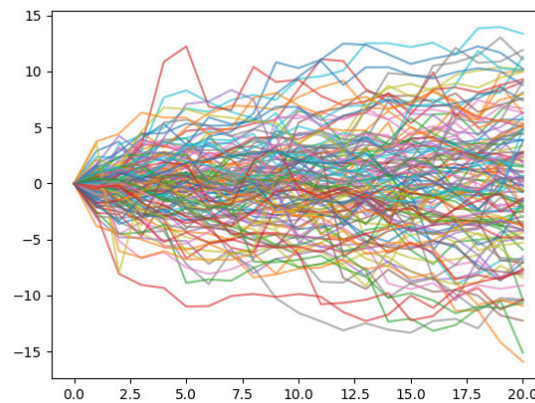
Dieser erste Zeitraum von 20 Börsentage zeigt die kurzfristige Entwicklung der Aktienkurse und soll einen optischen Vergleich zu random walks der gleichen Länge bieten. Die Abbildungen auf der linken Seite zeigen jeweils die Diagramme der Aktienkurse aus drei verschiedenen Zeitperioden. Rechts befinden sich drei Simulationen der modifizierten random walks gleicher Länge, welche mögliche Ergebnisse des Modells abstecken sollen.

Deutlich wird die Ähnlichkeit bezüglich der Schrittlängen. Es kann beobachtet werden, dass einerseits durchschnittliche Schrittgrössen relativ häufig auftreten. Auch grosse Schrittgrössen sind sowohl bei den Aktienkursen als auch bei den modifizierten random walks beobachtet werden. Diese grossen Schrittgrössen kommen in einer ähnlichen Häufigkeit und in einem ähnlichen Ausmass vor. Die Abbildung der random walks zeigt, dass die modifizierten random walks etwas breiter verteilt sind im Bezug auf die maximal und minimal erreichte Position. Sowohl Aktienkurse und random walks verteilen sich ungefähr in einem Bereich von 15 bis -15, wobei diese Verteilungen bei den random walks leicht breiter ausfällt.

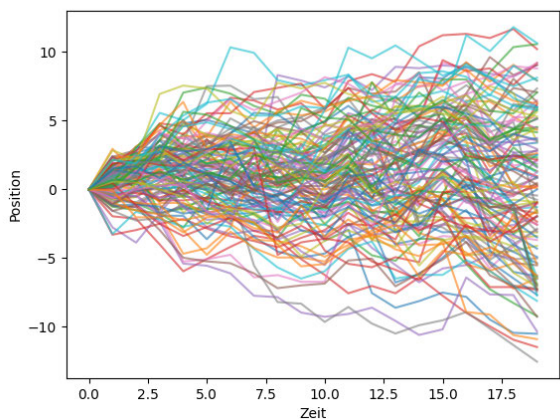




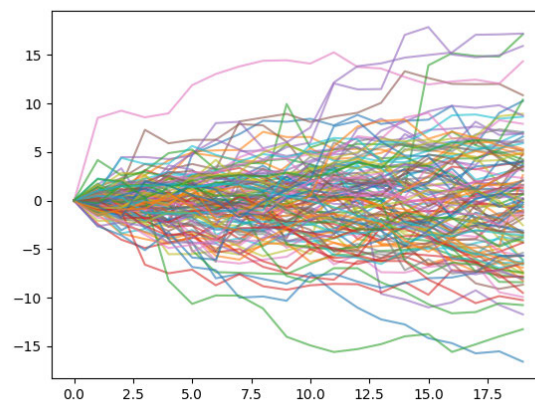
(a) Aktienkurse im Zeitraum vom 29.11. - 30.12.23



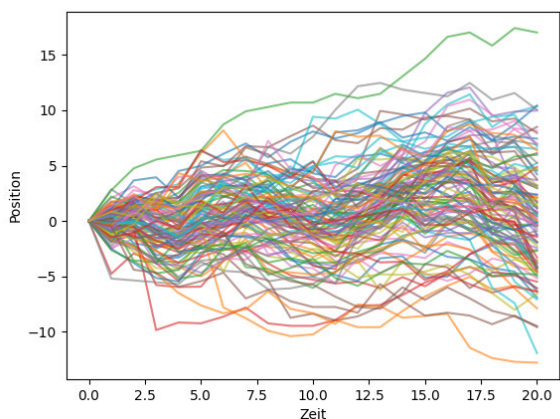
(b) gleiche Anzahl random walks mit 20 Schritten



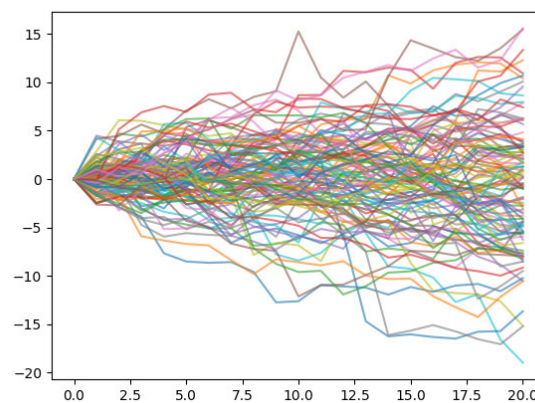
(c) Aktienkurse im Zeitraum vom 10.02. - 12.03.23



(d) gleiche Anzahl random walks mit 20 Schritten



(e) Aktienkurse im Zeitraum vom 10.06. - 10.07.19



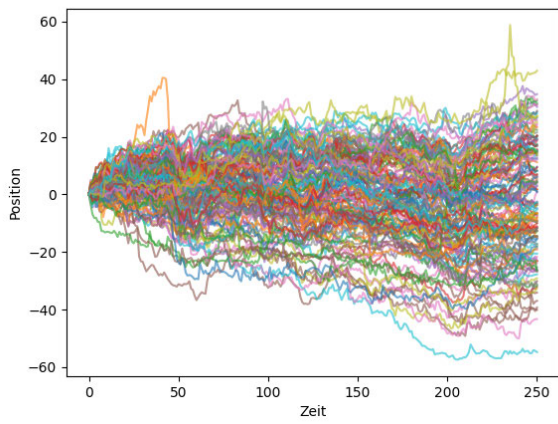
(f) gleiche Anzahl random walks mit 20 Schritten

**Abbildung 4.4** Visualisierung des Vergleichs zwischen kurzzeitigen Aktienkursbewegungen und modifizierten random walks gleicher Länge

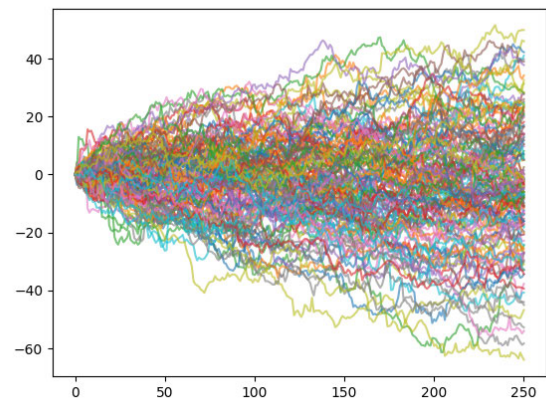
### 4.1.2 Zeitraum: 1 Jahr

Als mittelfristiges Zeitintervall wird die Dauer eines Jahres gewählt. Der visuelle Vergleich wird jeweils für das Jahr 2018, 2022 und 2023 durchgeführt. Durch die Visualisierung drei verschiedener Jahre sollen verschiedene Beispiele aufgezeigt werden, wie ein solcher Abschnitt aussehen kann. Dadurch ist auch möglich, abzuschätzen welche Unterschiede innerhalb der Aktienkurse von Jahr zu Jahr zu erwarten sind und somit auch die Beurteilung des Vergleichs bezüglich der modifizierten random walks verbessert werden kann. Als Gemeinsamkeit können 'Ausreisser' beschrieben werden. Damit sind random walks oder eben Aktienkurse, die sich relativ weit vom Grossteil der random walks oder Aktienkurse wegbewegen. Bei einem visuellen Vergleich fallen diese einzelnen Linien im Diagramm besonders ins Auge. Wie bereits im visuellen Vergleich der kurzen Abschnitten von 20 Tagen beschrieben, fällt auch hier auf, dass sich random walks und Aktienkurse in einem ähnlichen Bereich bewegen. In diesem Fall liegt dieser Bereich bei ungefähr -60 bis 60. Die Symmetrie der Abbildung ist in allen Diagramme in einem gewissen Ausmass vorhanden.

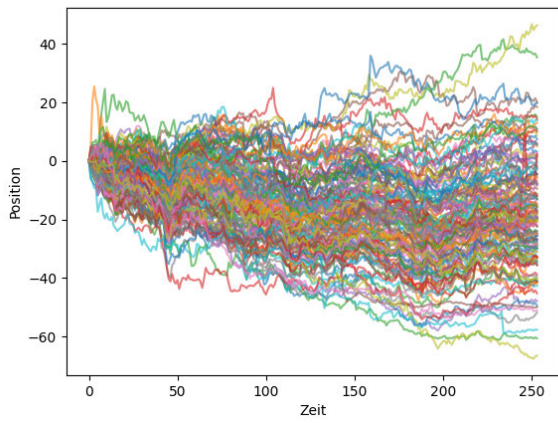
Dieser konkrete Zeitraum von einem Jahr spielt eine wichtige Rolle, weil er auch in den beiden anderen Vergleichsmethoden zur Anwendung kommt. Es fällt ähnlich zum langfristigen Zeitraum eine gewisse Abhängigkeit der Einzelaktien untereinander auf, wodurch eine leichte Gesamtbewegung des Marktes sichtbar wird. Es ist wichtig zu verstehen, dass vom Random-Walk-Modell primär impliziert wird, dass die einzelnen Schritte eines Aktienkurses unabhängig sind. Dies ist klar zu unterscheiden von der Unabhängigkeit der verschiedenen Aktienkurse.



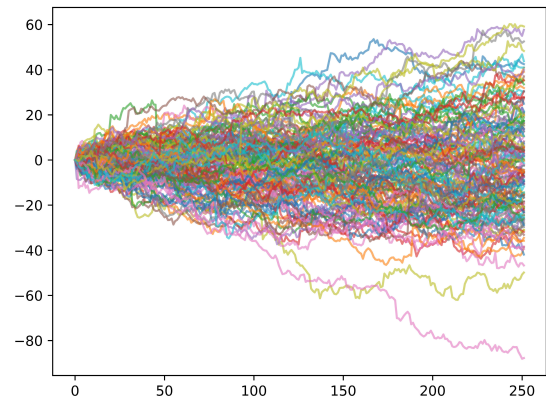
(a) Aktienkurse im Zeitraum vom 31.12.22. - 30.12.23



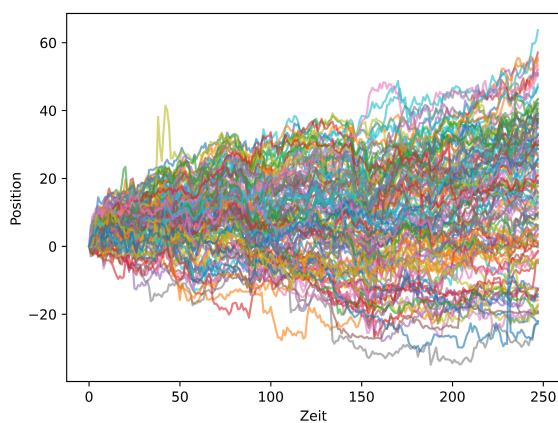
(b) modifizierte random walks



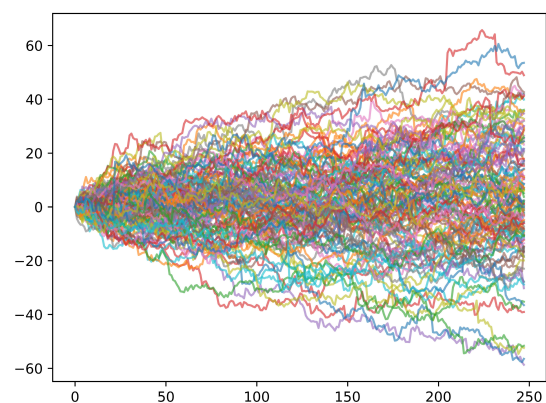
(c) Aktienkurse im Zeitraum vom 31.12.21. - 30.12.22



(d) modifizierte random walks



(e) Aktienkurse im Zeitraum vom 31.12.17. - 30.12.18

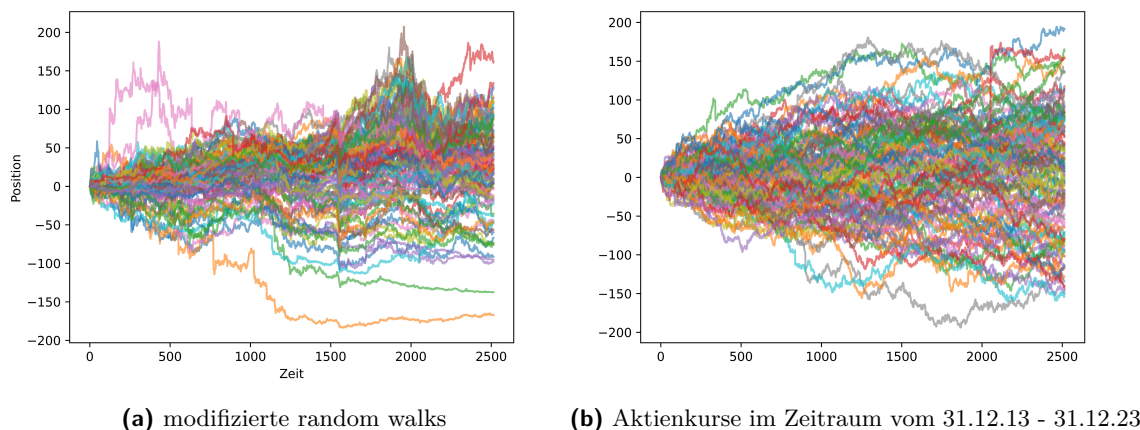


(f) modifizierte random walks

**Abbildung 4.5** Visualisierung des Vergleichs zwischen Aktienkursbewegungen innerhalb eines Jahres und modifizierten random walks gleicher Länge

### 4.1.3 Zeitraum: 10 Jahre

Es folgt nun ein langfristiger Vergleich über einen Zeitraum von 10 Jahren. Der Zeitraum wird vom 31.12.2013 bis zum 31.12.2023 gewählt. Zuerst soll bemerkt werden, dass über diesen langen Zeitraum einige Aktien entfernt werden im Prozess der Normalisierung (4.3b). Bei einem längeren Zeitraum gibt es mehr Aktien, die nicht über den ganzen Zeitraum an der Schweizer Börse gehandelt wurden und deshalb aussortiert werden. Wie bereits in den kurz- und mittelfristigen Vergleichen fällt auf, dass die einzelnen Aktien in einem gewissen Zusammenhang stehen. Es werden einige Tage sichtbar, wo sich ein Grossteil der Aktien in die selbe Richtung mit ähnlicher Schrittlänge bewegen. Erklären kann man solche Stellen mit Ereignissen in Politik oder Wirtschaft, welche sich über den gesamten Aktienmarkt auswirken. Bei den simulierten Random walks ist dies nicht der Fall, weil sich die einzelnen random walks unabhängig voneinander bewegen. Eine Gemeinsamkeit der beiden Abbildungen besteht, wenn man den Bereich betrachtet, indem sich ein grosser Teil der random walks oder Aktienkurse bewegt. In beiden Fällen liegt dieser Bereich bei etwa -150 bis 150. Was im Vergleich zu den vorherigen Vergleichen mit den kürzeren Zeitintervallen auffällt, ist die verminderte Symmetrie bei den Aktienkursen. Vermehrt ist eine gewisse Gesamttendenz vorhanden, welche der Aktienmarkt verzeichnet. Visuell gesehen fallen hier die stärksten Unterschiede zwischen den modifizierten random walks und den Aktienkursen auf.



