The Ergodicity Problem in Economics

The Advent of Ergodicity Economics

Doctoral Thesis

submitted in partial fulfillment of the
requirements for the degree of

Doctor rerum politicarum (Dr. rer. pol.)

to the Faculty of Business Management and Economics
at TU Dresden

by

Mark Kirstein, Dipl. Wirtsch.-Inf.,
born 26th August 1985 in Leipzig, Germany

Submission: 15th March 2019
Defence: 2nd Juli 2019

The Ergodicity Problem in Economics

The Advent of Ergodicity Economics

This version was compiled on 20th May 2020.

Mark Kirstein

✉ m.kirstein@lml.org.uk
@nonergodicMark

This version of my PhD thesis has been produced for the open access open peer review platform researchers.one. I am interested in reviewer feedback. Please feel free to upload your reviews, (dis)agreements, typos, errors, *etc.* directly to researchers.one or email me. Compared to the original submission this version contains only minor corrections with regard to *e.g.* typos, misplaced citations and some resolved ordering issues in the bibliography.

This dissertation was typeset in LaTeX.

Für meine Familie,

insbesondere meinem Großvater

Bruno H. Gahler (1930-2018) gewidmet.

Acknowledgements

> In a lot of mature fields in academia what people tend to do is they stop working on the big foundational problems and they work on this very detailed problems, where it is very hard to make progress. Heading to the foundational center of difficulty is actually a lot easier.
>
> STEPHEN WOLFRAM

I am not sure whether I agree with the above quotation. Over the period of working on this thesis I came to realise, the closer one moves towards the foundational centre of a discipline the stronger the resistance becomes especially from the direction of the sociology of science. Therefore I gratefully acknowledge the freedom and encouragement of the research environment my supervisor Prof. MARCO LEHMANN-WAFFENSCHMIDT provided and his insistence in demanding a thesis which is intelligible for the economics community. During many internal workshops and discussions my colleagues at the Chair of Managerial Economics gave helpful advice on my premature thoughts.

The maturation of my thoughts and the completion of my thesis would have been orders of magnitude harder if I had not have the privilege of being surrounded by people who constantly set me an example of what real science is all about. This was done first and foremost by my mathematician father, BERND KIRSTEIN, to whom I am deeply thankful. Without his engagement in guiding me through various mathematical texts and concepts – such as KIRSTEIN (2007) – I would have not been able to digest the deep subtleties of ergodic theory. Equally important are my mother HANNELORE KIRSTEIN, my brother NIELS, my grandparents, SOFIE KIRSTEIN, IRENE and BRUNO GAHLER, and especially the unconditional support by my wife BETTINA KIRSTEIN, who realised when it was necessary to cheer me up and always did – I will always love you. Altogether they created a domestic environment most supportive for my scientific endeavour.

I received further most important support from ALEXANDER ADAMOU and OLE PETERS from the London Mathematical Laboratory (LML). Since 2013 we established a close connection and I am honored to collaborate with such first rate scientists. They both serve as role models for me of the intrinsically motivated scientist and I enjoy their stimulating views on academia, science, and life. What they are creating with LML is an outstanding environment for research freed from any distracting burdens, a place of deep thoughts and a very welcome alternative to institutionalised academic science.

During the course of the doctoral studies I benefited in various ways and on various occasions from the support of the Young Scholars Initiative (YSI) of the Institute for New Economic Thinking (INET). It began at a conference at Leipzig University in October 2012 where Perry Mehrling encouraged me to create what became the first YSI working group. The complexity economics working group (CEWG) was co-founded and jointly ran by Franziska Schütze, Johannes Tiemer and me for several years and I benefited immensely from the reading groups and events we organised in the early stages of my doctorate. It was a perfect setting for us young scholars to explore and subsequently implement our own understanding of how complexity ideas can be fertile for economics.

Even for a PhD student in economics money is not important – as long as he has enough of it. Over the time of my doctorate I benefited from many conference grants by the Economics Department at TU Dresden. My attendance of the Complexity Science Summer School 2017 at the Santa Fe Institute (SFI) was jointly supported by a travel grant from the Deutsche Bundesbank and a travel award from the Graduate Academy at TU Dresden, financed by the excellence initiative of the German Federal Ministry of Education and Research (BMBF). Furthermore, a four months fellowship for the completion phase of the doctorate from the Graduate Academy is gratefully acknowledged. I thank Prof. Frank Schirmer and PD Dr. Gabriele Fassauer for providing and sharing office space during the crucial last part of the writing phase and thank all my colleagues at the Chair of Organisation for adopting me so pleasantly.

I owe special thanks to Alex Adamou, Bettina Kirstein, Bernd Kirstein, Thomas Krause, Ole Peters, André Roncaglia de Carvalho and in particular to Sven Resnjanskij, who have carefully read and discussed various parts of the manuscript with me. You all contributed much to eliciting in greater clarity what I sometimes vaguely put down in ink. I draw much inspiration from listening to the fantastic EconTalk podcast by Russell Roberts. This dissertation was typeset using LaTeX and the numerical simulations were done using Python, I want to thank the developers and the communities as well. I received first class IT and administrative support form Marco Barthel, Kameliya Bogdanova, Birgit Hergert and Chloe Scragg.

I dedicate the thesis to my grandfather. Sadly it was not granted him to witness the completion of my doctoral studies. This thesis is devoted to his remembrance and in recognition of what he made possible for me, which was everything a grandson could wish for.

Preface – How I went down the rabbit hole

> If I had more time, I would have written a shorter thesis.
>
> paraphrasing Mark Twain

During my studies, which started in 2005, I experienced the onset of the financial crises and the subsequent sovereign debt crisis. In the aftermath of what turned out to become a series of crises, I grew up in an intellectual climate where economics was constantly challenged from many vantage points. Among the suggested candidate theoretical frameworks to overcome some of the obvious shortcomings of standard economic theory was the conception of the economy as a complex evolving adaptive system. Having been in contact with cybernetics from my earliest childhood days, the transdisciplinary complexity view was most convincing and somehow deeply appealing to me. Complexity has been promoted since the mid 1980s by the SFI, an independent institute devoted to interdisciplinary research on complex systems located in New Mexico, USA. A dream came true when I attended the SFI summer school in 2017. Browsing over the final pages of the thesis, it is delightedly surprising how strongly the SFI influence shines through.

My diploma thesis (Kirstein 2012) was inspired by Benoît B. Mandelbrot's work on fractals. Considering that it dealt with the applicability of generalised limit theorems to power law distributed random variables and stable distributions in finance and realising that the St. Petersburg lottery is a fractal with a power law payout distribution, my route may seem predestined in retrospect. However, it never felt that way at every present moment in time. Nevertheless, following my interests prepared me insofar as to be highly receptive for interdisciplinary injections even if *e.g.* the mathematical level surpasses my current skill level.

I have vivid memories of the immediate fascination when I made first contact with what turned out to become the topic of my dissertation. It was in December 2012 and I read the short piece by Peters (2009) which appeared in the Santa Fe Institute Bulletin. I was struck by the article's unique combination of interesting topics like gambling, the nature of time, parallel universes and concepts I already encountered in lectures before like utility theory, risk management, expected return, portfolio theory. In addition illustrious personalities appear like Diogenes, the Bernoullis, Ludwig Boltzmann or John L. Kelly. The article employs the famous St. Petersburg lottery as a well-functioning narrative and is obviously written by an autodidact, who – if the publicly available information were reliable – has enjoyed apparently no formal economics education at all. This resonated strongly with me, as I consider myself not at all as an economist, partly because I majored in an information systems programme and took only voluntary minor courses in economics & finance. My interests do not stop at arbitrary university faculty borders, too. From this time on I was dragged down the rabbit hole by a tractor beam, trying to understand what the ergodic hypothesis and ergodic theory is all about and how it all relates to the treatment of randomness, risk and decision-making.

This journey was fueled by meeting Alexander Adamou at a conference in Edinburgh in April 2013, who had joined Ole Peters in the meantime as a fellow of the newly founded LML, an institute where researchers are commissioned to follow their own curiosity. After my talk in Edinburgh, two colleagues from TU Dresden, Alex and I dined together and he invited me to give a seminar at LML in June 2013. This visit turned out to be what I have been looking for since I entered academia, an environment purely devoted to reason and knowledge without any bureaucratic busywork. My seminar lasted for five hours excluding the lunch break and has been attended by Alex, Ole and Nicholas Moloney. I was fascinated by the intense and collaborative research environment, and that three bright scientists took the time to listen to my disorganised thoughts on *Non-Ergodicity – Its Role in Economics & Finance*. I do not know what they saw in me back then, but eventually this event led to the thesis in your hand on *The Ergodicity Problem in Economics*.

Dresden, Mark Kirstein
March 2019

Contents

List of Figures

List of Tables

Glossary

A set from the σ-algebra 56, 57

C_{fair} fair price of a lottery ticket 88, 105, 106, 109, 132, 143

C fee, price or cost of a lottery ticket xvi, 88, 90, 96, 115, 117, 120, 121, 122, 123, 127, 128, 129, 139, 141, 142, 143, 147, 150, 151, 157, 161, 164

F cumulative probability distribution a random variable 141

N ensemble size, number of realisations of a random variable xiv, 27, 28, 40, 67, 68, 87, 89, 90, 101, 109, 141, 142, 143, 144, 149, 150, 154, 161

R growth factor, a random variable, whereby the wealth changes in one round of a gamble are realisations of 261

S phase space xv, xvi, 31, 34, 54, 55, 56, 57, 58, 61, 62, 63, 64, 97, 98, 140

T number of sequential iterations of a gamble, so that $T\delta t$ is the total duration of a repeated gamble, also finite length of a measurement xiii, xiv, 27, 28, 32, 34, 40, 62, 63, 64, 68, 141, 148, 152, 153, 154, 155, 156, 159, 161, 263, 264

W wealth of a gambler is a realisation of the random variable W; one of our main observables 115, 118, 261

X random variable, the payout x from a gamble is the realisation of the random variable X and if the ticket fee is subtracted yields the change in the gambler's wealth or, in discrete time, the wealth increment; one of our main observables xiii, xvii, 86, 115, 136, 137, 140, 141, 142, 143, 256

Z generic random variable xiv, 61, 62, 63, 64, 67, 68, 69, 141

Z accumulated payouts 141, 142

Δ difference operator, for example Δw is a difference of two values of w, for instance at two different times xiv, 115, 116, 117, 118, 119, 120, 121, 122, 125, 126, 127, 128, 129, 131, 136, 137, 138, 139, 140, 148, 152, 153, 158, 159, 170, 256, 261, 262, 264

Φ generator transformation in discrete time xiv, xv, 55, 56, 57, 59, 61, 62, 63, 64

δt a time interval corresponding to the duration of one round of a gamble or, mathematically, the period over which a single realisation of the constituent random variable of a discrete-time stochastic process is generated. xiii, 27, 115, 116, 117, 118, 120, 121, 125, 126, 127, 128, 139, 146, 147, 148, 150, 152, 153, 154, 157, 158, 159, 261, 262, 263, 269

ℓ leverage or fraction of wealth invested in a gamble xiv, 156, 260, 261, 262, 263, 264, 265, 270, 271, 272, 275

Δt a general time interval 148, 152, 153, 264

Δu^{+} random change in utility after one round of the lottery is played, but irrespective of the price of the lottery ticket 121, 122, 140, 256

Δu^{-} sure change in utility after the purchase of the lottery ticket 120, 121, 122, 140, 256

Δw change in wealth 115, 116, 118, 119, 125, 128, 131, 136, 137, 138, 158, 159, 170, 256, 261, 262

$g_{\langle\rangle}$ exponential growth rate of the expectation value 149, 150, 151, 160, 262, 263

ℓ_{opt} optimal leverage 265, 275

$n_N^{\max}$ maximal number of consecutive coin tosses that occurred in an ensemble of size N 149, 150

$n_T^{\max}$ maximal number of consecutive coin tosses that occurred in a round of length T 154, 263

u_{B} candidate logarithmic utility function due to D. Bernoulli 126, 127, 129, 133, 158

u_{C} candidate utility function due to Gabriel Cramer 131, 132, 133, 139

$\left(\Phi^k(s)\right)_{k\in\mathbb{N}_0}$ single trajectory in discrete time starting at s or time series 59

$\left(f\circ\Phi^k\right)_{k\in\mathbb{N}_0}$ all possibly observable trajectories in discrete time 59

$\left(f\circ\Phi^k(s)\right)_{k\in\mathbb{N}_0}$ single observed trajectory in discrete time starting at s or time series 59

$(f\circ\varphi_t(s))_{t\in\mathbb{T}}$ observed trajectory 58

$(\varphi_t\,(s))_{t\in\mathbb{T}}$ single trajectory starting from point s, sequence of numbers 55, 58

$(\varphi_t)_{t\in\mathbb{T}}$ all possible trajectories, ensemble of all trajectories 55, 56, 57

$(Z_k)_{k\in\mathbb{N}_0}$ stochastic process in discrete time or a family of random variables indexed by time $k\in\mathbb{T}=\mathbb{N}_0$ 62

$(Z_t)_{t\in\mathbb{T}}$ stochastic process or a family of random variables indexed by time 61, 67

$\left(\Phi^k\right)_{k\in\mathbb{N}_0}$ all possible trajectories in discrete time 59

$(S, \mathcal{A}, \mathcal{P})$ probability space 54, 61, 140

$(S, \mathcal{A}, \mathcal{P}, \Phi)$ measurable dynamical system with a probability measure in discrete time 57

$(S, \mathcal{A}, \mu)$ measure space 54, 56, 57

$(S, \mathcal{A}, \mu, \Phi)$ measurable dynamical system in discrete time 57, 63, 64

$\mathbb{N}_0$ set of non-negative whole numbers, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ xiv, xv, 54, 55, 56, 57, 59, 61, 62, 63, 67, 261

$\mathbb{N}$ set of positive integers, $\mathbb{N} = \mathbb{Z}^+ = \{1, 2, 3, \ldots\}$ xvi, 8, 89, 97, 106, 147, 148

$\mathbb{R}^+$ set of non-negative real numbers $\mathbb{R}^+ = [0, +\infty)$ xv, 54, 55

$\mathbb{R}$ set of real numbers, $\mathbb{R} = (-\infty, +\infty)$ 31, 32, 34, 38, 55, 58, 61, 62, 110, 118, 125, 126, 257, 260

$\mathbb{T}$ domain of the time parameter $t \in \mathbb{T}$, for continuous time $\mathbb{T} = \mathbb{R}^+$, for discrete time $\mathbb{T} \in \mathbb{N}_0$ xiv, xv, 54, 55, 56, 57, 58, 59, 61, 62, 67

$\mathbb{Z}$ set of integers, $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ xv, 55

$\mathcal{A}$ σ-algebra xv, 54, 55, 56, 57, 58, 61, 63, 64, 140

$\mathcal{B}$ Borel σ-algebra 58

$\mathcal{I}$ identity map, $\mathcal{I} : S \to S$ via $s \mapsto s$ 55

$\mathcal{L}$ $\mathcal{L}^p$ is the Lebesgue function space of p-integrable measurable functions 63, 64

$\mathcal{N}$ normal distribution 8, 9, 234

$\mathcal{P}$ probability measure, a distribution function of a random variable X is denoted by $\mathcal{P}_X$ xv, 14, 54, 57, 61, 140

$\mathfrak{V}$ value of a gamble in prospect theory 136, 137, 256

$\mathfrak{v}$ value function applied to the payout of a gamble in prospect theory 136, 137

μ measure xv, 33, 34, 54, 56, 57, 62, 63, 64

E expectation operator or expectation value of a random variable $\mathrm{E}[\cdot] = \langle\cdot\rangle$ 63, 64, 87, 88, 89, 90, 94, 95, 97, 105, 106, 107, 109, 115, 116, 117, 118, 127, 138, 139, 170

Pr probability of the realisation of event $X = x$ is denoted by $\Pr\{X = i\} = p(i)$ 64, 86, 141, 142 return return of an investment is the payout minus the costs, return $= x - C$ xvi, 88, 89, 95

π probability weighting function in prospect theory 135, 136, 137

τ specific round in a gamble, $\tau \in \mathbb{N}$ 147, 148, 152, 153, 233

φ transformation or mapping xiv, 55, 56, 57, 58

f generic function, also $f(\cdot)$ is an observable, also outcome of a physical measurement xiv, 32, 33, 34, 58, 59, 62, 63, 64

g growth rate xiv, 148, 149, 150, 151, 154, 157, 158, 159, 160, 256, 262, 263, 264, 265, 270, 271, 272, 275

j label of a particular outcome 264

k_n frequency a given number of consecutive coin tosses occurred in the sequence of T rounds 149, 150, 154

n a natural number, $n \in \mathbb{N}$, often the number a die is thrown or a coin is tossed, n_j is the number of times outcome j is observed in an ensemble xiv, xvi, 69, 85, 86, 87, 88, 89, 90, 91, 93, 94, 96, 97, 104, 105, 106, 107, 109, 110, 115, 117, 123, 127, 128, 129, 132, 136, 137, 138, 139, 140, 141, 149, 150, 151, 153, 154, 157, 160, 161, 233, 263, 264

n ensemble element, label for particular realisation of a random variable 67, 68, 95, 149, 150, 151, 261, 263

p probability, $p(i)$ is the probability to observe event i xv, 86, 87, 88, 89, 93, 101, 105, 106, 107, 109, 127, 128, 132, 135, 136, 137, 138, 139, 141, 147, 150, 154, 160, 256, 260, 261, 262, 263, 264, 265

r random growth factor whereby wealth changes in one round of a gamble 147, 148, 149, 150, 151, 152, 153, 154, 158, 159, 160, 260, 261, 262, 263, 264, 268, 269, 270, 271

s point in phase space S xiv, xv, 55, 58, 59, 61, 62, 63, 64

t time xiii, xiv, xv, 15, 27, 32, 55, 56, 57, 58, 61, 67, 68, 69, 115, 116, 117, 118, 120, 121, 122, 123, 125, 126, 127, 128, 129, 131, 139, 146, 147, 148, 150, 151, 152, 153, 154, 157, 158, 159, 161, 234, 256, 260, 261, 262, 263, 264, 268, 269

u utility function, u : money $\rightarrow$ usefulness xiv, 85, 116, 117, 118, 119, 120, 121, 122, 123, 125, 126, 127, 129, 131, 132, 133, 137, 138, 139, 140, 158, 256

v stationary mapping function, ergodicity transformation, so that $v(\cdot)$ has stationary increments 158, 159, 160

w a realisation of the random variable W which denotes the wealth of a gambler; one of our main observables xiv, 115, 116, 117, 118, 119, 120, 121, 122, 123, 125, 126, 127, 128, 129, 131, 133, 136, 137, 138, 139, 146, 147, 148, 150, 151, 152, 153, 157, 158, 159, 160, 161, 170, 256, 260, 261, 262, 264, 268, 269

x the payout from a gamble is the realisation of the random variable X and if the ticket fee is subtracted yields the change in the gambler's wealth or, in discrete time, the wealth increment; one of our main observables xiii, xvi, 85, 86, 87, 88, 89, 90, 93, 94, 95, 97, 101, 105, 106, 107, 109, 110, 115, 117, 119, 121, 122, 127, 128, 129, 131, 132, 136, 138, 139, 140, 141, 147, 148, 150, 151

z generic realisation or value of a random variable X 67, 68, 69, 141

Acronyms

ABM agent-based model 251

ABS agent-based simulation 196

ACF auto-correlation function 201

AEP asymptotic equipartition 77

a.s. almost surely 177

a.e. almost everywhere 62, 63

ARA absolute risk aversion 133

ARFIMA auto-regressive fractionally integrated moving average 201, 202

BG Boltzmann-Gibbs 202, 203, 204, 205, 206

BM Brownian motion 201

CAPM capital asset pricing model 180

CDF cumulative probability distribution function 141

CEWG complexity economics working group ii

CLT central limit theorem 141, 178, 195, 204, 205

CoVAR cointegrated vector autoregressive 198

CPI consumer price index 197, 230

CPT cumulative prospect theory 112, 168, 252, 253, 258

CRRA constant relative risk aversion 134

CTRW continuous-time random walk 201, 208

DARA decreasing absolute risk aversion 134

DAX German stock index 230

DGP data generating process 7, 183, 187, 193

DN deductive-nomological 225

DRRA decreasing relative risk aversion 134

DSGE dynamic stochastic general equilibrium 198

EMH efficient markets hypothesis 9

ESS evolutionarily stable strategy 179, 214

EUT expected utility theory 18, 21, 22, 100, 113, 114, 116, 117, 119, 121, 123, 124, 125, 126, 131, 134, 135, 137, 139, 140, 145, 159, 168, 169, 226, 227, 228, 246, 250, 252, 256, 267

fBm fractional Brownian motion 201, 202

fGn fractional Gaussian noise 201, 202

FPUT Fermi-Pasta-Ulam-Tsingou 30, 72, 73

GARCH generalised autoregressive conditional heteroskedasticity 195, 202

GBM geometric Brownian motion 160, 234

GDP gross domestic product 7, 31, 58, 65, 174, 223, 230

GLM general linear model 195

GUT grand unified theory 99

IARA increasing absolute risk aversion 134

IIASA International Institute for Applied Systems Analysis 182

iid identically and independently distributed 66, 77, 141, 178, 201, 204, 205

IKE Imperfect Knowledge Economics 182, 198, 235

INET Institute for New Economic Thinking ii

IRRA increasing relative risk aversion 134

KAM Kolmogorov-Arnol'd-Moser 74, 75

LHS left-hand side 32, 33, 50, 147, 149, 157, 158

LLN law of large numbers 28, 77, 78, 104, 142, 154, 178, 193, 196, 264, 267, 270

LML London Mathematical Laboratory i, ii, iv, 20

LRD long-range dependence 200, 201

MEP maximum entropy principle 254, 257, 281

MSD mean squared displacement 207

MWI many world interpretation of quantum mechanics 212, 213

PDF probability distribution function 38, 73, 108, 118, 207

PKE Post Keynesian Economics 184, 188, 189

PMF probability mass function 136

PNAS Proceedings of the National Academy of Sciences of the USA 219, 220

$\mathcal{P}$-measure probability measure 172

PT prospect theory 134, 135, 136, 137, 168

RBC real business cycle 198

REH rational expectations hypothesis 176, 198, 237, 250

RGBM reallocating geometric Brownian motion 234

RHS right-hand side 38, 50, 149, 151, 158

SDE stochastic differential equation 159

SEU subjective expected utility 255

SFI Santa Fe Institute ii, iii, 20

SLLN strong law of large numbers 66

SSB spontaneous symmetry breaking 234

SSK sociology of scientific knowledge 242

SVAR structural vector autoregressive 198

TOE theory of everything 99

VAR vector autoregressive 198

VSR variation, selection, and retention 179

WEB weak ergodicity breaking 201

WLLN weak law of large numbers 140, 141

YSI Young Scholars Initiative ii

1 Introduction

At the 2016 annual meeting of the Verein für Socialpolitik[1] its former president Michael Burda chaired a panel on the topic *Teaching economics: Is there a need for reform?* In his opening statement Burda summarised the accusations put forth by critics of the economic discipline and enumerated several problems economics and economics teachers are occasionally confronted with in public either from the media, from students or other stakeholders. The exact wording of his opening reads as follows:

> "A lot of students and student teachers had begun to criticize the way we do our thing. Part of what we do is research, part of what we do is teaching. People said:
>
> - we use too much maths,
> - we have too much rationality in our models,
> - we have too much conformity in our agents, these representative agents we have,
> - we have too much of a streamlined view of what they do,
> - we have *too much ergodicity*, too much system independence or independence of initial conditions,
> - and then we are also told we have too little psychology, too little human element,
> - we know a lot about people and human existence from other fields that we don't integrate,
> - we don't use enough real-world examples,
> - we don't use heterogeneity of types and,
> - we have little respect for our own past."[2]

As can be seen from this quote, the demands and requested advancements of economic theory are manifold and taken together of an enormous complexity. But they have not prevented the never dwindling source of caustic remarks on economics, Ariel Rubinstein, to get something out of this situation: '[It is a good time to become an economist.] Economics has never

[1] The association of German speaking economists from Switzerland, Austria and Germany.
[2] Our emphasis. See https://livestream.com/accounts/2759289/events/5988143/videos/137065860 for a video recording of the panel discussion with quickly improving audio quality.

been in a worse state. This is unfortunate for humanity but fortunate for [PhD students].'[3] Thus, BURDA's list offers a rich reservoir of problem sets to work on. Quite astonishingly, the main topic of this thesis *non-ergodicity* has made it to point five in the list, too, and into the recent canon of the economics critique and even more surprising into the opening statement of such an event. It is therefore notable, that BURDA includes ergodicity in his list. Unfortunately, it is also representative for the whole discourse in economics, that the term *ergodicity* or the respective topic is not mentioned again in this panel session or for the rest of the conference.

Although a prominent and influential economist has even pedestalled the ergodic axiom:

> "Finally, there was an even more interesting [. . .] assumption implicit and explicit in the classical mind. It was a belief in unique long-run equilibrium independent of initial conditions. I shall call it the '*ergodic hypothesis*' by analogy to the use of this term in statistical mechanics. [. . .]
> Now, Paul Samuelson, aged 20 [. . .] as an equilibrium theorist he naturally tended to think of models in which things settle down to a unique position independently of initial conditions. Technically speaking, we theorists hoped not to introduce hysteresis phenomena into our model, as the Bible does when it says 'We pass this way only once' and, in so saying, takes the subject *out of the realm of science into the realm of genuine history*."[4]

Since the time of this quote, ergodicity has become hidden as an implicit assumption in mathematical economics which the community is largely unaware of. Thus we have identified not only a gap in the literature but a blind spot of the whole discipline.[5] Even more so the above quote documents, that ergodicity is portrayed as unquestionable if one seeks to do economics as a 'proper' science. This is in stark contrast to the advancements in other disciplines *e.g.* in non-equilibrium physics of the 20$^{\text{th}}$ century, which notably SAMUELSON treasured so highly. This thesis is devoted to the fundamental importance and ramifications of the non-ergodic case for economics. From an increasing number of contributions from the critics of economics end of the spectrum we discern that the ergodic axiom and the ignorance of non-ergodicity are key issues to understand the process of the formalisation of economics since the second world war. The following statement of a representative of the critical camp and prominent author contradicts the quote above and is of course music to our ears: 'The Ergodic Property, [. . .] the Most Important Property to Understand in Probability, in Life, in Anything'.[6] Furthermore, it marks the opposite extreme end of the spectrum regarding the role which is attributed to ergodicity.

[3] RUBINSTEIN 2013.
[4] See SAMUELSON (1968, pp. 11, our emphasis). We comment on SAMUELSON's role and the exact context of this quote in more detail in chapter 6.
[5] BYERS 2011.
[6] TALEB 2014.

How we can bring the two opposing mindsets harmoniously together? A glance back in the history of science teaches us that such stages in the evolution of a scientific discipline mark crucial turning points and often generate a sweeping momentum for new ideas, which get born into or should we say rise out of such circumstances. What has seemed as irreconcilable anomalies dissolve on closer inspection once they had been formulated in the right (often mathematical) language. Several recent promising results have led to the advent of *Ergodicity Economics*, which provides such a needed terminology. We approach this tension sketched above by identifying and demonstrating the content of the ergodicity problem in economics and in pointing into the direction where further progress awaits.

1.1 The Ergodicity Problem in Economics

Khinchin (1949, p. 47) showed that the thermodynamic quantities of interest in statistical mechanics are – from a mathematical point of view – sum functions which are defined via averages. Thus, he identified the ergodicity problem as a fundamental challenge for the mathematical foundations of statistical mechanics and stated it as follows.

Ergodic[ity] Problem in Statistical Mechanics

The problem of a theoretical justification of the replacement of time averages by ensemble averages, is [. . .] the ergodic[ity] problem.

Initially, the terminology of time and ensemble averages may be unfamiliar to an economics readership. However, both quantities are of utmost importance to economics and are easily understood once we relate to different labels that are familiar in economics. Peters (2011c) showed that the ensemble and time averages used in statistical mechanics correspond to two different conceptualisations of risk. In addressing both concepts of risk one by one we are naturally led to the ergodicity problem in economics.

Ensemble Perspective – Embedding Randomness within the Ensemble

The term ‘ensemble average’ may sound unfamiliar to an economics audience, but only until we reveal the following secret. The most famous guise of an ensemble average is the expectation value of a random variable, whose value is understood as the outcome of a random experiment. The expectation value conceptualises the risk or uncertainty associated with a random experiment as an average over the set of all these possible outcomes. This set is denoted as the ensemble. Hence, by averaging over all possible outcomes one computes what is called an ensemble average. The ensemble average has the mathematical structure of a probability weighted arithmetic mean. In the reminder, we will refer to this conceptualisation of risk as an embedding of randomness within the ensemble. The expectation value is the core quantity that underlies the dominant concept to quantify risk and as such enters economics at many places and in particular in decision theory. Thus the embedding of risk in the ensemble manifests itself in the overwhelming majority of the preexisting theories of decision making under uncertainty. More precisely, ensemble-averaged quantities (or variations of it) enter decision criteria in normative decision theories and through them the embedding of risk in the ensemble becomes highly performative. However, the ensemble which is behind the expectation operator is in many real-world applications only a virtual timeless construct that is inaccessible to a single decision maker. It consists of counterfactual states of the world and has no physical meaning for him. In the remainder, we will clarify the virtuality of the ensemble in more detail.

Time Perspective – Embedding Randomness within Time

A different conceptualisation of risk rejects the idea of the virtual ensemble and instead embeds the uncertainty of an outcome within time, *i.e.* along one factual path of events through historical time, how they are experienced for example by a real living human decision maker. Thus averaging along such a factual path is denoted as a time average. The time average may sound more familiar to economists if we simply label it as the average of some time series of an economic quantity over a period of time. By averaging over one such path, we compute a quantity that is physically meaningful to the decision maker in the single world he experiences. A crucial and so far largely overlooked determinant for economics is the dynamics of the environment in which individuals make decisions and under which the economic observables evolve. Whereas ensemble-average quantities remain unaffected from the dynamics, time averages of observables depend crucially on the dynamics. To give a glimpse of the flavour of our analysis, note that the time-average of the growth factors of wealth appears as a geometric mean of the period-wise growth factors under multiplicative dynamics, which can be transformed into an arithmetic mean of the logarithm of the period-wise growth factors of wealth. A large part of the thesis is devoted to the nature of such transformations and their origin. Not least because much confusion arose in the history of economics between the relation of such transformations (like the logarithm) and the dynamics of the situation, simply because in the ergodic case time averages may look like ensemble averages, although both are fundamentally different objects. Or after the application of an ergodicity transformation the time average resembles the known ensemble average, which led to the false impression to have always operated with a *correct* quantity. This led to the ergodic fallacy which we later explain. We also discuss the role of dynamics and observables in detail in the remainder.

From the Ergodicity Problem to Ergodicity Economics

When both conceptualisations of risk, *i.e.* the time perspective and the ensemble perspective, for the same observable lead to identical results this is denoted as the ergodic case. The ergodic case requires strong mathematical assumptions, which makes it the absolute exception. Consequently, it is of utmost importance to examine the ergodicity of those random observables which are involved in the assessment of risk. In the majority of cases, the economist will encounter the situation where the two conceptualisations of risk lead to different results, *i.e.* an investment seems favourable in one but not in the other conceptualisation. This is denoted as the non-ergodic case, which is the generic case. This leads us to paraphrase KHINCHIN's formulation in the context of economics as the focal problem of this thesis.

Ergodicity Problem in Economics

The problem of a theoretical justification of the replacement of time averages by expectation values is the ergodicity problem in economics and a key issue in ergodicity economics.

The general problem – to assess the relation between the two conceptualisations of risk – is referred to as the ergodicity problem in economics throughout the thesis. Kick-started by the seminal paper PETERS (2011c) and in the meantime supported by growing lecture notes[7] a whole new way to think about risk emerged and led to the new scientific theory of *Ergodicity Economics*, which tackles the ergodicity problem in economics head-on and proactively studies the – as it turns out very likely – possibility of broken ergodicity in many economic puzzles. The ergodicity problem poses deep and fundamental challenges for mathematical economics and financial mathematics of which a first selection could be resolved already with the help of the newly developed tools. Consequently, this thesis understands itself as a contribution to the *Ergodicity Economics* research programme and as a correction to what has been aptly coined the SAMUELSONIAN vice by FOLEY (2010):

> "if one uses sophisticated mathematical methods to analyze a complex adaptive system far from equilibrium under the prior assumption that it is an equilibrium system, the sophistication of the mathematics is not going to correct the fundamental conceptual error. Inherent in the application of mathematical methods to economics is the risk of what I will venture to call the Samuelsonian vice, which is my name for the temptation to change the formulation of the abstract problem to fit the mathematical tools available rather than to seek mathematical tools that are appropriate to the actual problem at hand."

[7] PETERS and ADAMOU 2018b.

The correction of the Samuelsonian vice in economics leads over to the dominant set of instruments used in economics to analyse stochasticity and random variables.

Soul-Searching in Economics

The critiques of standard economics models publicised at prominent places such as the Financial Times by proficient people often focus on the eschewal of a proper treatment of stochastic effects, which manifests in the Samuelsonian vice of simply assuming away non-linearity, non-stationarity or non-deterministic effects. For instance the following list of deficiencies by Buiter (2009):

> "[...] any potentially policy-relevant model would be highly non-linear, and that the interaction of these non-linearities and uncertainty makes for deep conceptual and technical problems.
> [...] stripping the model of its non-linearities and by achieving the transsubstantiation of complex convolutions of random variables and non-linear mappings into well-behaved additive stochastic disturbances
> Threshold effects, critical mass, tipping points, non-linear accelerators – they are all out of the window. Those of us who worry about endogenous uncertainty arising from the interactions of boundedly rational market participants cannot but scratch our heads at the insistence of the mainline models that all uncertainty is exogenous and additive.
> Technically, the non-linear stochastic dynamic models were linearised (often log-linearised) at a deterministic (non-stochastic) steady state. The analysis was further restricted by only considering forms of randomness that would become trivially small in the neig[h]bourhood of the deterministic steady state. Linear models with additive random shocks we can handle – almost!
> Much of this work is numerical – analytical results of a policy-relevant nature are few and far between – but at least it attempts to address the problems as they are, rather than as we would like them"[8]

Further George (2011, p. 634) identifies a 'degree of soul-searching' within the economics discipline as one consequence of the recent financial crisis. He states the

> "increasing irrelevance of much mainstream economics provoked some economists to re-examine their discipline. Linear or linearized models with well-behaved additive stochastic disturbances, based on 'microeconomic foundations' are no

[8] Buiter 2009.

longer anywhere near adequate. Non-linearity, complexity and randomness cannot be avoided."

This thesis shows how the *Ergodicity Economics* framework allows to coherently understand and study the non-linearity, non-stationarity, complexity and non-additivity of randomness in the relationship between economic quantities. To see what is meant by the linear(ised) models with additive randomness, let us take a look at the general structure of the dominant economic models in a stylised way in the next section.

1.2 The Nature of the Data Generating Process

In many situations real-world economic quantities evolve not in a fully predictable way. We can roughly classify two approaches to study such evolutions. In the first approach the data is thought to be the product of a deterministic data generating process (DGP), thus the focus lies on deterministic relationships. Stochastic influences are enclosed in a noise term which causes mere perturbations of the pure deterministic dynamics. In a second approach the stochastic behaviour itself and properties of the stochastically generated dynamics is seen as the primary research object. Both approaches necessarily work with different models to describe real-world phenomena and find different mathematical tools helpful. Let us consider both approaches in more detail.

Deterministic Data Generating Process

Within the first approach, an economist is usually interested in the value of some economic quantity, y, at some time in the future, *e.g.* for the next time step, $y(t+1)$. The quantity could be for example gross domestic product (GDP), the unemployment rate, an interest rate or the price of some financial security traded on an exchange. Following along the lines of econometrics, 'classical time series analysis assumes that the systematic components, *i.e.* trend, business cycle and seasonal cycle, are not influenced by stochastic disturbances and can thus be represented by deterministic functions of time. Stochastic impact is restricted to the residuals ...'.[9] If the endeavour is not simple data mining, but initiated to test the relationships postulated by some economic theory, the economic quantity is commonly explained and therefore understood deterministically as being the result of a weighted sum of different independent components, denoted by the functionals φ_i in the following generic

9 Kirchgässner et al. 2013, p. 3.

way

$$(1.1) \qquad y(t+1) = \text{const.} + \underbrace{\beta_1 \cdot \varphi_1(x_1) + \beta_2 \cdot \varphi_2(x_2) + \ldots + \beta_n \cdot \varphi_n(x_n)}_{\text{deterministic 'explanatory' part}} + \underbrace{\varepsilon(t)}_{\text{stochastic residual}}$$

$$(1.2) \qquad y(t+1) = \text{const.} + \sum_{i=1}^{n} \beta_i \cdot \varphi_i(x_i) + \varepsilon(t)$$

$$(1.3) \qquad y(t+1) = \text{const.} + \varphi(x_1, x_2, \ldots, x_n) + \varepsilon(t) \; .$$

The functionals, φ_i, encode in principle arbitrary transformations or functional forms, that can be themselves again weighted sums or moving averages or autoregressive functions of the independent variables x_i and past values of it. Such functionals often involve the lag or backshift operator, $B^k[x(t)] = x(t-k)$, $k \in \mathbb{N}$. The explanatory variables, x_i, can in principle also be values of the explained variable at an earlier time, *e.g.* $x_i = B^k[y(t)]$. Finally, all the functionals can be further pooled in an envelope functional, φ in eq. (1.3). In the framework of linear regression analysis, the weights in the sum, the βs, are called the regression parameters and the constant is called the intercept. Often, the used lag orders are identified via a fitting exercise on a training data set[10] and are unfortunately in the practice of empirical economics only rarely backed by theoretical considerations beyond seasonality and the like.

In econometrics the functionals are understood as the deterministic or *systematic* part in the explanation. The part of the variation that is not explained in this deterministic fashion is pooled in an *unsystematic* residual, which is simply added for convenience to the deterministic part.[11] This additive stochastic term is denoted by ε in the equations (1.1-1.3). This implies, the random variable ε does not contribute to the outcome of the explained variable y, in any systematic way. A common and convenient distributional assumption for ε is that it follows a so called white or Gaussian noise spectrum, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, which yields the standard normal distribution if $\sigma^2 = 1$. This assumption is convenient because if the noise terms are additive (as criticised above by Buiter), independent and distributed identically following some distribution of finite variance certain limit theorems apply, *e.g.* the central limit theorem. All together this serves the purpose of predictability which is facilitates if the noise is in general easy to handle, often it even allows the noise to be neglected because of its zero mean.

Already the choice of the name of the stochastic *residual* is indicative. Traditionally, random

10 Of course what is called mere fitting exercises may involve sophisticated identification strategies such as certain speeds of decaying (auto)correlations.

11 Kirchgässner et al. 2013, p. 3.

variables, that have in principle only negligible effects, are denoted with the symbol ε, because the symbol was used to label measurement errors. Certain measurement errors can be made arbitrarily small by taking further measurements and averaging over them. The standard Gaussian distribution was discovered (among others) by Carl Friedrich Gauss and (among other ways also) through repeated independent measurements in *e.g.* land surveying. The standard Gaussian distribution, $\mathcal{N}(0,1)$, was thereby identified as a powerful attractor in the space of probability distributions for a wide class of probability distributions with finite variance.

Mild Stochasticity in Neoclassical Finance Theory

It is not surprising that certain aspects of this general development in science spill over to economics. However, what is surprising is the specific form in which it appears. Following the efficient markets hypothesis (EMH) as the central pillar of neoclassical finance theory, the formation of prices on financial markets takes place in such a way, that prices are susceptible to the arrival of new information. In such a market populated with participants holding rational expectations they seek to process all available information in their investment reasoning. Only the arrival of unanticipated thus unpredictable new information generates the resulting stock market movements which are hence modelled stochastically. An apriori classification of which new information will influence the price and which will not, seems to be difficult to almost impossible in practice, and is not covered by the EMH. But the postulated results of Gaussian distributions or Gaussian noise was aptly dubbed by Mandelbrot as an ad hoc restriction to only 'mild' forms of randomness. 'Wilder' forms of randomness involve heavy tailed distributions and pose a problem for neoclassical finance theory, because certain key quantities depend on the existence of higher moments which do not exist for example in the case of α-stable distributions.[12]

Instead the stochastic universe perspective is based on the study of the inherent stochastics of the market and not restrict the specific form of stochasticity of the result from the outset.

A prime example for the conception of stochasticity in post 1945 economics is provided by the following extended quote of Paul Samuelson. It is apparent that he was aware of the early work on ergodic theorems of the 1930s and thus acknowledged the need for a stochastic turn,

> "The same formal laws of Newtonian-Lagrangean-Hamiltonian mechanics were now to be applied to a vast assemblage of molecules. It turned out, after such writers as G. D. Birkhoff, B. O. Koopman, J. von Neumann, N. Wiener, and E. Hopf had,

12 Kirstein 2012.

> by means of the 'ergodic theorem', cleared up some of the difficulties of logical analysis, that the classical mechanics of this extended assemblage of molecules still had the intrinsic time reversibility properties of frictionless pendulum. [...]
> At this point, classical thermodynamics became less convenient than kinetic theory and statistical mechanics. Boltzmann interpreted entropy in terms of probability, much as is done today in information theory.
> [Which led to] abandoning tight causality in favour of a probability model."[13]

But the sketched approach (which coincides with our first approach to stochasticity) is clearly based on the restriction of the problems to the ergodic case. In his usual ambiguous manner:

> "Why should a person interested in economics for its own sake spend time considering conservative oscillations of mechanics? Experience suggests that our dynamic problems in economics have something in common with those of the physical and biological *sciences*. But, as I long ago indicated, it is not useful to get 'bogged down in the research for economic concepts corresponding to mass, energy, momentum, force, and space.' And I may add that the sign of a crank or half-baked speculator in the social sciences is his search for something in the social system that corresponds to the physicist's notion of 'entropy'.
> [...] Just as Ehrenfest and other physicists had to add probability to the causal systems of physics to get around the time-irreversibility feature of classical mechanics that was so inconsistent with the time asymmetry of the Second Law of Thermodynamics, so we must, in the interests of realism, add stochastic or probability disturbances to our economic and biological causal systems.
> [...] Thus, as time passes and more exogenous shocks are experienced, we reach an 'ergodic state', where the new shocks are just enough to balance out the forgetting of old shocks. [...] [The distributions] do not widen indefinitely; instead they approach a limiting distribution (which is quite independent of where I started). [...]
> [In footnote 7] Now adding a little stochastic disturbance [...] will lead to an ergodic distribution."[14]

It will not be the last appearance of Samuelson in the remainder of this thesis, and always are his words like a two-edged sword.

[13] Samuelson 1965, pp. 117–118.
[14] Samuelson 1965, p. 121-123, original emphasis.

Stochastic Data Generating Process

However, this view of the role of minor random influences in an overall deterministic law-like relationship has been overhauled by a conception of stochasticity at the very core of the universe referred to as the stochastic cosmos[15], which is where we locate the second approach. If we study the world and economic dynamics, there are at least three reasons which impose the use of stochastic methods on us:

1. First, we can not follow the purposeful behaviour of individual economic agents in a model if they surpass a certain number – more formally, there are too many degrees of freedom.

2. Second, the non-linear interactions between the elements of systems tend to bring forth chaotic dynamics even from complete determinism on the individual level.

3. Ultimately, the basic understanding of the microcosmos relies on quantum mechanics, which is at its very core probabilistic, thus encloses a form of insurmountable or fundamental uncertainty.

To invoke the quantum revolution is actually unnecessary for our purpose, but gives further validity to the view of a stochastic organisation of the universe and adds weight on the scientificity of a probabilistic approach. The list of reasons for stochasticity shows the complementarity and compatibility of complete determinism at an intermediate, say mesoscopic, level of (condensed) matter and the everyday world how it surrounds us, and the need for probabilistic treatments of whole economies and dynamic systems on the macroscopic level on the one hand as well as at the microscopic level of individual decision makers and subatomic particles on the other hand. We find determinism and pure causality embedded in an unstable Goldilocks state between layers of contingency and stochasticity.

Randomness as a Measure of Ignorance

Already with the analogy of an omniscient intelligence later referred to as the LAPLACIAN demon, LAPLACE defined the limits of determinism and predictability:

> "In ignorance of the ties which unite such events to the entire system of the universe, they have been made to depend upon final causes or upon hazard [chance], according as they occur and are repeated with regularity, or appear without regard to order; but these imaginary causes have gradually receded with the widening bounds of knowledge and disappear entirely before sound philosophy,

[15] See *e.g.* NICKEL (2000) for a careful dissection of determinism and change in the context of the mathematical theory of dynamical systems.

> which sees in them only the *expression of our ignorance of the true causes.* Present events are connected with preceding ones by a tie based upon the evident principle that a thing cannot occur without a cause which produces it. This axiom, known by the name of the principle of sufficient reason [...]
>
> We ought then to regard the present state of the universe as the effect of its anterior state and as the cause of the one which is to follow. Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it – an intelligence sufficiently vast to submit these data to analysis – it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes."[16]

Later POINCARÉ motivated the probabilistic analysis by our ignorance due to the impossibility to know the initial conditions with infinite precision, which quite naturally leads to chaotic behaviour from purely deterministic principles

> "A very small cause which escapes our notice determines a considerable effect that we cannot fail to see, and then we say that that effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But, even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation approximately. If that enabled us to predict the succeeding situation with the same approximation, that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous [random] phenomenon.
>
> [On roulette and games of chance] It is, accordingly, impossible for me to predict what the needle I have just spun will do, and that is why my heart beats and I hope for everything from chance. The difference in the cause is imperceptible, and the difference in the effect is for me of the highest importance, since it affects my whole stake."[17]

In this sense a probabilistic theory is still based on a mechanistic explanation but includes the ignorance of the model builder about things he simply can not know or control for. Thus

[16] LAPLACE 1902, pp. 3-4, our emphasis.
[17] POINCARÉ 1914, pp. 67–68, 70.

we recognise the straw man argument behind presentations of stochastic models as acausal theories. Khinchin gives a similar account which is closest to the setting in which stochasticity and ignorance appear in this thesis:

> "For instance, under normal conditions we consider the number of tickets drawn in a lottery as a random variable. However, if we succeed in studying the mechanism of the drawing to such extent that we shall be able to determine this number beforehand, or, still more, if we succeed in drawing the number as we desire, then all elements of randomness disappear, although the mechanism of drawing is the same in both case."[18]

All these statements are expressions of the development of science which led to a modified view from the clockwork universe to a stochastic cosmos in stark contrast to the biblical meaning of chaos as a state of irregularity. Thus it is possible to discern order and regularity but on another level. Norbert Wiener's remarks carry us over to the next section, which introduces the Wiener-Kolmogorov conception of the stochastic organisation of the universe:

> "The errors of observation are part of the observation itself, and any separation between precise observations and errors will lead in the case of the semi-exact sciences to a methodology which actually is less accurate for the results we really desire than a broader, less obviously precise method."[19]

1.3 Stochastic Cosmos

> The fact that we are living in a stochastic universe
> has barely penetrated [economics],
> least to the standard textbooks where the idea of
> uncertainty is almost never mentioned.[20]
>
> Oskar Morgenstern

Both of the two preeminent probabilists of the 20$^{\text{th}}$ century, Norbert Wiener and Andrei N. Kolmogorov, incorporated randomness due to ignorance and 'the idea of stochasticity of nature because of its extreme complexity.'[21] While the former was driven by his desire for a

[18] Khinchin 1949, p. 53.
[19] Wiener 1956a, p. 252.
[20] Morgenstern 1972, p. 702.
[21] Molchanov 1997, p. 3; See also Masani 1990, Ch. 12C.

rigorous mathematical understanding of Brownian motion[22] which also led to *Cybernetics*,[23] the latter approached stochasticity with the related aim to axiomatise probability theory.[24] One major progress in the treatment of random variables by Kolmogorov was augmenting the *unmeasurable sample space* with an *event space* to a *measure space*. The sample space can contain ex ante inconceivable events which are thus immeasurable, but probability theory focused as of now on the measurable event space.[25] These achievements substantiate the idiom of the Wiener-Kolmogorov conception of the stochastic organisation of the universe. The thesis at hand follows in this tradition of the idea of the Wiener-Kolmogorov conception of the stochastic organisation of the universe and is addressing a correct handling and the identification of suitable forms of stochasticity in the realm of economics. The acknowledgement of stochasticity as fundamental at same microscopic level and the study of the regularities at the macroscopic level which arose from interactions at the microscopic level is the main methodological difference between our approach compared to traditional econometric approaches as mentioned above, that tend to favour the identification of purely deterministic mechanisms.

The movement, to which Wiener and Kolmogorov contributed so profoundly, changed the scientific enterprise to a great extent. Mumford (2000) describes how the inclusion of the random variable as a fundamental object freed science from the Aristotelian 'straight-jacket' of a binary logic and an unsubtle understanding of causality. It opened the space for contingency and the possibility of probability and statistics at its foundation. Therefore, Mumford (2000, p. 217) concludes

> "stochastic methods will transform pure and applied mathematics in the beginning of the third millennium. Probability and statistics will come to be viewed as the natural tools to use in mathematical as well as scientific modeling. The intellectual world as a whole will come to view logic as a beautiful elegant idealization but to view statistics as the standard way in which we reason and think."

A key challenge in this new way of probabilistic reasoning is the identification of a suitable probability measure $\mathcal{P}$ of the natural phenomenon which is modelled as a random experiment. To model a real world phenomena as a random and in principle repeatable experiment is the central point, as Molchanov (1997, p. 4) accentuates: 'the probability measure P in the basic probability space (Ω, F, P) is postulated, but what is the relation between the real statistical experiment and the measure P?'[26] Especially, Molchanov emphasises the role of ergodicity in the relation of the model to the real world which manifests the underlying assumption of ergodic and stationary increments at all times t and not only asymptotically

[22] Wiener 1920, 1921a,b, 1923.
[23] Wiener 1985.
[24] Kolmogoroff 1933.
[25] Johnson 2016, p. 193.
[26] Rigorous formalism of probability space, probability measures and the like follows in chapter 2.

(for $t \to \infty$).[27] Instantaneous ergodicity thus demands even more than usual ergodicity in same asymptotic limit. Wiener emphasised the role of contingency in a stochastic cosmos in the following way:

> "My early work on probability theory, as exemplified in my studies of the Brownian motion, had convinced me that a significant idea of organization cannot be obtained in a world in which everything is necessary and nothing is contingent [...]
> Organization we must consider as something in which there is an interdependence between the several organized parts but in which this interdependence has degrees. [...] [T]he internal interdependence is not complete and that the determination of certain quantities of the system leaves others with the chance to vary. This variation from case to case is a statistical one and nothing less than a statistical theory has enough freedom in it for the notion of organization to be significant.
> I was driven back on the work of Willard Gibbs, and on the conception of the world not as an isolated phenomenon but as one of many possible phenomena with an allover probability distribution. I was forced to consider causality as something of which there can be either more or less rather than as something which is either there or absent. [...]
> I formed a new respect for the irregular and a new concept of the essential irregularity of the universe."[28]

This thesis is in line with a general research programme that has been utilising contingency as a general precondition of thinking scientifically, similar to a methodology in modern non-equilibrium statistical physics, modern biology[29] and in particular in evolutionary economics.[30] It is in line with it in so far as to break with simple mono-causality even if we do not strive towards exploring the whole spectrum of degrees of causation between zero and one. This thesis differs from the aforementioned contingency approach in evolutionary economics in so far, that due to the immense number of factors which could influence economic phenomena our approach is a probabilistic treatment of aggregates.

Probabilistic Treatment of Aggregates

Boltzmann always believed that nature is deterministic in principle. The physics of his times became 'stochastic' due to a rigorous derivation of probabilistic laws of collective wholes from the deterministic laws governing individual molecules through the method of statistical mechanics, to which he contributed so enormously. The probabilities that appear in statistical

27 Molchanov 1997, *e.g.* see p. 20.
28 Wiener 1964a, pp. 322–323.
29 Monod 1971.
30 Lehmann-Waffenschmidt 2010.

mechanics are therefore the result of a rigorous derivation from deterministic effects at the level of molecules or even atoms. Thus probabilities have an ultimate source and arise from deterministic relationships. In the stochastic cosmos, a stochastic order is created out of deterministic chaos[31]. Hence, the development of statistical mechanics is seen by many scholars not as a real probabilistic revolution even if often labelled in such a way.[32] A true probabilistic revolution in science appears not before the era of quantum theory, which is based on KANTIAN apriori probabilities in its foundational mathematical structure.[33] These quantum probabilities cannot be thought of as the result of a derivation from some microlevel determinism, because quantum mechanics is already a description of the microlevel. Seeing these developments as revolutionary or not has nevertheless led to an enlarged acceptance of probabilistic tools in science.

Among the earliest to realise the analogy between the revolution in quantum physics and the social sciences was ETTORE MAJORANA around 1937. In his article on ‘The value of statistical laws in physics and social sciences’[34] he follows the ideal of the unity of science, which has proved extremely useful ‘as a powerful stimulus to the progress of science many times.’[35] This ideal was assured in the success of the application of mechanics to heavenly bodies, which became the celestial mechanics of KOPERNIKUS, BRAHE, GALILEO, KEPLER and NEWTON.[36] MAJORANA recasts in a sense the argument of the LAPLACIAN demon of determinism in the light of the recent development of quantum mechanics:

> “According to this point of view, which has produced the mechanistic conception of nature, the entire material universe evolves obeying an inflexible law, where the state of the universe at a given instant of time is completely determined from its state at the previous instant of time. This is a sign of the fact that the future is implicit in the present. In other words the future can be predicted with absolute certainty provided that the actual state of the universe is completely known.”[37]

He continues with the importance of fundamental limitations to the ‘mechanistic conception of nature’. One of the limitations is to obtain initial conditions with infinite precision from measurements. *E.g.* in physics due to complex translational and rotational movements of atoms and molecules. Much in the same way as in the social sciences where we analyse ‘certain global characters by deliberately renouncing the investigation of additional information, such

31 See PRIGOGINE and STENGERS (1984) and KAUFFMAN (1993) for the order out of chaos metaphor
32 KRÜGER et al. 1987a.
33 KRÜGER et al. 1987b.
34 MAJORANA (2005), see especially the introduction by ROSARIO MANTEGNA, who puts it in the historical context of the time the article was written around 1937 shortly before MAJORANA disappeared.
35 MAJORANA 2005, p. 136.
36 In order to get a better idea of the historical context of these mentioned scientists see their birth and death dates NIKOLAUS KOPERNIKUS (1473-1543), TYCHO BRAHE (1546-1601), GALILEO GALILEI (1564-1642), JOHANNES KEPLER (1571-1630) and ISAAC NEWTON (1642-1727).
37 MAJORANA 2005, p. 136.

as, for example, the biography of all individuals composing the society under investigation.'[38] Hence, already at a microscopic level in both physics and the social sciences, we are confronted with descriptions of a statistical character of the system constituents, which nevertheless do not render them acausal. Following this line of thought the free will of humans is reconcilable with a probabilistic framework with deterministic microfoundations of causal or even purposeful behaviour.

For example the mortality rate of humans can be modelled using a constant probability independent of age, although people clearly do age and irrespective of many causal influences that one can think of to influence this rate over the lifetime. This is very similar to the decay of a radioactive particle, which however shows no sign of ageing when time elapses, but is well described via a constant mortality rate, too. In particular, from the constant mortality rate we can derive the probable lifetime and the specific exponential form of the survival curve for a collective of people or particles. The prediction of the natural death and what exactly triggers the natural death of a human is hard to pin down precisely, but surely it is not an acausal phenomenon, it is due to our ignorance of the state of the body up to microscopic level that we label it in certain contexts as a random event. The decay of a particle, however, is as unpredictable as the time of death but can additionally be conceived as an acausal event that happens abruptly and in isolation of any possible causes.[39]

The combination of only imperfectly known initial conditions of the constituents of an observed system together with fully deterministic physical laws introduces probabilistic methods into the sciences of the inanimate as well as the animate world. Majorana acknowledges that the 'exceptional credit of physics evidently comes from the discovery of the so-called exact laws',[40] but that physics has been forced to move on from this static conception of the universe because of new scientific evidence. Ultimately, the probabilistic approach is a necessary consequence of any deterministic approach of treating aggregates (societies or molecules) about which we have only imperfect knowledge with regard to the possibly relevant causal factors, which all have deterministic causal effects in isolation. The probabilistic ansatz is therefore still a causal explanation, although it does not allow to identify single causes, but produces only statements about collective causes of a collective of entities, *i.e.* aggregates.

38 Majorana 2005, p. 138.
39 Majorana 2005, p. 139.
40 Majorana 2005, p. 135.

1.4 Gambles, Bets and Lotteries as a Tool of Understanding Dynamics

Life is a gamble, at terrible odds –
if it was a bet you wouldn't take it.[41]

TOM STOPPARD

Since the origins of probability theory gambling problems are central to its development. Many famous paradoxes arose from the discussion of lotteries. Also in economics, lotteries play a fundamental role as a model of situations which involve a random outcome. VON NEUMANN and MORGENSTERN (1967, ch. 3.3., 3.7.) restated mathematical economics based on lotteries when they axiomatised expected utility theory (EUT), thereby bridging the gap from ordinal to cardinal utility. Utility as a metric variable was thought to be necessary to compare several lotteries by means of their transformed expectation values as the decision criterion. The introduction of lotteries is associated with a change of the object of the preferences of the economic agents. Before VON NEUMANN and MORGENSTERN preferences were built over bundles of products and services, after them agents built preferences over probability distributions in the context of lotteries.[42] This is what is thereafter known as VON NEUMANN-MORGENSTERN utility. We carefully scrutinise utility theory in the remainder.

Closely interlinked with utility theory is the St. Petersburg lottery which posed a paradox again and again in the history of probability theory. DUTKA (1988, p. 18) attributes the importance of the paradox surrounding the St. Petersburg lottery 'far beyond probability theory, in which it plays only a technical role. It challenged the view, already well established in the early eighteenth century, that the calculation of the mathematical expectation of contingent events could serve as a reliable guide for rational action in the real world.' That is why the particular structure of any specific lottery, *e.g.* the St. Petersburg lottery, is not of central importance, but the strength of the *Ergodicity Economics* ansatz – which we follow in this thesis – is the more general approach of evaluating gambles using dynamics. The dynamic can be generated via lotteries or by completely other generators of stochasticity. As a motivational example in order to illustrate the power of gambles, consider the following real-world business models.

41 STOPPARD 2003.
42 SELTEN 2001, pp. 3–4.

Gambling Business Models

Let us consider venture capital and pharmaceutical companies. Their business model has very much in common with betting. Usually, these companies have many start-ups/drugs in their portfolio, that are individually both very risky and costly. The bet these companies take is that they hope to have invested their money in at least a single winner among all the investments they entered. The winner will hopefully refinance the losses accrued by all the other unsuccessful investments and they thus break even. In more favourable scenarios the return of a single investment exceeds the accrued losses of the other unsuccessful investments by far. However, in order to survive, they need to make sure, that no single unsuccessful investment has the potential to ruin them. If one of such bets could become lethal, this is called overbetting in the gambling terminology. What such companies fear is therefore the risk of overbetting. These companies have to make sure that they do not depend on the success of a single definite start-up or drug. But by their decision to invest in a particular drug or start-up they actually buy the long option position that is otherwise not publicly traded. The here depicted business model shares many similarities to a rational gambling strategies and necessarily involves many missed opportunities in the sense of a rejection to provide seed capital for companies or drugs that flourished afterwards. For example there are many reports even of successful venture capitalists that were offered the opportunity to enter early in Google, Facebook, Skype, Apple or AirBnB but did refuse the offer. However, it is the avoidance of entering lethal gambles which is important and not catching every favourable one.

We conclude the introduction with on overview of the structure of thesis and the contributions[43] in the respective chapters.

1.5 Contributions and Structure of the Thesis

> [I]t is also crucial to express your ideas in a way that other people, who have not spent the last few years wrestling with your problems and are not eager to spend the next few years wrestling with your answers, can understand without too much effort.[44]
>
> Paul Krugman

[43] Therefore we switch only in the next section from the 'we'-form to the personal pronoun 'I'.
[44] Krugman 1993.

With the decision to devote my thesis to the ergodicity problem I not only encountered a gap in the literature but I identified a blind spot in economics, which so far ignored the dynamic of the context in which a decision takes place. This is first and foremost a pressing and a conceptually deep and fundamental research problem.[45] So far the economics community absorbs the insights of *Ergodicity Economics* started by the seminal paper PETERS (2011c) slowly, although the results have gained some momentum outside of economics. This additionally motivates my choice of the ergodicity problem in economics as a dissertation topic. The reason for the slow absorption is twofold. First, a full penetration of the subject requires a very special skill set that is usually not imparted in a traditional economics programme with regard to the necessary mathematics of measure theory, dynamical systems theory and ergodic theory needed to handle random variables and stochastic processes appropriately. During my doctorate I acquired a profound understanding of several scientific disciplines besides economics, including statistical mechanics, information theory, evolutionary biology and not least ergodic theory. I acquired these skills especially during several research stays at the LML and a summer school at the SFI and thus had the opportunity to get in contact with the leading researchers of the respective fields. This allowed me to use several tools from these disciplines from which I derive insights which are highly relevant for economics. I questioned the dominant way stochasticity is modelled in economic models and realised the two different approaches mentioned above. However, the conceptual depth which is associated with a switch from embedding the randomness in the ensemble to an embedding of randomness within time is also an insurmountable obstacle for some. One contribution of my thesis in its totality is that I aspire to write it in a language and level of detail which facilitates a super diffusion of the ideas of *Ergodicity Economics* through the medium of scientists with all kinds of backgrounds but with a special emphasis on economists. For example, I cultivate the narrative of the embedding of randomness either within the ensemble or within time throughout the thesis, which I first come across in PETERS (2011c). In later publications, the latter is often referred to as the time interpretation or the time perspective. With as little polemics as possible I outline why the new perspective is sometimes radically different when compared to traditional approaches in economics.

In the following I address the contributions I make in every chapter, which at the same time serves as a description of the structure of the thesis.

Contribution in Chapter 1

In this introductory chapter I mainly prepare the stage and the reader for what is to come. I present two different approaches to deal with stochasticity. The predominant way in economics embeds the randomness within the ensemble and treats stochasticity as a negligible residual

[45] PETERSEN 1996; PETERS and WERNER 2017.

error term. *Ergodicity Economics* belongs clearly to the second and distinctly different perspective of the stochastic cosmos. In the stochastic cosmos the stochasticity plays a constructive role in generating order and structures on higher levels and the randomness is embedded within historical time.

Contribution in Chapter 2

In chapter 2 I provide a profound analysis of ergodic theory, which can serve as an entry for economists in Sec. 2.4, especially the combination of Sec. 2.1, Sec. 2.2 and Sec. 2.3 is helpful if one does not intend to spent as much time on the foundations of statistical mechanics and ergodicity as I did. The key contribution is that I establish connections between the mathematical fields of ergodic theory, the theory of dynamical systems and stochastic processes. In particular, I derive how the stochastic processes I study in later chapters are generated by an ergodic measurable dynamic system. So the data generating process is not a deterministic relationship as I indicate above, but an ergodic measurable dynamical system. This key contribution is in Sec. 2.4. In the overview of some results in ergodic theory in Sec. 2.6 I show interesting relations to other fields of physics and mathematics in which ergodic theorems are used. The content of this chapter is crucial for the chapters 4 and 8.

Contribution in Chapter 3

I chapter 3 I introduce the St. Petersburg lottery using the new terminology of *Ergodicity Economics*, which is mostly preliminary for the next chapter.

Contribution in Chapter 4

The publication of Peters (2011c) motivates chapter 4, which contains one of the main contributions of this thesis. My presentation follows the categorisation of the preexisting solution strategies by the mathematical transformation they introduced and a clear analysis of their weak spots and ad hocness. The key contributions are in Sec. 4.3 and Sec. 4.5.

In Sec. 4.3 I identify the innovations in Bernoulli (1738) and more importantly what was not in the paper. I derive in great detail the importance of the introduction of wealth instead of unescorted payouts. The introduction of utility is a spurious innovation, because it introduces an arbitrary psychological element in a theory of rational decision making under uncertainty, moreover it is mathematically flawed in the original source. Furthermore, I analyse what is the true content of EUT and assign the proper role to Laplace as the actual founder of EUT in the form it is used in economics today. I clearly show why utility is not a solution to the St. Petersburg paradox and which decision criterion D. Bernoulli actually derived.

In Sec. 4.5 I present the solution using the non-ergodic dynamics. In meticulous detail I introduce the difference between additive and multiplicative dynamics in Subsec. 4.5.1, I derive the ensemble-average growth rate in Subsec. 4.5.2 and the time-average growth rate in Subsec. 4.5.3. Finally, I derive in Subsec. 4.5.5 that not only the St. Petersburg lottery breaks ergodicity, but multiplicative dynamics is in general non-ergodic. I draw the connections between EUT and pick up many threads from Sec. 4.3. The most important thread is the role of the logarithm, which also appears in utility theory as a candidate functional form. It turns out that the logarithm has nothing to do with a utility function but is an ergodicity transformation. I devote the Subsec. 4.5.4 to this key issue.

Of course, the analysis profited enormously from the most recent publications and some unpublished material by Peters (2011a, 2014, 2018a) and Peters and Adamou (2018c), but nevertheless it must be stated explicitly that there is a lot of confusion about these topics in the literature. This makes it necessary to digest the topic from different angles, *e.g.* the perspective I choose in this chapter traces the different solution strategies in the history of the St. Petersburg lottery and combines them in a coherent discussion guided by the introduction of particular mathematical operations. Furthermore, the majority of the papers and results I work with that did not appear in economics journals but mostly in physics, biology or complexity outlets and I carve out their relevance for economics.

Contribution in Chapter 5

The literature review in chapter 5 is a contribution in a threefold way. The first contribution in Sec. 5.1 offers a review of the self-contained field of *Ergodicity Economics*, to the extent that the results are not already covered in other parts of the thesis in greater detail. In Sec. 5.2 I offer a genuine collection of topics which are discussed in economics and have a direct relation to the ergodicity problem, even if this is not stated explicitly in these fields or not fully comprehended. Besides the processing of very recent contributions to *Ergodicity Economics* and the exposition of the relations to research in economics, another genuine contribution of mine lies in the identification of relations to other research areas in Sec. 5.3. Here, I demonstrate the connection to certain topics in statistical mechanics, evolutionary biology, information theory and process philosophy. At the same time I offer an invitation to become acquainted with a variety of fields all connected via the ergodicity problem.

Contribution in Chapter 6

In this chapter I contribute to the field of science theory and intellectual history, sometimes called the history of economic thought. This chapter is partly based on previous work[46] which I greatly expanded. I trace the genealogy of the idea of ergodicity as a precondition for scientific conduct in post 1945 mathematical economics from Gibbs via Wilson to Samuelson. I identify the ergodic axiom as a key driver of the particular direction of mathematical economics and the mathematical tools that were introduced since 1945 to economics from mathematics, especially differential topology. This mathematisation process is accompanied with a rhetorical spillover from the natural sciences to economics, which I analyse using the instruments provided by McCloskey (1983). I find a rhetoric which surrounds the reliance on the ergodic axiom that is mostly silent. This corroborates the earlier argument that the ergodicity problem is a blind spot in economics.

In Sec. 6.4 I explain the connection between (non-)ergodicity and different conceptions of time which culminates in the ergodic mirror hypothesis. In Sec. 6.4.2 I introduce the ergodic fallacy into the literature and show their relevance in applications of economic and scientific insights in general for policy purposes. Additionally, I identify potential drivers of a broken symmetry in time, with which the broken ergodicity is associated. This hints at deeper connections of the role of (broken) symmetries in science.

Contribution in Chapter 7

The main contribution in chapter 7 is condensed in **Table 7.1**, which presents an overview of the evolution of decision criteria. The evolution prior to the advent of *Ergodicity Economics* is basically a maintenance programme of ensemble-average approaches. Embedding the randomness within time and the decision criterion of a time-average growth rate of wealth constitutes the current climax of the level of maturity in this evolution. This contribution is supported with a discussion of rationality in sciene and economics which I already touch upon in previous chapters.

Contribution in Chapter 8

Chapter 8 contains two contributions. In Sec. 8.3 I show that interpreting the Kelly criterion in the framework of utility theory is tantamount with forcing the Kelly criterion in a Procrustean bed. I show explicitly, that the Kelly criterion has nothing to do with utility, contrary to repeated claims in the literature, although Kelly himself was very explicit on this point. What Kelly actually derived is a maximisation of the time-average growth

[46] Kirstein 2016.

rate of the gambler's wealth, which is the optimisation of the decision criterion derived in Sec. 4.5 and chapter 7. Furthermore, I present an outlook to ongoing cutting edge research of mine in *Ergodicity Economics*, which incorporates the ignorance of the gambler about the exact parameters of the gamble. This research project continues the work I started in this thesis.

2 Ergodic Theory

> Was die Axiome der Wahrscheinlichkeitsrechnung angeht, so scheint es mir wünschenswert, daß mit der logischen Untersuchung derselben zugleich eine strenge und befriedigende Entwickelung der Methode der mittleren Werte in der mathematischen Physik, speciell in der kinetischen Gastheorie Hand in Hand gehe. Ueber die Grundlagen der Mechanik liegen von physikalischer Seite bedeutende Untersuchungen vor; ich weise hin auf die Schriften von Mach, Hertz, Boltzmann und Volkmann; es ist daher sehr wünschenswert, wenn auch von den Mathematikern die Erörterung der Grundlagen der Mechanik aufgenommen würde. So regt uns beispielsweise das Boltzmannsche Buch über die Principe der Mechanik an, die dort angedeuteten Grenzprocesse, die von der atomistischen Auffassung zu den Gesetzen über die Bewegung der Continua führen, streng mathematisch zu begründen und durchzuführen.[47]
>
> David Hilbert

When Hilbert included the clarification of the foundations of physics into his renowned speech on prime mathematical problems for the next century, he ennobled both Boltzmann and any engagement dedicated to the foundations of statistical mechanics. This chapter is devoted to a particular important element in the pursuit for a solid foundation of statistical mechanics known as the ergodic hypothesis and some basic aspects of a field in mathematics derived from it known as ergodic theory. Whereas the general focus of this thesis lies on decision theory under uncertainty, the presentation in this chapter extends beyond what is needed for this immediate purpose for two reasons. First, possible applications of a relatively recent mathematical field (ergodic theory) beyond what is utilised already in some other field (economics) may not be easily conceivable from the outset of a new theory (*Ergodicity Economics*). Second, this thesis is unparalleled in its attempt to convey a general idea of the ergodicity problem, the ergodic

[47] Hilbert 1900, p. 272.

hypothesis and ergodic theory to an economics audience. Therefore, this chapter enriches the endeavour of this thesis with the fascinating and intricate history, philosophy and mathematics behind the ergodic hypothesis. To prepare an understanding of ergodicity, special importance lies on the definition of ensemble and time averages and on sound interpretations of these abstract concepts that were originally developed for statistical mechanics purposes. In the chapters 3 and 4, we show the importance of these concepts for economics in transferring the correct interpretations to real-world problems of risk-taking.

This chapter is organised as follows. We begin with an intuitive motivation of time-averaging and ensemble-averaging in Sec. 2.1 by the classic example of tossing a coin. Sec. 2.2 states the ergodic hypothesis in a nutshell version on less than four pages. The ergodic hypothesis has been pivotal in the transition from classical to statistical mechanics. This is why a nutshell version is both ambitious and maybe enough for the non-physicist to decide how much time to spent on it, before we reconstruct the role of the ergodic hypothesis in the transition from classical to statistical mechanics in greater depth in Sec. 2.3. This section clarifies why the ergodic hypothesis was contrived and is a cornerstone in this transition. Sec. 2.3.2 clarifies the etymology of the name of the mathematical field *ergodic* theory, if we understand it as being derived from the *ergodic* hypothesis in physics. In Sec. 2.4 we introduce some formal concepts that constitute modern ergodic theory and in Sec. 2.6 we refer to some mathematical results in ergodic theory as starting points from where a more in-depth analysis could commence.

2.1 Intuition of Ergodicity – The Coin Toss Experiment

This section gives an intuitive notion of ergodicity via a simple random experiment of tossing a coin, the probabilist's drosophila. The question arises how to embed the randomness of the outcome of the coin flip. Before the experiment starts, we know nothing about the coin other than that it has two faces, 'heads' and 'tails'. The goal is to obtain the apriori unknown probabilities that an arbitrary face appears after a toss.[48] As the experimenter, we have now two ways to embed the randomness when we perform measurements of the probabilities.

[48] Even the number of faces could be in principle unknown and could be discovered by means of the experiment. I utilise an illustrative example from Peters and Mauboussin (2012), who give the same example with throwing dice and determining experimentally the probability that a certain number of pips shows up. Apt wording (of ensemble-average or time-average notion of probability) for this section is not completely rigorous but practical for our purpose and adopted from von Plato (1987). The wording of embedding randomness either in the ensemble or within time was first used in Peters (2011c).

2.1.1 Embedding Randomness in the Ensemble

We let N experimentees independently of each other each toss one coin exactly once. Then we count how many times k_i did an arbitrary face i show up. *E.g.* let k_{heads} be the absolute frequency of 'heads'. With an increasing number of experimentees the randomness, *i.e.* the uncertainty about the unknown probability, in the relative frequency k_{heads}/N decreases and converges to the true probability of 'heads', $p^{\text{EA}}_{\text{heads}}$,

$$\lim_{N\to\infty} \frac{k_i}{N} \to p_i^{\text{EA}}. \tag{2.1}$$

After performing the experiment, we calculated what is called an ensemble average over all tosses indicated by the superscript 'EA' in the true probability p_i^{EA}, which gives rise to the ensemble-average notion of probability. The ensemble consists of the N tosses performed by the N experimentees. Actually, we are interested of the probability of the faces of a specific coin. Though what we have calculated in eq. (2.1) could also be conceived of as the average of N independent copies of a coin and its coin tosser, who each toss their coin once in their parallel world. We will return often to this interpretation in the remainder of this thesis. In order to calculate the probabilities, we could have proceeded also in a different way.

2.1.2 Embedding Randomness within Time

Now we let one experimentee repeatedly toss a single coin once for T times. Then we count how many times m_i did an arbitrary face i show up. *E.g.* let m_{heads} be again the absolute frequency of 'heads'. With an increasing number of tosses by a now single one experimentee the randomness, *i.e.* the uncertainty about the unknown probability, in the relative frequency m_{heads}/T decreases and converges to the true probability of 'heads', $p^{\text{TA}}_{\text{heads}}$,

$$\lim_{T\to\infty} \frac{m_i}{T} \to p_i^{\text{TA}}. \tag{2.2}$$

Thereby we calculated what is called a time average over T consecutive tosses or, synonymously, over a single sequence of tosses of length δt such that $T\delta t$ gives he total duration of the observation. We indicate this by the superscript 'TA' in the true probability p_i^{TA}, which gives rise to the time-average notion of probability. This length can also be understood temporally as a duration over which we have tracked the outcomes of repeated tosses with the same coin by the same coin tosser.

In this simple example of a coin toss, our observable[49] has been the probability p of an arbitrary face i, which we sought to approximate with increasing precision by the limits of the relative frequencies either via the ensemble k_i/N or via the evolution over time m_i/T. Both averages converge in this example to the same quantity p_i,

$$\text{time average of the probability} = \text{ensemble average of the probability} \tag{2.3}$$

$$\lim_{T\to\infty} \frac{m_i}{T} = \lim_{N\to\infty} \frac{k_i}{N} \qquad \text{for } N, T \to \infty \tag{2.4}$$

$$p_i^{\mathrm{TA}} = p_i^{\mathrm{EA}} \qquad = p_i \ . \tag{2.5}$$

Hence, a correct mathematical model of the probability that an arbitrary face i shows up as the observable in the coin toss turns out to be ergodic. The ergodicity of an observable implies that the ensemble average and the time average of that observable coincide in the respective limits and can thus be used interchangeably under certain conditions.

With this intuitive example we have illustrated that randomness can be conceptualised in two different ways, either we embed the randomness in the ensemble or we embed it within time. What is the 'correct' embedding depends on the process we are studying. That both approaches are equivalent in our simple example, hence that the observable 'probability of a side of the coin' is an ergodic observable, is not a generic fact, but constitutes rather a special outcome of the specific combination of the system, the observable and the dynamics we were studying. The ensemble average and time average of an observable do not need to coincide by any means. The assumption of the ergodicity of an observable is a very strong statement no matter if it is explicitly stated or – as it is often the case – only implicitly assumed.

Fundamentally, we applied the law of large numbers (LLN) in both eq. (2.1) and eq. (2.2), which are interpreted as different embeddings of the randomness about the outcome. In doing so, we used two different types of averaging in the limits eq. (2.1) and eq. (2.2). Mathematically they are both arithmetic means and look identical, but conceptionally they are something very different, which turns out to be crucial. One is an ensemble-averaged quantity and the other a time-averaged quantity and the two (limits) need not coincide.

[49] An observable is the quantity we measure. In our example the measured quantity is the absolute frequency m of face $i \in \{\text{heads, tails}\}$, from which we derive the relative frequency and eventually the probability of face i. Therefore, to say the probability is our observable is in need of some lenience from the cognoscente.

Ergodic Hypothesis

The unjustified assumption of the equality of the ensemble average and the time average of an observable, *e.g.* in eq. (2.3), is called the ergodic hypothesis,

$$\text{time average of an observable} \overset{?}{=} \text{ensemble average of an observable}\,. \tag{2.6}$$

It is therefore problematic, that statistical calculation often require ergodicity, which is often only stated implicitly. Sometimes an equivalent expression to eq. (2.6) is used, because the ensemble average of an observable equals its expectation value

$$\text{time average of an observable} \overset{?}{=} \text{expectation value of an observable}\,. \tag{2.7}$$

However, when we use the label of the ensemble average, then the type of average that is applied is made explicit.

Simply assuming ergodicity may simplify the treatment of the system significantly, *e.g.* if we have either not much time for the actual experiment and opt for the ensemble-average notion or if we do not have much room to accommodate a lot of experimentees and opt for the time-average notion. It is important to note, however, that for the same system of coin tosses one can easily create observables that are non-ergodic. *E.g.* if we couple the outcome of the coin tosses to specified dynamic according to which the payouts of the coin tosser accumulate. Then the accumulated payouts contribute to the wealth of the coin tosser. Given very realistic dynamics, wealth will then be a non-ergodic observable. Examples for the latter case of non-ergodic observables and their proper treatment will be the main focus of the remainder of this thesis.

Having Stanislaw Ulam's famous quote in mind: 'Using a term like non-linear science is like referring to the bulk of zoology as the study of non-elephant animals',[50] the set of non-ergodic observables is much larger and richer than the set of ergodic observables. This fact is in no way special for ergodicity, but is always the case for rigorously defined mathematical properties and their complements. The mathematical branch of ergodic theory is thus concerned with the singling out of the exact conditions under which ergodicity is given for certain averages. *I.e.* if we interpret one index in the averages denoting time, then ergodic theory states when are both approaches, the ensemble and the time approach, indeed equivalent. 'Ergodicity' emerged from the ergodic hypothesis in physics, where the substitution of time and ensemble averages is pivotal for the foundations of the statistical approach to thermodynamics. Therefore, we present the ergodic hypothesis as one candidate solution to the ergodicity problem in statistical mechanics in a nutshell version in Sec. 2.2, and continue to shed more light on the origins

[50] Campbell et al. 1985, p. 374.

of ergodicity in Sec. 2.3. After these topics we turn to the pure mathematical aspects and formalise in Sec. 2.4 what has been presented so far either in an intuitive manner or from a physics perspective.

2.2 Ergodicity Problem in Statistical Mechanics in a Nutshell

> Ergodic theory has its roots in statistical mechanics. Both in the older Maxwell theory and in the later theory of Gibbs, it is necessary to make some sort of logical transition between the average behavior of all dynamical systems of a given family or ensemble, and the historical average behavior of a single system.[51]
>
> Norbert Wiener

Section 2.1 provided an intuitive understanding of ergodicity, the ergodicity problem, and two possible ways for embedding the randomness associated with the coin toss. In a similar vein, this section provides a basic understanding of the ergodic hypothesis as one candidate solution to the ergodicity problem in statistical mechanics. In the introduction of this chapter the endeavour of a presentation of the ergodicity problem in statistical mechanics on less than four pages has already been labelled ambitious. This is underscored if we bear in mind that a brilliant effort took such an eloquent writer like Khinchin almost 70 pages in a piece directed at an expert audience.[52] The reader familiar with the matter is requested to generously connive that important aspects remain largely unmentioned[53] and the necessary brevity of our effort. Legitimate suggestions for improvement should recede behind the acknowledgement of our higher aim, which is geared towards the non-physicist to arrive at a basic understanding of the ergodic hypothesis as one candidate solution to the ergodicity problem in statistical mechanics. In doing so it is necessary to make stronger use of statistical mechanics' terminology here than in other sections.[54]

The ergodic hypothesis is a core component of statistical mechanics and tries to connect the dynamics of a system with its statistics. Therefore, we imagine a collection of N

[51] Wiener 1939, p. 1.
[52] If not the whole book, see Khinchin (1949, ch. I - III). See also Masani (1990, ch. 11 & 12). Other classic book-length accounts are in Tolman (1938, ch. II-VI, X) and Sklar (1993, ch. 5 & 7).
[53] Such as the equipartition theorem, metric indecomposability, metrical transitivity, ...
[54] This section was induced by the reading of Patrascioiu (1987), who assembled information on the Fermi-Pasta-Ulam-Tsingou (FPUT) experiment, equipartition and non-linearity in a review combining the mathematics and physics associated with the ergodic hypothesis.

particles confined in a closed container and call this a thermodynamic system. For the purpose of statistical mechanics the particles are indivisible and interaction between them and the container walls follows entirely elastic collisions.[55] The dynamics are described via the behaviour of an observable of system, which often describes a macroscopic state *e.g.* temperature or pressure of the thermodynamic system as a whole. The macroscopic collective state is brought about by the microscopic states of all constituents of the system, here the gas molecules. The statistics are described via the specific arrangement of the N particles in the container, which is usually called their microscopic configuration or, in short, their microconfiguration. The microconfiguration of a thermodynamic system is uniquely specified by the HAMILTONIAN or phase function, which is a function that assigns a state of the system to pair of momenta (velocities) $\boldsymbol{q} = (q_1, q_2, \dots, q_N)$ and positions $\boldsymbol{p} = (p_1, p_2, \dots, p_N)$ of its N constituent particles. Later we denote the phase space by S. The two parameters for a single particle $(\boldsymbol{p}_n, \boldsymbol{q}_n)$ are also vectors in a d-dimensional space each with d coordinates. Different microconfigurations denoted by different positions in phase space can in principle result in the same value of the observable. The value of the observable is synonymously called the outcome of a measurement, the macroscopic state of the system or the phase of the system. *E.g.* repeated measurements of the room temperature at an interval of two minutes measure each time different microscopic arrangements of the particles, but still lead to similar values for the temperature of the room. Similar repeated measurements of the GDP *p.c.* at intervals of one year can lead to the similar values even if the economy has completely changed in the meantime. The phase function which assigns every point in phase space a macroscopic state is thus not bijective. Conversely, this implies the frequency of specific values of the macroscopic observable are brought forth by collections of different sizes of the different microscopic configurations. Given the simple assumption that all states are a priori equally distributed the equilibrium state can be derived easily. It is defined as the most likely system state and is that macroscopic state to which the most points in phase space map to.