**Abbildung 4.6** Graphischer Vergleich zwischen modifizierten random walks und Aktienkursen im Zeitraum von 10 Jahren

## 4.2 Vergleich der Eigenschaften bezüglich Nullstellen und Extremstellen

Im Gegensatz zu rein visuellen Vergleichen, welche im vorherigen Abschnitt gemacht werden, soll nun eine Vergleichsmethode folgen, welche konkreter auf die einzelnen Eigenschaften eingeht. Dies erlaubt ein differenzierteres Bild bezüglich der konkreten Gemeinsamkeiten und Unterschiede. Die verglichenen Eigenschaften stimmen mit den vorgestellten Eigenschaften der random Walks in Kapitel 2 überein.

### 4.2.1 Vergleich der Durchschnittswerte

Im Kapitel 'Random Walks' (2) wurde stets das Modell einfacher, symmetrischer random walks verwendet und die Eigenschaften solcher random walks beschrieben und begründet. Im darauffolgenden Kapitel wurde die Modifikation der Schrittlänge eingeführt, was den Vorteil einer besseren Repräsentation der Realität mit sich bringt. Im Gegenzug führt es aber auch dazu, dass die ent-

sprechenden Eigenschaften kaum noch theoretisch hergeleitet werden können. Im nachfolgenden Abschnitt werden im Bezug auf die modifizierten random walks ausschliesslich auf Daten aus einer Realisierung einer gleichen Anzahl random walks mit gleicher Länge wie bei den entsprechenden Aktienkursen betrachtet. Als zusätzlicher Richtwert, welcher auch die Auswirkungen der Modifikation aufzeigt, wird ebenfalls der theoretische Erwartungswert der entsprechenden Eigenschaft aufgeführt, welcher zum ursprünglichen Modell der unmodifizierten random walks gehört.

Beim Vergleich des Mittelwerts spielt besonders der Standardfehler eine bedeutende Rolle (4.4), weil dieser die Unsicherheit der Schätzung des erhaltenen Mittelwerts absteckt. Sowohl bei den modifizierten random walks als auch die Aktienkurse stellen nur eine Stichprobe aus der Grundgesamtheit dar.

|                                       | random walk | modifizierter random walk | Aktienkurse |
|---------------------------------------|-------------|---------------------------|-------------|
| <b>Anzahl Nullstellen:</b>            |             |                           |             |
| Mittelwert:                           | 11.754      | 9.177                     | 6.631       |
| Standardabweichung:                   | 8.977       | 7.714                     | 9.996       |
| Standardfehler:                       |             | 0,650                     | 0.841       |
| <b>Position erste Nullstelle:</b>     |             |                           |             |
| Mittelwert:                           | 13.372      | 15.372                    | 64.296      |
| Standardabweichung:                   | 30.227      | 34.418                    | 66.908      |
| Standardfehler:                       |             | 2.899                     | 5.635       |
| <b>Position letzte Nullstelle:</b>    |             |                           |             |
| Mittelwert:                           | 127.000     | 123.858                   | 111.085     |
| Standardabweichung:                   | 90.155      | 91.541                    | 74.935      |
| Standardfehler:                       |             | 7.709                     | 6.311       |
| <b>Maximum:</b>                       |             |                           |             |
| Mittelwert:                           | 12.229      | 17.788                    | 19.364      |
| Standardabweichung:                   | 9.604       | 13.916                    | 9.996       |
| Standardfehler:                       |             | 1.172                     | 0.842       |
| <b>Minimum:</b>                       |             |                           |             |
| Mittelwert:                           | -12.229     | -15.830                   | -12.856     |
| Standardabweichung:                   | 9.604       | 12.393                    | 13.121      |
| Standardfehler:                       |             | 1.044                     | 1.105       |
| <b>Position erstes Maximum:</b>       |             |                           |             |
| Mittelwert:                           | 121.123     | 136.440                   | 104.716     |
| Standardabweichung:                   | 89.931      | 89.763                    | 84.721      |
| Standardfehler:                       |             | 7.559                     | 7.135       |
| <b>Position erstes Minimum:</b>       |             |                           |             |
| Mittelwert:                           | 121.123     | 120.809                   | 137.929     |
| Standardabweichung:                   | 89.931      | 90.304                    | 94.628      |
| Standardfehler:                       |             | 7.605                     | 7.969       |
| <b>Durchschnittliche Endposition:</b> |             |                           |             |
| Mittelwert:                           | 0.000       | 2.148                     | 1.335       |
| Standardabweichung:                   |             | 22.772                    | 20.038      |
| Standardfehler:                       |             | 1.918                     | 1.688       |

**Tabelle 4.2** Vergleich der theoretischen Werte für einen unmodifizierten random walk mit den Werten der Aktienkurse und den Werten der Simulation der modifizierten random walks

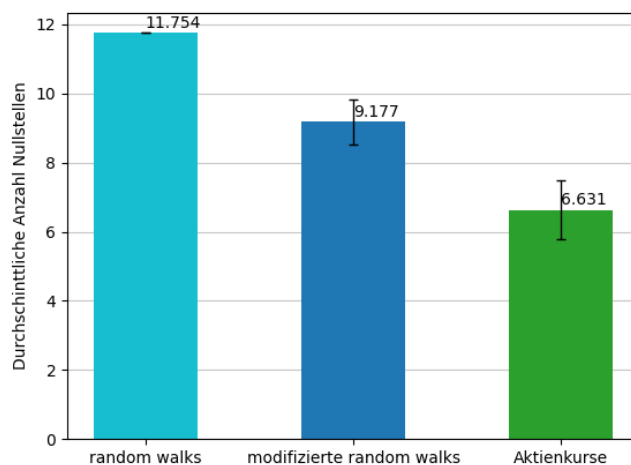
In diesem Abschnitt wird betrachtet, in welcher Ausprägung die präsentierten Eigenschaften der random walks in den normalisierten Aktienkurse vorkommen. Es werden jeweils der theoretische Erwartungswert für einfache, symmetrische random walks der Länge  $2n = 254$  mit den Werten aus

einer Simulation von modifizierten random walks von identischer Länge und dem Mittelwert der normalisierten Aktienkurse verglichen. Die Simulation wurde in der Programmiersprache 'Python' geschrieben und es wurden entsprechend der Anzahl ausgewerteter Aktien 141 modifizierte random walks simuliert und dessen Nullstellen analysiert. Es werden jeweils der theoretische Erwartungswert mit den Durchschnittswerten der Simulation und der Aktienkurse in einem Säulendiagramm dargestellt. Für die beiden Durchschnittswerte der Simulation und der Aktienkurse werden jeweils Fehlerindikatoren ergänzt. Dieser Balken bildet den Standardfehler ab. Der Standardfehler wird verwendet, weil er ein Mass dafür ist, wie genau der errechnete Durchschnittswert der Stichprobe dem tatsächlichen Durchschnittswert der Grundgesamtheit entspricht. Der Standardfehler wird mit der Formel [5]

$$SF = \frac{s_x}{\sqrt{n}} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}}{\sqrt{n}} \quad (4.4)$$

berechnet, wobei  $s_x$  die Standardabweichung ist und  $n$  die Anzahl Elemente der Stichprobe.[1] Es kann beobachtet werden, dass der Standardfehler der Simulation immer sehr gering ausfällt, was auf eine hohe Genauigkeit der Simulation hindeutet. Der Standardfehler der Durchschnittswerte der Aktienkurse fällt sehr viel grösser aus, was auf die kleine Stichprobe von 20 gewählten Aktienkursen zurückzuführen ist. Nun kann betrachtet werden, ob der theoretische Erwartungswert innerhalb des Intervalls [*Durchschnittswert* ± *Standardfehler*] liegt.

#### 4.2.2 Anzahl Nullstellen



**Abbildung 4.7** Diagramm zu den durchschnittlichen Anzahl Nullstellen

Man kann erkennen, dass der Wert der durchschnittlichen Anzahl Nullstellen der Aktienkurse tiefer ist als der Erwartungswert. Auch wenn das Intervall mit dem Standardfehler betrachtet wird, liegt der Erwartungswert nicht in diesem Intervall. Die Verteilung der Nullstellen soll nun betrachtet werden. Die Abbildung zur theoretischen Verteilung wird in Abbildung 2.4 gezeigt.

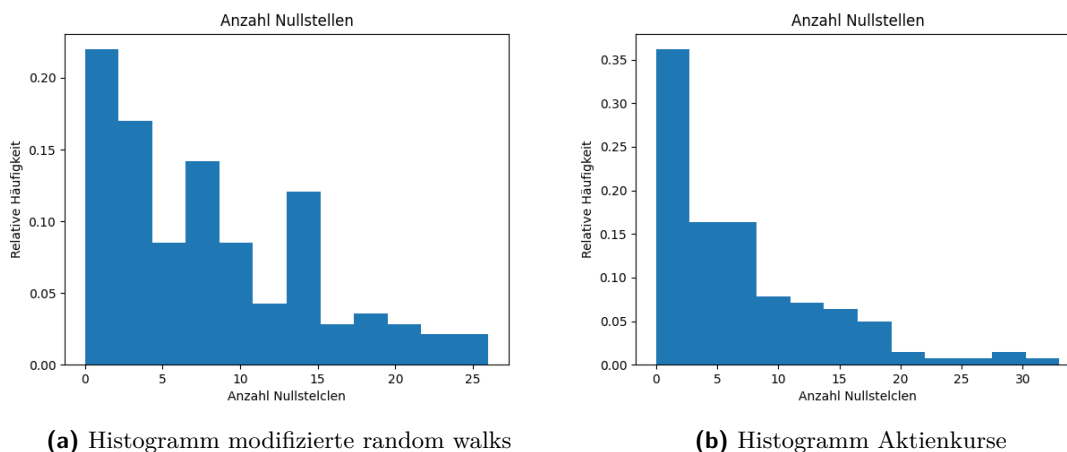


Abbildung 4.8 Diagramme zur Verteilung der Anzahl Nullstellen

### 4.2.3 Die erste Nullstellen

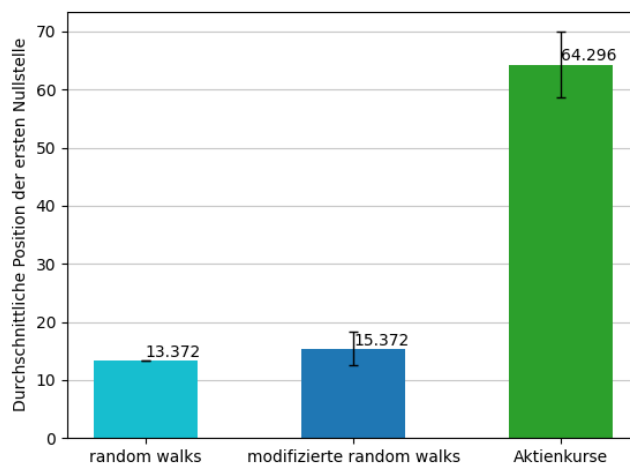
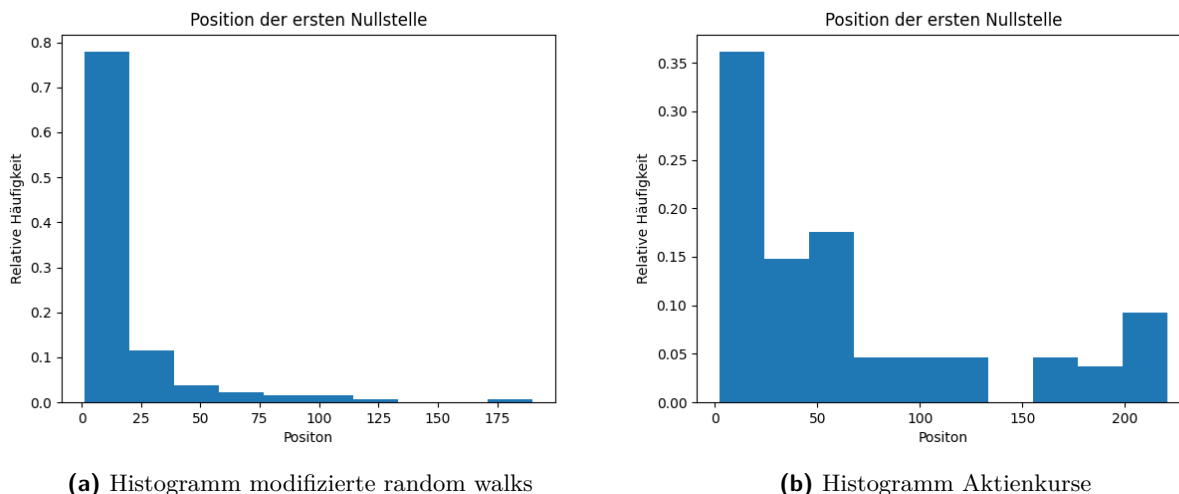


Abbildung 4.9 Diagramm zur durchschnittlichen Positionen der ersten Nullstelle

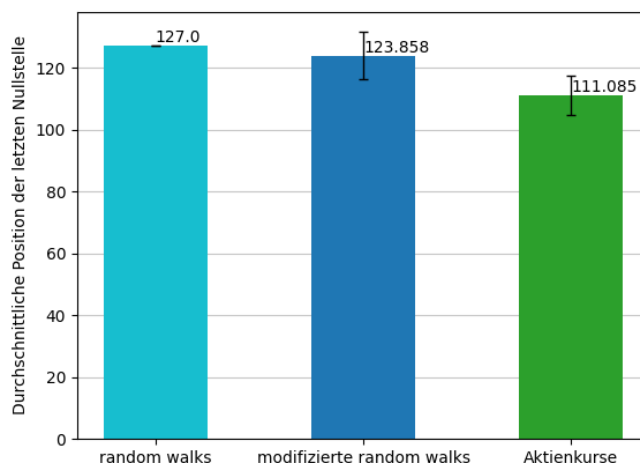
Man kann erkennen, dass der theoretische Erwartungswert für die durchschnittliche Position der ersten Nullstelle bei weitem nicht innerhalb des Konfidenzintervalls des Mittelwerts der Aktienkurse liegt. Wie sich dieser riesige Unterschied des Mittelwerts erklären lässt, soll nun ergründet werden. Eine mögliche Erklärung wäre, dass sich die Aktienkurse tendenziell schneller und stärker von der x-Achse entfernt haben und im Anfangsbereich wenige Nullstellen auftreten. Diese Erklärung, welche gestützt auf dem Verteilungsdiagramm 4.10b plausibel erscheint, würde diesen hohen Wert begründen. Es scheint verhältnismässig viele random walks zu geben, welche im Bereich von 200 Schritten eine erste Nullstelle aufweisen. Dies zeigt auch die Verteilung, dargestellt als Histogramm 4.10b. Ganz allgemein haben nur etwa halb so viel Aktienkurse wie modifizierte random walks eine erste Nullstelle in den ersten 25 Schritten.