What is called a dynamics of such a system is simply the time evolution of an observable of the system and commonly represented as a trajectory through its phase space. In principle, the phase space is given by all possible combinations of microconfigurations, that is by the combinations of all possible values of the pairs $\{\boldsymbol{p}, \boldsymbol{q}\}$. The phase space of a single particle with the two parameters momentum and position in three-dimensional space has already $2 \cdot d = 2 \cdot 3 = 6$ dimensions. The total phase space of N particles of the thermodynamic system each with their momentum $\boldsymbol{q}$ and position $\boldsymbol{p}$ is a $2 \cdot d \cdot N = 6N$-dimensional hyperspace or $\mathbb{R}^{6N}$. However, the phase space of a real existing system only forms a subregion of this hyperspace, because a finite amount of energy in the system constrains it from covering the whole hyperspace. Therefore, the phase space can be thought of as a manifold (of often much

[55] Although the particles can be molecules or atoms and through particle physics we know that they are collectives themselves.

lower dimension) embedded in the $\mathbb{R}^{6N}$ hyperspace, which is frequently referred to as the energy shell of the system. If the total hyperspace is a hyper-onion, then the energy shell is comparable to one layer of the skin of that hyper-onion but possibly with much more complicated overlaps and intersections than the regularly stratified onion skins.[56] For example if an office room is a real-world example of the model gas system, it contains a number of particles of the order of AVOGADRO's constant, $N_A \approx 6 \cdot 10^{23}$. The classical approach to study such a system relies on NEWTONIAN mechanics. It involves the tracking of all the entirely elastic collisions between the particles which obey the conservation of momentum and is called kinetic theory of gases. Any attempt of an experimental verification of realistic systems of the size $N \sim N_A$ via the kinetic approach is clearly out of reach.

2.2.1 Time Average

However, we can define some arbitrary function $f(\boldsymbol{p}, \boldsymbol{q})$ and call it an observable f of the system, *e.g.* its temperature, energy, average velocity per particle, pressure, number, *etc.* The time-average of an observable, $\overline{f}$, starting at the initial conditions $(\boldsymbol{p}(0), \boldsymbol{q}(0))$ is defined as the integral over the trajectory of $\overline{f}$ in phase space from $t = 0$ up to the end of the finite measurement period T,

$$\overline{f}_T\Big(\boldsymbol{p}(0), \boldsymbol{q}(0)\Big) := \frac{1}{T} \int_0^T f\Big(\boldsymbol{p}(t), \boldsymbol{q}(t)\Big)\, \mathrm{d}t \,. \tag{2.8}$$

Over the course of time, the system evolves and the trajectory of $f(\boldsymbol{p}(t), \boldsymbol{q}(t))$ meanders through phase space. From the shape of this trajectory or path we can derive information on the system's behaviour and the nature of how it explores the phase space. Note, in eq. (2.8) the time-average of the observable still depends on the initial conditions, which is in general true. The integral in eq. (2.8) measures the duration the system spends at specific values of the observable. If the measurement time approaches infinity and the system domain is bounded, we expect the time average of suitably chosen observables to arrive and eventually stay at some value, that is associated with a position in phase space, which is called the thermodynamic equilibrium. Thermodynamic equilibrium is understood here in the sense of a static state $(\boldsymbol{p}, \boldsymbol{q})$, *i.e.* unchanging and time-independent, which is also the simplest appearance of an equilibrium,

$$\overline{f} := \lim_{T \to \infty} \overline{f}_T\Big(\boldsymbol{p}(0), \boldsymbol{q}(0)\Big) \,, \tag{2.9}$$

In eq. (2.9) the time-average observable on the left-hand side (LHS) is not a random variable, but a constant (or a degenerate random variable whose distribution function is the DIRAC delta function). However, more realistic and hence of greater interest is the dynamic case of

[56] The phase space can even be a fractal.

an equilibrium. In a dynamic equilibrium the time-average observable on the LHS in eq. (2.9) is a random variable with a nondegenerate invariant limit distribution. This limit (probability) measure $\mu(\boldsymbol{p}, \boldsymbol{q})$ gives the frequencies towards which the durations spent in specific regions in phase space converge. The existence of such a limit measure was shown by BIRKHOFF (1927). So far, the time-average behaviour describes the dynamics of the system, now we link it to the statistics.

2.2.2 Ensemble Average

The neighbourhood of some state $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ in phase space forms a subvolume of the manifold. The size of this subvolume is measured by multiplying the length of the edges of the hypercube using $\Delta\boldsymbol{p}\Delta\boldsymbol{q}$. In order to derive the probability to find a system in the neighbourhood of some state we need to perform two steps. First, we need a time-independent (limit probability) measure denoted by $\mu(\boldsymbol{p}^*, \boldsymbol{q}^*)$, which gives a *reliable* number of microconfigurations in the specified neighbourhood in phase space. Second, we go to the infinitesimal limit and use the differential operators $\mathrm{d}p_i \cdot \mathrm{d}q_i$ to measure the volume observed at any instance of time. In brief, this probability is a statement over the mentally constructed ensemble of possible systems, it simply gives the ratio of the number of systems whose microconfigurations are in that neighbourhood divided by all systems in the ensemble. This specific version of the probability is commonly referred to as the micro-canonical ensemble probability distribution in statistical mechanics. Hence, the probability to find a system in the neighbourhood of state $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ is:

$$\left(\prod_{i=1}^{N} \mathrm{d}p_i \cdot \mathrm{d}q_i\right) \mu\left(\boldsymbol{p}^*, \boldsymbol{q}^*\right) \quad . \tag{2.10}$$

The ensemble average of the observable, $\langle f \rangle_\mu$, with respect to the measure μ is given by the product of the probability introduced in eq. (2.10) for an arbitrary state and the value of the observable in this states,

$$\langle f \rangle_\mu := \underbrace{\left(\prod_{i=1}^{N} \int \mathrm{d}p_i \, \mathrm{d}q_i\right)}_{\text{specific way of summation}} \quad \underbrace{\mu(\boldsymbol{p}, \boldsymbol{q})}_{\text{probability}} \quad \underbrace{f(\boldsymbol{p}, \boldsymbol{q})}_{\text{value}} \quad . \tag{2.11}$$

As eq. (2.10) and eq. (2.11) show, an ensemble is completely specified by a (limit) probability measure, no statement on dynamics is needed here. Furthermore and as indicated by the annotations, eq. (2.11) can be correctly interpreted as the expectation value of the observable, it is the frequency-weighted value of an observable. By the BIRKHOFF-KHINCHIN theorem[57]

[57] A detailed analysis of the theorem (BIRKHOFF 1931b; KHINTCHINE 1933; KHINCHIN 1949) follows in Sec. 2.4.6.

the equality of the time-average observable in eq. (2.8) and the ensemble-average observable in eq. (2.11) – or in short the ergodicity of the observable – is established via the proof that there exists a limit measure if the measurement time tends to infinity for which ensemble and time average of the observable coincide for almost all initial conditions $(\boldsymbol{p}(0), \boldsymbol{q}(0))$ except those of measure zero. Therefore, the dependence on the initial conditions of the time average that has been present in eq. (2.8) drops out and we arrive at the following equality:

$$\lim_{T\to\infty} \overline{f}_T(\boldsymbol{p}, \boldsymbol{q}) = \langle f \rangle_\mu \qquad \forall \; (\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{R}^{6N} \; . \tag{2.12}$$

The validity of eq. (2.12) justifies the assumption of ergodicity in statistical mechanics or more precisely it justifies the crucial substitution of the time averages by the ensemble average (or vice versa). Furthermore, the ergodic hypothesis entails the statement that the time a system spends in a certain state or region r – also called sojourn time – is proportional to the measure of this region in phase space $\mu(r)/\mu(S)$ with regard to the invariant ensemble measure. This implies that a single trajectory already covers the whole energy surface. The unreasonableness of this assumption was quickly realised and replaced a weakened assumption of the quasi-ergodic hypothesis, which is untenable, too. It states that a single trajectory does not fill the surface completely, but only covers it everywhere densely. From the times of Maxwell and Boltzmann physics still struggles to justify what exactly it is if not the ergodic hypothesis that makes statistical mechanics work. From the fact that an invariant probability distribution over the ensemble completely predetermines the behaviour of almost all system evolutions ensues a philosophical discussion on the nature of time and the role of purpose, to which we resume in chapter 6.

We conclude this section with the remark that ergodicity describes an apriori unnatural situation, whose absurdity is obvious if we reformulate in a medical context. If a physician reports the life expectancy at birth of a given cohort is *e.g.* 75 years, then he does not think for a second that life of almost everyone from this cohort really ends at the age of 75. Another illustrative example of two different conceptualisations of (death) risk that do not coincide and where the number of 75 may be completely useless to an individual other than to cheer him up. The existence of an unique invariant (probability) measure for certain transformations is a field of study on its own in mathematics and called ergodic theory, which operates rather independently of the special case studied in physics, when $(\boldsymbol{p}, \boldsymbol{q})$ belong to a Hamiltonian of a physical system. In what follows we continue with the historical origin of the ergodic hypothesis, which can be read as a dialog between Maxwell, Boltzmann and others on the sound microfoundations of thermodynamics.

2.3 Historical Origin of the Ergodicity Problem and the Ergodic Hypothesis

> Good mathematicians work on good problems and good problems have histories. [. . .] [I]t is often the history that makes it a good problem, because it establishes the connections with the solutions of other problems and with applications both inside and outside mathematics.[58]
>
> HAROLD M. EDWARDS

In this section[59] we reconstruct the genesis of the ergodic hypothesis as it circulated amongst the physics community during the late 19th and early 20th century. The reconstruction serves the purpose to provide the reader with the original intent in the minds of the founders of statistical mechanics, the conception of a virtual ensemble, and to make the reader familiar with the nativeness with which the time perspective comes up in other fields than economics. Based on this section's portrayal, we will apply the time perspective in Sec. 4.5 to decision making under uncertainty. This involves the replacement of the embedding of randomness in the ensemble with the embedding of the randomness within time. Along the way, unexpected similarities appear between problems of optimal transport of heat in physics and problems of optimal allocation of resources in economics.

The Austrian physicist LUDWIG BOLTZMANN contributed crucially to the foundations of statistical physics during the second half of the 19th century. Thereby he impelled previous work by RUDOLF CLAUSIUS[60] and JAMES C. MAXWELL.[61] Their common goal was to provide a dynamical basis for statistical mechanics, *i.e.* a theory that derives the known properties of physical systems at a macroscopic level (such as pressure, number of particles, volume or temperature) from the microscopic interactions between individual particles of the gas. The macro level has been successfully described already by the phenomenological approach of thermodynamics. It was soon 'recognized that the replacement of mean values in time by mean values in phase-space is an indispensable operation'[62] and poses the ergodicity problem. This substitution is justified if and only if the ergodic hypothesis is true, which appears in MAXWELL's writings as *The Principle of Continuity of Path*.

58 EDWARDS 1981, p. 108.
59 The symbols and notation in this section are rather self-contained. To avoid confusion we have not included them in the glossary, but every symbol is explained once we introduce it for the first time.
60 CLAUSIUS 1857.
61 MAXWELL 1860b,a.
62 BIRKHOFF and KOOPMAN 1932, p. 279.

2.3.1 The Search for Sound Microfoundations in Physics and Economics

The physics community around the turn of the century was striving for something that is quite familiar to economists – a rational microscopic foundation of macroscopic physics. In physics terminology, *coarse-graining* denotes the process to get from a description of the microscopic level to a description of the macroscopic level. Thus, physics was striving for a rational method of coarse-graining and macroeconomics appears as a method of coarse-graining from detailed descriptions of the microeconomic operations. Having the economics research programme of a sound microfoundation of macroeconomics in the back of one's mind, it is instructive to read again not only the title of GIBBS famous textbook *Elementary Principles in Statistical Mechanics*, but especially its subtitle which reads *Developed with Especial Reference to the Rational Foundation of Thermodynamics.* The wording *rational foundations* is interesting. First, it simply means a *scientific* explanation of thermodynamics which is in principle amenable to be replicated, understood by reason and tested by experiment. It is important to note, that *rational* has no metaphysical meaning. Instead it is the exact opposite, a scientific explanation without reference to metaphysical or divine and (so far) unobservable powers like the ether, destiny or spooky action at a distance.[63] Second, rationality is not necessarily a property of the system components in the theory[64], but a property of the way the argument of the theory is laid down.

Let us compare the conception of rationality to how it is treated in economics. Economic rationality is often given a behavioural spin and defined as a property of the behaviour of economic agents, be it people, firms, households or countries. This odd misconception of economic man has led to a distorted representation of him as an omniscient cold-blooded automaton.[65] This seems to be a mistake of category in attaching (especially strong forms of) rationality that is used within a (mathematical) model of the world to the behaviour of real-world agents. Instead of attaching rationality to the general line of argumentation. The introduction of a behavioural component can partly be saved on descriptive grounds, but does not help in solving normative problems *e.g.* of decision making. At this point, we refer to Sec. 7.1, whose focus starts from the discussion of rationality in decision making and where we elaborate in greater depth on this topic. We resume now with the reconstruction of the emergence of the ergodic hypothesis.

BOLTZMANN started statistical mechanics with his kinetic gas theory,[66] that is the study of

63 See also Sec. 6.2 for a discussion which goes into greater detail of this particular issue. The meaning of rational here is precisely the one that looms when talking to today's physicists that favour mechanistic explanations of a real-world phenomena.

64 In physics these are the particles or molecules.

65 Some label economics with this conception of economic man more as a study of the interactions between automata on abstract markets, although in their understanding economics should deal with real markets on which human fallible actors bargain, see KIRMAN (2006) and MIROWSKI (2007).

66 BOLTZMANN 1898b, 1964.

collisions between material particles. Statistical mechanics turned out to become this longed for microfoundation of macrophysics. Since then statistical mechanics has been an immensely powerful tool that not only spurred strongly the development of physics and mathematics, but also motivated analogue efforts in economics with the above mentioned research programme of finding a microfoundation for macroeconomics. See for example the work of IRVING FISHER, who started his career in economics with a PhD under the supervision of WILLARD J. GIBBS.[67] In 1867 MAXWELL derived the velocity distribution $f(x, v)$ of molecules in an ideal mechanical system in thermodynamic equilibrium, whereby x denotes the position and v the velocity,

$$f(x, v) = M_{\rho,u,T}(v) = \rho(x)\frac{\exp\left(-\frac{|v-u(x)|^2}{2T(x)}\right)}{\left(2\pi T(x)\right)^{3/2}} , \tag{2.13}$$

with three parameters defining the field environment for the hydrodynamic limit with scalar density ρ, scalar temperature T and the vector of the velocity u, that all depend on the position x. Due to MAXWELL's findings, the equilibrium distribution is named after him – MAXWELL distribution – and sometimes even denoted and referred to as the MAXWELLIAN M.[68] BOLTZMANN advanced the study of gas systems. On the one hand, the articles BOLTZMANN (1868, 1872) contain a derivation of the uniqueness of the MAXWELL distribution (eq. (2.13)) using a kinetic or mechanistic approach. The kinetic approach is the non-statistical mechanics approach, in which every single particle is tracked and the impact of collisions calculated explicitly. On the other hand, in BOLTZMANN (1877) he presented a derivation of the uniqueness of the MAXWELL distribution (eq. 2.13) using, quite remarkably, combinatorial rather than a kinetic techniques. Hereby, the many-particle limit allowed the application of limit theorems and other tools from probability theory. This is one of the first attempts to approach gas systems statistically. The combinatorial approach of statistical mechanics uses averaging procedures such as how many collision happen on average per unit time, what are the average collision-free path lengths between two consecutive collisions also called mean free paths and so on. The equilibrium state in its statistical mechanics conception is defined as that macroscopic state, which combines the largest number of distinct microscopic configurations of the individual particles. Then the equilibrium state is the most likely state if all microconfigurations are equally probable.[69] By the use of these by then new techniques BOLTZMANN (1877) also gave a probabilistic interpretation to the second law of thermodynamics, which states the impossibility of decreasing entropy for isolated thermodynamic systems. The famous

67 FISHER (1892, 1906). In chapter 6 we elaborate on the advent of mathematical economics, IRVING FISHER and others contributed to.

68 MAXWELL 1867.

69 A central assumption in MAXWELL's work and statistical mechanics is that of equipartition, *i.e.* the apriori equiprobability of every microscopic configuration. The probability to find a given system in a specific microscopic configuration is given by the so called partition function, which is a distribution over the microconfigurations in the ensemble. In equilibrium this distribution is an exponential distribution.

Boltzmann equation (2.14), an integro-differential equation, describes the time evolution of an ideal gas and its convergence in probability towards the Maxwell distribution, and is given by

$$\frac{\partial f(\boldsymbol{v}_1,t)}{\partial t} = \int_0^\infty \int_0^{\boldsymbol{v}_1+\boldsymbol{v}_2} \mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{v}_1'\psi(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_1')\left[f(\boldsymbol{v}_1',t)f(\boldsymbol{v}_2',t) - f(\boldsymbol{v}_1,t)f(\boldsymbol{v}_2,t)\right], \tag{2.14}$$

whereby $\boldsymbol{v}$ denotes the particle velocity before and $\boldsymbol{v}'$ after a collision took place between two particles, denoted by the subscripts 1 and 2, $f(\boldsymbol{v}_1,t)\,\mathrm{d}\boldsymbol{v}_1$ represents the probability density also known as the distribution function of the gas, that gives the number of particles within the velocity cell $(\boldsymbol{v}_1,\boldsymbol{v}_1+\mathrm{d}\boldsymbol{v}_1)$ at time t, and ψ is the differential cross section. The collision kernel in square brackets vanishes in equilibrium in the Maxwellian (eq. 2.13).[70] If function f is normalised appropriately it becomes a probability density function, hence, eq. (2.14) describes the time evolution of a probability distribution function (PDF). The convergence of the Boltzmann equation (eq. 2.14) towards the Maxwell equation (eq. 2.13) is the reason that this equilibrium distribution is called the Maxwell-Boltzmann distribution.[71]

However, it is important to note that the Boltzmann equation (eq. 2.14) describes a non-equilibrium situation and is a cornerstone of non-equilibrium physics,[72] which contains the equilibrium state as a special case. Similar approaches that describe the evolution of distribution functions are the Fokker-Planck equation[73] or the master equation approach.[74]

Depending on the particular interest, the Boltzmann equation is sometimes given in a different form, that makes it easier to see the components with regard to transport problems,

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\cdot\nabla_x f = Q(f,f)\ , \tag{2.15}$$

whereby $\boldsymbol{v}\cdot\nabla_x f$ represents the transport operator along a gradient at position $x \in \mathbb{R}^d$ and $Q(f,f)$ is the bilinear collision operator that represents the probability of the collisions between exactly two particles in a specified volume (which is basically the part in square brackets of the right-hand side (RHS) in eq. (2.14)). In essence, eq. (2.15) equates the transport with the collisions.[75] Typical transport of (probability) mass problems in physics deal with collisions between particles that lead to changing densities in the system. If the density is constant, the

70 Badino 2011, eq. (9).
71 The Boltzmann-Maxwell distribution gives the probability to find a molecule with a certain velocity in an ideal gas in equilibrium. Whereas the Boltzmann-Gibbs distribution gives the probability to find a system in a certain state as a function of the temperature and energy of this state.
72 Lebowitz and Montroll 1983.
73 Risken 1996.
74 Oppenheim et al. 1977; Weidlich 1991, Ch.8.
75 Gallagher et al. 2014.

gas has reached an equilibrium. In such states there can still be a transfer of energy due to the collisions which constitute heat conduction and give rise to viscosity.[76] Optimal transport problems have found many applications in economics, mostly under terms of optimal allocation of resources or optimal planning studied by KANTOROVICH.[77] One important assumption in eq. (2.15) on optimal transport is that the gas is sufficiently dilute. This ensures on the one hand that the collisions can be thought of as to take place only between two particles at a time and on the other hand justifies that the particle behaviour can be understood as sufficiently random.

Stoßzahlansatz or Molecular Chaos

These assumptions are part of BOLTZMANN's *Stoßzahlansatz* or *molecular chaos*, a particular method used in coarse-graining. The *Stoßzahlansatz* enters the time evolution of the velocities, $\frac{\partial f(v_1,t)}{\partial t}$ (eq. 2.13), and entails the specific way of the counting of the molecules, the set up of collisions between exactly two particles at certain differential cross-sections and the equipartition assumption, which collectively justify the independence of the sufficiently irregular and thus random behaviour of the particles, thus sometimes called molecular chaos. The assumption of molecular chaos emphasises the sufficient lack of correlation between the collisions of the particles which ensure the irregularity of their movement. Collectively this set of assumptions is a necessary precondition in order to apply limit theorems from probability theory to draws of random variables which are stochastically independent.

Molecular chaos has a different meaning than the *chaos* in chaos theory or deterministic chaos. The latter two or synonyms and characterise those phenomena with sensitive dependence of initial conditions. Chaos theory started by some foundational papers in the 1960s,[78] made big advances with the advent (of the visualisation power) of digital computers in the 1980s.[79] Some of the fundamental results of chaos theory use techniques like recurrence maps or homoclinic orbits that were developed around the time of the ergodic hypothesis.[80]

A mathematical proof of the unique convergence of the time evolution of the BOLTZMANN equation to the MAXWELL distribution rests on certain assumptions, the essential one is

[76] Due to recent progress the field of optimal transport is an active research topic in mathematics and physics (see *e.g.* VILLANI (2009)). The range of applicability of the BOLTZMANN equation (and other kinetic theories) to transport problems scales from the 'radiation transport in stars to metabolic transport across living cell membranes' (LEBOWITZ and MONTROLL 1983, p. v).

[77] For the MONGE-KANTOROVICH transportation problem see KANTOROVICH (1939, 1942). Recently, the KANTOROVICH metric became a powerful tool as a distance measure for probability distributions. It is sometimes referred to as the VASERSHTEIN or WASSERSTEIN metric (RÜSCHENDORF 1998; VERSHIK 2006, 2013).

[78] LORENZ 1963; LI and YORKE 1975.

[79] FEIGENBAUM 1978; MANDELBROT 1987.

[80] POINCARÉ 1890, 1899; CARTWRIGHT and LITTLEWOOD 1945.

the ergodic hypothesis in form of the H-theorem[81]. The H-theorem hinges on his so-called *Stoßzahlansatz*, some asymptotic conditions and on Liouville's theorem on the conservation of phase space volume over the time evolution.[82]

Boltzmann's *H*-Theorem

Boltzmann's H-theorem is a special functional of f given by

$$H(f,t) = \int \underbrace{f(\boldsymbol{v},t)}_{\text{probability of state}} \underbrace{\log f(\boldsymbol{v},t)\,\mathrm{d}\boldsymbol{v}}_{\text{quantity of state}} \;. \tag{2.16}$$

The eq. (2.16) was a further step towards an understanding of the role of entropy through its emergence in statistical mechanics, because it relates the frequency of a state of a gas system with the logarithm of the frequency of that state.

Furthermore, Boltzmann could show that the H-function does not change if f is the Maxwell distribution given in eq. (2.13). In other words, the time derivative of the H-function (eq. 2.16) is zero in the Maxwellian equilibrium case, and is strictly decreasing otherwise, thus

$$\frac{\mathrm{d}H}{\mathrm{d}t} \leq 0 \;. \tag{2.17}$$

The H-function eq. (2.16) must monotonically decrease to a minimum value until the Maxwell-Boltzmannequilibrium distribution is reached for which the equality in its time derivative holds. If the ergodic hypothesis in Boltzmann's kinetic approach were true it consists of two different notions of probability that could then be used interchangeably as already mentioned in Sec. 2.1. Most often, the probability is defined 'as the fraction

[81] Following Boltzmann (1872, pp. 335, starting with equation (17)) the correct term would be the 'E-theorem', but subsequent authors established the term 'H-theorem ' (*e.g.* Burbury 1894; Ehrenfest and Ehrenfest 1907, 1911). In later writings Boltzmann himself obeyed this convention (*e.g.* in Boltzmann 1895c, 1898a), see esp. the references in footnote 95 on the Nature debate. Furthermore, there was a discussion on whether the switch in notation from the E-theorem to the H-theorem, involves actually the Greek capital letter 'Eta'. Unfortunately, capital 'Eta' is indistinguishable from the Latin capital 'H' (Chapman 1937). As regards content, the 'Eta' would be closer to energy, which already occupies the symbol E, or entropy which is commonly denoted by S or *e.g.* in information theory by H (*e.g.* Shannon 1948, p. 393). However, until 30 years later Brush (1967a) founds no evidence for the 'Eta' claim.

[82] See Khinchin (1949, chap. II.4) for a detailed explanation and a proof of Liouville's theorem, which is of no further interest for our purposes at this point. The usual assumptions in gas theory are all of asymptotic nature and subsumed under the term of the thermodynamic limit:

- sufficiently many molecules ($N \to \infty$),
- sufficiently long observation time ($T \to \infty$),
- sufficiently increasing volumes V such that the ratio of $N/V < \infty$, which ensures the sufficient dilution of the gas and thus ensures the exclusive study of collisions of only two particles at a time.

of time that a particle has velocity v, but occasionally he defined the same probability as the fraction of the total number of particles having that velocity.'[83] The first notion is the time-average notion of probability and the latter is closer to an ensemble-average notion of probability.

Additionally, one must assume that the particles pass through all accessible states in phase space. The system trajectory which meets this requirement is said to cover the phase space densely. If the trajectories of particles would not pass through every point in phase space – *e.g.* because large enough energy barriers lead to a compartmentalised phase space – that in turn would lead to broken ergodicity of the system. This would make it much harder to infer valid statements on the uniqueness of the equilibrium state of the macroscopic system if only a proper subset of phase space is actually explored. An often made assumption in measuring the time average of observables is that the measured behaviour is ergodic in at least a subcomponent and the subcomponent could become arbitrary large if the measurement duration would be increased.[84] Due to some exceptional initial conditions the system could be stuck in a proper subset of the phase space where the system exhibits some very special properties such as extreme temperature or exceptionally ordered microscopic configurations. Such strongly ordered states would then possess extraordinary low entropy. Thus '[b]roken ergodicity requires a modification of conventional statistical mechanics, which normally computes averages over the whole phase space'.[85] Broken ergodicity is discussed in some detail later in Subsec. 5.3.1.2 with regard to new developments in non-extensive statistical mechanics and in Sec. 6.4.2 in terms of the ergodic fallacy.

Furthermore, we can see that eq. (2.16) already contains the famous structure of entropy, S. In fact eq. (2.16) gives the negative entropy, also called negentropy. It is due to the brilliant insights in Shannon (1948) that we now know of the much more general character of the entropy concept as a measure of information or surprise that extends far beyond the notion of disorder in thermodynamics and applies not only to atoms and physics. As can be seen from eq. (2.18), it depends only on a probability distribution and is applicable whenever a probability can be defined. Shannon (1948) and Khinchin (1957) derived from an axiomatic foundation the uniqueness of the logarithm as the proper functional form of f in the general structure in $\int p \cdot f(p)\,\mathrm{d}p$, which finally yields the Shannon entropy of the distribution p,

$$S(p) = -H(p) = -\int p \cdot \log(p)\,\mathrm{d}p \;, \tag{2.18}$$

[83] Badino 2011, p. 358.

[84] Sometimes non-ergodic behaviour of a system can be transformed into the analysis of a compartmentalised phase space that gives rise to ergodic behaviour within subcomponents each characterised by an ergodic measure on this subcomponent. This technique is called ergodic decomposition.

[85] Palmer 1989, p. 276.

if we keep in mind that the function f can easily be normalised and thus truly becomes a probability p.[86]

This concludes the reconstruction of the historical situation from which the ergodic hypothesis grew out of as Boltzmann's brainchild. We now turn towards the etymological origin of the exact term 'ergodic'.

2.3.2 Etymology of Ergodic Theory

This section explores the etymological background of the term *ergodic theory* if it is meant to be derived from the *ergodic hypothesis*. For this purpose it is helpful to stay close to the exact original formulations. The section therefore contains many direct quotations, the majority of them given in the original German to prevent us from getting lost in translation. Throughout the section, we have added our annotations in square brackets whenever rephrasing something in modern terminology eases an understanding.

Before we present two narratives about the emergence of the term *ergodic theory*, we make some remarks regarding problems of a proper translation of the original German terms into other languages. In the early writings on the topic, Boltzmann speaks of a noun called in singular *Ergode* and in the plural *Ergoden*.[87] This leads directly to the term *das Ergodenproblem*[88], from which the *Ergodenhypothese* and *Ergodentheorie* are derived. As we will discuss in a moment in more detail, the original German word was borrowed from Greek and seems as such hard to translate literally. However, the translation of these words into other languages is often done using an adjective instead of a noun, *e.g.* in *ergodic hypothesis*, *ergodic theory* or *ergodic problem*. Such translations bear a one-sided connotation that is misleading for the simple reason that a problem, a hypothesis or a theory by themselves can neither be ergodic nor non-ergodic. In all cases the usage of the noun *ergodicity* would seem more appropriate. In Russian texts the situation is similar. *E.g.* Корнфелд et al. (1980) which is the original of Cornfeld et al. (1982) uses an adjective in the title as well, Эргодическая теория which reads *ergodic theory*, but it contains as well formulations like критерии зргодичности (ch. 14.2) for *criteria for ergodicity* or проблема зргодичности which translates into our favoured *ergodicity problem*. Thus, throughout the thesis we only speak of the *ergodicity problem* and deviate from the English translation in Khinchin (1949) where it reads the *ergodic problem*. The other terms, ergodic theory and the ergodic hypothesis, are established to such a degree that it is not worth introducing a different notation here.

86 Villani 2008; Bais and Farmer 2008.
87 Boltzmann 1884; Ehrenfest and Ehrenfest 1911.
88 For example in Khintchine (1933) and also in the German translation Chintschin (1964) of Khinchin (1949).

The naming of *ergodicity economics* is carefully considered and comparable to *e.g.* complexity economics. Complexity economics is a new complementary approach to economics that harnesses methods used in (the admittedly hard to define field of) complexity science to tackle economic problems. Some authors even go so far to say that complexity economics as a theory covering also non-equilibrium situations is a generalisation of traditional economics, because it contains *e.g.* general equilibrium theory as a special case.[89] Surely it is true that ergodicity economics contains the case of ergodic observables in economics as a special case. It is not called *complex economics*, because this would give the unintended connotation of being a version of economics but just more complex or a version of economics which only deals with the subset of those economic problems that are complex. If this were the case every student would clearly prefer a subject called *trivial economics*. In a similar way the name *ergodic economics* would convey the impression of an economic theory that only covers the ergodic cases. Exactly the opposite is aimed for. *Ergodicity Economics* is a way of tackling the ergodicity problem in economic puzzles head-on and proactively study the – as it turns out very likely – possibility of broken ergodicity.

Now we continue with the presentation of the two narratives about the emergence of the term *ergodic theory*.

Boltzmann (1871, p. 284) states the basic assumption of his approach:

> "Von den zuletzt entwickelten Gleichungen können wir unter einer Hypothese, deren Anwendbarkeit auf warme Körper mir nicht unwahrscheinlich scheint, direkt zum Wärmegleichgewicht mehratomiger Gasmoleküle, ja noch allgemeiner zum Wärmegleichgewicht eines beliebigen mit einer Gasmasse in Berührung stehenden Körpers gelangen. Die große Unregelmäßigkeit der Wärmebewegung und die Mannigfaltigkeit der Kräfte, welche von außen auf die Körper wirken, macht es wahrscheinlich, daß die Atome derselben vermöge der Bewegung, die wir Wärme nennen, alle möglichen mit der Gleichung der lebendigen Kraft vereinbaren Positionen und Geschwindigkeiten durchlaufen, daß wir also die zuletzt entwickelten Gleichungen auf die Koordinaten und Geschwindigkeitskomponenten der Atome warmer Körper anwenden können."

Two statements are important in this quote. First, he remarks the vast irregularity of the motion due to the variety of forces acting upon the particles. And second, this justifies the assumption that the particles attain every position and velocity that is compatible with the energy of the system ('mit der Gleichung der lebendigen Kraft vereinbaren Positionen und Geschwindigkeiten'), *i.e.* the particles pass through every point in phase space on the accessible energy shell. While the first is the justification of the general probabilistic

[89] Arthur 2013, 2015.

approach to thermodynamics of gases also known as the molecular chaos hypothesis.[90] The latter assumption has later been called the ergodic hypothesis. Similar statements about the assumption of the ergodic hypothesis can be found in the works of Maxwell on mechanical systems, *e.g.*

> "The only assumption which is necessary for the direct proof is that the system, if left to itself in its actual state of motion, will, sooner or later, pass through every phase which is consistent with the equation of energy."[91]

Boltzmann himself called the introduction of the ergodic hypothesis simply a mathematical artifice – in the German original literally a 'Kunstgriff' – which is worth citing here due to some confusion about this fact in the literature:

> "Das bisher Vorgenommene ist nun allerdings nichts weiter als, ein mathematischer Kunstgriff, um einen Satz strenge zu beweisen, dessen exakter Beweis bisher nicht gelungen ist. Es gewinnt aber sehr an Bedeutung durch seine Anwendbarkeit auf die Theorie mehratomiger Gasmoleküle."[92]

He expounded the meaning of the trick several times in later rejoinders to critiques such as in Boltzmann (1884, p. 123):

> "Die Bewegung ist hier offenbar keine monozyklische; wir können sie aber mittels eines Kunstgriffes, den ich zuerst [. . . in Boltzmann (1871) . . .] anwandte, und welchen Maxwell weiter verfolgt hat, in eine monozyklische verwandeln"

and similar in a longer passage from Boltzmann (1887, pp. 263–265, German annotations in square brackets):

> "Der präzise mathematische Ausdruck gerade dieses Umstandes [der Gültigkeit der Ergodenhypothese] bietet aber Schwierigkeiten, und dürfte am besten mittels folgenden Kunstgriffes geschehen [. . .] [264]
> Statt eines einzigen Systems fingieren wir unendlich viele, vollkommen gleich beschaffene Systeme [Ensemble], in deren jedem auch gleich viel Energie enthalten ist, die aber im übrigen alle möglichen Anfangszustände besitzen. Alle sollen dieselbe Energievermehrung erfahren, und für alle sollen sich auch die äußeren Bedingungen in derselben Weise verändern. Alle Eigenschaften, welche von den zufälligen Anfangsbedingungen unabhängig sind, müssen nun auch diesem Inbegriffe von Systemen [Ensemble] in gleicher Weise zukommen. [. . .] Wenn aber die Werte dieser Größen vom Anfangszustande nicht in merklicher Weise abhängen, so wird dieser Mittelwert [Ensemblemittel] gleich sein müssen dem Werte derselben

[90] For an explanation of molecular chaos see section 2.3.1 on page 39.
[91] Maxwell 1890, p. 714, originally published just before he died in 1879.
[92] Boltzmann 1872, p. 345.

> Größe für jedes einzelne System [individuelle Zeitmittel]. Es ist also durchaus nicht notwendig, daß wir die Werte dieser Größen [Zeitmittel] für jedes einzelne bestimmten Anfangsbedingungen unterworfene System berechnen, sondern es genügt, ihre Mittelwerte für den ganzen Inbegriff von Systemen [Ensemble] zu berechnen. [Substituierbarkeit von Zeitmittel durch Ensemblemittel] [...] Bei passender Wahl dieser Funktion können wir es nun ermöglichen, daß für Inbegriff von Systemen [Ensemble] Gleichungen von derselben Einfachheit gelten, wie für ein System, das für sich allein alle möglichen, mit der Gleichung der lebendigen Kraft [Energieschale] vereinbaren Zustände durchläuft. Da wir nun [die Gültigkeit der Ergodenhypothese] nachgewiesen haben, daß die Werte derjenigen Größen, welche von den Anfangsbedingungen unabhängig sind, für jedes einzelne System [Zeitmittel] gleich dem Mittelwerte [Ensemblemittel] derselben Größen für einen beliebig gebildeten Inbegriff von Systemen [Ensemble] sind, so genügt es, die Mittelwerte derartiger Größen für denjenigen Inbegriff zu bestimmen, für welchen die Rechnung [265] möglichst einfach wird. Wir wollen nun zunächst eine derartige Wahl treffen.
> Wir denken uns also nicht ein einziges, sondern unendlich viele (N) gleich beschaffene Systeme [Ensemble] gegeben."

In the third last sentence he stresses explicitly the increased mathematical tractability of his artifice. It can be confusing, that the trick of the ergodic hypothesis was used long before it received its name. The exact term 'Ergode' or 'ergodisch' was used for the first time in Boltzmann (1884, p. 132, p. 140, respectively) and several times after that, *e.g.* in Boltzmann (1887, p. 226): 'Ich habe für derartige Inbegriffe von Systemen den Namen Ergoden vorgeschlagen'. In the same article Boltzmann defined a whole zoo of other terms sharing the same suffix of *-ode* indicating specific trajectories of ensembles through phase space to which we turn in a moment. Similar to 'Inbegriff von Systemen' the term *Holode* could nowadays be understood in a Gibbsian sense as an *ensemble*: 'german'[93]Wir denken uns nun sehr viele N genau gleich beschaffene derartige Systeme vorhanden; jedes System von jedem anderen völlig unabhängig. [...] [p. 132] Der Inbegriff aller dieser Systeme bildet eine Monode im früher definierten Sinne [...] und ich will die so definierte Gattung von Monoden mit dem Namen Holoden bezeichnen. Jedes der Systeme nenne ich ein Element der Holode.

From these facts diverge two narratives on the etymology and origin of the notion of the *ergodic hypothesis* and *ergodic theory*. These two narrative are the one by the Ehrenfests and the one which claims *isodic theory* to be the correct name if ergodic theory is to be derived from the ergodic hypothesis. Both will be discussed in the following.

93 Boltzmann 1884, pp. 131–132.

2.3.2.1 The Ehrenfests Narrative

Boltzmann's initial publications started a big debate among the scientific community that continued after his untimely death. The discussion had several focal points, to name just the most important ones:

- the so called Stoßzahlansatz, a collection of several assumptions necessary for the probabilistic treatment (see also the box on p. 39), and
- objections to the so called H-theorem and the ergodic hypothesis.

These issues are intricately entangled with each other. The careful disentanglement follows.

From the first law of thermodynamics for closed systems follows the conservation of energy. Thus, basic thermodynamics starts to study systems isolated from their environment. It is important to note, that the system is analysed by an external observer, for example a closed box of gas molecules that does not interact with the scientist, who studies it.[94] Additionally, the assumption of no dependence of the system state to its initial conditions would simplify the mathematical tractability of the state of such a system tremendously, because many of the forces that could potentially influence the state of the system are thus assumed away or assumed to cancel each other out in the aggregate. Hence, the final state, where the system comes to rest (also called its thermodynamic equilibrium or state of maximum entropy) is solely defined by the (internal and) in principle observable properties of the particles like their velocities and positions in phase space.

However, if the gas system consists of a vast number of molecules their experimental tracking becomes increasingly difficult, thus the method of kinetic gas theory becomes virtually impossible to apply in an experimental setting. Assuming the system is ergodic right away simplified the mathematical tractability of the problem, but triggered a big debate on the famous H-theorem among the science community culminating in the statement by Culverwell (1894a): 'Will some one say exactly what the H-theorem proves?' A back and forth of letters was published in Nature and attracted a lot of attention, which is why the debate on the H-theorem is also called the Nature debate.[95]

It was only when a much-noticed article appeared in 1911 jointly written by one of Boltzmann's students Paul Ehrenfest and his wife Tatyana Ehrenfest-Afanassjewa who was an

[94] This assumption has been completely reformulated in the context of quantum mechanics where the measurement procedure is a crucial part of the mathematical apparatus. See von Neumann (1932a, ch. VI).

[95] See Burbury (1894, 1895a,b), Culverwell (1894b, 1895), Watson and Culverwell (1894) and Bryan (1895) and the rejoinders in Boltzmann (1895a,c,b). See also Brown et al. (2009) and the references therein.

acknowledged mathematician herself, that brought some clarification into the discussion, esp. after BOLTZMANN's untimely death in 1906. More importantly, the article established the term of the '(quasi-)ergodic hypothesis', this specific term had not been used by BOLTZMANN (who had only used 'Ergode'), but should occupy the science community hereafter.

Since its first exploitation, the ergodic hypothesis or the H-theorem was much disputed and of hypothetical character, because the ergodicity of a gas system was and still is not amenable to an experimental verification.

The EHRENFESTS emphasised the hypothetical character of the assumption of ergodicity most explicitly, *e.g.* in the central chapters 10 and 11. See the following quote from the end of chapter 10:

> "Es empfiehlt sich, die speziellen Dichtenverteilungen [...] als *ergodische* zu bezeichnen, um daran zu erinnern, dass bis jetzt keinesfalls ein anderer Rechtfertigungsversuch für ihre Wahl vorliegt als die Berufung auf die Ergodenhypothese."[96]

And shortly afterwards, in chapter 11 they noted: 'Die fundamentale Voraussetzung dieser Untersuchung ist die Hypothese, dass die Gasmodelle ergodische Systeme sind [...] Mit ihrer Hilfe berechnete Boltzmann [diverse] Zeitmittel [...]'.[97] And to quote a longer central passage in the exposition of BOLTZMANN's approach they stated:

> "Die hier skizzierte Art der Berechnung der *Zeitmittel* [time average] [folgt eher Maxwell, Lord Rayleigh und Jeans] [...] doch weicht nur *formal* von der ursprünglichen Darstellung *Boltzmanns* ab. Diese entwickelt aus der Fiktion der ergodisch verteilten Systemschaar und unter Benutzung der Hypothese, dass das Gasmodell ein ergodisches System ist [...]
>
> $$\text{(34)} \qquad \lim_{T\to\infty} \frac{\mathrm{d}t}{T} = \frac{\sigma\,\mathrm{d}S}{\int \sigma\,\mathrm{d}S} \qquad [\ldots]\,.$$
>
> **11b. Kritik und Bedeutung des Boltzmannschen Resultates.**
>
> Bezüglich der Gl. (34) [reference to the equation establishing ergodicity, which is number 34 in the original and corresponds *e.g.* to eq. (2.39) in this thesis] ist hervorzuheben:
>
> 1. Akzeptiert man die Ergodenhypothese, so beansprucht Gl. (34) ein rein mechanisches Theorem zu sein, unabhängig von irgendwelchen „Wahrscheinlichkeits"-Überlegungen.

[96] EHRENFEST and EHRENFEST 1911, p. 33.
[97] EHRENFEST and EHRENFEST 1911, p. 34.

2. Verwirft man die Ergodenhypothese ganz, oder sucht man sie in modifizierter Form festzuhalten, so fehlt zur Zeit jeder Anhaltspunkt zur Behauptung, dass die Gl. (34) zu Recht besteht oder auch nur eine irgendwie brauchbare Annäherungsformel darstellt."[98]

Despite the lack of experimental verifiability of the ergodic hypothesis noted under 2., the advantageous ramifications include its intuitive comprehensibility and its agreement with equilibrium calculations. In addition, it was noticed that the enormous success of statistical mechanics in general and BOLTZMANN's Stoßzahlansatz in the kinetic theory of gases rests on the ergodic hypothesis. The article by EHRENFEST and EHRENFEST (1911) is important because it brought forward a consistent interpretation of BOLTZMANN's ideas regarding the ergodic hypothesis and it addressed popular objections to the H-theorem, which turned out to be the key for the validity of BOLTZMANN's ansatz.[99] Primarily, EHRENFEST and EHRENFEST (1911) reacted to the reversibility objection brought forward by JOSEF LOSCHMIDT[100] and the recurrence objection by ERNST ZERMELO.[101]

Finally, in a footnote in EHRENFEST and EHRENFEST (1911, p. 30, footnote 8) a possible etymological origin is given. Following them, the coinage of the word has similar roots like 'energy' and originates from Greek έργον [*ergon*] for *work* and οδός [*odos*] in modern Greek or οδος [*hodos*] in ancient Greek for *way*, *road* or *path*. It is meant to describe the path or trajectory of a particle on a shell of conserved energy, *i.e.* its phase space. This etymology of the word 'Ergode' is considered to be wrong and the correct etymology is given in the next section. But because of the wide influence of their article published in the prestigious *Encyklopädie der mathematischen Wissenschaften* edited among others by FELIX KLEIN this became the standard explanation of the name of the mathematical field ergodic theory.

2.3.2.2 The Isodic Theory Narrative

As already quoted above the first mention of 'Ergode' appeared in BOLTZMANN (1884). Whereas the EHRENFESTS omitted BOLTZMANN (1884) in their bibliography and erroneously traced the first mention to BOLTZMANN (1887). This omission is the origin of the narrative of

[98] EHRENFEST and EHRENFEST 1911, p. 35.

[99] See EHRENFEST and EHRENFEST (1907), therein one finds the reactions to both objections to the H-theorem. In the explanation of the validity of BOLTZMANN's ansatz in the sense that is true in probability except for states of measure zero. In this explanation the EHRENFESTS argue with the help of the famous dog-flea model.

[100] Also known as the *Umkehreinwand* in LOSCHMIDT (1876a,b, 1877, 1878), an earlier critique using similar arguments was given by Lord KELVIN (THOMSON 1874).

[101] Also known as the *Wiederkehreinwand* in ZERMELO (1896a,b), who used earlier ideas from HENRI POINCARÉ's recurrence theorem. The results in POINCARÉ (1899, ch. XXVI: Stabilité á la Poisson) had been reformulated in measure-theoretic terms by CONSTANTIN CARATHÉODORY, who used sets of zero measure in the LEBESGUE sense (CARATHÉODORY 1914, 1919).

the EHRENFESTS.[102] In what follows, we will carefully evaluate BOLTZMANN (1884), which leads us to a narrative that is opposing the standard narrative which is to a large extent due to the exposition of the matter in EHRENFEST and EHRENFEST (1911).

As mentioned before, the ergodic hypothesis is associated with certain requirements concerning the nature of the movement of the molecules in phase space. The different types of movements were given distinct names by BOLTZMANN like *Ergode*, which is a special form of *Monode*. BOLTZMANN (1884) introduces the term *Monode* (p. 129-130):

> "Ich möchte mir erlauben, Systeme, deren Bewegung in diesem Sinne stationär ist, als monodische oder kürzer Monoden zu bezeichnen. Sie sollen dadurch charakterisiert sein, daß [p. 130] die in jedem Punkte derselben herrschende Bewegung unverändert fortdauert, also nicht Funktion der Zeit ist, solange die äußeren Kräfte unverändert bleiben, und daß auch in keinem Punkte und keiner Fläche derselben Masse oder lebendige Kraft oder sonst ein Agens ein- oder austritt."

MATHIEU (1988, p. 375) gives the etymology of 'Monode', which describes

> "a closed system moving only in a finite region of the phase space whose motion always, independently of time, looks the same. A simple example is a mathematical pendulum. To try to understand the notion 'Monode' we recall that μόνος [*monos*] means *unique* and εἶδος [*eidos*] stands for [*kind of, appearance, species*]."

BOLTZMANN (1884, p. 134) continues:

> "... so wollen wir den Inbegriff aller N Systeme als eine Monode bezeichnen, welche durch die Gleichungen $\varphi_1 = a_1$ beschränkt ist. Die Größen a können entweder ganz konstant sein oder zu den langsam veränderlichen Größen gehören. Die Funktionen φ werden im allgemeinen durch die Veränderlichkeit der p_a immer langsam ihre Form ändern. Jedes einzelne System heißt wieder ein Element. Monoden, welche nur durch die Gleichung der lebendigen Kraft beschränkt sind, will ich als Ergoden, solche, welche außer dieser Gleichung auch noch durch andere beschränkt sind, als Subergoden bezeichnen. [...]
> Für Ergoden existiert also nur ein φ, welches gleich der für alle Systeme gleichen und während der Bewegung jedes Systems konstanten Energie eines Systems $\chi + \psi = (\Phi + L)/N$ ist,"

[102] Independently of this thesis this has been recognised by STEPHEN G. BRUSH who is the translator of BOLTZMANN (1964), see p. 282 therein. And a longer quotation of BRUSH is given in this thesis on p. 50. See also BRUSH (1971, p. 296, endnote 11), MATHIEU (1988), GALLAVOTTI (1995, 1999, §1.9), GALLAVOTTI et al. (2004, pp. 1–2) and DERNDINGER et al. (1987, pp. 1–4).

whereby the lower case Greek symbols χ and ψ denote the potential and kinetic energy of one individual system, and the upper case symbols Φ and L denote the potential and kinetic energy of the ensemble of N systems, respectively.[103] Therefore, the RHS of the equation in the last line of the quotation denotes an average over the ensemble of N systems and the LHS represents an average of a single system realisation over time, hence its time average.

BOLTZMANN (1884) contains further definitions of other terms which define more and more specific ensembles or distributions of the ensemble. We find *Orthode* (p. 130), *Planode* (p. 135), *Isomonode* and *isodisch* (p. 137).[104] The argument of the narrative of isodic theory starts from here. MATHIEU (1988, p. 375) interprets BOLTZMANN's conception as indeed stemming from the Greek ἔργον [*ergon*] for *work*, as the 'systems which were described by equations involving only energy, which were "worklike"', but traces this back to the Greek adjective ἐργ-ώδης [*erg-odeis*] which means *laborious*, which in turn is derived from ἔργον [*ergon*] and again εἶδος [*eidos*] for *aspect, appearance* instead of ὁδός [*hodos*] for *path*.

One can invoke BOLTZMANN (1884, p. 147) himself to argue for the case that if 'ergodic theory' is the theory derived from the 'ergodic hypothesis' then it ought to be called 'isodic theory' with the following quote:

> "Jedesmal, wenn jedes einzelne System der Monode im Verlaufe der Zeit alle an den verschiedenen Systemen gleichzeitig nebeneinander vorkommenden Zustände durchläuft, kann an Stelle der Monode ein einziges System gesetzt werden. [...] Für eine solche Monode wurde schon früher die Bezeichnung „isodisch" vorgeschlagen."

For BOLTZMANN the term 'Isoden' described what we today call ergodic behaviour in phase space with ἴσος [*isos*] standing for *same* and οδός [*odos*] for *path*.[105]

In the translation of BOLTZMANN's *Lectures on Gas Theory* the translator, STEPHEN G. BRUSH, corroborates the isodic narrative and even refrains from translating the word 'Ergoden' at all. He gives the following explanation in the notes:

> "This word is left untranslated since the English equivalent, 'microcanonical ensemble', had not yet come into use when Boltzmann wrote this book. (See [Gibbs, 1902, chap. x.]) *Ergoden* should not be confused with 'ergodic systems', *i.e.*, hypothetical mechanical systems that have the (impossible) property that their coördinates and momenta eventually take on every value consistent with

103 See also MATHIEU 1988, p. 375.

104 SINGH (2011, p. 6) gives BOLTZMANN's neologisms the following equivalent using modern terminology: *Monode* as a stationary distribution, *Orthode* as a stationary distribution which verifies the heat theorem, *Ergode* as the microcanonical ensemble, and the *Holode* as the canonical ensemble. Unfortunately, he does not provide an analog to *Isode*.

105 GALLAVOTTI et al. 2004, p. 2.

> their fixed total energy. Boltzmann never used the word Ergoden for such systems, but 'ergodic' came to be applied to them following the discussion published by P. and T. Ehrenfest in the Encyklopädie der mathematischen Wissenschaften (1911). Although the Ehrenfests made a valuable contribution by their critical analysis of the foundations of gas theory, they unfortunately misrepresented the opinions and even the terminology of Boltzmann and Maxwell. Boltzmann did discuss ergodic systems without calling them that ([Boltzmann, 1871, 1887]), but he did not make the foundations of gas theory depend on their existence, nor did he even make a clear distinction between going through every point on the energy surface, and going infinitely close to every point. Ergoden were first introduced explicitly in 1884 ([Boltzmann, 1884, 1885]) although both Maxwell and Boltzmann had previously used the same device in making calculations. Boltzmann used the word 'isodic' for the systems that we now call ergodic. Their impossibility was proved independently by Plancherel [1913] and Rosenthal [1913]."[106]

With the preceding statement the section on the historical origin of the ergodic hypothesis and why it was given this particular name comes to an end. We want to conclude with a wink.

Highly-Educated 19th Century German-Speaker

During private conversation about the general research topic and the keyword of 'non-ergodicity' with researchers from other disciplines as well as non-scientists we realised that 'non-ergodicity' obviously does not roll off the tongue easily. The American statistician Cosma Shalizi has a much simpler explanation of Boltzmann's naming: 'Being a highly-educated nineteenth century German-speaker, Boltzmann knew far too much ancient Greek, so he called this the "ergodic property" [...] The name stuck.'[107] This tongue-in-cheek quote nevertheless contains a grain of the spirit of the convention to use Greek identifiers. Interestingly, the articles Boltzmann (1871, 1875) contain a precursor of 'Ergode' called 'Ergal', which he draw from a series of papers from Clausius (1875a,b,c). We note, the words 'entropy', 'energy' and 'ergodicity' all share the same Greek root, but apart from that we spare the details of the etymology of 'Ergal'. Nonetheless it shows some proclivity for Greek language seemed to have been in the air back then, because Clausius was another 'highly-educated nineteenth century German-speaker' and expressed this proclivity explicitly when he coined the word 'entropy':

> "Sucht man für S [Entropie/entropy] einen bezeichnenden Namen, so könnte man, änlich wie von der Größe U [innere Energie/internal energy] gesagt ist,

[106] Boltzmann 1964, 282, with our references given in square brackets.
[107] Shalizi 2017, p. 575.

> sie sey der Wärme- und Werkinhalt des Körpers, von der Größe *S* sagen, sie sey der Verwandlungsinhalt des Körpers. Da ich es aber für besser halte, die Namen derartiger für die Wissenschaft wichtiger Größen aus den alten Sprachen zu entnehmen, damit sie unverändert in allen neuen Sprachen angewandt werden können, so schlage ich vor, die Größe *S* nach dem griechischen Worte ἡ τροπή [n trop], die Verwandlung [transformation/conversion], die *Entropie* des Körpers zu nennen. Das Wort *Entropie* habe ich absichtlich dem Worte *Energie* möglichst ähnlich gebildet, denn die beiden Größen, welche durch diese Worte benannt werden sollen, sind ihren physikalischen Bedeutungen nach einander so nahe verwandt, daß eine gewisse Gleichartigkeit in der Benennung mir zweckäßig zu seyn scheint."[108]

With this we leave the situation to the ancient Greek philologist and lead over to the ensuing section. After the birth of the ergodic hypothesis, the subsequent period is best characterised as an emancipation from its origin in statistical mechanics and the emergence of ergodic theory as a branch of mathematics. Two proofs of ergodic theorems by two eminent mathematicians, John von Neumann and George D. Birkhoff, kick-started the exploding interest by mathematicians in the field. Ergodic theory then grew into a thriving branch of mathematics, that now links probability theory, stochastic processes and dynamical systems theory.[109] The following section introduces some basic mathematical aspects of ergodicity and discusses different point of views on ergodicity.

2.4 Ergodic Measurable Dynamical Systems and Ergodic Theorems

Like every mathematical concept also ergodicity can be given a precise definition. With the help of assumptions and sometimes even axioms a definition divides between what contains to it and what not. This restrictiveness allows to generate true statements about the objects of interest. However, we sometimes forget that the defined objects are very exceptional objects, in the sense that they are well-behaved and possess the defined property. They may very well be the objects about which we can say something at least from a mathematical standpoint, but they may not represent what we are confronted with in the ill-behaved real-world. Take for example, the mathematical property of an object to be round. There are far more non-round objects conceivable than there are round objects – even if both sets are infinite. Therefore, we have to keep in mind Ulam's non-elephants and that the mathematical field of ergodic theory

[108] Our translation of the symbols in square brackets from Clausius (1865, p. 390).
[109] Auyang 1998, p. 2.

makes statements only about the smaller part of what is imaginable. The far greater set falls into the terrain of non-ergodicity.

In applications of ergodic theory outside of mathematics, for example in our case to economics, we have to be aware of the fact that the empirical scientist has to accomplish even more than the non-empirical mathematician. In economic reality we may not always encounter well-behaved objects that conveniently possess a certain property. We may encounter ergodic observables as well as non-ergodic observables in economics. One strategy to handle such ill-behaved objects is to reduce them with some transformations to well-behaved tractable objects. Therefore, our ultimate goal will be to apply the material presented in this section to economic problems where we are often confronted with non-ergodic observables.

2.4.1 Non-Ergodicity is the Rule, Ergodicity is the Exception

In the introduction to his article on 'Remarks on Non-Markov Processes' van Kampen (1998) reminds us about what we just discussed:

> "**Definition of non-Markov processes.**
>
> The term 'non-Markov process' covers all random processes with the exception of the very small minority that happens to have the Markov property.
>
> FIRST REMARK. Non-Markov is the rule, Markov is the exception.
>
> It is true that this minority has been extensively studied, but it is not proper to treat non-Markov processes merely as modifications or corrections of the Markov processes as improper as for instance treating all non-linear dynamical systems as corrections to the harmonic oscillator."[110]

Van Kampen's distinction between the rule and the exception is an essential insight. We could indeed introduce non-ergodicity in much the same way as van Kampen did introduce Non-Markov processes, with the single substitution of 'Markov' by 'ergodic' and it would produce the following true and important frame for this section:

> "**Definition of non-[ergodic] processes.**
>
> The term 'non-[ergodic] process' covers all random processes with the exception of the very small minority that happens to have the [ergodic] property.
>
> FIRST REMARK. Non-[ergodic] is the rule, [ergodic] is the exception.

[110] Van Kampen 1998, p. 90.

> It is true that this minority has been extensively studied, but it is not proper to treat non-[ergodic] processes merely as modifications or corrections of the [ergodic] processes as improper as for instance treating all non-linear dynamical systems as corrections to the harmonic oscillator."[111]

After having considered the historical origin of ergodicity stemming from physics in Sec. 2.3 we continue with formal definitions which facilitate a comprehension of the scope of ergodic theorems and their relation to economics. We continue to give physical[112] interpretations of the mathematical notation which enables an understanding of their meaning in the context of economics and decision making under uncertainty. In physics and mathematical economics it is common to model the evolution of a real-world system via the mathematical theory of dynamical systems. Consequently, a dynamical system is our object of study, which consists of a measure space and at least one transformation, which we introduce step by step in the following.

2.4.2 Measure and Probability Spaces

Let the triple $(S, \mathcal{A}, \mu)$ be a measure space. The set S is interpreted as the phase space of the system[113], $\mathcal{A}$ denotes a σ-algebra of measurable subsets of S, which are our observable events, and the finite measure μ is a mapping from the possible events to the non-negative reals $\mu : \mathcal{A} \mapsto \mathbb{R}^+$, and $\mu(S)$ will be interpreted as the volume of the phase space. In this thesis, we are almost exclusively interested in probability measures denoted by $\mathcal{P}$ for which $\mathcal{P}(S) = 1$ and $\mathcal{P} : \mathcal{A} \mapsto [0, 1]$. Although everything discussed for probability spaces also applies to more general measure spaces with nor or only little additions. Hence, the measure space $(S, \mathcal{A}, \mu)$ in our dynamical system becomes an ordinary probability space denoted by $(S, \mathcal{A}, \mathcal{P})$.

2.4.3 Time, Transformations and Trajectories

In the modelling framework of dynamical systems theory, the change or evolution of the system through time is modelled via mappings. Let us begin with defining proper time domains on which these mappings operate and in which the evolution takes place. Let $\mathbb{T}$ be one of the sets $\mathbb{R}^+ = [0, \infty)$ or $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. The set $\mathbb{T}$ will be interpreted as the time domain. If

111 Paraphrasing VAN KAMPEN (1998, p. 90).

112 The adjective 'physical' does not mean to apply only to objects from physics but is rather meant in the sense of 'real-world' objects, similar to the physical meaning of ensemble and time averages given in Sec. 2.1.

113 For a discussion of synonyms of S see Subsec. 3.5. For the purpose of the thesis we will mostly refer to the ensemble, but S is labelled synonymously the phase space in physics, the sample space in mathematics and the state space in economics.

we study phenomena in continuous time then $\mathbb{T} = \mathbb{R}^+$, if we study phenomena in discrete time then $\mathbb{T} = \mathbb{N}_0$.[114] Often real-world physical systems such as economies are analysed in discrete time. Then the measurements take place at regularly spaced equidistant intervals and only the subset of the natural numbers is selected from the real numbers as instants of time when the measurements are made.[115]

The evolution of the system through time manifests itself as a movement through phase space. Therefore, we consider a family $(\varphi_t)_{t\in\mathbb{T}}$ of $\mathcal{A}$-$\mathcal{A}$-measurable self-mappings of S, $\varphi : S \to S$, which describes this time evolution of the dynamical system. Thereby, $\varphi_0 = \mathcal{I}$ is the identity map, $\mathcal{I} : s \mapsto s$. More precisely, if the system is at time 0 in state $s \in S$ then the system will be at time t in state $\varphi_t(s)$. The specific value of the state $\varphi_t(s)$ is modelled as the outcome of a random experiment and thus $\varphi_t(s)$ is a random variable. The sequence of positions visited during such a movement through phase space motivates the ensuing definition.

Definition 2.1 (Trajectory). *For all $s \in S$ the sequence of points the system visits in phase space over time is denoted by $(\varphi_t(s))_{t\in\mathbb{T}}$ and called the trajectory or orbit of the system starting from point s. The family $(\varphi_t)_{t\in\mathbb{T}}$ denotes the ensemble of all possible trajectories for all initial conditions s from S.*

The trajectory of an initial point $s \in S$ represents a complete history of the system, *i.e.* all the states the system visits over the time. It is common to assume a functional relation between the members of the family $(\varphi_t)_{t\in\mathbb{T}}$. In the continuous case $\mathbb{T} = \mathbb{R}^+$ we assume that we have the following relation using the operator '$\circ$' for the composition of mappings:

$$\varphi_{t_1+t_2} = \varphi_{t_1} \circ \varphi_{t_2} \qquad \forall\ s \in S,\ t_1, t_2 \in \mathbb{R}^+ \ . \tag{2.19}$$

The assumption in eq. (2.19) encodes a fundamental semigroup property about the special nature of the dynamics of the system. In a way it is assumed that 'the laws governing the behavior of the system do not change with time'.[116] Furthermore, it is assumed that the time evolution is decomposable into commutative parts. In continuous time this leads to the fact that $(\varphi_t)_{t\in\mathbb{T}}$ is a flow. In the discrete case $\mathbb{T} = \mathbb{N}_0$, we suppose that the family $(\varphi_t)_{t\in\mathbb{T}}$ is generated by a single $\mathcal{A}$-$\mathcal{A}$-measurable self-mapping Φ of S. More precisely, we assume that for all $k \in \mathbb{N}_0$ we have the following relation:

$$\varphi_k = \Phi^k \ . \tag{2.20}$$

[114] An additional indicator of the time domain used at a certain place is the index variable. In this chapter we will use t only in the continuous case $\mathbb{T} = \mathbb{R}^+$, and *e.g.* k in the discrete case $\mathbb{T} = \mathbb{N}_0$.

[115] We could in principle study the case of $\mathbb{T} = \mathbb{R}$ or $\mathbb{T} = \mathbb{Z}$ as well, but limit the presentation to the physical meaningful applications of transformations forward in time.

[116] Petersen 1989, p. 1.

Here a single mapping Φ is the generator of the dynamics and completely governs the dynamics of the system. The transition from time step k to time step $k+1$ for all $k \in \mathbb{N}_0$ is only determined by Φ. The majority of our discussion of ergodic theory focuses on the action of powers of a single generator transformation Φ in discrete time. All our statements and most of the results on ergodic theory straightforwardly extend to the case of continuous time, *i.e.* the action of flows.[117]

2.4.4 Measure-Preserving Transformations and Ergodic Dynamical Systems

Now we presuppose a particular connection between the measure space $(S, \mathcal{A}, \mu)$ and the family $(\varphi_t)_{t\in\mathbb{T}}$ of $\mathcal{A}$-$\mathcal{A}$-measurable self-mappings of S. For this reason we first introduce the notion of μ-preserving transformation.

Definition 2.2 (Measure-Preserving Transformation). *Let be $(S, \mathcal{A}, \mu)$ a measure space and let φ be the self-mapping $S \to S$ be $\mathcal{A}$-$\mathcal{A}$-measurable. Then φ is called μ-measure-preserving if for every $A \in \mathcal{A}$ the equality*

$$\mu\left(\varphi^{-1}(A)\right) = \mu(A) \tag{2.21}$$

holds, i.e the set A and its preimage $\varphi^{-1}(A)$ have the same μ-measure.

Measure-preserving transformations are synonymously called μ-invariant and are the key component of measurable dynamical systems.[118] In the discrete case $t \in \mathbb{N}_0$, it is enough to assume that the generator mapping Φ preserves the measure μ, which means that for all $A \in \mathcal{A}$

$$\mu\left(\Phi^{-1}(A)\right) = \mu(A) \tag{2.22}$$

holds. Definition 2.2 jointly with eq. (2.20) implies for all $k \in \mathbb{N}_0$ and all $A \in \mathcal{A}$ the relation

$$\mu\left(\varphi_k^{-1}(A)\right) = \mu(A)\ . \tag{2.23}$$

The mapping φ is in a dynamical systems context often referred to as a transformation, because the real-world change that happens is modelled as a transformation which acts

117 The continuous case is related to local ergodic theorems, see especially the work on the maximal ergodic theorem starting with Wiener (1939), Yosida and Kakutani (1939) and Hopf (1954). See also the discussion in Einsiedler and Ward (2011, ch. 8.6) and Viana and Oliveira (2016, ch. 3.4).

118 Measure-preserving transformations are subject of an important theorem for statistical mechanics by Liouville, who proved the conservation of the phase space volume of a Hamiltonian flow of unitary operators in Hilbert space.

upon the physical system from one time step to another. Hence, the resulting quadruple $\left(S, \mathcal{A}, \mathcal{P}, (\varphi_t)_{t\in\mathbb{T}}\right)$ is our model of a real-world system, in the case of discrete time $\mathbb{T} = \mathbb{N}_0$ the notation can be made even more compact and we write $(S, \mathcal{A}, \mathcal{P}, \Phi)$. Furthermore the transformation and all powers Φ^k of it are also an endomorphism of $(S, \mathcal{A}, \mu)$ for all $k \in \mathbb{N}_0$

$$\mu\left(\left(\Phi^k\right)^{-1}(A)\right) = \left[\Phi^k(\mu)\right](A) = \mu(A)\ . \tag{2.24}$$

Definition 2.3 (Measurable Dynamical System). *A measurable dynamical system is the quadruple $(S, \mathcal{A}, \mu, \Phi)$ with the measure space $(S, \mathcal{A}, \mu)$, and Φ is a $\mathcal{A}$-$\mathcal{A}$-measurable self-mapping of S, which leaves the measure μ invariant.*

At the centre of ergodic theory are now special measurable dynamical systems of the following type.

Definition 2.4 (Ergodic Measurable Dynamical System). *A measurable dynamical system $(S, \mathcal{A}, \mu, \Phi)$ is called ergodic, if for every set $A \in \mathcal{A}$ that meets the condition $\Phi^{-1}(A) = A$ either $\mu(A) = 0$ or $\mu(S \setminus A) = 0$.*

Intuitively, Def. 2.4 states the impossibility to decompose Φ into two ergodic transformations. Let us give three real-world examples of transformations that leave the measure invariant.

1. First, a simple measure-preserving transformation one can think of is a permutation on a finite set with uniform distribution. *E.g.* opposite sites of a regular dice add up to seven, if we change the position of pips such that opposite sites do not add up to seven anymore, this would not change the probability of $1/6$ of a particular number of pips.
2. Second, an even simpler illustration is the cream poured in a cup of coffee. The cream has a certain measure, *e.g.* its weight. If we pour it in a coffee cup the measure (weight) of the cream stays invariant under the transformation of stirring – even if the initial drop of cream becomes infinitely thin eventually, because of the perfect homogeneous mix of coffee and cream after the enough stirring. Here we already see the property of mixing of the transformation stirring, which is even stronger than ergodic. This will be discussed in Subsec. 2.6.7 in the context of a hierarchy of properties of which ergodic is the weakest.
3. A third example of a measure-preserving transformation for countably infinite sets is the baker transformation. A baker has a finite amount of white dough to which he adds a smaller amount of some dark, say chocolate, dough. Suitable measures for the two sorts of dough would be their weight. Let us assume the light dough weighs 900 g and

the dark chocolate dough weighs 100 g, respectively. With every folding and stretching, which is the transformation that the baker applies to the total amount of dough, the two doughs become more and more mixed. But if we would measure the two sorts of dough after every iteration of continued transformations, there would still be the same amount of dark and white dough of 900 g and 100 g, respectively. Hence, the baker transformation is measure-preserving, ergodic and strongly mixing, see Subsec. 2.6.7.

In contrast, examples of transformations that are not measure-preserving and thus not ergodic are urn models that do not hold the number of balls in the urn constant, *e.g.* drawing balls without replacement of the drawn ball or PÓLYA urn models, were the drawn ball is replaced by more than one ball of the same colour. Here, every draw creates an accelerating feedback and path dependence. The probability to draw a ball of a certain colour changes with every transformation or time step. For references see Subsec. 5.2.4.

2.4.5 From Trajectories to Observables to Stochastic Processes

Most often we are not able to analyse the trajectory in (high-dimensional) phase space directly, but measure only a quantity of interest to us and that quantity depends on the state of the system. We refer to such a quantity as an observable of the system.

Definition 2.5 (Observable). *An observable is a $\mathcal{A}$-$\mathcal{B}$-measurable function $f : S \to \mathbb{R}$, where $\mathcal{B}$ denotes the set of* BOREL *σ-algebras of $\mathbb{R}$.*

An observable is simply a function that takes the state of the system (expressed by a point in phase space) as its input and transforms the input in such a way that a real number is the output. One advantage of the distinction between the realisation of a trajectory and the observation of a realised trajectory is that the coarse-graining measurement procedure is made explicit. Often an observable is a coarse-grained aggregate quantity of our system, and the question arises about the amount of information that gets lost during such an aggregation. Examples of coarse-grained aggregates are *e.g.* in physics the temperature of a system, in economics the GDP of an economy. In later chapters of this thesis, the wealth of a gambler will be our prime observable.

As we are interested in the evolution of a system we perform repeated measurements of the system. From Def. 2.1 we have expressions for realisations of trajectories. One specific trajectory through time starting at s – independent of the observation or measurement though unobserved in a sense – is denoted by $(\varphi_t(s))_{t\in\mathbb{T}}$. Combining Def. 2.1 of a trajectory and Def. 2.5 of an observable, we derive respective expressions for *observed* trajectories. We denote an observed trajectory through time starting at s by $(f \circ \varphi_t(s))_{t\in\mathbb{T}}$.

Explicit Representation of the Sequences

In the following we will concentrate on the discrete case $\mathbb{T} = \mathbb{N}_0$. More explicitly, in the case of an evolution in discrete time, we get the following sequences. The trajectory $\left(\Phi^k(s)\right)_{k\in\mathbb{N}_0}$ of a system starting at s consists of the following sequence

$$\left(\Phi^k(s)\right)_{k\in\mathbb{N}_0} = \left\{s, \Phi(s), \Phi^2(s), \ \ldots\right\} \ . \tag{2.25}$$

The observed trajectory starting at s is then denoted by $\left(f \circ \Phi^k(s)\right)_{k\in\mathbb{N}_0}$ and contains the following sequence

$$\left(f \circ \Phi^k(s)\right)_{k\in\mathbb{N}_0} = \left\{f\left(s\right), f\left(\Phi\left(s\right)\right), f\left(\Phi^2\left(s\right)\right), \ \ldots\right\} \tag{2.26}$$

$$= \left\{f \circ s, \left(f \circ \Phi\right)(s), \left(f \circ \Phi^2\right)(s), \ \ldots\right\} \ . \tag{2.27}$$

Eq. (2.27) can be interpreted as the outcome of repeated observations or measurements of the states of a system given by the trajectory in eq. (2.25). The family of all possible trajectories $\left(\Phi^k\right)_{k\in\mathbb{N}_0}$ of a system consists of the following sequence

$$\left(\Phi^k\right)_{k\in\mathbb{N}_0} = \left\{\Phi^0, \Phi^1, \Phi^2, \ \ldots\right\} \ . \tag{2.28}$$

The family of all observed possible realisations of trajectories are denoted by $\left(f \circ \Phi^k\right)_{k\in\mathbb{N}_0}$ the sequence

$$\left(f \circ \Phi^k\right)_{k\in\mathbb{N}_0} = \left\{f, f\left(\Phi\right), f\left(\Phi^2\right), \ \ldots\right\} \tag{2.29}$$

$$= \left\{f, f \circ \Phi, f \circ \Phi^2, \ \ldots\right\} \ . \tag{2.30}$$

Distinction between Trajectory and Observable in Economics

By now we have developed a notation with the explicit reference to an observable, which emphasises that we do not directly observe the true value or the system in all its microscopic details but only perform coarse-grained macroscopic measurements on our stochastic system, from which we can then again derive compatible microscopic configurations. All this is encoded in the measurement operation $f(\cdot)$. Advantages that derive thereof are that we are able to distinguish neatly between trajectories and observables and that we see the transformation of the underlying dynamical system and how it generates the dynamic. Though one downside is that such a notation is less common in economics. The distinction between a trajectory and an observable is usually not made explicit in economics. Although the measurement procedure is especially in macroeconomics very indirect and economists are very far from having a tangible experience of the states of the economic system. Instead, the study object is simply referred to as a time series of a quantity y of economic interest, such as $\{y_t, y_{t+1}, y_{t+2}, y_{t+2}, \ldots\}$. As another downside our notation grew quite cumbersome. We will often simply refer to the sequences in eq. (2.25) and in eq. (2.26) as *observables* where in econom(etr)ics and other empirical sciences one commonly refers to *time series*. However both terms signify the somewhat bulky terms of *observed trajectory* or *all observed possible trajectories*. To downsize the notation we utilise the fact that the sequences in eq. (2.25) and eq. (2.26) are sequences of scalar numbers, which are generated by a random experiment and thus are realisations of the stochastic processes in eq. (2.28) and eq. (2.29), respectively. For the purpose of this thesis, our observable will be the wealth of a gambler. A gambler's wealth at a certain moment in time will then be understood as the realisation of a random variable or the realisation of a stochastic process, because it depends on a random event such as the number of consecutive coin flips in a (St. Petersburg) lottery, to which we turn to in chapter 3. Thus our observable is a stochastic process. Therefore the introduction of stochastic processes follows quite naturally and facilitates a more compact notation. Additionally, the connection between stochastic processes and dynamical systems which generated them becomes more transparent in our exposition.

Stochastic Processes

> Gegeben sei uns der Zustand unseres Gases zu Anfang der Zeit, also $f(x,0)$. Gefunden soll werden der Zustand nach Verlauf einer beliebigen Zeit t, also $f(x,t)$. Der Weg, den wir da einschlagen werden, ist derselbe, den man in ähnlichen Fällen immer einschlägt. Wir berechnen zuerst, um wieviel sich die Funktion $f(x,t)$ während einer sehr kleinen Zeit τ verändert; hierdurch erhalten wir zunächst eine partielle Differentialgleichung für $f(x,t)$; dieselbe muß dann so integriert werden, daß f für $t=0$ den gegebenen Wert $f(x,0)$ annimmt.[119]
>
> LUDWIG BOLTZMANN

In the framework of random dynamical systems the trajectories (or the time series) are modelled as sequential realisations of the same random variable at successive moments in time. Consequently, the repeated measurements of an observable can be interpreted as one realisation of a stochastic process. An arbitrary function of a random variable is also a

119 BOLTZMANN 1872, p. 322.

random variable and an arbitrary function of a stochastic process is as well a stochastic process. Therefore the family of observables in eq. (2.28) and eq. (2.29) are stochastic processes.

Definition 2.6 (Stochastic Process). *The family $(Z_t)_{t\in\mathbb{T}}$ of real-valued random variables on the probability space $(S, \mathcal{A}, \mathcal{P})$ is called a stochastic process with the time domain $\mathbb{T}$ and range $\mathbb{R}$.*

We can consider a stochastic process also in a different but natural way as a mapping of a pair of variables into the reals. The first variable is a point in phase space and the second variable is the time parameter. Formally, let $(Z_t)_{t\in\mathbb{T}}$ be a stochastic process then we introduce the mapping $Z : S \times \mathbb{T} \to \mathbb{R}$ which is defined by

$$Z(s, t) \coloneqq Z_t(s) \ . \tag{2.31}$$

A function of two variables generates in a natural way two functions of one variable. Namely, we can fix the first variable and let the second variable vary or vice versa. For a fixed point in phase space $s \in S$ the mapping $Z_{s,\bullet} : \mathbb{T} \to \mathbb{R}$ is defined by

$$t \mapsto Z_t(s) \ . \tag{2.32}$$

Thereby $Z_t(s)$ describes a single trajectory (interchangeably realisation or sample path) of the stochastic process Z_t which is associated with s, see also **Fig. 2.1**.

For a fixed moment in time $u \in \mathbb{T}$ the mapping $Z_{\bullet,u} : S \to \mathbb{R}$ is defined by

$$s \mapsto Z_u(s) \ . \tag{2.33}$$

Thereby, we are arrive again at the mapping Z_u belonging to the original stochastic process $(Z_t)_{t\in\mathbb{T}}$. Looking at a stochastic process by this means elucidates why they are sometimes referred to as random functions, because the mapping acts on a probability space.[120]

Now we consider the discrete case $\mathbb{T} = \mathbb{N}_0$ and use the just introduced terminology to rephrase the stochastic processes in eq. (2.28) and eq. (2.29). For $s \in S$ and $k \in \mathbb{N}_0$ the stochastic processes in eq. (2.28) can be rephrased as

$$Y(s, k) = \Phi^k(s) \ . \tag{2.34}$$

[120] For a detailed introduction of stochastic processes see *e.g.* Mazo (2002, ch. 3) or van Kampen (2007, ch. 3). See especially Mazo (2002, pp. 26–27) for the alternative used notations in the literature.

For $k \in \mathbb{N}_0$ the stochastic processes in eq. (2.29) is rephrased if

$$Z_k := f \circ \Phi^k \tag{2.35}$$

and the mapping $Z : S \times \mathbb{T} \to \mathbb{R}$ according to $Z(s,k) := Z_k(s)$ then

$$Z(s,k) = \left(f \circ \Phi^k\right)(s) \tag{2.36}$$

denotes one sample path of the stochastic process in eq. (2.35) associated with the point s.

2.4.6 Ergodic Theorems

Time Averages of Observables

In many disciplines such as economics, statistical mechanics and information theory the long-term average of observables is among the potentially interesting quantities of the system. For example, Khinchin (1949) showed that the thermodynamic observables of interest – from a mathematical point of view – are sum functions which are defined via averages. In later chapters of this thesis, our observable is the wealth of a gambler (who is our decision maker) or derivatives thereof like the growth rate of wealth. In this particular economic application we study the long-term development of wealth if certain lotteries are played repeatedly under a certain dynamic. Therefore, we now introduce a key component in ergodicity economics, namely the concept of a time average.

Definition 2.7 (Finite-Time Average). *Let $(Z_k)_{k\in\mathbb{N}_0}$ be the stochastic process given by eq. (2.35). Then for $T \in \mathbb{N}_0$ we call*

$$\overline{Z}_T := \frac{1}{T}\sum_{i=0}^{T-1} Z_i \tag{2.37}$$

the T^{th}-time average of the sequence $(Z_k)_{k\in\mathbb{N}_0}$.

Clearly, $\left(\overline{Z}_T\right)_{T\in\mathbb{N}_0}$ is itself a stochastic process, because every function of a random variable is itself a random variable, and every function of a stochastic process is itself a stochastic process. However, what we are interested in is the asymptotic behaviour of the finite-time averages $\overline{Z}_T$. Doing this we first have to decide which type of convergence we should take into account. We will focus on μ-almost everywhere (a.e.) convergence, see for example the pointwise ergodic theorem by Birkhoff (1931b). It should be mentioned that in the theory of dynamical systems the convergence with respect to the $\mathcal{L}^1(\mu)$-norm is also of interest, see *e.g.* the mean ergodic theorem by VON NEUMANN (1932b).

To investigate the asymptotics, let S_0 be the set of all points $s \in S$ for which the sequence or sample path $\left(\overline{Z}_T(s)\right)_{T\in\mathbb{N}_0}$ converges. In the general case we can not expect that the set S_0 is 'large'. However, under appropriate assumptions, the set $S \setminus S_0$ will turn out to be 'very small', namely a set of μ-measure zero. In this case, the next step consists of a thorough investigation of the infinite-time average which is the limit function

$$\overline{Z} := \lim_{T\to\infty} \overline{Z}_T \,. \tag{2.38}$$

Especially interesting is the case, when the limit function $\overline{Z}$ turns out to be μ-a.e. constant. This case is a basic situation studied in ergodic theory, and leads to the pointwise ergodic theorem.

Theorem 2.1 (Pointwise Ergodic Theorem, Birkhoff 1931b). *Let $(S, \mathcal{A}, \mu, \Phi)$ be an ergodic measurable dynamical system with probability measure μ. Then for all μ-integrable functions f and it holds*

$$\overline{Z} = \int_S f \,\mathrm{d}\mu \quad \mu\text{-}a.e. \tag{2.39}$$

Here, the number to which the time averages converge is just the expected value of the random variable f,

$$\mathrm{E}\,[f] := \int\limits_S f \,\mathrm{d}\mu \,. \tag{2.40}$$

Therefore, theorem 2.1 justifies the substitution of time averages by ensemble averages. Finding such a vindication is exactly one possible solution to the ergodicity problem in statistical mechanics as aptly coined in KHINCHIN (1949, p. 47) which was initiated by the ergodic hypothesis and BOLTZMANN's H-theorem. However, the first proof of an ergodic theorem goes back to the mean ergodic theorem in VON NEUMANN (1932b). The pointwise ergodic theorem in BIRKHOFF (1931b) constitutes the second proof.[121] Sometimes theorem 2.1 is also referred to as the *individual ergodic theorem* because eq. (2.39) states the convergence for μ-almost all time averages associated with μ-almost all of the individual points $s \in S$. BIRKHOFF's result has been generalised in KHINTCHINE (1933) to abstract finite measure

[121] See also ZUND (2002) and BERGELSON (2004) for the chronology of priorities in the first proof of an ergodic theorem. Although BIRKHOFF's proof appeared first in print, it should be clear that VON NEUMANN's proof was first, since BIRKHOFF (1931b, p. 656) and BIRKHOFF and KOOPMAN (1932, p. 281) also refer explicitly to an unpublished manuscript version of VON NEUMANN (1932b) from which they got their inspiration. See also HOPF (1932). Furthermore, the earlier article VON NEUMANN (1929) deals with the ergodic hypothesis and the H-theorem in quantum mechanics. See also the remarks in ULAM (1976, p. 98).

spaces which is why some authors also refer to the pointwise ergodic theorem as the BIRKHOFF-KHINCHIN ergodic theorem.[122] Convergence almost everywhere is a more general form of convergence than the convergence in $\mathcal{L}^1(\mu)$-mean, proved in the following mean ergodic theorem.[123]

Theorem 2.2 (Mean Ergodic Theorem, von Neumann 1932b). *Let $(S, \mathcal{A}, \mu, \Phi)$ be an ergodic measurable dynamical system with probability measure μ. Then for all μ-integrable functions f and μ-almost all $s \in S$ it holds*

$$\lim_{T\to\infty} \int_S \left| \overline{Z}_T - \mathrm{E}[f] \right| \mathrm{d}\mu = 0 \ . \tag{2.41}$$

The mean ergodic theorem takes into account a more realistic understanding of the measurement process. As mentioned already in Sec. 2.2 in the context of statistical mechanics, in real-world applications many observations or point measurements are in fact finite-time averages. This is the case if the time scale of the measured microscopic effects is much faster than the time scale of the macroscopic measurement procedure itself. Given a sufficient speed of convergence the outcome of a point measurement may yield a sufficiently reliable approximation of the 'true' infinite-time average, *i.e.* it will be sufficiently close to the 'true' infinite-time average. The standard example in statistical mechanics are molecular collisions of a gas during the duration of a measurement of macroscopic quantities like pressure or temperature of the gas. The collisions happen on such tiny microscopic time scales compared to the measurement, that any macroscopic measurement is a long-term average.[124] VON NEUMANN explicitly stressed the point that the mean ergodic theorem more closely resembles the physical measurement process in a follow up paper[125] to the publication of the proof of the mean ergodic theorem – another intention was to emphasise why the *mean* ergodic theorem is advantageous. One seeks to choose long enough measurements in order to get arbitrarily small statistical dispersions of the time average. More precisely, there exists a measurement duration $T > T^*(\epsilon, \delta)$ for which the following inequality, which contains an expression from eq. (2.41), is satisfied in terms of probability

$$\Pr\left\{ \left| \overline{Z}_{T^*} - \mathrm{E}[f] \right| > \delta \right\} = \Pr\left\{ \left| \frac{1}{T^*} \sum_{i=0}^{T^*-1} Z_i - \mathrm{E}[f] \right| > \delta \right\} \leq \epsilon \ . \tag{2.42}$$

The situation is similar in economics, especially for macroeconomic quantities. Consider the point measurement of the observable 'unemployment' on February 1$^{\text{st}}$ 2019. Then it is

[122] Based on results in KOLMOGOROFF (1937) and YOSIDA and KAKUTANI (1939).
[123] The mean ergodic theorem can be further generalised to convergence in $\mathcal{L}^p(\mu)$-mean, $1 \leq p \leq \infty$, see *e.g.* WALTERS (1982, p. 36).
[124] VON NEUMANN 1932b, p. 80; PETERSEN 1996, p. 175.
[125] VON NEUMANN 1932c, pp. 264–265.

implicitly assumed that the unemployment on this day is in some sense representative for a period that exceeds a day, maybe that of a month or the winter season. Actually, on this very day there was a certain influx of some newly unemployed and a certain outflux of some newly employed people, leading to different measurements for the observable 'unemployment' if performed in the morning, during the working day or in the evening. The subsequent use of this number assumes a sufficient convergence to the true number on this day or extrapolated on the respective time period. The same reasoning applies to the measurement of GDP p.c. During the measurement of GDP p.c. at an instant in time, the quantity that actually gets measured is a finite-time average of the GDP p.c. during the duration of the measurement. Even if the measurement takes only an instant in time at our human time scale or the time scale of the measurement procedure, expressed in the time scale of the measured effect the instant may become a 'long' duration.[126]

Speed of Convergence

It is important to note that neither the pointwise ergodic theorem in 2.1 nor the mean ergodic theorem in 2.2 nor any other ergodic theorem contains any information on the speed of the convergence. The convergence is assured only in the asymptotic limit. Petersen (1989, ch. 3) even contains a proof that no statement about the speed of convergence in ergodic averages is possible. This has been a major point of criticism whenever ergodicity is assumed in applications like the measurement process just mentioned. Furthermore, when real-world systems like the economy are studied for policy reasons, the timescale of *e.g.* the speed of a mathematical convergence is of utmost importance. See the comments in Subsec. 5.2.5 and especially Subsec. 5.2.5.4 on the *Ergodicity Problem in Econometrics* and the challenge to create tests for ergodicity of time series. Simulation results are further discouraging in this context as expressed in the following quote: '[N]othing in the definition of "ergodic" depends on [a particular] timescale; it might in fact take much longer than the age of the universe for an actual ergodic system to [cover the whole phase space]'.[127] For more details on the speed of convergence in ergodic theorems see Krengel (1978, 1985, pp. 14–15).

[126] Too long durations of the measurements complicate the issue even further. For example the measurement of quantities of two individuals may cover two smaller but disjoint intervals during a longer total measurement procedure, but still they both enter with the same time index into the calculation. Such issues have been touched upon in Marx (2018, pp. 64, 70) in the context of the calculation of correlations between financial indices at different exchanges. As soon as the shifts in calender time get big enough, *e.g.* between the exchanges in Tokyo and New York, daily closing prices at one exchange can become part of the information set on which trading happens at another exchange, although both closing prices are indexed the same day.

[127] Palmer 1989, p. 278.

Ergodic Theory and the Strong Law of Large Numbers

The strong law of large numbers (SLLN) states the convergence of the arithmetic mean of sequences of identically and independently distributed (iid) random variables to their expectation value. From Def. 2.7 we can see that ergodic theory and especially the pointwise ergodic theorem yield conditions under which the arithmetic means in eq. (2.37) – which are the time averages from our point of view – converge to the expectation value in eq. (2.40) – which is the ensemble average from our point of view. Ergodic theory thus yields conditions under which the SLLN is satisfied for sequences of random variables with some dependencies and which fulfil additional time stationarity conditions.

Both ergodic theorems provide for the first time clarity about which mathematical model corresponds to the ergodic hypothesis in physics, namely semigroups of measure-preserving transformations. Subsequently, these two theorems led to a rise in interest in ergodic theory and started its emancipation into an independent mathematical field, which has ever since led to a cross-fertilisation between such fields as measure and probability theory, combinatorics, topology, set theory, functional analysis, number theory and even logic.[128]

2.5 Intuition on Ergodicity and Stochastic Processes

> Possibly when the dynamics of biological and social change shall have advanced to an adequate state of development some future Birkhoff will find and prove some super-ergodic theorem that will be a basic bit of mathematics in the statistical mechanics of such systems.[129]
>
> Edwin B. Wilson

In order to highlight the key terminology for the purpose of this thesis in particular and *Ergodicity Economics* in general, we make explicit the distinction between two types of averages, *i.e.* an ensemble average and a time average. The procedure of ensemble averaging relies on the averaging over a static timeless ensemble or phase space and does not rely on any evolution over time or any factual realisation of a stochastic processes. However, the concept of a time average hinges on an evolution over time and thus is a dynamic understanding of the object of research. In the remainder of this thesis, our observable is the wealth of a gambler (who is our decision maker) and derivatives thereof like the growth rate of wealth.

[128] Eisner et al. 2015, p. vii; Dumas 2014.
[129] Wilson 1945, p. 580

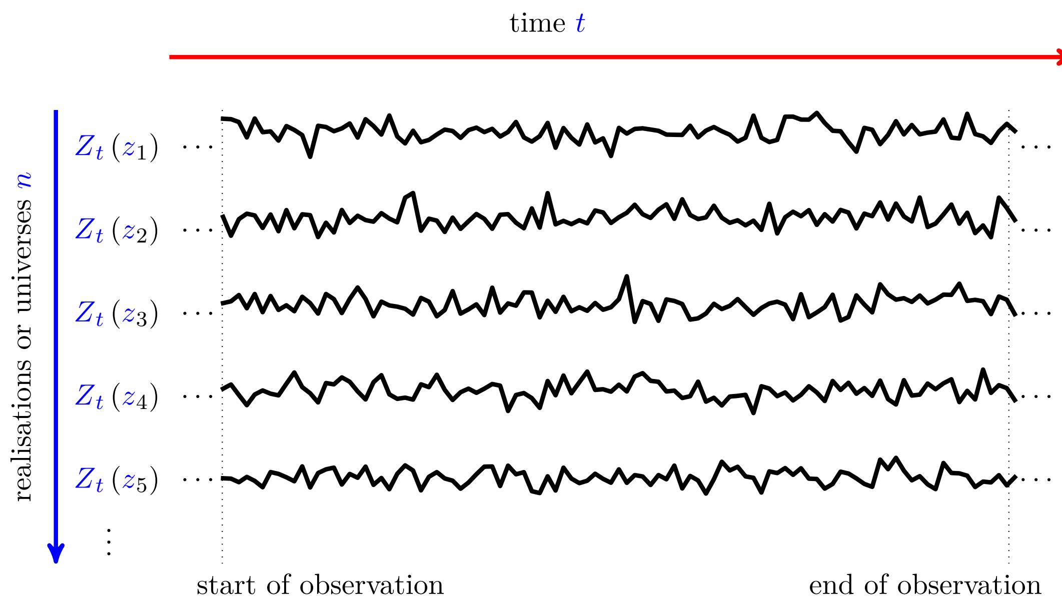


Figure 2.1: Stochastic Process Grid. Visualisation of the nature of a stochastic process $(Z_t)_{t\in\mathbb{T}}$. Depicted is an ensemble of five exemplary sample paths of the stochastic process over some observation period (Adapted from PETERS and ADAMOU 2018b, p. 11).

The following visualisations are meant to facilitate an intuitive understanding of the two averaging procedures and are not fully rigorous. We consider a random variable Z in discrete time $t \in \mathbb{T} = \mathbb{N}_0$[130] and we assume for the moment also a discrete ensemble or state space of size N which has only a finite or at most a countably infinite number of elements, *i.e.* $Z = \{z_1, z_2, \ldots, z_N\}$. We visualise a generic stochastic process $(Z_t)_{t\in\mathbb{T}}$ to facilitate a better understanding of the nature of this mathematical object.

We can understand any stochastic process to exist in the plane generated by the time domain and a spatial domain, whereby the spatial domain inhabits the realisations of the stochastic process, see **Fig. 2.1**. Therefore, the spatial domain can be called interchangeably the ensemble domain. In **Fig. 2.1** we plot five exemplary realisations of a generic stochastic process over a finite observation period. Any realisation has a history and future beyond the observation window which is indicated by the three dots before the start and after the end of the observation. Using the terminology of ergodic theory, a stochastic process is the ensemble of all its realisations.

[130] As an exception we denote the time variable in this section and in this discrete setting by t, too.

In the following we distinguish between two averaging procedures. A first procedure takes an average along the time evolution of a single sample path of the stochastic process. A second procedure averages over the different possible sample paths of the stochastic process. In **Fig. 2.2** time averaging is indicated by the red shade covering the third sample realisation and ensemble averaging is indicated by the blue shade covering the five depicted sample paths of the stochastic process at a specific moment in time denoted by t^*. Thereby we can utilise the notation of $Z_{\bullet,t^*}$, the mappings were introduced in eq. (2.33) and eq. (2.32).

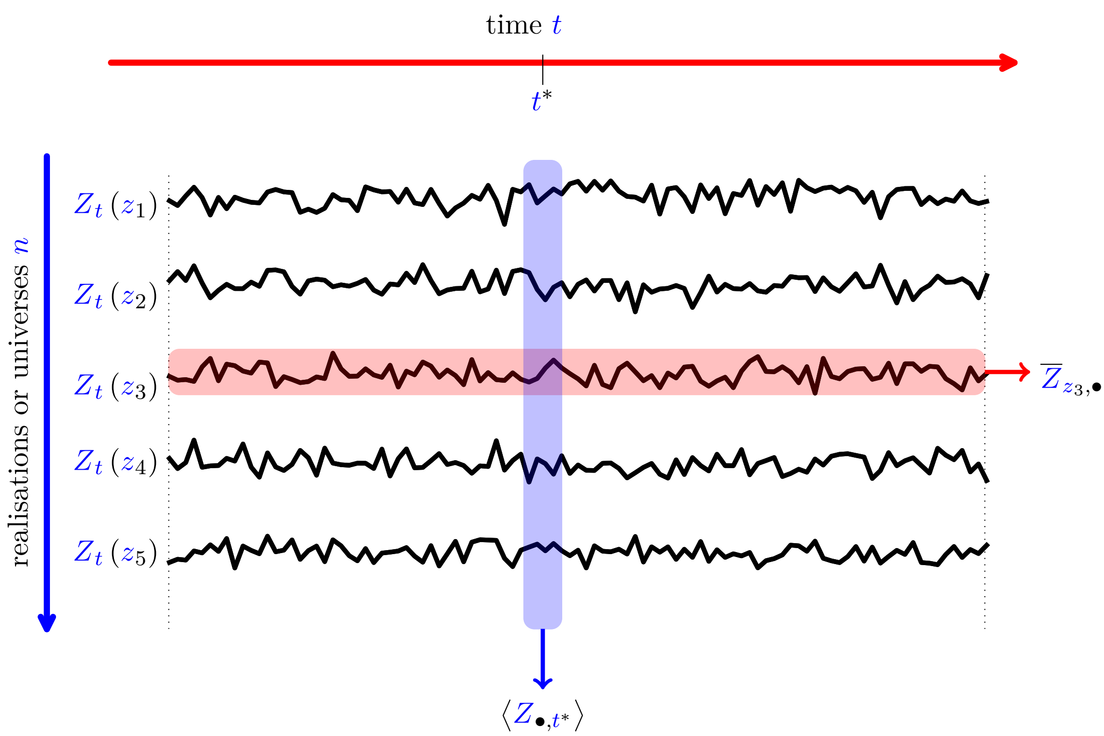


Figure 2.2: Two Averaging Procedures. Averaging over all realisations at a fixed time, *e.g.* at time t^*, yields the ensemble average at that time (blue shaded): $\langle Z_{\bullet,t^*} \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} Z_{z_n,t^*}$. Averaging over one realisation of a stochastic processes, *e.g.* the sample path associated with z_3, yields the time average (red shaded): $\overline{Z}_{z_3,\bullet} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} Z_{z_3,t}$ (Adapted from Peters and Adamou 2018b, p. 22).

The computation of ensemble averages at different moments in time as indicated in **Fig. 2.3** is in principle possible, but should not lead to different results, because our observable is a stochastic process that is a random variable defined in Def. 2.6 on the same probability space indexed by time. However, it may be instructive to think about ergodic theorems as statements about the convergence of all possible time averages and all possible ensemble averages. In fact even more can be said, the convergence of the time average of all ensemble averages and the ensemble average of all time averages to the identical value.[131]

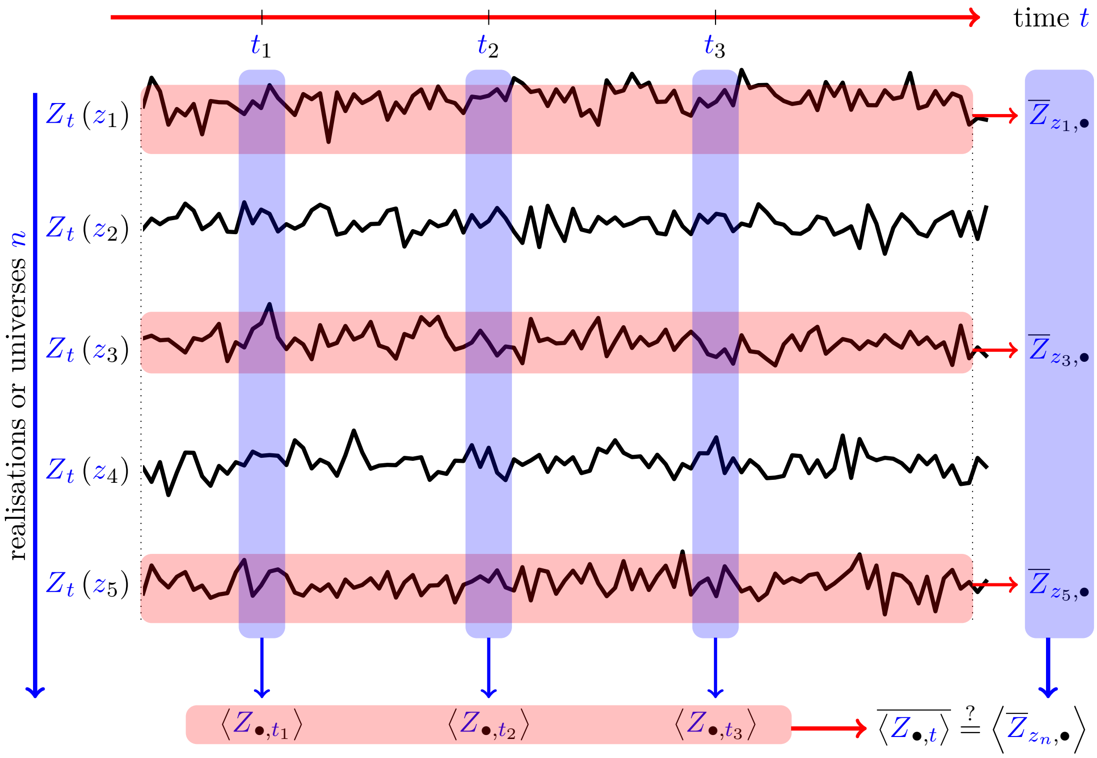


Figure 2.3: Intuition on Ergodic Theorems and Stochastic Processes. Ergodic theorems describe a situation when the time average of all the ensemble averages at all times denoted by $\overline{\langle Z_{\bullet,t} \rangle}$ converges towards the ensemble average at all times of all time averages denoted by $\left\langle \overline{Z}_{z_n,\bullet} \right\rangle$.

2.6 Some Results in Ergodic Theory

The first two proofs of ergodic theorems rendered the ergodic hypothesis mathematically tractable and thus sparked renewed interest. The proofs came from two eminent mathematicians, the mean ergodic theorem by John von Neumann and the point-wise ergodic theorem by George D. Birkhoff and started a vivid branch of mathematical research called ergodic theory. The two proofs and together with other publications supported an emancipation of ergodic theory from its origins in statistical mechanics[132] and shifted the topic sufficiently away from physics to attract interest of further first-rate pure mathematicians.

131 See also Sklar (1993, p. 62) on a related motivation in the context of the nature of measurements in statistical mechanics.

132 See for example Weyl (1916) and von Mises (1920a,b) and the dissertation under von Mises' supervision in Höflich (1927).

In this section we give a very brief and selective overview of some important results in the mathematical branch of ergodic theory. The caveat may be allowed, we do not give any theorems, definitions let alone proofs, rather direct the attention of the interested reader to the special literature. In general the mathematics literature on ergodic theory assumes more or less ergodicity and only to a lesser extent examines the ergodicity problem. Therefore, we review the mathematics literature here and the literature on the ergodicity problem as an open problem in chapter 5.

2.6.1 Metrical Transitivity and Metric Indecomposability

The study of invariant measures has become an active research area of ergodic theory. Especially dissipative systems are of interest to describe not only isolated systems far from reality but more realistic natural phenomena like turbulence with the tools of dynamical systems theory. At around the same time when the first ergodic theorems appeared in print, it became clear that the problem of the ergodic hypothesis in physics is associated with the mathematical property of metrical transitivity of the respective maps or flows.[133] Intuitively, a map or flow is ergodic, if the phase space can not be further decomposed into subsets of the phase space of non-zero measure on which the map or flow has an invariant measure. Then the trajectories over an ergodic region of the phase space are so strongly mixed that they are inseparable and the whole phase space is the ergodic region.[134]

2.6.2 Statistical Mechanics and the Ergodic Hypothesis

The Shift from Classical Mechanics to Statistical Mechanics

About whether or not the objections to the H-theorem led to a shift from a mechanical phraseology in Boltzmann (1868, 1872) to a more probabilistical phraseology in Boltzmann (1877), is still discussed controversially in the literature without any consensus[135] but eventually it led to an advancement of the mathematics in ergodic theory. Before this advancement could really start, two studies suspended it when they proved the impossibility of ergodicity for some mechanical systems[136] and some gas systems[137] and brought the discussion on the foundations of statistical mechanics to a halt.[138]

133 Birkhoff 1928, 1931a, 1932; Hopf 1932; Oxtoby and Ulam 1941.
134 Sklar 1993, pp. 164–167.
135 Von Plato 1994; Renn 2008; Brown et al. 2009; Badino 2011.
136 Plancherel 1913.
137 Rosenthal 1913.
138 Brush 1971.

Two Camps

Singh (2011) identifies at least two camps in the foundation of physics community, for the first camp the ergodic hypothesis is necessary for the foundations of statistical mechanics and for second camp it is not. Since there are strong arguments on both sides[139] the issue is still unresolved, and is therefore keeping physicists, mathematicians and philosophers of science interested in the foundations of statistical mechanics busy.[140]

Why Statistical Mechanics Using Gibbsian Ensembles Really Works?

For Ruelle (1973, pp. 613–614) the role of the ergodic hypothesis for the foundation of statistical mechanics is as follows: 'while one would be very happy to prove ergodicity because it would justify the use of Gibbs' microcanonical ensemble, real systems perhaps are not ergodic but behave nevertheless in much the same way and are well described by Gibbs' ensembles.' Since the debate in Nature (see footnote 95) the ergodic hypothesis is still controversially discussed, partly because it is generally not true and yet statistical mechanics using the Gibbsian ensemble approach works empirically very well.[141] The meaning of the statement 'statistical mechanics works empirically well' is the following. Actual experimental measurements are not done in a hypothetical ensemble but are performed over a finite amount of time and thus time-average quantities are what is measured. The ergodic hypothesis is then one possible justification of the equality of the time-averaged quantities from the measurement by the ensemble-average quantities. Pressure or temperature are macroscopic observables among others. Statistical mechanics seeks to explain the outcome of a measurement of a macroscopic observable which was brought about by molecular dynamics at the microscopic level beneath. Such measurements are surprisingly often correct, which is not yet fully understood. This means not for all cases exist convincing theories of the coarse-graining procedure. If we take for example the case of the observable *pressure of a gas*. Pressure is defined as the rate of energy transfer from the particles to the container walls averaged over an area that is large compared to the distance travelled by a particle before it collides with the wall.[142] The duration taken to measure the time average of the observable *pressure* has to be long in comparison to the microscopic time scale of the collision intervals in that part of the container wall in order to elapse enough time for the convergence to take place. It is here where the ergodic hypothesis is routinely invoked, which justifies the calculation of the

139 For the former camp see Ehrenfest and Ehrenfest (1911) and Khinchin (1949) and for the latter camp see Tolman (1938), Landau and Lifshitz (1980), Jaynes (1971) and Schwartz (2008) among others.

140 Blatt 1959; Brush 1967b; Sklar 1973; Jaynes 1971; Mackey 1974; Friedman 1976; Lavis 1977; Quay S. J. 1978; Leeds 1989; von Plato 1982, 1991, 1994; Navarro 1998; Batterman 1998; Badino 2006; Uffink 2007; van Lith 2001; Hagar 2005; Frigg and Werndl 2011; Werndl 2013; Werndl and Frigg 2015, 2017; Moore 2015.

141 Shlesinger 2000; Werndl and Frigg 2017.

142 If the collision happens between two particles then this distance is called the mean free path (length).

ensemble-average observables instead of the time-average observables. A frequent critique concerns the questionable approximation of the infinite time averages by finite time averages during the measurement process.[143] Following an explanation along the lines of the ergodic hypothesis, however, in order for the convergence of the time-average quantity take place and to really sample the whole phase space would make the measurement duration incredibly large, *i.e.* longer than the age of the universe $\sim 10^{17}$ sec.

Due to such and other critiques the role and the relevance of ergodic theory for statistical mechanics is debated time and again. *E.g.* in the article of Malament and Zabell (1980) titled 'Why Gibbs Phase Averages Work–The Role of Ergodic Theory' , they attribute some explanatory capacity to ergodic theory in its relation to the microcanonical measure, which is why Gibbs ensemble averaging works. On the contrary, Earman and Rédei (1996, Section 4) ascertain the explanatory irrelevance of ergodic theory and entitle their article 'Why Ergodic Theory Does Not Explain the Success of Equilibrium Statistical Mechanics'. They summarise the situation as follows (p. 70): 'all of the debates about how ergodicity would explain why phase averages work is purely academic since most of the systems in question are not ergodic.' In order to reconcile the fact that obviously many systems are not ergodic but statistical mechanics still works Vranas (1998) developed a framework to measure the distance of arbitrary systems to ergodicity, called epsilon-ergodicity. Vranas' conclusions rest upon two facts. First, he finds only tiny distances for many of the studied systems, which implies many of the non-ergodic systems are almost ergodic or epsilon-ergodic and thus statistical mechanics works. Second, some systems approach ergodicity if the degrees of freedom increase. Although this is not in general true, which is exemplified by the FPUT experiment to which we therefore turn in a moment.[144] Though, the exact meaning and implication of systems to be 'just a little bit non-ergodic'[145] is still not convincing and it is unclear when a system is just non-ergodic and not merely epsilon-ergodic with a large ε.

2.6.3 Fermi-Pasta-Ulam-Tsingou Experiment

Intriguingly, the physicist Enrico Fermi started his career with an early interest in the ergodic hypothesis. However, his proof is considered incomplete.[146] Even 30 years later in the 1950s, still not really convinced of the general critique of his proof, he initiated the famous FPUT experiment.[147] The goal of the experiment was to study the behaviour of a continuous system, *e.g.* one-dimensional atoms in a crystal. The authors used a numerical model with $2D$ nearest neighbour interaction between at most 64 points coupled by strings

143 Farquhar 1961, p. 17.
144 To be fair, Vranas himself lists this objection, too.
145 Vranas 1998, p. 694.
146 Fermi 1923, 1924.
147 Fermi et al. 1955.

as an approximation of a continuous lattice, which constitutes the first numerical simulation of a scientific problem.[148] '[T]he partial differential equation defining the motion of this string is replaced by a finite number of total differential equations. We have, therefore, a dynamical system of 64 particles with forces acting between neighbors with fixed end points.'[149] The FPUT experiment was an attempt 'to check the idea that essentially any nonlinearity [in the interaction between system elements] would lead to a system satisfying the ergodic hypothesis. [...] The outcome was quite against the conjectured conclusion and a new era was opened to understand why the tempting explanation of ergodicity as due to "any" (reasonable) nonlinearity failed.'[150] It was expected that thermalisation (the approach to equilibrium due to the mutual interactions) quickly leads to the equipartition of the energy of the system. However, this did not happen. Instead, the experiment showed periodic returns to the original mode. Therefore, the results were 'providing intimations that the prevalent beliefs in the universality of "mixing and thermalization" [and thus the ergodic hypothesis] in non-linear systems may not be always justified'.[151] The result was considered paradoxical, and subsequently called the FPUT paradox, *i.e.* that non-linear interactions do not guarantee equipartition and thermalisation automatically.[152]

2.6.4 Ergodic Chaos and the Study of Billiards

Important progress in ergodic theory from the mathematical point of view has been made with the help of two idealisations in mathematical models of billiards. The particles are modelled as point masses or hard spheres (sometimes disks), that travel along straight lines in a container of some shape until an interaction takes place, either among each other or with scatterers or the container walls. The interactions are modelled as elastic collisions. Similar to a trajectory, the union of all straight lines is called an orbit of the flow of the billiard. As it turned out, the geometry of the container crucially determines whether or not a point mass starting at some initial condition inside the container will follow an orbit that is chaotic.[153] Consequently, the study of billiards gave rise to the notion of ergodic chaos,[154] if there exists a unique invariant measure of a chaotic orbit. Ergodic chaos found applications in economics as well.[155] Although, the majority of economic studies on ergodic chaos do not give advice in the non-ergodic case and are focused on the existence of an invariant and unique PDF (an ergodic probability measure), *e.g.* of an optimal policy function if the economic variable

148 Weissert 1997.
149 Fermi et al. 1955, p. 4.
150 Gallavotti 2008, p. 1.
151 See Stanislav Ulam in the introduction to the reprint of Fermi et al. (1955) in Fermi et al. (1965). See also Ulam (1976, ch. 12).
152 Patrascioiu 1987; Weissert 1997; Dauxois 2008; Bagchi and Tsallis 2018.
153 Wightman 1985.
154 Eckmann and Ruelle 1985.
155 Jakobson 1981; Benedicks and Carleson 1985; Day and Shafer 1987; Carleson 1991.

behaves chaotically[156] or on the existence of ergodic distributions of economic equilibria, *e.g.* in discounted dynamic optimisation models, the logistic map used to explain Cobweb models or sunspot equilibria.[157] The study of billiards began with simple container shapes such as the Lorentz gas,[158] the Sinai billiard[159], the Bunimovich stadium[160], the Bunimovich mushroom.[161] So far the trajectories of the particles had been broken straight lines on flat manifolds.[162] Further studies extended the analysis to the behaviour of orbits on curved manifolds, when the line segments become geodesics.[163]

The typical real-world systems are far from being that simple or in other words molecules are not like ideal hard spheres and only occasionally has the environment in which these molecules move the shape of an ideal stadium or mushroom. But of course it is reasonable to start with the analysis of simple shapes, which were also the simplest idealisations of the gas containers that have been the original interest of Boltzmann and Maxwell. Further progress is based on universal behaviour of diverse shapes that depends only on a limited set of container properties, *e.g.* the number or size of focusing or dispersing elements. Many specific container shapes indeed fall into more general universality classes, which then give rise to specific features of the orbits.[164] *E.g.* Sinai billiards are ergodic and obey even stronger properties in what is called th ergodic hierarchy, which we mention in Subsec. 2.6.7. The boundary in phase space between initial conditions which lead to regular and those that lead to chaotic trajectories is often irregular or even fractal. The regions in phase space of initial conditions that lead to chaotic or periodic orbits can have a fractal dimension, too. Deeply connected to such questions is Kolmogorov-Arnol'd-Moser (KAM) theory and the discovery of so called KAM islands and KAM tori in phase space, which we briefly discuss in the next section.

---

156 Nishimura et al. 1994.

157 Majumdar and Mitra 1994; Araujo and Araujo 2000.

158 Lanford III. 1973; Chernov 1991; Szász 2000b.

159 The Sinai billiard is a square table with a circular scatterer at its centre (Sinai 1963, 1970). The Sinai billiard on a plane leads to the Lorentz gas. See also Altmann (2007).

160 The Bunimovich stadium is a rectangle with two semicircular caps (Bunimovich 1975, 1979; Sinai and Chernov 1987). See Bunimovich stadium for an animation of two balls in a stadium with a small difference in their initial conditions.

161 Bunimovich mushrooms are idealised mushroom shapes with semicircular cap attached to a number of stems of some shape(Bunimovich 2001).

162 The segments between two reflections are also called the free paths in the terminology of diffusion research, see also Subsec. 5.3.1.2.

163 See Hopf (1939, 1971), Anosov (1967), Anosov and Sinai (1967), Bunimovich and Sinai (1973) and Bunimovich (1974, 1990, 2007) and the surveys Hedlund (1939) and Jacobs (1960).

164 Other aspects that complicate the situation are repulsive bouncing, *i.e.* in real collisions additional short-range repelling forces between the molecules take place, which hopefully only have a negligible net effect; gas molecules are not spherical, but vibrate and move with some rotational momentum, thus interactions are not fully elastic; in liquids and solids the dilute gas approach is obviously debatable; and there are small gravitational attracting forces acting between the molecules which create a small long-term coupling (Strevens 2003, pp. 316–319).

2.6.5 KAM Theory

Why statistical mechanics works after all seems to have much more complicated reasons, and most likely KAM theory[165] will play an important role in the explanation of its success.[166] Kolmogorov, Arnol'd and Moser, from whose initials the name of KAM theory is derived, could prove the existence of special islands in phase space from which it is impossible for the system to escape. If the initial conditions are within such islands the systems is trapped. As a consequence, the system will not explore the whole phase space, indicating non-ergodic parts of the flow, so-called invariant KAM tori. The mere existence of such islands has long been known, but the KAM theorem proved that they are not of measure zero, which is a necessary outcome of proofs of ergodic theorems for almost everywhere convergence in phase space except initial conditions of measure zero. It is easy to accept, that the existence of islands in phase space, prevents trajectories from passing through every point in phase space and thus imply non-ergodicity. These KAM tori block certain regions in phase space, but itself become smaller with increasing system size and therefore allow the particles to explore increasing parts of the phase space. This suggests that the role of ergodic hypothesis for the foundations of statistical mechanics is more of the approximate kind in the large system size limit.[167]

If any conclusion could be drawn from the research programme on KAM theory, it would be that Hamiltonian system and their elliptic dynamics are in general non-ergodic. However, results from a parallel research programme on smooth ergodic theory or hyperbolic dynamics continuously find Hamiltonian systems that are ergodic. The programme was initiated by the work on geodesic flows on manifolds with negative curvature,[168] extended to arbitrary dimensions to so called ergodic Anosov systems[169] and their Gibbs measures[170] and further generalised to uniformly hyperbolic systems.[171][172]

2.6.6 The Study of Ergodic Averages

As it is visible from the pointwise ergodic theorem 2.1, ergodic theory in general can also be understood as the study of the convergence behaviour of averages. Here the connection of ergodic theory to its original focus of the ergodic hypothesis in statistical mechanics on the substitutability of ensemble and time averages becomes blurred. *E.g.* because initially

165 Kolmogorov 1954; Arnol'd 1963a,b; Moser 1962, 1966b,a.
166 Dumas 2014.
167 Sklar 1993, ch. 5.II.3; Bais and Farmer 2008, p. 634.
168 Hopf 1939, 1971.
169 Anosov and Sinai 1967; Anosov 1967.
170 Sinai 1963, 1972.
171 Smale 1967; Bowen and Ruelle 1975; Eckmann and Ruelle 1985.
172 Viana and Oliveira 2016, pp. xi–xii.

the change in dynamical systems has been embedded within time, but one can also study the behaviour of an average along a different dimensions than time. In FURSTENBERG (2017) this is done along the degree of magnification of fractal objects, wherein the fractal dimension is understood in the sense given by ergodic averages in an appropriately defined measure-preserving system.[173]

2.6.7 Ergodic Hierarchy

The ergodic hierarchy provides a hierarchy of properties with increasing demands with regard to the system's behaviour of which ergodicity is the broadest property. The hierarchy of properties is the following

$$\text{BERNOULLI} \subset K\text{-system} \subset \text{mixing} \subset \text{weak mixing} \subset \text{ergodic}\ . \tag{2.43}$$

We can think of the ergodic hierarchy as specific properties of the measure of an observable, intuitively let this measure have the initial shape of a ball in phase space but the shape mutates due to the transformations. The system is ergodic, if over the course of time the initial size of the ball maintains its measure even if its shape changes. The set of weakly mixing systems form a proper subset of the ergodic systems, and of those again a proper subset is mixing or strongly mixing, such as the baker transformation or the cream in your coffee. A system is ergodic and mixing if the over the course of time the ball (*i.e.* the measure of an observable of the system) loses its specific shape while it gets homogeneously dispersed over the whole phase space, but the total measure *e.g.* volume of the ball stays constant. Usually physicists analyse systems that are ergodic and mixing, because it is easier to integrate over homogeneous than irregular phase space occupancies and the stronger property of mixing has been thought to be essential before the introduction of quasi-ergodicity, because under a transformation that is mixing every point in phase space is visited asymptotically by a single trajectory. The K-systems[174] are a proper subset of the mixing systems and the property of a system to be BERNOULLI is the strongest statement. For the latter two elements in the ergodic hierarchy no intuitive interpretation is known. *E.g.* the SINAI billiard has the strongest property of being BERNOULLI.[175] For more information start with FRIGG et al. (2014).

[173] See also SCHMELING (2009) on the relation between ergodic theory and fractal geometry.
[174] K-systems go back to KOLMOGOROV and sometimes are referred to as quasiregular systems (SHIRYAEV 2000, p. 56).
[175] SINAI 1970.

2.6.8 Textbooks and Surveys

Besides the classic textbooks on the mathematics of ergodic theory,[176] most modern textbooks on probability theory[177] or dynamical systems theory[178] contain chapters on ergodic theory, often in the context of the LLN for iid random variables. Since then more recent textbooks with the main focus on ergodic theory appeared[179] and some have a more specialised focus *e.g.* on smooth ergodic theory[180] or ergodicity and stochastic processes.[181] *Ergodic Theory in the Perspective of Functional Analysis* is presented in the unpublished manuscripts Derndinger et al. (1987, 1998), which finally evolved into Eisner et al. (2015) with a focus on *Operator Theoretic Aspects of Ergodic Theory*. See especially their extensive bibliography of seminal publications. Einsiedler and Ward (2011) treat ergodic theory 'with a view towards Number Theory'. Choe (2005) is an introduction to ergodic theory by means of computer programmes. The book review in Kitchens (2007) of Choe (2005) is itself an instructive review of the field. Krengel (1985) is a comprehensive book-length survey on ergodic theorems, whereby he generalises the transformations firstly to operators in function space and even further to limits of averages of semigroups of operators, and studies primarily norm and almost everywhere convergence. We like to draw particular attention to the notes at the end of each section, which are sometimes more extensive than the actual section and contain most valuable information. In Sec. 2.4 our exclusive interest was in finite measures and most often even only in probability measures. The scope can be expanded to infinite measure spaces for which the standard reference is Aaronson (1997). Dumas (2014) discusses the ergodic hypothesis and the *The KAM Story* in the context of 'A Friendly Introduction to the Content, History, and Significance of Classical Kolmogorov–Arnold–Moser Theory'. In the study of entropy and large deviations ergodic theorems naturally play a fundamental role.[182]

Besides pure mathematics and theoretical physics, ergodicity is an important topic in information theory and electrical engineering as well. These fields study properties like the entropy and entropy rates of stationary and ergodic stochastic processes as sources of information and the capacities of channels through which the information is transmitted. The stochastic processes are commonly defined on probability spaces with finite-valued sample spaces, which are also called finite alphabets. Standard references are Shannon (1948), Khinchin (1957), Billingsley (1965) and Cover and Thomas (2006). One inportant result in information theory is the Shannon-McMillan-Breiman theorem on asymptotic equipartition (AEP),

[176] Hopf 1937; Halmos 1956; Jacobs 1960; Arnol'd and Avez 1968; Sinai 1977; Cornfeld et al. 1982; Walters 1982; Petersen 1989.
[177] Feller 1957, 1971; Breiman 1992, ch. 6; Klenke 2014, ch. 20.
[178] Arnold 1998; Jost 2005; Schuster and Just 2005; Stoop and Steeb 2006.
[179] Mañé 1987; Silva 2007; Nadkarni 2013; Viana and Oliveira 2016.
[180] Pei-Dong and Qian 1995; Barreira and Pesin 2013.
[181] Lasota and Mackey 1994; Borovkov 1998; Gihman and Skorokhod 2004; Shalizi and Kontorovich 2010.
[182] Petersen 1989; Ellis 2006; Downarowicz 2011.

which is an analogue to the LLN.[183] Billingsley (1999) treats ergodicity in chapter 4 in the context of different modes of convergence of sequences of random variables. Gray (2011b,a) are introductions to *Entropy and Information Theory* and *Probability, Random Processes, and Ergodic Properties* for (electrical) engineers. A related text on ergodicity and information theory is Shields (1996).

Early survey articles on contributions to ergodic theory are Birkhoff and Koopman (1932), Kakutani (1952) and Jacobs (1965). Halmos has produced several nicely written papers on the subject,[184] *e.g.* for a survey on the relationship between ergodic theorems and information theory see Halmos (1961) and Shields (1998), an extensive bibliography is in Halmos (1949), Halmos (1958) is a review of von Neumann's work on measure and ergodic theory, and Halmos (1939) links in particular chapters 2 and 8. Sinai (1989) gives an overview on 'Kolmogorov's work on ergodic theory'. Szász (2000a) is a good overview on the development from Boltzmann's ergodic hypothesis to the theory of ergodic billiards of Sinai, Bunimovich and co-workers. Accounts on the historical development of probability theory and with increasing degrees also of ergodic theory are found in Daston (1995), Krüger et al. (1987a,b), Sklar (1993), von Plato (1994, ch. 3.2) and Masani (1990, especially chapters 9-12).

Many important contributions to ergodic theory have not (yet) found applications in (ergodicity) economics that are necessary for our purpose and are thus far beyond the scope of this thesis, such as the multiplicative ergodic theorem[185] or the ramifications of Szemerédi's theorem[186], and many more.

183 Shannon 1948; McMillan 1953; Breiman 1957, 1960a; Barron 1985; Algoet and Cover 1988.
184 Ambrose et al. 1942; Halmos 1941, 1944, 1946, 1947.
185 Oseledets 1968, 2008; Arnold 1998, ch. 3.
186 See Szemerédi (1969, 1975) and for various ramifications Furstenberg (1977), Furstenberg and Katznelson (1978), Host and Kra (2005a,b), Kra (2005), Green and Tao (2008), Tao (2006, 2007, 2008, 2015) and Austin et al. (2011).

2.7 Conclusion

> Gewiß wird niemand derartige Spekulationen für wichtige Entdeckungen oder gar, wie es wohl die alten Philosophen taten, für das höchste Ziel der Wissenschaft halten. Ob es aber gerechtfertigt ist, sie als etwas völlig Müßiges zu bespötteln, könnte noch fraglich sein. Wer weiß, ob sie nicht doch den Horizont unseres Ideenkreises erweitern und durch Erhöhung der Beweglichkeit der Gedanken auch die Erkenntnis des erfahrungsmäßig Gegebenen fördern?[187]
>
> LUDWIG BOLTZMANN

This chapter's discussion of the ergodic hypothesis and ergodicity started with the physics' origin and presented some mathematical basics of ergodic theory. Nevertheless, the ergodicity problem is of utmost importance to economics. The notions of risk and expectations had been conceptualised around the 17^{th} and 18^{th} century. Around the same time were the origins of probability theory developed, too. Basic concepts like expectation values therefore predate most of the above debate on the ergodicity problem by a century or more. This led to a contingent conceptualisation of randomness in economics in a specific way. Namely the randomness of outcomes of the unique realisation of the factual world is embedded within a virtual ensemble of all possible realisations or parallel universes. The retention of embedding the randomness in the virtual ensemble caused the situation economics and decision making under uncertainty find themselves in and which ultimately led to the ergodicity problem in economics. In contrast to the embedding of randomness within a virtual ensemble, randomness can also be embedded within historical irreversible time. As it turns out, this approach yields helpful guidance for the individual decision maker. To show how this comes about is the focus of the following chapters.

187 BRODA 1955, p. 82.

3 The St. Petersburg Paradox Redux

How wonderful that we have met with a paradox.
Now we have some hope of making progress.

NIELS BOHR

In medieval times, commerce started growing in volume and number of trading partners, and trade expanded over ever greater parts of the globe. Thus, traders were faced with increasing uncertainty of the physical integrity of their overland and seaborne cargo, and the balance of their pending accounts. Naturally the need of supporting (financial) technologies arose. Early on, there circulated several books to facilitate the algebra-intensive trading activities of merchants. Among the widely used handbooks of their times were *Liber Abbaci* by FIBONACCI (1202), *Summa de arithmetica, geometria, proportioni et proportionalità* by PACIOLI (1494) and *Liber de Ludo Aleae*, which was written around 1550 and circulated long before it finally got published posthumously in CARDANO (1633).[188] These books contained roughly the collected knowledge of their times and therefore served as encyclopedias, but were also used as manuals by practitioners. Especially the *Liber Abbaci* teaches 'written procedures of calculation, algebra, and practical mathematics [... that] were known [...] as abaco'.[189] All these books are a rich source of what concerned the people at the time and give insight into the original motivation of their authors.

For example, the *Liber Abbaci* contains 'a wealth of applications of mathematics to all kinds of situations in business and trade, conversion of units of money, weight, and content, methods of barter, business partnerships and allocation of profit, alloying of money, investment of money, simple and compound interest.'[190] PACIOLI's *Summa* had already a stronger focus on mathematics and as the title suggests contains *everything on arithmetic, geometry and certain ratios*, and CARDANO's *Book on Games of Dice* carries in its title first explicit references to gambling. On the other hand, a discussion of gambling like usury often exposed the writer to

188 This can not be a comprehensive list, instead we focus on the subject of the origin of probability theory. Other texts circulating at the time were *e.g.* TARTAGLIA (1556) and FORESTANI (1682).

189 To be explicit here *abaco* means calculated without an abacus, 'A *maestro d'abbaca* was a person who calculated directly with Hindu numerals without using the abacus, and *abaco* is the discipline of doing this' (SIGLER 2002, p. 4; JOHNSON 2015), *abaco* was also the name of the reckoning schools that taught it.

190 SIGLER 2002, p. 5.

the danger of blasphemy. Astute writers could have made use of the shared mathematical structure and may simply dressed gambling problems as problems that naturally come up in commerce. Whereas in the 16th century, when Cardano (1633) circulated, games of chance are accepted to a greater degree, and furthermore, a very modern mindset finds its way into the zeitgeist. Games of chance are seen now as natural models of time-wise condensed commercial activities. In everyday commerce the concept of fair transactions were as important as fairness is today, that is why in Cardano (1633) early accounts of the concept of 'just' or 'fair' games surface.[191] It is an astounding fact, that one of the commerce related problems, called the 'unfinished game problem', treated in Pacioli (1494) and later taken up again in Cardano (1633) became seminal for the foundation of probability as an early ancestor of the St. Petersburg lottery.

3.1 The Unfinished Game

The cover story surrounding the unfinished game problem, treated in Pacioli (1494) and Cardano (1633), deals with an interrupted game and asks, how to split the pot in a fair manner. A similar problem occupied three Frenchmen. The mathematician Blaise Pascal came from a second generation nobility of office family, and had thus access to Paris' bourgeoisie. At some time in his life he met Antoine Gombaud, who is better known as Chevalier de Méré. Being a notorious gambler equipped with the ample amount of recreational time of a nobleman, de Méré approached Pascal with a question about a game of chance, then and now a popular pursuit of men. This launched the famous exchange of letters between Pascal and Pierre de Fermat which marks the beginning of probability theory. Fermat was another first-rate French mathematician and first generation *noblesse de robe*. The problem with which the Italian scholars and the French nobleman wrestled can be put like this[192]:

> Two players, Peter and Paul, agree on a game with a sequence of rounds. In each round they toss a coin. Peter gets a point whenever 'heads' shows up, Paul whenever 'tails' shows up. Both players stake an equal amount of money, say their wager is 50 €. Whoever scores seven points first wins the game which means, he gets the whole pot.
>
> Now, the game is abandoned before it is finished and can not be continued, *e.g.* because the authorities rush in and stop the illegal gambling activity. At the

[191] Schneider 1985, pp. 239–243; Johnson 2015, pp. 48.

[192] There is some evidence that the 'problem of the points' has been known at least 100 years before it was discussed in Pacioli (1494) as the 'problem of division', see Rigatelli (1985) for further information.

> moment of abandonment, Peter has five points and Paul has four points. How to divide the total pot of 100 € in a way both appreciate as *fair*?

This problem became known as *The Unfinished Game*[193] or synonymously as *The Problem of the Points*.[194] Interestingly, at the very beginning of what later evolves into probability theory is an attempt to find a solution to a moral question, a matter of 'ethical business practice [...] embodied in commercial mathematics.'[195] This emergence of mathematical probability from the context of vivid ethical business practice predates the shift of probability towards inanimate relative frequencies due to de Moivre (1718). To put it in the words of Sylla (2003, p. 314), it was 'equity among associates or partners rather than [...] relative frequencies [which] provided the foundation of the earliest mathematical probability theory.'

Eventually, the cover story of the unfinished game encompasses two common cases also relevant today, the fair price or value of the option to buyin or buyout: what is a fair compensation for a business partner who drops out of a joint venture prematurely and what would be a fair price for possible new partners to join the venture if some of the payouts have already accrued. Thus, a seemingly lax game represents a common business decision on how much equity to bring into an or take out of an ongoing joint venture.

Sylla (2003, p. 316) explains how the exact same problem is discussed and encoded in ancient Islamic laws[196] in the context of the distribution of inheritances to descendants, and goes on hypothesising whether this kind of mathematical problems migrated from Islamic to Christian texts. Leonardo Pisano also known as Fibonacci was an extensive traveller and could have got into contact with it. On the other hand, there are reports that Islamic texts underwent several cleansings of blasphemy to which gambling issues clearly belonged to. This could be one possible explanation why Pacioli (1494) always writes about innocuous situations in business, chess or crossbow tournaments and never about dice.[197]

A detailed discussion of the several solution attempts to the problem of the points, however, is not the focus of this thesis. Instead, we continue with the original St. Petersburg problem.

193 Devlin 2009, 2010, who chronicles the exchange of letters between Pascal and Fermat.
194 Or *Das Teilungsproblem*, *e.g.* in Schneider 1985.
195 Sylla 2003, p. 309.
196 It can be also found in Islamic mathematics textbooks such as the one by al Khwarizmi (1915), from which the terms 'algebra' and 'algorithm' are derived.
197 Schneider 1985, pp. 238.

3.2 The St. Petersburg Problem

> We are unwilling to be Paul, partly because we do not believe Peter will pay us if we have good fortune in the tossing, partly because we do not know what we should do with so much money or sand or hydrogen if we won it, partly because we do not believe we ever should win it, and partly because we do not think it would be a rational act to risk an infinite sum or even a very large finite sum for an infinitely larger one, whose attainment is infinitely unlikely.[198]
>
> JOHN M. KEYNES

In 1713 the Swiss mathematician NIKOLAUS I. BERNOULLI posed initially in an exchange of letters with the French mathematician PIERRE RÉMOND DE MONTMORT[199] and later also to his mathematician cousin DANIEL BERNOULLI the following problem.

> "Peter [the banker] tosses a coin and continues to do so until it should land 'heads' when it comes to the ground. He agrees to give Paul [the gambler] one ducat if he gets 'heads' on the very first throw, two ducats if he gets it on the second, four if on the third, eight if on the fourth, and so on, so that with each additional throw the number of ducats he must pay is doubled. Suppose we seek to determine the value of Paul's expectation."[200]

To reformulate the problem in a more modern and neat way[201], see the following box and **Table 3.1**.

St. Petersburg lottery

What is a reasonable fee to charge for a ticket for the following lottery?
A fair coin is tossed.

1. On heads, the lottery pays 1 € and the game ends. On tails, the coin is tossed again.

[198] KEYNES 1921, p. 364.
[199] For the exchange of letters see either DE MONTMORT (1713) or the edited collected works of BERNOULLI (1975, pp. 555).
[200] BERNOULLI 1954, p. 31.
[201] GABRIEL CRAMER, a Swiss mathematician and student of NIKOLAUS BERNOULLI's uncle JOHANN BERNOULLI, simplified the original game of throwing a six-sided die into a game of throwing a two-sided die also known as tossing a coin (BERNOULLI 1975, p. 560).

2. On heads, the lottery pays 2 €, and the game ends. On tails, the coin is tossed again.

3. On heads, the lottery pays 4 €, and the game ends. On tails, the coin is tossed again.

⋮ ⋮

n. On heads, the lottery pays 2^{n-1} €, and the game ends. On tails, the coin is tossed again.

∞ ⋮

Here n denotes the number of consecutive coin tosses until 'heads' appears for the first time. D. Bernoulli worked on this problem, while he was appointed member of the academy of sciences in St. Petersburg (1725-1731) and presented a solution to the academy in the academic year 1730/1731, which later got published as Bernoulli (1738) in the transactions of the academy,[202] for which reason this lottery is called the St. Petersburg problem.[203] In subsequent chapters several transformations of the payouts and the probabilities of the lottery become important, hence it is useful to introduce its basic components here. The payout x as a function of the length of a 'tails' sequence for the St. Petersburg lottery is depicted in **Fig. 3.1a** and denoted by[204]

$$x(n) = 2^{n-1} . \tag{3.1}$$

[202] See Bernoulli 1954, for the standard English translation.

[203] An alternative explanation for the naming is that N. Bernoulli put Paul and Peter in a casino in St. Petersburg in the original description of the problem, but we know of no reference that confirms this story. For further details on access to the preserved complete exchange of letters between Nikolaus I. Bernoulli, Pierre Rémond de Montmort, Gabriel Cramer and Daniel Bernoulli see the introductory remarks by Otto Spiess on p. 557 in Bernoulli (1975, K 9, p. 557–567). Astonishingly, the exchange of letters on the St. Petersburg paradox between Nikolaus and Daniel survived in the collected works of their uncle Jakob I. Bernoulli. Jakob I. Bernoulli is the author of one of the first books on probability theory *Ars Conjectandi* (Bernoulli 1713), which was edited and posthumously published by N. Bernoulli, who had a close relationship. Through the edition, Nikolaus was well acquainted with the knowledge of the time contained in *Ars Conjectandi*, which later motivated the first solution to the St. Petersburg paradox, see Sec. 4.1. The letters of Daniel Bernoulli that contained his progress on the St. Petersburg lottery appear in the collected works of his uncle Jakob Bernoulli although they were intended for Nikolaus. Daniel and his father, Johann I., had a strained relationship that made it impossible for Daniel to address his correspondence directly to Nikolaus I. This made it necessary in order to circumvent his father to address the correspondence officially to the unsuspicious brother of his father Jakob Bernoulli. And it is in Johann Bernoulli's correspondence, especially with his brother Jakob, where the letters therefore survived and are edited in volume three of five on probability of Jakob Bernoulli's œvre (Bernoulli 1975). See also NZZ (2002).

[204] For convenience and better aesthetics we will in later chapters sometimes prefer to write x_n instead of $x(n)$, because otherwise we had to write unsightly statements like $u\big(x(n)\big)$ all the time. For the moment the notation $x(n)$ makes explicit the dependence of the payout on the number of coin tosses.

The probability p in the St. Petersburg lottery for a sequence of 'tails' before the stopping condition 'heads' appears at the n^{th} coin tosses is depicted in **Fig. 3.1b** and denoted by[205]

$$p(n) = \left(\frac{1}{2}\right)^n = \frac{1}{2^n} = 2^{-n} \; . \tag{3.2}$$

The equations 3.1 and 3.2 can be combined to the probability of the specific event $x(n)$ happens as a realisation of the random variable X and denoted by

$$\Pr\left\{X = x(n)\right\} = \Pr\left\{X = 2^{n-1}\right\} = 2^{-n} = p(n) \; . \tag{3.3}$$

By eq. (3.3) the St. Petersburg lottery is completely specified. For graphical overview of payouts and probabilities depending on the number of coin tosses involved in the St. Petersburg lottery see also **Fig. 3.1**. **Table 3.1**. Now that we introduced the St. Petersburg lottery, the next step consists of a derivation of what would be reasonable price for a ticket of this lottery. This is done in the next section and we present what can be considered paradoxical about it. From stochastic point of view, we can already infer from eq. (3.3) that the random variable X follows a geometric distribution, in other words X is a power-law distributed random variable with a diverging first moment, see especially **Fig. 3.1b**. This divergence caused historically a lot of trouble.

3.3 The St. Petersburg Paradox

> Gambling is the act of exchanging something small and certain for something large and uncertain.[206]
>
> WILLIAM A. WHITWORTH

> Investment is exchanging one random variable for another.[207]
>
> THOMAS M. COVER

---

205 Similar like for the payouts, for convenience we will in later chapters sometimes prefer to write p_n instead of $p(n)$. For the moment the notation $p(n)$ makes explicit the dependence of the probability on the event of a certain number of coin tosses.

206 WHITWORTH 1886, p. 208.

Table 3.1: Structure of the complete St. Petersburg lottery. Overview of the payout $x(n)$, and expected payout $\mathrm{E}[x]$ dependent on when 'heads' appears for the first time at the n^{th} toss and their respective probabilities $p(n)$ in the St. Petersburg lottery. The last line summarises the lottery.

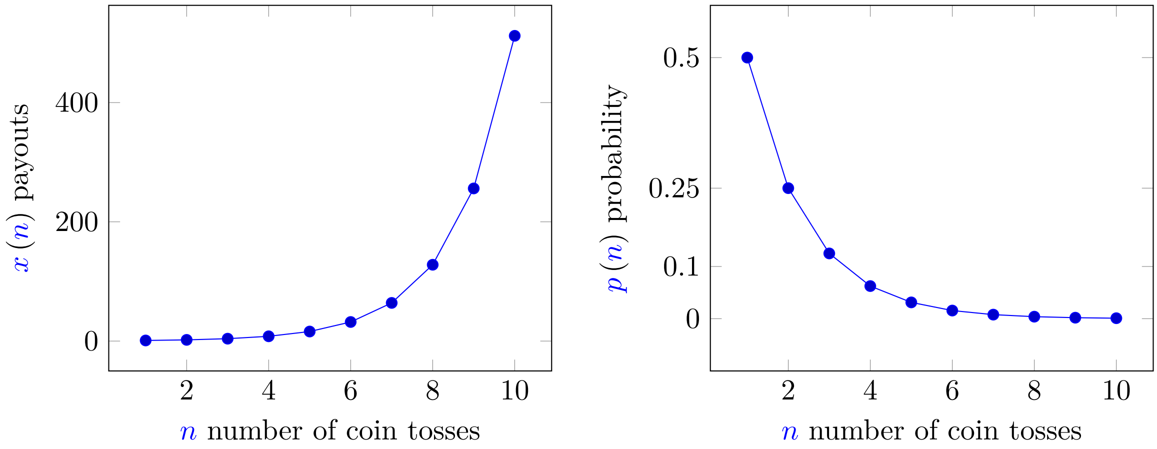
n	$p(n)$	$x(n)$	$\mathrm{E}[x]$
1	$2^{-1} = 1/2$	$2^{1-1} = 1$ €	0.5 €
2	$2^{-2} = 1/4$	$2^{2-1} = 2$ €	0.5 €
3	$2^{-3} = 1/8$	$2^{3-1} = 4$ €	0.5 €
4	$2^{-4} = 1/16$	$2^{4-1} = 8$ €	0.5 €
5	$2^{-5} = 1/32$	$2^{5-1} = 16$ €	0.5 €
$\vdots$	$\vdots$	$\vdots$	$\vdots$
10	$2^{-10} = 1/1024$	$2^{10-1} = 512$ €	0.5 €
$\vdots$	$\vdots$	$\vdots$	$\vdots$
100	$2^{-100} = 1/2^{100}$	$2^{100-1} = 1.27 \cdot 10^{30}$ €	0.5 €
$\vdots$	$\vdots$	$\vdots$	$\vdots$
1000	$2^{-1000} = 1/2^{1000}$	$2^{1000-1} = 1.07 \cdot 10^{301}$ €	0.5 €
$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	$1/2^N$	2^{N-1} €	$\sum_{n=1}^{N} 2^{n-1}\frac{1}{2^n} = \frac{N}{2}$

In order to derive a reasonable price Christiaan Huygens advocated for the – what seems in modern days self-evident – use of the expectation value as the proper decision criterion in risky situations:

> "if someone hides three similar coins in one hand and seven in the other, without my knowing which hand has which amount, and he offers to let me have the money in whichever hand I choose, then this offer is worth as much to me as if he gave me five coins."[208]

Let us follow his advice and see what constitutes the St. Petersburg *paradox*. That the

[207] Cover 2012.
[208] Huygens cited in Sylla 2003.



(a) Payout Function of the St. Petersburg Lottery. **(b)** Probability Distribution Function of the St. Petersburg Lottery.

Figure 3.1: Payout Function and Probability Distribution Function of the St. Petersburg Lottery. The payout function $x(n)$ (a) and the probability distribution function $p(n)$ (b) for the original St. Petersburg lottery for up to $n = 10$.

expectation value procedure actually calculates ensemble averages is the root cause of the whole confusion. The scientific consensus in premodern probability theory determined a reasonable cost of a lottery ticket as that price which equals the expected return $\mathrm{E}[\text{return}]$ from the lottery[209]

$$\mathrm{E}[\text{return}] = C_{\text{fair}} \;. \tag{3.4}$$

Interestingly, the moral concept of fairness was thereby given a mathematical meaning involving the pricing of a random variable which is based on an ensemble average.[210] Equation (3.5) shows, how the expected return depends on the expectation of the difference between the payout x and the cost C.[211] The value of the fair price is our ultimate desideratum needed to evaluate if one should enter a particular gamble or not. Therefore the fair price has to be known already when the gambler receives the offer to enter the lottery for an arbitrary asked price. The gambler then bases his decision to enter or not in comparing the asked price with the fair price. For this latter reason the asked price is not uncertain but a constant and can be taken out of the expectation operator $\mathrm{E}[\cdot]$ in the following equation,

$$\mathrm{E}[\text{return}] = \mathrm{E}[x - C] = \mathrm{E}[x] - C \;. \tag{3.5}$$

The following quations (3.6 - 3.7) particularise the exact computation and the composition of

209 See *e.g.* Feller 1957, ch. X.3; Chernoff and Moses 1986, p. 97.
210 See also the remarks on the role of morals in Peters (2011c, p. 4914).
211 Some authors prefer the word ‘winnings’ and not ‘payout’ to chime in with the time-honoured notation emphasising the gambling origin of much of early probability theory, see *e.g.* Peters (2011c).

the expectation value of our observable *payout*[212]

$$\mathrm{E}[x] = \sum_{n=1}^{\infty} \text{payout in state } n \cdot \text{probability of state } n$$

$$= \sum_{n=1}^{\infty} x(n) \cdot p(n) = \sum_{n=1}^{\infty} 2^{n-1} \cdot 2^{-n} = \sum_{n=1}^{\infty} 2^{n-1} \cdot \frac{1}{2^n} \tag{3.6}$$

$$= \underbrace{\underbrace{1\frac{1}{2}}_{=\frac{1}{2}} + \underbrace{2\frac{1}{4}}_{=\frac{1}{2}} + \underbrace{4\frac{1}{8}}_{=\frac{1}{2}} + \cdots + \underbrace{\cdot}_{=\frac{1}{2}} + \cdots}_{\infty \text{ many times}} \quad , \tag{3.7}$$

which basically shows, that the expectation value is computed as a sum of the state-dependent payouts weighted by the probability of the respective state.[213] The states differ in their state-dependent stopping condition, which uniquely identifies them. The stopping condition is the first appearance of 'heads' in a sequence of coin tosses. The state n is uniquely identified as a sequence of coin tosses, that showed 'tails' on every coin toss up to $(n-1)^{\text{th}}$ and the stopping condition, 'heads', on the n^{th} and thus last coin toss.[214] Let $N \in \mathbb{N}$ be an arbitrary but finite natural number. Then in principle, the stopping condition could never be reached for any finite number of coin tosses, N, and this is what makes the lottery actually tempting, because this entails the chance of an infinite payout. By agreement the payout in this case is simply that of the state with the maximal payout of all the N distinct states, which usually amounts to the last term in the sum, 2^{N-1}. Without any loss of generality, 'heads' appears for the first time after some N consecutive coin tosses took place, hence the payout in every single state, n, amounts to 2^{n-1}. As every single product of the infinitely many in the sum in eq. (3.7) equals 1/2 the infinite sum is diverging, see also **Table 3.1**. Hence, equation (3.8) continues where eq. (3.7) has stopped

$$\mathrm{E}[x] = \sum_{n=1}^{\infty} 2^{n-1} \cdot 2^{-n} = \sum_{n=1}^{\infty} 2^{n-1-n} = \sum_{n=1}^{\infty} 2^{-1} = \sum_{n=1}^{\infty} \frac{1}{2} = \infty \cdot \frac{1}{2} \to \infty. \tag{3.8}$$

Thus in general the expectation value of the observable payout – that is a random variable – diverges, but this is not what created the paradox, to which we turn to in a moment. Following premodern probability theory and the associated premodern decision theory, an infinite expected return of a game, $\mathrm{E}[\text{return}] \to \infty$ as in eq. (3.8) is sometimes treated as

212 The 'expectation value of a quantity' is interchangeable to the 'expected quantity'.

213 For the term 'state' exist several synonyms like 'state of nature', 'state of the world' or more commonly 'scenario'. See Subsec. 3.5 for a brief discussion of the synonyms and their origins.

214 Some authors, *e.g.* Feller (1957), do not follow the original description of the St. Petersburg lottery, which uses the length of the 'tails' sequence. Instead these authors track the first appearance of 'heads'. Thus the event *'heads' appeared at the first toss* marks the case of $n = 1$ for them and a sequence of 'tails' of length $n - 1 = 0$ in our version. In their version of eq. (3.7) the terms of the sum all equal unity. Ultimately, the expectation value of both versions diverges and they are equivalent.

if it were a number.[215] This implies it is reasonable to enter the game as long as the ticket cost is finite, because the statement $\infty - C = \infty$ is considered sensible. See for example how a standard textbook on decision theory by Chernoff and Moses (1986, p. 98) put it:

> "If a game is 'favorable' from the point of view of the expectation [value] of money and you have the choice of repeating it many times, then it is wise to do so. For eventually, your amount of money and, consequently, your utility are bound to increase (assuming that utility increases if money increases)."

For a finite sequence of coin tosses of length N the expected payout is

$$\mathrm{E}[x] = \sum_{n=1}^{N} \frac{1}{2} = \frac{N}{2}. \tag{3.9}$$

Basic Structure of the *Problem* of Lotteries with Infinite Expectation

In common parlance, the basic structure of the paradox is: *What is the value of a lottery ticket, offering a tremendous profit with a very low chance?* and thus a question of the correct valuation of the lottery. From a probabilistic point of view, a lottery is an exchange of two random variables, the payout and the ticket price, whereby the latter is a degenerate one. More precisely, lotteries of the St. Petersburg type involve the treatment of random variables with an infinite expectation.

Valuation problems as the one just presented are fundamental in economics and finance, as can be seen by the impact of the discovery of the Bachelier-Thorp/Black-Scholes-Merton formula for the valuation of options.[216] Standard decision theory should be a guide towards the optimal behaviour for such investment decisions. Furthermore, problems of these kind have to deal with a high level of uncertainty regarding events associated with very low probabilities of occurrence (so-called rare or extreme events), that often have a huge impact once they realise.[217] On the other hand, rare events are by definition rare and therefore difficult to estimate in a reliable way. This has led several authors in the field to question the general possibility of constructing valid statistics of rare events at all.[218]

215 Feller 1945, p. 302.

216 See Bachelier (1900), Thorp (1969), Black and Scholes (1973) and Merton (1973), the results of Bachelier & Thorp in the context of non-dynamic hedging allow for any type of distribution and are not restricted to well behaved distributions where all higher moments must exist like for the Gaussians, Thorp (1973) will be insightful as well. For a discussion on the priority issue of option pricing formulas see Schachermayer and Teichmann (2008), Haug and Taleb (2011) and Thorp (2017, pp. 166–184), as well as Hafner and Zimmermann (2009).

217 See for example the literature on natural catastrophes like Albeverio et al. (2006) among others.

218 Taleb 2005, 2010a,b; Johnson 2016.

Paradoxical Empirical Behaviour

Empirically, *i.e.* if people were offered St. Petersburg lotteries, they quickly reject the lottery before the ticket price reaches double digits. This is the reason why the St. Petersburg *problem* became the St. Petersburg *paradox*. This illustrates that the St. Petersburg paradox is not paradoxical on logical grounds per se. Nothing is paradoxical about random variables with infinite expectation, indeed they can be considered the rule rather than the exception, especially in economics.[219] The paradox arises because a normative decision theory does not predict well the empirically observed behaviour. Thus the descriptive and normative aspects of the (so far presented) decision theory are in conflict. Our focus is on finding a more adequate rational theory of decision making under uncertainty. However, the conflict between descriptive and normative domain of decision theories has motivated many empirical studies especially in psychology and the related field of behavioural and experimental economics, on which we briefly comment in the next section.

3.4 Empirical and Experimental Evidence

As already mentioned, real participants in experiments that involve St. Petersburg lotteries are not willing to pay more than a small amount, e. g. 5 €, for a lottery ticket, although the criterion of a positive expectation value tells one to play for any finite fee. In experimental studies situations are designed where the participants usually have the choice between some uncertain payouts from lotteries or certain payouts. Subsequently, their observed decision making behaviour is analysed, *e.g.* the price at which the change their behaviour from accepting to rejecting the offer. Of course, experimental studies have the drawback that one can never study the lotteries with an infinite expectation for the obvious reason that the supplier – the scientific lab – faces an unlimited downside risk of having to pay 2^{n-1} €, in case the experimentee is lucky and achieves a long series of successive tail throws. However, several clever study designs have been developed, but are not the focus of our thesis, which is why we only refer to them and some numerical simulation studies here.[220] These experimental studies will not enlighten the mathematics of the problem in decision theory directly, but can only hint at inconsistencies. Our goal is finding a better mathematical representation of the dynamics of decision making and in general the dissemination of the results of *Ergodicity Economics* for a better normative treatment of decisions under uncertainty which involves random variables irrespective of the existence of certain finite moments. In the end, we believe a better theory on normative grounds must pass the empirical test of observed behaviour

[219] See especially the work by Mandelbrot on α-stable distributions (Pareto 1964; Mandelbrot 1987; Mandelbrot 1997; Mandelbrot and Hudson 2004; Nolan 2018).

[220] Bottom et al. 1989; Erev et al. 2008; Hayden and Platt 2009; Neugebauer 2010; Cox et al. 2009, 2018; Silva 2016.

in the lab or more importantly in the real world. Otherwise the question arises: How to refute a normative theory at all? Hereafter, we briefly discuss the first empirical analysis by Buffon in the following paragraph and subsequently scrutinise the computation involved in the expectation operator.

Buffon's Empirical Investigation of the St. Petersburg lottery

In an exchange of letters Cramer – who will shortly reappear – draw the attention of his contemporary French scientist Georges-Louis Leclerc better known as the Comte de Buffon to the St. Petersburg lottery. Buffon conducted for the first time an empirical analysis. Therefore, he used a very special kind of a random number generator, namely he let a child play the game 2048 times. Buffon recorded the results and calculated the empirical expectation value of the lottery for his observation (see **Table 3.2**).[221]

This first empirical study by Buffon is remarkable. He concluded that 5.5 is a reasonable value of the expected payout of the lottery. However, we can not see from **Table 3.2** if and how the total payouts (and associated with it the sample average payout per game) depends on the number of total games played. For this reason **Table 3.3** is more informative. Here we stated Buffon's average payout per game for an increasing number of games k played. We can see the expected result of an increasing payout per game, which settles at 1/2 for every power of two.

As we will see, it is the particular way of computing the mathematical expectation in the decision criterion as an ensemble average, that creates much confusion in the conception and embedding of randomness in general and makes it impossible to resolve the originally stated St. Petersburg paradox in particular. The solution strategies to it are discussed in detail in chapter 4.

221 Reproductions of Buffon's tables can be found in Jorland (1987, Table 1, p. 168) and Dutka (1988, Table 1, p. 35).

Table 3.2: Buffon's empirical frequencies of game lengths for 2048. Buffon's experiment for 2048 games of the St. Petersburg lottery.

Sequence $T^{n-1}H$	Game length (in tosses) n	theoretical payout $x(n)$	theoretical frequency	BUFFON's empirical frequency	BUFFON's empirical payout
H	1	1	1024	1061	1061
TH	2	2	512	494	988
T^2H	3	4	256	232	928
T^3H	4	8	128	137	1096
T^4H	5	16	64	56	896
T^5H	6	32	32	29	928
T^6H	7	64	16	25	1600
T^7H	8	128	8	8	1024
T^8H	9	256	4	6	1536
T^9H	10	512	2	0	0
$T^{10}H$	11	1024	1	0	0
Total number of games			2047	2048	
Total payouts		11 264			10 057
Average payout per game		5.5			4.91

Notes: The sequence is denoted by how many times n in a row 'tails' = T appeared, *e.g.* T^3H denotes the sequence TTH. In 2048 theoretical games one game is not counted, because it did not end in time and the sequence consists of only 'tails' tosses. This sequence would need an additional apriori agreement on the payout associated with this sequence. Though with the number of games increasing the one omitted/unfinished game has only a decreasing effect on the (sample) average payout per game. Games of length longer than 8 did not occur. The theoretical absolute frequency of a sequence of length n in column 4 is given by the product of the true probability and the number of games, $p(n) \cdot 2048$.

Table 3.3: Total and average payout per game for different observation lengths

Number of Games	2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}
k	2	4	8	16	32	64	128	256	512	1024	2048	4096	8192	16384
	1	2	4	8	16	32	64	128	256	512	1024	2048	4096	8192
		1	2	4	8	16	32	64	128	256	512	1024	2048	4096
			1	2	4	8	16	32	64	128	256	512	1024	2048
				1	2	4	8	16	32	64	128	256	512	1024
					1	2	4	8	16	32	64	128	256	512
						1	2	4	8	16	32	64	128	256
							1	2	4	8	16	32	64	128
								1	2	4	8	16	32	64
									1	2	4	8	16	32
										1	2	4	8	16
											1	2	4	8
												1	2	4
													1	2
														1
Games Count	1	3	7	15	31	63	127	255	511	1023	2047	4095	8191	16383
Total Payout		4	12	32	80	192	448	1024	2304	5120	11264	24576	53248	114688
Payout per game	1	1.33	1.71	2.13	2.58	3.05	3.53	4.02	4.51	5.00	5.50	6.00	6.50	7.00
Diff		0.33	0.38	0.42	0.45	0.47	0.48	0.49	0.49	0.50	0.50	0.50	0.50	0.50

Notes: For k games there is always one sequence that consists of only 'tails' for which the game did not end yet, therefore the counted games in the fourth last row arise as a result from the number of games in the first row decreased by one, $k-1$. In the infinite game limit, $k \to \infty$, the marginal payout per game (and therefore the expected payout of finite games of the St. Petersburg lottery) approach 0.5 €, *i.e.* the expected payout grows by 1/2 for every increase in the power of two of the number of games, $\lim_{k\to\infty} \frac{\mathrm{d}\,\mathrm{E}[x]_k}{\mathrm{d}n} = \frac{1}{2}$, where $\mathrm{E}\left[x\right]_k$ denotes the expected payout for k games.

3.5 The Role of the Ensemble for the Expectation Operator

As already mentioned, real participants in experiments that involve St. Petersburg lotteries are not willing to pay more than a small amount, e. g. 5 €, for a lottery ticket. This hints at a weakness of the decision theory which considers the calculation of the expected return as meaningful. Calculating the expected return seems no to be what people do. In this section we show why people's intuition is by no means wrong-headed, let alone a manifestation of irrational behaviour on their side. Therefore, a look at the calculation of the expected payout in eq. (3.7) will be instructive. We scrutinise the role of the different states. What are the states in eq. (3.7)? Every single state can be interpreted as a possible future state of the gambler. In this sense all the possible future states are lined up like pearls on a string in eq. (3.10) and can thus be interpreted as parallel worlds at a fixed moment in time (the moment after the lottery is played) in the future over which the averaging is carried out,

$$\mathrm{E}\left[x\right] = \overbrace{\underbrace{1\frac{1}{2}}_{\text{world}_1} + \underbrace{2\frac{1}{4}}_{\text{world}_2} + \underbrace{4\frac{1}{8}}_{\text{world}_3} + \cdots + \underbrace{2^{99}\frac{1}{2^{100}}}_{\text{(lucky) world}_{100}} + \cdots}^{\text{ensemble average over all possible parallel worlds}} \tag{3.10}$$

$$= \sum_{n=1}^{\infty} \text{payout in world } n \cdot \text{weight of world } n \tag{3.11}$$

$$= \text{ensemble average} \quad . \tag{3.12}$$

What is made explicit here in equations (3.10 - 3.12) is the absurdity of the ensemble perspective. Now it is easy to see, that the ensemble average acts as if the gambler in lucky world_{100} wins 2^{99} € and interacts with his other copies in the virtual ensemble, which leads to the situation that his enormous payout superimposes all the other less lucky copies of the gambler in the other parallel worlds. Furthermore, it seems as if the parallel counterfactual copies of the gambler could pool and share all their payouts in order to make the average of them a physical relevant quantity for an individual gambler, for whom decision theory shall be a guide. Such an interaction is against the initial no-interaction principle on which the conception of an ensemble was built.

The key point is easily missed, because the following may appear remote from traditional economic reasoning. It helps to be precise on what is actually written in equations (3.5 - 3.11). $\mathrm{E}\left[r\right]$ is calculated, as if there is an interaction within parallel worlds. The expected return, $\mathrm{E}\left[\text{return}\right]$, is calculated, as if the payout from one lucky world, *e.g.* world_{100}, could compensate for all other worlds below the break-even point. Thus, a former problem in the realm of decision theory is treated, as if it were a mere portfolio problem. As if any participant in the lottery could split his ticket to buy at least one lucky world besides all the other unfortunate worlds. To be distinct, this clearly is a violation of space and time, in which the original problem was proposed. The use of the ensemble though may seem elegant, but can't be a proper solution to the paradox at least for two reasons.

1. First, we live in time! Which means, we live in just one universe. If some other copy of ourselves in a parallel universe wins a huge amount (because his world is for example lucky world_{100}) is of no importance to us, since we can not interact with this copy of ourselves and ask him to please share his/our money. Therefore, we want to know, whether we should partake in the lottery once or several times in a row in one and the same universe.

2. Second, if we use the concept of a multiverse, then we also have to accept the no-interaction premise between parallel worlds. They are independent copies, see definition 3.1, if pooling of resources occurs they are no longer independent.

If we assume the cardinality of the ensemble to be infinite, as in equations (3.5 - 3.12), say infinitely many parallel worlds, it is immediately visible, no matter how high a ticket price C may be, it can not exceed a diverging sum, However, in almost all real-world situations the ensemble is a finite set. One reason among others, eventually there is a limited amount of money on the planet.

In case of the St. Petersburg paradox, it could be simplified in assuming the ensemble being a finite set, not only because there is a limited amount of money on the planet, which means, both the provider of the lottery (banker) and the gambler, accepting the wager, posses only a finite amount of money. Hence, the provider can not afford to pay every possible winnings, and the gambler cannot afford every ticket price. A provider of this lottery, say a casino, would of course charge a huge ticket price, hence, not everybody could partake. But if willing participants could borrow short term, they can always argue to the creditor in their own favour with the unlimited expected return and therefore should always obtain credit. If we assume an ensemble of finite cardinality $n^{\max}$, this is equivalent to a finite possible wealth of the provider. *E.g.* in the luckiest world $n^{\max} - 1$ consecutive 'tails' throws took place before the first 'heads' appeared, leading to a maximal winning of $2^{n^{\max}-1}$ €. Besides, a finite ensemble set is conceivable, $n^{\max} - 1$ could define the maximum wager the provider is willing to stake. On the other hand the gambler could be conscious that the supplier's wealth is limited, too, because the supplier has an incentive to offer the bet for a ticket price that matches his wealth level.

To drive the message of this absurdity home come what may, we present a lottery that is exaggerated to the extreme but not completely unrealistic. Therefore, we change the payout scheme in the following way, the payout in every state is -1 €, *i.e.* the gambler has to pay the supplier of the lottery 1 €, except if one tosses 100 times in row 'tails'. In this and only in this case the gambler gets the payout of $2^{99\,999}$ €. If the gambler and the supplier agree on the length of the round of not more than 100 tosses, we get

$$\mathrm{E}[x] = \underbrace{-1\frac{1}{2}}_{\text{world}_1} + \underbrace{-1\frac{1}{4}}_{\text{world}_2} + \underbrace{-1\frac{1}{8}}_{\text{world}_3} + \cdots + \underbrace{2^{99\,999}\frac{1}{2^{100}}}_{\text{(extremely lucky) world}_{100}} . \tag{3.13}$$

Presented this way we see that it is easy to construct a lottery with a positive expectation value if only we inject a large enough payout of a gazillion euros – completely irrespective of its (im)probability, if $1/2^{100}$ is too unlikely simply increase the power of the payout. At this point we refer to the statements of the virtuality of the ensemble in the context of statistical mechanics on page 210.