**Abbildung 4.10** Histogramme zur Verteilung der Position der ersten Nullstelle

In beiden Histogrammen kann man erkennen, dass die erste Nullstelle mit grosser Wahrscheinlichkeit ganz am Anfang des random walks auftritt. Es muss zusätzlich erwähnt werden, dass random walks nicht nur nach einer geraden Anzahl Schritte eine Nullstelle aufweisen können. Dass alle ersten Nullstellen relativ früh im Verlauf der Aktienkurse auftreten, stimmt sehr gut mit dem Histogramm (a) überein, da auch dort gesehen werden kann, dass im späteren Verlauf des random walks nur noch mit sehr geringer Wahrscheinlichkeit eine erste Nullstelle auftritt. Zusätzlich gilt zu beachten, dass nullstellenfreie random walks auch keine erste Nullstelle besitzen.

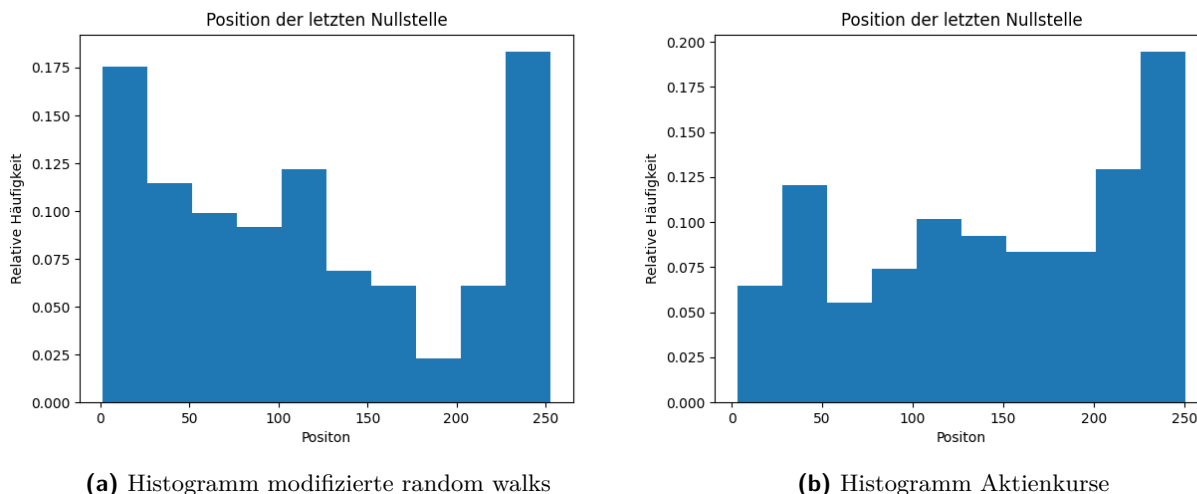
#### 4.2.4 Die letzte Nullstelle



**Abbildung 4.11** Diagramm zur durchschnittlichen Positionen der letzten Nullstelle

Wie bereits bei der Betrachtung der Anzahl Nullstellen liegt der Erwartungswert für die Position der letzten Nullstelle über dem Intervall des Mittelwerts der Aktienkurse. Die Verteilung der letzten Nullstelle nach dem Arcus-Sinus-Gesetz wurde bereits beschrieben, nun werden diese theoretischen Erkenntnisse mit der Simulation und mit den Aktienkursen verglichen.





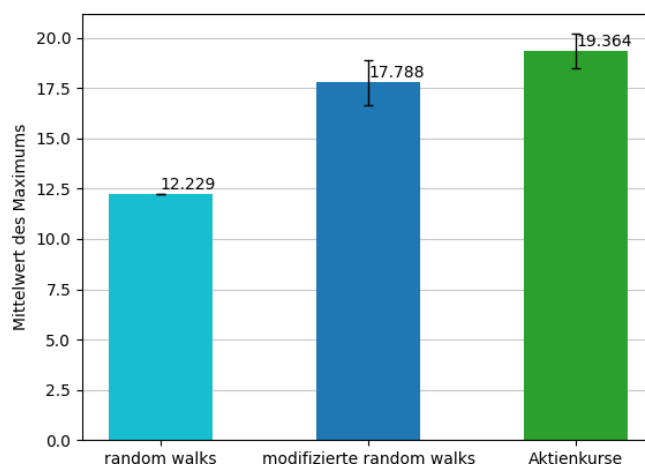
**Abbildung 4.12** Diagramme zur Verteilung der Position der letzten Nullstelle

Betrachtet man das Histogramm (b) erkennt man, dass die linke Seite stark an die linke Seite der Abbildung (a) erinnert, wobei am Ende der Aktienkurse eine viel geringere Häufigkeit der letzten Nullstelle auftritt als durch die theoretischen Erkenntnisse erwartet. Dadurch lässt sich auch der tiefe Durchschnittswert für die Position der letzten Nullstelle erklären.

### 4.2.5 Maximum

Analog zur Analyse der Nullstellen soll nun ein Vergleich zwischen der präsentierten Theorie, der Umsetzung in der Simulation und der Auswertung der Aktienkursen gezogen werden. Auch bezüglich der Extremstellen sind die Erkenntnisse aus der Theorie und der Simulation übereinstimmend, daher liegt der Fokus primär auf dem Vergleich dieser Erkenntnisse mit den Werten und Verteilungen der Aktienkursen.

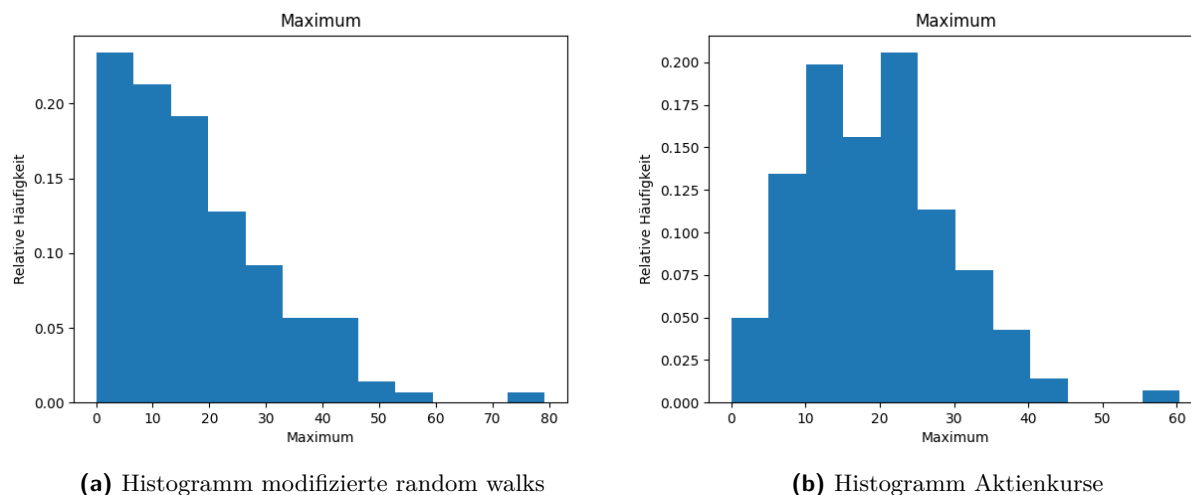
Bezüglich der Theorie konnte festgestellt werden, dass sich Maximum und Minimum vom Betrag her nicht unterscheiden. In diesem Abschnitt soll die Ausprägung der Extremstellen in der Theorie, in der Simulation und in den Aktienkursen verglichen werden.



**Abbildung 4.13** Diagramm zum Mittelwert der Maxima

### Verteilung des Maximums

Werden die Werte der ersten Spalte mit der letzten Spalte verglichen, fällt auf, dass gerade der Durchschnittswert ähnlich ausfällt. Die letzte Spalte zeigt die Auswertung der Beträge aller Extremstellen der Aktienkurse. Es fließen also sowohl Maximum als auch Minimum in diese Werte ein. Dass Maximum und Minimum bei den Aktienkursen so unterschiedlich ausfällt ist vermutlich darauf zurückzuführen, dass die Aktienkurse im Gegensatz zur Annahme in der Theorie eine negative, durchschnittliche Endposition aufweisen. Die erwartete Verteilung einer Extremstelle wurde bereits durch das Diagramm 2.10b präsentiert. Nun soll diese Verteilung mit der Verteilung des Maximums der Aktienkurse verglichen werden.



**Abbildung 4.14** Diagramme zur Verteilung des Maximums

Primär fällt die Verschiebung des Maximalwerte für die relative Wahrscheinlichkeit des Maximums auf. Dabei entspricht die Verteilung der modifizierten random walks in der linken Abbildung stärker der theoretischen Verteilung für einfache, symmetrische random walks. Die genau Begründung für diesen Sachverhalt zu finden, gestaltet sich relativ schwer. Im Grunde genommen besagt diese Abbildung, dass in diesen Zeitabschnitt wenige Aktienkurse den Anfangswert nie überschreiten oder dies nur sehr leicht tun. Bei den modifizierten random walks tritt dies häufiger auf. Diese Verteilung kann auch zur Begründung des höheren Mittelwerts der Aktienkurse genutzt werden, es wird deutlich, dass dieser Mittelwert grösser sein wird, als der entsprechende Mittelwert der modifizierten random walks. Das Histogramm zeigt auch die unterschiedliche Form der Verteilung, wobei die random walks eine Verteilung aufweisen, wo die relative Häufigkeit mit zunehmendem Maximum ständig abnimmt. Die Verteilung des Maximums der Aktienkurse beschreibt eine Form, wo die relative Häufigkeit erst zunimmt, ihr Maximum erreicht und schliesslich wieder abfällt.

Position des Maximums

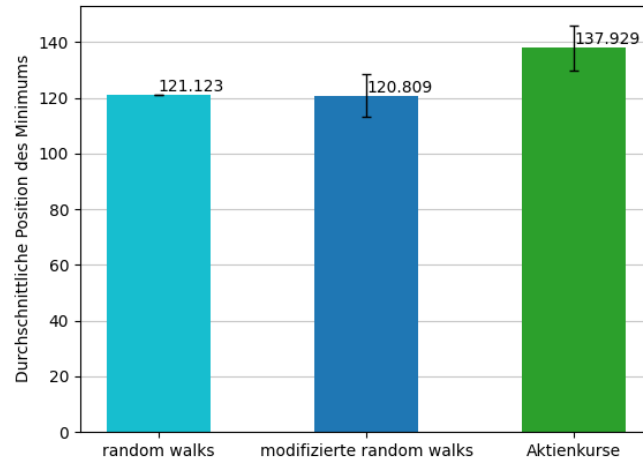
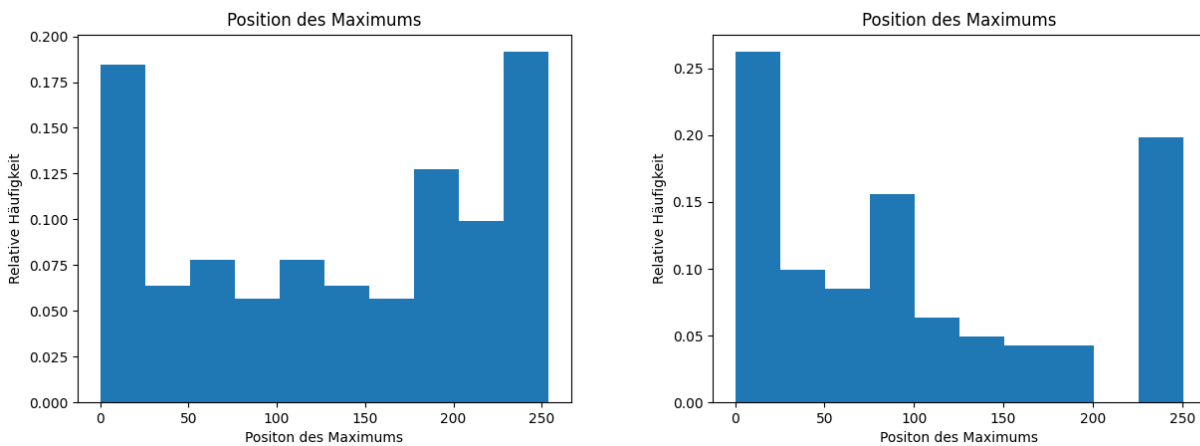


Abbildung 4.15 Diagramm zur durchschnittlichen Position des Maximums

Betrachtet man die Abbildung 4.15 und die beiden Diagramme (4.16) zur Position des ersten Maximums, kann die gleiche Feststellung gemacht werden, wie im Abschnitt zur Verteilung der letzten Nullstelle. Beide Verteilungen, die der letzten Nullstelle und die des ersten Maximums, haben aufgrund des Arcus-Sinus-Gesetz eine U-förmige Verteilung. Nun fällt auf, dass die Aktienkurse in der ersten Hälfte durchaus dieser Verteilung ähneln aber am Ende keine Anhäufung von letzten Nullstellen oder ersten Maxima haben. Auch dieses Phänomen könnte eventuell durch die durchschnittlich negative Endposition der Aktienkurse erklärt werden. Zusätzlich wird durch diese Verteilung sichtbar, weshalb der Durchschnittswert in Abbildung 4.15 unter dem theoretischen Erwartungswert liegt.



(a) Histogramm modifizierte random walks

(b) Histogramm Aktienkurse

Abbildung 4.16 Diagramme zur Verteilung der Position des Maximums

Die beiden Verteilung der Positionen des Maximums ähneln sich stark und weisen in beiden Fällen die erwartete 'U-Form' auf. Bei den modifizierten random walks liegt das Maximum etwas häufiger im letzten Bereich. Allgemein sind diese Unterschiede aber vermutlich primär auf die Zufälligen Abweichungen aufgrund der begrenzten Stichprobengröße zurückzuführen.

### 4.2.6 Minimum

Die Mittelwerte des Minimums zeigen, dass die Aktienkurse und die einfachen, symmetrischen random walks in einem ähnlichen Bereich liegen und die modifizierten random walks im Durchschnitt ein etwas tieferes Minimum aufweisen.