Now we reiterate what has been stated in throughout chapter 2 in a more applied context. What the calculation of the expected return actually computes, is the average of all realisations of a random variable, *i.e.* the average over all possible worlds, which are possibly (countably) infinitely many, like in the example of the St. Petersburg lottery, $n \in \mathbb{N}$. Using the language of ergodic theory, introduced in chapter 2, that is called an ensemble average of an observable, in this case an ensemble average of the payouts. Or put differently, in order to emphasise the specific way the expectation value is computed: the expected payout is computed as an ensemble average over parallel worlds. Associated with the expectation operator is the implicit embedding of the randomness of the outcome within a virtual ensemble. The distinction between formalising expectations as ensemble or as time averages has been brought forward for the first time in Peters (2011c) also in the context of the St. Petersburg lottery. Peters (2011c) recognises that it is very natural that a gambler's wealth dynamic is non-ergodic – for the St. Petersburg lottery in particular but also for most wealth dynamics in general. The following terms can all be used synonymously for the ensemble:

- states,
- states of the world,
- states of nature,
- state space (common in economics),
- phase space (common in physics, S in our notation),
- sample space (common in mathematics, often denoted by Ω),
- scenarios,
- parallel universes,
- multiverse or
- parallel worlds.

However, if the observable is not ergodic, then mathematical robust statements about the ensemble are either much harder to derive or such statements about the ensemble simply do not matter too much, because then the evolution of the system over time is not necessarily depicted by any of the snapshot ensembles. We cast the meaning of an ensemble in the following definition.

Definition 3.1 (Ensemble). *Given a random experiment, the phase space or sample space, denoted by S, contains all possible outcomes of a random variable and is called the ensemble or the ensemble set. The individual elements of the ensemble are independent from each other. The cardinality of the ensemble set can be infinite.*

In order to emphasise the absurdity of interaction with virtual parallel universes, we will often use the term parallel worlds or even parallel universes when having the ensemble in mind. Given the above list, the term somebody uses thus tells more about his background and the name of the structure in which he conceptualises randomness, than it gives away information about the subject itself.

From the independence of the elements in Def. 3.1 follows, that the elements cannot interact with each other. The ensemble must not be connected to an actual random experiment but can be connected to a random experiment perceived as a Gedankenexperiment, like it was done in the context of the ergodic hypothesis by Maxwell, Boltzmann and Gibbs in Sec. 2.2 and Sec. 2.3. The individual elements of the ensemble can thus be perceived as being part of parallel universes that are connected via the conceptual idea of an ensemble. The elements of the ensemble can be interpreted as independent copies of each other that differ in exactly one characteristic, *i.e.* in statistical mechanics their microconfiguration. Often in economic applications and for the case of the St. Petersburg lottery as well, the individual elements of the ensemble differ only in the payout that realised as a result of the random experiment. It is in this sense that we speak of the possible payouts as parallel worlds. The idea of parallel worlds is (physically) meaningless to everybody who is bound to live in just one world, in which several realisations of random experiments materialise as a sequence of events over time. The realisations of a random experiment, that may have materialised in all other parallel worlds but the one we live in, are in this sense irrelevant to us who live in just one world and not across many parallel worlds.

Modal Logic and Contingency

In reality and in a random experiments the future state of world or the system under study is only rarely fully predetermined, instead many future states are possible, but an arbitrary single possible future state need not occur necessarily. Such circumstances go beyond the scope of simple (mono-)causality and thus are beyond the usual operators of a binary logic, *i.e.* among others condition A caused *later* state B, $A \to B$, or condition A did not cause *later* state B, $A \nrightarrow B$.[222] Nevertheless, such possible, but not necessary developments can be treated formally within the framework of modal logic. Central concepts of modal logic are the logical consequences of the terms 'possible' and 'necessary', and the associated operators, *i.e.* $\Diamond e$ is used to express an event e is possible, and to denote if an event e is necessary $\Box e$ is used. Some future states that are *possible*, but do *not necessarily* need to become the subsequent state, are called *contingent*, formally this would be denoted by $\Diamond e \wedge \Diamond \neg e$, *i.e.* both e and not e are possible.[223] In a sense modal logic would provide a language for the analysis of the evolution of historical processes through contingent states complementary to the language of stochastics.[224] This thesis employs the latter.

States as Parallel Worlds

Following the notion of different states as parallel worlds, every different state is contingent, *i.e.* it realises or could realise in a different realisation of the random experiment. This is the reason why this state is one element of the sample space, which is in probability theory usually denoted by Ω. In other words, the sample space is the union of all possible outcomes or the sample space is the union of all parallel worlds. The same general idea has many different names in different scientific disciplines. In economics and game theory the convention is to call this the *state space*. In finance, the term *scenario* is often used, *e.g.* the standard case of two scenarios, the first scenario of the booming economy and the second scenario of an economic recession. The same term is sometimes used in mathematics and physics, too. More common is the term *phase space* to denote the space of all possible states a (dynamical) system can adopt.

In the foundations of physics interesting debates on methodological questions arose especially in the interpretation of quantum mechanics and around possible grand unified theories (GUTs) of quantum (field) theory and the general theory of relativity.[225] These debates are of great significance also for the methodology of economics. What should bring the two distinct disciplines closer together in their methodological discussions are the relevance of stochasticity, contingency and historicity in both of their research subjects. Some physics theories are based on the conception that the contingent evolution of the very small as well as on the cosmic scale could be conceived of as a branching into a myriad of parallel universes

[222] It may be noted, that the logical operators, $\rightarrow$ or $\nrightarrow$, by themselves are not associated with any temporal interpretation which justifies the use of the adjective *later*, nevertheless this is how we make sense of the real-world phenomena we study. We understand them as being embedded in historical time.

[223] Garson 2013.

[224] Lehmann-Waffenschmidt 2010.

[225] Candidate theories for theories of everything (TOEs) or GUTs are among others (supersymmetric) string theory, M-theory, quantum gravity, loop quantum gravity, causal set theory, Wheeler–DeWitt equation or E_8 theory.

at every point in time.[226] The collection of these parallel universes is referred to as the multiverse.

Conclusion

Ever since its creation the St. Petersburg lottery has occupied great minds of mathematics[227], statistics, moral philosophy,[228] and economics,[229] such that a list of originators of tentative solutions to the St. Petersburg paradox reads like a who is who of science.[230] It repeatedly fueled discussions on the logical and ontological foundations of probability theory. Also, the St. Petersburg paradox led to the invention of utility theory in economics and to EUT as a new theory of rational decision making that needs serious reevaluation as we show in the next chapter.

Although the origins of EUT go back as far as to Bernoulli (1738), it was no immediate success. The first awakening of EUT happened when the marginalists W. S. Jevon, L. Walras and C. Menger revolutionised economics from 1870 onwards. The great breakthrough of EUT was due to the seminal publication of von Neumann and Morgenstern (1955, pp. 28, 38) in 1944 and attracted de novo interest in the topic.[231] The game changer in von Neumann and Morgenstern (1955) was its axiomatic foundation of EUT. The axiomatic approach was especially fashionable in economics in the first half of the 20$^{\text{th}}$ century, see Sec. 6.3 for a detailed discussion of this particular aspect. The decisive difference in their approach was that it solely relies on axioms about preference relations. This convinced the scientific community which quarrelled over the almost impossibility to observe individual utility functions and their curvatures. In von Neumann and Morgenstern (1955) the St. Petersburg lottery is used as a motivation for the introduction of utility functions in general and expected utility of certain lotteries in particular. This practice is still common in textbooks on standard microeconomics[232] or on classical and behavioural finance.[233]

226 On the theory of cosmological natural selection see for example Smolin (1992, 2004). Of a wider and by no means less interesting scope are Smolin (1997, 2006, 2013) and Mangabeira Unger and Smolin (2015).

227 N. Bernoulli 1713, D. Bernoulli 1738; Euler 1862; Czuber 1882; Feller 1945.

228 Hacking 1980.

229 Keynes 1921; Menger 1934; von Neumann and Morgenstern 1955; Arrow 1951; Samuelson 1960; Aumann 1977; Shapley 1977b; Gigerenzer and Selten 2002a.

230 The given references can only serve as an incomplete list. For further contributions by Buffon, D'Alembert, Condorcet, J. F. Fries, Poisson, Cournot, de Morgan, Bertrand, von Bortkewitsch, É. Borel, Fréchet, Marschak, F. P. Ramsey, de Finetti, Friedman, Savage and others and survey material on the subject see especially Todhunter (1865), Samuelson (1977), Jorland (1987) and Dutka (1988).

231 Friedman et al. 2014, Ch. 1; Hens and Rieger 2016.

232 Mas-Colell et al. 1995.

233 Hens and Rieger 2016, Ch. 2.

Every era of probability theory came up with time and again new tentative solutions to the paradox (that are partly contradictory to each other). Jorland (1987, pp. 157–158) identified three categories of tentative solutions in order to let eq. (3.6) converge, that are based on either one of the following three strategies:

1. substituting absolute payouts, x, for some other function of the payouts, $f(x)$, to implement slower growth or even an upper bound to possible extreme payout, or
2. substituting absolute probabilities p for another function $g(p)$, to implement requested behaviour especially in the tails for extremely small probabilities and/or even a lower bound to the probability (known as truncation), or
3. by limiting the maximum number of possible coin tosses, N.

All these solutions are united by the belief 'that it is the rules of the game, by allowing for an infinite number of tosses, that make the series diverge'.[234] In what follows, we will show that the treatment of the averages of a random variable and a dynamics, defined by the mode of repetition over time – are key and lead to the broken ergodicity of our observables. Non-ergodic dynamics were not tractable with any mathematics before early 20$^{\text{th}}$ century. Today, we have new tools from ergodic theory at our disposal, which contribute to a resolution of the paradox and more generally enable a new treatment of decision-making under uncertainty in economics based on time averages.

[234] Jorland 1987, p. 157.

4 Solution Strategies to the St. Petersburg Paradox

After the initial exposition of the St. Petersburg paradox by Nikolaus Bernoulli in 1713 and first solution proposals by Cramer in 1728, Daniel Bernoulli in 1738 and others, the paradox spurred many scientists to conciliate the paradoxical results with their disciplines and develop a solution. Hereafter, it became very famous as a part of almost every textbook on probability and decision theory until today and is mentioned in many influential textbooks for mathematics and economics, to name just a few, e.g. Todhunter's *A History of the Mathematical Theory of Probability*[235], Keynes' *A Treatise on Probability*[236], von Neumann and Morgenstern's *Theory of Games and Economic Behavior*[237], Feller's *An Introduction to Probability Theory and Its Applications*[238], and in a myriad of papers in mathematics, statistics, economics, psychology and decision theory, some of which will be discussed or mentioned below if they are adding valuable insights for the advancement of this thesis.

The structure of this chapter follows a classification of the solution strategies based on the mathematical operation they implement. This has the advantage that we do not refer to mere labels but state precisely what has been introduced:[239]

- Sec. 4.1 contains solution strategies which rely on a truncation of small probabilities in the ensemble average,
- Sec. 4.2 contains solution strategies which rely on truncation of the payouts in the ensemble average,
- Sec. 4.3 contains solution strategies which introduce non-linear transformations of wealth or probabilities in the ensemble average,
- Sec. 4.4 introduces a solution strategy which limits the number of coin tosses,

235 Todhunter 1865.
236 Keynes 1921, pp. 62.
237 Von Neumann and Morgenstern 1955, p. 28.
238 Feller 1957, Ch. 10.4.
239 The survey article Jorland (1987, p. 157) is structured in a related manner.

- and finally Sec. 4.5 contains a solution strategy which embeds the randomness of the lottery in time and acknowledges the non-ergodicity of the wealth dynamic.
- concludes the chapter and leads over to the literature review in the next chapter.

Let us state the key result up front to keep it in mind, when going through the chapter. The solution strategies 1 to 4 are all based on the embedding of randomness in a virtual ensemble. None of these strategies yields a convincing evaluation of gambles in general and with respect to the St. Petersburg lottery in particular as it has been originally posed by N. Bernoulli. They all critically rely on additional ad hoc assumptions on the rules of the game or idiosyncratic characteristics of a gambler (his utility function). Thereby the solution strategies 1-4 redefine the problem and then solve this redefined version of the problem. The reader will immediately identify at least the first three solution strategies as ad hoc. The embedding of the randomness within time is made explicit for the first time in Peters (2011c), which constitutes solution strategy 5 and justifies a new framing of former solution strategies provided in this chapter. This solution strategy takes into account the non-ergodicity of wealth which is a consequence of treating the wealth dynamic properly. The introduction of a dynamic is by no means an ad hoc assumptions and the method is capable to deal with arbitrary dynamics. The solution strategy using non-ergodicity is covered in detail in Sec. 4.5 and builds the core of this chapter if not of the whole thesis. The resolution using non-ergodicity stands momentarily as the final step in the history of the St. Petersburg paradox, but if the history of this problem teaches us one thing, it is not to believe in ultimate resolutions. Nevertheless it opened a door to the new paradigm of *Ergodicity Economics*, which seeks growth-optimality. The expositions in this chapter will prepare the reader for the new paradigm of growth-optimal decision theory which we discuss in more depth in chapter 7.

4.1 Solution Strategy 1: Truncation of the Probabilities

The very first solution traces back to N. Bernoulli enclosed in his letter to Montmort. Basically, his solution is based on a truncation of the tails of the probability distribution, but how did he justify the ad hoc truncation as a solution? N. Bernoulli was also the editor of his uncle's, Jakob Bernoulli, posthumously published *Ars Conjectandi*[240]. In this book J. Bernoulli introduces the notion of 'moral certitude' as well as 'moral impossibility'. The meaning of these ad hoc notions is the following. 'Moral impossibility' is closest to absolute or mathematical impossibility, *i.e.* the very low probabilities of a sequence of a very high number of consecutive coin tosses n until 'heads' appears for the first time, of say

[240] Bernoulli (1713), the Latin title translates as 'The Art of Making (Probabilistic) Conjectures'. It contains the first version LLN. See also Bernoulli (1899).

$p(n) < 1/1000$, make it look almost impossible to encounter such a sequence within in a single game or even a lifetime. Though it is not completely impossible, not mathematically impossible, e.g. one could, in principle and physically toss a coin 50, 100 or a 1000 times in a row and always see 'heads'. The probability would be $\left(\frac{1}{2}\right)^{1000} \approx 10^{-300}$, which is so tiny to lead N. BERNOULLI to say that such events basically never happen. But N. BERNOULLI considers such low probability events as sufficiently unlikely such that they can be 'safely' neglect. Following this line of thought, moral impossibility cancels mathematical possibility. It becomes 'morally certain' to omit these terms from the expectation computation. This idea of the truncation of low probabilities reappears in the history of probability e.g. as COURNOT's principle.[241]

Bernoulli and Moral Impossibility

It is in this sense the gamble is considered unrealistic beyond a lower threshold of probabilities that renders it 'morally impossible'. Following in his uncle's spirit, NIKOLAUS BERNOULLI denied the practical relevance of very low probabilities. For example, if N. BERNOULLI considers all probabilities $p(n) < 0.001$ as 'morally impossible', then his version of the St. Petersburg lottery is changed into a truncated version of it given in **Table 4.1**, where only the first $n^{\max} = 9$ terms enter the sum in eq. (3.7). In other words, within the infinite ensemble the gambler is only admitted to be only as lucky as his copy in world_9 in eq. (3.10) and in such cases a finite-ensemble average or more precisely an artificially limited ensemble average is computed. In this version of the game a fair ticket price is the finite-ensemble average (or artificially limited ensemble average)

$$C_{\text{fair}} = \mathrm{E}[x] = \sum_{n=1}^{n^{\max}} \frac{2^{n-1}}{2^n \text{ €}} = \sum_{n=1}^{n^{\max}} \frac{1}{2} = n^{\max} \cdot \frac{1}{2} = \frac{n^{\max}}{2} = \frac{9}{2} = 4.5 \text{ €}\,. \tag{4.1}$$

To illustrate the ad hocness of any truncation of the probabilities let us think of a N. BERNOULLI with a greater amount of imagination, *i.e.* a N. BERNOULLI for whom only events of the probability lower than one in a million, $p(n) < 1/1\,000\,000$, would start to become 'morally impossible'.[242] This would still result in another truncated version of the original game, which is given in **Table 4.2**. In **Table 4.2** every coin toss up to the biggest integer

241 For references to COURNOT's principle see COURNOT (1843), SHAFER and VOVK (2006), SHAFER (2010) and JOHNSON (2016). A brief discussion follows on 107.

242 See also footnote 243 on page 107 on different scales of moral impossibility.

Table 4.1: N. Bernoulli's truncated version of the game. The truncation appears at the first event with a probability of occurrence $p(n) < 1/1000$, here after $n^{\max} = 9$ indicated by the dashed line. Strictly speaking p are no probabilities, because $\sum_n p(n) \neq 1$.

n	$p(n)$	$x(n)$	$\mathrm{E}[x]$
1	1/2	1 €	0.5 €
2	1/4	2 €	0.5 €
3	1/8	4 €	0.5 €
⋮	⋮	⋮	⋮
9	1/512	256 €	0.5 €
- - -	- - -	- - -	- - -
10	0	512 €	0
11	0	1024 €	0
⋮	⋮	⋮	⋮

$n^{\max} \leq k \in \mathbb{N}$ enters the sum, for k satisfying

$$\frac{1}{2^k} < \frac{1}{1\,000\,000} \tag{4.2}$$

$$2^k < 1\,000\,000 \qquad |\log()$$

$$k \ln 2 < \log 1\,000\,000$$

$$k < \frac{\log 1\,000\,000}{\log 2} \approx 19.93 \tag{4.3}$$

$$\hookrightarrow n_{\max} = 19.$$

Thus using eq. (3.9) and eq. (4.1), the fair ticket price for the more imaginative N. Bernoulli amounts to the finite-ensemble average (or artificially limited ensemble average)

$$C_{\text{fair}} = \sum_{n=1}^{n^{\max}=19} \frac{1}{2^n} 2^{n-1} = 19 \sum_{1}^{19} \frac{1}{2} = \frac{19}{2} = 9.5\ . \tag{4.4}$$

From eq. (4.4) it follows that for an exponentially decreasing (im)probability p^{-n} and an exponentially increasing payout x^{n-1} the ticket price still rises in the end, even if only arithmetically for every allowed probability or coin toss, respectively. This is what ultimately causes the sum to diverge. Viewed from the ensemble perspective, Bernoulli limits the size of the ensemble set or how lucky the gambler is allowed to become. In other words, a truncation of the probabilities is equivalent to limiting the allowed maximum length or duration of a

Table 4.2: Imaginative N. Bernoulli's truncated version of the game. The truncation appears at the first event with a probability of occurrence $p(n) < 1/1\,000\,000$, here after $n^{\max} = 19$ indicated by the dashed line. Strictly speaking p are no probabilities, because $\sum_n p(n) \neq 1$.

n	$p(n)$	$x(n)$	$\mathrm{E}[x]$
1	1/2	1 €	0.5 €
2	1/4	2 €	0.5 €
3	1/8	4 €	0.5 €
⋮	⋮	⋮	⋮
18	1/262 144	131 072 €	0.5 €
19	1/524 288	262 144 €	0.5 €
20	0	524 288 €	0
21	0	1 048 576 €	0
⋮	⋮	⋮	⋮

single game.[243] In Sec. 4.5 we introduce the duration of a single round irrespective of its outcome.

Cournot's Principle

The ad hoc and therefore arbitrary assumption of truncating very low probabilities to analyse lotteries and the like is also known in the history of probability as 'Cournot's principle'.[244] In the days of Cournot and especially during the principle's heyday in the late 1940s and early 1950s the rationale for its utilisation was, that it is the only way to connect probability theory to events in the real world.[245] Shafer and Vovk (2006, p. 18) present 'Cournot's principle' as a test of a probabilistic theory, where one picks an event that has a very low probability according to the probabilistic theory. The occurrence of such an unlikely event

243 Buffon and Borel pursued similar solution strategies and have interesting justifications for the truncation thresholds. While the former argues with the negligibility of probabilities smaller than 1/10 000, because a 56 years old man with self-assessed good health would neglect the probability of dying within the next day, albeit the relative frequency for a 56 years old man of Buffon's times is close to 1/10 000. Whereas the latter distinguished between various scales of negligibility at the human, the terrestrial and the cosmic scale. At the human scale the improbable begins at 1/1 000 000, because that is the probability of a lethal traffic accident within a day that every Parisian neglects and pursues his normal daily activities outside the home. (See Dutka 1988, pp. 33–34).

244 Cournot 1843, p. 58; Shafer 2010, p. 2.

245 Shafer 2010.

is interpreted as a reason to reject the probabilistic theory. It is indeterminate how such an interpretation of the tail of a probability distribution extends to continuous probability densities, e.g. how it explains cases when all possible events form a continuum and individually have a very low probability or are even of measure zero. The limit case of infinitely many possible events is described by a continuous PDF. Imagine for example the probability for a dart to land on a specific but arbitrary point on the circular disc. The probability is zero for every point, because the circular disc consists of a continuum of infinitely many such points. An example for the discrete case are federal lotteries, in which every possible (combination of) numbers has the same but very low probability to win. By increasing the length of the draw of numbers one can fall below any (im)probability threshold. If we were forced to speculate, what is really meant by this invocation of the principle of COURNOT could be something closer to convergence in probability, which utilises that all the behaviour is described correctly by the probabilistic theory up to a set of measure zero, which thus can justly be disregarded.[246]

4.2 Solution Strategy 2: Truncation of the Payouts

The second solution of the St. Petersburg lottery appeared in a letter[247] in 1728 by Swiss mathematician GABRIEL CRAMER[248] to N. BERNOULLI, he writes:

> "The mathematical expectation is rendered infinite by the enormous amount which I can win if the cross does not fall upward until rather late, perhaps at the hundredth or thousandth throw. Now, as a matter of fact if I reason as a sensible man, this sum is worth no more to me, causes me no more pleasure and influences me no more to accept the game than does a sum amounting only to ten or twenty million guilders."[249]

He proceeded in very similar vein of N. BERNOULLI who truncated the probabilities at some point, whereas CRAMER's solution truncates possible payouts. In a similar line of reasoning as above the amount of money involved becomes for CRAMER 'morally inconceivable' or of no more additional value or we could even put it in the words of the BERNOULLIS 'morally impossible'. For CRAMER any amount above 2^{24} € $>$ 16 000 000 € is 'morally inconceivable' and therefore it does not matter if a gambler wins 2^{25} € or even 2^{35} € ducats, both amounts give him the same pleasure.[250]

---

[246] See also the presentation SHAFER (2006) on 'Why did Cournot's principle disappear?'
[247] BERNOULLI 1975, pp. 557–567.
[248] CRAMER's rule to solve systems of linear equations is also named after GABRIEL CRAMER.
[249] Translation taken from DEHLING (1997, p. 225).
[250] JORLAND 1987, p. 159; DEHLING 1997, p. 225.

The maximum winning to enter into eq. (3.7) is at most 2^{24}, which results in yet another version of the game given in **Table 4.3**, where the truncation appears in the payouts. A fair price for this version of the lottery is the finite-ensemble average (or artificially limited ensemble average)

$$C_{\text{fair}} = \sum_{n=1}^{N=25} \frac{1}{2^n} 2^{n-1}\,€ + \sum_{n=25}^{\infty} \frac{2^{24}}{2^n\,€} = \frac{25\,€}{2} + 1\,€ = 13.5\,€\ . \tag{4.5}$$

Table 4.3: Cramer's truncated version of the game. All payouts bigger than $x > 2^{24}$ yield the same pleasure as 2^{24}. Here indicated for all payouts below the dashed line.

n	$p(n)$	$x(n)$	$\mathrm{E}[x]$
1	$1/2^1$	$1\,€$	$0.5\,€$
2	$1/2^2$	$2\,€$	$0.5\,€$
3	$1/2^3$	$4\,€$	$0.5\,€$
⋮	⋮	⋮	⋮
24	$1/2^{24}$	$2^{23} = 8\,388\,608\,€$	$0.5\,€$
25	$1/2^{25}$	$2^{24} = 16\,777\,216\,€$	$0.5\,€$
25	$1/2^{25}$	$2^{24} = 16\,777\,216\,€$	$0.25\,€$
26	$1/2^{26}$	$16\,777\,216\,€$	$0.125\,€$
27	$1/2^{27}$	$16\,777\,216\,€$	$0.0625\,€$
⋮	⋮	⋮	⋮

The Ultimate Limit is the Globally Obtainable Number of Coins

Whereas the truncation of the probabilities can be thought of as a limitation attributed to the imagination of the gambler, the truncation of the payouts can be seen as a limitation of the banker's ability to fulfil the gambler's claim. No supplier of the lottery could keep the promise implicitly given by acknowledging the rules (*i.e.* meet the claim of extreme payouts) of the game if they surpass his individual wealth, the liquid capital in the bank's vault or, say, the total amount of coins in the world. It is in this sense that Cramer is considering the gamble unrealistic beyond a certain upper threshold of payouts and therefore renders it 'morally impossible'. In this line of thought a truncation seemingly resolves the paradox in such a way, that everything above that threshold simply doesn't enter the calculation anymore because it is unthinkable. This is clearly ad hoc.

However, there is a similar but less restrictive albeit pseudo-solution to the paradox not by truncation but by introducing upper bounds to the (usefulness) of payouts. CRAMER's reasoning is illustrated nicely together with the longer quote on page 108 and the following statement: 'Mathematicians value money in proportion to its quantity, commonsense men in proportion to its use'.[251] This marks a crucial point in the history of science. Therefore, CRAMER's conception of a truncation of a gambler's absolute payout may be regarded as a transformation which corresponds to a linear utility function $h : \mathbb{R} \to \mathbb{R}$ with an upper bound and in its most general form is given by:

$$h(x) = \begin{cases} x\,(n)\ € & \forall\ n < n^{\max} \\ x\,(n^{\max} - 1)\ € & \forall\ n \geq n^{\max} \end{cases}, \tag{4.6}$$

whereby the payout function given in eq. (3.1) still applies for $x\,(n)$. Thus, the truncated payout function h is a piecewise linear function of the absolute payouts and is depicted in **Fig. 4.1**.

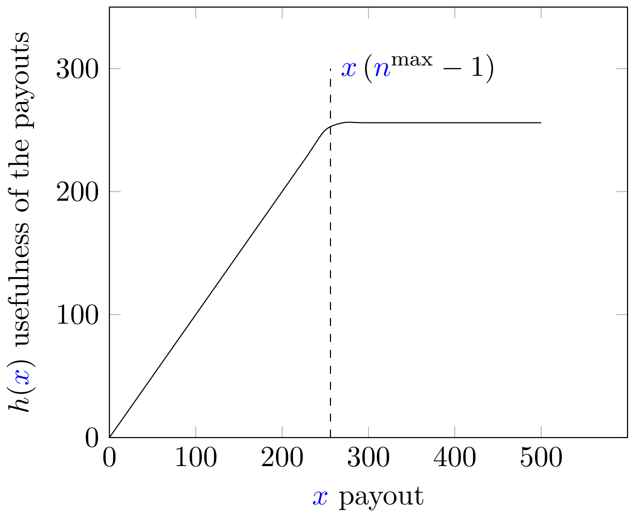


Figure 4.1: Cramer's truncated payout function h. Truncated payouts $h(x)$ given in eq. (4.6) with $n^{\max} = 9$. The usefulness of the payouts stays constant for any values greater than $x\,(n^{\max} - 1) = 256$ or, put differently, there is zero marginal usefulness of payouts above $x\,(n^{\max} - 1)$.

CRAMER's line of reasoning is a tentative solution of the paradox and a precursor of what later became known as utility theory, which will be discussed in more detail in the next Sec. 4.3.

[251] BERNOULLI 1975, pp. 560–561.

Other truncaters

Several authors have justified the truncation of the payouts by all kinds of mundane reasons and thereby reiterated in a sense CRAMER's argument of the genuineness of the banker's promise. E.g. SHAPLEY questioned an (in his opinion) unstated implicit assumption of the gullibility of the gambler: 'One assumes that the [gambler] believes the offer to be genuine, *i.e.*, believes that he will actually be paid, no matter how much he may win';[252] he put forth the house limit of any casino;[253] and the remoteness from everyday experience of the extreme possible payouts in St. Petersburg lottery: 'unreasonably large amounts of money must also be disallowed if our everyday experience is to be brought to bear in deciding whether someone's supposedly rational.'[254] Another ad hoc restriction of the class of random variables. However, SHAPLEY (1977a, p. 449) also gives an amusing real-world analogy of the St. Petersburg lottery in order to determine the value of a 'blank cheque', a signed cheque with the money amount left blank, e.g. cashable at the *First Rational Bank of St. Petersburg*. Because the bank account at the First Rational Bank of St. Petersburg will contain an admittedly large but finite amount, the fee for this blank cheque must be lower or equal than that amount and therefore finite.

All the arguments to limit the divergence of the possible payouts, that are brought forward by SHAPLEY, introduce additional rules or environmental conditions that are absent from the original formulation. The original St. Petersburg lottery was posed by BERNOULLI as a thought experiment. As such it poses on the one hand an interesting problem in the abstract PLATONIC heaven of mathematics, but on the other hand depicts in an abstract sense distinct real-world situations as described in Sec. 3.1. Even if some may find it implausible in its original formulation for all kind of reasons, we prefer to keep striving for a solution of the abstract mathematical problem as it is given to us and look for a solution within the original setting. A premature introduction of arbitrary worldly imperfections of the mundane sphere makes it easy to derive ad hoc pseudo-solutions. Therefore, we strive for a resolution of the problem on a higher level of abstraction than e.g. SHAPLEY does. Because the insights gained from our analysis will help us to guide the direction to the solution of a whole array of problems that are relevant to economics and decision making in everyday situations and related to the embedding of randomness within time, but sometimes not easy to convey.

For COWEN and HIGH (1988) the time to consume an infinite amount of money or goods is ultimately limited by the gambler's finite lifetime. Similar, KIM (1973, p. 149) remarks that the lifetime of the gambler ultimately limits the relevant planning and therefore playing horizon of the lottery, which he addresses with an *augmented-income approach* that is of no

[252] SHAPLEY 1977b, p. 440.
[253] SHAPLEY 1977a, p. 448.
[254] SHAPLEY 1977a, p. 449.

further interest here. With the exception that he relates a possible solution to the paradox to not only fair gambling tickets but also fair insurance fees: 'The basic requirement for any theory of uncertainty is that the model must be capable of accounting for the purchase of both lottery and insurance by the same decision maker' and a dependence of the lottery ticket on the wealth level of the gambler.[255] Especially, the dependence of the gambler's initial wealth is often missing in decision theories, such as cumulative prospect theory (CPT).[256] Kim (1973) draws also the following lesson from the results on the St. Petersburg lottery, namely 'to reject the assumption of linear utility of income or wealth (and, therefore, the mathematical expectations approach)' but unfortunately tries to rescue the programme by the 'search for other types of utility function[s].'[257]

Brito (1975) divides the solutions to the St. Petersburg lottery into three classes. First, those who ascribe the paradox to the conditions of the game itself, e.g. impossibility to disburse an infinite amount to the gambler, the impossibility to consume an infinite quantity of something material, which can only be overcome by the 'only possible candidates for infinite consumption in a finite time period [which] are the more abstract pleasures; knowledge and power. It can be argued, however, that both of these activities are very time intensive';[258] second, truncation of the probabilities (due to moral impossibility), and third, non-linear distortions of the payout function such as the utility concept, which he sees as ad hoc. Similar to the invocation of the abstract pleasures of knowledge and power, Aumann (1977, p. 444) reflects upon which possible experiences could yield infinite rewards:

> "For example, the lottery ticket that Paul is considering might be some kind of open-ended activity–one that could lead to sensations that he has not heretofore experienced. Examples might be religious, aesthetic, or emotional experiences, like entering a monastery, climbing a mountain, or engaging in research with possibly spectacular results."

Aumann concludes from this that sensations which increase (decrease) indefinitely with the stimuli are only possible for spiritual returns (losses) like those considered in Pascal's wager for salvation (or eternal condemnation).[259] Furthermore, Aumann (1977) completely adheres to the capability of bounded utility functions to make the whole paradox and the argument of Brito (1975) and Shapley (1977b) disappear, which was first introduced in Menger (1934).

255 Kim 1973, p. 155.

256 In CPT every decision is made in relation to some reference point, which is most often the entry price. The entry price and therefore also the reference point has no clear relation to the gambler's initial wealth. Both aspects of insurance and the dependence of the lottery ticket of the wealth level of the supplier/bank and the gambler are addressed in Peters and Adamou (2017a, 2018b, ch. 2.9) and Peters (2011c, figure 1), respectively.

257 Kim 1973, p. 148.

258 Brito 1975, p. 125.

259 Lengwiler 2009.

We introduce EUT in detail in the following section and discuss the issue with bounded utility functions as well.

4.3 Solution Strategy 3: Non-Linear Transformation of Payouts

The ordering principle used throughout this chapter is the mathematical operation that is introduced into the evaluation of gambles. In the two preceding sections hard truncations had been applied to the two components of the ensemble average, *i.e.* to the probabilities in Sec. 4.1 or to the payouts in Sec. 4.2. The main purpose behind the truncations is to counteract the divergence of the expectation value. The solution strategy we present in this section counteracts the divergence, too, but involves less radical actions. The idea is to introduce a non-linear transformation of the wealth to counteract the exponential growth. This non-linear transformation is simply a function of wealth, which is called a utility function. Utility is a psychological concept and a continuation of the idea of usefulness, which was already present in CRAMER's work. The favoured rationale in economics for the introduction of utility is the decreasing marginal usefulness of consumption, but utility appeared first in context of wealth. At the same time utility is also injecting a subjective or idiosyncratic element about the nature of the gambler into the objective mathematical expectation, which is why expected utility is also called moral expectation. The necessary specification of the functional form of the utility functions adds an additional degree of freedom to the theory.

The hope was that this psychological element turns out to be universal for all gamblers. Role models for universal psychological relationships have been discovered in the 19th century in a field called psychophysics. E.g. the WEBER-FECHNER law relates diverse stimuli and their sensation in humans.[260] Empirical evidence, however, does not lend support to this hope of a universal utility function.[261]

It was hoped for that this idiosyncratic function may be universal across different gamblers at least up to linear transformations. Then this function would encode a characteristic trait of people in the model. This turned out to be untenable case.[262]

Throughout the section it is important to keep in mind, that the utility concept still operates on the same base as the mathematical expectation, *i.e.* an ensemble average over the possible states, hence an embedding of the randomness within the ensemble. When utility was first formalised, it was intended to amend the concept of 'mathematical expectation' and thus referred to as 'moral expectation'. It was not intended to escape from an embedding of randomness within the ensemble, which had not been comprehended back then.

260 WEBER 1834; FECHNER 1869.
261 FISHBURN 1988; FRIEDMAN et al. 2014; PETERS 2011c; PETERS and ADAMOU 2018b.
262 FRIEDMAN et al. 2014.

Almost every textbook on decision theory and utility theory attributes EUT to the seminal paper by BERNOULLI (1738). However, this is at most partly true as it will turn out. Nevertheless the paper contains at least two major innovations. The first is the switch from an evaluation of the payouts from a gamble to consider changes in the gambler's wealth which we discuss in Subsec. 4.3.1. A survey of the literature on decision theory reveals considerable confusion. This teaches us the lesson that it is essential to start with a clear notation. From this solid basis we continue with the presentation of EUT and derive its decision criterion in Subsec. 4.3.2 and compare it in Subsec. 4.3.3 with a completely different criterion D. BERNOULLI derived. In a series of papers by PETERS[263] presents in detail what BERNOULLI's decision criterion really is. The astonishing result is that BERNOULLI's criterion is inconsistent with the criterion in EUT. Surprisingly, it was LAPLACE who gave EUT its correct form, we discuss this historical incidence in Subsec. 4.3.4. In Subsec. 4.3.5 we discuss proposed functional forms of utility.

4.3.1 The Introduction of Wealth

One major advancement in BERNOULLI (1738)[264] is that he realised the necessity to incorporate the initial wealth levels of gamblers as a reference point in the evaluation of gambles.

[263] See PETERS and GELL-MANN (2016), PETERS (2017), PETERS (2018c) and especially PETERS (2014, 2018a). The reader may find it helpful to consult PETERS (2017, 2018c) in which BERNOULLI's derivation and its flaws are meticulously reproduced using the above mentioned figure. As the author of this thesis, I am in the fortunate state to have discussed this subject several times with OLE PETERS and ALEX ADAMOU over the years. I am even more privileged, to having been granted access to the unpublished manuscripts PETERS (2014, 2018a). Unfortunately, until the present day the public is deprived of these highly relevant and lucid expositions from which we draw an immense amount of understanding of the subtleties of the subject of this thesis and also a huge amount of inspiration. Because they are unpublished, let us comment on them in a bit more detail.

PETERS (2014) is wonderfully written compilation of discussions and email exchanges on the subject with KENNETH ARROW and MURRAY GELL-MANN, which actually spread over many weeks, even years. The paper reproduces a most interesting back and forth of arguments and counter-arguments and is a rich source of how key concepts of economics, information theory and statistical physics are probed to the limit and finally merged into a coherent entirety.

PETERS (2018a) is a comment on the text by BERNOULLI (1738). He locates it in the sphere of other decision theories by HUYGENS and LAPLACE, and restates the arguments in the modern terminology on random variables and stochastic processes, which is largely developed by the same author in the context of *Ergodicity Economics*. Furthermore, the relation between the decision criteria in HUYGENS', LAPLACE's and BERNOULLI's theories is portrayed.

In citing and commenting on these texts here, we hope to exercise a mild force of extortion, which may lead to the publication and dissemination of these papers. They definitely deserve to be read so that more people can benefit from them. [In the meantime, PETERS (2018a) is available online.]

[264] The Latin original (BERNOULLI 1738) was first translated into German and commented by the mathematician ALFRED PRINGSHEIM in BERNOULLI (1896). The original text is set in a historical context by LUDWIG FICK's accompanying introduction. Later a translation into English was provides in BERNOULLI (1954) by LOUISE SOMMER with the help of KARL MENGER. The Latin original did not contain the famous graph of the concave utility function of wealth. It was inserted in the German translation (p. 32) to visualise BERNOULLI's reasoning and kept in the English version (p. 26). Page numbers in the following always refer to the BERNOULLI (1954) is not explicitly stated otherwise.

Bernoulli (1896, p. 26) derived this insight from the famous comparison of the different usefulness of the same absolute payout for a poor man and a rich man and when he later evaluates the specific case of the St. Petersburg lottery:

> "Thus there is no doubt that a gain of one thousand ducats is more significant to a pauper than to a rich man though both gain the same amount.[265]
>
> If Paul owned nothing at all the value of his expectation would be [...] two ducats, precisely. If he owned ten ducats his opportunity would be worth approximately three ducats; it would be worth approximately four if his wealth were one hundred, and six if he possessed one thousand. From this we can easily see what a tremendous fortune a man must own for it to make sense for him to purchase Paul's opportunity for twenty ducats.[266]"

Formally this implies that the study of $\mathrm{E}[x]$ is no longer considered to be sufficient, but the expected change in wealth $\mathrm{E}[\Delta w]$ is the new object of interest. Well this is close but not exactly how Bernoulli thought about the topic, but let us be precise at this point and introduce all the notation before we continue with the second innovation in Bernoulli (1738). Very much as before, let C be the ticket price and x the payout e.g. obtained from a gamble. Then the *final wealth* – after the gambler paid the fee C and received an uncertain payout according to some payout scheme, e.g. $x(n)$ as defined in eq. (3.1) for the St. Petersburg lottery – is given by

$$w_{t+\delta t} = w_t + x_t - C \ , \tag{4.7}$$

whereby δt denotes the time interval it takes to play a single round of a gamble. In this generic setting of a gamble, the initial wealth w_t and the fee C are known for sure. The final wealth $w_{t+\delta t}$ in eq. (4.7) depends also on the payout x_t in the round, which is a realisation of the random variable X_t. Consequently, $w_{t+\delta t}$ is a realisation of the random variable $W_{t+\delta t}$, too. As eq. (4.7) is an additive representation of the change in wealth, we call Δw a wealth *increment*,

$$\Delta w = \underbrace{w_{t+\delta t}}_{\text{final wealth}} - \underbrace{w_t}_{\text{initial wealth}} = x_t - C \ . \tag{4.8}$$

In what follows it will turn out to be crucial to plug in the correct expressions for the initial wealth and the final wealth. Let us continue with Bernoulli's second innovation.

---

265 Bernoulli 1954, p. 26.
266 Bernoulli 1954, p. 32.

4.3.2 Expected Utility Theory – The Introduction of Moral Expectation

The second innovation in Bernoulli (1738) is the introduction of a utility function u, which usually encodes the criterion of marginal decreasing usefulness of money.[267] Changes in utility of wealth are given by

$$\Delta u(w) = \underbrace{u(w_{t+\delta t})}_{\text{utility of final wealth}} - \underbrace{u(w_t)}_{\text{utility of initial wealth}} . \tag{4.9}$$

Thus the effectively interesting quantity for expected utility theory is not the expected change in wealth $\langle \Delta w \rangle$ but the expected change in the utility of wealth, which is interchangeably denoted by $\mathrm{E}[\Delta u(w)]$ or $\langle \Delta u(w) \rangle$ and given by

$$\langle \Delta u(w) \rangle = \langle u(w_{t+\delta t}) - u(w_t) \rangle \tag{4.10}$$

$$= \underbrace{\mathrm{E}\Big[\overbrace{u(w_{t+\delta t})}^{\text{uncertain utility of final wealth}}\Big]}_{\text{Expected utility of final wealth}} - \underbrace{u(w_t)}_{\text{certain utility of initial wealth}} . \tag{4.11}$$

Importantly, the concept of expected utility is still tied to the embedding of randomness within the ensemble as can be seen from this quote:

> "If the utility of each possible profit expectation is multiplied by the number of ways in which it can occur, and we then divide the sum of these products by the total number of possible cases, a mean utility [moral expectation] will be obtained, and the profit which corresponds to this utility will equal the value of the risk in question."[268]

In order to derive the decision criterion in the framework of EUT we have to agree on expressions for the initial wealth w_t and the final wealth $w_{t+\delta t}$ according to a generic gamble and plug them into eq. (4.10).

---

[267] In the general case utility functions have decreasing marginal utility, but sometimes utility functions with increasing marginal utility appear in the literature to model risk seeking behaviour or so called exotic preferences.

[268] Bernoulli 1954, p. 24.

Decision Criterion in Expected Utility Theory

Now if plug eq. (4.7) into eq. (4.10) this yields the expected change in utility

$$\langle \Delta u(w) \rangle = \langle u(w_{t+\delta t}) - u(w_t) \rangle \tag{4.12}$$

$$= \langle u(w_t + x - C) - u(w_t) \rangle \ . \tag{4.13}$$

Let $C_{\max}$ be the value of the fee for which the expected change in utility in eq. (4.13) is zero. For any $C > C_{\max}$ the expected change in utility is negative and EUT recommends not to take the gamble. For any $C < C_{\max}$ the expected change in utility is positive and EUT recommends to take the gamble. The exact value of the maximum fee for a given gambler therefore depends on the specific gamble, *i.e.* payouts & probabilities, and his initial wealth w_t. If a gambler has the choice between several gamble, it is rational to choose the one which maximises $\langle \Delta u(w) \rangle$.

Let us apply the decision criterion according to EUT to the St. Petersburg lottery.

Expected Utility of the St. Petersburg Lottery

Now we plug in the respective payouts from eq. (3.1) and probabilities from eq. (3.2) for the St. Petersburg lottery without any further specification of the functional form of the utility function for the moment, we get the expected change in utility from the St. Petersburg lottery as follows,

$$\langle \Delta u(w) \rangle = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n u\left(w_t + 2^{n-1} - C\right) - u(w_t) \ . \tag{4.14}$$

The convergence of $\langle \Delta u(w) \rangle$ now depends on the functional form of u, which we discuss in a little while in Subsec. 4.3.5.

The replacement of (changes in) absolute wealth by (changes in) non-linear transformations of absolute wealth marks a momentous step in the history of economics. Because in mathematical economics wealth must be treated as a random variable. It is worth repeating, that every function of a random variable is itself a random variable and every function of a stochastic process becomes itself a stochastic process. The additional treatment of a function of our random variable using an expectation operator $\mathrm{E}[\cdot]$ is mathematically just the application of another functional. Thus we arrive at the situation where economics studies functionals of functions of random variables. To illustrate this fact we annotate eq. (4.15) in a similar way as we annotated eq. (4.11) already,

$$\mathrm{E}\left[\Delta u\left(w\right)\right] := \overbrace{\underbrace{\mathrm{E}}_{\text{a function of}}}^{\text{a functional of}} \underbrace{\left[\Delta \overbrace{u}^{\text{a function of}} \left(\underbrace{w}_{\text{a random variable}} \right)\right]}_{\text{a random variable}} . \tag{4.15}$$

Associated with this doubly transformation of the random variable is the hope the functional takes the PDF as an argument and ultimately yields a degenerate random variable or simply a plain number, which applies universally as a decision criterion. As it turns out this hope is untenable. We have to remember, that the proper mathematics to handle random variables did to a large extent not exist at the time of DANIEL BERNOULLI. Probability theory was in its infancy (and the BERNOULLI family would make important contributions to it) not to speak of stochastic processes, which were only fully understood in the 20^{th} century. One important fact is that the two functions are not commutative, *i.e.* the expected change in utility of wealth $\mathrm{E}\left[\Delta u\left(w\right)\right]$ can not be derived in a trivial way from the expected wealth $\mathrm{E}\left[\Delta w\right]$. To show the non-commutativity we exploit the special nature of utility functions. They have been introduced to encode non-linear marginal usefulness of money, thus they are explicitly introduced not to be mere linear transformations of wealth, which would formally read as $u\left(w\right) = aw + b$ for arbitrary $a, b \in \mathbb{R}$. Here helps JENSEN's inequality, a well-known result in mathematics.

Jensen's Inequality, Expectation, and Changes in Utility and Wealth

Given w is a realisation of an integrable random variable W and a concave function u, we can derive the following inequality due to JENSEN (1906), thereby from the second line on we take advantage of the fact that the initial wealth w_t is not a realisation of a random variable in our context but known to us:

$$\left\langle \Delta u\left(w\right)\right\rangle < \Delta u\left(\left\langle w\right\rangle\right) \tag{4.16}$$

$$\left\langle u\left(w_{t+\delta t}\right) - u\left(w_t\right)\right\rangle < u\left(\left\langle w_{t+\delta t}\right\rangle\right) - u\left(\left\langle w_t\right\rangle\right) \tag{4.17}$$

$$\left\langle u\left(w_{t+\delta t}\right)\right\rangle - u\left(w_t\right) < u\left(\left\langle w_{t+\delta t}\right\rangle\right) - u\left(w_t\right) \tag{4.18}$$

$$\left\langle u\left(w_{t+\delta t}\right)\right\rangle < u\left(\left\langle w_{t+\delta t}\right\rangle\right) \;. \tag{4.19}$$

Put differently, from JENSEN's inequality follows in general the non-commutativity of the expectation operator and the utility mapping. In order to achieve commutativity, we can not change the expectation operator which is set in stone. Thus only the utility mapping is open to manipulation. However, only for the excluded case of linear utility functions the two operations would become commutative.

In economics concave utility functions are associated with risk averse preferences for which eq. (4.19) applies, convex utility functions are associated with risk seeking preferences. Both cases are discussed in undergraduate (micro)economics textbooks, however, too few textbooks contain a reference to the simple underlying result by JENSEN's inequality.[269]

Remark on the Correct Observable

One more remark on the quantity $u(\Delta w)$ one finds repeatedly in the literature, e.g. in publications with more than 50 000 citations, which is an additional source of confusion on the already intricate evaluation of gambles.[270] The only sensible quantity is a change in utility denoted by Δu, note also the position of the Δ in eq. (4.19).[271] The inconsistent notation of $u(\Delta w) = u(x)$ runs into a problem. Implicitly such a notation encodes a situation in which the payout of a gamble generates the same usefulness to every player irrespective of his wealth. *I.e.* as if 500 € create the same usefulness to a billionaire as they would for a poor person. This is exactly the state of the theory BERNOULLI tried to overcome with the introduction of a dependence on the initial wealth levels in the evaluation of gambles. The principal idea is to ask: What is the effect of an additional amount of money given an initial amount of money? *Ergodicity Economics* takes the question as it stands. BERNOULLI decided to rephrase the question and replaced effect with utility. Even in the world of utility theory the utility of an unescorted payout $u(x)$ is clearly nonsense.

Anticipating the content on functional forms at this point a bit, we can ask the question how to evaluate negative payouts? E.g. $u(x = -8)$ can be handled with neither logarithmic nor square root utility functions, who are both only defined on the positive real numbers. In order to evaluate $u(-8) = \log(-8)$ or $u(-8) = \sqrt{-8}$, we would need imaginary numbers to be able to solve these equations. It is therefore imperative to always analyse changes of wealth Δw (eq. 4.8) or changes in utility of wealth Δu e.g. induced by the payout of a gamble.[272] Even more so it teaches us that we have to interpret a quantity in its domain of application and explicitly ask for the physical meaning of any (mathematical) expression – irrespective whether it comes from a well-know mathematical theorem or not.

By now we have compiled the relevant notation, derived the decision criterion according to EUT and continue with an analysis of the true decision criterion, which DANIEL BERNOULLI

[269] For instance ch. 12 on uncertainty VARIAN (2010) contains no reference to JENSEN's inequality.
[270] KAHNEMAN and TVERSKY 1979, 263, tenet (i).
[271] PETERS and ADAMOU 2018b, pp. 18–19.
[272] This topic is important in dimensional analysis (BARENBLATT 2003) and picked up again in the context of the ergodicity transformation in Subsec. 4.5.4.

derived.

4.3.3 Bernoulli's Decision Criterion

After D. Bernoulli elucidated his first innovation of the introduction of initial wealth levels in the evaluation of gambles, he made the following reasoning to derive his decision criterion: '[I]n a fair game the disutility to be suffered by losing must be equal to the utility to be derived by winning.'[273] The elaborations surrounding this quote are what is supported by the famous figure in later translations. The basic lapse Bernoulli made is that he did not plug in the correct expressions for the initial wealth and final wealth of a gamble, which we derived in eq. (4.7). Instead we can think of Bernoulli's procedure in the following way. He starts with dividing the original gamble (defined in eq. (4.7) and eq. (4.9)) into two independent subgambles. The first subgamble is the 'disutility suffered by losing'. This subgamble is strange, because it actually describes a gamble for which we buy a ticket to partake, but never receive a payout, hence a sure loss. The second subgamble is the 'utility of winning', which is a gamble for which we never bought a ticket to partake but out of nowhere receive an uncertain payout. If presented this way, it is of course not surprising that we do not arrive at the original gamble if we merge the two subgambles. Let us go through it in detail.

First, we denote what Bernoulli called the 'disutility suffered by losing' by Δu^-, because we have no a priori information on it we have to treat it as uncertain, and operate with $\left\langle \Delta u^- \right\rangle$. In words, this quantity is the (expected) change in utility by the sole purchase of the lottery ticket, this is given formally by

$$\left\langle \Delta u^- \right\rangle := \left\langle \underbrace{u\left(w_{t+\delta t}\right)}_{\substack{\text{utility of}\\ \text{final wealth}}} - \underbrace{u\left(w_t\right)}_{\substack{\text{utility of}\\ \text{initial wealth}}} \right\rangle \tag{4.20}$$

$$= \left\langle \underbrace{\overbrace{u\left(w_t - C\right)} - \overbrace{u\left(w_t\right)}}_{\text{all quantities certain}} \right\rangle \tag{4.21}$$

$$= u\left(w_t - C\right) - u\left(w_t\right) < 0 \qquad \forall\, C > 0\,. \tag{4.22}$$

The last step is justified because u is a monotonic increasing function. At first we notice that Bernoulli used strange initial and final wealths for this quantity, which will cause problems shortly. The quantity Δu^- is not a random variable, because all inputs are known. Put differently, eq. (4.22) encodes the disutility from solely buying the ticket, irrespective of

[273] Bernoulli 1954, p. 27.

any realisation of a possible payout, hence Δu^{-} must be clearly negative for positive ticket prices.[274]

Second, let what BERNOULLI called the 'utility of winning' be denoted by Δu^{+} or rather treat it as a random variable $\left\langle \Delta u^{+} \right\rangle$ again, for the same reasons as above. In words, this quantity is the expected change in utility induced by the realisation of a payout. Formally this is given by

$$\left\langle \Delta u^{+} \right\rangle := \left\langle \underbrace{u\left(w_{t+\delta t}\right)}_{\substack{\text{utility of} \\ \text{final wealth}}} - \underbrace{u\left(w_{t}\right)}_{\substack{\text{utility of} \\ \text{initial wealth}}} \right\rangle \tag{4.23}$$

$$= \left\langle \overbrace{\underbrace{u\left(w_{t}+x\right)}_{\text{random variable}}} - \overbrace{u\left(w_{t}\right)} \right\rangle , \tag{4.24}$$

which is the expected change in utility with no fee. Again we have to notice the strange initial and final wealths that enter the quantity $\left\langle \Delta u^{+} \right\rangle$. The quantity Δu^{+} really is a random variable. By now we have simplified the components involved enough to state BERNOULLI's decision criterion.

Bernoulli's Criterion

According to BERNOULLI a gambler should buy a ticket if the absolute value of his 'disutility suffered by losing' Δu^{-} is less than his expected 'utility to be derived by winning' Δu^{+} or formally

$$\left\langle \Delta u^{+} + \Delta u^{-} \right\rangle = \left\langle \Delta u^{+} \right\rangle + \Delta u^{-} \tag{4.25}$$

$$= \left\langle u\left(w_{t}+x\right) - u\left(w_{t}\right) \right\rangle + u\left(w_{t}-C\right) - u\left(w_{t}\right) \tag{4.26}$$

$$= \left\langle u\left(w_{t}+x\right) \right\rangle - u\left(w_{t}\right) + u\left(w_{t}-C\right) - u\left(w_{t}\right) > 0 \ . \tag{4.27}$$

Interestingly, it follows that BERNOULLI's criterion in eq. (4.27) is in general not equivalent to the criterion which we strictly derived from EUT in eq. (4.13) (only valid for linear utility), thus

$$\left\langle \Delta u \right\rangle \neq \left\langle \Delta u^{+} + \Delta u^{-} \right\rangle \ . \tag{4.28}$$

As a matter of fact this should not be too surprising, because the quantities Δu^{-} and $\left\langle \Delta u^{+} \right\rangle$ are just absurd. The quantity $\left\langle \Delta u^{+} \right\rangle$ corresponds to the expected change in utility a gambler

274 Note that in PETERS (2018a) the quantity $\delta u^{-} = -\Delta u^{-}$ in our notation. Because $\Delta u^{-} < 0$ we could make the following substitution: $\left|\Delta u^{-}\right| = u\left(w_{t}\right) - u\left(w_{t}-C\right)$. This substitution would need to be carried through our computations and change the first '+' into a '-' in eq. (4.25).

experiences for which he never purchased a ticket and the quantity Δu^- corresponds to a gamble where the gambler has purchased a ticket but never realises a payout.[275] Once again let us remark, that the quantities Δu^+ and Δu^- are strictly derived from the original BERNOULLI (1738), which later got transformed coherently in the aforementioned figure in the German and English translations of his original publication.

Nevertheless we can analyse if there exist any cases at all for which eq. (4.28) turns into an equality. Therefore we simplify the involved terms

$$\langle \Delta u \rangle \stackrel{?}{=} \left\langle \Delta u^+ + \Delta u^- \right\rangle \tag{4.29}$$

$$\langle u(w_t + x - C) - u(w_t) \rangle \stackrel{?}{=} \langle u(w_t + x) \rangle - u(w_t) + u(w_t - C) - u(w_t) \tag{4.30}$$

$$\langle u(w_t + x - C) \rangle \stackrel{?}{=} \langle u(w_t + x) \rangle - u(w_t) + u(w_t - C) \tag{4.31}$$

The first equality of the criteria happens for the trivial case of linear utility $u(\cdot) = \cdot$. Linear utility corresponds to the absurd case where a utility is introduced just to not use it and has actually been excluded implicitly from the outset,

$$\langle \Delta u \rangle = \left\langle \Delta u^+ + \Delta u^- \right\rangle \tag{4.32}$$

$$\langle u(w_t + x - C) \rangle = \langle u(w_t + x) \rangle - u(w_t) + u(w_t - C) \tag{4.33}$$

$$\langle w_t + x - C \rangle = \langle w_t + x \rangle - w_t + w_t - C \tag{4.34}$$

$$w_t + \langle x \rangle - C = w_t + \langle x \rangle - C \; . \tag{4.35}$$

The second possible way an equality can be achieved is the trivial case when the fee is zero $C = 0$ which yields

$$\langle \Delta u \rangle = \left\langle \Delta u^+ + \Delta u^- \right\rangle \tag{4.36}$$

$$\langle u(w_t + x - C) \rangle = \langle u(w_t + x) \rangle - u(w_t) + u(w_t - C) \tag{4.37}$$

$$\langle u(w_t + x) \rangle = \langle u(w_t + x) \rangle - u(w_t) + u(w_t) \tag{4.38}$$

$$\langle u(w_t + x) \rangle = \langle u(w_t + x) \rangle \; . \tag{4.39}$$

When D. BERNOULLI finally tries to apply his criterion to the St. Petersburg lottery in §19., he must have realised the problems in eq. (4.27) in some way. Otherwise it is hard to explain why he reintroduced linear utility at this point – which is actually a de-introduction of utility – after he argued for non-linear utilities in the remainder to encode decreasing

[275] Two examples which serve as further counter-examples against the use of BERNOULLI's decision criterion are given in PETERS (2018c). First, it is possible that this criterion does not recommend to enter a gamble for which the $C_{\max}$ is less than the lowest possible payout. Thus a guaranteed net profit of an uncertain extent without a downside is still rejected. Second, a sure payout of an arbitrary non-zero amount is worth less than that amount.

marginal utility. Furthermore, on p. 32 in BERNOULLI (1954) he set both initial wealth and the ticket fee to zero and basically stripped away everything from his criterion to just arrive at the second case for which the two criteria just happen to coincide, see eq. (4.36). As we have seen in eq. (4.34), for linear utility BERNOULLI's strange criterion just happens to coincide with the EUT criterion. In this case the formula in BERNOULLI (1954, p. 33) includes a positive ticket fee and presupposes large wealths and linear utility, the respective lines read as follows: 'The amount which the buyer ought to pay for this proposition differs somewhat from the amount it would be worth to him were it already in his possession. Since, however, this difference is exceedingly small if a (Paul's fortune) is great we can take them to be equal.'[276]

BERNOULLI's formulas correspond to our eq. (3.7) or eq. (3.8), though he used a different notation. Let us transform them into our notation to show that BERNOULLI used the EUT criterion in eq. (4.13) and not his own. BERNOULLI used a notation that denotes the probabilities not as ratios as we do, e.g. in eq. (3.2), but as higher-order roots. The reason behind that we will become explicit when we introduce it in Sec. 4.5, too. 'Bernoulli1954'33 wrote

$$\sqrt[2]{w_t + 1 - C} \cdot \sqrt[4]{w_t + 2 - C} \cdot \sqrt[8]{w_t + 4 - C} \cdot \ \cdots \ = w_t \ . \tag{4.40}$$

This expression in eq. (4.40) can be simplified if we take the logarithm

$$\ln w_t = \frac{1}{2}\ln\left(w_t + 1 - C\right) + \frac{1}{4}\ln\left(w_t + 2 - C\right) + \frac{1}{8}\ln\left(w_t + 4 - C\right) + \cdots \tag{4.41}$$

$$= \sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n \ln\left(w_t + 2^{n-1} - C\right) = \left\langle \ln\left(w_t + 2^{n-1} - C\right)\right\rangle \tag{4.42}$$

$$0 = \left\langle \ln\left(w_t + 2^{n-1} - C\right)\right\rangle - \ln w_t \ . \tag{4.43}$$

The last step in eq. (4.43) sets the expected change in utility to zero, which finally exactly determines the maximum fee a gambler should pay for the ticket according to the criterion of EUT in eq. (4.13) for logarithmic utility $u(w) = \ln w$.

Conclusion

The fact that BERNOULLI did not use his own criterion is concealed and hard to spot, most likely he himself did not realise it, but fortunately we can make full use of PETERS (2014, 2018a,c). BERNOULLI (1954, p. 32) concluded that:

---

[276] BERNOULLI 1954, pp. 32-33.

> "From this formula [on pp. 32-33] which evaluates Paul's prospective gain it follows that this value will increase with the size of Paul's fortune and will never attain an infinite value unless Paul's wealth simultaneously becomes infinite."

We would have liked if the conclusion had fully acknowledged the nature of the random variable involved and would read instead similar to:

> "From this formula which evaluates Paul's prospective gain it follows that this value will increase with the size of Paul's fortune and *can only be made finite in the framework of embedding the randomness within the ensemble if we introduce ad hoc assumptions, e.g. about the functional form of a utility function, particular values of the fee and/or initial wealth or truncations of either the probability distribution which governs the payout scheme or the permitted amount of coins in the game/world.*"

The discussion of BERNOULLI's criterion points towards an incoherence in BERNOULLI (1738). The tension arises between the use of what in fact is the decision criterion according to EUT and BERNOULLI's own different criterion. Both can not be transformed into each other without great loss of generality. Nevertheless, most sources refer to BERNOULLI (1738) as the origin of EUT. The next section clarifies how this came about

4.3.4 Laplace's Noblesse

As mentioned in the motivating paragraph at the beginning of this section, almost all textbooks attribute the EUT decision criterion DANIEL BERNOULLI, see for instance the appraisal in TODHUNTER (1865):

> "390. We will now give Daniel Bernoulli's application of his theory of Moral expectation to the Petersburg Problem. [. . .] Laplace reproduces this part of Daniel Bernoulli's memoir [. . .]
> 393. [. . .] We may remark that Laplace adopts Daniel Bernoulli's view.
> 490. Laplace speaks very highly of Daniel Bernoulli [in the context of another problem].
> 1042. [Chapter X entitled De l'esperance morale] may be described as mainly a reproduction of the memoir by Daniel Bernoulli."[277]

How it happened is a fascinating piece in the history of science. When the theory of probability proceeded the concept of moral expectation developed into an important topic which was taken up by the textbooks of the time. A widely read classic textbook of the time was

[277] TODHUNTER 1865, pp. 220, 222, 228, 609.

LAPLACE (1812), to whom we owe e.g. the definition of probability as the ratio of favourable cases to all possible cases (of equal probability). Therein Chapter X is titled *De l'esperance morale* or *On moral expectation*. Here LAPLACE introduced the term *l'avantage moral* (p. 439) similar to *moral expectation* and refers on pp. 439 explicitly to 'le principe de Daniel Bernoulli' and his logarithmic utility function.

Unfortunately, for the (economic) science community, LAPLACE was too much of a noble character to stress all too clearly that he in fact corrected BERNOULLI and gave the EUT decision criterion in its correct form on p. 440 which corresponds to eq. (4.13). LAPLACE (1812) gave the reader no reason to not think of D. BERNOULLI as the originator of this theory. Maybe it seemed obvious to LAPLACE what BERNOULLI intended to do, but simply got astray along the path. Maybe LAPLACE focused only on certain parts of it, we do not know, but since then this story stuck.

4.3.5 Functional Forms of Utility

4.3.5.1 Bernoulli's Logarithmic Utility

The first innovation in BERNOULLI (1738) was the introduction of wealth and led decision theory to the study expected changes in wealth $\langle \Delta w \rangle = \langle w_{t+\delta t} - w_t \rangle$. The second innovation led to the analysis of expected changes in utility of wealth $\langle \Delta u(w) \rangle = \langle u(w_{t+\delta t}) - u(w_t) \rangle$, which required the specification of the functional form of the utility function u. BERNOULLI reasoned about this in the following way: 'any increase in wealth, no matter how insignificant, will always result in an increase in utility which is inversely proportionate to the quantity of goods already possessed.'[278] Let the 'quantity of goods already possessed' be denoted by our symbol for initial wealth w_t, then formally this would read as

$$\Delta u \propto \frac{1}{w_t} \Delta w \ , \tag{4.44}$$

note that it is a statement about the difference in utility, not the utility of a difference. BERNOULLI (1954, pp. 27–28) wrote the assumption not as difference equation but as a differential equation with a proportionality factor $a \in \mathbb{R}$ as

$$\mathrm{d}u = a \, \frac{\mathrm{d}w}{w_t} \ , \tag{4.45}$$

[278] BERNOULLI 1954, p. 25.

whose solution is a logarithmic function[279] with w as its argument and w_t enters as a constant, which is subsumed in the constant b in the second last step

$$u(w) = a \int \frac{1}{w_t}\,\mathrm{d}w = a \log \frac{w}{w_t} = a \log w - \underbrace{a \log w_t}_{\text{const.}=b} = a \log w - b \qquad \forall\ a, b \in \mathbb{R}\ . \tag{4.46}$$

As EUT considers changes in utility the b cancels out, $u(w_{t+\delta t}) - u(w_t) = a \log w_{t+\delta t} - \not{b} - \left[a \log w_t - \not{b}\right]$. Often for convenience $a = 1$ and $b = 0$, for which the dependence on the initial wealth w_t disappears from the logarithmic utility equation and it may get forgotten quickly that there is a reference point dependence on the initial wealth at all. *I.e.* for $b = 0$ follows $0 = \log w_t \rightarrow w_t = 1$ and for arbitrary initial wealth $u(w = w_t) = \log w_t - \log w_t = 0$.[280] Therefore the expression $u(w)$ is actually already encoding a difference or a change in utility, namely a change from some initial wealth. As we see, EUT does contain a reference level dependence. Thus EUT studies the expected changes in the logarithm of wealth

$$\langle \Delta u(w) \rangle = \langle u(w_{t+\delta t}) - u(w_t) \rangle = \langle \log w_{t+\delta t} - \log w_t \rangle\ , \tag{4.47}$$

and the decision criterion involves the maximisation of $\langle \Delta u(w) \rangle$. Consequently, all translations of Bernoulli (1738) got augmented by the famous figure of the logarithmic utility curve. **Table 4.4** lists the utility of the payouts in the St. Petersburg lottery for a utility function using the natural logarithm and an initial wealth of $w_t = 1$. In what follows we denote logarithmic utility functions by u_B.

279 Here and throughout, $\log_b$ is used for the logarithm to the base b, we use the common notation for natural logarithm (*logarithmus naturalis*) $\ln = \log_e$ and the common or decadic or decimal logarithm $\lg = \log_{10}$. Often especially in information theory the binary logarithm (*logarithmus dualis*) is used due to the binary character of bits. Unfortunately for newcomers, this is sometimes not made explicit. For convenience, we use base 2 as default, $\log = \log_2$. For logarithmic utility the base of the logarithm is usually chosen in a way to ensure convenience, which thus differs from purpose to purpose and from discipline to discipline.

280 And not the wrong statement that the utility of some states or even a negative payout can lead to negative utility, $u(w = w_t) = \log w_t < 0$. One can suffer an decrease in utility but a negative utility makes no sense given the state of the theory. See also the **Remark on the Correct Observable** on p. 119. An often claimed advancement by prospect theory or other theories of behavioural economics is the introduction of a *new* reference point (Kahneman and Tversky 1979; Tversky and Kahneman 1992), which as we see is already in EUT.

Table 4.4: Logarithmic utility from expected payout in the St. Petersburg game. Adaption of **Table 3.1** with logarithmic utility, $u_{\mathrm{B}}(w) = \ln(w)$, of the payouts and expected logarithmic utility of the wealth for $w_t = 1$.

n	$p(n)$	$x(n)$	$\Delta u_{\mathrm{B}}(w) = \ln(w_{t+\delta t})$	$\mathrm{E}[\Delta u_{\mathrm{B}}(w)]$
1	1/2	1 €	0	0
2	1/4	2 €	0.693	0.173
3	1/8	4 €	1.386	0.173
4	1/16	8 €	2.079	0.130
5	1/32	16 €	2.773	0.087
⋮	⋮	⋮	⋮	⋮
10	1/1024	512 €	6.238	0.006
⋮	⋮	⋮	⋮	⋮

Expected Utility of the St. Petersburg Lottery with Logarithmic Utility

With the exposition of BERNOULLI's ideas that led to the logarithmic utility function, we can evaluate the St. Petersburg lottery and therefore specify what we presented in the box surrounding eq. (4.14).

$$\langle \Delta u_{\mathrm{B}}(w) \rangle = \langle u_{\mathrm{B}}(w_{t+\delta t}) - u_{\mathrm{B}}(w_t) \rangle \tag{4.48}$$

$$= \langle \log w_{t+\delta t} - \log w_t \rangle = \left\langle \log \left(\frac{w_{t+\delta t}}{w_t} \right) \right\rangle \tag{4.49}$$

$$= \sum_{n=1}^{\infty} \left(\frac{1}{2} \right)^n \log \left(\frac{w_t + 2^{n-1} - C}{w_t} \right) . \tag{4.50}$$

The step from eq. (4.48) to eq. (4.49) reveals that only the *logarithmic* utility function allows a representation of the difference in utility as a ratio of final and initial wealth – *square root* utility would not do the trick. To interpret the logarithm of this ratio in terms of utility is actually an abuse of concepts that becomes apparent once the language of *Ergodicity Economics* is used. We do not want to anticipate too much of what is derived in great detail in Sec. 4.5, but discipline ourselves in referring to Subsec. 4.5.4 where we pick up this thread. As a next step the convergence of the *moral expectation* in eq. (4.50) has to be established in order to see whether the trick with utility works at all in producing a converging expected utility.

Convergence Behaviour for Logarithmic Utility

To remind us about the original intention behind the introduction of utility, Bernoulli wanted to derive a decision criterion which does not diverge as the expected payout $\langle \Delta x \rangle$ or the expected change in wealth $\langle \Delta w \rangle$ does. Let us therefore check whether the expected change in utility converges. The convergence of eq. (4.48) is determined by the convergence of its first term, let us expand the term

$$\langle \log w_{t+\delta t} \rangle = \Big\langle \log \left(w_t + x\left(n\right) - C \right) \Big\rangle = \sum_{n=1}^{\infty} \log \left(w_t + x\left(n\right) - C \right) p\left(n\right) \tag{4.51}$$

$$= \sum_{n=1}^{\infty} \log \Big(w_t + 2^{n-1} - C \Big) \, 2^{-n} \; . \tag{4.52}$$

To determine the convergence of the series in eq. (4.52) let the series be reduced to

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \log \Big(w_t + 2^{n-1} - C \Big) \, 2^{-n} \; , \tag{4.53}$$

with two subsequent terms of the series given by

$$a_n = \log \Big(w_t + 2^{n-1} - C \Big) \, 2^{-n} \tag{4.54}$$

$$a_{n+1} = \log \left(w_t + 2^{n} - C \right) 2^{-(n+1)} \; . \tag{4.55}$$

Then we can apply the ratio test and see that the series does indeed converge for any finite initial wealth and finite fee

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\log \left(w_t + 2^{n} - C \right) 2^{-(n+1)}}{\log \left(w_t + 2^{n-1} - C \right) 2^{-n}} \right| \tag{4.56}$$

$$= \left| \frac{\log \left(w_t + 2^{n} - C \right) 2^{n}}{\log \left(w_t + 2^{n-1} - C \right) 2^{n+1}} \right| \tag{4.57}$$

$$= \left| \frac{\log \left(w_t + 2^{n} - C \right)}{2 \log \left(w_t + 2^{n-1} - C \right)} \right| \tag{4.58}$$

avoid negative convergence in the logarithm $\forall \; w_t + 1 \text{€} < C < \infty n \to \infty$

$$= \left| \frac{\log 2^{n}}{2 \log 2^{n-1}} \right| \tag{4.59}$$

$$\lim_{n \to \infty} \left| \frac{\log 2^{n}}{2 \log 2^{n-1}} \right| = \frac{1}{2} \; . \tag{4.60}$$

We have thus established the absolute convergence of the series,

$$\lim_{n \to \infty} \sup \left| \frac{\log 2^{n}}{2 \log 2^{n-1}} \right| < 1 \; . \tag{4.61}$$

The exact value of $\langle \Delta u_{\mathrm{B}} (w) \rangle$ depends on the initial wealth and the ticket price and can be positive or negative, for gains or losses of utility with respect to the utility of the initial wealth. Finally, the solution strategy using logarithmic utility seems to has reached the ultimate goal, namely BERNOULLI could produce a finite moral expectation value of the St. Petersburg lottery, whereby $C_{\max}$ in the expression

$$\left\langle \log \left(w_t + x(n) - C_{\max} \right) \right\rangle = \log w_t \ , \tag{4.62}$$

denotes the maximum fee a gambler with initial wealth w_t should be willing to pay for a lottery with the payout scheme $x(n)$. Thereby the logarithm transforms the exponentially increasing wealth into a linearly increasing utility of wealth. But let us reflect on how BERNOULLI arrived at this point. What seems to be a *solution* comes at the price of the introduction of a postulated function which is unobservable but at the same time is meant to encode a universal psychological trait of humans. This is an unsatisfactory state and can hardly be considered a theoretical advancement. For BERNOULLI the general idea to introduce arbitrary concave functions of wealth as utility seemed appropriate and sufficient, e.g. he did not held his logarithmic functional form in any higher esteem than other concave functions such as CRAMER's square root utility function,[281]section. In doing so, therein lies the strong hint if not proof that BERNOULLI did not recognise the special physical meaning underlying the logarithm.

Logarithmic Utility is Not a Hard-Wired Human Trait

What contributed to the confusion regarding the real role of the logarithm was a self-enforcing feedback loop between BERNOULLI's findings and its repeated appearance in the context of psychophysical experiments. This gave the impression as if the logarithm would constitute a natural choice among all the possible subjective reweightings of absolute wealth since WEBER deduced from experiments a constant logarithmic scaling relationship between objective intensity of a physical stimulus and its sensory perception of weight. Later, FECHNER performed additional experiments for e.g. temperature, brightness and sound intensity.[282] These findings from the natural sciences are often used to put additional weight on the credibility of the general utility concept in economics[283] and are often used to corroborate logarithmic utility in particular.[284] For instance, STIGLER (1950, p. 373) wrote about how the 'Bernoulli hypothesis [...] merged with the Weber-Fechner law'.

281 BERNOULLI 1954, p. 31.
282 WEBER 1834, 1851; FECHNER 1869.
283 SINN 1989, ch. 3 A.1 & A.2, ch. 4 A.3.
284 SINN and WEICHENRIEDER 1993; SINN 2003.

This exchange of ideas really was a reflexive loop and has not been a one way street[285], there is evidence for the mutual stimulation. E.g. FECHNER referred to BERNOULLI's findings already in his preface:

> "Das erfahrungsmässige Gesetz, welches die Hauptunterlage der psychischen Masslehre bildet, ist schon vorlängst von verschiedenen Forschern in verschiedenen Gebieten aufgestellt und in verhältnissmässiger Allgemeinheit namentlich von E. H. Weber, den ich überhaupt den Vater der Psychophysik nennen möchte, ausgesprochen und experimental bewährt worden. Die mathematische Function anderseits, die den allgemeinsten und wichtigsten Fall der Anwendung unseres Massprincips bildet, ist ebenfalls schon vorlängst von verschiedenen Mathematikern, Physikern und Philosophen, wie Bernoulli (Laplace, Poisson), Euler (Herbart, Drobisch), Steinheil (Pogson) für besondere, der Psychophysik zuzueignende, Fälle auf dieses Gesetz gegründet und von anderen Forschern reproducirt oder acceptirt worden."[286]

And even more explicitly in chapter VIII, where FECHNER used the LAPLACIAN expressions *fortune morale* for moral expectation and *fortune physique* for mathematical expectation:

> "Die mathematische Function anderseits, welche die Grösse des Reizes mit der Grösse der Empfindung verknüpft, ist nach particulären Gesichtspuncten schon vor mehr als hundert Jahren von Euler, später wiederholt von Herbart und Drobisch, für die Abhängigkeit der Empfindung der Tonintervalle von den Verhältnissen der Schwingungszahlen; noch etwas vor Euler von Daniel Bernoulli, später von Laplace und Poisson, für die Abhängigkeit der *fortune morale* von der *fortune physique*, endlich von Steinheil und von Pogson für die Abhängigkeit der Sterngrössendifferenzen, die nichts Anderes als Differenzen von Empfindungsgrössen sind, von der photometrischen Intensität der Sterne aufgestellt worden."[287]

From the derivation in this section with its extensive annotations and comments together with the results in Sec. 4.5 we draw the conclusion that the concept of (logarithmic) utility can not be justified on biological grounds as a somehow hard-wired human trait of subjective sensing of changes in wealth levels. Apart from this refutation, the so called Super St. Petersburg lottery is another way to refute utility theory, which we discuss in eq. (4.74). The main result states that the mathematical operation of applying the logarithm to wealth as a non-linear

[285] As the situation is often depicted, see for example SINN (1989, Ch. 3 A), SINN and WEICHENRIEDER (1993, p. 77), SINN (2003, pp. 98–99) and SEIDL (2013, p. 250).
[286] FECHNER 1869, p. viii.
[287] FECHNER 1869, p. 65.

transformation only works if the payouts increase at most exponentially. The flawed solution strategy of using utilities fails for more generally defined lotteries, e.g. if the payouts diverge super-exponentially. Thus EUT is incapable of providing a solution to the general problem of assigning a value to the offer of an arbitrary gamble.

Although a logarithmic utility function makes the series of the ordinary St. Petersburg lottery converge to a finite expectation value in eq. (4.48), we have see utility for what it is and acknowledge that it is only a trick. The purported success of utility in rationalising some empirical decision behaviour on a descriptive level, is simply the exploitation of an additional degree of freedom – psychological idiosyncrasies of the individual gambler, which can take arbitrary form. Utility as such is conceptionally empty in the normative sense of solving the decision problem. The mathematical operation of the non-linear transformation of wealth has nothing to do with the structure of the mathematical problem in general. This solution strategy is therefore far from being a general solution to decision problems, for the one reason that it is still based on virtual ensemble averages. Seen from this angle it is a trick in much the same sense as BOLTZMANN's ergodic hypothesis was a trick that happens to achieve the desired results, but can not be derived from more fundamental aspects of the problem at hand, *i.e.* the dynamics of the observable wealth in case of gambles.

Historically, there exists another functional form of utility that is worth to be mentioned which had been suggested by CRAMER. However, we can discuss it briefly, because in the words of BERNOULLI 'I have found his theory so similar to mine that it seems miraculous that we independently reached such close agreement on this sort of subject'.[288]

4.3.5.2 Cramer's Square-Root Utility

Besides his solution using a truncation of the payouts which we discussed in Sec. 4.2, CRAMER also proposed another solution that does not truncate extreme payouts but applies a non-linear transformation to wealth in a similar spirit to BERNOULLI. CRAMER proposed as the functional form of the utility function the square root function which we denote by $u_{\mathrm{C}}(w) = \sqrt{w}$.[289] In fact, CRAMER analysed only the payouts and did not introduce wealth, which is the same to say that CRAMER studied only the case for $w_t = 0$. Therefore, all the remarks on proper observables and the quantity $u(\Delta w) = u(x)$ apply. Formally, CRAMER tried to find a ticket

[288] BERNOULLI 1954, p. 33.

[289] Similar to the solution strategy of truncation, CRAMER's original solutions proposed in an exchange of letters survived in the collected works of JAKOB BERNOULLI (1975), but D. BERNOULLI cited extensively from the letters in his original text. See also the remarks in footnote 203 on p. 85.

price C_{fair}, which satisfies the following equation

$$u_{\text{C}}(C_{\text{fair}}) \stackrel{!}{=} \langle u_{\text{C}}(x)\rangle = \sum_{n=1}^{\infty} u_{\text{C}}(x_n)p_n \tag{4.63}$$

$$\sqrt{C_{\text{fair}}} = \sum_{n=1}^{\infty} \frac{\sqrt{2^{n-1}}}{2^n} \ . \tag{4.64}$$

Convergence Behaviour for Square Root Utility

In order to establish the convergence of the series the following simplifications are helpful, and as shown in the eq. (4.66) utilise an existing law for the series in the last step render the application of a convergence criterion needless

$$\sum_{n=1}^{\infty} \frac{\sqrt{2^{n-1}}}{2^n} = \sum_{n=1}^{\infty} \frac{\sqrt{2^{n-1}}}{2 \cdot 2^{n-1}} = \sum_{n=1}^{\infty} \frac{\sqrt{2^{n-1}}}{2\sqrt{2^{n-1}}\sqrt{2^{n-1}}} \tag{4.65}$$

$$= \frac{1}{2}\sum_{n=1}^{\infty} \frac{1}{\sqrt{2^{n-1}}} = \frac{1}{2}\left(2+\sqrt{2}\right) = 1 + \frac{1}{\sqrt{2}} \ . \tag{4.66}$$

Thus the square root function then leads as well to the convergence of the series in the expectation, and a numerical simulation yields the limit of the series in eq. (4.66)

$$C_{\text{fair}} = \left(1 + \frac{1}{\sqrt{2}}\right)^2 \approx 1.707^2 \approx 2.9 \ . \tag{4.67}$$

Thus, a gambler with zero initial wealth and a square root utility function has a willingness to pay a little less than 3 €. In **Fig. 4.2** we plot CRAMER's square root utility function and BERNOULLI's logarithmic utility function with an upper bound. Note the difference in the scales compared with linear (or synonymously no) utility in **Fig. 4.1**, they differ by an order of magnitude, which shows the impact of different non-linear concave transformations.[290]

Curvature of Utility Functions

If we summarise the two variants of functional forms, we see that different concave utility functions can in principle deliver arbitrarily different expected changes in utility, e.g. for square-root utility the expected utility of the St. Petersburg lottery is 2.9 € (see eq. (4.67) and for expected logarithmic utility the fair price depends on the specific logarithm and more importantly on the initial wealth. Still for both functional forms exists a finite and positive

[290] If the non-linear transformations are interpreted as utility functions, then the scale (absolute values) of utility has no meaning, because the concept of utility is ordinal and not cardinal, *i.e.* a utility function is only specified up to a class of affine transformations.

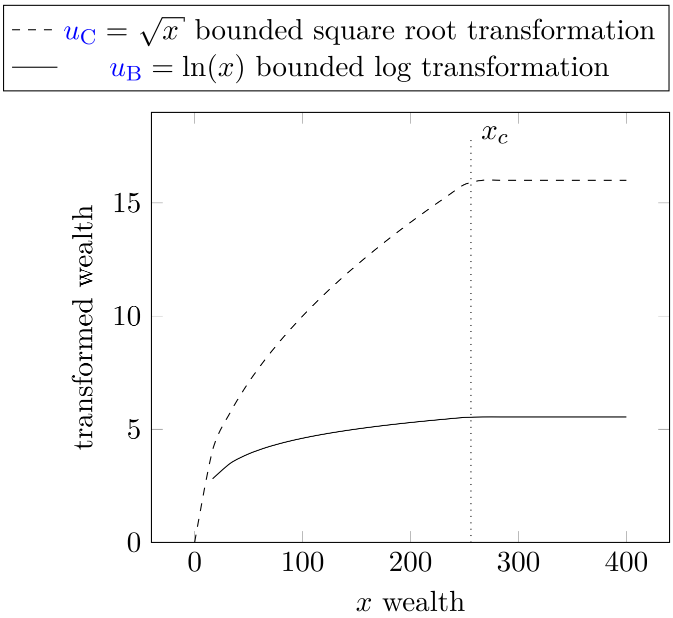


Figure 4.2: Non-Linear Transformation of the Payout Functions with an Upper Bound. Two examples of non-linear bounded and concave transformation functions are given here, the logarithmic function $u_B = \ln(x)$ (solid line) and square root function $u_C = \sqrt{x}$ (dashed line). An upper bound of nine coin tosses, $n_{\max} = 9$, gives an upper bound for both functions $\ln(x \geq x^*) = \ln(256) \approx 5.54$ and $\sqrt{x} \geq x^*) = \sqrt{2^{n_{\max}-1}} = \sqrt{2^8} = \sqrt{256} = 16$. Both transformed payout functions stay constant for any values greater than $x^* = 256$. Hence, there is zero marginal payout from gains above x_c.

expectation value, which advises to enter the gamble if offered at a price less than the fair price. However, utility functions are only evaluated cardinally, not in comparison of the absolute value they produce. Thus, the decision whether a gambler should enter the lottery solely depends on the offered price and does not depend on the specific functional form of utility up to affine linear transformations.

As a reaction to cardinal utility and belief that the results in Menger (1934) hold up, economics dismissed the unbounded logarithm and concentrated on the study of properties of families of affine transformations of a particular functional form for utility instead, most importantly their curvature. Within this retreat economics is still occupied with maintenance work of the ad hoc concept of utility. Surprisingly, it is again the logarithmic transformation of wealth, which has the convenient property within conventional economics of a constant curvature also known as the Arrow-Pratt measure of absolute risk aversion (ARA),[291]

(4.68) $$-\frac{u''}{u'} = 1 \qquad \text{for } u(w) = \log w \; .$$

[291] Arrow 1965, p. 37; Pratt 1964.

This measure stays constant with respect to affine transformations of utility functions for the class of logarithmic utility functions and is therefore referred to as constant relative risk aversion (CRRA). Of course, also increasing and decreasing absolute and relative risk aversions are possible in principle, which gives rise to the respective terminology of increasing absolute risk aversion (IARA), increasing relative risk aversion (IRRA), decreasing absolute risk aversion (DARA) and decreasing relative risk aversion (DRRA), but the CRRA is the standard type of utility functions taught to every undergraduate in economics.

4.3.6 Non-Linear Transformation of Probabilities

In this section we briefly mention the most popular non-linear transformation applied to the probabilities, which is called the probability weighting function in prospect theory (PT), which already involves a highly non-linear transformation of the (usually) payouts (and not wealth).[292] Kahneman and Tversky sought to understand empirical decision behaviour observed in lab experiments, which routinely led to violations of the axioms of EUT. Using a rhetorical trick, they often pretended to be less concerned on the derivation of a normative decision theory or what constitutes rational behaviour (here they more or less accepted EUT), and only strove for the best possible fit of a (very general) descriptive model to the empirical data.[293] Let us give a quintessential quote:

> "Because these [axioms of EUT] are normatively essential but descriptively invalid, no theory of choice can be both normatively adequate and descriptively accurate.[294]
>
> Prospect theory differs from the other models mentioned above in being unabashedly descriptive and in making no normative claims. It is designed to explain preferences, whether or not they can be rationalized.[295]"

Thereby they planted the seed for a strange dichotomy of descriptive and normative domains of scientific theories on which we elaborate in greater detail in 7. For Kahneman and Tversky EUT always served as a reference point of rationality. Deviations from EUT are thus perceived as irrational biases. In their opinion they devised a hypothetical decision

[292] Kahneman and Tversky 1979; Tversky and Kahneman 1992.
[293] See also the abstract, introduction and conclusion in Kahneman and Tversky (1979) or Tversky (1975). The general argument of th rhetorical trick is also acknowledged in Heukelom (2015): 'But Kahneman and Tversky's use of normative and descriptive can also be seen as a very clever way of trying to convince economists of the relevance of Prospect theory for economics. In Prospect theory, Kahneman and Tversky did not tell the economists that their theory was complete nonsense or useless. Instead, they claimed to understand economics as using one theory to cover both the normative and the descriptive realm. For the normative part they fully agreed with economists, which fitted in neatly with practice in behavioral decision research. But, Kahneman and Tversky argued, economists had been mistaken in using that same theory in the descriptive domain.'
[294] Tversky and Kahneman 1986, p. S271.
[295] Tversky and Kahneman 1986, p. S272.

theory on a descriptive level. On the one hand they denied much relevance of rationality on the empirical domain, but one the other hand both pointed repeatedly to irrational empirical behaviour and thereby making tendentious references to the respective domain of decision theories.

However, a whole research programme of *bias & heuristics* stem from their work and much of today's experimental and behavioural economics can be seen as a continuation of it. One element in the decision theory which this research programme devised is called PT and contains both a non-linear transformation of payouts and according to (cumulative) prospect theory[296] decision makers non-linearly transform the objective probabilities involved. The non-linear transformation of wealth in EUT and in PT constitutes already a subjective weighting of wealth and as we have shown introduces an additional (psychological) degree of freedom which is hard to pin down empirically. A non-linear transformation of the objective probabilities suffers from the same problems. Here a subjective weighting of the probabilities introduces several additional degrees of freedom, because the weighting function usually contains at least two additional parameters sometimes even more. The weighting of the probabilities is as ad hoc as the truncation were.

Prospect Theory as Anti-Bernoullian

More precisely, the behaviour of the participants seemed to Kahneman & Tversky as if they overrate small probabilities close to zero and underrate large probabilities close to unity, which leads to the characteristic inverse S-shaped probability weighting function depicted in **Fig. 4.3** and denoted as π.

Decision makers in PT treat probabilities in exactly the opposite way as N. Bernoulli's strategy using 'moral impossibility', which was a truncation of low probabilities close to zero. By the time of the arrival of probability weighting the combinatorial space of truncations and non-linear transformations has been exhaustively explored. Furthermore, we see ad hoc transformations of probabilities in opposite directions over the course of the history of decision theory, which justifies the labelling of the behaviour of probability weighting as anti-Bernoullian.

The functional form used in Tversky and Kahneman (1992, p. 309) and depicted in **Fig. 4.3** is

(4.69) $$\pi\left(p\right)^{*} = \frac{p^{\gamma}}{\left(p^{\gamma} + (1-p)^{\gamma}\right)^{-\gamma}} \ ,$$

[296] Kahneman and Tversky 1979, p. 283; Tversky and Kahneman 1992, p. 313.

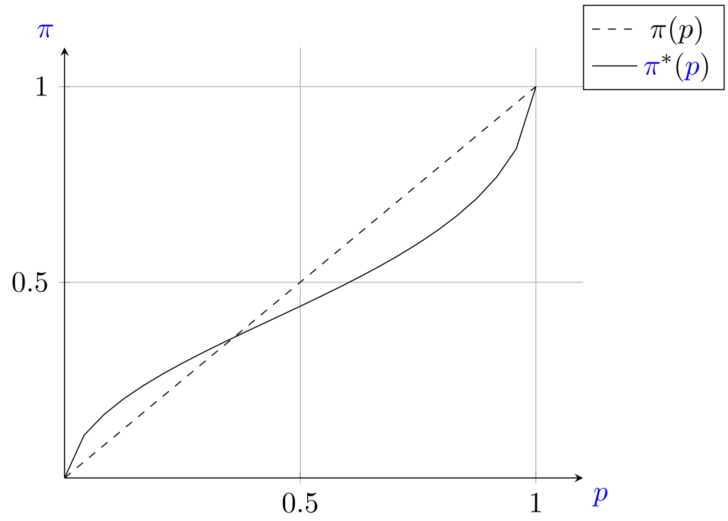


Figure 4.3: Probability Weighting Function in Prospect Theory. Function $\pi(p) = p$ gives the objective probability (dashed line). Function $\pi^*(p)$ shows the fit of a typical experimentally observed probability weighting function (solid line). The functional form follows eq. (4.69) with $\gamma = 0.65$.

and introduces one additional parameter, γ. For parameter values $\gamma < 1$ eq. (4.69) produces the postulated overweighting of low probabilities and underweighting of moderate to high probabilities. Different functional forms of the probability weighting function have been proposed in the literature, usually they are associated with at least two or sometimes even more additional parameters.[297]

An arbitrary gamble is denoted in PT by the enumeration of the payouts and respective probabilities $f_X = (p_1, x_1; \ldots; p_n, x_n)$, which is in essence the probability mass function (PMF) f_X of the discrete random variable X. Then PT evaluates this gamble by computing the quantity $\mathfrak{V}$ also called a prospect

$$\mathfrak{V}(f_X) = \sum_{n=1}^{\infty} \Delta\mathfrak{v}(\Delta w)\,\pi(p_n) = \sum_{n=1}^{\infty} \mathfrak{v}(x_n)\,\pi(p_n) = \sum_{n=1}^{\infty} \mathfrak{v}_n \pi_n \,, \tag{4.70}$$

whereby the value function $\mathfrak{v}$ is a kind of non-linear transformation similar to utility. If we analyse the structure of $\mathfrak{V}(f_X)$ further, we realise that PT arrived at the most general form to write a weighted arithmetic mean. Note, that for several reason PT signifies a

297 For example Prelec (1998, 2000) derived a functional form with in total six additional parameters that account for so called compound invariance. Compound invariance is composed of the two principles subproportionality and subadditivity of the decision weights, besides the above mentioned over-/ & underweighting. Usually the parameter for losses is different from the one for gains, such that a weighting function with n parameters actually introduces $2n$ parameters through the backdoor. See also Glimcher and Fehr (2014) and for different functional forms and a survey on empirical parameter fits especially Fox and Poldrack (2014, Table A3).

step backwards e.g. the dependence on initial wealth is lost again. This evolution is wrong directed.

Wrong Directed Evolution of Ensemble-Based Decision Theories

Eq. (4.71) visualises the wrong directed evolution of decision theories as carrying the generalisation of ensemble averages to the extremes. Originally the ensemble average was a sum of the state dependent product of the probabilities and respective payouts. When the truncations in Sec. 4.1 and Sec. 4.2 can be interpreted as applying an ad hoc function to either the probabilities or the payouts, PT tops that and wraps a function around every single component.

$$\text{(4.71)} \quad \text{Huygens:} \quad \langle \Delta w \rangle = \sum_{n=1}^{\infty} \underbrace{\Delta w}_{\text{absolute wealth}} \cdot \underbrace{p_n}_{\text{objective probability}}$$

$$\text{(4.72)} \quad \text{EUT:} \quad \langle \Delta u(w) \rangle = \sum_{n=1}^{\infty} \underbrace{\Delta u(w)}_{\text{function of wealth}} \cdot \underbrace{p_n}_{\text{objective probability}}$$

$$\text{(4.73)} \quad \text{PT:} \quad \mathfrak{V}(f_X) = \sum_{n=1}^{\infty} \underbrace{\mathfrak{v}(\Delta w)}_{\text{function of wealth}} \cdot \underbrace{\pi(p_n)}_{\text{function of probability}}$$

This thesis turns away from this wrong directed evolution towards an alternative decision theory along the lines of *Ergodicity Economics* which rises beyond the presented evolution due the idea of embedding the randomness within time. We refer to chapter 7 and Sec. 4.5.

4.3.7 Super St. Petersburg Lotteries

As described in the preceding section EUT was one attempt to tame the paradoxical situation of infinite mathematical expectation by using the logarithm to transform wealth which is driven by the exponentially increasing payouts into a linearly increasing utility of wealth. However, Menger (1934) showed that it is always possible to find a payout function of a lottery for which the expectation value and the expected logarithmic utility is infinite as well. In other words, we can create a gamble with a specifically designed payout function that leads to a failure of EUT even after the utility transformation is applied. In other words, the decision maker is confronted anew with the original paradoxical decision problem of infinite expectation after applying any non-linear transformation to wealth. In the literature this type of modified St.

Petersburg lottery is often referred to as *Super St. Petersburg lottery.*[298] Using BERNOULLI's criterion MENGER (1934, p. 468) falsely concluded the following:

> *For every unbounded function u which transforms wealth increments Δw in the calculation of the mathematical expectation $\mathrm{E}\left[\Delta u\left(w\right)\right] = \sum_n \Delta u\left(w\right) p_n$, one can create a payout scheme $x\left(n\right)$ of a lottery (similar to the St. Petersburg lottery), such that the expectation value of the gambler's transformed wealth diverges again, $\mathrm{E}\left[\Delta u\left(w\right)\right] = \infty$.*

This implies one can set up the rules of the lottery in such a way that the increase in the payouts will exceed any concaveness of a function u unless it is bounded. The general idea of the Super St. Petersburg lottery trick is to let the payout function grow at least as fast as the probabilities shrink with an increasing number of coin tosses, even after a non-linear transformation was applied to the payouts, e.g. the logarithm. If the payouts and probabilities scale identically except the sign, then the infinite series in the sum in e.g. eq. (4.50) keeps diverging. More specifically, we have to set up the rules in such a way that the payouts escape or at least balance the force of the convergence towards zero applied in eq. (3.8) by the decreasing probabilities, 2^{-n}, on the n^{th} coin toss. In physics, this is known as a problem that involves different scaling behaviours.

Therefore, MENGER assumed the gambler is offered a similar lottery as before but where the rate of divergence of the payouts is even higher than exponential, e.g. an iterated exponentiation[299], then he is faced with what is known in the literature as a *Super St. Petersburg Lottery.*[300] Therefore, let us assume $x\left(n\right) = e^{2^{n-1}}$ and a logarithmic utility function $u\left(w\right) = \ln w$, then

[298] See for example SAMUELSON (1977).

[299] An iterated exponentiation is called a tetration, one possible notation is ${}^{n}a := \underbrace{a^{a^{\cdot^{\cdot^{\cdot^{a}}}}}}_{n}$, which is attributed to RUDY RUCKER.

[300] MENGER 1934, p. 467.

the expected change of the utility of wealth is given by

$$\mathrm{E}\left[\Delta u\left(w\right)\right] = \sum_{n=1}^{\infty} \Delta u\left(w\right) p\left(n\right) = \sum_{n=1}^{\infty} u_{\mathrm{C}}\left(w_{t+\delta t}\right) p\left(n\right) - u\left(w_t\right) \tag{4.74}$$

$$= \sum_{n=1}^{\infty} \frac{\ln\left(w_t + x\left(n\right) - C\right)}{2^n} - u\left(w_t\right) \tag{4.75}$$

$$= \sum_{n=1}^{\infty} \frac{\ln\left(w_t + e^{2^{n-1}} - C\right)}{2^n} - \underbrace{u\left(w_t\right)}_{\text{const.}} \tag{4.76}$$

$\forall\ C, w_t < \infty$ and $n \to \infty$ the series is dominated by the exponential terms

$$\lim_{n\to\infty} \frac{\ln\left(w_t + e^{2^{n-1}} - C\right)}{2^n} \approx \lim_{n\to\infty} \frac{\ln e^{2^{n-1}}}{2^n} = \lim_{n\to\infty} \frac{2^{n-1}}{2^n} = \frac{1}{2} \tag{4.77}$$

$$\hookrightarrow \mathrm{E}\left[\Delta u\left(w\right)\right] = \sum_{n=1}^{\infty} \frac{\ln\left(w_t + e^{2^{n-1}} - C\right)}{2^n} - u\left(w_t\right) \to \infty\ . \tag{4.78}$$

The series in the expectation diverges again and we have another refutation of utility theory. The gambler or decision maker is confronted with the exact same situation as before, for which EUT was introduced. As a result the logarithmic function in EUT constitutes no solution of a generalised version of the paradox of which the St. Petersburg lottery can only be considered a special case.

4.3.7.1 The Alleged Necessity of Unbounded Utility Functions

From this result Menger (1934) concluded incorrectly the additional and much more renowned claim, that therefore only bounded utility functions can successfully resolve paradoxes of this kind, which can be considered as a solution strategy that rests again on a kind-of-truncation. E.g. a candidate utility function would be $u\left(w\right) = 1 - e^{-w}$, which is bounded from above by 1. Peters and Gell-Mann (2016, p. 8) point out that 'Menger implicitly ruled out the all-important logarithmic function that connects utility theory to information theory and provides the most natural connection to the ergodicity argument.' The logarithm appears in this thesis in exactly the mentioned contexts in Sec. 5.3, Sec. 4.5 and chapter 8. Another surprising fact is that the Super St. Petersburg lottery is known by fewer economists than the false ad hoc restriction to bounded utility functions. Subsequently, bounded utility functions flourished among the economics community, and Menger (1934) became an influential paper held as 'a modern classic that [...] stands above all criticism' by many of the leading figures in economics.[301]

[301] Quote from Samuelson (1977, p. 49), similar appraisals can be found in Arrow (1951), Markowitz (1976), Shapley (1977b,a), Aumann (1977) and Seidl (2013).

However, PETERS (2011a) realised that MENGER worked not with LAPLACE's correct(ed) version of EUT but started from BERNOULLI's flawed subgambles given in eq. (4.25), which is plausible given that as the becoming translator he may have known BERNOULLI (1738) even better than LAPLACE (1812, Ch. X). MENGER began his derivation with BERNOULLI's strange quantity $\left\langle \Delta u^{+} \right\rangle$ and completely ignores the negative divergence in the second term Δu^{-}.[302] PETERS' analysis scrutinises both convergence behaviours. The flaw in MENGER (1934) is indicated after eq. (4.76), but to derive it in full we had to anticipate too many components which are explained in greater detail only in Sec. 4.5. For the detailed presentation of the argument we thus refer to the original references[303] and refrain from a deeper immersion in the topic at this point.

All solution strategies discussed so far, required the ad hoc change of a rule of the original lottery. The solution strategy discussed in Sec. 4.4 introduces kind of a mode of repetition to the evaluation of gambles through the backdoor, contained in an agreement between the gambler and supplier of the gamble on the truncation. The effect of reasonable modes of repetition will be a key point when we handle the non-ergodic wealth dynamics explicitly in Sec. 4.5. In Sec. 3.3 the notion of fairness was derived in eq. (3.4). FELLER (1945) derived a notion of fairness using the modern language of probability theory, which will reveal some surprising results. In doing so we follow the classic FELLER (1957, ch. X.3 and X.4) to a large extent.

4.4 Solution Strategy 4: Limiting the Number of Rounds

In the language of modern probability theory the outcome one round of the lottery is conceived as a realisation of a random experiment, thus the payouts x are realisations of a random variable X. The driving random variable in X is of course the number of coin tosses n, but as we have emphasised before X is here a function of a random variable and thus a random variable as well. Furthermore, because the coin toss has only the two outcomes 'heads' or 'tails', thus the coin toss is a specific random experiment called a BERNOULLI trial.[304] This information generates the probability space of the coin toss, that consists of the sample space $S = \{\text{heads, tails}\}$, a set of events $\mathcal{A}$ and a probability measure $\mathcal{P}$, which is a function that assigns values between 0 and 1 to the events, $\mathcal{P} : \mathcal{A} \to [0, 1]$. Thus a random experiment is uniquely defined via the probability space $(S, \mathcal{A}, \mathcal{P})$.[305] Now, the derivation of the fair ticket price can be generalised using the weak law of large numbers (WLLN).[306] Therefore,

302 MENGER 1934, p. 465.
303 PETERS 2011a; PETERS and GELL-MANN 2016.
304 Referring to JAKOB BERNOULLI.
305 The concept of the probability space together with the axiomatisation of probability theory traces back to KOLMOGOROFF (1933).
306 The theorem of the weak law of large numbers goes back to JAKOB BERNOULLI (1713).

we consider N identical coin tosses as our random experiment, where N such repeated coin tosses generate a sequence of payouts $\{Z_k\}_{k\in[1,2,\ldots,N]}$, which are iid random variables with the same cumulative probability distribution function (CDF) $F_X(x)$ or simply $F(x)$. Let the expectation value of the payout $\langle X\rangle$ exist and given by

$$\langle X\rangle = \int_{-\infty}^{\infty} x\,\mathrm{d}F(z) \ , \tag{4.79}$$

or for discrete random variables

$$\langle X\rangle = \sum_{n=1}^{N} x_n \cdot p_n \ . \tag{4.80}$$

We denote by Z_T the accumulated payouts after N coin tosses until the stopping condition of the lottery happens

$$Z_N = X_1 + X_2 + \cdots + X_N \ . \tag{4.81}$$

If the number of trials tends to infinity $N \to \infty$, then the WLLN[307] assures for every $\varepsilon > 0$, that the accumulated payouts Z_N will in probability be of the order of $N\langle X\rangle$

$$\Pr\left\{\left|Z_N - N\langle X\rangle\right| < \epsilon N\right\} \to 1 \ , \tag{4.82}$$

or, put differently, for large N the accumulated net gain $Z_N - N\langle X\rangle$ is small compared to N.[308] Hence, the modern definition of $\langle X\rangle$ as a 'fair' price or entrance fee. If the supplier of a lottery offers an entrance C smaller than $\langle X\rangle$, then in the long run, *i.e.* for many trials N, the gambler very likely accumulates a positive net gain and the game is said to be *favourable* to him and vice versa for the *unfavourable* case. It seems appropriate to note at this critical juncture, that Feller (1957, p. 234) points out the decisive difference between 'fair' and 'favourable' that has been overlooked in classical probability theory. Therein a factual entrance fee $C = \langle X\rangle$ was labelled a fair price and respectively a fair game. This caused much confusion. Eq. (4.82) is in fact a statement about the size of the net gain in relation to the number of trials *in the long run* and nothing is revealed about its sign, which would render it favourable or not. Implicitly, the WLLN in eq. (4.82) is supplemented by the central limit theorem (CLT), which takes into account the fluctuations of the random variable as well, *i.e.* expressed in the variance $\mathrm{Var}[X]$. Therefore, the validity of the statement only applies in the sufficiently *long run*. From this follows that fair games can very well be favourable or unfavourable to the gambler. Feller (1957, p. 235) even refers to Feller (1945), which illustrates the case of a fair game which is unfavourable for the gambler by the example of the St. Petersburg lottery.

307 Feller 1945.
308 Feller 1957, p. 234.

FELLER (1957, p. 235) concludes 'when the variance is infinite, the term "fair" game becomes an absolute misnomer.'

Furthermore, the LLN only applies to random variables with finite expectation values (eq. (4.80)). Because premodern probability theory had no available tools to handle infinite expectation values, they posed insurmountable difficulties. In total this led to decision theories with absurd conclusions, e.g. the one given in the opening quote of this section or the ad hoc concept of moral impossibility. Random variables with infinite expectation are far from being mere pathologies, that could be neglected in practical circumstances, but play an important role in applications in modern economics, physics and statistics. E.g. such random variables are found as waiting and first-passage times in the simplest stochastic processes. In fact, FELLER among others developed generalised versions of limit theorems and generalised laws of large numbers to include such cases.

FELLER's solution proceeds as follows. He noted that it is the assumption of a fixed ticket cost throughout the repeated trials, which leads classic probability theory and thus standard decision theory to the advice of always entering a game with infinite expected payouts. The solution proceeds from eq. (4.82) and aims to correct for the treatment of '∞' as if it were an ordinary number. The infinity of the possible payouts in eq. (3.8) will sooner or later surpass the accumulated fixed costs and thus it seems as if the infinite mathematical expectation is realised in the long run. He thus concluded, that a fair price must vary with the number of trials.[309] Therefore, another variable is introduced C the entrance fee in FELLER (1945), which increases with the number of trials, but does not change between the trials. Let the accumulated costs be denoted by C_N,

$$C_N = N \langle X \rangle \ , \tag{4.83}$$

then a notion of a 'fair' cost is retained which satisfies the adjusted limit theorem

$$\Pr\left\{|Z_N - C_N| < \epsilon C_N\right\} \to 1 \tag{4.84}$$

$$\Pr\left\{\left|\frac{Z_N}{C_N} - 1\right| < \epsilon\right\} \to 1 \ . \tag{4.85}$$

In order to satisfy eq. (4.85) the accumulated costs have to keep up with the increasing accumulated gains, such that the ratio is close to unity. This is again a scaling problem. The magnitude of the expected payout from a single trial is $\log N$, because the expected payouts increase if the number of trials increases by powers of two. To put it another way, the gains scale with the logarithm of the length of the longest trial or the number of trials. Thus for N repeated trials, the accumulated gains are of the order of magnitude of $Z_N \sim N \log N$.

309 FELLER 1945, p. 302.

Table 4.5: Feller's fair ticket prices of the St. Petersburg lottery. The table shows how the fair ticket price scales with the increasing number of rounds which is equivalent to say with exhaustively sampling the ensemble of size N.

N	$C_N = N\langle X\rangle = N\log N$	$C_{\text{fair}} = \langle X\rangle$
1	0.0	0.0
2	2.0	1.0
3	4.8	1.6
4	8.0	2.0
5	11.6	2.3
6	15.5	2.6
7	19.7	2.8
8	24.0	3.0
9	28.5	3.2
10	33.2	3.3
20	86.4	4.3
50	282.2	5.6
100	664.4	6.6
1000	9965.8	10.0
10000	132877.1	13.3
100000	1660964.0	16.6
1000000	19931568.6	19.9

Therefore the accumulated costs have to scale in the same way to yield a fair price for the (St. Petersburg) lottery. The variable cost C is set in such a way that the accumulated costs scale with the order of $C_N \sim N\log N$.[310] The fair price has to scale at most like $N\log N$ but more than $\log N$ and is thus given by

$$N < 2^{C_{\text{fair}}} \leq N\log N \quad |\ \log() \tag{4.86}$$

$$\log N < C_{\text{fair}} \leq \log(N\log N) \tag{4.87}$$

$$\log N < C_{\text{fair}} \leq \log N + \log\log N \quad . \tag{4.88}$$

Fig. 4.4 depicts the corridor for the fair ticket fee defined by eq. (4.88). *I.e.* the range of the costs in comparison with the logarithm of the rounds as a lower bound and the logarithm of the rounds plus an additional tiny term as the upper bound. **Table 4.5** shows the scaling of the fair tickets costs for increasing trials.

The fair price thus depends on the longest trial observed in all the trials, which is reasonable but not apriori knowable. However, how to apriori derive the longest trial? FELLER's solution

[310] The proof utilises CHEBYSHOV's inequality and can be found e.g. in FELLER (1945, p. 303, 1957, p. 237).

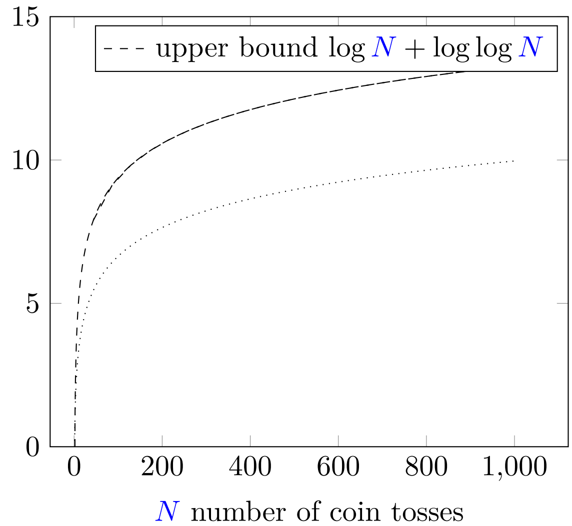


Figure 4.4: Feller's Corridor of a Fair Cost. From eq. (4.88) the depicted corridor is derived in which the fair fee of lottery ticket scales with the number of coin tosses allowed. The upper bound is given by the dashed line and the lower bound by the dotted line. *I.e.* the gambler pays more for more coin tosses.

derives a fair price for the version of the lottery, where it is assumed that the banker and the gambler agree upon a fixed and finite maximum number N of coin tosses in a single trial or in a sequence of trials prior to the execution. In fact, FELLER's solution can be reinterpreted as a situation in which the gambler and/or the supplier agree on the truncation of the probabilities/payouts. The gambler pays for every extra number of coin tosses. More coin tosses bear the chance of a higher payout and therefore justify a higher price. But again it was necessary to introduce an additional assumption to the original problem in order to let the expectation value converge.

So far all solution strategies follow the conception of embedding randomness in the ensemble, because they all rely on the calculation of expectation values or derivatives thereof. We now turn to a resolution of the paradox which relies on the other possible embedding of randomness within time, which will turn out to be relevant for the individual decision maker.

4.5 Solution Strategy 5: Non-Ergodic Dynamics

At the beginning of the solution strategy 5 is the acknowledgement of the ergodicity problem in economics, which has been completely ignored by the strategies 1-4 so far. Associated with a resolution using non-ergodicity is the rejection of the idea that the ensemble or parallel universes are relevant and the acceptance of the irreversibility of time. That is to say, what is relevant for an individual decision maker is what happens to his wealth over the course of time and not possible incidents in parallel universes within a virtual ensemble. Thus this strategy constitutes the new paradigm of an embedding of the randomness within time. The solution strategy 5 is due to the seminal paper in Peters (2011c), which started a new field now called *Ergodicity Economics*. Peters follows a completely different track than all the solution strategies presented earlier, inasmuch as no ad hoc assumptions are required. *I.e.* no non-linear distortions of strategy 3 (no utility), no truncation of neither payouts nor probabilities nor bending of the original rules as it is done in strategies 1, 2 and 4.

Solution strategy 5 relies only on the introduction of a mode of repetition which is by no means ad hoc. Multiplicative dynamics turn out to be a completely natural choice in the context of wealth accumulation. Indeed the most common manifestation of multiplicative growth in economics and beyond is the compound interest effect in investment. Furthermore, understanding the dynamics of wealth following a stochastic multiplicative growth process is natural, because many evolutionary processes of cooperation amongst diverse entities ranging from 'cells, organisms, families, herds, companies, institutions'[311] up to whole economies are modelled via multiplicative growth processes.[312] Multiplicativity in this context simply means that the resources available for possible investments in a given round are determined by the outcome of the previous round(s), and that the payouts are proportional to the amount wagered in every round. Mathematically, the key novelty is the computation of time-average quantities, e.g. the growth rate of wealth that is generated by repeated gambles.

The section is organised as follows. We commence with a comparison of additive and multiplicative wealth dynamics and introduce all the necessary new quantities in Subsec. 4.5.1 first and foremost the growth rate. Equipped with the newly acquired tools we derive the growth rate of the expected wealth for gambles in general and the St. Petersburg lottery in particular in Subsec. 4.5.2. Analogously we derive the time-average growth rate of gambles in general and the St. Petersburg lottery in particular in Subsec. 4.5.3. Thereby we are led to ergodicity transformations and the role of the time scale in Subsec. 4.5.4, and finally combine the insights to derive the non-ergodicity of multiplicative dynamics in Subsec. 4.5.5. Here we contrast the key results with seeming similarities to the EUT and elaborate what exactly breaks the ergodicity. In Subsec. 4.5.6 we conclude this key section of the thesis. Our notation follows as strictly as possible the notation in Peters (2011c) and Peters and Adamou (2018b), which facilitates a smooth transition when proceeding to the key original sources on *Ergodicity Economics*.

---

311 Peters and Adamou 2018a.
312 Aitchison and Brown 1963; Lewontin and Cohen 1969; Redner 1990; Nowak 2006.

4.5.1 Additive and Multiplicative Dynamics

Let us imagine a bookie offers the following gamble at zero cost for the moment.[313] Contingent on the outcome of tossing a fair coin the gamble offers the payout of either +10 € or -10 €. We already introduced generic gambles with an additive dynamic in eq. (4.7), where we interpreted the payout (minus the cost) as a wealth increment. Because the coin toss has only two possible outcomes, it is still manageable to synonymously express the gamble in the following way,

$$w_{t+\delta t} = w_t \quad + \quad \begin{cases} +10\,€ & \text{if 'heads'} \\ -10\,€ & \text{if 'tails'} \end{cases} . \tag{4.89}$$

An additive dynamic encodes a situation in which the payout from a gamble is independent from the amount wagered. Gambles with additive dynamics are almost exclusively what is studied in economic lab experiments and the initial wealth level of the subjects remains an uncontrolled variable, because it is not relevant in the additive regime. The increments in additive dynamics have no memory and are thus not path dependent. But a bookie could also offer the following gamble. Contingent on the outcome of tossing a fair coin the gambler either receives a multiple of what he has wagered, let it be 10% or pay a multiple of what he has wagered, let it be -10%. This gamble is more realistic because it introduces an additional dependence of the payouts, which are not only contingent on the realisation of a random event but also on the amount wagered. Thus the wealth increments under multiplicative dynamics become path dependent. The fact that payouts are proportional to the wagered amount can be denoted as follows

$$w_{t+\delta t} = w_t \quad \cdot \quad \begin{cases} +10\% & \text{if 'heads'} \\ -10\% & \text{if 'tails'} \end{cases} . \tag{4.90}$$

Gambles of the type of eq. (4.90) generate multiplicative wealth dynamics and are also equivalently represented in a slightly different way,

$$w_{t+\delta t} = w_t \quad \cdot \quad \begin{cases} 1.1 & \text{if 'heads'} \\ 0.9 & \text{if 'tails'} \end{cases} , \tag{4.91}$$

using the growth factors 1.1 or 0.9, which express the rate of change and replace the additive increment. E.g. a gambler wagers all his fortune of 100 € and wins (loses) 10% by the first

[313] The cost of a gamble can always be factored in the state-dependent payouts and does not need to appear as a term explicitly. E.g. the outcomes in eq. (4.89) can be interpreted as containing a cost, which is already subtracted from the stated payouts. For example a real-world bookie would offer a coin toss gamble with a tiny difference in the absolute amounts of the payouts, eg. 10 € and -9.9 €.

round of the gamble given in eq. (4.91) and decides to play again. For the second round he is now endowed with 110 € (90 €) available for investing. In the second round, he can not get from 110 € to the initial 100 € as would be possible in the additive gamble, but he will get either $110\,€ \cdot 0.9 = 99\,€$ in the unfavourable case or $110\,€ \cdot 1.1 = 121\,€$ in the favourable case and so on and so forth ... . This dynamic regime is different to an additive regime, where the wealth increment of one round does not affect the wealth increment of another round.

In order to generically study the evolution of wealth generated by repeated gambles, we have to take the relevant dynamic explicitly into account, in case of wealth this is most likely a multiplicative dynamic. This consideration of a dynamic automatically embeds the randomness within time. For the case of wealth it follows quite naturally that it takes place in a regime of multiplicative dynamics for which the following derivations apply.

Repeated Gambles under Multiplicative Dynamics

Compared to an additive wealth dynamic in eq. (4.7), a multiplicative dynamic of the wealth of a gambler is given by

$$w_{t+\delta t} = w_t \cdot r_{t+\delta t} \,, \tag{4.92}$$

with the rate of change r or synonymously the growth factor. For a generic binary gamble we can write

$$w_{t+\delta t} = w_t \quad \cdot \quad \begin{cases} r_1 & \text{with probability } p_1 \\ r_2 & \text{with probability } p_2 = 1 - p_1 \,, \end{cases} \tag{4.93}$$

with the contingent growth factors denoted by r_1 and r_2. This looks now familiar to the gamble from the example in the beginning of this section which we get for $p_1 = p_2 = 1/2$ and $r_1 = 1.1$ and $r_2 = 0.9$ in eq. (4.91).[314] Because the growth factors are random variables this is also called noisy or stochastic multiplicative growth. Taking eq. (4.92) and moving the growth factor to the LHS and with the help of eq. (4.7), we derive the growth factor for a single round for lotteries in its general form as

$$r_{t+\delta t} = \frac{w_{t+\delta t}}{w_t} = \frac{w_t + x_{t+\delta t} - C}{w_t} \,, \tag{4.94}$$

here $x_{t+\delta t}$ denotes the payout in that round of the gamble. It also reveals how the sequence of per round growth factors $(r_{t+\tau\delta t})_{\tau\in\mathbb{N}}$ is derived from the sequence of per round

[314] Note that the subscript of the growth factors is not their time index, because eq. (4.93) considers only the elements in the ensemble of possible growth factors.

payouts $(x_{t+\tau\delta t})_{\tau\in\mathbb{N}}$. In fact under multiplicative dynamics the change is expressed using a random wealth multiplier instead of a random wealth increment. Here we see that the latter is a stochastic process and thus the former as well, because it is a function of the latter. Taking the logarithm of eq. (4.94), we derive a relation of which we make strong use of

$$\ln r = \ln w_{t+\delta t} - \ln w_t \tag{4.95}$$

$$\ln r = \delta \ln w \; . \tag{4.96}$$

Getting from the Growth Factor to the Growth Rate

Because we deal with repeated gambles under multiplicative dynamics we are interested in what happens over time, *i.e.* for several rounds T of the gamble over a period Δt and a single round has the duration δt, thus $\Delta t = T\delta t$. We are especially interested in the conversion of the growth factor per time interval Δt into a growth rate for this specific period, denoted by g, such that $\exp(gT)$ yields the growth factor if T rounds of the gamble are played. Transposing the following equation using the logarithm on both sides and remembering to consider the duration of a single round δt this yields the exponential growth rate,

$$r_{\Delta t} = \prod_{\tau=1}^{T} r_{\tau\delta t} = e^{g\Delta t} = e^{gT\delta t} \tag{4.97}$$

$$\sum_{\tau=1}^{T} \ln r_{\tau\delta t} = gT\delta t \tag{4.98}$$

$$g = \frac{\sum_{\tau=1}^{T} \ln r_{\tau\delta t}}{T\delta t} \; . \tag{4.99}$$

Now we consider only a single round $T = 1$ of length δt in time units and use the just derived equation for the growth rate (4.99) and the equation for the growth factor (4.94) this yields the single round exponential growth rate per round of the gamble

$$g = \frac{1}{\delta t} \ln r = \frac{1}{\delta t} \ln \left(\frac{w_{t+\delta t}}{w_t} \right) = \frac{(\ln w_{t+\delta t} - \ln w_t)}{\delta t} = \frac{\delta \ln w}{\delta t} \; . \tag{4.100}$$

Especially from the last two terms in eq. (4.100) we re-derived eq. (4.96) we see more clearly what a growth rate is, namely the difference in the logarithms of a variable per some unit of time and we established a direct relation between our main observable *growth rate of wealth* and the observable *changes in the logarithms of wealth*. Meanwhile we have introduced the concepts of the growth factor and growth rate. In the following two subsections we appropriate both concepts to scrutinise the ergodicity of the observable *growth rate of wealth*. We start in Subsec. 4.5.2 with the derivation of the exponential growth rate of the ensemble average of wealth $g_{\langle\rangle}$ and continue with the derivation of the time-average growth rate of wealth $\overline{g}$ in Subsec. 4.5.3.

4.5.2 Ensemble-Average Growth Rate

Let $\langle r \rangle_N$ denote the ensemble-average growth factor for an ensemble of finite size of N realisations. Then the ensemble-average growth factor is given by the arithmetic mean of the individual growth factors of the N realisations

$$\langle r \rangle_N = \frac{1}{N} \sum_{n=1}^{N} r_n \ . \tag{4.101}$$

As simple as eq. (4.101) may look at a first glance, it is instrumental to our topic to understand what is actually computed here. The N realisations can be conceived as independent copies of the single gambler, who each realise one of the N possible outcomes of the random variable in their parallel universe. The gamblers are each playing the gamble in their isolated universe and generate their individual r_n on the RHS. But the quantity on the LHS pretends that – by averaging – all the copies of the single gambler in their N distinct universes on the RHS somehow pool their growth factors and split them evenly among each other. See also the annotation in the equations (3.10-3.12). The quantity $\langle r \rangle$ on the LHS is then treated as 'the expected' growth factor – as if it were meaningful for an individual gambler. In the majority of cases, when the ensemble-average growth factor can not be achieved by any individual it is still called the 'expected' growth factor.[315] Think of the fact that nobody has ever rolled a regular dice and thrown the expected value of 3.5. In this sense the quantity called 'expected value' 'fails to capture the reality of the situation'[316] and can become a hazardous misnomer. This is exactly the embedding of the randomness of the outcome (of the random variable r in this case) in the ensemble, which has been introduced in Sec. 2.1.1.

In the next step of the derivation of the growth rate of the expected wealth, the summation is changed from running over the N parallel worlds in eq. (4.101) to run over the length of consecutive coin tosses n in eq. (4.102). The frequency of a particular sequence of consecutive coin tosses is geometrically distributed, up to the largest sampled sequence of consecutive coin tosses, $n_N^{\max}$. In other words, $n_N^{\max}$ is the largest sequence of consecutive coin tosses that has been observed in a sample of N gamblers. Almost surely the length of this sequence is much smaller than the size of the ensemble, $n_N^{\max} \ll N$, and thus always finite. Therefore, every growth factor r_n in the sum in eq. (4.102) associated with a length of the sequence being n is weighted by its absolute frequency k_n in the sample of N parallel universes, which is in our

[315] The majority of cases is the set of non-ergodic dynamics.
[316] Peters 2011c, p. 4916.

case the number of copies of gambler all playing in their universe,

$$\langle r \rangle_N = \frac{1}{N} \sum_{n=1}^{n_N^{\max}} k_n r_n \;. \tag{4.102}$$

This implies that not every possible value of r_n must realise, for many the weight will be zero, especially the very high growth factors generated by extremely long sequences of 'tails' could not realise above some threshold, for which one would have to toss the coin even more often than (only) N (finite) times. In principle we could and we do study the infinite case in a moment. Similar to what has been observed in Sec. 2.1.1, if the number of gamblers N factored in approaches infinity, the law of large numbers has two effects. First, the relative frequencies converge to the true probabilities of n coin tosses, $k_n/N \to p_n$, because we sample the state space effectively, and second, the highest sampled coin toss in the infinite size ensemble diverges to infinity[317], $\lim_{N\to\infty} n_N^{\max} \to \infty$, which yields the infinite-ensemble average growth factor

$$\langle r \rangle \coloneqq \lim_{N\to\infty} \langle r \rangle_N = \lim_{N\to\infty} \sum_{n=1}^{n_N^{\max}} \frac{k_n}{N} r_n = \sum_{n=1}^{\infty} p_n r_n \;. \tag{4.103}$$

Finally, we got rid of the subscript and call this quantity simply the ensemble-average growth factor, $\langle r \rangle$. The ensemble-average growth factor is a regular scalar number and no random variable anymore. Similarly, using eq. (4.100), we get the infinite-ensemble-average growth rate or simply the ensemble-average exponential growth rate, $g_{\langle\rangle}$,

$$g_{\langle\rangle} = \frac{1}{\delta t} \ln \langle r \rangle = \frac{1}{\delta t} \ln \left(\lim_{N\to\infty} \langle r \rangle_N \right) = \frac{1}{\delta t} \ln \left(\sum_{n=1}^{\infty} p_n r_n \right) \;. \tag{4.104}$$

Ensemble-Average Growth Rate of the St. Petersburg Lottery

Plugging in the probabilities p_n as well as the payout x_n of the St. Petersburg lottery (equations 3.2 and 3.1), we obtain the ensemble-average exponential growth rate of the St. Petersburg lottery as

$$g_{\langle\rangle} = \frac{1}{\delta t} \ln \left(\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \frac{w_t - C + 2^{n-1}}{w_t} \right) \;. \tag{4.105}$$

Within the sum in eq. (4.105) everything behaves as before, *i.e.* the payouts diverge

317 Although the highest observed coin toss diverges towards infinity, $\lim_{N\to\infty} n_N^{\max} \to \infty$, the rate of divergence is slower compared to the divergence of the sample size, hence $\lim_{N\to\infty} \left(n_N^{\max}/N\right) \to 0$.

to infinity and are counterbalanced by the probabilities that converge towards zero. If the sum inside the logarithm is diverging again, then also the logarithm of a diverging number is diverging. Even if the logarithm is a concave function it still grows without bound – as we have seen in the derivation that led to the diverging expectation value (computed as an ensemble average) of a single round in eq. (3.8). Here it is the corresponding ensemble-average exponential growth rate for a single round of the St. Petersburg lottery given in eq. (4.105) that diverges.

The divergence of $g_{\langle\rangle}$ has led decision theories to recommend entering the gamble for any finite ticket cost $C < \infty$. We see from eq. (4.105) that irrespective of the finite initial wealth w_t and finite fee C, the terms of the infinite sum are dominated by the exponentially increasing payout in the large-sample limit $n \to \infty$.

In drawing an interim conclusion, we ask in principle for the real-world meaning of a number. What does an abstract number represent in the real world? E.g. what does the 'expected' growth rate (factor) of the expected wealth in the RHS of eq. (4.105) (RHS of eq. 4.101) represent for an individual decision-maker? We record that the pooling and sharing behind the 'expected' growth rate (factor) of the expected wealth bears no physical meaning for an individual – the pooling and sharing is an illusion. The meaning of $g_{\langle\rangle}$ is that it is the growth rate of the ensemble average of wealth. More verbosely, the growth rate of the function whose value at each time is the ensemble average of the values of all possible wealth trajectories at that time. Following the growth rate of the ensemble average as a decision criterion would recommend to accept any arbitrary high but finite ticket price to play the lottery, because there always exists an n^* such that for all $n > n^*$ the payouts $x(n)$ positively distort the ensemble average and therefore also the growth rate (factor) of the ensemble average. Thus using a quantity which is derived from the ensemble average of wealth will rationalise any finite ticket price $C < \infty$.

From eq. (4.105) we see, that in order to derive a fair ticket price the growth rate of the expected wealth would have to equal zero, $g_{\langle\rangle} = 0$, which implies that on average the gambler's wealth neither grows nor shrinks $\langle r \rangle = 1$. As has been said earlier already, the sum in eq. (4.107) diverges such that is impossible to find a finite fee which would satisfy the following equation

$$g_{\langle\rangle} = 0 \stackrel{!}{=} \ln \left(\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \frac{w_t - C + 2^{n-1}}{w_t} \right) \tag{4.106}$$

$$1 = \left(\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \frac{w_t - C + 2^{n-1}}{w_t} \right) . \tag{4.107}$$

From eq. (4.107) we see again that the fair ticket price increases in the initial wealth level.

4.5.3 Time-Average Growth Rate

After having derived the growth rate of the expected wealth in the preceding section, we now derive the time-average of that same observable with a view to the ergodicity or non-ergodicity of this observable. In a first step, we therefore construe the gamble as a repeated gamble over an arbitrary number of rounds denoted by T, whereby the wealth available to invest in the gamble in a given round is determined by the outcome of the previous round(s), and ultimately by the initial wealth. Hence, the dynamic is multiplicative and the process is irreversible. Irreversible here means that if the gambler has lost so much of his wealth at some point in time, that it becomes impossible for him to enter the gamble again in any later round (*i.e.* the remaining wealth is lower than the price of the lottery ticket), our gambler can not just go back in time and choose a different luckier realisation of himself and start afresh from a copy of himself in parallel universe. Also it is impossible for the gambler to ask an independent copy of him in a parallel 'lucky' universe for a loan. As absurd as this sounds, this is exactly what is behind the computation of ensemble-average quantities. Strictly speaking, the gambler always experiences a particular temporal sequence of realisations of the random variable *payout in the gamble*, and went bankrupt at some moment in time in the just described case.

At first we generalise eq. (4.92), because when taking a time average we are interested in what happens to the wealth after repeated gambles when an arbitrary number of rounds have been played. Let T denote the number of rounds of a gamble such that an arbitrary time interval Δt is given by $\Delta t = T\delta t$. Compared to an additive wealth dynamic in eq. (4.7), a multiplicative dynamic of the wealth of a gambler for T arbitrary rounds is given by

$$w_{t+\Delta t} = w_{t+T\delta t} := \underbrace{\overbrace{\underbrace{w_t \cdot r_{t+\delta t}}_{=\,w_{t+\delta t}} \cdot r_{t+2\delta t}}^{=\,w_{t+2\delta t}} \cdot r_{t+3\delta t}}_{=\,w_{t+3\delta t}} \cdot \cdots \cdot r_{t+T\delta t} = w_t \prod_{\tau=1}^{T} r_{t+\tau\delta t} \,, \tag{4.108}$$

whereby $r_{t+\tau\delta t}$ is the growth factor of round τ. When averaging over a factual realisation over time we always denote the time index variable by τ to clearly distinguish from averaging over all realisations in the ensemble. For a single round we set $T = 1$ and it reduces again to eq. (4.92). In the next step, we derive the average growth factor over a period of T rounds, or synonymously the mean per-round growth factor of T rounds, which is the finite-time average growth factor denoted by $\overline{r}_T$ and defined by taking the T^{th} root of the product in eq. (4.108).

This yields

$$\overline{r}_T = \left(\prod_{\tau=1}^{T} r_\tau\right)^{\frac{1}{T}} = \sqrt[T]{\prod_{\tau=1}^{T} r_\tau} \ . \tag{4.109}$$

With regard to its mathematical structure, the finite-time average growth factor $\overline{r}_T$ in eq. (4.109) is a geometric mean. Contrary to the physically meaningless ensemble-average growth factor whose mathematical structure is an arithmetic mean or a weighted arithmetic mean, see equations (4.101-4.103), respectively. We can express geometric means via fractional exponents, as in the middle term in eq. (4.109), or via higher-index roots, as in the last expression in eq. (4.109), which is the form typically encountered in the economics literature.[318] What the geometric mean of the growth factors given in eq. (4.109) yields is exactly the *equivalent constant growth factor per round* of the gamble over a period of T rounds. The following equation eq. (4.110) emphasises this interpretation of the *equivalent constant growth factor per round* in eq. (4.108) by the use of additional annotations we get the time-average wealth

$$\overline{w}_{t+\Delta t} = \overline{w}_{t+T\delta t} = w_t \prod_{\tau=1}^{T} r_{t+\tau\delta t} = w_t \overbrace{\underbrace{\left(\left(\prod_{\tau=1}^{T} r_{t+\tau\delta t}\right)^{\frac{1}{T}}\right)}_{\substack{\text{finite-time average} \\ \text{growth factor } \overline{r}_T}}^{T}}^{T\text{-times}} = w_t \cdot (\overline{r}_T)^T \ . \tag{4.110}$$

Note that the T in the subscript of $(\overline{r}_T)^T$ simply labels the *finite*-time average growth factor, whereas the T in the superscript denotes the mathematical operation of raising the finite-time average growth factor to the T^{th} power and signifies the T time periods. The superscript acts in the sense of the inverse of the Tth root. This heavy notation will not bother us for long, because we soon proceed with the infinite-time average, for which we can suppress the subscript.

The sequence of realisations of the random payouts generate a random wealth trajectory, which is a sample path of a stochastic process. In order to derive justified statements about the effects of the gamble on the wealth dynamic, we have to shift the level of analysis from any single realisation of the stochastic process to the general underlying stochastic process. The goal of the next step is therefore to apply the law of large numbers in the time domain similar to what we did before in the ensemble domain. For this purpose we introduce the frequencies of different lengths of consecutive coin tosses n. The number of coin tosses can be interpreted

[318] See for example the reference Markowitz (1959, Ch. 6), in which higher-index roots are used.

as the duration of a round δt or waiting time as hitherto. Along these lines we change the index of the product in eq. (4.109) to run over the coin tosses, n, and up to the largest sampled sequence of consecutive coin tosses, $n_T^{\max}$, in the sequence of T rounds in eq. (4.111). Because we always observe only a finite number of tosses, also $n_T^{\max}$ will be finite. Thus, we can rewrite eq. (4.109) for the finite-time average growth factor as

$$\overline{r}_T = \left(\prod_{n=1}^{n_T^{\max}} r_n^{k_n} \right)^{\frac{1}{T}} , \tag{4.111}$$

where k_n denotes again the absolute frequency of n consecutive coin tosses but now in a sequence of T rounds.[319] In a single time step the outcome is purely random, if a few time steps are considered there is still a lot of noise which we have to separate from the signal of the real effect of the gamble on wealth. Therefore we consider the infinite round limit in the next step, *i.e.* we let the number of rounds T go to infinity, then the law of large numbers has again two effects. First, the noise is separated from the signal, i.e the relative frequencies converge to the true probabilities of n consecutive coin tosses, $k_n/T \to p_n$. Second, the highest sampled length of one round measured by $n_T^{\max}$ becomes infinite, too, $\lim_{T\to\infty} n_T^{\max} \to \infty$, which yields

$$\overline{r} := \lim_{T\to\infty} \overline{r}_T = \lim_{T\to\infty} \left(\prod_{n=1}^{n_T^{\max}} r_n^{k_n} \right)^{\frac{1}{T}} = \lim_{T\to\infty} \prod_{n=1}^{n_T^{\max}} r_n^{\frac{k_n}{T}} = \prod_{n=1}^{\infty} r_n^{p_n} . \tag{4.112}$$

Finally, we arrived at the infinite-time average growth factor or simply the time-average growth factor, $\overline{r}$, that is a regular scalar number and no random variable anymore like the finite-time average $\overline{r}_T$, because its stochasticity vanishes in the probabilistic limit. As a remark, as soon as we arrived at the probabilities as powers in the calculation the higher-index roots disappear. This enables us to use the time-average growth factor or the time-average growth rate as a criterion to base our decision making on. We get the time-average exponential growth rate by taking the logarithm of the time-average growth factor,

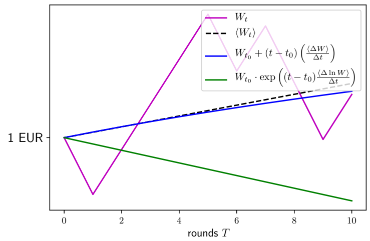
$$\overline{g} = \frac{1}{\delta t} \ln \overline{r} = \frac{1}{\delta t} \ln \left(\prod_{n=1}^{\infty} r_n^{p_n} \right) = \frac{1}{\delta t} \sum_{n=1}^{\infty} p_n \ln(r_n) . \tag{4.113}$$

Fig. 4.5 visualises the effect of the LLN in repeated gambles under multiplicative dynamics for the specific gamble of either $r_1 = 1.5$ or $r_2 = 0.6$ with $p = 1 - p = 1/2$.[320] Every subfigure is a log-linear plot and contains four curves. The dashed line indicates the evolution of the

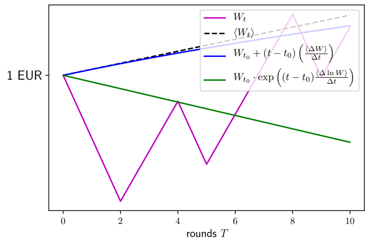
[319] In the derivation of the ensemble average k_n denotes the frequency in the fictitious ensemble of N parallel universes. *I.e.* how many independent copies of the gambler, k_n, have sequences of length n in the entirety of the N parallel worlds of the ensemble.

[320] The +50%/-40%-gamble will be the basis for the numerical simulations in Chapter 8.

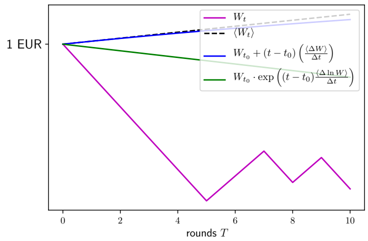
expected wealth. The blue line denotes the expected rate of change, which is the ergodic observable under additive dynamics. The random wealth trajectories are depicted in magenta. The time-average (exponential) growth rate is the ergodic observable under multiplicative dynamics and is depicted in green. Every row contains three realisations of random outcomes (coin tosses) of the same length to indicate the amount of inherent noise at the respective stage with increasing number of rounds $T \in \{10, 100, 1000, 10\,000\}$. **Fig. 4.5** indicates the decreasing probability of large deviations from the time-average growth rate (green line) at the scale of $T = 10\,000$ compared to the in part large deviations from the green line e.g. in the second ($T = 100$) and third row ($T = 1000$).



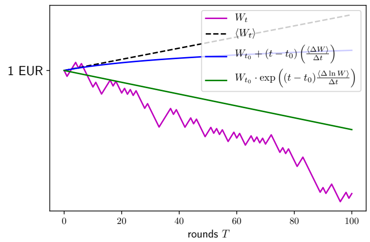
(a) $T = 10$.



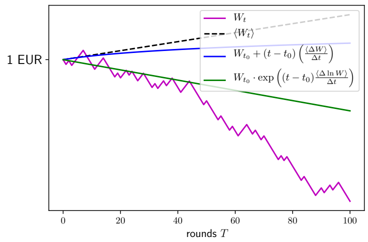
(b) $T = 10$.



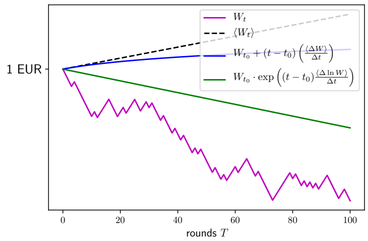
(c) $T = 10$.



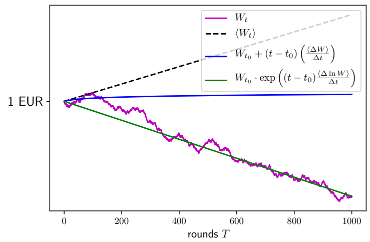
(d) $T = 100$.



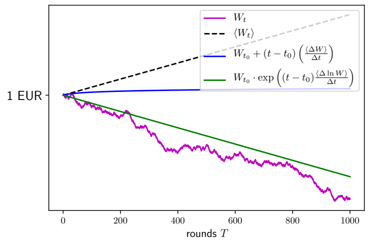
(e) $T = 100$.



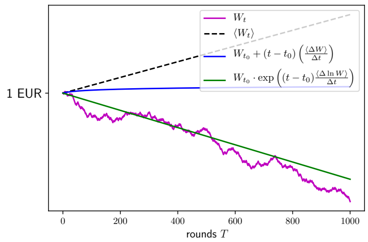
(f) $T = 100$.



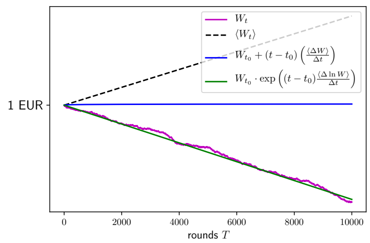
(g) $T = 1000$.



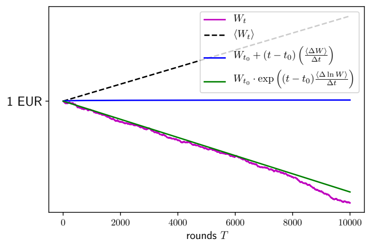
(h) $T = 1000$.



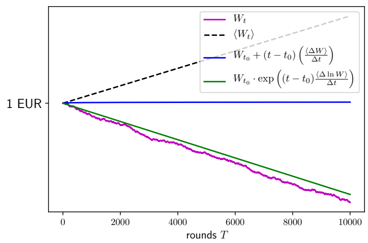
(i) $T = 1000$.



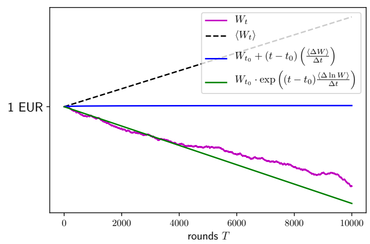
(j) $T = 10\,000$.



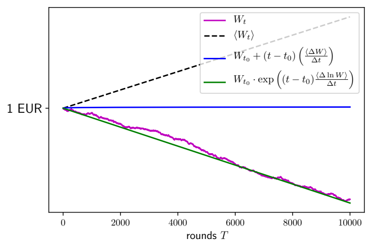
(k) $T = 10\,000$.



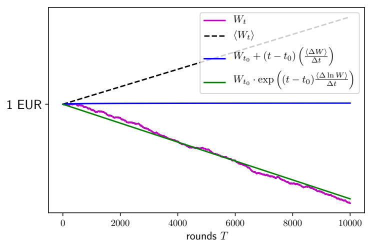
(l) $T = 10\,000$.



(m) $T = 10\,000$.



(n) $T = 10\,000$.



(o) $T = 10\,000$.

Figure 4.5: Noise vs Law of Large Numbers in +50%/-40%-Gamble With increasing number of rounds (realisations) of the +50%/-40%-gamble in eq. (8.6) and $\ell = 1$, the stochastic process in magenta increasingly aligns to the time average growth rate in green. The dashed line shows the ensemble-average growth rate. Depicted in blue is the expected rate of change, which is the ergodic observable under additive dynamics. The vertical axes are logarithmic. The fifth row depicts three additional realisations and depicts otherwise identical results like the fourth row.

Time-Average Growth Rate of the St. Petersburg Lottery

For the special structure of the St. Petersburg lottery we again plug eq. (3.1) and eq. (3.2) into eq. (4.113) and get the time-average growth rate

$$\overline{g} = \frac{1}{\delta t} \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{n} \ln\left(\frac{w_t + 2^{n-1} - C}{w_t}\right) , \tag{4.114}$$

because the probabilities (have to) sum up to one, $\sum_{n=1}^{\infty} 2^{-n} = 1$, and the denominator $\ln w_t$ does not depend on the index number n, we can restate eq. (4.114) as before by

$$\overline{g} = \frac{1}{\delta t} \left[\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{n} \ln\left(w_t + 2^{n-1} - C\right) - \ln w_t\right] . \tag{4.115}$$

Thus the time-average growth rate depends on the ticket cost C and the gambler's wealth w_t at the start of each round.

Interestingly, our time-average growth rate in eq. (4.115) exactly resembles the expected change in logarithmic utility as given by LAPLACE in eq. (4.50). The meaning behind this coincidence is explained in the next subsection.

4.5.4 Ergodicity Transformation

In this section we have derived the time-average growth rate for repeated gambles in general in eq. (4.113) and explicitly for the St. Petersburg lottery in eq. (4.115) purely from the setup of repeated gambles under multiplicative dynamics. At no time did we need to make use of any ad hoc concepts such as utility, but still eq. (4.115) coincides with the expected change in logarithmic utility in eq. (4.50). Below the box that contains eq. (4.50) we already hinted at the true nature of the logarithm. Now is the time to pick up the thread. Therefore we comment on the annotated eq. (4.116) using the proper language of *Ergodicity Economics*.

Hidden Information in Logarithmic Utility of Wealth Changes

On the LHS of eq. (4.116) we have the expected change in *logarithmic* utility augmented by the duration δt of one round of a gamble, which gives the expected change in logarithmic utility of wealth per round. Avoiding the utility terminology, this is the expected change of the logarithms of wealth denoted earlier by $\delta \ln w$. And a change in the logarithms of some variable per unit time is the growth rate of this variable, see eq. (4.100).

$$\underbrace{\frac{1}{\delta t}\overbrace{\langle \Delta u_{\mathrm{B}}(w)\rangle}^{\substack{\text{expected change}\\ \text{in log utility}}}}_{\substack{\text{expected change in}\\ \text{log utility per } \delta t}} = \underbrace{\frac{1}{\delta t}\left\langle \overbrace{\log \overbrace{\frac{w_{t+\delta t}}{w_t}}^{\text{growth factor } r}}^{\log r, \text{ is ergodic}} \right\rangle}_{\substack{\text{expected time average growth rate}\\ \hookrightarrow \text{time average growth rate}}} . \tag{4.116}$$

At the core of the RHS we have a wealth ratio, then the logarithm of this wealth ratio, which gives the logarithm of the growth factor r according to eq. (4.94). The logarithm of the growth factor per unit of time $\frac{\log r}{\delta t}$ in turn gives the exponential growth rate g according to eq. (4.99), whereby the unit of time is the duration of a single round of the gamble δt. Under multiplicative dynamics the changes in wealth Δw are not an ergodic observable, because they grow proportionally with the initial wealth. However, the logarithm of the wealth ratio is an ergodic observable, hence $\delta \ln w$ is ergodic or to be more precise and state it as it appears in eq. (4.116) $\delta \ln w = \ln w_{t+\delta t} - \ln w_t = \ln (w_{t+\delta t}/w_t)$ is ergodic. The ergodicity of the growth factors $\frac{w_{t+\delta t}}{w_t}$ arises from the multiplication and is inherited by $\ln\left(\frac{w_{t+\delta t}}{w_t}\right)$. Thus the ensemble average of the ergodic growth rate of log wealth can be expressed as a time average of the growth rate of wealth.

Now importantly, only for logarithmic utility $u_{\mathrm{B}}(w) = \log w$ the quantity on the LHS can be written equivalently as the quantity on the RHS. Thus the expected change in logarithmic utility of wealth per proper unit time happens to coincide with the time-average growth rate of wealth. However, in the terminology of utility this is a mere coincidence and reveals no deeper insight on the idiosyncratic utility function of a gambler. The following statement in 4.117 makes this even more apparent.

$$\text{non-ergodic observable of wealth} \xrightarrow[v(\cdot)]{\text{some transformation}} \text{ergodic observable of wealth}$$

$$v(w_{t+\delta t}) - v(w_t) \quad \longrightarrow \quad v\left(\frac{w_{t+\delta t}}{w_t}\right) \tag{4.117}$$

$$\hookrightarrow v(w) = \ln w \tag{4.118}$$

$$v\left(\frac{w_{t+\delta t}}{w_t}\right) = v(r) \to \ln r \tag{4.119}$$

We know that the observable *changes in wealth* $\Delta w = w_{t+\delta t} - w_t$ is not ergodic. It can be shown, however, that the rate of change or growth factors $r = \frac{w_{t+\delta t}}{w_t}$ are ergodic. Does a transformation v exist such that maps the difference in eq. (4.117) into a ratio? Well, the logarithm is the sought for transformation does exactly this, converting a difference in a ratio and for this reason Peters and Gell-Mann (2016) and more explicitly so Peters and Adamou (2018c) coined this an ergodicity transformation and thereby identified the true role of the logarithm in decision making under uncertainty. The logarithm understood as an ergodicity transformation that converts a non-ergodic observable (*changes in wealth* Δw) in an ergodic observable (*changes of the logarithms of wealth* $\delta \ln w$). Therefore we can use the expected change of this transformation per unit time $\frac{v(w)}{\delta t}$ as a decision criterion such that

$$\frac{\delta v\left(w\right)}{\delta t} = \frac{v\left(r\right)}{\delta t} = \frac{\delta \ \ln w}{\delta t} \tag{4.120}$$

$$\bar{g} = \begin{cases} < 0 & \rightarrow \quad \lim\limits_{T\to\infty} w \rightarrow 0 \\ > 0 & \rightarrow \quad \lim\limits_{T\to\infty} w \rightarrow +\infty \end{cases}, \tag{4.121}$$

i.e. if $\frac{\delta v(r)}{\delta t}$ is positive, then wealth w grows in the long run and if $\frac{\delta v(r)}{\delta t}$ is negative, wealth w changes negatively, *i.e.* shrinks in the long run. Therefore the profitability of a gamble is determined by the time average growth rate of wealth, which is identical to say the change of the ergodic growth factor of wealth per unit time.

Thus an ergodicity transformation justifies the substitution of time averages by the ensemble average, which gives an answer to our main topic the ergodicity problem and relates to why the quantity derived from EUT looks similar by sheer coincidence only. The quantities of a growth factor and a growth rate only make sense once a dynamic is specified which embeds the randomness of the gamble within time. Time, dynamics, and growth rates are instrumental ingredients in *Ergodicity Economics*, but are absent in the evaluation of gambles using utility theory. In other words the logarithm is in fact not a functional form of utility, which is a flawed concept after all. To repeat the labelling used in Subsec. 4.3.5, this is the reason why a framing of the ergodicity transformation in the terminology of utility is an abuse of concepts. In addition, this mediates between the parlance of utility theory and the terminology of *Ergodicity Economics* in the following way. What is referred to as different utility functions actually encodes different non-ergodic dynamics and thus is freed from idiosyncratic characteristics of the particular gambler.

Furthermore, Peters and Adamou (2018c) developed a general method to derive what would be the specific utility function given an arbitrary dynamics using Itô's calculus. For example imagine a dynamic that is neither purely additive nor purely multiplicative but something in between, for which e.g. the square root (utility) function is the sought-for ergodicity transformation. This step also demonstrates the relation to stochastic differential equations

(SDEs) such as geometric Brownian motion (GBM) as a model of stochastic growth processes such as wealth accumulation. GBM is the standard model in the field of mathematical finance. We can not explain all the important considerations involved in the fascinating derivation of the ergodicity transformation, which involves the monotonicity of v in w in combination with a proper dimensional analysis but refer to the special literature.[321]

From the discovery of the role of ergodicity transformations PETERS and ADAMOU (2018c) deduce the hypothesis that decision makers behave rational, when rationality means the optimisation of their time-average growth rate of wealth. From this hypothesis emerges the new cosmos of *Ergodicity Economics* with a new rationality concept. A decision maker then seeks to optimise the time-average growth rate of his wealth under the given dynamics. From these insights arises a whole array of new fascinating research question.

4.5.5 Non-Ergodicity of Multiplicative Dynamics

What is left in order to determine the ergodicity of our observable wealth under a multiplicative dynamic is that we have to compare the growth rate of the expected wealth $g_{\langle\rangle}$ from eq. (4.105) with the time-average growth rate $\overline{g}$ in eq. (4.114) – similar to what we have done in Sec. 2.1.2 in eq. (2.5)

Non-Ergodic Growth Rate of Wealth for Multiplicative Dynamics

From eq. (4.122) we see that the growth rate of wealth under multiplicative dynamics is in general non-ergodic or synonymously the ergodicity is broken under multiplicative dynamics. But more generally, we see from eq. (4.123) that the reason the ergodicity gets broken is the non-commutativity of the operations of the logarithm and summation or between taking the logarithm of a sum and summing over logarithms. Thus the non-ergodicity of the growth rate of wealth is a consequence of JENSEN's inequality.

$$g_{\langle\rangle} \neq \overline{g} \, , \tag{4.122}$$

$$\sum_{n=1}^{\infty} p_n \ln\left(r_n\right) \neq \ln\left(\sum_{n=1}^{\infty} p_n r_n\right) \, . \tag{4.123}$$

If we plug in the specific probabilities and payouts for the St. Petersburg lottery again we can establish the broken ergodicity in this special case of a gamble under multiplicative dynamics,

[321] See especially PETERS and ADAMOU (2018b, pp.18) PETERS and GELL-MANN (2016) and PETERS and ADAMOU (2018c) who additionally recommend BARENBLATT (2003) on the treatment of dimensional analysis. The discussion nicely relates to our **Remark on the Correct Observable** on p. 119.

$$\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \ln\left(\frac{w_t + 2^{n-1} - C}{w_t}\right) \neq \ln\left(\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \frac{w_t + 2^{n-1} - C}{w_t}\right) . \tag{4.124}$$

In the context of incorrect use of statistics in many areas, Bauer et al. (2015, pp. 21–22) alert about the common error in economics to compute arithmetic means of growth rates, where the geometric mean is the only sensible mathematical operation. Although the authors are unaware of the ergodicity problem, they understand that there is a problem. This is a typical manifestation of the ergodicity problem as the nonchalant but unjustified substitution of time averages by ensemble averages without scrutinising the dynamic not to speak of the ergodicity of the observable. Computing arithmetic means of growth rates misses the essential part of the multiplicative dynamics of the phenomena and produces misleading ensemble averages of temporal processes.

4.5.6 Conclusion

In this section we discussed the procedure of averaging over an ensemble and averaging over time, which lead to fundamentally different averages studied in ergodic theory. By averaging over N possible realisations of a random variable and letting the number of realisation tend towards infinity, we sample the timeless state space of possible realisations with increasing precision. By averaging over one factual realisation of a stochastic process over T rounds and letting the number of rounds tend towards infinity we sample a representative factual realisation of the stochastic process with increasing precision. Thus the difference amounts to an average over something possible and an average over something factual. It is the latter which has importance for individual decision makers.

Meanwhile we are well prepared to absorb the related literature on the ergodicity problem which we review in the ensuing chapter 5. We start with a discussion of further results from within the core of the *Ergodicity Economics* research programme and continue to review the literature related to the ergodicity problem in economics and beyond. In the chapters 7 and 8 we pick up the thread of *Ergodicity Economics* again and continue with an exploration of the potential of this new approach.

4.6 Concluding Remarks on the Solution Strategies

The sheer mass of the literature on the St. Petersburg lottery forced us in the presentation of solution strategies to make a choice. This choice is guided by the ordering criteria of which mathematical operation was implemented by the chosen strategies. In presenting the solution strategies to the St. Petersburg lottery this way, we covered the main steps in the history of the problem that are relevant for the purpose of this thesis to arrive at *Ergodicity Economics*. Needless to say, we can not cover the material exhaustively here, but we want to at least mention an array of interesting contributions that were more or less on the right track.

Euler and Geometric Means

For example Euler (1862) derives a geometric mean as a proper quantity to evaluate a gamble and directly emphasised to the difference between his rule and traditional mathematical expectation:

> "Following the usual method to this point one ought to combine together all statuses [scenarios/states of the world], which are able to happen in each of the cases, into one sum and divide this (sum) by the number of cases. Thus the difference between these two methods consists in this, in so far as ours uses multiplication, when the other addition, likewise the exponent, when this uses multiplication itself; if we determine the operations geometrically, those men in fact (determine them) arithmetically, thus as these men adjust those operations on the statuses themselves, we would transfer the same (operations) onto the logarithms of the statuses, and the same, which it brings forth, the number corresponding to the logarithm of this, indicates to us the sought status of him playing."[322]

Whitworth on Dynamics, Average Multipliers, the Gambler's Ruin Problem

William Allen Whitworth was a mathematician, who published several recognised textbooks at his time such as *Choice and Chance*[323] and later became a Victorian priest.[324] The early editions of Whitworth (1870) contain an appendix *On the Disadvantage of Gambling*, which later became an independent chapter e.g. in Whitworth (1886, Ch. IX, 1901, Ch. XI). His books contain lucid explanations and derivations very similar to those in Sec. 4.5 without using the terminology of time and ensemble averages, of course, but the mathematics is correct and very close to ours. He clearly had the time perspective or long run in mind. For instance, he noticed the necessity to consider the mode of repetition – that is the dynamic – to be able to evaluate gambles properly: 'Whether a single fair bet, regarded by itself and without consideration of its being repeated, can be said to be disadvantageous is

[322] Euler 1862, p. 3.
[323] Whitworth 1870.
[324] Irwin 1967.

a question open to argument.'[325] And further '[b]ut in truth the advantage or disadvantage of a venture must be judged by its effect in the long run when the operation is indefinitely repeated. If it is advantageous to perform the operation once it must be advantageous to perform it again. Its tendency is seen in the ultimate result.'[326]

Furthermore, he explained the difference between repeated gambles under additive dynamics in which sequential decision making is not *con*-sequential and multiplicative dynamics in which the decisions are also consequential in the literal meaning of the word – *with consequences.*

> "The absolute value of a mathematical expectation is not the prize which a man of limited means ought to pay for the prospect. It expresses the value of the expectation to a man who is able to repeat the venture indefinitely without the risk of his operations being ever terminated by lack of means."[327]

The range of applicability of the expectation value renders Whitworth in the following successful way: 'The speculator's fund, to begin with, must be infinite in comparison with the stakes involved, before he may venture to give the absolute value of the mathematical expectation for any contingent prospect which he may desire to purchase.'[328] The time-average growth rate is aptly coined the average multiplier, 'Hence the average multiplier is less than unity, or the general tendency of the man's operation is to decrease his funds.'[329]

Whitworth considers explicitly the gambler's ruin problem in repeated gambles,

> "This is the basis of the disadvantage of gambling there is (as it were) a pendulum swinging between gain and loss with oscillations of varying amplitude : the pendulum oscillates fairly enough between gain and loss, but when once it reaches a certain degree of loss it is held there and swings no more. This stop, existing only on the side of loss with no compensating stop upon the side of gain, is the 'disadvantage' of gambling."[330]

The gambler's ruin problem reframed in the terminology of *Ergodicity Economics* is strong ergodicity breaking due to the absorbing barrier of bankruptcy at e.g. (without debt) zero wealth. Hence the objective of the gambler is to prevent bankruptcy while maximising his time-average growth rate, otherwise he can no longer play the repeated gambles. Finding a so called optimal bet size to wager at repeated gambles is the main goal behind the Kelly criterion, to which we turn to in chapter 8. Finally, we get the impression that some of the

325 Whitworth 1901, p. 223.
326 Whitworth 1901, p. 224.
327 Whitworth 1901, p. 244.
328 Whitworth 1901, p. 244.
329 Whitworth 1901, p. 243.
330 Whitworth 1901, p. 225.

necessary ingredients and already some of the solutions of *Ergodicity Economics* were already known to WHITWORTH. In his conclusion he compares his average multiplier with the solution strategies that appeared in the history of the St. Petersburg lottery. It is worth quoting here in its entirety:

> "The result at which we have arrived is not to be classed with the arbitrary methods which have been again and again propounded to evade the difficulty of the Petersburg problem and other problems of a similar character. Formulae have often been proposed, which have possessed the one virtue of presenting a finite result in the case of this famous problem, but they have often had no intelligible basis to rest upon, or, if they have been established on sound principles, sufficient care has not been taken to draw a distinguishing line between the significance of the result obtained, and the different result arrived at when the mathematical expectation is calculated.
> We have not assigned any new value to the mathematical expectation ; we have not substituted a new expression for the old ; but we have deduced a separate result, which without disturbing the mathematical expectation has a definite meaning of its own. We have found not the fair price at which a contingent prospect may be transferred from one man to another, but the value which such a prospect has to a man in given circumstances."[331]

Cover and Information Theory

Another closely related and present-day line of thought of involves solutions of the St. Petersburg lottery which originated from information theory.[332] COVER (2011) draws the conclusion of the misleading role of the expectation value of wealth, too, and instead argues in accordance to the *Ergodicity Economics* approach for the geometric mean as an appropriate decision criterion. However, the interpretation differs in so far, that a gambler is always in favour of the some of St. Petersburg lottery but wants less of it with increasing fee. It is equivalent to say, the gambler always wants to invest a fraction of his wealth in fractions of the St. Petersburg lottery. Thus in general the St. Petersburg lottery is found to be attractive for all fees C. Interestingly, COVER derives the harmonic mean as the critical fee, below which the gambler wants all of it.

One lesson to be learned from the preceding solution attempts seems to be that it is impossible to solve the St. Petersburg paradox without further qualifications. This suggests that the St. Petersburg problem as a mathematical problem is not well-posed. However, there emerged time and again new approaches to deal with this seeming ill-posedness that are interesting on

331 WHITWORTH 1901, p. 246.
332 BELL and COVER 1980; COVER and THOMAS 2006, e.g. as a textbook exercise ch. 6 Ex. 6.17.

their own, but none of them are convincing final resolutions. Whereas the first four solution strategies rely on ad hoc assumptions, the fifth strategy relies only on the assumption of a reasonable mode of repetition and in doing so embedded the randomness within historical time. Thus, the time perspective that *Ergodicity Economics* provides, competes as the next candidate of a final resolution in the history of this intriguing problem.

5 Literature Review on the Ergodicity Problem in Economics and Beyond

> [T]he pricing model is stationary ergodic whereas the market, generating radically emergent events, is not stationary ergodic. [...] I would contend the issue is more fundamental: that the markets, in reality, cannot be associated with any stationary ergodic processes, on which he bases his subsequent arguments.[333]
>
> TIMOTHY C. JOHNSON

This thesis is a contribution to the research programme of *Ergodicity Economics*, which is a very recently developed way to conceptualise risk and inform decision-making under uncertainty. The following literature review is organised in three parts, which serve the following three purposes. First, there is the purpose of the traditionally very narrow focus of a literature review, which is to refer to work that stands in close relation to our topic. This is fulfilled in Sec. 5.1, which gives an overview of previous results of the *Ergodicity Economics* approach. Note that we additionally direct to specific references throughout the thesis when we discuss special topics in the several chapters and sections. The aim of the second part of the literature review is to present elective affinities to related work from within economics in Sec. 5.2. In general we have to note that economic publications only rarely use the terms 'ergodic' or 'non-ergodic'. In fact, this constitutes a first finding: both of these two terms are used inversely proportional to their importance in the economics community. A second finding is that 'non-ergodic' or 'broken ergodicity' is mentioned even less often than 'ergodic', which corroborates the conclusion that the ergodic problem is strongly disregarded and deserves greater attention. The aim of the third part is to present elective affinities to research fields outside of economics in Sec. 5.3. We present work related to the ergodicity problem beyond economics, especially in physics (5.3.1), information theory (5.3.2) and biology (5.3.3). The topics we present in the second and third part are selected under the guiding theme of their potential for synergies between the fields and related economic questions. *E.g.*

333 JOHNSON 2016, p. 197.

sometimes the terminology may be different, but we seek to make transparent if the nature of the studied phenomena are nevertheless related. After having read the preceding chapters that in mathematics and the sciences both terms – 'ergodic' or 'non-ergodic' – are much more common comes with no surprise. Therefore, Sec. 5.3 is on the one hand intended as an invitation to dive further into the presented topics with the aid of the cited literature and on the other hand evidence how artificial the boundaries between academic disciplines are.