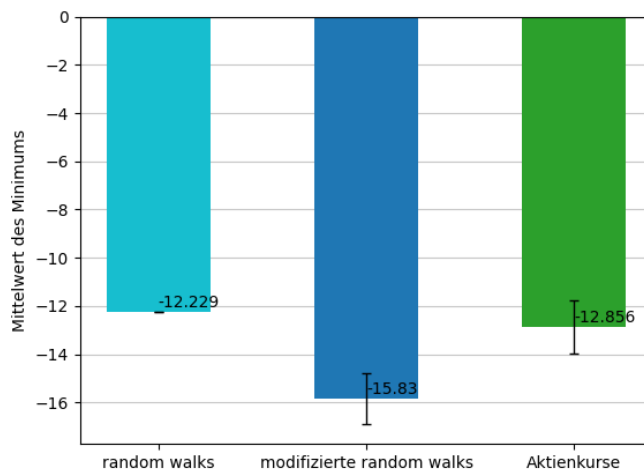


Abbildung 4.17 Diagramm zum Mittelwert der Minima

### Verteilung des Minimums

Die beiden Verteilungen des Minimums, dargestellt als Histogramm, weisen grosse Ähnlichkeit bezüglich ihrer Form auf. Die relative Häufigkeit nimmt ab -50 fast kontinuierlich zu und erreicht ihr Maximum kurz vor 0. Auch hier tritt natürlich eine gewisse Abweichung auf, die aber durchaus zu erwarten ist. Alleine die begrenzte Anzahl an ausgewerteten Aktien und random walks können diese Unterschiede herbeirufen.

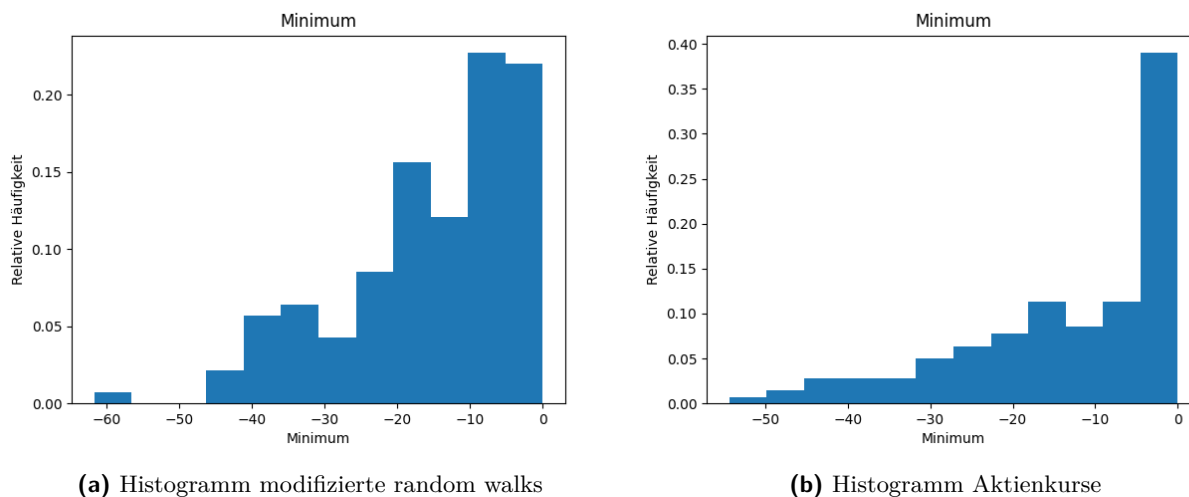
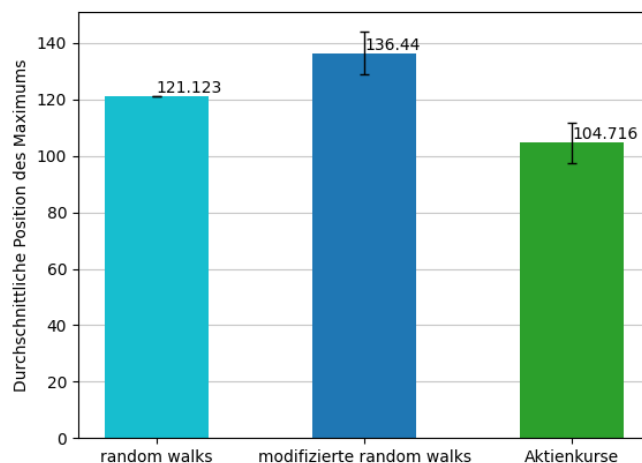


Abbildung 4.18 Diagramme zur Verteilung des Minimums

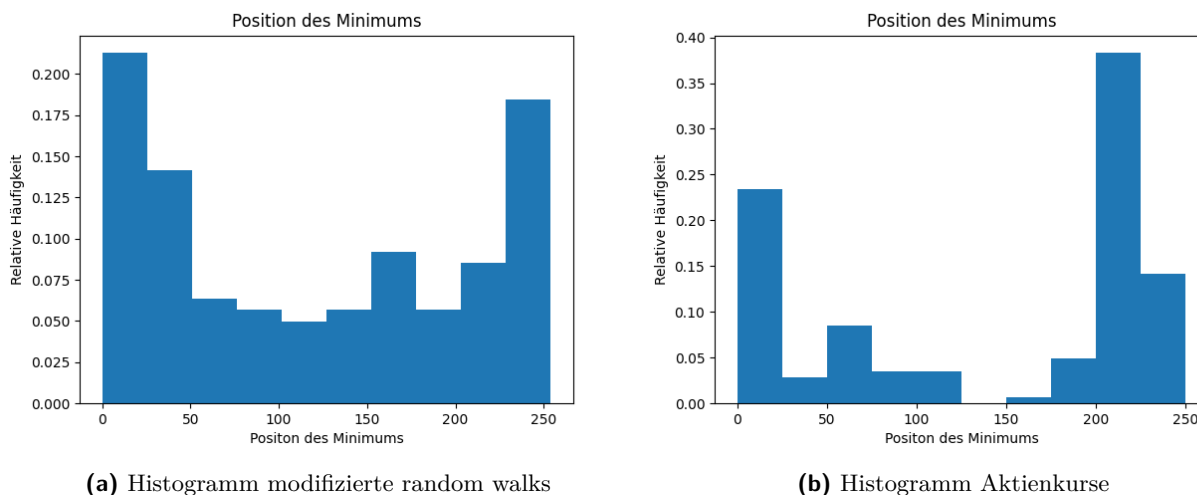
### Position des Minimums

Bezüglich der Position des Minimums liegt der Mittelwert der Aktienkurse unter dem Mittelwert der einfachen, symmetrischen, aber auch den modifizierten random walks. Diese beiden Werte liegen nicht innerhalb des Intervalls  $[\text{Mittelwert} \pm \text{Standardfehler}]$ , was für einen tatsächlich bestehenden Unterschied spricht.



**Abbildung 4.19** Diagramm zur durchschnittlichen Position des Minimums

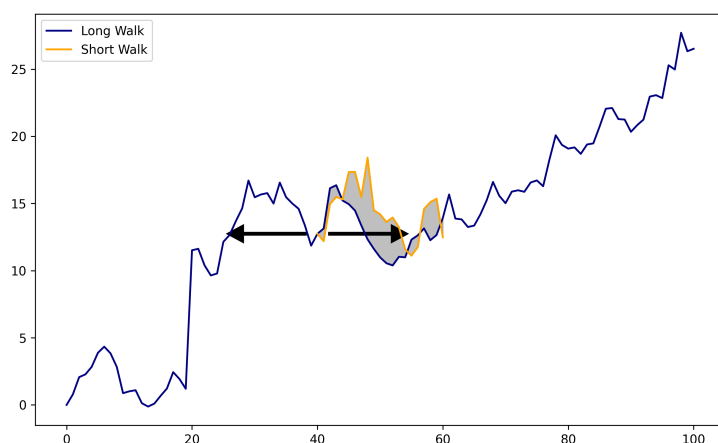
Die Verteilung der beiden Größen ähnelt jeweils der erwarteten 'U-Form'. Diese 'U-Förmigkeit' zeichnet sich jedoch bei den modifizierten random walks etwas deutlicher ab. Die Grundform ist aber in beiden Fällen klar erkennbar.



**Abbildung 4.20** Diagramme zur Verteilung der Position des Maximums

### 4.3 Vergleich durch eine Abschnittsuche innerhalb eines langen random walks

Als weitere Vergleichsmethode soll der normalisierte Kursverlauf einer Aktie mit Abschnitten eines sehr langen random walks verglichen werden. Als Gegenvergleich wird ein random walk einer dem Aktienkurs entsprechenden Länge generiert und auf gleiche Weise als ähnlicher Abschnitt auf dem langen random walk gesucht. Der random walk oder der Aktienkurs als Abschnitt wird jeweils entlang des langen random walk um einen Schritt verschoben, wobei der kurze Pfad mit dem Abschnitt des langen random walks einen Ausgangspunkt teilt. Es wird die Fläche zwischen den beiden Polygonzügen als Indikator für die Ähnlichkeit der beiden Abschnitte berechnet.



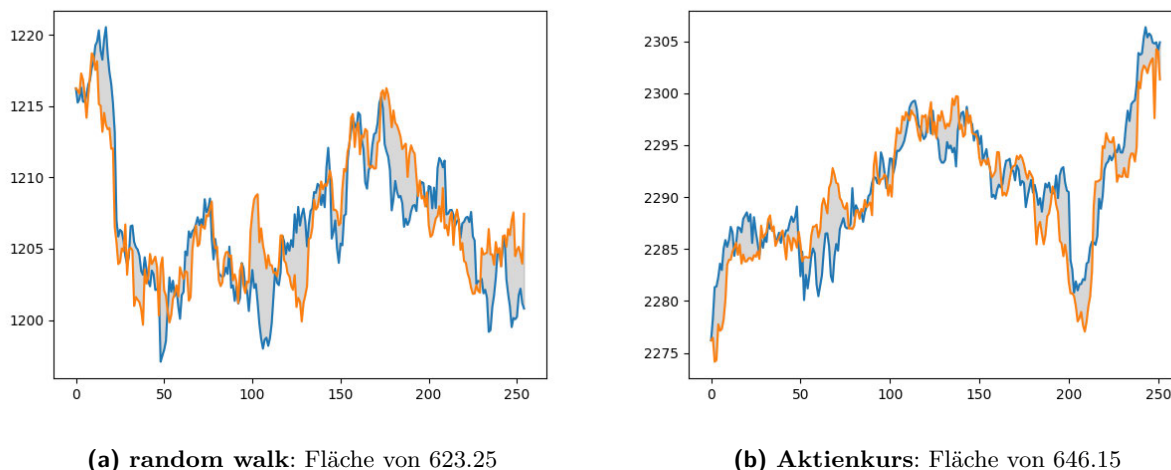
(a) Grafik zur Abschnittsuche auf einem langen random walk



(b) Animation der Abschnittsuche

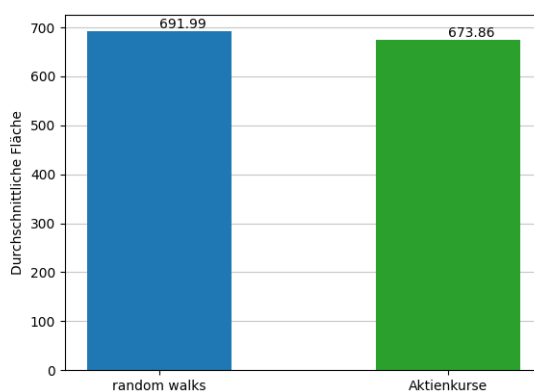
**Abbildung 4.21** Visuelle Erklärung des Prinzips der Abschnittsuche auf einem langen random walk

Betrachtet wird nach Durchlaufen des gesamten langen random walks die minimale Fläche zwischen einem Abschnitt des langen random walks und des kurzen random walks beziehungsweise dem normalisierten Aktienkurs. Der Betrag der Fläche alleine hat natürlich wenig Aussagekraft und hängt auch stark von der Länge des Aktienkursabschnitts und des langen random walks ab. Mit zunehmender Länge des langen random walks steigt die Wahrscheinlichkeit, dass ein Abschnitt gefunden wird, der dem gesuchten Abschnitt stärker ähnelt, gemessen an der dazwischenliegenden Fläche. Um diese Methodik zu verbessern, ist es sinnvoll den langen random walk mit einer möglichst hohen Schrittzahl zu wählen, was in der Realität jedoch mit der Laufzeit des Vergleichsalgorithmus abgewägt werden muss. Von Interesse ist schlussendlich, ob und wie sich die jeweilige Fläche im Vergleich 'random walk mit random walk' und 'Aktienkurs mit random walk' unterscheidet, besonders über mehrere Aktienkurse hinweg.



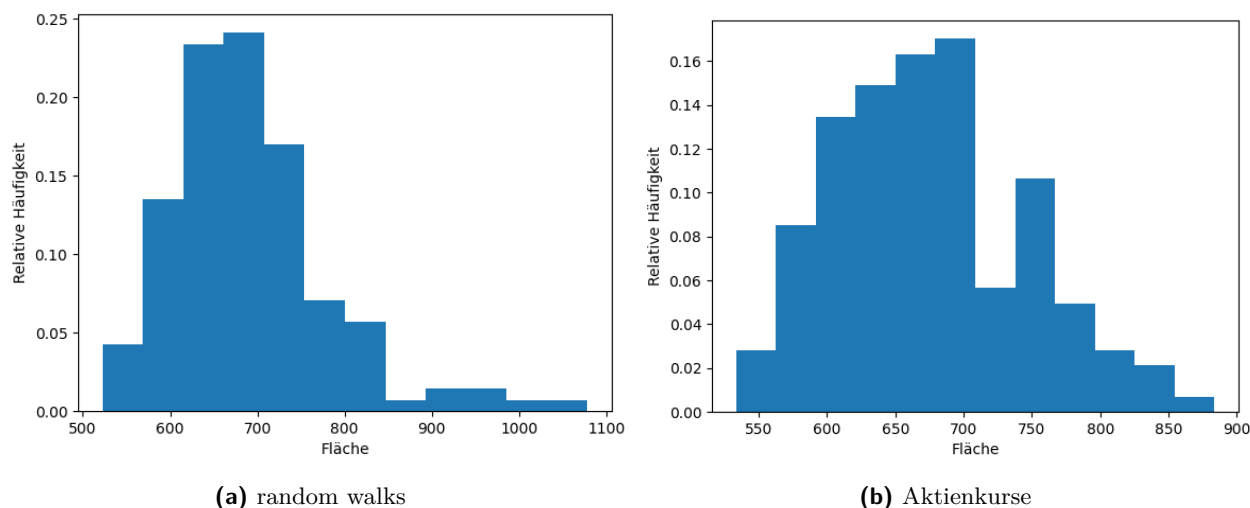
**Abbildung 4.22** Beste Abschnittsübereinstimmung bezüglich Fläche auf einem Random Walk mit  $10^7$  Schritten

Es fällt auf, dass trotz der grossen Länge  $10^7$  Schritten des langen random walks und der dadurch langen Rechenzeit, noch einen beträchtlichen Unterschied zwischen den Abschnitten besteht. Eine Möglichkeit, um allgemein einen besser passenden Abschnitt zu finden, wäre die Verkürzung des zu vergleichenden random walks oder des entsprechenden Zeitraums des Aktienkurses. Im allgemeinen ist es aufgrund der hohen Berechnungszeit kaum umsetzbar, diese Abschnittsuche für alle Aktienkurse im Vergleich zu einem random walk der Länge von 10 Millionen Schritten durchzuführen. Nun muss entweder die Anzahl der zu vergleichenden Aktienkurse oder die Schrittlänge des langen random walks reduziert werden. Eine Reduktion des letzteren Faktors hat sich als bessere Option gezeigt und daher wird die Länge des langen random walks um den Faktor 10 vermindert. Der random walk, auf dem die passenden Abschnitte gesucht werden hat nun noch eine Länge von einer Million Schritte. So kann die Abschnittsuche für alle 141 Aktienkurse, welche nach der Normalisierung noch vorhanden sind, durchgeführt werden. Der gleiche Prozess wird für 141 random walks mit identischer Schrittlänge von 254 Schritten durchlaufen. Auch an dieser Stelle wird eine weitere Optimierung vorgenommen um die Komplexität des Python-Codes bezüglich Laufzeit zu reduzieren, indem der lange random walk nur einmal generiert wird und alle Vergleiche an diesem random walk durchgeführt werden. Durch die beschriebene Methodik wurden neben dem Mittelwert der Flächen auch die Verteilung der Flächen in einem Diagramm visualisiert.