5.1 Recent Work on Ergodicity Economics

The *Ergodicity Economics* research programme continues the work started in the seminal papers Peters (2011c,b). Much of it is covered in detail in Sec. 4.3, Sec. 4.5, and chapter 8. Here we briefly refer to results which are covered to a lesser extent in this thesis. Peters (2011c) identified the ergodicity problem in a broader scope and particular its consequences for economics and finance for the first time and developed the so-called time resolution. The time resolution accounts for the different effects of randomness if conceptually embedded within historical time in contrast to the traditional embedding within a virtually constructed ensemble. The ergodicity problem raises the need for the awareness of the conceptual difference and careful choice between time averages and averages over stochastic ensembles as this choice can cause the sure ruin of an economic agent. Thereby, Peters used the time-average growth rate of the gambler's wealth dynamic generated by repeated participations in St. Petersburg lotteries to resolve what has been known since 1713 as the St. Petersburg paradox.[334] Peters (2011b) picked up the insights from the time resolution to answer questions in finance regarding the optimal leverage of individual investors and showed the dominance of growth-optimal leverage over the standard mean-variance framework and risk measures associated with it like the Sharpe ratio. Peters (2011a) identified a mathematical error regarding different speeds of convergence towards infinity in Menger (1934)[335], which invalidate the conclusion that utility functions are needed at all let alone bounded utility functions to evaluate gambles like the St. Petersburg lottery.

In Peters and Gell-Mann (2016) the solution to the gamble-selection problem is given, which lays at the foundation of economics, how to choose optimally between two uncertain outcomes. This paves the way for a new decision theory, the growth-optimal decision theory, and is there to replace EUT and modern cousins of it like PT or CPT.[336] The results of Peters and Gell-Mann (2016) have been generalised recently in Peters and Adamou

[334] For a detailed presentation of the time resolution see Sec. 4.5.

[335] The error in Menger (1934) is a significant one, which interestingly can be traced back to an easily overlooked inaccuracy in the calculation of the expectation value in Bernoulli (1738). For a detailed analysis also see the blog post Peters (2018c). See also the discussion in Sec. 4.3.7.

[336] A deep discussion of the evolution of decision theories ensues in Sec. 7.4.

(2018c) and involve a detailed discussion on the relation to EUT and how arbitrary functional forms of utility map on to specific wealth dynamics. We discuss ergodicity transformations in Subsec. 4.5.4 in detail.

All these and other results have been further developed into the coherent framework of *Ergodicity Economics*, which is available in the form of growing lecture notes Peters and Adamou (2018b). *Ergodicity Economics* allows the resolution from a coherent angle of many famous puzzles in economics and finance like the insurance puzzle,[337] wealth inequality[338] and cooperation.[339] Specifically, in Peters and Adamou (2013, 2017b) a new concept of efficiency labelled *stochastic market efficiency* or synonymously *leverage efficiency* is proposed, which successfully tests the hypothesis that the optimal leverage is attracted to 1. In Peters and Adamou (2017b) the equity premium puzzle[340] is resolved, an explanation of leverage-driven bubbles is offered and a theory of noise in financial markets is derived. In addition, policy decisions on how to set interest rate rationally are deduced.[341]

Before we continue with the review of work related to the ergodicity problem in economics, we briefly review a debate on geometric means, which is very close to *Ergodicity Economics*. In fact it is so close that it could well have anticipated a lot of its results.

5.1.1 Geometric Mean Debate

The time average of different observables can take different appearances. *E.g.* for growth factors the time average has the mathematical structure of a geometric mean. Geometric means have been proposed in the literature before also as decision criteria, but only on rare occasions. Among them are publications by the famous mathematician Leonhard Euler[342] and the mathematician turned Victorian priest William Allen Whitworth,[343] which are discussed in Sec. 4.6. Around the specific appearance of time averages as geometric means arose a debate in economics and finance.

The debate on the superiority of the geometric mean is a prime example for a wasted chance in the history of economic thought, because the door opened by Peters (2011c) could have been walked through already in the late 1950s or 1960, if there had not been such prominent critics like Paul Samuelson. Unfortunately, the rhetoric of authority (*e.g.* in Samuelson (1979a), in the sense of McCloskey (1983)) was too imperative. The problem was that

337 Peters and Adamou 2017a.
338 Berman et al. 2016, 2017.
339 Peters and Adamou 2018a.
340 Mehra and Prescott 1985.
341 For a description of the *Ergodicity Economics* programme see also http://lml.org.uk.gridhosted.co.uk/research/economics/.
342 Euler 1862.
343 Whitworth 1870.

neither the critics nor the proponents saw the revolutionary potential of the time average behind the mere appearance of the geometric mean. Thus a collection of synonyms emerges in different subdisciplines of economics and finance, for instance the necessity to maximise $\mathrm{E}[\log \text{wealth}]$. With the exception of BREIMAN (1961), who is aware that $\max \mathrm{E}[\log \Delta w]$ is the maximisation of a growth rate of an exponentially growing wealth w, the true nature behind the geometric mean was missed.

There exists a long list with prominent economists as its members[344] that are at least aware of the geometric mean approach in *e.g.* KELLY (1956), LATANÉ (1959) and MARKOWITZ (1959), but fail to see the time average behind it and the theoretical advancement of embedding randomness in time. Clearly for SAMUELSON, we have to state that he rejected the geometric mean approach on the false ground, that the geometric mean approach yields an expression that is actually a time average, that by mere chance looks like a familiar quantity called expected logarithmic utility. By chance we mean if luckily the logarithmic utility function is chosen. In the papers what is actually a time average (exponential growth rate) of a multiplicative growth process in a way is then interpreted in a way that is conform with and thus corroborating the existing framework of embedding randomness in the timeless ensemble using utility theory.[345]

5.2 Related Work to the Ergodicity Problem in Economics

After we reviewed the most recent advancements in *Ergodicity Economics*, we present work related to the ergodicity problem in further branches of economics. We start to review the coverage of ergodicity in dictionaries, encyclopedias and handbooks of economics. We continue the review with discussions on related work to the ergodicity problem in finance, general equilibrium economics, and econometrics.

5.2.1 Ergodicity in Dictionaries, Encyclopedias and Handbooks

Articles in handbooks and encyclopedias or entries in dictionaries are natural starting points to get acquainted with a new topic. Therefore, we briefly review the treatment of 'ergodicity' in *The New Palgrave Dictionary of Economics*. In the ensuing subsection we discuss the treatment of (non-)ergodicity in three handbooks. The presentations are meant to be exemplary for the general treatment of the ergodicity problem in economics so far.

344 SAMUELSON 1969, 1971b, 1979b; MERTON and SAMUELSON 1974; SINN 1989; SINN and WEICHENRIEDER 1993; SINN 2003.

345 See for instance MARKOWITZ (1959, Part II, Section VI Return in the Long Run, Part IV, Section XI).

5.2.1.1 Ergodicity as a Well-Behaved Equilibrium

The New Palgrave Dictionary of Economics contains an entry on 'Ergodicity and Non-Ergodicity in Economics' by HORST (2008). As a general result, the Palgrave entry by HORST (2008) and the references therein are following a typical understanding of ergodicity as a well-behaved economic equilibrium.[346] Ergodicity is a particular strong requirement to such an equilibrium, which is often not made explicit but referred to with the help of following (partly overlapping) properties:

- the equilibrium distribution exists,
- the equilibrium distribution is a limit distribution,
- the limit distribution is invariant or stationary, and
- the invariant limit distribution is unique.

Hence, an ergodic equilibrium in dynamic models of the economy is a unique invariant limit distribution. A violation of the ergodicity (also called broken ergodicity or breakdown of ergodicity) is associated with the occurrence of path-dependence, but not mentioned in great detail, but we acknowledge the tight constraints a rather short dictionary entry is subject to. The Palgrave entry focuses on the following topics given in the original order of appearance:

1. path dependence,[347]
2. endogenous preference formation[348] and stochastic strategy revision in strategic interaction,[349]
3. non-ergodic economic growth,[350]
4. models of social (mean-field) interaction,[351]
5. models of social (local & global) interaction.[352]

However, the exposition is not close to connect ergodicity to the ergodicity problem and its relation to the understanding of risk and decision making under uncertainty, which is our focus and was first given in PETERS (2011c). Nevertheless, the way HORST (2008) presented ergodicity in this dictionary entry is very representative for the general treatment of ergodicity in economics so far. In the remainder of this chapter, we will refer to this interpretation of ergodicity as the existence of some unique invariant limit distribution or well-behaved equilibrium. Therefore, we confine ourselves to discuss only the third item of the list in a bit more detail.

346 See for example KORINEK (2015).
347 DAVID 1985.
348 HILDENBRAND 1971; FÖLLMER 1974.
349 FÖLLMER 1974; BLUME 1993.
350 BROCK and MIRMAN 1972; BLUME 1993; DURLAUF 1993.
351 BLUME 1993; DURLAUF 1993, 2001; BROCK and DURLAUF 2001.
352 BISIN et al. 2006; HORST and SCHEINKMAN 2006, 2009.

Ergodicity and Macroeconomic Growth

In the paper by Durlauf (1993) titled 'Nonergodic Economic Growth' multiple steady states are studied in a model of technological innovation in different industries. Every industry's output and investment decisions are modelled via a stochastic process conditional on past production technology choices. Multi-dimensional random processes or random fields then give rise to non-ergodic growth, in the sense of path-dependent growth. The non-ergodicity of growth manifests itself in the existence of multiple equilibria. *E.g.* in the simplest case of two equilibria, a high- and a low-output equilibrium, coordination on the low-output equilibrium is seen as a coordination failure on an inferior equilibrium distribution of investments. The actions of each microeconomic agent are modelled via a probability measure ($\mathcal{P}$-measure) and collectively form a set of $\mathcal{P}$-measures. An aggregate equilibrium interpreted 'as a joint probability measure characterizing many agents'[353] conditional on the history of the economy exists, if the joint $\mathcal{P}$-measure over all agents is compatible with the set of individual $\mathcal{P}$-measures. Stated differently, if the process is ergodic an equilibrium exists, i. e. a unique $\mathcal{P}$-measure compatible with all conditional $\mathcal{P}$-measures at any given point in time. This proves to be useful to 'equate the invariant probability measure over all agents [ensemble measure] with the long-run equilibrium of the economy [time measure]',[354] which manifests the elimination of time and use of ergodicity. The connection between the necessity of the process to be ergodic for the existence of a unique equilibrium is unfortunately not much elaborated on. But nevertheless, the dictionary entry selected the paper to relate non-ergodicity to economic growth.

In the paragraph on non-ergodic economic growth, also optimal macroeconomic growth models are discussed. In Sec. 4.5 we commented already briefly on growth-optimality. Growth-optimality is the main focus of the chapters 7 and 8, but the two understandings of growth-optimality differ strongly. In *Ergodicity Economics* the growth-optimality arises from the maximisation of the time-average growth rate of wealth, whereas the mentioned macroeconomic growth models strive for an optimal consumption pattern in one-sector toy economies.[355]

---

353 Durlauf 1993, p. 350.

354 Durlauf 1993, p. 350, our additions in square brackets.

355 Models of optimal macroeconomic growth can be traced back to Ramsey (1928), von Neumann (1937), Cass (1965) and Koopmans (1965). These models can be understood in the following sense: 'The cornerstone of one-sector optimal economic growth models is the existence and stability of a steady-state solution for optimal consumption policies. The optimal consumption policy is the stable branch of the saddle point solution of the system of differential equations governing the dynamics of the economy' (Brock and Mirman 1972, p. 479). A frequent identifier in this line of research is the 'golden rule of XYZ' as a synonym for optimal in the Ramsey sense, *e.g.* the optimal rate of savings is also called the 'golden rule savings rate', other instantiations of the golden rule concept are the golden rule path, golden rule of accumulation, golden rule of research, golden rule of education and even the golden rule of procreation (Phelps 1967, p. ix). Especially Edmund Phelps produced a whole anthology of golden rules (Phelps 1961, 1965, 1966).

Interesting extensions of strategic interaction models to non-ergodic behaviour are not covered in the Palgrave entry, but begin with work published later among others by the same author and co-authors in Horst and Wenzelburger (2007), which we discuss in Subsec. 5.2.2.

5.2.1.2 Encyclopedia of Complex Systems in Finance and Econometrics

The *Encyclopedia of Complexity and Systems Science*[356] lists ergodic theory among the mathematical and modeling tools to analyse complex systems and grants a whole chapterto ergodic theory. Spread over the whole encyclopedia, there are in total 20 articles on aspects of ergodic theory, from very general topics like

- 'Ergodic Theory: Basic Examples and Constructions'[357] or
- 'Ergodic Theorems'[358]

to very sophisticated and abstract mathematical contributions on

- 'Smooth Ergodic Theory'[359] or
- 'Ergodic Theory: Interactions with Combinatorics and Number Theory'.[360]

To the ergodicist's great surprise, the encyclopedia of *Complex Systems in Finance and Econometrics*[361] – which is a proper subset of the *Encyclopedia of Complexity and Systems Science* – contains not a single article on a topic related to ergodic theory, and only 12 rather incidental mentions of 'ergodic', and three mentions of 'non-ergodic' throughout the whole book – quite astonishing for a handbook of almost 900 pages and such an important topic.

5.2.1.3 Ergodicity Problem in Handbooks of Economics

Handbook of Financial Markets

The *Handbook of Financial Markets – Dynamics and Evolution*[362] is a splendid overview on cutting-edge research in finance from an evolutionary perspective. Especially the contributions by Evstigneev et al. (2009) and Wenzelburger (2009) are interesting for our focus. Wenzelburger (2009) develops the results gained in Horst and Wenzelburger (2007)

356 Meyers 2009.
357 Nicol and Petersen 2009.
358 Del Junco 2009.
359 Wilkinson 2009.
360 Ward 2009.
361 Meyers 2011.
362 Hens and Schenk-Hoppé 2009.

on non-ergodic asset price behaviour further. For the moment, we postpone the discussion of WENZELBURGER (2009) to Subsec. 5.2.2.2, where we embed it in the larger context of evolutionary finance. The exposition of the KELLY[363] criterion in EVSTIGNEEV et al. (2009) as a cornerstone of evolutionary finance is unmatched in the literature. Also the authors present the concept of the growth-optimal KELLY criterion in a language that is very accessible to economists. Furthermore, the relevance the authors attribute to the KELLY criterion goes beyond to what has been done in previous work especially by the same group of authors.[364] We review both the two specific articles and evolutionary finance in general in greater detail in Subsec. 5.2.2.2.

Two other contributions to the handbook, BLUME and EASLEY (2009) and KURZ (2009), relate to ergodicity, but still restrict themselves to the analysis of ergodic dynamics of markets. Whereby BLUME and EASLEY (2009) follow along the lines of their previous collaboration on strategic interaction,[365] to which we return in the context of mathematical finance in Subsec. 5.2.2.1 in a little while. In the other article the following observation is made:

> "Repeating the experiment [of forecasting GDP growth and change in GDP deflators] over time, we find that the distribution of forecasts fluctuates in two ways. First, it exhibits changes in the cross-sectional variance of the forecasts, reflecting changes in degree of disagreement. Second, it exhibits large fluctuations over time in relation to the stationary forecasts reflecting correlation in forecasters' views about unusual conditions at the time. Sometimes forecast distributions are below the stationary forecast, whereas in May 2000 the distribution was above the stationary forecast. [...]
> We thus note that for any variable, individual forecasts are correlated and the average market forecast fluctuates around the stationary forecast. This nonjudgmental forecast is a central yardstick in the work reviewed later. Also, observe that the cross-sectional variance of forecasts fluctuates over time."[366]

However, KURZ does not go so far as to simply conclude the non-ergodicity of the observable growth rate of GDP and the GDP deflator, but clings to the idea of the existence of a stationary forecast. In addition, we find the following often made – but in our opinion dubious – claim: 'We also note that in the nonergodic case the data requirement is greater, since then we need data for many alternative sequences [of the observable ...] with different starting points, but the basic theory remains unchanged.'[367] First, we doubt the claim that all results hold

363 KELLY 1956.
364 For work by this group of authors with less emphasis on the KELLY approach see the references in Subsec. 5.2.2.2.
365 BLUME and EASLEY 1992, 2002.
366 KURZ 2009, p. 441.
367 KURZ 2009, p. 459 footnote 3.

unchanged also in the non-ergodic case. Second, if economic processes happen in irreversible historical time and there is only exactly one realisation of a process it remains unclear from where to take additional realisations. The ensemble stays a mental construct that is inaccessible to us. A similar statement taken from the lecture notes of a statistics course: „Im Fall eines nichtergodischen Prozesses sind viele Prozessrealisationen erforderlich.“[368] Although the principal idea – to long for additional realisations of the process to get a better overall picture – is understandable, but still it remains impossible to get additional realisations of a singular trajectory. We return to this issue in the context of (time series) econometrics in Subsec. 5.2.5. Clearly, the approach of time-average quantities offers a road ahead without having to sacrifice mathematical tractability.

Handbook on the Kelly Capital Growth Investment Criterion

The collection of reprints in MacLean et al. (2011g) is devoted to the Kelly criterion in the theory of optimal growth. Besides the seminal papers of the field, MacLean et al. (2011g) brings together thoughtfully elected contributions with the clear focus on asset allocation from a variety of different research communities that resemble the topics in this literature review, *e.g.*

- related work from operations research and mathematics, which we review in Sec. 2.6,
- related work from economics and finance, which we review in Sec. 5.2,
- related work from evolutionary finance, which we review in Subsec. 5.2.2.2,
- related work from physics, which we review in Subsec. 5.3.1, and
- related work from information theory and electrical engineering, which we review in Subsec. 5.3.2.

Every part is enriched by very readable introduction, which also serve as excellent literature reviews.[369] The key difference to the *Ergodicity Economics* approach is that the editors and papers cling to the utility concept. We will see the same pattern in evolutionary finance in Subsec. 5.2.2.2 in a moment. Recently, the nature of the transformation applied to the payouts called utility has been revealed as a mapping, which people apply with respect to the dynamics of their wealth[370] This is in stark contrast to Kelly (1956, p. 918), who argued rather explicitly against the utility concept and saw no need to invoke it. We elaborate on the Kelly approach and its relation to utility in detail in chapter 8.

We could continue to list further collections or articles in handbooks[371] – especially by

[368] Huschens 2017, p. 210.
[369] MacLean et al. 2011b,a,e,d,f,c.
[370] Peters and Adamou 2018c.
[371] Rachev 2003; Durlauf et al. 2005; Cont 2010; Ait-Sahalia and Hansen 2010.

proponents of the Kelly optimal-growth community[372] – that are more or less loosely connected to our specific focus. The examples we gave should be enough to convey the message that the identification of the ergodicity problem in economics – which includes the embedding of risk within historical irreversible time, the focus on time-average quantities together with a rejection of both expectation values and the concept of utility – is a genuine contribution of the ergodicity economics programme and explored in this thesis. It can not be found elsewhere in the literature in this explicit form.

5.2.2 Finance

The discussion in this subsection picks up the thread where the Palgrave discussion come to an end. Although it is not always possible to keep the two branches of evolutionary finance and mathematical finance strictly apart, the explicit incorporation of Darwinian ideas in the former serves as a dividing line for the moment. We present the different approaches in the way they tackle topics close to the ergodicity problem in economics. Nevertheless, they share the common theme to go beyond the rational expectations hypothesis (REH) paradigm[373] and increasingly study the effects of strategic interaction among heterogeneous agents. The study of the interaction of heterogeneous agents in (financial) markets was precluded by the REH, because in dynamic general equilibrium models of the economy the REH paradigm reduced the agents' heterogeneities to a single infinitely lived representative agent in an economy without the explicit need to model a financial market. Dissatisfied with the unrealistically strong assumptions on agents' information processing capabilities associated with the REH, several authors started to develop models with heterogeneous agents, in which the ability to correctly anticipate the distribution of future asset prices is the exception rather than the general case.[374] Thereby the agents are allowed to differ *e.g.* with regard to their initial endowment, preferences and beliefs. Further, simply assuming rational expectations does not illuminate the process of how agents in the economy form their expectations, which is now widely accepted to be important.

Therefore, two types of models in finance arose. In a first type of models there are sets of strategies from which the boundedly rational agents choose and are sometimes allowed to switch. Often the distinction is made between two types of agents, *e.g.* fundamental and noise traders or chartists.[375] [376] A second type of models studies populations of agents that apply hard-wired strategies, similar to competition between biological species endowed

[372] Ziemba and Vickson (2006, especially Part IV), for more hands-on advice to investing see Ziemba and Ziemba (2008).
[373] Muth 1961.
[374] Wenzelburger 2009, pp. 346–347.
[375] De Long et al. 1990; Shleifer and Summers 1990; Gode and Sunder 1993, 1997; Lux and Marchesi 1999, 2000; Sunder 2006.
[376] We discuss related issues in the context of the evolution of rationality in decision theories in Sec. 7.1.

with a fixed genome or hard-wired trait in evolutionary biology and ecology.[377] Interesting observables are the performance of individual agents or strategies measured by their wealth dynamics or returns; and measured observables for populations are their market share and the general dynamics of the resulting market behaviour, *e.g.* the behaviour of asset prices in the context of financial markets. However, early models of social interaction of heterogeneous agent models kept the interaction strength between the agents low in order to keep certain equilibria well-behaved. They thus confined themselves to the ergodic case of *e.g.* asset price behaviour, which means future asset prices converges to an ergodic equilibrium distribution that is the same for every agent and every initial distribution of the agents' characteristics. For that reason the rational expectation in such models is unique.[378]

5.2.2.1 Mathematical Finance

An interesting extension of the strategic interaction approach by Horst and Wenzelburger (2007) allows for stronger interaction between the heterogeneous agents. This increased level of interaction triggers then a causal chain and finally results in non-ergodic behaviour. The causal chain runs as follows:

1. The heterogeneous agents choose between two mediators, to whom they delegate their portfolio transactions.
2. The mediators are able to attract funds of the agents depending on their past performance which defines their market share.
3. In his target function the mediators follow a myopic optimisation scheme and have to make estimations of the ensemble mean and variance. (Ensemble) Mean-variance optimisation is the standard in (behavioural) finance literature.[379] The ensemble mean-variance optimising mediator embeds the risk in the ensemble and it is at this point that the standard 'rationality' injected into the mediators is spurious as shown throughout this thesis. It is only a matter of time until a mean-variance optimiser will be caught flat-footed by aiming for assets with seemingly favourable ensemble means of the returns (or expected returns) and will get swept away by a large fluctuation, similar t the St. Petersburg lottery with its seemingly favourable expectation value but unfavourable time-average growth rate of wealth, see Sec. 4.5. This would almost surely (a.s.) not happen to a Kelly-type mediator who is maximising his time-average growth rate of wealth.[380] A 'non-rational' mediator, in the sense that he does not follow a mean-variance strategy, will then attain larger market shares.

377 See the citations in the Subsec. 5.2.2.2.
378 Blume and Easley 1992, 2002; Brock and Durlauf 2001; Föllmer et al. 2005; Böhm and Chiarella 2005; Horst and Scheinkman 2006, 2009.
379 Markowitz 1952; Böhm and Chiarella 2005.
380 In later chapter we refer to this strategy as temporal optimisation.

4. Because a strategy which maximises the time-average growth rate is not necessarily mean-variance efficient,[381] the performance measures of the mediators are therefore no reliable signal for rational beliefs or that the mediator is managing *the* efficient portfolio.
5. As soon as one of the mediators temporarily significantly surpasses the opposing mediator, this creates incentives for imitation and strong herd behaviour and the agents' choices become correlated.
6. In the case of uncorrelated behaviour the influence of external randomness alone would uniquely converge to some distribution due to the iid stochastic effects, the LLN and some version of the CLT on which the probabilistic approach[382] relied so far.
7. However, the correlated behaviours are a source of additional endogenous randomness on a continuous basis. The additional amount of endogenous randomness inhibits the behaviour to settle down to a unique distribution.
8. In contrast, strongly path-dependent trajectories of the distributions emerge. This implies, that the time-average asset price distributions admittedly do convergence, but not uniquely and are thus highly path-dependent.
9. Thus, allowing for higher levels of interaction between market participants leads to non-ergodic market behaviour, in this case non-ergodic asset price behaviour.

The handbook article Wenzelburger (2009) developed the approach in Horst and Wenzelburger (2007) into a common framework that seeks to combine the probabilistic approach to market dynamics[383] and the deterministic approach to market dynamics[384]. From the probabilistic approach it was possible to obtain rigorous mathematical results, but the assumptions precluded many interesting phenomena such as non-ergodic dynamics which are overcome in Wenzelburger (2009, p. 349).

In another recently published text on mathematical finance[385] the growth-optimal portfolio is studied in the Black-Scholes-Merton model. However, the demanding terminology used therein and in similar sources from mathematical finance shares little commonalities with our approach and no reference to the ergodicity problem, time averages or Kelly (1956) is made. Quite the contrast, the text pegs itself to the logarithmic utility case. We therefore transition to the discussion of issues related to the ergodicity problem in evolutionary finance, where we deepen the discussion of another fine article from the *Handbook of Financial Markets*.

381 Hakansson 1971.
382 Föllmer and Schweizer 1993; Föllmer et al. 2005; Horst 2005.
383 The probabilistic approach is based on stochastic processes, see Föllmer and Schweizer (1993), Föllmer et al. (2005) and Horst (2005).
384 The deterministic approach to market dynamics is based on deterministic recursive functions such as the logistic map and cobwebs, see Brock and Hommes (1997, 1998).
385 Schachermeyer 2017, Ch. 3.

5.2.2.2 Evolutionary Finance

The approach of evolutionary finance studies the dynamics of financial markets and the dynamics of wealth distributions of the market participants and is inspired by the Darwinian scheme of variation, selection, and retention (VSR). The field has roots in diverse areas such as evolutionary economics, financial economics, economic theory, mathematical finance, and dynamical systems theory. A prolific team of authors consisting of Igor V. Evstigneev, Thorsten Hens, Rabah Amir and Klaus Reiner Schenk-Hoppé has produced many interesting publications[386] that seek to augment standard approaches to mathematical economics and finance with behavioural and evolutionary insights. We direct special attention to the very useful handbook article Evstigneev et al. (2009). The exposition of the Kelly criterion and the ascription of its proper weight in an evolutionary theory of finance in Evstigneev et al. (2009) is extraordinary even for this group of authors. 'This approach lets actions speak louder than intentions and money speak louder than happiness',[387] because the authors shun to inject any notions of utility or equilibrium concepts in this article and therefore the selective dominance of the optimal growth strategy and the property of it as an evolutionarily stable strategy (ESS) are established.[388] Thus they make the Kelly criterion *the* central pillar for the whole field. Applying the Kelly criterion in an investment strategy is temporal maximisation of wealth or equivalently maximising the time-average growth rate of wealth.

Already from Breiman (1961) follows that the Kelly strategy outperforms any other competing strategy with probability one in the asymptotic time limit, which is interpreted as its selective dominance. Following the philosophy of Alchian (1950) and Friedman (1953), the rational market participants will acquire increasing market shares and drive out other competitors in the long run. Thus the maximisation of the time-average growth rate of wealth – to which we will also refer to as temporal maximisation – qualifies a new evolutionarily reasonable rationality criterion. While this result is only an asymptotic result and for that reason seen by some authors as a too strong statement[389], the previous ensemble mean-variance decision models are not even in the asymptotic time limit superior. We devote chapter 8 to an extension of the Kelly criterion to realistic circumstances, which contains a genuine contribution to prevent the risk of overbetting in the case of uncertain gamble parameters.

Furthermore, Evstigneev et al. (2015) and Hens and Rieger (2016) are rare exceptions in the landscape of textbooks. The former treats the classical topics of modern finance

386 Evstigneev et al. 2002, 2008, 2011a; Hens and Schenk-Hoppé 2004, 2005; Amir et al. 2005, 2011.

387 Evstigneev et al. 2009, p. 510.

388 Amir et al. 2005, 2011, 2012; Lensberg and Schenk-Hoppé 2007; Evstigneev et al. 2002, 2008, 2011b; Lo et al. 2018.

389 See for example Lv and Meister (2010) and the geometric mean debate in Subsec. 5.1.1.

theory such as efficient markets, mean-variance portfolio analysis, capital asset pricing model (CAPM), derivative securities pricing, and general equilibrium models of asset markets, but additionally covers a 'less standard but very important topic, which to our knowledge has not previously been covered in elementary textbooks, is capital growth theory (Kelly, Breiman, Cover and others).'[390] However, the material on *Ergodicity Economics* as presented in Peters and Adamou (2018b) moves beyond the material in Evstigneev et al. (2015) and Hens and Rieger (2016), because *Ergodicity Economics* does derive the results without utility theory. This critique applies even more to the recent latter textbook Hens and Rieger (2016). Nevertheless growth-optimality is discussed briefly on eight pages,[391] but the conflicting though classic topics of utility theory and behavioural 'insights' to decision-making are presented extensively on the first 100 pages.

Although these authors are evidently aware of the general possibility of the non-ergodicity of financial dynamics, they do not use the term 'non-ergodic' explicitly in their textbooks,[392] but did refer to it in earlier publications[393] When they identify evolutionarily superior portfolio rules, they basically find that the optimal strategies are Kelly strategies, but they do not identify the role of time averages explicitly, especially that the Kelly criterion in the theory of optimal capital growth is a time average. However, they clearly identify the theoretical potential for a renewal of the foundations of financial mathematics that lies dormant in the optimal growth literature. In summary, the material quoted on evolutionary finance shows close relation to the approach of *Ergodicity Economics*.

5.2.3 General Equilibrium Economics

Publications on general equilibrium economics aim at well-behaved equilibria in the sense given above. In the framework of competitive Walrasian exchange economies or Arrow-Debreu economies[394] the phase space known from physics becomes the space of preferences and time is replaced by prices. In response to the negative results regarding a unique existence of economic equilibria due to Sonnenschein (1972, 1973), Mantel (1974) and Debreu (1974), a group of mathematical economists based in Germany formed around Werner Hildenbrand[395] and extended the traditional microeconomic foundations of market demand.[396] Built on earlier results of continuously differentiable mean demand functions,[397] Trockel (1984) treated individual demand as a random variable within the Arrow-Debreu framework. In order

[390] Evstigneev et al. 2015, p. vi.
[391] The subject is given only 8 pages in section 5.7 under the headings of *Evolutionary Portfolio Theory* and the *Evolutionary Portfolio Model*.
[392] Evstigneev et al. 2015; Hens and Rieger 2016.
[393] Hens and Schenk-Hoppé 2005; Evstigneev et al. 2002, p. 330, 2006, pp. 460, 462, 2011a, pp. 194, 198.
[394] Arrow and Debreu 1954; Debreu 1959; Arrow and Hahn 1971.
[395] Hildenbrand 1974.
[396] Hildenbrand 1994, p. ix.
[397] Dierker et al. 1980a,b, 1984; Trockel 1983; Hildenbrand 1980.

to smooth the market demand function he used an ergodic approach when aggregating over an ensemble of agents' characteristics. Ultimately, TROCKEL's ergodic approach enables the replacement of the price (time) averages of individual agents' demand by the ensemble average of all agents demands. The resulting ensemble average demand is unique and continuous. Thus, a nice mean demand behaviour emerges and simplifies the mathematical tractability very much in the same sense as the original *Kunstgriff* of the ergodic hypothesis simplified the mathematics for BOLTZMANN.[398]

This strand of mathematical economics advanced through the application of many results and techniques from differential topology to economic problems.[399] This development is complementary to the relatively recent econophysics movement, which applies results and techniques from dynamical systems theory, stochastic processes and statistical physics to economic problems. Even if in some contributions to general equilibrium economics the term 'ergodic' appears[400] – sometimes even prominently in the title – they are prime examples of ergodicity used in the above mentioned interpretation of a well-behaved equilibrium. Therefore, this strand of research shares no direct relation to the approach of *Ergodicity Economics*, because it does not relate the ergodicity to the treatment of risk and decision making under uncertainty, and the non-ergodic case is mostly circumvented.

5.2.4 Innovation Economics and Increasing Returns

To a large extent innovation economics makes no reference to the ergodicity problem. In his PhD thesis LAUER (1996) gives an evolutionary view of markets of consumer goods. Therein, he mentions the term 'nonergodicity' six times and recognizes the necessity of non-ergodicity as an important criterion to differentiate evolutionary economics from equilibrium economics,[401] explains verbally the meaning of non-ergodicity as a property of a process, that is not converging towards a well-behaved equilibrium, and therefore puts it in the proximity of chaos theory. LAUER urges the reader to drop the assumption of ergodicity in economic models. Finally, he argued that the introduction of psychological reasoning creates indeterminateness of the outcomes, leading to non-ergodic chains of action.[402] In summary, LAUER (1996) is aware of the importance of non-ergodicity for an economic discipline to be evolutionary, but gives neither a detailed explanation nor a rigorous (formal) definition besides the discussion on a phenomenological level.

In the early work on increasing returns by W. BRIAN ARTHUR urn models were used to successfully generate the empirical facts of path-dependence or technological lock-ins in

398 TROCKEL 1984, p. 64 and Ch. 6.
399 SMALE 1967, 1973, 1974a,b,c,d, 1976, 1980b,a, 2000; DIERKER 1974.
400 SAMUELSON 1976b; BLUME 1979; DAY and SHAFER 1987; DUFFIE 1987; DUFFIE et al. 1994.
401 LAUER 1996, pp. 31, 35.
402 LAUER 1996, pp. 38, 81.

the diffusion process of technological standards like the famous QWERTY keyboard[403] or various multimedia standards. The locked-in standards are not necessarily technologically superior and thus do not denote Pareto optimal welfare situations. Also Pólya urns – which we mentioned already in Subsec. 2.4.4 as examples for transformation that are not measure-preserving – had been among the first models of non-ergodic dynamics in economics of technological lock-ins. These kind of models were intensively studied by a group of authors that met at the International Institute for Applied Systems Analysis (IIASA) in Laxenburg, Austria.[404] Though, no reference to the conception of risk and the concept of time-average quantities is made, but nevertheless Arthur, Ermoliev and Kaniovski explicitly refer to the non-ergodic behaviour of the market shares.

5.2.5 Econometrics

In this section we discuss the relationship between econometrics as it is today and the ergodicity problem. We begin with the origin of econometrics as it was conceived by Tinbergen, Frisch and others, which led immediately to a famous debate with Keynes. From this we derive two cultures of dealing with statistics that are recently of utmost relevance in the face of the ubiquitous buzzwords big data and machine learning. From here we are naturally led to the discussion of econometrics and non-ergodicity in Post Keynesian Economics, where a recent debate on the Davidson's ergodic/non-ergodic approach is ongoing. We carefully review the coverage of ergodicity in econometrics and statistics textbooks, which mostly circumvent the ergodicity problem. The exception is a branch called dynamic econometrics, which we present at the end of this larger section jointly with a spin-off theory called Imperfect Knowledge Economics (IKE). Finally, we conclude this section as well as the larger section on related work on the ergodicity problem in economics.

5.2.5.1 Keynes-Tinbergen Debate on Econometrics

Tinbergen (1939) is one of the earliest attempts to use statistical methods to study the relationships between economic observables and especially the business cycle dynamics. The approach of using classic linear regression models led to the development of the new discipline of econometrics. In a back and forth of articles J. M. Keynes and Jan Tinbergen exchanged opposing views on the capability of this approach.[405] Their debate is known since as the Keynes-Tinbergen debate on econometrics.[406]

[403] David 1985, 2001.
[404] Arthur et al. 1983, 1986, 1987c,a,b, 1988; Arthur 1989; Dosi et al. 1994.
[405] Tinbergen 1940; Keynes 1939, 1940.
[406] Leeson 1998; Louçã 1999; Caspari 2014.

Keynes' critique on the first econometric approaches identified some problems on a mostly phenomenological level. For Keynes, there is simply too large a mismatch between economic reality and the restrictions demanded by the statistical apparatus. In his opinion, the method was not (yet) able to handle realistically modelled economic processes, because it abstracted from much of the core properties of economic data. Keynes was especially afraid to base policy recommendations on the spurious results of this method. The following quotes report in an exemplary way Keynes' stance:

- 'the factors are not independent, or the correlations involved are not linear, or there are other relevant respects in which the economic environment is not homogeneous over a period of time (perhaps because non-statistical factors are relevant)'[407] and
- 'the most important condition [in Tinbergen's approach] is that the environment in all relevant respects [. . .] should be uniform and homogeneous over a period of time. We cannot be sure that such conditions will persist in the future, even if we find them in the past.'[408]

In using modern terminology, Keynes' position can be summarised in the following way: On the one hand, economic reality is driven by and consists mainly of innovations, structural breaks in the data over time, and changes in expectations that are impossible to anticipate. On the other hand, following Syll (2015, 2018), Tinbergen's method of econometrics rests on rather strong (partly overlapping) assumptions among others:

- completeness (of having identified all relevant factors),
- temporal homogeneity of the relationships (needed for the application of econometric results in the future),
- stability (presupposition of invariant causal mechanisms that relate economic variables)
- measurability (all relevant factors are quantifiable, *e.g.* expectations and psychological factors),
- independence (of the explaining variables), and
- linearity (*i.e.* only additive contributions of error terms).

All the symptoms Keynes listed – dependence of the factors, nonlinear correlation structure, inhomogeneity of causal relationships over time, *etc.*– are indeed easily shared by non-ergodic economic observables built on the DGP. Note, even an ergodic DGP is enough to generate such symptoms *e.g.* under multiplicative dynamics, a fact we demonstrated in Sec. 4.5. A striking parallel is indeed that Keynes referred exactly to the non-linearity of effects outside the domain of additive dynamics when elaborating on linear changes in a growth rate of a geometric

407 Keynes 1939, p. 560.
408 Keynes 1939, p. 560.

series which lives in the domain of multiplicative dynamics, '[i]f the influence of changes is linear, it follows that the influence of the absolute rate is not linear.'[409]

The congruence between Keynes's critique and symptoms of non-ergodic observables has led some authors astray to repeatedly claim that Keynes was kind of the first non-ergodic economic theorist. First and foremost this applies to the writings of Paul Davidson, which we discuss in greater detail in the context of Post Keynesian Economics (PKE) in Subsec. 5.2.5.3. Of course, this is nothing but a gross exaggeration. To be distinct, Keynes did never use the word 'ergodic'. At the time he has been a professional economist, economic consultant and speculator for too long to be enough of a professional mathematician as would be necessary to know enough about the technicalities of the very abstract and remote subject of ergodic theory, let alone (fore)see its relation to economics. That is not to say he would not have owned the capabilities to do so. Furthermore, in the 1930s, a clear formulation of the foundations of ergodic theory was just underway. To summarise, there are no signs that Keynes was aware of 'ergodicity', but his stance is in principle in line with a non-ergodic approach which motivated great parts of PKE.

It is an irony of history that Tinbergen's PhD supervisor was no less a figure than the Austrian physicist Paul Ehrenfest[410]. Ehrenfest together with his mathematician wife Tatyana published the famous review of Boltzmann's approach to statistical mechanics in the encyclopedia of the mathematical sciences[411], where they coined the term of '(quasi-)ergodic hypothesis', see Sec. 2.3. We therefore assume that Tinbergen was aware of ergodicity but did not see the relations to the ergodicity problem in economics, but developed an econometric approach that is largely constrained to only handle ergodic cases. Later Tinbergen passed on the ergodic legacy to his PhD student Tjalling Koopmans and both became central figures in the development of econometrics.

However, similar debates on the capability of econometrics take place to this day and still revolve around similar arguments that were given by Keynes, see for example the recent debate revolving around the so called 'The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics'[412] which leads us over to discuss two cultures of statistical modelling.

5.2.5.2 Two Cultures of Statistical Modelling

Since the days of the Keynes-Tinbergen debate the increasing formalisation of economics was paralleled by an explosion of computing power. This starts time and again a methodological

---

409 Keynes 1939, p. 564.
410 See for the genealogy `http://genealogy.math.uni-bielefeld.de/id.php?id=110468`.
411 Ehrenfest and Ehrenfest 1911.
412 Lawson 1988, 2009; Leamer 1983, 2010; Angrist and Pischke 2010; Sims 2010; Syll 2018.

debate about the use of statistics for causal inference and prediction purposes, recently spurred in combination with the availability of massive amounts of data.[413]

In the recent debate the usefulness and soundness of the predominant methods of data analysis for inference or prediction and the attitude towards data are questioned in the natural sciences and also specifically in economics in three papers by Breiman (2001), Gelman and Shalizi (2012) and Giannakouros and Chen (2015), respectively. Each contribution juxtaposes two modelling cultures but they use different words. We condense the following three dichotomies therefore into two cultures:

- cultures of 'data modeling' versus 'algorithmic modeling',[414]
- the approach of 'classical (frequentist) statistics' versus 'Bayesian statistics/inverse probability',[415] and
- the 'econometric view' versus the 'problem-solving approach'.[416]

For the purpose of identifying two distinct cultures in statistical modelling we group the former parts of every dichotomy into a first culture of modelling philosophy and the latter parts into a second culture of modelling philosophy. Our discussion continues with the first cultural group and subsequently we discuss the second.

The first culture contains data modelling, classical frequentist statistics, and the econometric view and focuses on theory testing and parameter estimation of a static econom(etr)ic model given some sample data. In this domain, usual activities involve theory testing in general, and more specifically 'testing whether a deductively derived propositional model is rejected',[417] estimating sampling distributions, allot stars to Fisher's p-values, Neyman-Pearson hypothesis tests, Neyman's confidence intervals. Everything is framed as if it were the analysis of a repeated experiment, which is a hint to the origin of statistics, namely to support the decision making process in controlled agricultural experiments. This viewpoint can be nicely summarised: 'For Haavelmo Nature has conducted an experiment and we must guess its design so that from the given data we can test whether we have correctly hypothesized nature's laws.'[418] Furthermore, the statistician pretends to know exactly or to be able to restrict appropriately the parametric class of models as Breiman (2001, p. 202) remarks:

> "This enterprise has at its heart the belief that a statistician, by imagination and by looking at the data, can invent a reasonably good parametric class of models for a complex mechanism devised by nature. Then parameters are estimated and

413 Economist 2010a,b.
414 Breiman 2001.
415 Gelman and Shalizi 2012.
416 Giannakouros and Chen 2015.
417 Giannakouros and Chen 2015, pp. 94.
418 .Giannakouros and Chen (2015, pp. 95). The reference is Haavelmo 1944, p. 6.

> conclusions are drawn. But when a model is fit to data to draw quantitative conclusions:
>
> - The conclusions are about the model's mechanism, and not about nature's mechanism.
>
> It follows that:
>
> - If the model is a poor emulation of nature, the conclusions maybe wrong.
>
> [...] It is a strange phenomenon – once a model is made, then it becomes truth and the conclusions from it are infallible."

This reveals an underlying philosophy of static entities and relationships, namely a model of real-world phenomena which is amenable to identical preparations studied in repeated (and repeatable) experiments. In historical sciences the repeated measurement under controlled treatments of otherwise identical systems is only rarely possible, because most of the times the economist has only one realisation of the observables that interest him for the purposes of his study. In so far, this approach follows a 'philosophy of being',[419] which is an atomistic philosophy of objects and things that simply *exist* in a static state and stand in static relationships to each other, see also our brief statements on process philosophy on page 215.

In contrast, the second culture of the latter parts of the dichotomies consists of algorithmic modelling and a Bayesian problem-solving approach to statistics and is the philosophy for which both papers Breiman (2001) and Giannakouros and Chen (2015) advocate. The second modelling philosophy tries to a lesser extent to find *the* true model of nature, but can be compared to a pragmatic engineering attitude towards a multitude of approaches rather than a single particular technique.[420] Advisable activities include among others more computationally oriented techniques similar to those of 'code makers and breakers of [...] encoding, compression, and transmission of data, with no pretenses of discovering scientific laws of nature',[421] and to be aware of model uncertainty by using models in the Keynesian sense as an aid to thought. Furthermore, taking the open-endedness of the evolutionary process and the role of an observer-scientist as given is what eventually commands a different more flexible strategy. The problem-solving approach 'is a continuous process of resolving a problematic situation of which the inquirer is an integral part and that is in principle open to other inquirers.'[422] Therefore, this approach deals naturally with the reflexivity of processes of human design, where both the 'experiment' and the experimenter constantly

419 Prigogine 1980.
420 Giannakouros and Chen 2015, p. 88.
421 Giannakouros and Chen 2015, p. 97.
422 Giannakouros and Chen 2015, p. 94.

evolve. This attitude opposes a 'philosophy of being'[423] and corresponds more closely to a process philosophy, see also p. 215 and chapter 6. In this branch of philosophy processes are seen as the fundamental building blocks of reality. Hence, it is a 'philosophy of becoming'[424] where everything must be understood in a state of flux.

Interim Conclusion

Taking a broader perspective, this discussion on econometrics and data analysis can be interpreted as a continuation of the Keynes-Tinbergen debate or the Keynes-Tinbergen debate can be interpreted as being part of this bigger discussion, which is still not settled and of urging topicality for at least two reasons.[425] First, because every now and then a new revolution regarding the inferential and/or predictive capabilities and the credibility and reliability of statistical models is announced due to some 'new' data analysing techniques[426], which immediately triggers opposing views to pronouncing the 'con' in e*con*ometrics[427]. Previously, neural nets have triggered such a prediction hype around the 1990s. More recently, new techniques of machine/statistical/deep learning triggered another hype which led to the creation of a new discipline called data science and new journals.[428] Many of the mentioned techniques suffer from one fundamental problem. Better prediction usually comes at the price of lower interpretability[429]: 'Current accurate [machine learning] prediction methods are also complex black boxes. So we are facing two black boxes, where ours seems only slightly less inscrutable than nature's.'[430] Second, this discourse is important as nowadays a large proportion of PhD students in economics and data scientists all over the world are engaging in empirical research and rightly do so, but this comprises in a strong focus of building (econometric) models, thereby implicitly following a research programme of finding the 'true' DGP or causal structure. Often the underlying DGP is modelled as a deterministic process and the only way randomness enters is via a residual error term, which becomes negligible under appropriate assumptions *e.g.* ergodicity.[431] Thus, most students are not aware of the underlying implicit assumption of ergodicity.

Additionally, the demand for such (claimed) knowledge for the sake of policy or control has

423 Prigogine 1980.
424 Prigogine 1980.
425 Morgan 1990, Ch. 4.4.
426 Angrist and Pischke (2010) emphasise recent improvements in data availability, data quality and research designs and attribute the credibility revolution to several modern methods *e.g.* randomised controlled trials, fixed effects, difference-in-difference, instrumental variables, regression discontinuity *etc.*
427 Hendry 1980; McCloskey 1983; Leamer 1983, 2010; Sims 2010; Ziliak 2010.
428 The early neural nets and today's deep neural nets belonged to the transdisciplinary field of artificial intelligence.
429 See also Hutson (2018).
430 Breiman 2001, p. 210.
431 See also the introductory remarks on the nature of the DGP on p. Sec. 1.2 and Kirchgässner et al. (2013, pp. 1–5).

always been and will always be high. Hence, 'insights' derived from such data analysis and basing policy counselling on these 'insights' is done with great ease, once it is payed for. It is for exactly this reason, why it is especially important to understand, how deeply rooted and heavy loaded the assumption of the ergodicity problem in econometrics is. The Lucas critique[432] seen from the non-ergodicity perspective teaches us, that the use of econometrics and probabilities in economics and finance especially as a mean for economic policy decisions, can only end in an arms race between the econometrician/politician and the system. The politician is trying to steer the economy and the economic system under investigation. The economic system is continually (or rather often abruptly) changing its internal structure as a reaction to the econometric measurement or more general to the observation process, in striking resemblance of the famous observer effect that is invoked in quantum measurements. We continue to discuss broken ergodicity and the observer effect in greater detail also in chapter 6.

The Keynes-Tinbergen debate gave birth to a general critique of econometrics, and, more recently, to a debate in the *Journal of Post Keynesian Economics* on ergodic/non-ergodic approach by Paul Davidson. These issues are discussed in the following subsections. We start with the general ramifications of non-ergodicity for the approach of econometrics and then give an overview on Davidson's work, which allows us to understand the debate.

5.2.5.3 Non-Ergodicity in Post Keynesian Economics

The work of Paul Davidson is a great exception in the economics community, in as much as 'non-ergodicity' is key to his understanding of the economy and thus he often explicitly referred to the role of the ergodic axiom in economics. He often stressed the necessity to also consider non-ergodic stochastic processes for modelling economic phenomena and many.[433] Following Davidson (2003) the PKE community identifies the following three core axioms of neoclassical economics:

1. the gross substitution axiom,
2. the neutral money axiom, and
3. the ergodic axiom.

PKE is a school of thought in opposition to all three axioms. The gross substitution axiom asserts that every good can be substituted by any other good, therefore when the relative prices change economic agents will simply switch in their consumption from the now more expensive good to the now cheaper good. PKE studies situations where the substitution

432 Lucas 1976.
433 Davidson 1982, 1984, 1991, 1992, 1996, 2006, 2009a, 2012, 2014, 2015.

is not so smoothly, which leads to larger jumps in prices. The neutrality of money axiom assures that hording money as a mean in itself does not appear. In this view money is a mere neutral numéraire and not a good in itself. Many studies in PKE drop the axiom and explore the role of non-neutral money. Understandably, publications on the role of money experience increased public attention after periods of financial turmoil.[434] Finally, the role of the ergodic axiom in neoclassical economics and the neoclassical synthesis is pondered upon. In Davidson's understanding, ergodicity justifies drawing samples from the past to infer future behaviour of *e.g.* probabilities and prices, when actually the impossible task of drawing samples from the future would be the necessary thing to do. In his own words: 'Ergodicity implies that future outcomes are merely the statistical shadow of past and current market signals.'[435]

Keynesian Uncertainty

Several times in his writings, P. Davidson developed the concept of nonergodicity as a synonym for Keynesian uncertainty.[436] Davidson's concept of Keynesian uncertainty draws on the famous distinction of Knightian uncertainty, which we discuss in the context of broken ergodicity in 6.

To give a balanced account of Davidson's merits it also has to be stated that he overshot the mark to some extent. His repeated claims that J. M. Keynes already identified the ramifications of non-ergodicity for economics in the 1930s derived from his take on the nature of probability in *A Treatise on Probability*[437] and his later *The General Theory of Employment, Interest and Money*[438] seem a little exaggerated to say the least. To prove the limits of exegesis, we list some examples of the typical statements of Davidson:[439]

- 'Keynes produced an analysis that did not require the classical ergodic axiom',[440]
- 'Keynes's nonergodic uncertainty and animal spirits concepts',[441]
- 'Keynes rejected this ergodic axiom in developing his analysis of how a capitalist system operated in the real world when crucial decisions have to be made',[442] and
- 'Keynes argued that the economic future is uncertain and that, therefore, the classical ergodic axiom that is fundamental to any efficient market theory is not applicable to real-world financial markets'.[443]

434 Davidson 2002, 2009b; Lavoie 2011; Cynamon et al. 2012.
435 Davidson 1996, p. 480.
436 Davidson 1982, 2003, p. 248.
437 Keynes 1921.
438 Keynes 1936.
439 In Davidson (2010, p. 568) he makes himself guilty of the same misdemeanour towards Frank R. Knight.
440 Davidson 1996, p. 494.
441 Davidson 2003, p. 253.
442 Davidson 2009b, p. 37.
443 Davidson 2009b, p. 88.

The Mathematics of Non-Ergodicity Is Not of Measure Zero

For many critics of (standard) economics that follow DAVIDSON's interpretation of non-ergodicity, the implications too often involve the general impossibility of mathematical or analytical tractability of relevant models in economics. See for example the following titles of articles by DAVIDSON: 'Is Probability Theory Relevant for Uncertainty? A Post Keynesian Perspective'[444] or 'Is economics a science? Should economics be rigorous?'.[445] Nothing against a provocative title but here lies a deep misunderstanding. As Sec. 2.4 shows for ergodic theory and Subsec. 5.3.1 for ergodicity breaking, both questions are highly formal and mathematically rich subjects. It is just that we have to switch from the ensemble-centred view to the time-centred view and try to make use of time-average quantities. There is something in between the spectrum with the extreme ends of paradigmatic determinism and total disorder, and that is stochasticity. This is where a debate on ergodicity or non-ergodicity starts. Only rarely one finds more modest statements like the following in a footnote, 'Hence, although, in the 1930s, neither the classical theorists nor Keynes knew the term ergodic, Keynes's concept of uncertainty overthrows this classical postulate (which was implicitly rather than explicitly stated in classical theory).'[446]

The danger is to fall for the Post Keynesian line of a distorted interpretation of ergodicity. *E.g.* HESSE (2010) is a very interesting appraisal of the formalisation process in economics in the post-war German Federal Republic and surely is a contribution to the history of economic thought worth reading. Nevertheless, the text suffers from – what is for the purpose of the study only a minor inaccuracy – an inadvertently reversed understanding of 'ergodicity' and the false notion as if 'ergodicity' was developed by KEYNES,

> "Die „Post-Keynesianer" [...] fanden den Aspekt der „Ergodizität", das heißt der Irreversibilität [sic!] historischer Prozesse, wie er im Kapitel 12 der *General Theory* entwickelt wurde, in der neoklassischen Synthese vernachlässigt."[447]

See also MEHRLING (2008, p. 424):

> "[L]et us suppose for the sake of argument that everyone agrees that non-probabilistic uncertainty is an ever-present and important feature of our actual experience. Let us suppose further that everyone also agrees that the economy as a whole is nonergodic (as Davidson puts it) in the sense that the relative frequency of events experienced in the past is a poor guide to the relative frequency of events we will experience in the future. The question then is why an entire profession came to

444 DAVIDSON 1991.
445 DAVIDSON 2012.
446 DAVIDSON 2002, p. 36 in footnote 13.
447 HESSE 2010, p. 37.

think that it might be useful to expand the theory of choice from an environment of certainty (as Fisher) to an environment of probabilistic risk (as Arrow) [. . .].”[448]

Only Observables Can Be (Non-)Ergodic

The study of *Ergodicity Economics* teaches to discipline ourselves and strictly attach ‘ergodic’ or ‘non-ergodic’ only to observables. Otherwise we commit a philosophical fallacy known as *mistake of category* and ultimately fail to convey the main advancement in the ergodicity economics programme, namely to show a way how to handle non-ergodic dynamics in embedding their risk within time. Unfortunately, Davidson attached ‘ergodic’ and ‘non-ergodic’ to things that are no observables on various occasions, and as a result obfuscated the object under study. The following list shows in an exemplary way the problematic phrasing (our emphasis):

- ‘In an *ergodic universe*, any single event will appear to be unique to the observer only if he or she does not have a sufficient a priori or statistical knowledge of reality to properly classify this event with a group of similar conditional events.’[449]
- ‘The rare appearance of these black swans is already preprogrammed into *nature’s ergodic plan for the economy*’[450]
- ‘[. . .] often the decision makers do not have the computational power to process sufficient information about the presumed *ergodic future*.’[451]
- ‘Had behavioral theorists, post-Walrasians, and Taleb adopted Keynes’s general theory as their basic framework, irrational behavior can be explained as sensible behavior if the *economy is a nonergodic system*.’[452]

Like any other mathematical property, (non-)ergodicity can only denote a mathematical object, which is always virtual and ideal to some extent. Similarly, we would not say the bird is symplectic, we agreed on using it only for mathematical objects like groups or geometries. The real-world objects or systems may display certain phenomena and characteristics, *e.g.* evolutionary change is surely an integral part of our universe and our economies, but a universe, an economy or the future irrespective of its state can in no way be ergodic or non-ergodic nor symplectic. This would be the manifestation of a mistake of category in our terminology. Instead, the universe may evolve, the economy may exhibit structural breaks and there may be unforeseen and unforeseeable events, which altogether cause our observables of it to show non-ergodic behaviour in their path-dependent evolution.

448 Mehrling 2008, p. 424.
449 Davidson 2010, p. 568.
450 Davidson 2010, p. 567.
451 Davidson 2010, p. 570.
452 Davidson 2010, p. 570.

Although we sympathise very much with the general spirit in North (2005, Ch. 2), it is a prime example for the argument we make in this section:

> "A more fundamental reason is the *nonergodic nature* of the world we are continually altering. An *ergodic economy* is one in which the fundamental underlying structure of the economy is constant and therefore timeless. But the *world we live in is non-ergodic* — a world of continuous novel change; and comprehending the world that is evolving entails new theory, or at least modification of that which we possess."[453]

In summary, Davidson did rightly recognise the importance of non-ergodicity for economic theorising and deserves credit for this. At the same time Davidson puts a certain spin of his own interpretation on non-ergodicity that may not be helpful, if we recognise that economists have to deal with non-ergodic observables as the rule rather than the exception. But see also the acknowledgement by John R. Hicks, who was influential as one of the main protagonists in formalising many ideas from *The General Theory of Employment, Interest and Money* after Keynes' death in 1946.[454] In a quote from their personal correspondence the character of ergodicity as a mere label shines through: 'I have missed my chance, of labeling my own point of view as non-ergodic. One needs a name like that to ram a point home.'[455] 'Non-ergodicity' is of course more than only a label of somebody's philosophy.

Recent Debate on the Ergodic/Non-Ergodic Approach

Recently, Davidson's ergodic/non-ergodic approach sparked a discussion in the *Journal of Post Keynesian Economics*. The debate followed the classic structure of a thesis, antithesis and finally an attempt of a synthesis. Davidson posed the thesis of the relevance of the ergodic/non-ergodic distinction in economics.[456] The antithesis by O'Donnell dismisses the ergodic/non-ergodic approach as more or less irrelevant.[457] And finally a mediating synthesis is attempted in Rosser Jr. (2015) which is the most interesting contribution to the heated tempers.[458] Because of the described confusion surrounding the topic in the Post Keynesian context, we will not go into detail here. No final consensus seems to be possible between the different parties. Instead, we dare to conjecture that many issues will dissolve on closer inspection once they are formulated in the terminology of *Ergodicity Economics*, similar to what happened to the paradoxes involving Loschmidt's *Umkehreinwand* and Zermelo's

[453] North 2005, p. 16, our emphasis.
[454] Hicks 1980.
[455] Davidson 2014, p. 77.
[456] Davidson 1991, 1996, 2012, 2015, 2016.
[457] O'Donnell 2014; O'Donnell 2016a,b.
[458] Some spectators joined the debate on the ergodic/non-ergodic approach (Carrión Álvarez and Ehnts 2016; Nasir and Morgan 2018).

Wiederkehreinwand as well as the St. Petersburg Paradox once the language was found to properly formalise what has seemed paradoxical at first.[459]

5.2.5.4 Ergodicity Problem in Econometrics

In economics and finance there is often only one realisation of a particular time series, hence the econometrician is confronted with a situation where there is no statistical ensemble from which to compute anything like an ensemble-average observable – apart from situations with cross-sectional data, where this is possible and valid, of course. Hence, non-ergodicity poses a profound problem for econometrics and the analysis of time series in general. The ergodicity problem in econometrics is summarised as follows. Often the econometrician has one unique historical realisation of a stochastic process as a time series of an economically relevant observable. What justifies the inference from certain parameters of one realisation to the whole nature of a stochastic process? Some authors refer to this situation as non-experimental data.[460] The ergodic hypothesis can be understood as a methodological link between non-repeatable experiments and ex post statistical regularity which translate into ex ante forecasting accuracy.

Although econometricians produce techniques to deal with structural breaks and the like[461], usually this is not connected to 'non-ergodicity'. The ergodicity problem in econometrics applies to all realisations of a stochastic process, but does apply as well to the DGP – the underlying model representation in econometrics. Even though the stochastic term in a DGP is only conceived as an error term, the stochastic term makes any realisation at a fixed moment in time a random variable, thus the DGP is a stochastic process, too.[462] Only because it is usually assumed that by the LLN the stochastic error term becomes arbitrarily negligible, it is why time series in economics are traditionally not treated as a stochastic process but as a deterministic process.

Ergodicity in Econometrics Textbooks

In general, ergodicity is not treated appropriately in almost all econometrics textbooks. If the topic is touched upon at all, usually the stationarity of the first two moments is assumed. Thereby, silently sneaking in the additional assumption about the existence of the moments or the implicit restriction to distributions which have a finite first and second moment,

459 Petersen 1989, p. 35; Rothstein 1974; Peters 2011c; Peters and Adamou 2018b.
460 Wold 1969; Poitras 2018.
461 For example local linear regression or segmented regression which produce a polygon chain of regression line segments.
462 Van Kampen 2007, p. 52.

respectively. Consulting some of the most commonly used econometrics textbooks[463] on the ergodicity problem or the mere terms 'ergodicity', 'ensemble' and 'time average' will remain largely inconclusive in how to proceed in the non-ergodic case. At rare occasions, stationarity is ensured via mean-ergodicity and (co)variance-ergodicity.[464] Let us exemplary quote from some textbooks on time series econometrics:

- 'Economic applications of [an ergodic theorem] depend on whether it is reasonable to suppose that economic time series are stationary and ergodic. Ergodicity is often difficult to ascertain theoretically [...] and is impossible to verify empirically (since this requires an infinite sample). Stationarity, on the other hand, is a property that can be investigated empirically. However, many important economic time series seem not to be stationary [and thus are also not ergodic] but heterogeneous, exhibiting means, variances, and covariances that change over time.'[465]

- 'A further requirement is that the process be ergodic. Since this is a yet more difficult concept, which cannot be adequately explained without the introduction of considerable mathematical apparatus not required elsewhere. [...] [p. 5] Unfortunately, it is not possible to test for ergodicity using a single realization [...]. Ergodicity will be assumed to hold in all situations considered in later sections.'[466]

- 'We should emphasise that the procedure of using a single realisation to infer the unknown parameters of a joint probability distribution is only valid if the process is ergodic, which roughly means that the sample moments for finite stretches of the realisation approach their population counterparts as the length of the realisation becomes infinite. Since it is very difficult to test for ergodicity using just (part of) a single realisation, it will be assumed from now on that all time series have this property.'[467]

Of course, there exist textbooks which include a treatment of ergodicity, but usually these are more specialised texts often not tailored to an econom(etr)ics audience.[468] The exception proving the rule is Hamilton (1994, Ch. 3.2), who even uses the terminology of ensemble and time averages. Another exception applies to textbooks on non-linear time series analysis, see especially Kantz and Schreiber (2003, pp. 197–199).[469] Not exactly a textbook is McCauley (2009), but he dismantles many of the assumptions needed to apply standard methods of econometrics (white noise error terms in regression models, deriving ergodicity

463 Granger and Newbold 1986; Tsay 2002; Tsay and Chen 2019; Lütkepohl 2005; Mills 2015; Neusser 2016; Tanaka 2017.
464 Kirchgässner et al. 2013, p. 13; Huschens 2017, p. 209; White 1984, p. 43.
465 White 1984, p. 43.
466 Granger and Newbold 1986, pp. 4–5.
467 Mills and Markellos 2008, p. 10; Mills 2015, p. 6.
468 Hannan 1970; Basawa and Scott 1983; Davidson 1994; Kantz and Schreiber 2003.
469 Tsay (2002, p. 130) and Tsay and Chen (2019, pp. 42–43) only briefly mention ergodicity and the problem to infer from one realisation to the whole nature of a stochastic process.

from increment stationarity of $I(d)$ integrated non-stationary time series beyond the WIENER process, and cointegration as a generalisation of $I(d)$ integration built on too simple conceptions of noise) as invalid in the light of their conformity to properties of empirical data. DOMOWITZ and EL-GAMAL (2001, pp. 365–366) further corroborate the general result of our analysis:

> "The concept of ergodicity is fundamental in the analysis of economic time series and of dynamic models calibrated by time series data. It is, therefore, surprising that no general testing procedure has been proposed to examine [p. 366] this important hypothesis."

Circumventing the Ergodicity Problem

The implicit assumption of ergodicity usually enters into econometric models via an explicit stationarity assumption, *e.g.* of the stochastic error term – although the two are not trivially linked. *E.g.* a stochastic process can be stationary in the sense that the process is characterised by stationary distribution parameters but does not necessarily need to be ergodic. Vice versa, a non-ergodic process can be stationary. The application of the usual statistical methods requires the time series to be at least stationary and often ergodic. Since TINBERGEN's early general linear models (GLMs) with stationary error terms, further techniques were developed which include generalised least squares estimation and generalised autoregressive conditional heteroskedasticity (GARCH) models and tests for ergodicity of them.[470] WOLD and JURÉEN (1953, Ch. 9.4) contains an early econometric analysis of demand that clearly articulates the role of the (pointwise) ergodic theorem and the CLT in the study of moving averages of (co-variance) stationary time series.[471]

If an economic time series is not stationary from the outset, it is a *common practice* to apply differencing to the original time series which is also called an integrated time series $I(d)$ of order d. The original (not integrated) time series is denoted by $I(0)$ and the first differences are denoted by $I(1)$ and so on. NASIR and MORGAN (2018) criticise that the status of differencing as a *common practice* may lead to a situation where important facts such as the information loss due to differencing pass unnoticed and that the derivation of policies from integrated time series may not straightforwardly apply to the original time series. In further publications the econometrics community starts to explore non-stationary dynamics which possess a unit

470 MEITZ and SAIKKONEN 2008; KAPETANIOS and SHIN 2011.

471 Important to note, throughout the book WOLD and JURÉEN (1953) one finds the familiar notation of ensemble averages as space averages but also the (in modern terminology) confusing notation of time average as phase averages.

root, which means the process is not mean-reverting. First empirical tests are developed[472] also in the context of analysing agent-based simulations (ABSs).[473]

Another attempt to circumvent the ergodicity problem in econometrics is to create an ensemble out of slices of many non-overlapping time windows out of a single time series. This technique is called sliding window time averages.[474] Eventually, this technique does not lead to satisfactory results, because neither the convergence of time-average observables to the ensemble-average observable, nor the convergence of the time-average quantities at all can be proven.[475] However, the development of econometric tests for ergodicity of time series is not exactly our focus of embedding risk within historical time for decision making under uncertainty.

Dynamic Econometrics

As we have demonstrated, the usual way is either to not mention (non-)ergodicity and the problems that come about with it at all or to restrict the analysis in a way that solely ergodic processes are apriori allowed. A notable exception is the work by David F. Hendry and co-authors on dynamic econometrics, which is very much in line with the already mentioned philosophy of becoming. Nevertheless, in a book of more than 800 pages even Hendry draws the conclusion:

> "Whether economic reality is an ergodic process after suitable transformations is a deep issue which we cannot analyse rigorously in this [800+ pages] book. [...] [In later parts of the book] we will be concerned with regime shifts and other factors that might induce structural change in the economy, and which in practice make ergodicity untenable without careful formulation of the stochastic variables for which it is claimed."[476]

Nonetheless, Hendry's work on *Dynamic Econometrics*[477] evolves in a direction from where a merger with the *Ergodicity Economics* ansatz seems possible. Over the years, together with co-authors, he developed in a series of textbooks and papers a statistical framework of the invigorating sort, because he identifies clearly problematic aspects in the relation between the

---

472 Domowitz and El-Gamal 1993, 2001; Corradi et al. 2000; Meitz and Saikkonen 2008; Kapetanios and Shin 2011.

473 Windrum et al. 2007; Grazzini 2012; Grazzini and Richiardi 2015; Delli Gatti et al. 2018.

474 McCauley 2008, 2008, Ch. 3.7.8 & 7.3.

475 This is due to the fact, that no ergodic theorem exists for non-stationary processes which would allow the replacement of time averages by ensemble averages. And even for non-stationary processes with stationary increments the time windows are strongly correlated so that the LLN and, respectively, Tchebychov's theorem do not apply, which is needed to prove a particular convergence.

476 Hendry 1995, p. 100.

477 Hendry 1995.

assumptions of econometrics and real-world characteristics of the economic process. To name just a few aspects:

- intertemporal optimisation breaks down when structural breaks/unanticipated change abound, 'ceteris paribus assumptions are inappropriate in a non-stationary, or evolving, world',[478]
- most econometric models are timeless, in the sense that they 'emphasise 'deep parameters' of taste and technology that are claimed to be temporally invariant',[479]
- structural breaks render it very unlikely if not impossible for individuals (*e.g.* forecast policy effects) as well as for institutions (see *e.g.* performance of consumer price index (CPI) inflation forecasts) to generate forecast errors that are minimum mean squared.[480]

This all boils down to the implicit assumption of the ergodic hypothesis, although broken ergodicity is prevalent in almost every macroeconomic data sets as indicated by the above list. With their statistician glasses, Hendry and collaborators usually label this broken ergodicity of observables as 'structural breaks', 'unanticipated change' or 'location shifts', which are endemic in the real world. Their approach deals with such issues head-on and ultimately seeks to handle possible but unknown mis-specifications in econometric forecasting models of economic processes, where structural breaks in the data are the rule rather than the exception. The structural breaks manifest themselves in location shifts in the underlying (parameters of) distributions,that appear at unknown times by unknown magnitudes.[481] [482]

Hendry's main conclusion is that if the economic world we seek to describe is non-stationary, evolving and experiencing unanticipated breaks and changes, then modelling has to be an evolutionary process as well, with clear selection criteria and the models' ability to be quickly adjustable after such unanticipated changes. Which is in stark contrast, if in order to adjust a model, one needs to add a whole financial sector or introduce the concept of money to the model. The approach of dynamics econometrics thus is in the spirit of a philosophy of becoming.

---

478 Hendry and Mizon 2011b, p. 178.
479 Hendry and Mizon 2011b, p. 179.
480 Hendry and Mizon 2011b, p. 180.
481 Hendry 2012, p. 320.
482 See the textbooks Hendry (1995), Hendry and Morgan (1997), Clements and Hendry (1999) and Hendry and Nielsen (2007) or for articles Hendry (2000a,b) and Hendry and Mizon (2011a, 2014). For an overview of the whole approach and applications of some techniques see especially Hendry (2012).

Imperfect Knowledge Economics

The approach of IKE strives to merge two fundamental insights in the economic process. First, the dispersion of knowledge across the market participants.[483] Often their information processing leads to opposing views, the prerequisite for any market transaction. And second, the fundamental uncertainty of economic change.[484]. The latter is referred to as 'non-routine change' in IKE, where we would again refer to broken ergodicity due to the non-routine change. It is this combination that leads to the break with the tradition of the REH[485] and dynamic stochastic general equilibrium (DSGE) models especially in macroeconomics (and their special case of real business cycle (RBC) models). In an economic environment that experiences non-routine change, the market participants are simply unable to generate rational expectations[486], because the location shifts in the distributional parameters are highly path-dependent on the history of the economy.[487] Thus, 'rational expectations' are no longer rational in the presence of broken ergodicity. More specifically, if the economic actors in reality can have only imperfect knowledge about their environment, then the economic actors within the models of the economy also possess only imperfect knowledge. How can the economist, who builds a model that contains himself, can be free of this imperfection?[488] Thus, the even more important imperfection is on the side of the economist or econometrician, who models the economy with his imperfect knowledge about his research object. The general IKE approach has been first and foremost developed by Roman Frydman and Michael Goldberg,[489] but actually has a long history.[490] In recent years, the approach of IKE utilised dynamic econometrics through the cointegrated vector autoregressive (CoVAR) modelling technique. A Scandinavian group of statisticians and econometricians around Søren Johansen, Katarina Juselius and Morten N. Tabor took up the thread to further develop the theory started by Hendry's dynamic econometrics.[491] Primarily, the imperfect knowledge is modelled via trajectory-and-time-dependent parameters in a CoVAR framework.[492]

483 Which is often attributed to von Hayek (1945, 1975).
484 Which is often attributed to Keynes or Frank Knight.
485 Muth 1961.
486 In the sense of being able to generate minimum mean square expectation errors, see also Subsec. 5.2.2.
487 The econometrics literature on unemployment, exchange rates and trade often refers to the occurrence of path dependence as a consequence of some exogenous shock as hysteresis, see Amable et al. (1994), Røed (1997), Setterfield (1997), Blanchard and Summers (1986), Piscitelli et al. (2000) and Belke et al. (2014). For the underlying mathematics see Krasnosel'skii and Pokrovskii (1989) and Mayergoyz (1991).
488 Sargent 1995, Ch. 1.
489 Frydman and Goldberg 2011, 2012, 2013; Frydman and Phelps 2013.
490 Phelps 1970; Frydman 1982; Frydman and Phelps 1983; Frydman and Goldberg 2003.
491 Johansen 1995; Juselius 2006; Tabor 2014; Frydman et al. 2017.
492 Another recent variant of the class of vector autoregressive (VAR) econometric models are structural vector autoregressive (SVAR) models, which try to explicitly incorporate time-varying parameters. See Sims (1980, 2002).

5.2.6 Conclusion

Before we discuss work related to the ergodicity problem in economics but from outside of economics, we want to facilitate the transition by a tongue-in-check conclusion, which still conveys a message. Given the necessary level of formalism to arrive at a proper understanding of ergodicity and given the scanty discussion of non-ergodicity inside economics, every single mentioning of the term '(non-)ergodic' in articles and books outside of economics come as a surprise. Interestingly, authors who are critical of the econ tribe[493] of our times are sometimes aware of an abstract problem identified with the fancy term 'non-ergodic'. Astonishing remarks come *e.g.* from the angle of (economic) sociology writes:

> "[V]iele der Autoren [führen] die Krise der ökonomischen Theorie auf die Tatsache zurück, daß ihre Modelle die Bedeutung derZeit nicht angemessen berücksichtigen.
> [...]
> Gerade die Statistik führe zu einem „ergodischen“ Verständnis der Zeit. Die Zeit der Ökonomen sei jedoch historisch, da sie aus Entscheidungen hervorgehe, die Zäsuren markieren und so einen Unterschied machen."[494]

Also other publications by Elena Esposito contain stimulating discussions on the role of time, probability, contingency and an uncertain and open future in economics and are interesting in their connection to chapter 6.[495] Another unexpected example from a critical spectator of the econ tribe refers to ergodicity in the following way,

> "Die Wahrscheinlichkeit des Risikos gewinnt dadurch eine „ergodische“ Struktur und will, dass alle künftigen Ereignis-Populationen analog zu aktuellen Wahrscheinlichkeiten gestreut sind und sich nach einer Relation verhalten, in der eine Serie von tausend Würfen mit einem Würfel einem einzigen Wurf mit tausend Würfeln entspricht."[496]

Much to our amusement there is the branch of 'ergodic literature' in literature studies, which refers to a style of the fabrication of texts, which involves a feedback loop between the creator and reader of a text. One example is if a text includes hyperlinks the reader can decide whether to follow the link or not. Thus, the reading process becomes non-linear due to different stories that can evolve from the bifurcations arising at the hyperlinks. As a consequence ergodic literature requires non-trivial efforts by the reader to traverse the text.[497] For this reason we have the impression that it should be labelled non-ergodic literature.

493 Leijonhufvud 1973.
494 Esposito 2007, pp. 100–101.
495 Esposito 2010.
496 Vogl 2011, pp. 168–169.
497 Aarseth 1997, p. 1.

Apart from its name, ergodic literature has nothing in common with mathematical ergodic theory.

5.3 Ergodicity Problem Beyond Economics

Since ergodic theory is part of various courses in advanced mathematics (stochastics, dynamical systems theory, number theory, *etc.*) and physics curriculum (statistical physics and information theory) many examples given here to elucidate its importance stem from these fields.

5.3.1 Physics

5.3.1.1 Continuing Relevance of Mandelbrot's Non-Ergodic Models of Long-Range Dependence

> [G]ive full idea how 'erratic' the functions with [infinite expectation are. Many functions have time averages with] infinite variance. This means that enormous deviations from average behavior can be expected.[498]
>
> Benoît B. Mandelbrot

Two recent publications by Graves et al. (2017) and Watkins (2017) rediscovered a series of papers by Benoît B. Mandelbrot from the 1960s[499] in the context of modelling long-range dependence (LRD), which is a common property of many time series. Fueled by his visual imagination Mandelbrot saw the connection between the property of long memory in hydrology and economics. Thereby the central problem in hydrology is to store water in times of plenty to survive in times of drought. To solve this problem people started damming river valleys and the question of optimal heights of dams naturally arises. The optimal dam satisfies the following conditions:

> "Storage is ideal if (a) the outflow is uniform, (b) the reservoir ends the period as full as it started it, (c) the dam never overflows, (d) the capacity is the smallest compatible with (a), (b), and (c). [Because the random variable which denotes the river level has infinite expectation and scales with the observation length] Such

498 Mandelbrot 1967, p. 296.
499 Mandelbrot 1963, 1965a,b, 1967; Mandelbrot and Van Ness 1968.

> an ideal dam cannot be designed in advance, except by luck, since the data to design it are only available after the fact. Actual design is, nevertheless, helped by knowledge of the capacities of the ideal dams corresponding to every s-year period within the known record."[500]

This problem has striking resemblance to the economic problem of optimal capital reserves. For a given period what is the (d) smallest amount of money we need to accumulate during booms in order to survive times of crisis and to still generate (a) a constant cash flow to meet our expenses, (b) the amount of money at the beginning of a period equals the amount at the end, (c) we never run out of money in the meantime. Similar as the river levels also the largest losses in financial markets are dependent on the observation period and thus a random variable with an expectation value which is in a scaling relationship with the observation length.[501] As a consequence of the long memory the auto-correlation function (ACF) has no characteristic time scale, similar to the St. Petersburg lottery which has no characteristic payout scale. Such data exhibits a diverging or a power-law behaviour of the ACF which inherits to power-law behaviour at the origin of the Fourier spectrum. This phenomena is also known as the infrared catastrophe, $1/f$-noise or flicker noise. Now there are two paradigmatic classes of generators of $1/f$ noise or long memory. The first is the auto-regressive fractionally integrated moving average (ARFIMA) class of ergodic and stationary models. A convenient way to model LRD is to use the stationary ARFIMA processes which exhibit a rapid decay in their ACF. However, the ARFIMA models have LRD by construction and rely on an a priori stationarity and ergodicity assumption. The second class are the fractional Brownian motion (fBm) models, which are a generalisation of Brownian motion (BM) for generalised exponents of the backshift operator $H \in [0, 1]$, *i.e.* the model allows fractional exponents, which is why they are sometimes referred to as fractal Brownian motion models. fBm is in general non-ergodic and has a time-dependent conditional spectral density.[502] For a Hurst exponent of $H = 1/2$ the fBm is an iid sequence and is called fractional Gaussian noise (fGn) because it coincides with the stationary increments of BM. fBm models are an important example of weak ergodicity breaking (WEB) and belong to the class of fractional renewal models which again belong to the family of continuous-time random walk (CTRW) models.

Thus Mandelbrot's fBm and his papers from the 1960s provide a non-ergodic solution to the long memory problem, which is often ignored in comparison to the ergodic ARFIMA and fGn models. The built in LRD of fGn in the ergodic model class caused some confusion. From the observation of LRD in the data it was inferred that the data was generated by ergodic models. However, from observing LRD alone it is not sufficient to exclude the non-ergodic

500 Mandelbrot and Wallis 1969, p. 243, our remarks.
501 Mandelbrot dubbed such temporal scaling as the Joseph effect and spatial scaling as the Noah effect.
502 Nualart 2007.

class of models. The focus on a specific frequency range dictates the model choice. Whereas the ARFIMA approach has an inherent focus on the short-term and high-frequency spectrum which is relevant for forecasting, the fGn approach has an inherent long-term focus. This signifies another time an inherent one-sidedness of model choice in economics which is also present in the disregard for dynamics in decision theory, the study of one-shot games and on many other occasions.

Interestingly, from this origin developed two modelling approaches. In econometrics and hydrology the dominant approach is the stationary and ergodic fGn and even more the ARFIMA models. However, in physics the non-ergodic fBm and α-stable distributions prevail.[503] These findings are especially interesting, because two recent dissertations in finance were defended at the author's department which modelled conditional variances using ARFIMA and especially the GARCH family of models.[504] Unaware of Mandelbrot's non-ergodic models, we asked during one defence whether there exists a class of models which build a bridge between Mandelbrot's approach using heavy-tailed distributions such as the family of the α-stable distribution and the econometric approach presented by the two then PhD candidates which relied on distributions with finite first moments to model certain stylised facts of time series. Watkins (2017) and Graves et al. (2017) answered this question with their rediscovery of the continuing relevance of Mandelbrot's fBm models.

5.3.1.2 Non-Extensive Statistical Mechanics

Ergodicity is assumed throughout standard statistical mechanics, because the substitutability of time averages by ensemble averages is needed as a crucial operation. This yields entropies of the standard Boltzmann-Gibbs (BG) type, S_{BG}. As most complex systems are non-ergodic systems, the standard entropy concept does not apply, because of non-trivial internal correlation structures and their dynamics. Thus, it is hoped that generalised entropies apply to non-ergodic complex systems out of equilibrium.

> "Ergodicity, this is to say, dynamics whose time averages coincide with ensemble averages, naturally leads to BG statistical mechanics, hence to standard thermodynamics. This formalism has been at the basis of an enormous success in describing, among others, the particular stationary state corresponding to thermal equilibrium. There are, however, vast classes of complex systems which accommodate quite badly, or even not at all, within the BG formalism. Such dynamical systems exhibit, in one way or another, non-ergodic aspects."[505]

503 Graves et al. 2017, p. 18.
504 Walther 2017; Klein 2017.
505 Tsallis et al. 2003, p. 89.

From this quote we see the intricate connection between ergodicity and (in this case thermal) equilibrium concepts. Among such above mentioned complex systems which obey non-ergodic properties are surely economic systems, too. Within a new proposed framework of non-extensive statistical mechanics also non-equilibrium systems (that usually obey non-ergodicity) become manageable. Using generalised entropies that are applicable *e.g.* for option pricing purposes and to volumes and distributions of returns on financial markets in non-ergodic observables of economic systems.[506]

Since the seminal paper Tsallis (1988) there is an intense discussion about possible generalisations of statistical mechanics and its range of application.[507] It has been shown that the entropy functional (which connects thermodynamic entropy and the microscopic level) is not unique, but can be generalised in various ways. Of course, not every possible generalisation needs to have a sensible physical interpretation, which is why this issue is intensely discussed in the community. Examples for generalised entropies with a sensible physical interpretation are the Rényi entropy[508] or the Tsallis entropy.[509] What is determining the specific form of the entropy functional is the dynamical and geometrical correlation structure of the elements of the system.