**Abbildung 4.23** Mittelwert der Durchschnittsflächen der Abschnittsuche

Auffallend ist, dass der Mittelwert der Fläche der Aktienkurse geringer ausfällt als der entsprechende Mittelwert der random walks. Die Differenz zwischen den Mittelwerten beträgt 18.13 und fällt dadurch nicht allzu gross aus. Eine gewisse Zufallskomponente beeinflusst diesen Wert, was sich auch an den Grafiken 4.22 sichtbar wird. Die relativ grosse Schrittzahl des langen random walks reicht nicht aus, um einen nahezu übereinstimmenden Abschnitt zu finden. Ein Grund dafür sind die verschiedenen Schrittgrößen. Die Wahrscheinlichkeit, einen genau übereinstimmenden Abschnitt zu finden wäre grösser, entspräche die Schrittgrösse einheitlich der Länge 1. Trotzdem zeigt die Auswertung des Mittelwerts, dass sich Aktienkursabschnitte keine schlechtere Übereinstimmung zeigen. Neben dem Mittelwert soll auch die Verteilung der einzelnen Flächen visuell als Histogramm dargestellt werden. Entsprechendes wird in der folgenden Abbildung gezeigt.



**Abbildung 4.24** Verteilung der Flächen der besten Abschnittsübereinstimmung

Die beiden Histogramme in Abbildung 4.24 zeigen ein leicht unterschiedliches Bild, weisen jedoch grundlegende Gemeinsamkeiten auf. Der Mittelwert liegt, wie bereits diskutiert, an einer ähnlichen Stelle. Eine Anhäufung von entsprechenden Flächen um den Mittelwert ist ebenfalls bei beiden Diagrammen klar erkennbar. Das Histogramm der random walks und dessen bestpassender Abschnitt mit dazwischenliegender Fläche zeigt einige Werte, welche den Wert von 900 übersteigen. Bei den Aktienkursen liegt das Maximum der Fläche unter 900.

Allgemein fällt auf, dass die Aktienkurse im Vergleich mit dem random walk nicht 'schlechter' abgeschnitten haben, der Mittelwert der gefundenen Minimalwerte liegt tiefer. Andererseits kann am Verteilungsdiagramm erkennbar, dass die random walks gelegentlich grössere Zwischenflächen erzeugen, als der Vergleich mit Aktien.



## Kapitel 5

# Diskussion

Es ist von grosser Wichtigkeit, die Ergebnisse und Erkenntnisse dieser Arbeit richtig einzuordnen und zu diskutieren. Es wurden diverse Aspekte im Zusammenhang mit random walks im Allgemeinen und dem Random-Walk-Modell zur Modellierung von Aktienkursen beleuchtet. Die Datenanalyse der Aktienkurse, welche als Vergleichsgrundlage mit den modifizierten random walks dient, verhält sich bezüglich eindeutiger Interpretation am schwierigsten.

In diesem Kapitel werden nun die gewonnenen Erkenntnisse dieser Arbeit mit Bezug auf die Leitfragen diskutiert. Die erste Leitfrage bezieht sich auf Kapitel 2.2 und Kapitel 2.3, denn sie fragt nach den Eigenschaften der einfachen, symmetrischen random walks auf  $\mathbb{Z}$ . Es ist zu betonen, dass nur ein Anschnitt der Eigenschaften präsentiert werden konnte, da random walks eine weitreichende Vielfalt an spannenden Phänomenen aufweisen. Zudem wurde auf das komplette Beweisen aller Eigenschaften verzichtet, damit dieser Teil der Analysen der Eigenschaften von einfachen, symmetrischen random walks auf  $\mathbb{Z}$  nicht übermässig viel Raum einnimmt und mehr Gewicht auf die datengestützte Analyse des modifizierten Modells gelegt werden kann. Trotzdem ist es gelungen, ausgewählte Eigenschaften aufzuzeigen und zu begründen. Dabei lag der Fokus auf Eigenschaften, die im Zusammenhang mit Aktienkursen einen sinnvollen Vergleich boten.

Die zweite Leitfrage beschäftigt sich mit der Suche nach einer passenden Modifikation des Random-Walk-Modells und greift damit den Titel dieser Arbeit auf. Dazu werden erst die Gemeinsamkeiten zwischen Aktienkursen und random walks herausgearbeitet. Gleichzeitig fallen aber auch Unterschiede ins Auge, welche durch Modifikationen aufgefasst werden können. Anhand echter Aktienkursdaten konnte gezeigt werden, dass die einzelnen Schritte der Kursbewegungen, also die Veränderungen der Schlusskurse, nicht normalverteilt sind. Als passender stellte sich eine Student-t-Verteilung heraus, welche besonders Extremevents in Form von sehr grossen Änderungen des Schlusskurses von einem auf den nächsten Tag. Nach einer Normalverteilung, würden diese grossen Kursveränderungen viel seltener auftreten, als in Wirklichkeit beobachtet werden kann. Es wurden verschiedene Modifikationen für die random walks vorgestellt, diskutiert und schliesslich die sinnvollen Modifikationen in einer Simulation umgesetzt. Darauf folgend wurden die Ergebnisse der Modifikationen analysiert und es stellte sich heraus, dass einige Eigenschaften der Aktienkurse signifikant besser repräsentiert werden als durch die unmodifizierten random walks. Hingegen war für gewisse Eigenschaften keine Verbesserung erkennbar. Wie gut diese modifizierten random walks die Kursbewegungen wiedergeben, ist schwer quantifizierbar, weil auch hier ein relativ grosser Interpretationsspielraum existiert.

Eine Eigenschaft von Aktienkursen, welche als Unterschied zu den modifizierten random walks erkennbar wird, ist das sogenannte 'Volatility Clustering'. Es fallen einige Abschnitte auf, wo auf sehr grosse Schrittlängen weitere grosse Schritte folgen. Dieses Phänomen spricht gegen die Random-

Walk-Theorie, weil diese eine Unabhängigkeit der einzelnen Schritte prognostiziert. Diese Eigenschaft wurde nicht mit Hilfe einer Modifikation der random walks implementiert, weil eine entsprechende Modifikation sehr stark in das Konzept der random walks eingreifen würde.

Die vierte Leitfrage handelt von der Korrektheit der Random-Walk-Theorie und ist bezüglich Beantwortung eher offen. Mit Sicherheit bietet diese Arbeit keinen abschliessenden Beweis für die Random-Walk-Theorie, wenn auch dies gar nicht die Absicht dieser Arbeit war. Es konnte gezeigt werden, dass zwischen dem Verhalten des Random-Walk-Modells und echten Aktienkursen relevante Gemeinsamkeiten auftreten, welche für die Gültigkeit der Random-Walk-Theorie sprechen. Ander Faktoren hingegen sprechen wiederum dagegen, wie beispielsweise das beobachtete Phänomen des 'Volatility Clusterings'. Zudem sollte auch erwähnt werden, dass die Methoden der Datenanalysen gewisse Einschränkungen aufweisen, die Interpretationsspielraum lassen. Zudem handelt sich es immer um einen Datenausschnitt aller Kursverläufe an der Schweizer Börse sowie auch wird immer nur mit einer Stichprobe an modifizierten random walks als Vergleich gearbeitet. Daraus entsteht eine statistische Unsicherheit, welche beachtet werden muss.

Schliesslich sollte angemerkt werden, dass die Random-Walk-Theorie im Bereich der Finanzwissenschaft umstritten ist und es sowohl Befürworter als auch Gegner gibt. In der praktischen Anwendung sind Chart- und Fundamentalanalysen weit verbreitet und werden von vielen Anlegern unterstützt. Ob diese Analyseansätze tatsächlich einen Vorteil bringen, welcher ermöglicht, den Markt nachhaltig und langfristig zu schlagen ist offen. Die Random-Walk-Theorie spricht diese Möglichkeit ab. Sie hängt aber auch eng mit dem Konzept der Markteffizienz zusammen, welche die Reaktion des Aktienmarkts auf Informationen beschreibt. Auch hier gehen gewisse Aspekte der Komplexität der Märkte verloren, da einige Idealbedingungen angenommen werden. Beispielsweise wird davon ausgegangen, dass jeder Marktteilnehmer in jeder Situation rational handelt, was in Realität nicht immer der Fall ist, weil Menschen eben auch auf emotionaler Basis Entscheide treffen können. Weiter wird auch davon ausgegangen, dass der Markt 'friktionslos' ist. Es fallen also nirgends Gebühren für das Handeln von Aktien an, was in Realität ebenfalls nicht komplett erfüllt wird.

Allgemein gibt es eine Vielzahl an Modellen zur Darstellung von Kursbewegungen, jedes Modell bringt schliesslich Stärken und Schwächen mit sich und kann immer nur ein simplifizierte Reproduktion der Realität abbilden. Niemals aber wird ein Modell in der Lage sein, jeden Einflussfaktor, welcher auf den Aktienmarkt einwirkt einzeln und korrekt darzustellen, weil weder alle diese Einflüsse, noch dessen genau Auswirkung bekannt ist. Die Schwierigkeit besteht nun darin, ein Modell zu definieren, dass an der richtigen Stelle und im korrekten Ausmass vereinfacht.

Auch in der Fachliteratur kann die Random-Walk-Theorie nicht bewiesen werden und es gibt verschiedene Aspekte, die dafür und dagegen sprechen. Nach der Auswertung und den Vergleichen kann der Schluss gezogen werden, dass diese Arbeit eher dafür spricht, dass gewisse Eigenschaften von Aktienkursen durchaus durch modifizierte random walks abgebildet werden können und dies ein Hinweis darauf ist, dass es sich bei solchen Kursbewegungen tatsächlich um einen zufälligen Prozess handelt.

Zuletzt soll darauf eingegangen werden, ob diese Arbeit Kursvorhersagen ermöglicht. Der aufmerksame Leser konnte sich die Antwort auf jede Fragestellung sicherlich schon selbst ausmachen. Eine solche Vorhersage der Kursbewegungen durch modifizierte random walks ist ausgeschlossen, da dieses Modell auf der Random-Walk-Theorie basiert, dessen zentrale Aussage, die Unvorhersehbarkeit und Zufälligkeit von Aktienkurse ist. Unabhängig davon, ob Aktienkurse grundsätzlich durch modifizierte random walks modelliert werden können oder nicht, lässt sich durch diesen Ansatz keine Vorhersage treffen. Im ersten Fall, dass nachgewiesen werden könnte, dass sich Aktienkurse nicht durch modifizierte random walks modellieren lassen, erübrigt sich eine mögliche Vorhersage durch modifizierte random walks. Im zweiten Fall, dass ein Zusammenhang zwischen random walks und

Kursbewegungen festgestellt werden kann und eine passende Modifizierung gefunden wird, bietet sich daraus ebenfalls keine Möglichkeit zukünftige Bewegungen der Aktienkurse zu prognostizieren. Es wäre ein Hinweis darauf, dass eine Vorhersage nicht wegen der gewählten Methode unmöglich ist, sondern grundsätzlich durch keine erdenkliche Methode möglich wäre. Daraus liesse sich schliessen, dass auch keine weiteren Bemühungen zur Vorhersage erbracht werden sollten, da es sich um Zeitverschwendung handeln würde.

## Kapitel 6

# Schlusswort

Vor einem Jahr war ich der festen Überzeugung, meine Maturaarbeit auf keinen Fall im Bereich der Mathematik zu schreiben. Im Verlauf der Themenwahl hat sich diese Meinung plötzlich geändert und nun, am Ende dieser Arbeit kann ich bezeugen, dass es keine Fehlentscheidung war, mich für genau dieses Thema zu entscheiden. Diese Arbeit bot mir Gelegenheit, mich mit sehr interdisziplinären Themenbereichen auseinanderzusetzen. Besonders die Abwechslung zwischen dem Erarbeiten der Theorie, dem Programmieren der Simulationen und der Datenauswertung war sehr bereichernd. Gerade das Programmieren der Simulationen half extrem, die erarbeitete Theorie schneller nachvollziehen zu können. Zu Beginn war ich mir nicht der Grösse und Komplexität des Themenbereichs der random walks - etwas so simples, wie ein Zusammenhängen mehrer Münzenwürfe, bewusst. Das Erarbeiten der Theorie war für mich persönlich die grösste Herausforderung und forderte sehr viel Geduld. Gleichzeitig bot dies mir eine Gelegenheit, mich in diverse Themenbereiche der Mathematik einzulesen und ich kam erstmals mit wissenschaftlicher Literatur aus diesem Themenbereich in Berührung.

Der Aspekt des Praxisbezugs und der Verbindung zwischen Mathematik und Finanzökonomie bot Vielseitigkeit und Abwechslung. Ich konnte mich in den Bereichen der Arbeit zu vertiefen, welche mich besonders interessierten oder mir einfach mehr Freude bereiteten. Gerade die Datenanalyse und die damit verbundene Vertiefung in der Programmiersprache 'Python' gefiel mir sehr. Allgemein konnte diese Arbeit, welche ursprünglich als Maturaarbeit an der Kantonsschule Sursee entstand, im Rahmen von Schweizer Jugend forscht überarbeitet und erweitert werden, was sich meiner Meinung nach stark positiv auswirkte.

Nun möchte ich mich bei allen Personen bedanken, die zum Entstehen dieser Arbeit beigetragen haben. Besonderer Dank gebührt Herrn Michael Muri, der meine Arbeit betreut hat. Einerseits danke ich ihm für die ausführliche Beratung bezüglich der Themenwahl und den hilfreichen Vorschlägen und Einschätzungen bezüglich der Machbarkeit. Andererseits möchte ich mich für die hilfreichen Kommentare und Anregungen während dem Erstellen dieser Arbeit bedanken.

Weiter gebührt meinem betreuenden Experten Dr. Xavier Richard, welcher die Weiterentwicklung dieser Arbeit im Rahmen von Schweizer Jugend forscht tatkräftig unterstützt hat, grosser Dank. Ich möchte mich für die motivierenden Gespräche sowie die hilfreichen Vorschläge und Anregungen bedanken. Ebenfalls habe ich Dr. Xavier Richard die grundlegende Idee für den Vergleich durch eine Abschnittsuche in einem langen random walk (4.3) zu verdanken.

Natürlich möchte ich mich auch bei der Stiftung Schweizer Jugend forscht für die Möglichkeit, diese Arbeit weiterzuführen und viele Erfahrungen zu sammeln, bedanken.