Central to Tsallis' generalisation is a crucial analysis of the additivity and extensivity of the entropy. For any two probabilistically independent systems A and B, the entropy S is additive iff

$$S(A+B) = S(A) + S(B). \tag{5.1}$$

In classical statistical mechanics the BG entropy functional is correctly depicting the thermodynamic entropy with the probabilities, p_i, of the number of microconfigurations that constitute the macroscopic states W. For discrete numbers of states the entropy is defined as

$$S_{\mathrm{BG}}(p) = -k \sum_{i=1}^{W} p_i \ln p_i \ , \tag{5.2}$$

with the probabilities summing up to unity, $\sum_{i=1}^{W} p_i = 1$, and k being a constant (like Boltzmann constant k_B or $k = 1$ as usually taken in information theory). For the case that each macrostate has equal probability eq. (5.2) simplifies to Boltzmann's famous epitaph,

$$S_{\mathrm{BG}} = -k \ln W. \tag{5.3}$$

506 Tsallis et al. 2003.
507 Tsallis 2014; Beck 2009.
508 Rényi 1961.
509 Tsallis 1988.

Whereas extensivity is a statement of the dependence of the entropy on the system size as measured *e.g.* by the number of elements, N, in the system. If the entropy scales proportionally to an (infinitely) increasing system size, $S(N) \propto N$ for $N \to \infty$, then the entropy is said to be extensive. Thus, in comparison to additivity extensivity depends additionally on the nature of the correlations. The specific correlations could in principle lead to systemic effects that either cancel each other out or become reinforced. Standard BG statistical mechanics applies to the case of completely independent and thus uncorrelated particle movements and therefore has an extensive entropy.[510] As introduced already in eq. (2.18), the SHANNON entropy is denoted by a plain S and will reappear as a limit for both generalised entropies of TSALLIS and RÉNYI.

However, if entropies are non-additive, statistical mechanics is no longer extensive, and a non-extensive version of statistical mechanics is needed. Such a non-extensive statistical mechanics was proposed in TSALLIS (1988) and generalises entropy in the following way to so called q-entropies,

$$S_q(p) = k \, \frac{1 - \sum_{i=1}^{W} p_i^q}{q-1} \, , \tag{5.4}$$

for $q \neq 1$. In the limit of $q \to 1$ BG statistical mechanics and entropy are regained, and S_q reduces to the BG or SHANNON entropy, $S_{q=1} = S_{\text{BG}} = S$,

$$\lim_{q \to 1} S_q = S_{\text{BG}} = S. \tag{5.5}$$

In fact a whole apparatus of q-statistics arises using generalised forms of logarithms and exponentials – so called q-logarithms and q-exponentials.[511] The entropic index q can be interpreted. It is intimately connected to the 'geometry of the measure in phase space on which probability is concentrated',[512] which creates the link to our topic of non-ergodicity:

> "It should be clear that nonextensive statistical mechanics is by no means intended to replace BG statistical mechanics for systems such as those in stationary states characterized by thermal equilibrium consistent with ergodicity. The nonextensive alternative is proposed, instead, as a way of dealing, through mathematical methods that are quite similar to the usual ones, with anomalous systems. These include a wide class of nonergodic systems [. . .]."[513]

Non-extensive statistical mechanics has proven to be very useful to scientists of many disciplines as it is extending the range of applicability of both the classical CLT for iid random variables

510 Examples for other extensive physical quantities are number of molecules, mass, volume, energy, *etc.* Examples of intensive physical quantities are concentration, density, pressure, temperature, *etc.*
511 Instead of the prefix q often referred to as TSALLIS statistics or TSALLIS entropy.
512 BAIS and FARMER 2008, p. 659.
513 GELL-MANN and TSALLIS 2004, p. xiv.

with a finite second moment (*e.g.* the Gaussian standard case) and the generalised Lévy-Gnedenko CLT for iid random variables with infinite second moment, to accommodate also correlated random variables. This leads to the class of (q, α)-stable distribution, which yields for $q \geq 1, \alpha \leq 2$ the class of stable distributions of Lévy[514] and Gnedenko;[515] and as particular instances for $(q = 1, \alpha = 2)$ the Gaussian distribution family and so called q-Gaussians for $(q \geq 1, \alpha = 2)$.[516]

Some studies have rigorously shown that the usage of generalised entropies which are derived from non-extensive statistical mechanics is required, if systems are strongly interacting and therefore non-ergodic. This is the case because non-ergodicity leads to a violation of the fourth Shannon-Khinchin axiom (assuring the additivity of the sub-entropies) for such systems. A relaxation of this axiom then naturally leads to generalised entropies which uniquely characterise any system by two scaling exponents that define equivalence classes for all systems.[517] Several studies analyse explicitly the non-ergodic properties of economic systems using q-statistics and often find structures known from other non-equilibrium physical systems such as power law distributions,[518] non-Gaussian power grid frequency fluctuations[519] or apply q-statistics to the microeconomics of probabilistic choice.[520]

Animated living organic systems are systems that display exactly these pronounced internal correlations mentioned at the beginning and often classified as organised complexity,[521] which is why the application of this approach to economic systems seems promising. Unlike inanimate systems of unorganised complexity[522] or molecular chaos at the microscopic level to which BG statistical mechanics applies. Economic systems are organic wholes, in a sense even living systems. Therefore the social sciences should become acquainted with established results and follow recent developments within the field of non-extensive statistical mechanics, for example see Tsallis (2009, Ch. 7.3) on applications in economics.

Rényi generalised the Boltzmann-Gibbs-Shannon entropy. The Rényi entropy or entropy of order $\alpha, 0 < \alpha \neq 1$, is given by

$$S_\alpha(p) = \frac{1}{1-\alpha} \log \left(\sum_{k=1}^{n} p_k^\alpha \right). \tag{5.6}$$

514 Lévy 1937.
515 Gnedenko and Kolmogorov 1954.
516 See also Kirstein (2012) for a detailed analysis of the usefulness of stable distributions for economic processes.
517 Hanel et al. 2011, 2014; Thurner and Hanel 2012.
518 Duarte Queiros et al. 2005.
519 Schäfer et al. 2018.
520 Takahashi 2007.
521 Weaver 1948.
522 Weaver 1948.

Similar to the TSALLIS entropy, in the limit of $\alpha \to 1$ BG statistical mechanics and entropy are regained, and S_α reduces to BG entropy, S_{BG}, or SHANNON entropy, S, respectively,

$$\lim_{\alpha \to 1} S_\alpha = S_{\text{BG}} = S. \tag{5.7}$$

The RÉNYI entropy is heavily used in information theory, *e.g.* in the estimation of the dimension of multifractals[523] and to measure entanglement in quantum information theory, and generalises the KULLBACK-LEIBLER divergence.[524] The order parameter α can be interpreted in terms of the diversity in (*e.g.* ecological) systems.

Benefits from Generalised Logarithms and Exponentials for Ergodicity Economics

The computation of ensemble averages relies only on a probability distribution. To compute time averages, however, an additional assumption about the dynamic is required. A noisy multiplicative dynamic is often a natural choice to model real-world growth phenomena such as wealth accumulation or cell growth. In economics this is known as the effect of compound interest. Since multiplicative growth is a reasonable model for wealth accumulation, a logarithm enters naturally in the computation of the time-average growth rate such as in eq. (4.115). Of course also other forms than a pure multiplicative or additive dynamic are conceivable to govern real-world dynamics. Generalised logarithms may moderate between different dynamics and therefore studies on generalised logarithms in statistical mechanics may provide a stock of results to build upon.[525]

5.3.1.3 Ergodicity Breaking and Single Particle Tracking

Diverse forms of ergodic and non-ergodic behaviours are observed and studied in the field of anomalous diffusion.[526] Anomalous diffusion is a hypernym to encompass subdiffusive and superdiffusive behaviour. Superdiffusion explores space faster than normal diffusion of BROWNIAN motion and subdiffusion explores space slower. One is then interested in whether a suitably defined observable on this dynamic is ergodic and hence the usual apparatus of statistical mechanics based on ensemble statistics applies or whether the observable is non-ergodic and further analysis requires more care in the choice of suitable evaluation methods. In anomalous diffusion one studies the diffusion behaviour over time of a molecule, *i.e.* the position, x, of a molecule at a certain time, t, which gives a trajectory or time series, $x(t)$,

523 MANDELBROT 1987.
524 KULLBACK and LEIBLER 1951; EBELING et al. 1998, Ch. 3.8.
525 GELL-MANN and TSALLIS 2004; TSALLIS 2009; HANEL et al. 2011.
526 CHERSTVY et al. 2013.

that is then analysed. Now, an often studied observable is the mean squared displacement (MSD), $\left\langle x^2(t) \right\rangle$ or, in other words, the mean distance travelled by the molecule over a specific time span t,

$$\left\langle x^2(t) \right\rangle = \int_{-\infty}^{\infty} x^2 P(x,t)\,\mathrm{d}x, \tag{5.8}$$

if the particle is released at the origin at time $t_0 = 0$, $P(x,t)$ is a Gaussian PDF and the solution of the diffusion equation. The standard way of studying the MSD is to use its ensemble average version, $\left\langle x^2(t) \right\rangle$. Similarly, the time average mean square displacement, $\overline{\delta^2(\Delta)}$, can be studied too

$$\overline{\delta^2(\Delta)} = \frac{1}{t-\Delta}\int_0^{t-\Delta} \Big[x(t'+\Delta) - x(t')\Big]^2 \,\mathrm{d}t', \tag{5.9}$$

with the time series $x(t')$ of length t, the sliding window of width $\Delta \ll t$.

In the ergodic case, the time average MSD in equation (5.9) (at least) for long measurement times is equivalent to the ensemble average MSD in equation (5.8),

$$\overline{\delta^2(\Delta)} = \left\langle x^2(\Delta) \right\rangle. \tag{5.10}$$

However, in experiments distinct disparities between the time averages and the regularly computed ensemble averages are found, '[c]ontinuous-time random-walk anomalous diffusion violates the Boltzmann–Khinchin ergodic hypothesis and even long-time averages of an observable differ from the corresponding ensemble average',[527] which turns eq. (5.10) into an inequality in eq. (5.11) for many real life processes

$$\overline{\delta^2(\Delta)} \neq \left\langle x^2(\Delta) \right\rangle. \tag{5.11}$$

Furthermore, on these observables so-called *ergodicity breaking parameters* are constructed in order to analyse the distance from ergodic MSD of Brownian motion.[528] Bouchaud (1992) coined the terms and put forth the discrimination of 'weak ergodicity breaking' and 'strong ergodicity breaking'. Strong ergodicity breaking occurs if there is a compartmentalised phase space such that only proper subsets of the whole phase space are explored. The quasi-ergodic hypothesis – that every point in phase space will be arbitrary closely visited – is clearly violated in this case. The strong form of ergodicity breaking is in a sense a straightforward form of non-ergodicity. In the case of weak ergodicity breaking quasi-full coverage of the phase space is in principle possible or not prohibited, but does not occur due to several reasons. Therefore, weak ergodicity breaking is especially interesting and currently an active research topic in the anomalous diffusion community, partly spurred by the availability of

527 Metzler 2015, p. 114.
528 Burov et al. 2011.

novel experimental techniques such as single particle tracking,[529] single molecule tracking in living cells[530] or blinking quantum dots.[531]

The progress in anomalous diffusion triggered renewed interest in Schrödinger's famous question: *What is Life?* Schrödinger pointed out that the origin of life involves the non-equilibrium states of nature[532] of processuality of life and observes many non-ergodic dynamics to understand complex molecules, such as the building and folding of proteins.[533] Nucleic acids, mRNA and proteins are of intensified interest as candidate single molecules.[534] *E.g.* the process from folding to unfolding of proteins dissipates energy and manifests the out-of-equilibrium behaviour of the phenomenon, which contrasts traditional notions of thermodynamics.[535] Other application areas of non-ergodic subdiffusion involve non-living physical systems such as disordered semiconductors and groundwater motion in geophysical formations, and in the crowded interiors of living cells.[536]

The Relevance of Results on Anomalous Diffusion for Economics

The results from single particle tracking experiments are closely related to economic processes. The original interest in random walks followed mainly spatial aspects.[537] The CTRW models used to study anomalous diffusion instead is especially focused on the temporal aspects such as the distribution of waiting times between two molecular collisions.[538] In economics, the tracing of single particle trajectories corresponds to a time series of an economically relevant observable. Unlike in experimental sciences, an econometrician is confronted with exactly one factual realisation of an economic (stochastic) process. In other words, an econometrician has only one time series at his disposal. If ensemble statistical mechanics does not hold for single particles, economic theories that are based in a similar fashion on the idea that randomness is embedded within the ensemble do not hold as well. This problem can be construed as the unique factual realisation in time versus the ensemble of all possible realisation, which has been a recurring theme *e.g.* in Subsec. 5.2.5.

529 Metzler et al. 2014.
530 Weigel et al. 2011; Barkai et al. 2012.
531 Stefani et al. 2009.
532 Jarzynski 2011.
533 Ornes 2017.
534 He et al. 2008.
535 Jarzynski 2011, p. 591.
536 Sokolov 2008.
537 Brown 1828; Pearson 1905a,b; Rayleigh 1905; Einstein 1905.
538 Sokolov 2008; Metzler et al. 2009; Klafter and Sokolov 2011.

5.3.1.4 Pervasiveness and Virtuality of Ensemble Thinking

The goal of this section is to illustrate the virtuality of the conception of an ensemble of systems by exemplary quotations from eminent scientists. A rebuttal of idea of an unrestricted validity of ensemble thinking and ensemble-averaged quantities is important because the embedding of randomness in the ensemble is deeply ingrained and pervasive in the collective thinking that it may take a while before the mind is ready for a different embedding of uncertainty. As the Schumpeter section quote tells us, changing thinking habits is among the hardest things. The built-in error correction mode of science is nothing we can rely on, but scientists have to actively advocate for their convictions. It is not a new insight, that otherwise the fluctuations in the error correction mode of science can take longer swings than the time scale of the career of individual scientists.

As the preceding sections showed, the concept of the expectation value – that is an ensemble average – and with it the whole formalism of premodern probability theory stems from the 17^{th} century and predates any notion of ergodicity by centuries. Since then the tools to handle non-ergodicity wait to be used to their fullest extent in economics. On the one hand most theoretical structures cling to an embedding of randomness in the ensemble even in context where they are misleading such as in the case of decision under uncertainty. On the other hand this facts entails plenty of problems to work on.

Equipped with this finding, it is therefore hardly surprising that one can find ample manifestations of thinking that is frozen in the ensemble domain and basically equating the concept of a proper prospect with an ensemble average or expectation value. As one example serves the following quotation of Keynes who mainly speaks about radical uncertainty and the possibility to construct a probabilistic judgement. However, his conception of the uncertainty is fully embedded in the ensemble domain as becomes visible by the end of this quote:

> "By 'uncertain' knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty; nor is the prospect of a Victory bond being drawn. Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealthowners in the social system in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know. Nevertheless, the necessity for action and for decision compels us as practical men to do our best to overlook this awkward fact and to behave exactly as we should if we had behind us a good

> Benthamite calculation of a series of prospective advantages and disadvantages, each multiplied by its appropriate probability, waiting to be summed."[539]

The everyday usage of arguments that are implicitly based on parallel universes – especially in contexts where it is actually not relevant – shows how pervasive this wrong usage of this mathematical object *expectation value* is. That is why the ensemble and the time perspective have to be made explicit way beyond the St. Petersburg paradox in the general context of decision making under uncertainty which is a major part of economics. The pervasiveness of thinking in ensembles manifests if it is falsely applied to individuals. This happens every time when people deal with averages. A single person will never throw a 3.5 with a single die and no woman has 2.5 children.[540]

Ensembles as Fantasies

Many of the eminent physicists of the 20$^{\text{th}}$ century draw a clear demarcation line between physically relevant objects and virtual objects that are only mere temporary crutches to overcome a specific problem. For instance, in their famous series of textbooks LANDAU and LIFSHITZ circumvent the ergodic hypothesis at all as a justification for statistical physics:

> "It should be emphasised once more that this distribution is not the true statistical distribution for a closed system. Regarding it as the true distribution is equivalent to asserting that, in the course of a sufficiently long time, the phase trajectory of a closed system passes arbitrarily close to every point of the manifold defined by equations (4.3). But this assertion (called the ergodic hypothesis) is certain1y not true in general."[541]

Already in the preface of his seminal *Elementary Principles in Statistical Mechanics* GIBBS prepares the readership for the new approach:

> "The usual point of view in the study of mechanics is that where the attention is mainly directed to the changes which take place in the course of time in a given system. [...] Inquiries of this kind are often simplified by taking into consideration conditions of the system other than those through which it actually passes or is supposed to pass, but our attention is not usually carried beyond conditions differing infinitesimally from those which are regarded as actual.
> For some purposes, however, it is desirable to take a broader view of the subject. We may imagine a great number of systems of the same nature, but differing in

539 KEYNES 1937, pp. 213–214.
540 The latter statements concerning the abundance of children stems from POUNDSTONE (2005).
541 LANDAU and LIFSHITZ 1980, Preface and the second footnote on p. 12.

> the configurations and velocities which they have at a given instant, and differing not merely infinitesimally, but it may be so as to embrace every conceivable combination of configuration and velocities. And here we may set the problem, not to follow a particular system through its succession of configurations, but to determine how the whole number of systems will be distributed among the various conceivable configurations and velocities at any required time, when the distribution has been given for some one time."[542]

Obvious from this quote is the virtuality of the ensemble, it is only an artefact of the mind, but a useful one. Von Neumann (1929, p. 30) emphasised the same fact when he discussed the conditions of ergodicity, when 'die statistischen Eigenschaften der (fiktiven) mikrokanonischen Gesamtheit für das mangelhaft bekannte (wirkliche) System unterstellt werden dürfen.' Another clear treatment comes from the Austrian-Hungarian companion of von Neumann Erwin Schrödinger:

> "Here the N identical systems are mental copies of the one system under consideration – of the one macroscopic device that is actually erected on our laboratory table. Now what on earth could it mean, physically, to distribute a given amount of energy E over these N mental copies. The idea is, in my view, that you can, of course, imagine that you really had N copies of your system, that they really were in 'weak interaction' with each other, but isolated from the rest of the world.
> [. . . p. 9 . . .]
> the Gibbs point of view, namely, that we are dealing with a virtual ensemble, of which the single member is the system really under consideration."[543]

Also the work of Jan von Plato contains interesting passages on the ergodicity problem. He often discussed the virtuality behind the ensemble approach in the context of the ergodic hypothesis in statistical mechanics, the time-average interpretation of probability and related issues:

> "However, the ensemble cannot exist physically, even if its use in calculations leads to agreement with what is observed.[544]
> There is a definite connection between the probability-as-time-average and probability-as-limit-of-relative-frequency views[545]
> According to Gibbs's followers, the ensemble distribution is not a physical one. It is rather justified by its 'usefulness' in giving the right results. The main problem with this approach is how the unphysical a priori probabilities can lead to any

[542] Gibbs 1902, Preface, p. vii.
[543] Schrödinger 1948, pp. 3, 9.
[544] Von Plato 1987, p. 380.
[545] Von Plato 1987, p. 380.

> physically meaningful conclusions.[546]
> Late in 1931, the single-system approach experienced a major breakthrough, as John von Neumann and George D. Birkhoff established their ergodic theorems for classical systems fulfilling certain conditions. Ironically, this took place after the probabilistic (Born's interpretation of the wave function) and indeterministic (Heisenberg's relations) character of the new quantum mechanics was established.[547]"

Von Plato's comment on quantum mechanics lead us to the role of the ensemble in quantum mechanics.

The Multiverse in Cosmology and Quantum Physics

The notion of parallel worlds or parallel universes is also known as the multiverse. Theories, which include a multiverse are not alien to science. Although the genre of science fiction literature has dealt with multiverse concepts a long time before it spilled over into scientific debates, mainly due to the introduction of probability theory into the foundations of physics with the creation of quantum mechanics.[548] The dynamics in quantum physics is deterministic in general,

> "According to the Schrödinger equation, this wavefunction evolves over time in a deterministic fashion that mathematicians term 'unitary'. Although quantum mechanics is often described as inherently random and uncertain, there is nothing random or uncertain about the way the wavefunction evolves."[549]

However, interpretations of quantum mechanics differ on the issue of how Schrödinger's wave function, which describes the evolution of quantum states, is thought to collapse to macroscopic observable states. For a long time the dominant interpretation is attributed to the group around Niels Bohr and therefore called the *Copenhagen Interpretation*. The many world interpretation of quantum mechanics (MWI) is one alternative explanation to the famous Copenhagen interpretation, that gained renewed interest in the last decades.[550] Everett III (1957) gave birth to an ongoing discussion on the reality and possibility of parallel worlds. It took almost 20 years until a longer and therefore better amenable version of the many-world interpretation in Everett III (1973) was discussed on a broader range due to DeWitt (1970) and DeWitt and Graham (1973). Associated to the multiverse are well the known issues and paradoxa, which quantum mechanics introduced into the natural sciences, such as the possibility to travel through worm holes, time travel, the grandfather paradox,

[546] Von Plato 1987, p. 384.
[547] Von Plato 1987, p. 382.
[548] See e. g. the issue of *Nature* Vol. 448 Number 7149 and especially Nature (2007) and Wolfe (2007).
[549] Tegmark 2007, p. 23.
[550] Everett III 1957; Tegmark 2003, 2007.

etc. The MWI postulates that all states in the ensemble of possible states could realise in many parallel worlds. Thus at every instance of time a random experiment is performed on the quantum level to determine the factual macroscopic state all states realise, but we can not experience them. Therefore the MWI relates the microcosm to a macrocosm which is actually a multiverse, which also led to some popularity within the cosmology community. As a mere spectator from the outside these incidents document nevertheless an increased openness to competing interpretations of quantum mechanics, which can not be ruled out on formal grounds..[551]

5.3.2 Gambling & Information Theory

In information theory the potential of Kelly (1956) has been realised to the fullest extent. This is due to the fact that coding effectiveness depends on the channel capacity and the entropy. As Kelly showed entropy rate is just the optimal time-average growth rate.

Thomas Cover is the co-author of the widely used textbook Cover and Thomas (2006), which contains a chapter on Kelly gambling and horse races. Cover has a track record of papers combining Kelly's insights and questions in finance which lead to solutions to the St. Petersburg paradox and so called universal or growth-optimal portfolios.[552] A nice review on connections between ergodic theory and information theory is Shields (1998). Many insights in information theory have motivated studies in evolutionary biology.[553]

5.3.3 Evolutionary Biology and the Origin of Behaviour

In biology the study of repeated interactions arises quite naturally in the context of evolution. A key term in evolutionary biology is the fitness of an individual, a group or a whole species. This motivated analogous use of the fitness metaphor to evaluate the performance of individuals, firms or business strategies in the field of evolutionary economics.[554] In a static timeless view fitness is equated with the profits, but there is not much evolution associated with this view. A more evolutionary understanding of fitness considers naturally the performance over time. Therefore it is less a surprise that we find many publications which find a strategy which maximises the geometric mean of the growth factors of the fitness to perform better over time than strategies which maximise the expected absolute fitness. The earliest examples of this line of research in evolutionary biology are Cohen (1966) and Lewontin and Cohen (1969). In the years to come many publications in evolutionary biology have similar titles

551 Tegmark 1998, p. 855.
552 Bell and Cover 1980; Cover 1991.
553 Donaldson-Matasci et al. 2008, 2010.
554 Veblen 1898; Alchian 1950; Axelrod and Hamilton 1981; Selten 1991; Lindgren 1992.

which contain phrases similar to *evolutionary performance in randomly varying environments*. Sometimes papers with such titles even study decision making in uncertain environments explicitly and derive very similar equations like *Ergodicity Economics* and we did in this thesis. However, they are just re-derivations of the geometric-mean property and do not recognise the useful distinction between time and ensemble averages introduced by *Ergodicity Economics*.[555] However, this translates to what we referred to as noisy multiplicative growth using the *Ergodicity Economics* terminology. This is another indicator of the widespread appearance of multiplicative dynamics, which arise naturally in many contexts and it is not only the relevant dynamic which applies to wealth but also predominant in biology.[556]

As we already indicated in Subsec. 5.2.1.3 and Subsec. 5.2.2.2, the growth-optimal strategy can be understood as a dominant strategy which outperforms other strategies and an ESS. The evolutionary biology community labelled this as geometric-mean fitness, because the growth factors determining the number of offspring are just geometric means under multiplicative dynamics.[557]

Another branch of research in evolutionary biology models the same effect of geometric-mean fitness using the Kelly's bet-hedging scheme.[558] Yet another more recent branch operates under the heading of the *Origin of Behaviour* and argues similar to Alchian (1950) and Friedman (1953), that the selective dominance of growth-optimal behaviour is the origin of the widespread finding of it.[559]

Miscellaneous

Econophysics One would expect to find discussions of ergodicity or the ergodicity problem in the econophysics community, which is not the case. However, exceptions confirm a rule, and some publications use the term 'ergodicity',[560] but as well do not identify the ergodicity problem in economics or connect ergodicity to an embedding of risk within historical time as is done in *Ergodicity Economics*. As always, an exception proves the rule. Farjoun and Machover (1983) is an early classic in the application of statistical methods to political economy. In chapter 3 they discuss the profit rate as a random variable from which they derive a critique of the common equilibrium conception of the uniform rate of profit as a result of averaging over rates of profit in different industries. In fact very much in line with

555 Yoshimura et al. 2009, 2013a,b.
556 Peters and Adamou 2019.
557 Orr 2009; Hilbert 2017.
558 Seger and Brockmann 1987; Philippi and Seger 1989; Donaldson-Matasci et al. 2008; Beaumont et al. 2009; Simons 2009, 2011; Crean and Marshall 2009; Olofsson et al. 2009; Rees et al. 2010; Ripa et al. 2010; Starrfelt and Kokko 2012; Grimbergen et al. 2015.
559 Brennan and Lo 2011; Lo et al. 2018; Orr 2009; Zhang et al. 2014a,b; Orr 2018.
560 Kolesnikov and Rühl 2010; Zapart 2015; Poitras and Heaney 2015; Poitras 2018.

our goal, they are explicit on the role of ergodicity and made a careful distinction between time and ensemble averages that is worth reading:

> "This time average is conceptually very different from the space average of Z [ensemble average of a random variable Z], which we have called EZ [expectation value of a random variable Z]. In the former [time average perspective], [the realisation] i is held fixed, t is allowed to vary, and the averaging is performed over all t from 0 to T; in the latter [ensemble average perspective], time is held fixed and [the realisations] i allowed to vary, so that the averaging is performed over the whole sample space. (The term space average thus refers to the sample space, not to physical real space.)"[561]

Demography Some studies in demography research rely on invariant limit distributions and utilise some ergodic theorems.[562]

Geomorphology Methodological uniformitarianism is the reigning methodology in geomorphology but to some extent assumed in all sciences. In short, the uniformitarianism principle states that the same natural laws and processes hold at all times and at all places in the universe in the past. It explains the present as a logical consequence of certain laws at work in the past. For example in geomorphology, forces which uniformly acted upon rocks and sediments over centuries generated the shape and layering of today's landscapes. The same laws apply through time and space, *e.g.* the same law of sedimentation and erosion was active during ages and over the whole planet.[563]

Developmental Psychology In a series of papers[564] Peter Molenaar has indicated the importance of ergodicity in the analysis of empirical time series in developmental psychology and that most of test theory is based on the ergodicity assumption. In developmental psychological studies certain treatment effects trigger *e.g.* learning processes which generate non-stationary characteristics within a group of people. Thus intra- and inter-individual variation do in general not coincide and show different dynamics.

Process Philosophy The origin of a philosophy which acknowledges dynamics in the form of processes as the fundamental building blocks of reality goes at least back to Heraklit and his famous dictum 'pantha rhei' and several modern authors[565] of whom Nicholas

561 Farjoun and Machover 1983, p. 48, our additions in square brackets.
562 See for example Cohen (1976, 1977a,b, 1979b,a) and Lange (1979). We thank Adolf Wagner for indications which led to our awareness of the use of ergodicity in the field of demography.
563 See Paine (1985) and the course description Advanced Geomorphology.
564 Molenaar 2004, 2007.
565 Heidegger 2002; Whitehead 1978; Prigogine 1980.

Rescher[566] is one of the leading figures: 'Important though logic and language are [...] it is the mathematical language of process-of transformation functions and differential equations – that is of the greatest help in depicting the world's physical realities.'[567]

The reality of time in scientific or economic explanations is discussed explicitly in Smolin (2009, 2013, 2015) and Mangabeira Unger and Smolin (2015), in which the authors advocate for the philosophical stance of temporal naturalism which takes time seriously. '... the paradox of living in time and believing in the timeless'[568] is a central puzzle. Even more recently a friendly debate on the true nature of time between Carlo Rovelli[569] and Smolin[570] continues to this day. In the ensuing chapter 6 we deepen the discussion of the nature of time and its relevance for a methodology in economics.

566 Rescher 1962, 1996, 2000.
567 Rescher 2000, p. 20.
568 Smolin 2013, p. xiv.
569 Rovelli 2011, 2018b,a.
570 Smolin 2009, 2013, 2014, 2015; Mangabeira Unger and Smolin 2015.

6 From the Ergodic Hypothesis in Physics to the Ergodic Axiom in Economics

In this chapter shows how the status of the ergodic *hypothesis* central to statistical mechanics undetectedly morphed into the ergodic *axiom* in economics. The work of PAUL SAMUELSON who we already quoted in the introduction turns out to be crucial, yet not solely responsible for that.

To shed light on this status change, this section is structured in four main parts. In the Sec. 6.1 we reconstruct the intellectual lineage connecting the main figures GIBBS and SAMUELSON. Thereby a genealogy of an idea of ergodic economics is reconstructed. Sec. 6.2 analyses briefly the idea of rationality in the conduct of science as opposed to rationality of economic agents. In Sec. 6.3 we discuss some epistemological aspects of a possible superior form of a methodology for economics and analyse methodological spillovers from physics. In Subsec. 6.3.1 we sketch the epistemological and historical context of logical positivism during the time when the mathematisation of economics proceeded and at the same time the foundation of this epistemology became severely challenged by GÖDEL's theorems (Subsec. 6.3.2). We identify ergodicity among the key drivers in the process of mathematisation of economics. The methodological aspects lead over to Subsec. 6.4.4 and corroborate the general thrust of the literature review, i.e. we find only rare occasions where the fundamental assumption of ergodicity as a key driver of the mathematisation is made explicit. Section Sec. 6.5 concludes chapter.

6.1 From Boltzmann & Gibbs via Wilson to Samuelson

A first hint of who and what inspired much of the mathematisation of economics methodology from the 1930s onwards is given on the very first page of SAMUELSON's seminal *Foundations of Economic Analysis*. There SAMUELSON quotes the physicist and one of the founders of statistical mechanics J. WILLARD GIBBS directly below the book title with the statement 'Mathematics is a Language'. This quote alone has the power to explain most of the following and is the condensed form of the more elaborate section to come.

This section answers the following question: What role did the ergodic hypothesis play in the mathematisation of economics? To answer this question we repeat the quote we used in the introduction from Samuelson (1968, pp. 11), because it stands out from the literature and sets us on the right track:

> "[...] interesting [...] assumption implicit and explicit in the classical mind. It was a belief in unique long-run equilibrium independent of initial conditions. I shall call it the '**ergodic hypothesis**' by analogy to the use of this term in statistical mechanics. [...]
> Now, Paul Samuelson, aged 20 [...] as an equilibrium theorist he naturally tended to think of models in which things settle down to a unique position independently of initial conditions. Technically speaking, we theorists hoped not to introduce hysteresis phenomena into our model, as the Bible does when it says 'We pass this way only once' and, in so saying, takes the subject **out of the realm of science into the realm of genuine history**.
> [...] we envisaged an oversimplified model with the following **ergodic property**: no matter how we start [...] after a sufficiently long time it will become [...] a unique **ergodic state**."[571]

This quote is typical for Samuelson's rhetorical style and telling in several ways. First, it is a fully conscious loaning of the ergodic hypothesis from physics for economic theorising. There are very few other statements, making this fact so explicit, but no earlier one.[572] The following questions guide the analysis in what follows:

1. How is it that Samuelson was aware of the underlying assumption of ergodicity to (his) mathematical economics and became a fierce proponent of the ergodic case?
2. Is there a superior methodology of economics? What can be said about that methodology compared to those of historical disciplines and the natural sciences?
3. Is the ergodic hypothesis empirically validated to rightly assume it in economics?

We tackle question 1 first and question 2 in Sec. 6.3. The third question has been answered already, see e.g. Sec. 4.5.

Samuelson and the Ergodic Side of the Coin

The longer quote above by Samuelson signals his awareness of the necessity of the ergodic hypothesis for equilibrium economics to work. His awareness alone is exceptional as this quote

571 Quotation marks are in the original, only the bold emphasis is added.

572 The first papers being aware of this loaning process appear 30 years later by Paul Davidson (see Davidson 1991, 1996, 2009a, 2012).

from 1968 is the first[573] explicit mentioning of the precise term 'ergodic' in an economics context our literature analysis brought about. Although some justified voices have been raised about how to correctly interpret Samuelson's quote and whether it is not a tongue-in-cheek statement,[574] we explicitly observe that if Samuelson's publications are related to ergodicity or even carry 'ergodic' in its title or the term at least appears somewhere in the paper,[575] it is without exception that Samuelson focused on the ergodic side of the coin, and did not treat the ergodicity problem in its totality, which would include the non-ergodic case as well.

But what led Samuelson to be aware of at least this side of the coin? Well, he gave the answer himself, he 'was vaccinated early to understand that economics and physics could share the same formal mathematical theorems',[576] as he describes in 'How *Foundations* came to be'. We retrace the main reason for this physics inclination to Samuelson's education. Samuelson was a Ph.D. student at Harvard from 1936 to 1941 and from 1937-1940 even a junior fellow. Harvard Junior Fellows in these days were engaging in a 'Faustian bargain [...] free to work on whatever they liked, but forbidden for three years to work toward any degree or Ph.D. dissertation.'[577] During this year he became the only (self-declared) disciple of Edwin Bidwell Wilson. Wilson was a versatile mathematician, who taught statistics and a course on mathematical economics which was essentially a course on equilibrium theory to Harvard graduate students from 1922-1945.[578] As an indication of Wilson's widespread scholarship let us mention that he served as the president of the American Academy of Arts and Sciences from 1927-1931 and he was the first 'managing editor of the Proceedings for fifty years, from its first issue, dated January 15, 1915, until his death in December 1964.'[579] Not only were the first ergodic theorems published in Wilson's PNAS, the long-lasting discussion they triggered involved many publications of new ergodic theorems. The PNAS became *the* journal for early publications on ergodic theory.[580]and still are to the present day.[581][582] Importantly, Wilson was again a student of Gibbs, who himself did not have

573 Although in Samuelson (1962, p. 5) the 'strong ergodic theorem' is cited, but not in relation to economics. Nevertheless it proves his speaking acquaintance with ergodicity.

574 Carrión Álvarez and Ehnts 2016.

575 Samuelson 1962, 1971b,a, 1973, 1976b,a.

576 Samuelson 1998, p. 1376.

577 Samuelson 1998, p. 1377.

578 James Tobin reports of the influence Wilson's courses at Harvard's Public Health School had on him, too. See Colander (2007, p. 394).

579 Hunsaker and Mac Lane (1973, p. 300). See also Bogorad (1995) for information regarding Wilson's merits for the Proceedings of the National Academy of Sciences of the USA (PNAS).

580 Among others see Birkhoff (1931a), von Neumann (1932c), Hopf (1932), Wintner (1932), Birkhoff and Koopman (1932), Alaoglu and Birkhoff (1939), Halmos (1946), Kallianpur and Robbins (1953), Harris and Robbins (1953), Derman (1954) and Prigogine et al. (1976).

581 Moore 2015; Ashley 2015.

582 A query for `ergodic` in the title of all PNAS publications yields 28 hits, if `ergodic` can appear in the title or abstract the number increases to 45 hits and we get 280 hits if we further allow appearances in the text (23rd December, 2015). For such a broad journal that is not specialised on ergodicity it shows the importance and interest of the editor in this topic. However this is not reflected for the section attributed

many students but as one of the founders of statistical mechanics many followers.[583] This is the intellectual lineage Samuelson sees himself as 'perhaps [Wilson's] only disciple'[584] and elsewhere '[Wilson] had been the only protégé at Yale of Willard Gibbs. Since I was Wilson's main protégé, that makes me kind of a grandson to Gibbs'.[585] Thus we have established the intellectual connection linking Gibbs to Samuelson via Wilson.[586]

Weintraub (1991) provides many hints on possible topics from mathematics and the natural sciences of interest to economists. Many of them were brought to Samuelson's attention by Wilson, among others Le Chatelier's principle from chemistry,[587] quadratic forms[588], the relation between Lotka's *Elements of Physical Biology* and Wilhelm Ostwald's *Energetics*[589] and the work of George David Birkhoff on the mathematics of dynamical systems.[590] The link to Birkhoff is important, because he was one of the early founders and contributors to ergodic theory, and deliveres an early proof of the pointwise ergodic theorem (see 2.1).[591]

In his chapter *E. B. Wilson and the Gibbs tradition*[592] Weintraub reconstructs, how the work of Samuelson can be seen as the continuation of a Gibbsian tradition in other scientific disciplines. Weintraub (1991, p. 61) summarizes the essence of the Gibbsian tradition as

> "the prejudices of E. B. Wilson, a world-view shaped as a student of Willard Gibbs, were congenial to Samuelson's own program of making economics scientific by presenting the essential propositions of the subject in a mathematical, and thus clearly analyzable, form."

With this remarks, it becomes clear, why Samuelson was inclined to choose the Gibbs quote right below the title in his *Foundations*, because they shared a common worldview

to economics within PNAS, as 6 out of 280 hits are from the economic sciences category only. The search for `non-ergodic` or `nonergodic` on the broadest scope yields 67 hits, where none appeared in the economic sciences section.

583 Gibbs 1902; Lebowitz and Penrose 1973.
584 See Samuelson (1998, p. 1376). For self documented influence on him, see also Samuelson (1983).
585 Barnett 2007, p. 160.
586 See also Mirowski 1989.
587 Weintraub 1991, p. 46.
588 Generalisations of statistical variances to quadratic forms became the topic of Samuelson's first PhD student Lawrence Klein (Samuelson 1990, p. 255).
589 See Ostwald (1909) *Energetische Grundlagen der Kulturwissenschaft*. Ostwald is of further interest as a promoter of Gibbs' ideas and translator of Gibbs (1892) into German (Körber 1961, p. XV, pp. 87–104). Furthermore, Ostwald was a scientific opponent of atomism, which Boltzmann's kinetic theory of gases launched, but still proved his integrity in being the driving force behind the appointment of Boltzmann by Leipzig University from 1900-1902 and was later convinced by the experimental verification of atoms.
590 Weintraub 1991, pp. 47, 49–53, 64, 70.
591 Birkhoff 1931b.
592 Weintraub 1991, pp. 57–62.

in making disciplines more scientific through mathematisation.[593] Unfortunately, this contributed to the manifestation of a ill-conceived view on ergodicity and its relevance for economics.

Ludwig Boltzmann 1884, H-Theorem

  Henri Poincaré recurrence theorem

    J. Willard Gibbs (Yale 1871-1903) *Principles* 1902

      George D. Birkhoff *PNAS* 1931
      John von Neumann *PNAS* 1932

        Edwin B. Wilson (Yale 1901-1911, Harvard PHS 1922-45,
        PNAS Managing Editor 1915-1964)

          Paul Samuelson (Harvard 1936-1941)
          Foundations 1947 'Mathematics is a Language'
          Economics 1948

Figure 6.1: Genealogy of the Ergodic Hypothesis in Physics to the Ergodic Axiom in Economics. The figure establishes the intellectual lineage between Samuelson's mindset and the ideas of Boltzmann and especially Gibbs, which were moderated via Samuelson was under the direct influence of Wilson at MIT.

The lineage regarding ideas, that are central for becoming acquainted with the ergodic hypothesis, thus reads Gibbs-Wilson-Samuelson. In order to incorporate more scientists relevant with respect to the ergodic hypothesis, it could be justly expanded to Poincaré-Gibbs-Birkhoff-Wilson-Samuelson, see figure **Fig. 6.1**. Surely it was through Wilson that Samuelson came in to close contact with the mathematics of a physics' image of the world in those times, including the meaning of the ergodic hypothesis. Not only for Samuelson was Wilson a rich source of inspiration, many economists who were looking for mathematical structures and concepts from the natural sciences waiting to be applied elsewhere turned to him as he was strongly supporting the use of mathematical concepts in other disciplines such as in economics.[594] But there is evidence that Samuelson had direct contact with Birkhoff on the topic of ergodic theory, as he recalls in his reminiscences about Norbert Wiener:

593 Interestingly, if speaking of the mathematisation of economics Gibbs was the supervisor of Irving Fisher's Ph.D. thesis on 'Mathematical Investigations in the Theory of Value and Prices' at Yale University. Fisher is considered as one of the earliest American neoclassical economists and Fisher (1892) as 'a classic memoir on mathematical economics' (Samuelson 1990, p. 260).

594 A forerunner of this development took place around 1890s with the research programme of energetics by Ostwald (1909), which tried to base everything including cultural phenomena on the principle of energy conservation (Körber 1961, p. XV).

> "But I can testify from my personal experience of dining with Dean George Birkhoff about once in every eight weeks at the Society of Fellows during 1937-1940 that Birkhoff talked much about Wiener's rival efforts in connection with then hot ergodic theory and much else.[595] [...]
> Not infrequently Wiener's name would come up as when once I recalled him [Birkhoff] saying: 'Norbert did it again at the Joint Seminar. He was able to roll out the Weak Ergodic Theorem from the Fundamental Theorem of the Calculus.'[596]"

These historical remarks help to explain PAUL SAMUELSON's special role and to understand how he came to know about ergodicity and its fundamental relevance for the seeming success of (equilibrium) economics. In this regard, SAMUELSON's importance is hard to underestimate, that is why some scholars labeled the skein of mathematical economics as *Samuelsonian economics*.[597] To gain an understanding of SAMUELSON's enormous impact on economics taught and developed at MIT and even for the whole scientific discipline some authors even equate MIT economics with *Samuelsonian economics* and furthermore equate this again with mathematical economics plus the desire to control the economic system.[598] Especially the desire to control the economic system has to be understood in terms of the MIT as an engineering school. Usually, engineers try to understand systems and approach their research subject with the explicit wish to be in control of its behaviour. A project pursued at MIT very similar in spirit and under SAMUELSON's watchful eyes was WIENER's cybernetics, who also contributed to ergodic theory.[599]

6.2 From Rational Mechanics to Rational Economics

The subtitle of the classic textbook on statistical mechanics GIBBS (1902) reads *Developed with Especial Reference to the Rational Foundation of Thermodynamics*. What is meant by *rational* in this subtitle? It is science understood as an endeavour, which yields explanations via the use of reason and not metaphysical forces like ether, telepathy or ideological (especially religious) beliefs. In this way science is an attempt that strives for rationality and at the same time tries to work out to what extent observable natural phenomena can be explained by the means of reason and rational explanations.[600] This approach to the understanding of nature became a desideratum of many researchers like GALILEO GALILEI, RENÉ DESCARTES

595 SAMUELSON 1997, p. 38.
596 SAMUELSON 1997, p. 41.
597 WEINTRAUB (1991) and MCCLOSKEY (2002). For the enormous range of the parts in which SAMUELSON made an impact see SZENBERG et al. (2006).
598 CHERRIER 2014.
599 WIENER 1939, 1956a; WIENER and WINTNER 1941a,b, 1943, 1958.
600 Of course, this does in no sense imply, that science can answer everything or that its scope spans everything that is important.

and foremost Isaac Newton and big parts of the society, in order to free themselves from religious tutelage and/or royal dominion. This can not be analysed without the historic context in which these developments are embedded. Nowadays they are conceived in one way or another as the founders of modern science beginning roughly with the Enlightenment period from 1650 onwards.

The French Revolution initiated later intellectual developments of the Enlightenment that led to the rise of positivism and later empiricism e.g. in the Vienna Circle and ultimately in an unconditional trust in science. As we write these lines, we live in a world that is dominated by the ideology of the superiority of technology and scientific argument in public discourse. There may be good reasons for that, not least if science is the worst form of understanding nature and society except all those other forms that have been tried from time to time that turned out to be even worse[601]. In this sense, we nowadays experience a situation of great danger, since the most dangerous ideology could be the belief of being free of any ideology. Even if it is maybe the purest ideology of believing in (what one believes) science (is). It is in this sense calls for 'Disenlightenment' have to be understood.[602] The approach of *Ergodicity Economics* involves a disenlightenment of traditional economics and rebuilding it through re-enlightenment. Thus the proponents of disenlightenment do not vote for the relegation of science but for an enlightened handling also and especially of its limitations even if they are much harder to see. Such conduct has generated Sorcerer's Apprentices of evolution and with it the spectre of a repairable world.[603]

Occasionally, science and scientist claim, they form their conclusion in a manner free of any ideology. Often, this false image is attributed to science (or even worse consciously used) by non-scientists, e.g. journalists or politicians[604]. Scientists have to fend this allegation off. Scientists have to make a stand against the utilization of their work by interests alien to science and make up their minds, whether it is possible to use research findings outside of their lab-like context e.g. in political discourse. This is not an easy thing to do, as it requires the scientist to be clear about his own ideological biases, prejudices and presuppositions. The Economist concludes '[Ms Reinhart and Mr. Rogoff] have sometimes been less careful in media articles. This is perhaps their biggest mistake. The relationship between debt and growth is a politically charged issue. It is in these areas that economists must keep the most rigorous standards.' This is so hard for us, because it deals with disciplining not others but ourselves. This situation resembles the debate that followed Friedman (1953) on the naturalistic fallacy of the character of economic science being solely positive or also

601 Paraphrasing Churchill's famous saying about democracy.
602 Riedl (2004, p. 116) uses the German neologism 'Abklärung'.
603 Riedl 2004, p. 83.
604 See the debate on the 90% debt to GDP ratio (Economist 2013) on the existence of a certain harmful government debt thresholds (Reinhart and Rogoff 2010; Herndon et al. 2014), the great willingness of politicians to accept certain tendentious findings and what momentum this can create.

normative.

Geisendorf (2009, 2010) wonderfully points out the strangeness of the absorption of the idea of rationality in economics. A dominant branch of rational expectations methodology developed in the 1970s in economics and formed the foundation of the current neoclassical mainstream, which usurped a specific understanding of rationality also well described as the monopolisation of the term rationality by economics. Namely, not the discipline of economics being rational, in the sense that it uses reason in the form of logic, experiment, empirical evidence, measurability and replicability, but in defining the behaviour of elements of its subject area as being rational in the very limited sense of a myopic optimality under some constraints. A criterion which is sometimes reckless to follow as a guidance in a world with uncertainty and an open future. All that just to justify the use of concepts like complete or full information, which are a trick to simplify the mathematics needed to solve these economic problems in the models, but bear no resemblance to the real-world problem, similar to Boltzmann's trick of the ergodic hypothesis to simplify the mathematics involved in solving problems in statistical mechanics.

Although it has to be emphasised, there is nothing irrational in a scientific theory in which some or all of its actors behave boundedly rational when confronted with an uncertain open future. Hence, a theory of boundedly rational (not omniscient) or even zero-intelligence agents can be a rational scientific theory.[605]

6.3 Rhetoric of a Methodological Spillover from the Natural Sciences to Economics

This section addresses question 2: Is there a superior methodology of economics? What can be said about that methodology compared to those of history and science? More precisely, it targets two aspects. First, the methodological spillover of natural sciences methodology into economics, the so called *received view*. Second, the rhetoric of this spillover with respect to the crucial assumption of ergodicity in statistical mechanics.

6.3.1 Theory of Science in Economics – Logical Positivism

McCloskey attests that the majority of economics follows a

> "credo of Scientific Method, known mockingly among its many critics as the Received View, [that] is an amalgam of logical positivism, behaviorism, operationalism, and the hypothetico-deductive model of science. Its leading idea is that all

605 Gode and Sunder 1993.

> sure knowledge is modeled on the early 20th century's understanding of certain pieces of 19th century physics."[606]

A careful division of the relationship of logical positivism and the deductive-nomological (DN) model of science seems indicated.[607] McCloskey criticises the naïve believe in this amalgam as the exclusive scientific methodology of economics, that can be labeled *logical positivism* or *modernist methodology*. Logical positivism operates along the lines of empirically falsifiable experiences about worldly phenomena. Favouring reason and science over religion and authority started out during and after the age of Enlightenment from the work of Galilei, Descartes and Kant. In logical positivism the DN model is the preferred methodology. The DN is also known as the Hempel-Oppenheim model[608] and consists of a set of a priori laws or axioms (*nomos*) from which in a deductive manner further conclusions are drawn or, more precisely, implications are derived. From a philosophy of science standpoint the DN model is therefore the approach to follow for an axiomatisation of a field outside of mathematics. The axiomatisation of mathematics culminated in the attempts of logicians and mathematicians[609] to base mathematics on, what was perceived as the purest form of reason namely, pure deductive logic by preserving consistency. Thereby following the second problem on *Die Widerspruchslosigkeit der arithmetischen Axiome* or *the consistency of the acxioms of arithmetics* among the so-called Hilbert programme.[610] The results in Gödel (1931a,b) brought the work on the second problem to an end with a negative result. The philosophy of science most closely associated with the DN is logical positivism and was prominently advertised in the 1950s by Friedman (1953). However, logical positivism was put aside by philosophers of science already in the 1930s. We can ask which epistemological standpoint was not put aside from philosophers of science. However the case of logical positivism is especially interesting for the purpose of this study, because it was the favoured epistemology in the natural sciences and physics at the time when first cross-fertilisation between ergodic theory and economics could have started. We discuss briefly why this did not happen in the post world war II period and not for a long time after it in the next section.

606 McCloskey 1983, p. 484.
607 The DN model is used in Popper (1935).
608 Hempel and Oppenheim 1948.
609 Whitehead and Russell 1910, 1927a,b.
610 Hilbert 1900.

6.3.2 Gödel or the Mortal Blow to Logical Positivism and Hilbert's Programme

> In einem Fach, das mit einem derart schwierigen
> Gegenstand befasst ist wie die Ökonomik,
> ist ein Zustand der Ungewissheit nicht sehr angenehm,
> ein Zustand der Gewissheit aber ist bloß lächerlich.[611]
>
> Heinz D. Kurz (paraphrasing Voltaire)

Philosophers close to or even members of the Vienna Circle pushed the logical positivism movement further, among others Rudolf Carnap, Karl R. Popper, Moritz Schlick, and Ludwig Wittgenstein.[612] In the social sciences, logical positivism was incorporated by the founding father of sociology Auguste Comte[613] and across this way became the eminent epistemology for economics together with the attempt to its formalisation and axiomatisation. But through the work Gödel and his famous undecidability and incompleteness theorems this movement ended tragically. The only certainty left is the one about fundamental uncertainty.

Presenting Gödel and his undecidability theorems as witness to argue that all scientific inquiry which follows a deductive approach like neoclassical economics, is eventually doomed to fail, is always an extreme position if not a thought-terminating cliché, because it argues with the ultimate consistency. Of course, this does not imply, that there are not statements that can be proved or refuted. To be distinct, there are provable statements and consistent theories in economics. For economics this discussion in mathematics is relevant, whenever the attempt is undertaken to build large theories that try to retain consistency. The attempt to model an economy formally and deduce all statements from a priori laws and axioms seems thus at least to be prone to this undecidability issues. But half way to the ultimate there may well be decidable and (dis)provable statements, which go without contradiction with the rest of the theory.

At the core of mathematical economics lie attempts of axiomatisation such as the von Neumann and Morgenstern axioms of EUT or the Arrow-Debreu general competitive equilibrium view of the economic system. With such axiomatisation comes some intricate epistemological and ontological problems of the relevance of mathematical objects for empirical sciences. An axiomatisation tells us a lot about the assumptions needed that can give rise to

611 Kurz 2014, Preface, p. 5.

612 The Vienna circle pushed it so far, that some authors labeled this methodology as ultraempiricism. This particular and exclusive understanding of the substance of science fueled much of the debates in philosophy of science and scientific theory in the 20th century.

613 In connection to Comte this epistemology is often called logical empiricism.

a successful axiomatised theory, but not necessarily something about a real-world phenomena. A good example are existence theorems in mathematics. An existence theorem is mostly informative on the set of assumptions or axioms needed for the proof. It is not at all revelatory about the existence of this very object in reality and this is also not the aim of an existence theorem. Indeed, the ergodic theorems serve a perfect example. Mathematicians came up with first proofs of ergodic theorems, but from a physical point of view it was clear that the trajectories of particles do not behave according to the (quasi-)ergodic hypothesis. Even if there had been theorems about the principle constructability of ergodicity in mathematics, it does not imply in any way that these theorems describe a factual truth. In an economics context let us discuss the 'Existence of an Equilibrium for a Competitive Economy', which is the title of a well-known paper that gave rise to the Arrow-Debreu-McKenzie model of general competitive equilibrium.[614] Economics severely misinterpreted these results. The novelty this paper brought about lies in a careful study of the set of assumptions needed to prove the existence of a general equilibrium for a competitive economy on all of its partial markets simultaneously. Now, as economics wants to pursue scientifically, economists have to digest this purely mathematical statement and confront it with real-world economies. Sadly, the proof of the existence of this object, the general equilibrium in some Platonic heaven – that is largely detached from the reality of real markets – was used to justify the validity of the whole research programme of general equilibrium theory although actually only mathematical consistency is what has been justified. Because some equilibrium seemed to not be impossible in a purely mathematical sense, it was thought that is has to be out there in the real world somewhere and that we can reach it, if we just set the parameters accordingly (or leave the economic system untouched) and it is worthwhile trying to arrive at this divine situation. Such scientific attempts are sooner or later prone to the same fate that Gödel's results caused for Hilbert's second problem. As it was shown by the so-called Sonnenschein-Mantel-Debreu theorems the set of assumptions is almost arbitrarily large in terms of the functional form of the demand curves and individual utility functions needed to prove this theorem.[615] Which is not a sign of the universality of the result but instead of the dead end this research programme arrived at, which we already discussed in Subsec. 5.2.3.[616] Very much in the same sense did the successful axiomatisation of EUT by von Neumann & Morgenstern stop the questioning of the validity of this hypothetical model of decision making. Largely due to its internal consistency which was maybe even mistaken for beauty. Surely it is a kind of beauty and harmony coming from a theory that is internally consistent, but it can also cover up the lack of relatedness to reality. Hossenfelder (2018) finds similar issues in certain parts of physics today.[617]

614 Arrow and Debreu 1954; McKenzie 1959.
615 Sonnenschein 1972, 1973; Mantel 1974; Debreu 1974.
616 Kirman 2006.
617 Hossenfelder 2018, for the relation between the problem with beauty and mathematics in physics and economics see especially ch. 10.

This false belief in the definiteness of any one possible axiomatisation together with grand admiration for such mathematical work in economics[618] stifled new research avenues that deviate from the respective set of axioms or deviate even more in the general tool set. Even in the behavioural economics community the validity of EUT is largely accepted. In fact this leads to a state of cognitive dissonance in the community, because their research largely deals with evoking violations of the EUT axioms in lab experiments. This is justified with the help of the creation of a false dichotomy between the normative domain, where EUT remains more or less unquestioned and the descriptive domain of a decision theory.[619]

The axiomatisation idea is a key of the Bourbaki tradition in mathematics. Particularly through people like Gerard Debreu a very specific conception of mathematical beauty represented by the Bourbaki tradition was injected into mathematical economics.[620] This explains why many of the mathematicians entering economics between 1940-1990 had a background in topology.[621] This caused together with other influences[622] an increasing influx of physicists to economics in the subsequent period of roughly the 1970s until today, where they are largely united under the roof of complexity science. As the thesis seen in its totality proves, *Ergodicity Economics* is not against the use of mathematics in economics. in fact much to the contrary, but we have to find mathematical structures which correspond to the real-world phenomena and not rely solely on those which generate publishable *results*. Thus the question is about the right choice of which mathematics to utilise and not whether to use mathematics at all. In the end, the great strength of a mathematical approach is that we know where and how to probe the theory. But science has to do it.

Finally, McCloskey remarks 'strict logical positivism is dead',[623] which leaves a methodological vacancy behind, that he sees filled by postmodernism. Basic critique towards logical positivism are its obsessive hostility to metaphysics, which turns out to be metaphysically itself. The basic question arising is: Can mathematics really be perceived as a symbolic counterpart of the universe, if there are so many important phenomena, which up to now flee a satisfying every mathematical description? Of course, this question is to big to be answered in this thesis. Thus we close this section on the limits of logical positivism with the following quote, which summarises our view point on this issue:

618 Von Neumann and Morgenstern and Arrow and Debreu are in part meritedly seen as grand heroes in economics. As it is often the case, their followers may be more rigid than the originators themselves, as can be seen by the continuing support by Kenneth Arrow to establish the conception of the economy as an evolving complex adaptive system originally pursued by the Santa Fe Institute. This shines a light on the importance of the sociology of science.

619 See for example the Kahneman and Tversky research programme and especially the domain of prospect theory as a purely *descriptive* theory. We discuss this issue in greater detail in chapter 7.

620 Debreu 1991.

621 E.g. Gerard Debreu, Herbert Scarf, Steven Smale, Albert Tucker, etc.. See also the Subsec. 5.2.3 general equilibrium economics.

622 E.g. less funding of the space programmes in the US (Derman 2004) and the availability of increasing amounts of data, which resembles the data deluge physicists had to deal with already in particle physics.

623 McCloskey 1983, p. 486.

> "To a first approximation the method of science is 'find an explanation and test it thoroughly', while modern core mathematics is 'find an explanation without rule violations'. The criteria for validity are radically different: science depends on comparison with external reality, while mathematics is internal.
> The modern view is that Hilbert's proposal – that mathematical deduction might be a general prototype for science – is a failure."[624]

Methodological Remarks by Hayek

Von Hayek (1994, p. 58) sees the natural sciences prevailing over the social sciences due to a circular argument, that goes like this: Natural science dresses their findings in simple formulae, because the natural sciences define their research subject via the feasibility of dressing it in simple mathematical formula.

> "There is, however, no justification for the belief that it must always be possible to discover such simple regularities and that physics is more advanced because it has succeeded in doing this while other sciences have not yet done so. It is rather the other way round: physics has succeeded because it deals with phenomena which, in our sense, are simple."

In this sense, everything that is describable in simple mathematical terms is physics, everything else is not. This is a recursive definition of the subject of a field, contingent on its successful application. Blaming other disciplines, why they not succeed in finding simple mathematical regularities is actually not reasonable, if one follows Hayek's chain of thought. Such a statement is missing the point of the necessity of pluralism in method and model choice, if the research subject requires it, a point prominently stressed by Keynes[625] and Hayek.[626] This view explains why the mind and consciousness are not seen as a physics problem, because there are no convincing ways to describe the mind and consciousness in a mathematical formal statement (yet). Whereas, the pendulum is simple enough for human brains, to easily find mathematical expressions of its behaviour, therefore the pendulum is seen as a physics problem. Additionally, von Hayek (1989, p. 3) points out, that the success in physics rests on the stable existence of observables, whereas in economics and social sciences a variable enters economic theories and often becomes treated as if it were an (stable) observable[627], 'which happens to be accessible to measurement.' *Ergodicity Economics* explicitly deals with finding proper observables, thus to work with proper observables is in no way limited to physics. Finding the right observable is crucial and may be a hard thing to do. This again is

624 Quinn 2012, pp. 31, 35.
625 Keynes 1924, p. 322.
626 Von Hayek 1994.
627 See also Marx (2018) on the stability of observables.

a case, where a method is used reversely, but the reverse must not always be true as well, as a reversed logical implication does not work in opposite direction with changed direction of causality. For example, the mass of a solid body is an observables, and so mass is a potential element for physical theories. The German stock index (DAX), Dow Jones Industrial Average or any other stock indices are not an observable in the original sense, because the DAX on January, 1^{st} in 1988 consists not only of different stocks than the DAX on January, 1^{st} in 2015, most of the companies of these days do not even exist anymore. Furthermore, the meaning the whole system, any central bank or an arbitrary individual attaches to the DAX is likely to have completely changed in totally different ways during this time and is therefore likely to be different at arbitrary moments in time. The same holds for GDP, unemployment rate, money aggregates, CPI and many if not most economic variables. All these quantities are human artefacts. Humans attach meaning to it for pragmatic reasons and these reasons can change (for legitimate causes). For example, how to quantify what is understood at a time as a useful measure of unemployment changed tremendously in the 20^{th} century, owing a lot to economic debates regarding structural/voluntary unemployment following different ideas or theories of different economic thinkers like Marx, Keynes or neoclassical conceptions. This makes obvious, how problematic it is, to interpret a time series of say unemployment rates or CPI over long periods.

Hayek is also insightful in reflecting on how a virtue may become a vice in thinking beyond the Popperian principle of falsification, if the scientist is confronted with a complex phenomena:

> "The advance of science will thus have to proceed in two different directions: while it is certainly desirable to make our theories as falsifiable as possible, we must also push forward into fields where, as we advance, the degree of falsifiability necessarily decreases. This is the price we have to pay for an advance into the field of complex phenomena."[628]

This argument resembles Feyerabend (1976) and his response to Lakatos's and Popper's sometimes dictatorial decree on methodology. Nevertheless falsifiability and replicability are important for science, and can only be temporarily abstained from and not principally. For instance, the Higgs boson was conjectured decades before an experimental verification became feasible and could not refute its detection. Never should the falsification be impossible in principle like it is the case with utility, which is an ingredient in economic theories impossible to observe.

628 Von Hayek 1994, p. 58.

6.4 Non-Ergodicity and the Notion of Time

Equilibrium in economics is an empty concept.
It means, the model has a solution.[629]

STEVEN DURLAUF

Ramification of the Unconscious Assumption of Ergodicity

If we shorten the formal aspects given in Sec. 2.4 and the ergodicity problem to motivate the content of this section, then we state that the average behaviour of a system over all possible or conceivable systems (ensemble) does not necessarily correspond to a representative behaviour of one realisation of such a system over time – the finding of non-ergodic dynamics. This leads us to the question whether it is possible to arrive at truths in economics that reside outside of time? What could such timeless truths be? In physics there are the so-called natural laws.[630] Standard economics postulates the existence of such a timeless truth in form of a unique, invariant stationary equilibrium simultaneous on all markets. Because the majority of mathematical economics involves such an equilibrium, we coin this as *ergodic economic theory* or *ergodic economics*.

Hence, ergodicity allows to infer the system behaviour from one time series (observed trajectory in physical terms). Hence, following the conceptual understanding of ergodic processes, there is no need to study the history of a process, because there is no sensitivity to initial conditions, 'in an ergodic system time is irrelevant and has no direction'.[631] The elimination of time is what is most often swept under the carpet of equilibrium analysis, e.g. the speed of convergence, if an economic equilibrium is defined via the physical metaphor of an equilibrating mechanical rest point. In this wording lies some confusing potential, as is also noticed by DURLAUF's section quote and by NIEHANS (1997, p. 58), '[i]n der Nationalökonomie wird die Lösung eines solchen [Gleichungs]Systems oft ein »Gleichgewicht« genannt, obgleich die physikalischen oder normativen Konnotationen dieser Bezeichnung durchaus irreführend sein können.'

Even in cases where the focus on equilibrium simplifies an analysis a lot, the price that has to be paid for this simplification is the elimination of time. A careful analysis whether the equilibrium feature could also be a bug, is seldom carried out. Hence, the mathematisation in

629 Introductory statement by STEVEN DURLAUF to his talk at the 1st International Conference on 'Cliometrics and Complexity' in Lyon in June 2016.

630 E.g. law of gravity. Since its discovery by NEWTON its validity is assumed at all times – in the past and in the future – and at all places – on planet Mars as well as on planet Earth.

631 PETERS and MAUBOUSSIN 2012.

ergodic economics is so often solely seen as a triumph of science and an achievement per se, thereby leaving an important part, namely time, unconsidered.

6.4.1 Past Performance is not an Indicator of Future Results

Investment opportunities offered by commercial banks usually come with the following qualifier in the fine print: 'Past performance is not an indicator of future results'. In the following we substantiate the role of time under ergodicity and non-ergodicity using common terminology in the philosophy of time[632] and a recent line of research on temporal naturalism.[633] The key term is the *present moment* or the *Now*, as the current point in time which separates the past from the future. The present moment has different properties in the understanding of time in the ergodic case and the non-ergodic case, which we discuss in the following.

Analytical vs Historical Time

The true nature of time is an eternal mystery to mankind with many contradicting views. For Heraklit change was the essence of the universe, and his *pantha rhei* was cast in the familiar idiom 'No man ever steps in the same river twice'. A famous representative of an opposing belief was Albert Einstein. He thought of time as being only an illusion, expressed best in his own words: 'Für uns gläubige Physiker hat die Scheidung zwischen Vergangenheit, Gegenwart und Zukunft nur die Bedeutung einer wenn auch hartnäckigen Illusion.'[634] This opinion led Einstein into a serious philosophical debate with the French philosopher Henri Bergson, who was very influential in the early 20$^{\text{th}}$ century. We can categorise Bergson's thinking into process philosophy, which we mentioned briefly on page 215. For him the processuality is an essential qualia of the universe and life has to be understood as a process not as a state. Also Wiener was lifelong follower of Bergsonian conception of time.[635]

Figure 6.2 reveals the difference between an analytical understanding of time in the ergodic case and an historical understanding of time in the non-ergodic case. Our visualisation of the role of time utilises the work by Davidson which we reviewed in Subsec. 5.2.5.3.[636]

The left part of **Fig. 6.2** depicts the ergodic case. Here, the present moment is merely an arbitrary dividing line in time between an otherwise homogeneous medium, indicated by the

632 McTaggart 1908; von Weizsäcker 1939.
633 Smolin 2009, 2014, 2013, 2015; Mangabeira Unger and Smolin 2015.
634 Einstein wrote these lines in a letter of condolence to the family of his friend and physicist Michele Besso. We doubt that these lines gave much comfort to the family, nevertheless they express a prominent view on time.
635 Masani 1990, Ch. 12 F.
636 Davidson 1991, 2009a.

straight vertical line on top of the present moment. Homogeneous in the sense that there is nothing which distinguishes the past from the future, e.g. the probability distribution of a certain random variable is stationary and therefore the same before, during and after the present moment. Davidson (2009a, p. 328) famously described this situation as 'the future is merely the statistical shadow of the past', which is why instead of the word *Future* there is just a mirror image of *Past* depicted. In Davidson's thinking probabilities can be meaningfully applied under ergodicity, because what is ultimately needed is to draw samples from the future distribution, but because the future and the past are in fact identical drawing samples from the past to infer future behaviour of e.g. prices is valid. The concept of time used in the ergodic case is often referred to analytical or logical time, because it is just something like an index or a counter and nothing more. To be more precise, in the context of ergodic theory the ensemble average and time average coincide, i.e. the index to average over the ensemble (we used for example n) and the index over time (denoted by τ) are only a different labels for the same object. Time in this conception is fully reversible and can be thought of as an illusion. We have first visualised this in Kirstein (2016), recently this has been dubbed the 'Empiricist mirror thesis' in Nasir and Morgan (2018).

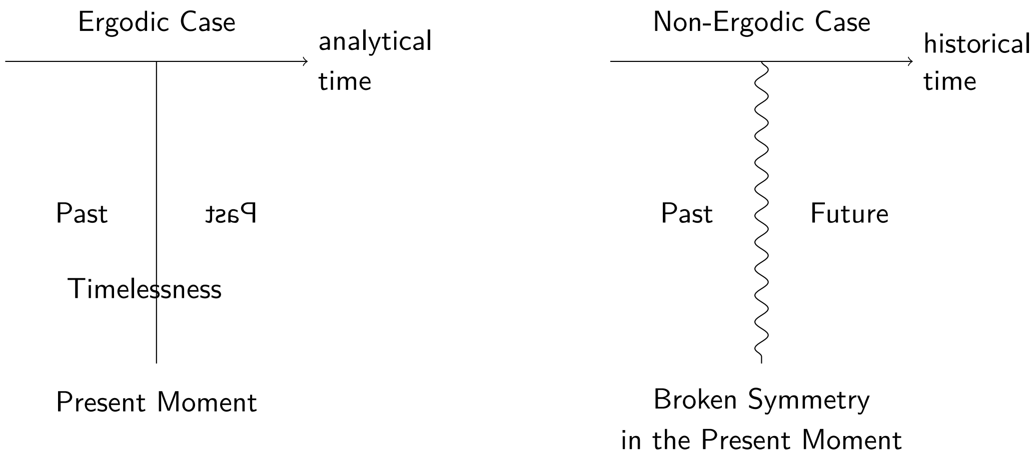


Figure 6.2: Mirror Hypothesis of Ergodicity. Left: Time in the ergodic case. Probabilities (drawn from the past) are meaningful for the future 'the future is merely the statistical shadow of the past.' Right: Time in the non-ergodic case. As probabilities can not be drawn from the future, they are 'meaningless' for the future according to Davidson.

Quite the contrary in the non-ergodic case on the right of **Fig. 6.2**. In the non-ergodic case the present moment carries the potential of a symmetry break, that effectively divides the past from the future and makes the two distinguishable, indicated by the snaky vertical line at the present moment. This broken symmetry of the system or the process occurs in such a way, that the future is clearly different from the past. Hence, the *Future* is not anymore just a mirror image of the *Past*, but owns unique new characteristics. The concept of time in the non-ergodic case is therefore historical or real time and matches the time concept we experience in our lifes. Often this is dubbed Bergsonian time. Therefore the mirror function is lost in the figure. We can think of a probability distribution of a certain random variable that may be different after the present moment than it has been before, e.g. a non-stationary distribution of a random variable due to a time dependence in their parameters. Berman et al. (2016, 2017) performed an empirical test of the ergodic hypothesis for US data and get exactly such a finding of an ever widening wealth distribution, which implies a broken ergodicity in the reallocating geometric Brownian motion (RGBM) model for the relevant realistic parameter regime.[637]

The broken symmetry described here is quite similar to certain theories about physical phenomena in condensed matter physics like superconduction or the Higgs particle in particle physics. Both are explained by spontaneous symmetry breaking (SSB).[638] Additionally, Palmer (1989, pp. 275–276) remarks that 'broken ergodicity' generalises 'broken symmetries', which is why 'non-ergodicity is an essential ingredient of broken symmetry, but is in fact much more general.' Furthermore, a connection between SSB and non-ergodic systems is given by Liu (2003, p. 599):

> "few systems, if any real ones, are proven ergodic and some are proven non-ergodic, and even for those proven ones, how does ergodicity gives [sic!] us any reason to believe the probability it gives for finding a single outcome at a certain moment? [. . .] SSB systems are obviously non-ergodic."

6.4.2 Ergodic Fallacy

Now all the terminology is introduced to understand the ergodic fallacy. The ergodic fallacy is comprised of the misbelief in a causal relationship, when there is none. The causal relationship is either not existent at all or could constantly change, which would put the kind of relationship closer to a contingent one than a causal.[639] The latter case proves most interesting, as it leads to questions like 'What is causality? Can causality change? If causality can change, is it then the same causality or something different after it changed?', and finally, 'Does causality exist at all? What is its ontological status?' Broadly speaking, the ergodic fallacy is the search for timeless (natural) laws if there can not be such laws. However, to answer questions on causality is beyond what we want to attempt. We focus on the ergodic fallacy in more detail.

---

637 For time-dependent distribution see also Peters and Klein (2013) for broken ergodicity in general GBM or the random acceleration model with $\mathcal{N}\left(\langle X\rangle\, t, \frac{3}{2}\,\mathrm{Var}\,[X]\,\sqrt{t}\right)$. Similarly, time-dependent variance is also studied in volatility models in finance using time-varying Gaussian distributions.

638 See especially Liu (2003). See also Goldenfeld (2005, Ch. 2.9 on SSB and Ch. 2.10 on Ergodicity Breaking) and on a discussion that when time reversal symmetry gets broken, ergodicity breaks as well.

639 See Lehmann-Waffenschmidt (2010) for a way to measure causality for contingent historical processes in a gradual manner.

To elucidate the ergodic fallacy, we explain its meaning with the help of **Fig. 6.3**. First, it embedded in an environment where time is real, thus the concept of historical time. In such an environment, we can imagine a scientist observing a process in a finite time window, the observation window (1). The scientist has no knowledge or data over any time periods before that window or is simply not interested in these time periods, so everything before that observation window is just ‘not analysed’ as indicated in the far left of the figure. A scientist falling prey to the ergodic fallacy, now believes it is possible to inductively infer a ‘(spurious) causal relationship’ from his observed time period, which is indicated by step (1). Furthermore, he projects the relationship into the future.

Naturally, the present moment marks the end of his observation period and carries the potential of a symmetry break.[640] If in a present moment a symmetry break occurs, the scientist is confronted with a situation, in which there is one causal relationship before the present moment and a different one after it. This is called a ‘non-causal window’ and indicated step (2) in **Fig. 6.3**. Non-causal is here understood in the sense of not mono-causal. It is possible that the occurrence of broken symmetry is not recognised immediately after it happened and therefore the causal relationship seems to continue for a while, indicated by the partial mirror image of the word ‘Ergodicity’ after the first symmetry break. Eventually, the causal relationship breaks down and instead of the reflection of the ergodic process ‘something new’ is realising and destroys the mirror image character of the future seen from the present moment or the past, indicated by step (3). Of course, there can be successive symmetry breaks, creating successive ‘non-causal windows’ indicated by step (4). Furthermore, a time window which spans one or more symmetry breaks is a non-causal window, non-causal in the sense of changing causal relationships.[641]

6.4.3 Potential Drivers of Broken Symmetry

The exposition of the ergodic fallacy leads us to the interesting question of what are the drivers of symmetry breaks in general and in economics specifically. The first candidate that comes into an economist’s mind is of course innovation.

Innovations

As we indicate in Subsec. 5.2.4, innovation economics is a rich source of identifying the conditions which determine the diffusion of innovations once they occurred and how innovations change the structures on the consumer and producer side of the market. E.g. Witt (2003,

[640] In the field of IKE they refer to what are symmetry breaks in our terminology as non-routine change, but intend similarly unpredictable change (Frydman and Goldberg 2007, 2011).

[641] See also Petersen (1996) and the relevance of ergodic theorems for empirical scientists confronted with the role of time and the measurement process.

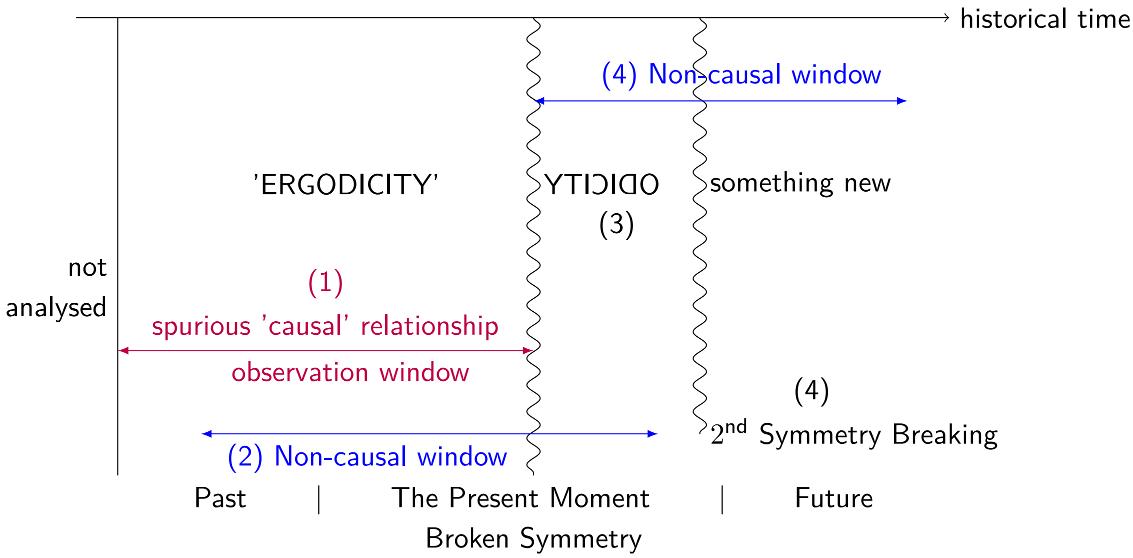


Figure 6.3: Ergodic Fallacy or Causality Fallacy.

p. 13) refers to the 'dissemination or post revelation analysis'. Innovation economics offers fewer insights on what drives these innovations other than entrepreneurial creativity. Therefore we concentrate on a more fundamental driver of symmetry breaks than entrepreneurial innovations, namely, the process of observation itself.

Imagine if science itself, the pure act of cognition, is changing the system in unforeseeable ways? This creates method of science, which destroys the stability of its research subject. In this case, scientific observations itself inhibit what can be discovered by them. Therefore, some epistemologies concentrate on the discovery of ultimate truths, that reside outside of time like a mathematical object (triangle, line, point, etc.) or a natural law. There are several insights related to the observation-perturbation aspect we just mentioned and possible drivers of broken symmetries. Some of them have been made already decades ago, but basically share the identical core of an evolving system that changes its inner causal structure in unforeseeable non-deterministic ways. Among them are Taleb's black swan argument, the Lucas critique, Soros' reflexivity, Goodhart's law or Wiener's strong coupling.,[642] which we briefly review now.

Taleb's Black Swan Argument

Other possible drivers of symmetry breaks are the so called black swans which were initially discussed in Taleb (2005) and gained popularity after the recent financial crisis was seen by many as a manifestation of black swan events. However, we think this misappropriates

642 Lucas 1976; Soros 2013; Goodhart 1975; Wiener 1985.

TALEB's original argument.[643] In our opinion TALEB's metaphor addresses the limits inherent in the model building process which is what science basically does. In the process of building a model of a real-world phenomena, the researcher has to abstract from certain aspects of reality otherwise he would build a one-to-one map of the territory which is not helpful.[644] Sooner or later it will be exactly these aspect of reality that were abstracted from, which will become relevant and invalidate the relationships in the model. As a consequence the symmetry the model is based on breaks down. TALEB compares the behaviour of the scientist to that of a turkey. Every morning and every evening the turkey is fed for 99 days. Thus he extrapolates that also on day 100 he will be fed in the morning. Only to realise that the symmetry breaks on day 100 and he is the next meal.

Lucas Critique

As one the founder of REH, ROBERT E. LUCAS is also the originator of the so called LUCAS critique. It states that an economic measure loses its power once it is publicly known that it is used e.g. by the treasury department or tax authorities for regulatory issues. Thus the use of a measure is destroying the causal relationship on which its effectiveness relies on.[645] This critique emphasises the importance of internally caused changes in the expectation formation and the change this causes in certain parameters in econometric models. However, preferences and production technologies are often considered as exogenously given in such models, even if they are used for forecasting purposes when the forecast is influencing the outcome. The public release of the use a certain measure can be interpreted as a symmetry break. Compare these effects with the use of impact factors to measure the performance and capabilities in science. It is no surprise that the behaviour with regard to the choice of journals and topics is influenced once their use in appointment procedures is publicly released. What does this say about the information content of impact factors?

Goodhart's Law

Closely related to the LUCAS critique – made in the context of macroeconometric models – is GOODHART's Law in the context of banking regulation, 'Any statistical relationship will break down when used for policy purposes'.[646] Regulation of the banking sector loses its steering ability once it focuses on single balance sheet items. As soon as banks know they are judged according to their equity, it is reasonable to make the necessary shifting between balance sheet items, maybe even create off balance sheet vehicles. Transparent regulatory rules therefore

643 TALEB 2010a,b.
644 KORZYBSKI 1933, p. 750.
645 LUCAS 1976.
646 GOODHART 1975.

create a self-destroying prophecy and sometimes even worse implement pro-cyclical measures. E.g. Daníelsson et al. (2001) and Daníelsson (2002, 2008) examined the self-reinforcing aspects in the Basel II and III framework. Especially they find no stable relation between the compulsory risk models neither in times of crisis nor in quieter times. Hence, Daníelsson rephrases Goodhart's law as follows, 'risk model breaks down when used for regulatory purposes'.[647] Again the public release of the use of certain measures leads to symmetry break and the disappearance of a causal relationship.