# Literaturverzeichnis

- [1] Sven Erick Alm. Simple random walk. [http://www2.math.uu.se/~sea/kurser/stokprocmn1/slumpvandring\\_eng.pdf](http://www2.math.uu.se/~sea/kurser/stokprocmn1/slumpvandring_eng.pdf), 2002. Besucht am 05.09.2023.
- [2] Artem Bogdanov. Technische analyse : Darstellung und kritik. Master's thesis, Karl-Franzens-Universität Graz, 2014.
- [3] Eugene F. Fama. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417, 1970.
- [4] William Feller. The numbers of zeros and of changes of sign in a symmetric random walk. *L'Enseignement Mathématique*, 3(3):229–235, 1957.
- [5] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1 of *Wiley series in probability and mathematical statistics*. John Wiley & Sons, Inc., 3 edition, 1968.
- [6] S.G. Foss Sergey, Serguei, Korshunov Dmitry, and Zachary Stan. *An Introduction to Heavy-Tailed and Subexponential Distributions*, volume 38 of *Springer Series in Operations Research and Financial Engineering*, chapter Heavy-Tailed and Long-Tailed Distributions, pages 7–38. Springer, New York, 01 1970.
- [7] Norbert Henze. *Irrfahrten-Faszination der Random Walks*. Springer, Wiesbaden, 2018.
- [8] D. Johnston. An introduction to random walks. <https://math.uchicago.edu/~may/VIGRE/VIGRE2011/REUPapers/Johnston.pdf>, 2011. Besucht am 20.08.2023.
- [9] R. Korn. Elementare finanzmathematik. Technical Report 39, Fraunhofer Institut Techno- und Wirtschaftsmathematik, 2002.
- [10] Markus Oestreich and Oliver Romberg. *Keine Panik vor Statistik!*, volume 1. Springer, 1 edition, 2009.
- [11] Godwin Chigozie Okpara. Stock market prices and the random walk hypothesis: Further evidence from nigeria. *Journal of Economics and International Finance*, 2(4):49, 2010.
- [12] Bhanu Pant and TR Bishnoi. Testing random walk hypothesis for indian stock market indices. In *Research Paper Presented in UTI Capital Market Conference Proceedings*, pages 1–15. Citeseer, 2001.
- [13] Svetlozar T Rachev, Young Shin Kim, Michele L Bianchi, and Frank J Fabozzi. *Financial models with Lévy processes and volatility clustering*. John Wiley & Sons, 2011.
- [14] Amar Shah, Andrew Wilson, and Zoubin Ghahramani. Student-t Processes as Alternatives to Gaussian Processes. In Samuel Kaski and Jukka Corander, editors, *Proceedings of the Seven-*

*teenth International Conference on Artificial Intelligence and Statistics*, volume 33 of *Proceedings of Machine Learning Research*, pages 877–885, Reykjavik, Iceland, 22–25 Apr 2014. PMLR.

- [15] Kin Sibanda Tafadzwa T. Chitenderu, Andrew Maredza. The random walk theory and stock prices: Evidence from johannesburg stock exchange. *International Business & Economics Research Journal*, 13(6), 2014.

# Abbildungsverzeichnis

|      |  |    |
|------|--|----|
| 1    | symmetrischer random walk mit t-verteilten Schrittlängen . . . . .   | 1  |
| 2.1  | Ein random walk auf $\mathbb{Z}$ der Länge $n = 6$ , dargestellt auf einem Zahlenstrahl . . . . .  | 8  |
| 2.2  | Ein random walk auf $\mathbb{Z}$ der Länge $n = 200$ . . . . .   | 8  |
| 2.3  | random walk mit Nullstelle bei $2k = 2$ und $2k = 6$ . . . . .   | 11 |
| 2.4  | Wahrscheinlichkeitsverteilung der Anzahl Nullstellen für einen random walk der Länge $2n = 254$ . . . . .  | 13 |
| 2.5  | Wahrscheinlichkeitsverteilung der ersten Nullstelle $F_{40}$ . . . . .   | 15 |
| 2.6  | Letzte Nullstelle eines random walks zum Zeitpunkt $L_{2n}$ [7] . . . . .  | 16 |
| 2.7  | Dichte- und Verteilungsfunktion der Arcus-Sinus-Verteilung . . . . .   | 18 |
| 2.8  | Wahrscheinlichkeitsverteilung der letzten Nullstelle $L_{40}$ . . . . .  | 18 |
| 2.9  | Spiegelung des Teilweges ab erstmaligem Erreichen der Höhe $k$ an der Achse $y = k$ [7] . . . . .  | 20 |
| 2.10 | Vergleich zwischen der Binomialverteilung (2.10a) und der Verteilung der Wahrscheinlichkeit des Maximums $M_n$ (2.10b) . . . . .                             | 23 |
| 2.11 | Stabdiagramm zur Verteilung von $E_{254}$ . . . . .  | 25 |
| 3.1  | Vergleich zwischen einem Aktienkurs und einem random walk mit gleicher Länge . . . . .   | 26 |
| 3.2  | Normalverteilung mit $\mu = 0$ und unterschiedlichen Standardabweichungen $\sigma$ . . . . .   | 27 |
| 3.3  | Histogramm zur Verteilung der Kursveränderungen der Aktienkurse und Graph der Normalverteilung . . . . .   | 28 |
| 3.4  | Vergleich zwischen der Modellierung von grossen Schrittgrössen im Randbereich der Normal- und t-Verteilung . . . . .   | 29 |
| 3.5  | Student-t-Verteilung mit verschiedenen $\nu$ . . . . .   | 30 |
| 3.6  | Histogramm zu den Schrittgrössen der Aktienkurse und Graph der t-Verteilung mit $\nu = 4$ . . . . .  | 31 |
| 3.7  | symmetrischer random walk mit t-verteilten Schrittlängen . . . . .   | 32 |
| 4.1  | Kursverlauf von 20 Aktien im Zeitraum von 30.12.2021 bis 30.12.2023 . . . . .  | 35 |
| 4.2  | Aussortierte Aktienkurse im Zeitraum vom 31.12.2020 bis zum 31.12.2023 mit einer Toleranz von 400 Tagen ohne Kursbewegung und 100 Tagen Verkürzung . . . . . | 36 |
| 4.3  | Vergleich zwischen dem originalen und dem normalisierten Aktienkurs der HBM Healthcare Investments AG (HBMN.SW) . . . . .                                    | 37 |
| 4.4  | Visualisierung des Vergleichs zwischen kurzzeitigen Aktienkursbewegungen und modifizierten random walks gleicher Länge . . . . .                             | 39 |
| 4.5  | Visualisierung des Vergleichs zwischen Aktienkursbewegungen innerhalb eines Jahres und modifizierten random walks gleicher Länge . . . . .                   | 41 |
| 4.6  | Graphischer Vergleich zwischen modifizierten random walks und Aktienkursen im Zeitraum von 10 Jahren . . . . .   | 42 |
| 4.7  | Diagramm zu den durchschnittlichen Anzahl Nullstellen . . . . .  | 44 |

|      |  |    |
|------|--|----|
| 4.8  | Diagramme zur Verteilung der Anzahl Nullstellen . . . . .  | 45 |
| 4.9  | Diagramm zur durchschnittlichen Positionen der ersten Nullstelle . . . . .                               | 45 |
| 4.10 | Histogramme zur Verteilung der Position der ersten Nullstelle . . . . .                                  | 46 |
| 4.11 | Diagramm zur durchschnittlichen Positionen der letzten Nullstelle . . . . .                              | 46 |
| 4.12 | Diagramme zur Verteilung der Position der letzten Nullstelle . . . . .                                   | 47 |
| 4.13 | Diagramm zum Mittelwert der Maxima . . . . .   | 47 |
| 4.14 | Diagramme zur Verteilung des Maximum . . . . .   | 48 |
| 4.15 | Diagramm zur durchschnittlichen Position des Maximums . . . . .  | 49 |
| 4.16 | Diagramme zur Verteilung der Position des Maximums . . . . .   | 49 |
| 4.17 | Diagramm zum Mittelwert der Minima . . . . .   | 50 |
| 4.18 | Diagramme zur Verteilung des Minimums . . . . .  | 50 |
| 4.19 | Diagramm zur durchschnittlichen Position des Minimums . . . . .  | 51 |
| 4.20 | Diagramme zur Verteilung der Position des Maximums . . . . .   | 51 |
| 4.21 | Visuelle Erklärung des Prinzips der Abschnittsuche auf einem langen random walk . .                      | 52 |
| 4.22 | Beste Abschnittsübereinstimmung bezüglich Fläche auf einem Random Walk mit $10^7$<br>Schritten . . . . . | 53 |
| 4.23 | Mittelwert der Durchschnittsflächen der Abschnittsuche . . . . .   | 53 |
| 4.24 | Verteilung der Flächen der besten Abschnittsübereinstimmung . . . . .                                    | 54 |

## Tabellenverzeichnis

|     |  |    |
|-----|--|----|
| 4.1 | Importierte Kursdaten der Aktie ' <b>ABBN.SW</b> ' . . . . .   | 34 |
| 4.2 | Vergleich der theoretischen Werte für einen unmodifizierten random walk mit den<br>Werten der Aktienkurse und den Werten der Simulation der modifizierten random walks | 43 |



## Anhang A

# Programmcode zur Datenanalyse

An dieser Stelle sollte darauf hingewiesen werden, dass die einzelnen Python-Files nur zusammenhängend mit dem ganzen Projektordner funktionsfähig sind.

### A.1 'main.py'

**Listing A.1** Das zentrale Python-File, welches direkt ausgeführt wird

```

1 import matplotlib.pyplot as plt
2 import statistics
3 import pandas as pd
4 from src.plot import Plot
5 from src.random_walk_generator import RandomWalkGenerator
6 from src.stock_data import StockDataImport, StockDataList
7 from src.search_algorithm import SearchAlgorithm
8 from src.histogram import Histogram
9
10 # -----
11
12 # TODO: Set options
13
14 plot_original_stock = True
15 plot_all_stocks = False
16 plot_all_random_walks = False
17 random_walk_search_algorithm = False
18 save_figs = False
19 # characteristics analysis
20 analysis_of_characteristics = False
21 option = 1 # option = 1 -> stock || option = 2 -> random walk
22 importNewData = False
23 if importNewData:
24     import src.stocks
25
26 # -----
27
28
29 ticker_symbols = pd.read_csv('new_ticker_list.csv')
30 ticker_list = ticker_symbols['Ticker'].tolist()
31
32 random_walk_1 = RandomWalkGenerator(10 ** 3)
33 random_walk_2 = RandomWalkGenerator(254)
34 random_walk_1.generate_random_walk()
35 random_walk_2.generate_random_walk()
36
37 walk1 = random_walk_1.get_random_walk()
38 # walk2 = random_walk_2.get_random_walk()
39
40 if option == 1:
41     option_type = 'stock'
42 else:
43     option_type = 'random_walk'
44
45 # two random walks are plotted and the area inbetween is calculated
46 # Plot(walk1, walk2, "Random Walk 1", "Random Walk 2").plot_functions()
47

```

```

48
49 if plot_original_stock:
50     # Download of stock data
51     walk = StockDataImport("ZURN.SW", "2022-12-30", "2023-12-30")
52     walk.plot_original_data()
53     if save_figs:
54         plt.savefig('original_stock.png', dpi=300)
55     plt.show()
56     walk.plot_data()
57     if save_figs:
58         plt.savefig('normalized_stock.png', dpi=300)
59     plt.show()
60     walk.plot_changes()
61     plt.show()
62
63     walk.print_data()
64     walk.get_data().to_csv('data_excerpt.csv')
65
66 if random_walk_search_algorithm:
67
68     results_area_stock = []
69     results_area_rw = []
70
71     #for ticker in ticker_list:
72     ticker = ticker_list[1]
73     random_walk_2 = RandomWalkGenerator(254)
74     random_walk_2.generate_random_walk()
75     walk2 = random_walk_2.get_random_walk()
76
77     result_random_walk = SearchAlgorithm(walk1, walk2).compare_with_same_starting_point()
78     Plot(result_random_walk[1], result_random_walk[2], "Random-Walk-1", "Random-Walk-2").
79     plot_functions()
80     print("min.-area-between-a-random-walk-and-another-random-walk:-", result_random_walk[0])
81     results_area_rw.append(result_random_walk[0])
82     # To be able to compare the stock data with the random walk, the stock data gets normalized: It
83     # needs an average
84     # step length of 1
85     stock1 = StockDataList(ticker)
86     stock_data = stock1.get_changes()
87     result_stock = SearchAlgorithm(walk1, stock_data.cumsum()).compare_with_same_starting_point()
88     print(f"min.-area-between-{ticker}-data-and-random-walk:-", result_stock[0])
89     results_area_stock.append(result_stock[0])
90     Plot(result_stock[1], result_stock[2], "Random-Walk-1", "Random-Walk-2").plot_functions()
91
92     Histogram(results_area_stock, True, 'Flaeche-zwischen-Aktienkurs-und-dem-besten-Abschnitt-des-
93     random-walks',
94     , option_type, 'Flaeche', bins=12, show=True).plot()
95     Histogram(results_area_rw, True, 'Flaeche-zwischen-Aktienkurs-und-dem-besten-Abschnitt-des-
96     random-walks',
97     , option_type, 'Flaeche', bins=12, show=True).plot()
98
99     print('Avarage-Area-between-stocks-and-section-of-the-long-random-walk:-',
100     sum(results_area_stock) / len(results_area_stock))
101     print('Avarage-Area-between-a-random-walk-and-section-of-the-long-random-walk:-',
102     sum(results_area_rw) / len(results_area_rw))
103 n = 250
104
105 if plot_all_stocks:
106     # Plot of multiple stock prices
107     for ticker in ticker_list:
108         StockDataList(ticker).plot_data()
109         n = len(StockDataList(ticker_list[0]).get_changes()) - 1
110         if save_figs:
111             plt.savefig('stocks_3.png', dpi=300)
112         plt.show()
113
114 # Plot of x random walks
115 if plot_all_random_walks:
116     # Plot of multiple stock prices
117     for i in range(len(ticker_list)):
118         walk = RandomWalkGenerator(n)
119         walk.generate_random_walk()
120         walk.plot_random_walk(False)
121         if save_figs:
122             plt.savefig('random_walks_20d_4.png', dpi=300)
123         plt.show()
124
125 if analysis_of_characteristics:
126     # create empty lists for the characteristics
127     number_of_zeros = []
128     first_zeros = []
129     last_zeros = []
130     minimum_list = []

```

```

128     maximum_list = []
129     positions_minimum = []
130     positions_maximum = []
131     end_positions = []
132     no_zeros = 0
133
134 # every stock or random walk gets analysed individually and the result gets stored in the list
135 for i in range(len(ticker_list)):
136     if option == 1:
137         walk = StockDataList(ticker_list[i]) # stock data of given ticker gets stored as 'walk
138
139     if option == 2:
140         walk = RandomWalkGenerator(254) # random walk object with chosen number of steps is
141             generated and stored
142     zero_crossings = walk.find_zero_crossings()
143     number_of_zeros.append(zero_crossings[1])
144     maximum, minimum, position_max, position_min, end_position = walk.find_min_max()
145     maximum_list.append(maximum)
146     minimum_list.append(minimum)
147     positions_maximum.append(position_max)
148     positions_minimum.append(position_min)
149     if zero_crossings[1] != 0:
150         first_zeros.append(zero_crossings[2])
151         last_zeros.append(zero_crossings[3])
152     if zero_crossings[1] == 0:
153         no_zeros += 1
154     end_positions.append(end_position)
155
156 # Calculate means
157 avg_number_of_zeros = sum(number_of_zeros) / len(ticker_list)
158 avg_no_zeros = no_zeros / len(ticker_list)
159 avg_first_zeros = sum(first_zeros) / len(first_zeros)
160 avg_last_zeros = sum(last_zeros) / len(last_zeros)
161 avg_maximum = sum(maximum_list) / len(maximum_list)
162 avg_minimum = sum(minimum_list) / len(minimum_list)
163 avg_position_max = sum(positions_maximum) / len(positions_maximum)
164 avg_position_min = sum(positions_minimum) / len(positions_minimum)
165 avg_end_position = sum(end_positions) / len(end_positions)
166
167 # Calculate standard deviation
168 stdev_number_of_zeros = statistics.stdev(number_of_zeros)
169 stdev_first_zeros = statistics.stdev(first_zeros)
170 stdev_last_zeros = statistics.stdev(last_zeros)
171 stdev_maximum = statistics.stdev(maximum_list)
172 stdev_minimum = statistics.stdev(minimum_list)
173 stdev_position_maximum = statistics.stdev(positions_maximum)
174 stdev_position_minimum = statistics.stdev(positions_minimum)
175 stdev_end_position = statistics.stdev(end_positions)
176
177 # histogram number of zeros
178 Histogram(number_of_zeros, save_figs, 'Anzahl-Nullstellen', option_type, xlabel="Anzahl-
179     Nullstellen",
180     bins=12).plot()
181
182 # histogram first zero
183 Histogram(first_zeros, save_figs, 'Position-der-ersten-Nullstelle', option_type, xlabel="
184     Positon", ).plot()
185
186 # histogram last zero
187 Histogram(last_zeros, save_figs, 'Position-der-letzten-Nullstelle', option_type, xlabel="
188     Positon", ).plot()
189
190 # histogram max
191 Histogram(maximum_list, save_figs, 'Maximum', option_type, xlabel="Maximum", bins=12).plot()
192
193 # histogram min
194 Histogram(minimum_list, save_figs, 'Minimum', option_type, xlabel="Minimum", bins=12).plot()
195
196 # histogram position_max
197 Histogram(positions_maximum, save_figs, 'Position-des-Maximums',
198     option_type, xlabel="Positon-des-Maximums").plot()
199
200 # histogram position_min
201 Histogram(positions_minimum, save_figs, 'Position-des-Minimums',
202     option_type, xlabel="Positon-des-Minimums").plot()
203
204 # Print averages
205 print('-----MITTELWERTE-----')
206 print('NULLSTELLEN')
207 print("durchschnittliche-Anzahl-Nullstellen:", avg_number_of_zeros)
208 print("durchschnittliche-Position-erste-Nullstelle:", avg_first_zeros)
209 print("durchschnittliche-Position-letzte-Nullstelle:", avg_last_zeros)
210 print("Anteil-nullstellenfreie-Walks:", avg_no_zeros)
211 print('-----')

```

```

207 print('EXTREMSTELLEN')
208 print("Durchschnittliches -Maximum:", avg_maximum)
209 print("Durchschnittliches -Minimum:", avg_minimum)
210 print("Durchschnittliche -Position -Maximum:", avg_position_max)
211 print("Durchschnittliche -Position -Minimum:", avg_position_min)
212 print("Durchschnittliche -Endposition:", avg_end_position)
213 print('-----STANDARDABWEICHUNG-----')
214 # Print standard diviation & error
215 print('NULLSTELLEN')
216 print("Standardabweichung -Anzahl -Nullstellen:", stdev_number_of_zeros)
217 print("Standardabweichung -Position -erste -Nullstelle:", stdev_first_zeros)
218 print("Standardabweichung -Position -letzte -Nullstelle:", stdev_last_zeros)
219 print('-----')
220 print('EXTREMSTELLEN')
221 print("Standardabweichung -Maximum:", stdev_maximum)
222 print("Standardabweichung -Minimum:", stdev_minimum)
223 print("Standardabweichung -Position -Maximum:", stdev_position_maximum)
224 print("Standardabweichung -Position -Minimum:", stdev_position_minimum)
225 print("Standardabweichung -Endposition:", stdev_end_position)

```

## A.2 'random walk generator.py'

**Listing A.2** Generierung und graphische Darstellung von random walks sowie Definiton der Funktionen zur Auswertung dessen Null- und Extremstellen

```

1 import matplotlib.pyplot as plt
2 import numpy as np
3 from scipy import stats
4
5
6 class RandomWalkGenerator:
7     def __init__(self, number_of_steps):
8         self.number_of_steps = number_of_steps
9         self.walk = None
10
11     def generate_random_walk(self):
12         self.walk = np.zeros((self.number_of_steps + 1))
13         for i in range(self.number_of_steps):
14             step_length = np.abs(stats.t.rvs(df=4))
15             direction = np.random.choice([-1, 1])
16             self.walk[i + 1] = self.walk[i] + direction * step_length
17
18     def get_random_walk(self):
19         if self.walk is None:
20             raise ValueError("Random walk not generated yet. Call generate_walk() first.")
21         return self.walk
22
23     def plot_random_walk(self, show):
24         x_coordinates = np.arange(0, self.number_of_steps + 1)
25         plt.grid(False)
26         plt.plot(x_coordinates, self.walk, alpha=0.6, label="Random walk")
27         if show:
28             plt.show()
29
30     def find_zero_crossings(self):
31
32         self.generate_random_walk()
33         changes_cumsum = self.get_random_walk().tolist()
34         last_change = changes_cumsum[1]
35         zero_crossings = []
36         num_zero_crossings = 0
37
38         for i in range(1, len(changes_cumsum)):
39             if last_change <= 0 < changes_cumsum[i]:
40                 zero_crossings.append(i - 1)
41                 num_zero_crossings += 1
42             elif last_change >= 0 > changes_cumsum[i]:
43                 zero_crossings.append(i - 1)
44                 num_zero_crossings += 1
45             last_change = changes_cumsum[i]
46
47         if num_zero_crossings != 0:
48             last_zero_crossings = zero_crossings[-1]
49             first_zero_crossings = zero_crossings[0]
50         else:
51             first_zero_crossings = None
52             last_zero_crossings = 0
53
54         return zero_crossings, num_zero_crossings, first_zero_crossings, last_zero_crossings
55
56     def find_min_max(self):

```

```

57         self.generate_random_walk()
58         changes_cumsum = self.get_random_walk().tolist()
59         maximum = max(changes_cumsum)
60         minimum = min(changes_cumsum)
61         position_maximum = changes_cumsum.index(maximum)
62         position_minimum = changes_cumsum.index(minimum)
63         end_position = changes_cumsum[-1]
64
65     return maximum, minimum, position_maximum, position_minimum, end_position
66

```

### A.3 stocks.py

**Listing A.3** Download der Kursdaten anhand einer Ticker-Liste, abspeichern der Daten in einem CSV-File und Aktualisierung der Ticker-Liste

```

1 import yfinance as yf
2 import pandas as pd
3 import csv
4 from src.stock_data import StockDataList, is_ticker_to_remove, is_length_too_short
5
6 download_new_data = True
7
8 if download_new_data:
9
10     ticker_list = pd.read_csv('ticker_symbols.csv')
11     ticker_list = ticker_list['ticker'].tolist()
12     ticker_list = [ticker + '.SW' for ticker in ticker_list]
13
14     # time range
15     start_date = "2020-12-31" # "jjjj-mm-dd"
16     end_date = "2022-12-31"
17
18     # Download of stock data
19     stock_data = pd.DataFrame()
20     for ticker in ticker_list:
21         df = yf.download(ticker, start=start_date, end=end_date)
22         df['Ticker'] = ticker # Add a column for the ticker symbol
23         stock_data = pd.concat([stock_data, df])
24
25     # Save the data as CSV file
26     stock_data.to_csv('stock_data.csv')
27
28     old_ticker_list = ticker_list[:] # creat copy of ticker_list
29
30     # remaining tickers
31     new_ticker_list = []
32
33     for ticker in old_ticker_list:
34         # get changes
35         changes = StockDataList(ticker).get_changes()
36         reference_length = len(StockDataList('ABBN.SW').get_changes())
37
38         if (not is_ticker_to_remove(changes, reference_length * 0.1)
39             and not is_length_too_short(changes, reference_length * 0.01)
40             and not changes == []):
41             # add ticker to the new list
42             new_ticker_list.append(ticker)
43
44     # Anzahl der entfernten Ticker ausgeben
45     print("Anzahl der entfernten Ticker:", len(old_ticker_list) - len(new_ticker_list))
46     print(len(old_ticker_list))
47
48     # save new ticker_list
49     with open('new_ticker_list.csv', 'w') as myfile:
50         wr = csv.writer(myfile, quoting=csv.QUOTE_ALL, delimiter='\n')
51         wr.writerow(['Ticker'] + new_ticker_list)
52
53     # Aktualisiere ticker_list mit der neuen Liste
54     ticker_list = new_ticker_list

```

### A.4 stock data.py

**Listing A.4** Auswertung und Normalisierung der Aktienkursdaten aus dem CSV-File, Definition der Funktionen zur Analyse der Null- und Extremstellen der Aktienkurse; Download von Einzelaktienkursen und dessen Auswertung und Visualisierung

```

1 import matplotlib.pyplot as plt
2 import numpy
3 import numpy as np
4 import pandas as pd
5 import yfinance as yf
6
7 stock_data = pd.read_csv('stock_data.csv')
8
9
10 # Works with every ticker | Data will be imported
11 class StockDataImport:
12     def __init__(self, ticker, start_date, end_date):
13         self.ticker = ticker
14         self.start_date = start_date
15         self.end_date = end_date
16         self.data = yf.download(self.ticker, start=start_date, end=end_date)
17         self.closing_prices_array = self.data["Close"].to_numpy()
18         self.changes = numpy.zeros(self.data["Close"].count())
19
20     def get_changes(self):
21         for closing_price in range(len(self.closing_prices_array) - 1):
22             change = self.closing_prices_array[closing_price + 1] - self.closing_prices_array[
23                 closing_price]
24             self.changes[closing_price + 1] = change
25
26         avg_step_length = numpy.mean(abs(self.changes))
27         self.changes = self.changes / avg_step_length
28         return self.changes
29
30     def get_data(self):
31         return self.data
32
33     def print_data(self):
34         print(self.data)
35
36     def plot_data(self):
37         x_coordinates = np.arange(len(self.data["Close"]))
38
39         y_coordinates = self.get_changes().cumsum()
40         plt.plot(x_coordinates, y_coordinates)
41         plt.xlabel('Boersentage')
42         plt.ylabel('Aktienkurspreis in CHF')
43         plt.show()
44
45     def plot_original_data(self):
46         x_coordinates = np.arange(len(self.data["Close"]))
47         y_coordinates = self.closing_prices_array
48         plt.xlabel('Boersentage')
49         plt.ylabel('Aktienkurspreis in CHF')
50         plt.plot(x_coordinates, y_coordinates)
51
52     def plot_changes(self):
53         x_coordinates = np.arange(len(self.data["Close"]))
54
55         y_coordinates = self.get_changes()
56         plt.plot(x_coordinates, y_coordinates)
57         plt.xlabel('Zeit')
58         plt.ylabel('Kursveraenderung')
59         plt.show()
60
61 # For working with all the Stock data already imported in stocks.py
62 class StockDataList:
63
64     def __init__(self, ticker):
65         self.ticker = ticker
66         self.data = stock_data[stock_data['Ticker'] == ticker]
67         self.closing_prices_array = self.data["Close"].to_numpy()
68         self.changes = numpy.zeros(self.data["Close"].count())
69         self.number_of_zeros = 0
70         self.position_of_zeros = []
71
72     def get_changes(self, zeros=True):
73         for closing_price in range(len(self.closing_prices_array) - 1):
74             change = self.closing_prices_array[closing_price + 1] - self.closing_prices_array[
75                 closing_price]
76             self.changes[closing_price + 1] = change
77
78         if zeros:
79             avg_step_length = numpy.mean(abs(self.changes))
80             if avg_step_length > 0:
81                 self.changes = self.changes / avg_step_length
82         else:

```

```

83         self.changes = self.changes[self.changes != 0]
84         avg_step_length = numpy.mean(abs(self.changes))
85         if avg_step_length > 0:
86             self.changes = self.changes / avg_step_length
87
88     return self.changes
89
90 def find_zero_crossings(self):
91
92     changes_cumsum = self.get_changes().cumsum()
93     last_change = changes_cumsum[1]
94     zero_crossings = []
95     num_zero_crossings = 0
96     last_zero_crossings = None
97     first_zero_crossings = None
98
99     for i in range(1, len(changes_cumsum)):
100         if last_change <= 0 < changes_cumsum[i]:
101             zero_crossings.append(i)
102             num_zero_crossings += 1
103         elif last_change >= 0 > changes_cumsum[i]:
104             zero_crossings.append(i)
105             num_zero_crossings += 1
106         last_change = changes_cumsum[i]
107     if num_zero_crossings > 0:
108         last_zero_crossings = zero_crossings[-1]
109         first_zero_crossings = zero_crossings[0]
110
111     return zero_crossings, num_zero_crossings, first_zero_crossings, last_zero_crossings
112
113 def end_position(self):
114     data = self.get_changes()
115     return data[-1]
116
117 def find_min_max(self):
118
119     changes_cumsum = self.get_changes().cumsum().tolist()
120     maximum = max(changes_cumsum)
121     minimum = min(changes_cumsum)
122     position_maximum = changes_cumsum.index(maximum)
123     position_minimum = changes_cumsum.index(minimum)
124     end_position = changes_cumsum[-1]
125
126     return maximum, minimum, position_maximum, position_minimum, end_position
127
128 def get_data(self):
129     return self.data
130
131 def print_data(self):
132     print(self.data)
133
134 def plot_data(self):
135     x_coordinates = np.arange(len(self.data["Close"]))
136     y_coordinates = self.get_changes().cumsum()
137     plt.xlabel('Boersentage')
138     plt.ylabel('Position')
139     plt.plot(x_coordinates, y_coordinates, alpha=0.6, label="Aktienkurse")
140
141 def plot_original_data(self):
142     x_coordinates = np.arange(len(self.data["Close"]))
143     y_coordinates = self.closing_prices_array
144     plt.xlabel('Boersentage')
145     plt.ylabel('Position')
146     plt.plot(x_coordinates, y_coordinates)
147
148
149 def is_ticker_to_remove(changes, zeros):
150     # Search of stocks with long sequences without any price change
151
152     num_changes = len(changes)
153
154     # Calculates the number of zeros (sections without any price changes) in a time series
155     num_nonzero = len(np.nonzero(changes)[0]) # number of non-zeros
156
157     # Calculate the difference between the number of changes and the number of non-zeros
158     num_zeros = num_changes - num_nonzero
159
160     # Check if the number of zeros is bigger than the set limit
161     if num_zeros > zeros:
162         return True
163     else:
164         return False
165
166

```

```

167 # Check if the stock price is defined in the whole time interval
168 def is_length_too_short(changes, length_tolerance):
169     return len(changes) < (len(StockDataList('ZURN.SW')).get_changes()) - length_tolerance)

```

## A.5 histogram.py

**Listing A.5** Definition von Klassen zur Erstellung von Histogrammen und Balkendiagrammen

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3
4
5 class Histogram:
6
7     def __init__(self, data, save, label, type, xlabel, ylabel='Relative-Haeufigkeit', bins=None,
8                 show=True):
9         self.data = data
10        self.bins = bins
11        self.weights = (np.ones_like(self.data) / len(self.data))
12        self.show = show
13        self.save = save
14        self.type = type
15        self.xlabel = xlabel
16        self.ylabel = ylabel
17        self.label = label
18        if bins is None:
19            self.bins = round((max(self.data) - min(self.data))/10)
20
21    def plot(self):
22        plt.hist(self.data, bins=self.bins, weights=self.weights, label=self.label)
23        plt.xlabel(self.xlabel)
24        plt.ylabel(self.ylabel)
25        plt.title(self.label)
26        if self.save:
27            plt.savefig(f'plots/{self.label}_{self.type}.png')
28        if self.show:
29            plt.show()
30
31 class BarChart:
32     def __init__(self, data, save, standard_errors, label, ylabel):
33         self.data = data
34         self.save = save
35         self.standard_errors = standard_errors
36         self.label = label
37         self.ylabel = ylabel
38         self.bars = ['random-walks', 'modifizierte-random-walks', 'Aktienkurse']
39
40     def plot(self):
41         bar_colors = ['tab:cyan', 'tab:blue', 'tab:green']
42
43         bar_plot = plt.bar(range(len(self.bars)), self.data, yerr=self.standard_errors, width=0.4,
44                           color=bar_colors,
45                           capsize=5, error_kw={'capsize': 3, 'elinewidth': 1, 'capthick': 1},
46                           zorder=2)
47
48         plt.xticks(range(len(self.bars)), self.bars)
49         plt.ylabel(self.ylabel)
50         plt.grid(which='major', axis='y', alpha=0.7)
51         plt.grid(which='minor', axis='y', alpha=0.4)
52         # plt.yticks(np.arange(math.ceil(max(self.data))/10, minor=True)
53
54         for bar in bar_plot:
55             height = bar.get_height()
56             plt.text(bar.get_x() + bar.get_width() / 2, height + 0.1, height, ha='center', va='
57                     bottom', fontsize=10)
58
59         if self.save:
60             plt.savefig(f'plots/{self.label}_mittelwerte')
61
62         plt.gca().set_axisbelow(True)
63         plt.show()
64
65 # This part plots the bar charts, if this file is run
66 if __name__ == '__main__':
67     # Anzahl Nullstellen
68     BarChart([11.754, 9.177, 6.631], True, [0, 0.650, 0.841], 'Mittelwert-der-Anzahl-Nullstellen',
69             'Durchschnittliche-Anzahl-Nullstellen').plot()

```



```

70
71 # Erste Nullstelle
72 BarChart([13.372, 15.372, 64.296], True, [0, 2.899, 5.635], 'Durchschnittliche-Position-der-
      ersten-Nullstelle',
73           'Durchschnittliche-Position-der-ersten-Nullstelle').plot()
74
75 # Letzte Nullstelle
76 BarChart([127.000, 123.858, 111.085], True, [0, 7.709, 6.311], 'Durchschnittliche-Position-der-
      letzten-Nullstelle',
77           'Durchschnittliche-Position-der-letzten-Nullstelle').plot()
78
79 # Maximum
80 BarChart([12.229, 17.788, 19.364], True, [0, 1.127, 0.842], 'Mittelwert-des-Maximums',
81           'Mittelwert-des-Maximums').plot()
82
83 # Minimum
84 BarChart([-12.229, -15.830, -12.856], True, [0, 1.044, 1.105], 'Mittelwert-des-Minimums',
85           'Mittelwert-des-Minimums').plot()
86
87 # Position Maximum
88 BarChart([121.123, 136.440, 104.716], True, [0, 7.559, 7.135], 'Durchschnittliche-Position-des-
      Maximums',
89           'Durchschnittliche-Position-des-Maximums').plot()
90
91 # Position Minimum
92 BarChart([121.123, 120.809, 137.929], True, [0, 7.605, 7.969], 'Durchschnittliche-Position-des-
      Minimums',
93           'Durchschnittliche-Position-des-Minimums').plot()
94
95 bar_colors = ['tab:blue', 'tab:green']
96
97 bar_plot = plt.bar([0, 1], [691.99, 673.86], width=0.5, color=bar_colors, zorder=2)
98
99 plt.ylabel('Durchschnittliche-Flaeche')
100 plt.grid(which='major', axis='y', alpha=0.7)
101 plt.grid(which='minor', axis='y', alpha=0.4)
102 plt.xticks([0, 1], ['random-walks', 'Aktienkurse'])
103 # plt.yticks(np.arange(math.ceil(max(self.data))/10, minor=True)
104
105 for bar in bar_plot:
106     height = bar.get_height()
107     plt.text(bar.get_x() + bar.get_width() / 2, height + 0.05, height, ha='left', va='bottom',
108             fontsize=10)
109
110 if True:
111     plt.savefig('plots/bar-abschnittsuche')
112
113 plt.gca().set_axisbelow(True)
114 plt.show()

```

## A.6 goodness of fit.py

**Listing A.6** Suche und Überprüfung einer passenden Verteilung zur Modifikation der Schritte im Random-Walk-Modell

```

1 from scipy.stats import t, norm
2 import numpy as np
3 from src.stock_data import StockDataImport, StockDataList
4 import matplotlib.pyplot as plt
5 from src.stocks import ticker_list
6 # In order to find a suitable modification of the random walk model, a distribution is to be found
      which can model
7 # the steps of real share prices appropriately.
8
9 all_changes = np.empty(0) # empty numpy array is created
10 for x in ticker_list: # creat a list with all changes from every downloaded stock
11     stock = StockDataList(x)
12     current_changes = stock.get_changes(False)
13     all_changes = np.append(all_changes, current_changes)
14
15 all_changes = all_changes.flatten() # flatten multi-dimensional numpy array
16
17 df, loc, scale = (4, 0, 1) # define degrees of freedom, location and the scale
18
19 x = np.linspace(-8, 8, 1000) # define the dimension of the plot
20 fig, ax = plt.subplots(1, 1)
21 rv = t(df)
22 ax.plot(x, rv.pdf(x), lw=2, label=f't-Verteilung-df={df}', color='r') # plot the graph of the
      Student-t-distribution

```

```

23 ax.hist(all_changes, density=True, bins=40, range=(-8, 8), alpha=0.7) # plot a histogram of the
    changes
24 ax.legend(loc='best', frameon=False)
25 plt.savefig('plots/goodness_of_fit_t.eps', format='eps')
26 plt.show()
27
28 mean, std = norm.fit(all_changes) # get mean and standard deviation of the data
29 x2 = np.linspace(-8, 8, 1000) # define the dimension of the plot
30 fig2, ax = plt.subplots(1, 1)
31 rv = norm(mean, 1.1)
32 ax.plot(x, rv.pdf(x), lw=2, label=f'Normalverteilung', color='g') # plot the graph normal
    distribution
33 ax.hist(all_changes, density=True, bins=40, range=(-8, 8), alpha=0.7) # creat a histogram of the
    changes
34 ax.legend(loc='best', frameon=False)
35 plt.savefig('plots/goodness_of_fit_norm.png', dpi=300)
36 plt.show()

```

## A.7 plot.py

**Listing A.7** Visualisierung des besten Abschnitts der Abschnittsuche im Vergleich mit dem gesuchten Abschnitt

```

1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 class Plot:
5     def __init__(self, function1, function2, label1, label2):
6         self.function1 = function1
7         self.function2 = function2
8         self.label1 = label1
9         self.label2 = label2
10
11     def plot_functions(self):
12         if len(self.function1) >= len(self.function2):
13             length = len(self.function1)
14         else:
15             length = len(self.function2)
16
17         plt.plot(range(length),
18                 np.c_[self.function1, self.function2], label=[self.label1, self.label2])
19         plt.fill_between(range(length), self.function1, self.function2, color='gray', alpha=0.3,
20                          label='Area between walks')
21         plt.xlabel('Schrittabschnitt des langen random walks')
22         area = np.trapz(np.abs(self.function1 - self.function2))
23         plt.legend()
24         plt.show()
25         #plt.savefig('plots/comparison-')

```

## A.8 Liste der Ticker-Symbole

| Ticker-Symbole |          |         |          |          |           |          |          |
|----------------|----------|---------|----------|----------|-----------|----------|----------|
| ABBN.SW        | ABBNE.SW | BCVN.SW | ESUN.SW  | HUBN.SW  | MED.SW    | RSGN.SW  | TECN.SW  |
| ALC.SW         | ACLN.SW  | BEKN.SW | EEII.SW  | HBLN.SW  | MEDX.SW   | RLF.SW   | TEMN.SW  |
| GEBN.SW        | ADXN.SW  | BEAN.SW | EFGN.SW  | IDIA.SW  | MTG.SW    | RIEN.SW  | TKBP.SW  |
| GIVN.SW        | ADEN.SW  | BELL.SW | EFGNE.SW | IMPN.SW  | MTGE.SW   | REHN.SW  | TIBN.SW  |
| HOLN.SW        | ADVN.SW  | BBN.SW  | ELMN.SW  | INA.SW   | METN.SW   | RSGW.SW  | TXGN.SW  |
| BAER.SW        | AEVS.SW  | BKW.SW  | EMMN.SW  | IFCN.SW  | MBTN.SW   | SANN.SW  | UBXN.SW  |
| KNIN.SW        | AIRE.SW  | BOSN.SW | EMSN.SW  | INRN.SW  | MIKN.SW   | SAHN.SW  | UBSGE.SW |
| LISP.SW        | ALLN.SW  | BUCN.SW | EDHN.SW  | ISN.SW   | MOZN.SW   | SCHN.SW  | VZUG.SW  |
| LOGN.SW        | ALPN.SW  | BCHN.SW | RUS.SW   | IREN.SW  | MOZN1.SW  | STRN.SW  | VLRT.SW  |
| LONN.SW        | ALPNE.SW | BRKN.SW | EPIC.SW  | VBSN.SW  | MOBN.SW   | SWTQ.SW  | VATN.SW  |
| NESN.SW        | ALSN.SW  | BVZN.SW | EVE.SW   | JFN.SW   | MOLN.SW   | SENS.SW  | VARN.SW  |
| NOVN.SW        | AFP.SW   | BYS.SW  | FTON.SW  | KARN.SW  | AERO.SW   | SFPN.SW  | VAHN.SW  |
| PGHN.SW        | APGN.SW  | CALN.SW | FHZN.SW  | KNRS.SW  | NBEN.SW   | SFSN.SW  | VETN.SW  |
| CFR.SW         | ARBN.SW  | CPEN.SW | FORN.SW  | KLIN.SW  | NESNE.SW  | SFZN.SW  | VILN.SW  |
| ROG.SW         | ARON.SW  | CMBN.SW | FREN.SW  | KOMN.SW  | NOVNEE.SW | SKAN.SW  | ROL.SW   |
| RO.SW          | ARYN.SW  | CIE.SW  | GALE.SW  | KUD.SW   | NREN.SW   | SNBN.SW  | VONN.SW  |
| SDZ.SW         | ASCN.SW  | CICN.SW | GAM.SW   | KURN.SW  | OFN.SW    | SWON.SW  | VZN.SW   |
| SCHP.SW        | ASWN.SW  | CFT.SW  | GAV.SW   | LLQ.SW   | OBSN.SW   | SWONE.SW | WKBN.SW  |
| SGSN.SW        | AUTN.SW  | CFTE.SW | GEBNE.SW | LAND.SW  | OERL.SW   | SOONE.SW | WARN.SW  |
| SIGN.SW        | AVOL.SW  | CLN.SW  | GF.SW    | LECN.SW  | ONE.SW    | SPEX.SW  | WIHN.SW  |
| SIKA.SW        | BANB.SW  | CLTN.SW | GLKBN.SW | LEHN.SW  | ODHN.SW   | SGKN.SW  | XLS.SW   |
| SOON.SW        | BALN.SW  | COTN.SW | GMI.SW   | LEON.SW  | ORON.SW   | SRAIL.SW | YTME.SW  |
| STMN.SW        | BARN.SW  | CPHN.SW | GRKP.SW  | LISN.SW  | PEAN.SW   | STGN.SW  | YPSN.SW  |
| UHR.SW         | BLKB.SW  | CLXN.SW | GURN.SW  | LISNE.SW | PEDU.SW   | SUN.SW   | ZEHN.SW  |
| SLHN.SW        | BSLN.SW  | DAE.SW  | HBMN.SW  | LISPE.SW | PMN.SW    | UHRN.SW  | ZUBN.SW  |
| SREN.SW        | BSKP.SW  | DCN.SW  | HBMNE.SW | LOGNE.SW | PMNE.SW   | SLHNE.SW | ZUGN.SW  |
| SCMN.SW        | BION.SW  | DKSH.SW | HELN.SW  | LONNE.SW | PLAN.SW   | SPSN.SW  | ZUGE.SW  |
| UBSG.SW        | BIOEE.SW | DOCM.SW | HIAG.SW  | LUKN.SW  | PPGN.SW   | STLN.SW  |          |
| VACN.SW        | BCGE.SW  | DOKA.SW | HLEE.SW  | MCHN.SW  | PEHN.SW   | SQN.SW   |          |
| ZURN.SW        | BCJ.SW   | DESN.SW | HOCN.SW  | MOVE.SW  | PSPN.SW   | THAG.SW  |          |

## Anhang B

# Digitale Dateiablage

Computergestützte Analysen von Daten und dessen Visualisierung machen einen Grossteil dieser Arbeit aus. Neben dem Auszug des Python-Programmcodes in Anhang A, sollen diese digitalen Inhalte in ihrer natürlichen Form zur Verfügung gestellt werden. Denn diese digitalen Inhalte können nie mit Detailgetreue oder Funktionalität in gedruckter Form wiedergegeben werden. So soll interessierten Lesern einerseits der funktionsfähige Programmcode, aber auch Animationen und Diagramme in höherer Auflösung nicht vorenthalten bleiben. Aus diesem Grund werden über den folgenden Link oder den QR-Code erwähnte Inhalte auf Google Drive zur Verfügung gestellt.

**Link:**

<https://drive.google.com/drive/folders/1uWd9f8i15A2zT25NTvXAt0-jsSfCDbtw?usp=sharing>

**QR-Code:**



# Deklaration

„Ich erkläre hiermit,

- dass ich die vorliegende Arbeit selbständig und nur unter Benutzung der angegebenen Quellen verfasst habe,
- dass ich auf eine eventuelle Mithilfe Dritter in der Arbeit ausdrücklich hinweise,
- dass ich eine allfällige Nutzung von Künstlicher Intelligenz (z.B. ChatGPT) in der Reflexion der Arbeit ausgewiesen und diskutiert habe,

**Ort:** Schenkon

**Datum:** 31. März 2024

**Unterschrift:** *C. Willmann*