Soros Reflexivity

The book Soros (1994) by Soros popularised the reflexive feedback loop between economic theories and their research object, which can be understood with the framework of the ergodic fallacy. The origin is a descriptive or positivistic theory of financial markets. The use of this theory in practice, because it is interpreted to be also a normative theory, makes a reformulation of the theory necessary in order to incorporate the effect that some agents in the model now use the theory, which starts an infinite regress of the type of the Morgenstern paradox.[648] Such self-enforcing feedback loops between economics and the economy by no means need to equilibrate. Some of the topics which involve reflexivity are also discussed under the topic of performativity.[649]

Wiener's Strong Coupling

> It is in the social sciences that the coupling between the observed phenomenon and the observer is hardest to minimize.[650]
>
> Norbert Wiener

Roughly 40 years before the word reflexivity become fashionable, cybernetics developed as a scientific discipline of feedback phenomena Or control and communication in the animal and the machine initiated by Norbert Wiener.[651] When asked about the applicability of the cybernetic ideas in the realm of the social sciences and economic planning Wiener emphasised time and again the greater level of the coupling between the researcher and

[647] Daníelsson 2002, p. 1276.
[648] Morgenstern 1935; Lehmann-Waffenschmidt 1990.
[649] Callon 1998; MacKenzie 2003; Hirte 2010, 2014.
[650] Wiener 1985, p. 163.
[651] Wiener 1985.

the research object especially in economics.[652] This increased level of coupling impedes any trivial transfer of insights from simple cybernetic systems like thermostats to much more purposeful actors on financial markets and in the economy.[653] At many places WIENER addresses ergodicity in his *Cybernetics*, but in the following quote the relation is more implicitly given:

> "In other words, in the social sciences we have to deal with short statistical runs, nor can we be sure that a considerable part of what we observe is not an artifact of our own creation. An investigation of the stock market is likely to upset the stock market. We are too much in tune with the objects of our investigation to be good probes. In short, whether our investigations in the social sciences be statistical or dynamic—and they should participate in the nature of both— they can never be good to more than a very few decimal places, and, in short, can never furnish us with a quantity of verifiable, significant information which begins to compare with that which we have learned to expect in the natural sciences. We cannot afford to neglect them; neither should we build exaggerated expectations of their possibilities. There is much which we must leave, whether we like it or not, to the un-'scientific', narrative method of the professional historian."[654]

WIENER makes a deeper investigation of ergodicity, the difference between analytical and historical time, the distinction between past, present an future in the context of time series analysis in *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series.*[655] In this publication WIENER makes the important statements for the purpose of this thesis, e.g. with regard properties of the relevant mathematical operators:

- 'The use of the complex plane is intimately allied to the fact that physically applicable operators of engineering allow us to work with the past of our data, but not with their future',[656]

- 'We have said already that, while the past of a time series is accessible for examination, its future is not. That means that our operators must have an inherent certain one-sidedness.'[657]

Or with regard to the occurrence of novelty,

---

652 WIENER 1985, esp. Ch. VIII.
653 ROSENBLUETH and WIENER 1945.
654 WIENER 1985, p. 164.
655 Used in the prediction of air combats it was classified during the second world war which explains the delayed publication of WIENER (1964b). Due to its yellow cover it had the reputation to be referred to as the *Gelbe Gefahr* (danger in yellow) using an alliteration in German.
656 WIENER 1964b, p. 8.
657 WIENER 1964b, p. 12.

- 'The simplest operation which we can perform is that of extrapolating them or, in other words, of prediction. This prediction, of course, does not in general give a precise continuation of a time series or message, for, if there is new information to come, this completely precludes an exact estimate of the future',[658]

and with regard to the relation between ergodicity and predictability

- 'Therefore, if we take any property of a Brownian motion independent of the absolute position of this motion, we shall generate a time series having the ergodic property.'[659]

Conclusion on the Drivers of Symmetry Breaks

All the presented potential drivers of broken symmetries show what can be explained with the help of the ergodic fallacy. Furthermore, a common theme is the complex feedback between theory and research object, mediated by the researcher. The most extreme answer to the question what drives the symmetry break is: It is the process of observation and measurement itself, which on the one hand creates the symmetries in order to identify them and destroys them at the same time in an eternal loop. To put it differently, the act of doing science is what creates and breaks symmetries. Only the researching economist gives birth to and destroys the symmetry in the world.

From this section many more questions can be derived on ergodicity, causality and quantum like observer effects, in fact too many so that we better conclude with the even more fundamental question this section raised: On the appropriateness of any timeless mathematical structure to describe time-bound economic phenomena?

6.4.4 The Silent Rhetoric of Ergodic Economics

As is shown in previous chapters the ergodicity problem is of fundamental importance for economic theorising, because the focus on the ergodic case led virtually to the elimination of time from economics as a creative force. It is then important, if this elimination is hardly ever made explicit in the literature. Thus we are interested in analysing if a special rhetoric is involved in this process. McCloskey forced economics and economists to start thinking about their rhetoric, '[e]conomics should become more self-conscious about their rhetoric [...]'[660] and elucidates what we can expect of a careful investigation of rhetoric: 'unexamined metaphor is a substitute for thinking – which is a recommendation to examine the metaphors, not to attempt the impossible by banishing them.'[661] Metaphors are used by everyone, we used

658 Wiener 1964b, p. 8.
659 Wiener 1964b, p. 21.
660 McCloskey 1983, p. 482.
661 McCloskey 1983, p. 507.

one when we refer to the ensemble or time perspective as different embeddings of uncertainty either within the ensemble or within time. Their use per se is almost unavoidable and not a problem in itself. However, in this section we carve out a special rhetoric in economics, which rests on the metaphor of *the equilibrium* and involves the implicit assumption of ergodicity. We show that not only *the equilibrium* guided most of 20$^{\text{th}}$ century reasoning in economics, it is the assumption of ergodicity which facilitated this development like almost no other model property.

> "A rhetoric of economics makes plain what most economists know anyway about the richness and complexity of economic argument but will not state openly and will not examine explicitly.
> The invitation to rhetoric, however, is not an invitation to irrationality in argument. Quite the contrary. It is an invitation to leave the irrationality of an artificially narrowed range of arguments and to move to the rationality of arguing like human beings."[662]

Quite the contrary of what McCloskey describes in the above quote is the case for ergodicity in economics. After making the experience of my own instruction in economics and statistics, after having done a careful literature review and after having talked to many scholars at conferences during and after giving talks on this topic, it is not at all obvious, that many economists know, that they are implicitly assuming ergodicity in their models, although it is necessary for their models to possess a (unique equilibrium) solution. Therefore, this section analyses what contributed to the current situation of what could be called a predominance of ergodic economic models, which are based or only focus on the ergodic case in contrast to *Ergodicity Economics*, which treats the full spectrum of the ergodicity problem in economics.

The key assumptions of mainstream neoclassical economics, that are stated explicitly over and over again especially by the many critics read as follows:

- use of the representative agent (firm/household/government),
- rational expectations,
- efficient markets.

Ergodicity is clearly missing on this list, which is why it surprised us even more when it appeared on Burda's list given in the introduction. Therefore, the rhetoric of ergodic economics consists of concealing this fundamental assumption, which we coin the *silent rhetoric of ergodic economics*. However, the loaning of metaphors from physics was not always and not in every detail concealed, 'the core of neoclassical research program is a mathematical metaphor approximated from physics in the 1870s which equates potential energy to "utility", forces to

[662] McCloskey 1983, p. 509.

"prices", commodities to spatial coordinates, and kinetic energy to the budget constraint. The early generations of neoclassical economists constructed their system from such metaphors, and openly acknowledged them.'[663] Such an open acknowledgement is clearly not the case for the assumption of ergodicity. Only in more recent epistemological analyses of economics, ergodicity is recognised as central for neoclassical and today's mainstream economics.[664] For Hirte and Thieme (2013, p. 71) ergodicity even counts to the 'Grundaxiome'.

McCloskey (1983, pp. 500) finds several mechanisms of sociological influence in the process of scientific discovery[665], among others appeal to authority, appeal to relaxation of assumptions, appeal to hypothetical toy economies and appeal to analogy. This is not the first time to recall the enormous amount of influence the impressive œuvre of Paul Samuelson had on economics. A considerable amount of this influence appears to be by appeal to authority, partly to personal authority partly to authority of his mathematics. The reason for this appeal can be seen in the 'air of easy mathematical mastery',[666] which was a long-cherished desideratum at a time where all the (social) sciences were envying the queen of the sciences, viz highly mathematised physics.[667] How else then as a faithful belief in the master of (at his time) modern economics can it be explained, that there is almost no reference stating explicitly the importance of the ergodicity problem for economics except the one from Samuelson (1968) himself, who has been '[t]hroughout his career, [...] the master of scientific rhetoric, continuously and consciously hinting at parallels between neoclassical theory and twentieth century physics, and just as consciously denying them, usually in the very same article.'[668] As a further example of Samuelson's rhetoric ambivalence he often brings up statements like the following: 'Is this black magic or is it only an economist's way of observing that there is no free lunch in the form of a perpetual motion machine of the first kind? You decide.'[669]

A longer demonstration of his mastery in the conscious and continuous referral to physics and the denial of its relevance even on the very same page is:

> "I have limited tolerance for the perpetual attempts to fabricate for economics concepts of 'entropy' imported from the physical sciences or constructed by analogy to Clausius-Boltzmann magnitudes. The monthly mail still brings grandiose schemes to replace the dollar as a unit of value by energy or entropy units. Superficial knowledge of thermodynamics, brought into contact with ignorance of economics, cannot even in the presence of the catalyst of noble intentions beget

663 Mirowski 1989, p. 176.
664 See Heise (2012, p. 85) and Hirte and Thieme (2013, pp. 28, 64).
665 Which developed nowadays into a research area of its own called sociology of scientific knowledge (SSK).
666 McCloskey 1983, p. 500.
667 See also Mirowski (1989) on the misuse of physics metaphors in economics.
668 Mirowski 1989, p. 186.
669 Samuelson 1990, p. 266.

> stable equilibrium of useful products. That is not a tautology, merely a finding of fifty-five years of reading the morning mail.
> The benefit to theoretical economics, it will be shown, comes from discerning the *mathematical isomorphisms* between the maximum-minimum systems of thermodynamics and the cost-profit-utility systems of classical Walras-Debreu economics.
> [...]
> Thus, recently I tried to model a theory of the asymptotic approach of the income distribution to an ergodic state; when the very rich and poor are more risk averse than the middle classes, the system evolves into an ergodic Markov state; still it displays all the Loschmidt-Poincaré-Zermelo paradoxes of being truly time-symmetric![670]"

Any of the above mentioned findings amplified what became a self-enforcing process of the direction of the mathematisation in economics. Even more so if the underlying mathematical field is not yet fully systematised and its ramifications not entirely understood as was the case with ergodic theory in the middle of the 20th century when SAMUELSON's publications initiated an intellectual path dependency.

6.5 Conclusion

After all, we ask with MCCLOSKEY (1983, p. 513) the joint question: 'What is to be gained by taking the rhetoric of economics seriously?' And find the answer there as well: 'The question can be answered by noting the burdens imposed by an unexamined rhetoric.' What are the burdens of this unexamined silent rhetoric of ergodicity in economics? Well let us stick with MCCLOSKEY 'The mathematical and statistical tools that gave promise in the bright dawn of the 1930s and 1940s of ending economic dispute have not succeeded, because too much has been asked of them.'514 What is true for the use of mathematics and statistics in economics is also true for mathematics itself as was demonstrated by the proofs of GÖDEL. Mathematics and statistics did not lead to an ultimate answer to all economic controversies as logic did not lead to the ultimate truth. In his Nobel prize lecture the physicist EUGENE P. WIGNER defines the objective of physics as the explanation of (inanimate) nature. An explanation is for WIGNER an understanding of nature which produces no surprises. Over the course of this thesis it became clear, economics as a discipline and its a domain is fundamentally driven by surprises. The surprise by *Ergodicity Economics* is the new light in which we now see a lot of older economics puzzles.

[670] SAMUELSON 1990, p. 256.

7 Growth-Optimal Decision Theory

> The end point of rationality is to demonstrate the limits of rationality.
>
> Blaise Pascal

In this chapter we discuss what is behind the concept of rationality in economics in Sec. 7.1. The critique of the rationality model in economics has led to the emergence of bounded rationality. We clarify this concept in Sec. 7.2 and visit its origins and discuss whether behavioural economics has achieved the original aims of the founders of bounded rationality. In Sec. 7.3 we discuss the inflation of utility functions as a container of preferences and optimisation as a general organisation principle in nature. In Section 7.4 we present an overview of the several decision theories, which did not yet appear in previous chapters and derive from it the direction in the evolution of decision theories so far. We discuss especially the maximin theory in greater detail, because the underlying ensemble averaging is not obvious and it is a very important and flexible model. The insights are condensed in **Table 7.1**.

7.1 Rationality in Economics

Rationality in economics is a term that needs to be rigorously defined, otherwise it gets confused with its everyday meaning of simply 'reasonable'. Rationality in economics is a term that *is* rigorously defined. In economics the term 'rationality' determines a special type of behaviour – a behaviour that is first and foremost consistent. Consistency means that there is no contradiction between a given set of behavioural axioms. Homo œconomicus or the economic man is the prime model of rational behaviour in economics. Thereby, homo œconomicus is a fictitious character, whose behaviour is rational in a very specific manner, namely economic man is solely driven by his self-interest. Although the method of optimising a utility function alone is capable of encoding very diverse forms of rationality, homo œconomicus is overwhelmingly often depicted as having only self-regarding preferences and in the majority of cases even only pecuniary goals. Albeit the so called model of MaxU (an abstract utility maximising agent and a representative of the species homines œconomici) is by no means restricted to self-regarding behaviour. In fact, MaxU can obey exclusively other-regarding preferences, too, without any change to the underlying theoretical conception. Simply the self-regarding ingredients of the utility function are replaced by other-regarding ingredients.

A problem arises if the flexibility inherent in the theoretical model is exploited in the following way:

- observe empirical behaviour,
- find some preferences under which the observed behaviour becomes optimal,
- claim that the so inflated utility function is *the* rationality of economic men.

This is equivalent to a fitting exercise, where the parameters that are adjusted and *estimated* are the preferences in this example. The problem with mere fitting exercises is, they do not explain anything, because they are flexible enough to encompass almost everything. For instance, in a moment we consider rational models of addiction and we see that there is simply nothing left, such a theoretical model could not rationalise. The problem is when an ex-post rationalisation is mistaken for an explanation.

Usually one camp, the pro-traditional-economics camp, sees no problem in holding on to a homo œconomicus with self-regarding preferences – the classical MaxU. The opposing camp of critics all to often wants to replace this MaxU with just another version of MaxU, one with other-regarding preferences. It turns out that the whole debate that criticises the prevailing homo œconomicus world view of traditional economics is misdirected. Much of the confusion stems from an interpretation of the self-interested baker and butcher of Adam Smith's *An Inquiry into the Nature and Causes of the Wealth of Nations*[671] which happens to be out of context and the context is given by Smith's *first book The Theory of Moral Sentiments*.

Basic research in normative decision theory deals with the identification of a minimal set of axioms, which bring about the consistency of the behaviour and secure a smooth flow of the optimisation procedure. Von Neumann and Morgenstern (1955) proved that the following four axioms are necessary and sufficient for an axiomatisation of EUT:

1. the completeness axiom,
2. the transitivity of preferences,
3. the axiom of continuity and
4. the independence axiom.

671 Smith 1904, Ch. 2, p. 16.

An especially compelling axiom is the transitivity of preferences for goods, consumption bundles or gambles. If this decision-maker has intransitive preferences, one could in principle construct an almost effortless business model, which is called an arbitrage opportunity in economics. One could in principle construct a money pump,[672] exploit the arbitrage opportunity and create the possibility for unlimited profits. The money pump in economics can be understood in analogy to a perpetuum mobile. Whereas the perpetuum mobile is a fictitious type of motion that continues without external energy supply. The money pump can be constructed in reality as soon as one of the axioms is violated.

If we imagine a decision-maker with intransitive preferences, he could have a preference table, which contains binary preference relationship, like the following. *I.e.* he strictly prefers a pear to an apple, an apple to a banana, and a banana to a pear. Formally, intransitive preferences are denoted as:

$$\text{pear} \succ \text{apple}, \tag{7.1}$$

$$\text{apple} \succ \text{banana}, \tag{7.2}$$

$$\text{banana} \succ \text{pear}. \tag{7.3}$$

However, if the decision-maker owns one of these fruits an arbitrageur could exploit his preferences indefinitely in the following way. The arbitrageur uses the preference table given by the equations (7.1-7.3) to offer a sequence of trades each for a small fee. If we assume the decision-maker is endowed with an apple to begin with, the arbitrageur chooses the fruit where the apple appears on the righthand side (eq. (7.1)), and offers a pear. Now the decision-maker owns a pear, the arbitrageur uses the preference table and offers again a trade, this time of a banana (eq. (7.3)). Now the decision-maker owns a banana, the arbitrageur uses again the preference table and offers the apple (eq. (7.2)), which closes the first cycle. Comprising, after one cycle of the money pump the decision-maker with the intransitive preferences is in the same state as in the beginning, but the arbitrageur earned three times the small fee.

Be transitivity as reasonable as it may, it is also not hard to imagine that people violate transitivity or other mathematical properties that ensure the consistency. Especially if the goods bundles or lotteries to choose from, increase in complexity, people may quickly violate transitivity empirically. Although consistency is plausible, it is easy to show in experiments and in real life violations of any of the axioms of rationality, where it comprises maximisation of expected utility, Bayesian updating, transitivity, dynamic consistency, *etc.* As anecdotal evidence let us quote Stanislaw Ulam, a superb Polish pure mathematician working among others on the most abstract fields of logic and set theory, who confessed in his autobiography:

672 Cubitt and Sugden 2001.

> "[A]t the age of eight or nine I tried to rate the fruits I liked in an order of 'goodness' [...] until I discovered to my consternation that the relation was not transitive – namely, plums could be better than nuts which were better than apples, but apples were better than plums. I had fallen into a vicious circle [of the money pump], and this perplexed me at that age."[673]

In more systematic manner SELTEN (2002, p. 15) finds empirical violations of all the basic ingredients of rationality, 'people do not obey Bayes's rule, their probability judgments fail to satisfy basic requirements like monotonicity with respect to set inclusion, and they do not have consistent preferences, even in situations involving no risk or uncertainty.' If even our best mathematicians do not behave according to the very restrictive rationality criteria, in sounds reasonable for ARKES et al. (2016) to ask 'How bad is incoherence?', if it doesn't hurt too much in reality and is actually beneficial in real decisions. This leads us back to the above mentioned concept of ecological rationality.

The model of the fictitious homo œconomicus is a showpiece model of rationality in economics, which is often used in introductory courses for pedagogic purposes. But beyond that there are other modi of rationality, that can in principle also include other-regarding preferences like some altruistic behaviour. Thus there are models of rationality with fairness, reciprocity *etc.* Such models of rationality with other-regarding preferences are no homo œconomicus models anymore but are still considered rational models in the literature, in as much as they can still be conform to some requirements on the preferences which ensure consistency. This is an important difference that gets overlooked all to often in public discourse on economics.

The following paragraphs directly build on the Sec. 2.3 and the remarks to the subtitle of GIBBS (1902). Common knowledge and experimental evidence have shown:[674] modern humans, *homo sapiens sapiens* or simply 'We', are very different from *homo œconomicus*, in part because we still have much of *homo sapiens* or even older ancestors in us. If we would correct for the mistake of category that conceives rationality solely as a property of the individual, it would enable us to recognise easily that weird conceptions of rationality lead straightly to rational fools.[675] A rational theory can very well contain behaviour that seems irrational to an observer (with an odd conception of rationality) as has been shown repeatedly in the literature on the following keywords: allocative efficiency of market outcomes with zero-intelligence participants or noise traders,[676] on rational inattention[677] or on rational bubbles.[678] Furthermore, it would correct the category mistake and remove the label 'irrational'

[673] ULAM 1976, p. 91.
[674] GIGERENZER et al. 2011; KAHNEMAN and TVERSKY 1979; TVERSKY and KAHNEMAN 1992.
[675] SEN 1977.
[676] DE LONG et al. 1990; SHLEIFER and SUMMERS 1990; GODE and SUNDER 1993, 1997; SUNDER 2006.
[677] SIMS 2003, 2006.
[678] BLANCHARD and WATSON 1982; DIBA and GROSSMAN 1987, 1988b,a; LUX and SORNETTE 2002.

from several stigmatised empirical behaviours identified in positive or descriptive studies that only fall under this category due to a quixotic conception of rationality. Rationality is a very fragile and observer-dependent notion that hinges extremely on its precise definition and the rule garbage in, garbage out applies. Thus our conclusion that rationality is overrated as an behavioural determinant.

7.2 Bounded Rationality and Behavioural Economics

We will cursory discuss two examples from microeconomics and one example from macroeconomics that fall in the category of the neoclassical repair shop. We start with the former.

7.2.1 Axiomatic Variation in Neoclassical Microeconomics

Adding new constraints (such as asymmetric information) to the otherwise unchanged neoclassical framework is seen by Kapeller (2011) as a form of 'axiomatic variation' and he identifies the utility function as the key object of neoclassical flexibility in order to encompass all kinds of behaviours that are actually at odds with it[679]. Thus, from a methodological viewpoint Güth (2008, p. 251) characterises the repairs and game-fitting exercises rightly as 'new developments in neoclassical microeconomics', which takes them most of their revolutionary charm, with which they are sometimes presented. As Peters (2011c, p. 4926) alluded

> "[t]his is problematic because the arbitrariness of utility can be abused to justify reckless behaviour, and it ignores the fundamental physical limits, given by time irreversibility, to what can be considered reasonable. [...] [That is why in ergodicity economics] arbitrary utility functions are replaced by the physical truth that time cannot be reversed."

Indeed, with a little help of absurd preferences and idiosyncratic discounting one can rationalise even the most reckless behaviour such as drug addiction. In that case, addicts are myopic utility maximisers of their infinite lifetime, which contains a preference for the immediate thrill and they discount payoffs far in the future so strongly that effectively only the immediate future is relevant for them, so why care for the future given these preferences – this would clearly be irrational.[680]

679 The prime example for this method is the application of 'the economic approach' to non-market interactions in Becker (1976).

680 Becker and Murphy 1988.

Another example, are the influential papers by FEHR and SCHMIDT (1999) and BOLTON and OCKENFELS (2000), who treat the then (for economics novel) factors fairness,[681] reciprocity and the like wholly within the neoclassical framework of utility maximisation. The contribution brought about through these and other papers in mainstream behavioural economics lies more in that important aspects of economic transactions like fairness, reciprocity, empathy, altruism, cooperation, trust, *etc.* have moved into the area of attention of mainstream economics and are no longer disregarded in its theoretical machinery. However, the relevance of such aspects for economic transactions is sometimes obvious to the man in the street. To put a finer point on it, what these papers actually achieved, was to find a way to introduce fairness concerns and the like into the existing mathematical framework of optimising utility functions. From the discussion in this thesis it follows, that this can only be a beginning and not the end in the search for better economic theories. First and foremost, while these attempts are inflating the utility function the free parameter problem hasn't been taken into account properly. Through ENRICO FERMI and folk memory the following remark on the free parameter problem survived: 'I remember my friend JOHNNY VON NEUMANN used to say, with four parameters I can fit an elephant, and with five I can make him wiggle his trunk'.[682] A conjecture that has recently been proven.[683]

If a utility function is constructed as an arbitrary concave and bounded function, such a function assigns lower values to high winnings in the infinite ensemble set than the absolute value of money would. Strangely enough, the arbitrariness of BERNOULLI's individual utility level is seen as a strength of EUT by others, because the solution leaves room for idiosyncratic utilities depending on the individual wealth levels.[684]

7.2.2 Bounded Rationality in Macroeconomics

Having spent some pages on microeconomics, let us briefly remark on a possible road for bounded rationality macroeconomics. A prominent although misleading example is SARGENT (1995). Despite its title, *Bounded Rationality in Macroeconomics*, SARGENT (1995) is still very close to the neoclassical programme. The approach suffers from a form of hyperrationality the problem with the REH: 'when implemented numerically or econometrically, rational expectations models impute much more knowledge to the agents within the model (who use the equilibrium probability distributions in evaluating their Euler equations) than is possessed by an econometrician, who faces estimation and inference problems that the agents in the model have somehow solved'[685]

681 RABIN 1993.
682 DYSON 2004.
683 MAYER et al. 2010.
684 Additionally, the quote illustrates how easily rationality in economics gets misunderstood, see GEISENDORF (2009, 2010).
685 SARGENT 1995, p. 3.

A different road towards a more bottom-up theory of macroeconomics with truly reasonable microfoundations is sketched in Colander et al. (2008) and De Grauwe (2010). In this respect, agent-based models (ABMs) are promising. The microeconomic foundation can be fleshed out explicitly via respective algorithmic coding of the agents. Its great advantage is the freedom of the modeller to endow the agents with an arbitrary amount of intelligence. Thereby the whole spectrum between zero-intelligence, bounded rationality and full rationality becomes treatable in a single framework.[686] Much in the way Arthur (1999, 2010, 2013, 2015) envisaged complexity economics not as an alternative or complement to neoclassical economics, but as a more general out-of-equilibrium economic theory, which includes neoclassical economics as a special case, *e.g.* for perfect competition and if the agents possess complete information.[687]

7.2.3 Original Intent of Bounded Rationality Research

The original motivation of the founding fathers of bounded rationality, like Herbert Simon or Reinhard Selten, was to refine the omniscient homo œconomicus who optimises without any internal cognitive bounds or external informational restrictions. At the time of early behavioural economics, such research was performed under the umbrella term of 'bounded rationality', because it explored forms of non-optimising behaviour as consistent with bounded forms of rationality like search processes to discover choice alternatives, satisficing and aspiration adaptation and questioned explicitly the role of utility functions and unobservable beliefs.[688] Although Selten (2002, p. 15) remarks that bounded rationality cannot be precisely defined, because it must be explored, he outlines it as 'the rational principles that underlie nonoptimizing adaptive behavior of real people'. Herein, early behavioural economics was more open in as much as that it did not equate rationality with optimisation in the literature.[689] Today, the most important research programme of truly bounded rationality deals with the identification of fast and frugal rules and heuristics that are employed by a certain species (humans, animals, institutions, *etc.*). This research analyses how to rationally understand non-optimising behaviour. Such research is done under the terms of the 'adaptive toolbox'[690] or 'ecological rationality'.[691] A nice example for a rationality that is ecological is the famous Linda problem or conjunction fallacy.[692] Humans utilise the information given

686 See among other sources Delli Gatti et al. 2008, 2011.

687 For the impossible sake of completeness, recently, with De Grauwe (2012) another behavioural macroeconomics theory appeared. Whether or not it falls prey to the same criticism as Sargent's approach, we can't judge at the moment.

688 Simon 1947, 1955, 1956, 1957; Sauermann and Selten 1962; Selten 1991, 1998.

689 Selten 2002; Gigerenzer and Selten 2002b.

690 Gigerenzer et al. 1999; Gigerenzer 2000a; Gigerenzer and Selten 2002a; Gigerenzer et al. 2011; Hertwig et al. 2013.

691 Gigerenzer 2000a, ch. II; Smith 2003, 2007; Todd et al. 2012.

692 Gigerenzer 2000b.

in the environment of the problem, *i.e.* in the description of Linda, regardless whether this implies contradictions with the laws of probability or logic, that describe ideal situation that do not correspond with the real-world situation.

In summary, the research on bounded rationality is exactly a proper form of dealing with rationality as it was in the mind of GIBBS and the general natural sciences. Now we turn to give a detailed critique of expected utility theory, that adds to the remarks we gave when we presented the solution strategy of moral expectation in Sec. 4.3.

7.3 On Utility and Maximisation

7.3.1 Inflating the Container – On Complex Utility Functions

There is considerable discontent with a misdirection of that part of research on bounded rationality that found it's way into mainstream behavioural economics.[693] For GÜTH (1995, 2008) too much of behavioural economics research focuses on defining a suitable optimisation or model game a posteriori to retrodict observed behaviour. He coined the terms 'neoclassical repair shop' and 'game-fitting exercises' for such activities. The method of the 'neoclassical repair shop' usually works like this, introduce 'additional arguments of utility functions like a desire for fairness, altruism, or envy, as well as specific forms of incomplete information',[694] all with the goal that the so adjusted model yields the observed behaviour as the optimal one. As a consequence utility functions become mere containers for more and more effects, and grow more complex. Sadly this 'offers no really satisfying explanations, but shift only the problem to another level of research questions, namely why people have such utility functions.'[695] This is particularly incoherent, as the original framework of neoclassical economics is meant to avoid psychology or starts 'after psychology', by pooling all psychological idiosyncrasies in the utility function of an individual.[696]

BERG and GIGERENZER (2010) use an even more extreme diction and characterise the evolution of decision theories from mathematical expectation to EUT and eventually to CPT as a 'Bernoulli repair program aimed at resuscitating the mathematical operation of weighted integration, based on the definition of mathematical expectation, as a theory of mind' and continue in the same vein as this thesis that the 'repair program is based largely on tinkering with the mathematical form of the mathematical expectation operator and cannot be described as a sustained empirical effort to uncover the process by which people actually choose gambles.'[697]

693 SELTEN 1991, 2002; GÜTH 1995, 2008; GIGERENZER and SELTEN 2002b; BERG and GIGERENZER 2010.
694 GÜTH 1995, p. 342.
695 GÜTH 1995, p. 342.
696 GÜTH and KLIEMT 2004.
697 BERG and GIGERENZER 2010, pp. 136, 138.

For both, behavioural economics is simply neoclassical economics plus new parameters with psychological names, such as loss aversion, other-regarding preferences such fairness and reciprocity, subjective probability weighting or hyperbolic discounting. The injection of such parameters transforms ordinary utility functions into behavioural utility functions which accommodate additional preferences (see also **Table 7.1**).

> "By virtue of this modeling strategy based on constrained optimization, with virtually all empirical work addressing the fit of outcomes rather than justifying the constrained optimization problem-solving process itself, behavioral economics follows the Friedman as-if doctrine in neoclassical economics focusing solely on outcomes. By adding parameters to increase the R-squared of behavioral models' fit, many behavioral economists tacitly (and sometimes explicitly) deny the importance of correct empirical description of the processes that lead to those decision outcomes."[698]

Instead of such correct empirical descriptions of the problem-solving process, the adherents of the adaptive toolbox and ecological rationality research programme further complain about the interpretation of deviations from (seemingly) rational behaviour, called 'biases' or 'fallacies', and the evaluation of people's judgement capabilities as poor or even 'irrational'. The rhetoric implies the status of mere optical illusions that need to be corrected. Ultimately, it is such a view of empirical results in combination the 'The Rhetoric of Irrationality'[699] that justifies a consulting industry in the name of *nudging* people to better decisions.

An empirical approach in economics can have two consequences, either a profound reformulation of the theory or disavow any disagreement with the theory by calling it anomalies from an otherwise correct theory for the majority of the cases. Standard behavioural economics, such as CPT, chose the latter. It is instructive to look at the development of the history of (mainstream) economic thought from an evolutionary perspective. Besides reaching leaves high up in trees, due to its long neck the giraffe has problems to lay down, to drink from a low source of water and on various other occasions. It is well known among evolutionary biologists that the giraffe would be much better of if it had more than the usual seven neck vertebra, that would make up its neck. But such a change in the mammal blueprint is very unlikely to happen in evolution, one could say it is too radical. It almost seems as if a profound reformulation of the core of economic theory is too radical, too. We propose such a change with the temporal optimisation scheme from Ergodicity Economics. Although this is still an optimisation scheme, the change of the optimand comes with profound changes on the notion of an evolutionarily reasonable goal or rationality. Before we continue with the evolution of

698 Berg and Gigerenzer 2010, p. 144.
699 Lopes 1991.

decision theories, we give some reasons why we shouldn't break with optimisation in general not too easily.

7.3.2 Optimisation as a General Organisation Principle in Nature

However, optimisation is not to be condemned per se. The optimisation principle is abound as an order principle in nature, see *e.g.* the least action principle[700] or maximum entropy principle (MEP)[701] and has found many applications in the theory of nonlinear programming. Besides the much more radical question: *What is a sensible embedding for randomness?* and before entering the wilderness of bounded rationality one can pose the admittedly less radical question: *What is a sensible maximand?* And it is in the combination of the radical and the less radical question that ergodicity economics differs from traditional economic approaches. The first goal in order to maximise anything and subsequently maybe strive for PARETO-efficient states is simply to survive the next period. An optimisation routine should reflect that and the optimisation has to ensure primarily the survival. This turns the optimisation into a problem to prevent the gambler's ruin and is thus much closer to a sensible rationality criterion also in evolutionary terms.

7.4 Evolution of Decision Theories

> After almost seven decades of intense attempts to generate and validate estimates of parameters for standard decision theories, it is perhaps time to ask whether the failure to find stable results *is* the result.
>
> FRIEDMAN ET AL. (2014, P. 3)

In **Table 7.1** a selection of alternative decision theories are given together with their decision criterion and its mathematical structure. It combines and considerably extends the tables in SCHOEMAKER (1982) and CAMERER (1999) under the theme of identifying their conception of risk or which embedding of randomness they use, either within time or traditionally within a virtual ensemble indicated in the last column 'mathematical structure'. The take away message can be gathered quickly, if skimmed over the last column. None of the decision theories so far presented in the literature explicitly takes the time perspective into account and the growth-optimal decision theory, the last entry, is the first attempt to deviate from virtual ensembles and instead considers effect on the time-average growth of the decision makers

700 SAMUELSON 1965, 1972.
701 JAYNES 1957a,b.

wealth that is induced by the decision problem. Usually the decision criterion undergoes an optimisation, in the spirit of being either a minimand or a maximand, we thus entitle the respective column 'optimand'. Good surveys on further alternative decision theories given in **Table 7.1** and beyond are Fishburn (1988, 1991) or Sinn (1989). In the following we comment on some items from the table.

Subjective Expected Utility

Savage (1972) axiomatised utility theory and subjective probability, which gave rise to the subjective expected utility (SEU) model. Therein, the decision maker performs Bayesian updating of his subjective probabilities as his prior. This model defined the rationality concept in economics for some time as the maximisation of subjective expected utility. At the same time SEU became the accepted rationality model, Allais showed experimentally violations of its axiomatic foundations if certain consequences in the choice between two otherwise similar gambles have been added or removed.[702]

Behavioural utility

As indicated above, the general mode of behavioural economics seems to be to inflate the utility function to accommodate arbitrary preferences such as preference for the equality of payouts.

Maximin rules

Wald's maximin rule can be illustrated in two ways, either in purely game theoretical terms as a two-person zero-sum game or as an interference problem between Nature and a statistician. In a two-person zero-sum game both players are completely ignorant about each other's decision, but are aware of the common payoff matrix. Player 1 chooses his strategy denoted by τ_1, player 2's choice is denoted by τ_2. As we deal with a zero-sum game, the payoff function for player 1, $K(\tau_1, \tau_2)$, is simply the negative of player 2's payoff function, $-K(\tau_1, \tau_2)$. As player 1 strives to maximise his payoff, $K(\tau_1, \tau_2)$, player 2 tries to minimise $K(\tau_1, \tau_2)$ at the same time. Von Neumann and Morgenstern (1955, section 14.5) proved the equivalence of the minimax and the maximin procedure in pure strategies

$$\max_{\tau_1} \min_{\tau_2} K(\tau_1, \tau_2) = \min_{\tau_2} \max_{\tau_1} K(\tau_1, \tau_2). \tag{7.4}$$

[702] Allais 1953.

Table 7.1: Evolution of Decision Criteria. This table lists diverse decision theories to elicit their evolution. Except for the growth-optimal model no approach uses temporal optimisation.

Decision Theory	Criteria/Optimand	Mathematical Structure
Mathematical Expectation (HUYGENS 1657)	$\max \langle \Delta w \rangle$	ensemble average
BERNOULLI's criterion (BERNOULLI 1738)	$\max \left\langle \Delta u^{+} + \Delta u^{-} \right\rangle$	ensemble average
EUT (LAPLACE 1812)	$\max \langle \Delta u\,(w) \rangle$	ensemble average
Variants of Expected Utility Theory		
Subjective Expected Utility (RAMSEY 1950; SAVAGE 1972; QUIGGIN 1982)	$\max \sum_i f\,(p_i) \cdot \Delta u\,(w)$	ensemble average
Behavioural Utility (FEHR and SCHMIDT 1999) (BOLTON and OCKENFELS 2000)	$\max \sum_i p_i \cdot u \left(\Delta w_i^1, \bullet \right)$	ensemble average
Maximin (Minimax) Rule (WALD 1945, 1950)	$\min_{\tau_2} \max_{\tau_1} K(\tau_1, \tau_2)$	ensemble average
(Cumulative) Prospect Theory (KAHNEMAN and TVERSKY 1979) (TVERSKY and KAHNEMAN 1992)	$\max \mathfrak{V}\,(f_X)$	ensemble average
Hyperbolic Discounted Utility (STROTZ 1955; AINSLIE 1975; THALER 1981)	$\max \beta \sum \delta^t u\,(\Delta w)$	ensemble average
Growth-Optimal (PETERS and ADAMOU 2018b)	$\max \overline{g}$	time average

In the conception of the game as an inferential problem, Nature has the role of player 1 and player 2 is the statistician. Nature chooses a state of the world, θ. The statistician bases his choice on a particular statistical decision function, $\omega(E)$.[703] The outcome, $K[\theta, \omega(E)]$, of the game poses a risk to the statistician given the decision function he chose, *e.g.* a negative payout $r[\theta|\omega(E)] = -K[\theta, \omega(E)]$, which he wishes to minimise. Without further information on Nature's intentions, it is now reasonable and actually most cautious of the statistician to base his statistical decision function and therefore his decision theory on the assumption that Nature chooses a state of the world that is the least favourable for him. Put in other words, if the statistician has no information to the contrary, he plays it safe if he assumes that Nature tries to maximise his risk, which he again seeks to minimise, $\min_{\omega(E)} \max_{\theta} K[\theta, \omega(E)]$,[704] which results in the analog of eq. (7.4).

In the inference conception as well as in the zero-sum game conception, the generalisation to mixed strategies is a realistic extension even if only one player uses them. For example it is reasonable to conceive Nature as playing not a particular pure strategy but a mixed strategy. The statistician does not necessarily play a mixed strategy, which would mean to randomise his choice over the space of all possible decision functions, but chooses a particular statistical decision function. Although being a non-probabilistic decision-making model in pure strategies, the extension of the maximin rule to mixed strategies is based on the minimisation of the average risk for a least favourable distribution in the statistical decision function. This average risk is again an ensemble average.

To formalise mixed strategies, a player i does not choose a particular value τ_i but a probability distribution, such that τ_i becomes a random variable. The variables τ_i thus become vectors, $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, with the probabilities of the respective strategies as their elements, $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T$ and $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_m)^T$. The average risk, $\langle K(\xi, \eta) \rangle$, is calculated then as an ensemble average over the probability-weighted possible outcomes,[705]

$$\langle K(\xi, \eta) \rangle = \sum_{\tau_1=1}^{n} \sum_{\tau_2=1}^{m} K(\tau_1, \tau_2) \xi_{\tau_1} \eta_{\tau_2} \; . \tag{7.5}$$

One famous solution arises if the space of possible solutions (*i.e.* the payoff function generated by the pairs of choices) spans a manifold that has the shape of the surface of a saddle, such as for the function $f(x, y) = x^2 - y^2, \quad x, y \in \mathbb{R}$ in **Fig. 7.1**. Here, the maximin solution is exactly at the saddle point, because this point is at the minimum of the manifold in one direction and at the same time it is the maximum in the orthogonal direction. The maximin rule thus generalises the von Neumann & Morgenstern theory and there exist several variations of it such as the Savage's minimax regret rule and (α) maxmin expected utility, which are all based on the same ensemble average risk notion.

The conception of the interference problem of the statistician about Nature's choices shares similarities with the MEP mentioned in the introduction. The MEP is also built on the principle of identifying the least unwarranted estimation. In both frameworks, Nature's choice can be estimated in the most reliable way if we interpret it as presenting us with the maximal amount of uncertainty given the information it actually conveys.

703 The statistical decision function, ω, assigns each observation E a distribution function that leads to the acceptance of the hypothesis that $\omega(E)$ is the correct distribution, $H_{\omega(E)}$. See Wald (1939, 1945, p. 265) for more details on the definition of the statistical decision function.

704 Wald 1945, p. 279.

705 Wald 1945, p. 280.

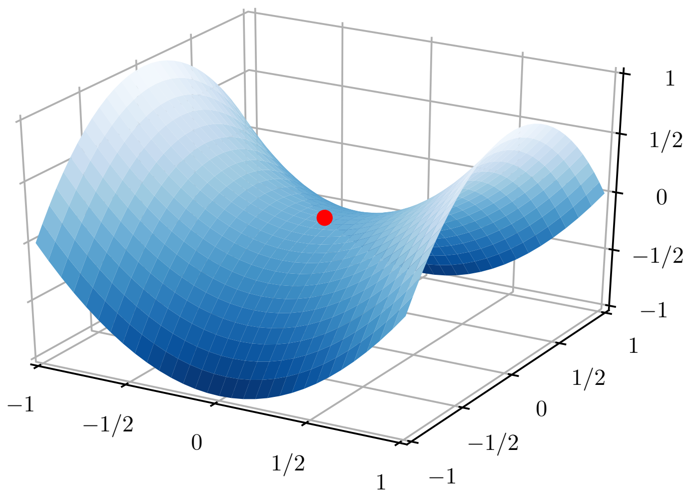


Figure 7.1: Saddle surface spanned by the function $f(x, y) = x^2 - y^2$ with $x, y \in$. The maximin procedure yields the saddle point, $s^* = (0, 0)$, as the optimal point.

We have seen already in Sec. 4.3 that the traditional mathematical expectation operator and expected utility theory are based on ensemble averages. The maximin rule is a generalisation of VON NEUMANN and MORGENSTERN expected utility, hence, WALD's maximin rule is based on an ensemble average risk conception, too. The same applies to the other variants of expected utility theory like the (α-) maxmin expected utility model,[706] CHOQUET expected utility,[707] subjective expected utility,[708] CPT,[709] minimax regret rule which is an application of the minimax theorem to regret, loss or opportunity costs.[710]

In summary, the outcome of **Table 7.1** does not surprise, rather it corroborates the ubiquity of the embedding of randomness within a virtual ensemble. This embedding is ubiquitous because the main tool to deal with randomness and uncertainty in all the listed decision theories is the expectation operator, which is – as mentioned many times by now – an ensemble average.

706 GILBOA and SCHMEIDLER 1989; GHIRARDATO et al. 2004.
707 SCHMEIDLER 1989.
708 RAMSEY 1950; SAVAGE 1972; QUIGGIN 1982.
709 KAHNEMAN and TVERSKY 1979; TVERSKY and KAHNEMAN 1992.
710 NIEHANS 1948; SAVAGE 1951.

8 Kelly Criterion under Uncertainty

"[F]ractional Kelly" is prudent to the extent the results of the Kelly calculations reflect uncertainties.[711]

EDWARD O. THORP

Ergodicity Economics is first and foremost a theory on the evaluation of uncertain prospects, which acknowledges the ergodicity problem. As the evaluation of uncertain prospects is at the formal basis of economics it constitutes a new approach to economic problems. One of the basic insights of the previous chapters is that in order to embed randomness within the time domain, the explicit incorporation of a dynamic is crucial. In the case of wealth this is a multiplicative dynamic which makes sequential decision making under uncertainty also *con*-sequential. In doing so *Ergodicity Economics* goes beyond the criterion of the expectation value (or some transformations of it) and instead we derived the time-average growth rate of wealth as a reasonable decision criterion. Besides the early treatments of multiplicative repetition mentioned in Sec. 4.6 by *e.g.* EULER and WHITWORTH, another most interesting connection to *Ergodicity Economics* can be traced back to the seminal publication by JOHN L. KELLY which appeared already in 1956. Possibly due to its title *'A New Interpretation of Information Rate'* and its information theory jargon the paper has not received the attention it deserves from the economics community, except in evolutionary finance as we discussed in Subsec. 5.2.1.3 and Subsec. 5.2.2.2. Here the community is recently coming to recognise the KELLY criterion as its central pillar. However, in evolutionary finance KELLY's decision criterion is still forced in the PROCRUSTEAN bed of utility theory.

So far the discussion focussed on whether to enter a gamble or not. It turns out that the achievement of the optimal time-average growth rate given a certain gamble is fully under control of the gambler. The decisive variable is the fraction of his wealth the gambler chooses to invest and is also known as the KELLY criterion. We present this in Sec. 8.2 which provides some preliminaries before we present KELLY's criterion In Sec. 8.3 we resolve the confusion surrounding the KELLY criterion and prove that it has nothing to do with utility. Therefore

[711] THORP 2006.

we augment the theory by taking the investment fraction as another variable explicitly into consideration. This makes transparent that Kelly in fact devised a criterion which maximises nothing else than the time-average growth rate of wealth. Thereby its true content is revealed with the help of the terminology provided by *Ergodicity Economics*. Moreover, we lay bare that a specific leverage is already assumed in the ensemble approach though only implicitly. Furthermore in Sec. 8.4 we present an outlook on advancement of the Kelly criterion under uncertainty, which is ongoing cutting edge research of ours in *Ergodicity Economics*, which we continue in a follow-up project. The advancements involve the realistic situation when a gambler takes into account his own ignorance about some of the parameters that define the gamble.

8.1 Preliminaries of the Gamble Setup

Very much as before in eq. (4.93), we consider a binary gamble or Bernoulli trial, where the wealth after one round of a gamble is the product of the initial wealth and the realisation of a random growth factor. Before we can state explicitly the two growth factors in this augmented setting, we introduce the dependence of the payout on the invested fraction of the initial wealth. We recall that under multiplicative dynamics the payout is already proportional to the gambler's wealth w_t at the start of each round.

Let $\ell \in \mathbb{R}$ denote the fraction of a gambler's initial wealth that he invests in a gamble, to which we refer to as the *leverage*. *I.e.* if $0 < \ell < 1$ the gambler invests only a proper fraction of his initial wealth, if $\ell = 0$ he invests nothing and if $\ell = 1$ he invest his total initial wealth, and $\ell > 1$ encodes a situation where the gambler invests a multiple of his initial wealth in the gamble by borrowing additional money. Even the situation of short selling a gamble could be treated by $\ell < 0$ in principal. Irrespective of the outcome of the gamble, the fraction $(1 - \ell)w_t$ is not invested and stays constant.[712] The gambler wagers the amount ℓw_t and with probability p the wager is multiplied by $r_{1,t}$ and the payout from the gamble is $r_{1,t}\ell w_t$. With probability $1 - p$ the growth factor $r_{2,t}$ applies and the payout is $r_{2,t}\ell w_t$. Often gambles are considered where $r_{1,t} > 1$ and $0 \leq r_{2,t} \leq 1$, where there occurs some loss to the wagered amount. If the second case shall denote an outcome where a total loss of the wager appears, then $r_{2,t} = 0$ and the wager $r_{2,t}\ell w_t = 0$ and is just lost. In such cases $r_{1,t}$ is sometimes called the payout

[712] In principle a growth factor could be attached to the risk-free cash, too, which is in our case implicitly equal to unity. See for instance Peters (2011b) in the context of continuous time.

odds. This yields the following two possible payouts from the gamble

$$w_{t+\delta t} = \begin{cases} \overbrace{(1-\ell)w_t}^{\text{risk-free asset}} + \overbrace{r_{1,t}\;\ell\;w_t}^{\text{risky gamble}} & \text{with probability } p, \\ \underbrace{(1-\ell)w_t}_{\text{risk-free asset}} + \underbrace{r_{2,t}\;\ell\;w_t}_{\text{risky gamble}} & \text{with probability } 1-p \end{cases} \tag{8.1}$$

Over the course of the repetitions of this gamble a particular sequence of $\{R_{r,\bullet}\}_{t\in\mathbb{N}_0}$ realises (using the notation we introduced in Subsec. 2.4.5 in the context of stochastic processes) and generates the gambler's wealth trajectory $\{W_t\}_{t\in\mathbb{N}_0}$. We assume to repeat the identical gamble at every time step or, put differently, we consider a repeated static gamble, thus the realisations of the growth factor $R_t \in \{r_1, r_2\}$ are time-independent, therefore we denote them by r_1 and r_2 in the remainder,

$$w_{t+\delta t} = (1-\ell)w_t + \begin{cases} r_1\;\ell\;w_t & \text{with probability } p, \\ r_2\;\ell\;w_t & \text{with probability } 1-p\;. \end{cases} \tag{8.2}$$

Now the question arises what is the optimal leverage? Put differently, what is the optimal fraction to invest in a gamble? We know already from Sec. 4.5 that the ensemble perspective does not yield credible decision criteria. However, a short analysis of the optimal leverage in the ensemble perspective reveals a hidden assumption. Therefore, we answer this question from both the ensemble perspective as well as the time perspective.

Ensemble Perspective

Let us compute the expected change in wealth $\langle \Delta w \rangle$ of the gamble in eq. (8.2). Therefore we take all known quantities out of the expectation and arrive at

$$\langle \Delta w \rangle = \langle w_{t+\delta t} \rangle - w_t = \langle (1-\ell)w_t + r_n \ell w_t \rangle - w_t \tag{8.3}$$

$$= \langle r_n \ell w_t \rangle + (1-\ell)w_t - w_t = \langle r_n \rangle\, \ell w_t - \ell w_t = (\langle r \rangle - 1)\, \ell w_t\;, \tag{8.4}$$

whereby the expected growth factor or ensemble-average growth factor $\langle r \rangle$ is simply the probability weighted mean of all possible growth factors given in eq. (4.103). From eq. (8.4) wee see that the expected change depends linearly on the leverage, thus in order to maximise the expected change the gambler should maximise his leverage as soon as the expected growth rate is greater unity. In fact, this perspective suggests to increase leverage indefinitely or at least up to a borrowing constraint. But we see even more, in order to achieve the ensemble-average growth factor or the expected change per round to its full extent the leverage must equal unity $\ell = 1$. We can also look at the the expected final wealth and see the same fact

$$(8.5) \qquad \langle w_{t+\delta t} \rangle = \underbrace{\langle r \rangle \, \ell w_t}_{\text{payout from the risky asset}} + \underbrace{1 \cdot (1-\ell) w_t}_{\text{payout from the risk-free asset}} .$$

Here, the gambler only receives the full effect of the expected growth factor $\langle r \rangle$ without any lesser returns from the safe asset if he is fully invested $\ell = 1$. To emphasise the implicitly assumed growth factor of 1 we have added the superfluous expression '$1 \cdot \ldots$' in the second summand in eq. (8.5).

Thus we have made transparent that the ensemble perspective is associated with an implicit assumption of a gambler being fully invested, which is a situation that pretends as if the wealth will grow on average exponentially with the ensemble-average growth factor $\langle r \rangle$ per round and that this situation can only be reinforced with higher leverage but not harmed. The strong stimulus to leverage as much as possible which is accompanied with the ensemble perspective incentivises reckless behaviour and clearly violates the ancient insight: Do no put all your eggs in one basket. Here the two baskets are represented by holding either risk-free cash (no gamble) and by the risky gamble. The linear dependence on the leverage is inherited by the ensemble-average growth rate $g_{\langle\rangle}$ where it enters in the logarithm, which is however still unbounded.

In conclusion, within the ensemble perspective the optimal leverage is the highest possible leverage or no optimal leverage exists at all.

Let us consider the following example.

Example 8.1.1 (+50%-40%-Gamble from the Ensemble Perspective). *A gambler is offered a coin toss with the equal chances of either a 50% win or a 40% loss on the invested money – should a gambler play? If yes, what is his optimal leverage?*

$$(8.6) \qquad w_{t+\delta t} = \begin{cases} (1-\ell)w_t + 1.5\ell w_t & \textit{with probability } p = \frac{1}{2}, \\ (1-\ell)w_t + 0.6\ell w_t & \textit{with probability } 1-p = \frac{1}{2} . \end{cases}$$

The expected change in wealth is

$$(8.7) \qquad \langle \Delta w \rangle = \left(\frac{1.5 + 0.6}{2} - 1 \right) \ell w_t = 0.05 \ell w_t$$

The expected change in wealth grows linear in ℓ, thus the gambler should maximise ℓ in order to maximise the expected change.

The ensemble average of the growth factors according to eq. (4.103) is

$$\langle r \rangle = \frac{1.5\ell + 0.6\ell}{2} = \frac{2.1\ell}{2} = 1.05\ell \ , \tag{8.8}$$

and yields an expected growth factor of 1.05, thus it pretends as if our wealth will grow on average exponentially at 5% per round. But as we can see this is only the case if the gambler has a leverage of $\ell = 1$, in other words, if the gambler is fully invested. In fact, this decision criterion suggests to increase leverage indefinitely or at least up to a borrowing constraints. Of course, the linear dependence is inherited to the ensemble-average exponential growth rate which according to eq. (4.104) is given by

$$g_{\langle\rangle} = \frac{1}{\delta t} \log \langle r \rangle = \frac{1}{\delta t} \log 1.05\ell = \frac{1}{\delta t} (\log 1.05 + \log \ell) = \frac{1}{\delta t} (0.049 + \log \ell) \ . \tag{8.9}$$

The same situation here, in order to maximise the exponential growth rate maximise the leverage indefinitely or up to a possible constraint.

Example 8.1.2 (+50%-40%-Gamble from the Time Perspective). *A gambler is confronted with the identical gamble, but now in order to maximise his wealth he computes the time-average exponential growth rate dependent on the leverage. At first, we have to compute the time-average growth factor to compute in the second step the time-average exponential growth rate given in eq. (4.113).*

The time-average growth factor according to eq. (4.112) is given by

$$\overline{r} = \lim_{T\to\infty} \prod_{n=1}^{n_T^{\max}} r_n = \prod_{n=1}^{\infty} r_n^{p_n} \tag{8.10}$$

$$= (1.5\ell)^{\frac{1}{2}} (0.6\ell)^{\frac{1}{2}} = \sqrt{1.5 \cdot 0.6}\,\ell = \sqrt{0.9}\,\ell = 0.95\ell \ . \tag{8.11}$$

Thus in order to arrive at positive growth for this gamble – that is growth factors greater than unity – the gambler needs to borrow additional money.

As we show in a moment, high leverage will lead to the gambler's ruin, because the fluctuations in the multiplicative dynamics are not taken properly into account. We can conclude that the leverage mediates between the time-average perspective and the ensemble-average perspective.

8.2 The Kelly Gamble

Kelly (1956) started to consider the long run exponential growth rate of a gambler's wealth

$$\lim_{T\to\infty} \frac{1}{T} \log\left(\frac{w_{t+\Delta t}}{w_t}\right) , \tag{8.12}$$

which clearly is the time-average exponential growth rate of wealth, so we denote it by the usual symbol and transform it using the higher-index roots again

$$\bar{g} = \lim_{T\to\infty} \frac{1}{T} \log\left(\frac{w_{t+\Delta t}}{w_t}\right) = \lim_{T\to\infty} \log\left(\frac{w_{t+\Delta t}}{w_t}\right)^{1/T} = \lim_{T\to\infty} \log\left(\sqrt[T]{\frac{w_{t+\Delta t}}{w_t}}\right) . \tag{8.13}$$

Kelly (1956) analysed a somewhat stylised gamble, he took a Bernoulli trial as his basis which pays either +1 or −1 and the gambler chooses the fraction of his wealth to invest in the gamble:

$$r_t = \begin{cases} r_{1,t} = 1+\ell & \text{with probability } p, \\ r_{2,t} = 1-\ell & \text{with probability } q = 1-p\ . \end{cases} \tag{8.14}$$

Then the gambler's final wealth is determined by

$$w_{t+\Delta t} = w_t \cdot r_j^{n_j} = w_t\ (1+\ell)^{n_1}(1-\ell)^{n_2}\ . \tag{8.15}$$

If we plug eq. (8.15) in eq. (8.12), apply the LLN in the second last row, it yields the time-average growth rate

$$\bar{g} = \lim_{T\to\infty} \frac{1}{T} \log \frac{w_{t+\Delta t}}{w_t} = \lim_{T\to\infty} \left[\frac{1}{T} \log\left(\frac{w_t\ (1+\ell)^{n_1}(1-\ell)^{n_2}}{w_t}\right)\right] \tag{8.16}$$

$$= \lim_{T\to\infty} \left[\frac{1}{T} \log\left((1+\ell)^{n_1}(1-\ell)^{n_2}\right)\right] \tag{8.17}$$

$$= \lim_{T\to\infty} \left[\frac{1}{T} \log\left((1+\ell)^{n_1}\right) + \frac{1}{T} \log\left((1-\ell)^{n_2}\right)\right] \tag{8.18}$$

$$= \lim_{T\to\infty} \left[\frac{n_1}{T} \log(1+\ell) + \frac{n_2}{T} \log(1-\ell)\right] \tag{8.19}$$

$$= p \log(1+\ell) + q \log(1-\ell)\ . \tag{8.20}$$

Finally, in eq. (8.20) Kelly arrived at an expression of the time-average growth rate which depends on the leverage ℓ. Now everything is prepared for a maximisation of $\bar{g}$ with respect to ℓ. In this maximisation Kelly used some algebraic tricks in which the information theory custom of using logarithms to the base two becomes handy. We spare the details of this maximisation, which involves some transformations whose exact knowledge is not important for the purpose of this chapter. The only property of the maximisation of $\overline{g}$ that is important is that it has a quadratic form, which is shown in **Fig. 8.3** for numerical simulations of specific gambles and such a symmetric parabola has a maximum. The resulting optimal leverage is

$$\ell_{\text{opt}} = p - q \qquad \forall\ p \geq q > 0\ , \tag{8.21}$$

which in gambling terminology is referred to as the *edge*, thus the if the edge is zero, no bet is placed. If the payout for a win is $+B$, then the odds are B and the edge is $Bp - q$ and the optimal leverage is

$$\ell_{\text{opt}} = \frac{\text{edge}}{\text{odds}} = \frac{Bp - q}{B}\ . \tag{8.22}$$

This is Kelly's famous criterion and we see that it maximises the time-average exponential growth rate. Many of the publications mentioned in Subsec. 5.1.1 built on this result.[713]

8.3 Kelly and Ergodicity Transformations

> Richtiges Auffassen einer Sache und Mißverstehen der gleichen Sache schließen einander nicht vollständig aus.
>
> Franz Kafka

Even from experts and in otherwise splendid sources we find statements like the following:

> "Kelly (1956) is given credit for the idea of using log utility in gambling and repeated investment problems, as such it is known as the Kelly criterion. [...] Not only does he show that log is the utility function which maximizes the long run growth rate, but that this utility function is myopic in the sense that period by period maximization based only on current capital is optimal."[714]

In this section we use Kelly's own words to demonstrate that he was very explicit in drawing a demarcation line between his results and utility theory. Let us start with the Kelly's introductory remarks where we see already the property of the ergodicity transformation well up between the lines: 'Furthermore, this cost function must be such that its expected value

[713] The most important are Latané (1959), Breiman (1960b, 1961) and Thorp (1969, 1973).
[714] MacLean et al. 2011b, p. 5.

has significance, *i.e.*, a system must be preferable to another if its average cost is less.'[715] This is exactly what we derive in Subsec. 4.5.4, the logarithmic changes in wealth are an ergodic observable, thus the expectation of it has meaning, namely it coincides with the time average. Further Kelly referred explicitly to von Neumann & Morgenstern's utility functions:

> "The utility theory of Von Neumann shows us one way to obtain such a cost function. Generally this cost function would depend on things external to the system and not on the probabilities which describe the system, so that its average value could not be identified with the rate as defined by Shannon.
> [...] The author believes that [utility as a cost function] is too general to shed any light on the specific problems of communication theory."[716]

We could not concur more. The injection of utility constitutes an ad hoc transformation. It introduces an external idiosyncratic element in decision theory which is neither justified by the gamble set up, the dynamics nor anything else inherent to the problem. Still on p. 918, Kelly adds:

> "A cost function, if it is supposed to apply to a communication system, must somehow reflect this feature [...] the author would not know how to test such an arbitrary combination [of a utility as a cost function and the dynamics of the gamble]."

Kelly's transition from his introduction to the main part of his article and emphasises that he renounces from the introduction of any arbitrary utility and reads as follows:

> "What can be done, however, is to take some real-life situation which seems to possess the essential features of a communication problem, and to analyze it without the introduction of an arbitrary cost function"[717]

In what follows Kelly derives his mathematical result of the Kelly criterion, *i.e.* the optimal fraction to invest in a repeated gamble which maximises the time-average growth rate of the gambler's wealth, and resumes to the discussion of the cost function in the conclusion. Again, we could not concur more with his assessment of the facts:

> "The gambler introduced here follows an essentially different criterion from the classical gambler. At every bet he maximizes the expected value of the logarithm of his capital. The reason has nothing to do with the value function which he

715 Kelly 1956, p. 918.
716 Kelly 1956, p. 918.
717 Kelly 1956, p. 918.

> attached to his money, but merely with the fact that it is the logarithm which is additive in repeated bets and to which the law of large numbers applies."[718]

He signifies the difference between his decision criterion of the optimal time-average growth rate of wealth and other decision criteria, such as maximising the expected wealth or maximising the expected transformed wealth according to EUT. From the last sentence quoted above, it is clearly evident that he understood the role of the logarithm in the application of the LLN which is the manifestation of the ergodicity in this context, even if he does not use this word. Next he tries to come up with a tongue-in-cheek example in which additive dynamics could be realistic and thus the maximisation of the expected wealth would become optimal. Recall that this is a manifestation of sequential decision making that is not *con*-sequential or in KELLY's words:

> "Suppose the situation were different; for example, suppose the gambler's wife allowed him to bet one dollar each week but not to reinvest his winnings. He should then maximize his expectation (expected value of capital) on each bet. He would bet all his available capital (one dollar) on the event yielding the highest expectation. With probability one he would get ahead of anyone dividing his money differently."[719]

Furthermore, KELLY foresees the wide applicability of his result beyond mere gambling. *Ergodicity Economics* and this thesis prove him right:

> "Although the model adopted here is drawn from the real-life situation of gambling it is possible that it could apply to certain other economic situations. The essential requirements for the validity of the theory are the possibility of reinvestment of profits and the ability to control or vary the amount of money invested or bet in different categories. The 'channel' of the theory might correspond to a real communication channel or simply to the totality of inside information available to the investor."[720]

KELLY concludes with summarising the relevant result from an information theory point of view:

> "If the odds are consistent with the probabilities of occurrence of the transmitted symbols (*i.e.*, equal to their reciprocals), the maximum value of this exponential rate of growth will be equal to the rate of transmission of information."[721]

718 KELLY 1956, p. 926.
719 KELLY 1956, p. 926.
720 KELLY 1956, p. 926.
721 KELLY 1956, p. 926.

In merely commenting on KELLY's original statements, we believe to have shown once and for all that his result are an early ancestor to the research programme of *Ergodicity Economics* and one does him wrong if put in the proximity of utility theory. KELLY hardly could have been more explicit.

8.4 Outlook on Repeated Perturbed Gambles

Let us now briefly indicate the content of an ongoing research project, which will occupy us after the work of this thesis is finished. **Fig. 8.1a** depicts the generic set up of the repeated gambles we studied so far. The gamble is driven by a sequence of growth factors, whereby the two possible realisations are denoted here by r_{up} and r_{down}.

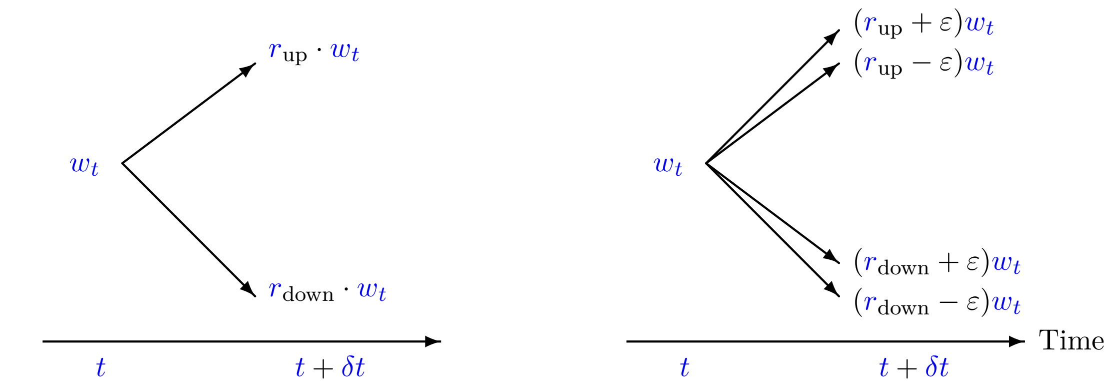
Fig. 8.1b depicts the advancement we have in mind, which incorporates an ignorance of the gambler about whether he really knows the exact parameters of the gamble, in this case the growth factors. The risk of overbetting is well-known in the community and a joint desideratum is to arrive at a proper reduced KELLY criterion which is called a fractional KELLY strategy. For instance

> "Long term compounders ought to avoid using a greater fraction ("overbetting"). Therefore, to the extent that future probabilities are uncertain, long term compounders should further limit their investment fraction enough to prevent a significant risk of overbetting."[722]

The ignorance is specifically designed to be expectation-neutral in the case of equal probability of the two outcomes. Expectation-neutral means the ignorance does not affect the expectation value of the growth factors, because an upward deviation $(r+\varepsilon)$ is as probable as a downward deviation $(r-\varepsilon)$. However, the perturbation becomes relevant in the time average of the growth factors. The fact that increased fluctuations affect the behaviour over time even if they are expectation-neutral is remarkable. Thus we consider two forms of uncertainty. First, the general randomness associated with the outcome of the random experiment. Second, an additional amount of uncertainty encodes the ignorance of the gambler and is denoted by ε.

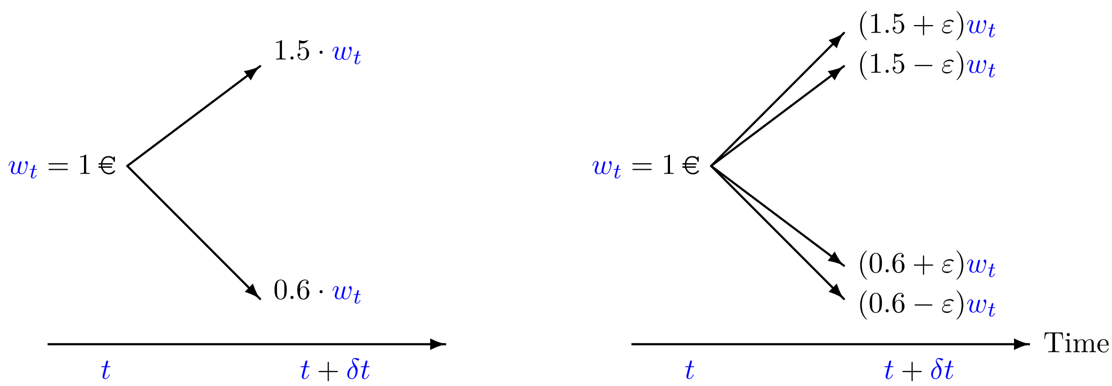
In the numerical simulations that follow, we use the parameters given in **Fig. 8.2**, *i.e.* with equal probability either $r_{\text{up}} = 1.5$ or $r_{\text{down}} = 0.6$ realises and the initial wealth of $w_t = 1$ €. This models a coin toss with the chance to win 50% of the wager or lose 40% of the wager. Clearly, the expectation value +5% is seemingly favourable.

722 THORP 2006, p. 36.



(a) Generic classic gamble. **(b)** Generic perturbed gamble.

Figure 8.1: Generic Classic and Generic Perturbed Gamble. Depiction of the generic classic gamble (a) in discrete time and for discrete payouts. The perturbed gamble (b) will be studied in the follow-up project. The fluctuation, ε, appears in the growth factors, but is expectation-neutral.



(a) Classic gamble with specific growth factors. **(b)** Perturbed gamble with specific growth factors.

Figure 8.2: Perturbed Gamble for Specific Growth Factors. Extending **Fig. 8.1**, here depiction of the classic gamble (a) and perturbed gamble (b) with specific growth factors. Perturbed gambles of this kind will be studied in the follow-up project. The numerical simulations always start at the initial wealth $w_t = 1\,€$. The fluctuation, ε, appears in the growth factors and is again expectation-neutral.

Fig. 8.3 contains the results of numerical simulations of the unperturbed and perturbed gambles. The parameters of the perturbed gamble given in **Fig. 8.2** are $r_{\text{up}} = 1.5$, $r_{\text{down}} = 0.6$ and $\varepsilon = 0.3$ and used throughout the numerical simulations. The left column contains graphs of the time-average growth rate in dependence of the leverage for $\ell \in [0, 1]$, which are appropriately scaled to cover the whole codomain. The graphs in right column are zoom-ins for the domain $\ell \in [0, 0.5]$ and codomain $\overline{g}(\ell) \in [0, 0.008]$, because only for a limited range of the domain the curves exhibit positive growth rates. We are ultimately interested only in positive growth rates, because we are looking for optimal growth rates. Gambles with negative time-average growth rates should not be entered at all. The top row contains the respective graphs for the unperturbed gamble **Fig. 8.3a** and **Fig. 8.3b**. The middle row contains the respective graphs for the perturbed gamble **Fig. 8.3c** and **Fig. 8.3d**. Finally, the bottom row contains the combined graphs of the perturbed and unperturbed gamble to demonstrate the extent of the shift of the curves with the additional uncertainty due to the incorporation of the gambler's ignorance in **Fig. 8.3e** and **Fig. 8.3f**.

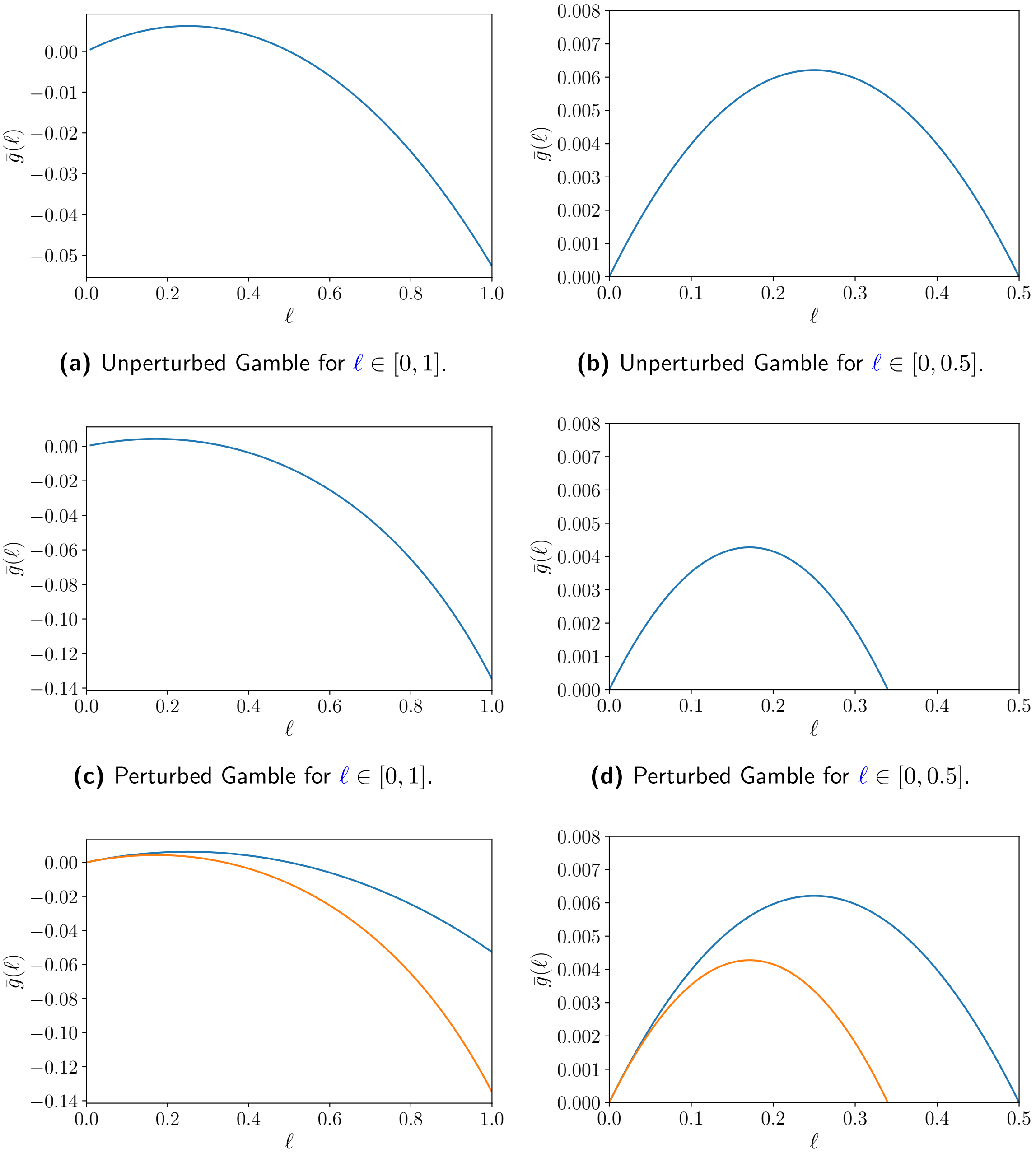
8.4.1 Non-Linearity in the Risk of Overbetting

The risk of overbetting manifest in the longer right tail of the curves in the left column if viewed from their maximum point. *I.e.* the loss in units of the optimal growth rate for deviations from the optimal leverage to the left is smaller than the loss in units of the optimal growth rate for deviations from the optimal leverage to the right. Indeed a deviation to the right can become quickly lethal and lead to the gambler's ruin, when the growth rate becomes negative, thus the term overbetting.

An additional property of the time-average growth rate is visible from **Fig. 8.3e** and **Fig. 8.3f**. The effect of overbetting is non-linear in amount of ignorance. *I.e.* for positive deviations from the optimal leverage have to be paid with increasing losses in units of the maximum growth rate. This expresses the non-linearity of the risk of overbetting.

8.4.2 Long-Term Performance of Different Leverages

Fig. 8.4 contains the wealth trajectories for different investment strategies which are determined by the leverage they choose. We depict six different leverages, $\ell \in \{0.05, 0.11, 0.16, 0.25, 0.3, 0.5\}$ and to better combine the wealth trajectories with their respective time-average growth rate, we added the graph **Fig. 8.3b** again on top as **Fig. 8.4a**. Because we model stochastic multiplicative growth it takes an initial phase for LLN to counteract the noise in the trajectories. After that phase the trajectories settle down on their particular ascent determined by their time-average growth rate in **Fig. 8.4b**. *E.g.* the final wealth for an optimal leverage for the unperturbed gamble is denoted by $\ell_{4,\text{opt}} = 0.25$ is clearly higher than



(a) Unperturbed Gamble for $\ell \in [0, 1]$.

(b) Unperturbed Gamble for $\ell \in [0, 0.5]$.

(c) Perturbed Gamble for $\ell \in [0, 1]$.

(d) Perturbed Gamble for $\ell \in [0, 0.5]$.

(e) Both Gambles for $\ell \in [0, 1]$.

(f) Both Gambles for $\ell \in [0, 0.5]$.

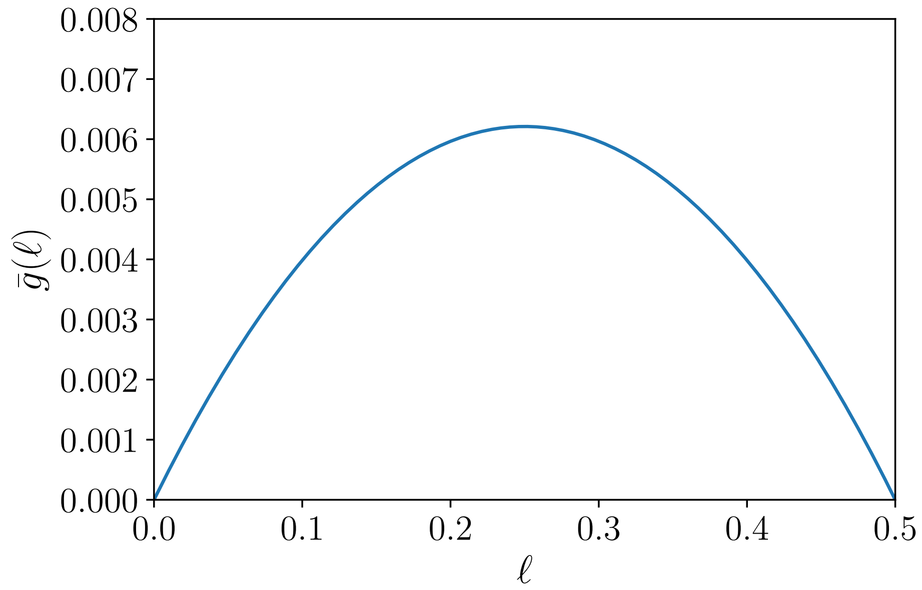
Figure 8.3: Effect of Ignorance about Gamble Parameters. The time-average growth rate $\overline{g}$ depending on the amount of leverage ℓ for the unperturbed gamble in subfigures (8.3a) and (8.3b), for the perturbed gamble in subfigures (8.3c) and (8.3d). Subfigures (8.3e) and (8.3f) combine the unperturbed and the perturbed gamble to demonstrate the effect of ignorance about the exact value of the gamble parameters on the time-average growth rate. Note, the non-linear effect of overbetting in the case with ignorance, which is the more pronounced the greater the leverage. This expresses the non-linearity in the risk of overbetting. The parameters of the perturbed gamble given in **Fig. 8.2** are $r_{\text{up}} = 1.5$, $r_{\text{down}} = 0.6$ and $\varepsilon = 0.3$.

the final wealth of a strategy which uses any other leverage, *e.g.* the optimal leverage of the perturbed gamble $\ell_{3,\text{opt}} = 0.16$. Furthermore, we see that a for example the variability in the wealth trajectory for conservative leverage $\ell_1 = 0.05$ seems to be very low compared to the more aggressive leverage of $\ell_6 = 0.5$, but have to keep in mind the logarithmic vertical axis. However, this observation contains some truth, lower variability in wealth is a consequence of less aggressive leverages. In contrast to the ensemble perspective, high leverages very quickly become lethal if analysed over time and are not advisable. Thus the time perspective unveils the impact of leverage in realistic way.

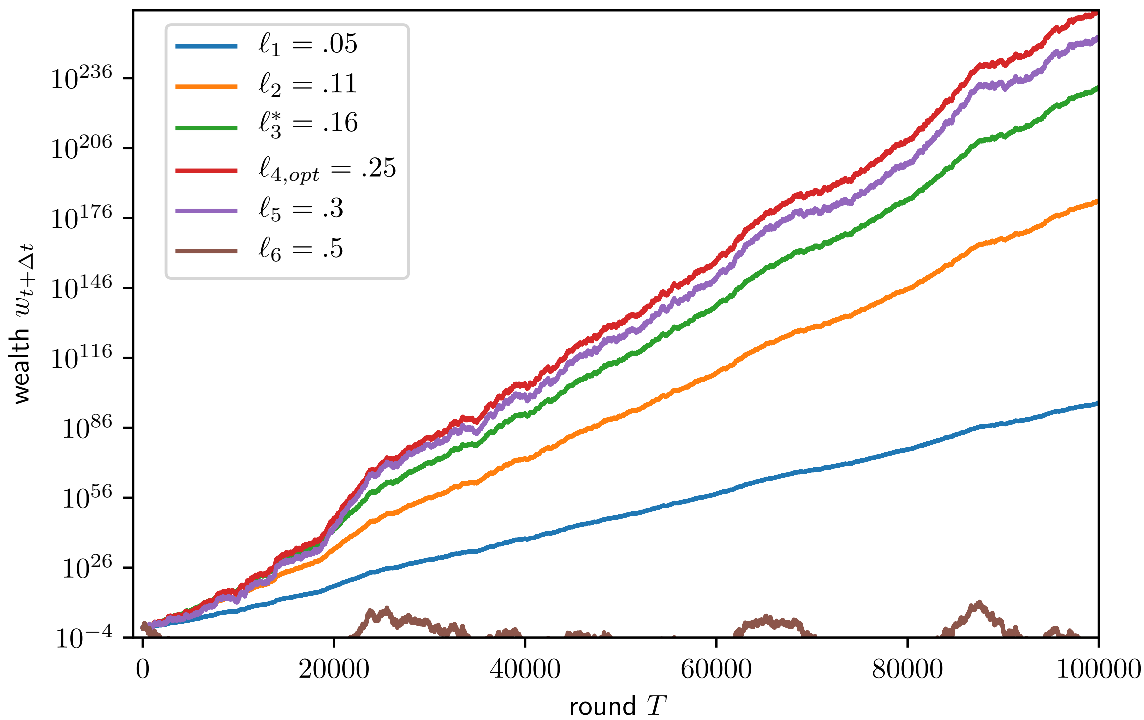
Furthermore, if we compare the trajectory for optimal leverage for the unperturbed gamble $\ell_{4,opt} = 0.25$ again with the trajectory using $\ell_3^* = 0.16$, we see that the unwarranted assumption of some ignorance about the parameter is not that costly even for $\varepsilon = 0.3$ in units of final wealth and certainly not lethal. The time-average growth rate of a leverage of ℓ_5 is zero, which nicely comes out with the used realisation of random numbers.

Similar to before **Fig. 8.5** depicts the wealth trajectories for the perturbed gamble and $\varepsilon = 0.3$. The effect of the additional uncertainty associated with the ignorance of the exact gambling parameters finds its expression in the significantly different scales in **Fig. 8.4b** and **Fig. 8.5b**, which is also contained in the lower possible growth rates according to **Fig. 8.5a** compared to **Fig. 8.4a**. The optimal leverage for the perturbed gamble is now $\ell_{3,\text{opt}}^* = 0.16$ which is clearly lower than the optimal leverage of the unperturbed gamble $\ell_{4,\text{opt}}$, and the respective wealth trajectories perform accordingly. *I.e.* the leverage which is optimal for the unperturbed gamble performs clearly worse than the optimal leverage for this perturbed gamble.

The strategy using the optimal leverage of the unperturbed gamble or, put differently, an overconfident gambler with $\ell_4 = 0.25$ performs significantly worse than *e.g.* an underconfident gambler with $\ell_2 = 0.11$. Again we see only little variability in the still very cautious strategy to invest $\ell_1 = 0.05$ compared to the more aggressive strategy using a leverage of $\ell_5 = 0.3$. From **Fig. 8.5a** we see that for $\overline{g}(\ell_1) \approx \overline{g}(\ell_5)$, which implies similar wealth dynamics. Nevertheless the more conservative leverage outperforms the more aggressive leverage in this realisation of random numbers, because of some fluctuations in the first 30 000 rounds, after that the ascents are similar which corresponds to their associated growth rates. The strategy using a leverage of $\ell_6 = 0.5$ is so aggressive that overbetting occurs almost immediately and a gambler with this strategy is ruined so quickly that the respective trajectory is hardly recognisable in **Fig. 8.5b**. From **Fig. 8.5a** we see that the growth rate for ℓ_6 is negative, thus the respective wealth trajectory is almost invisible in the figure.

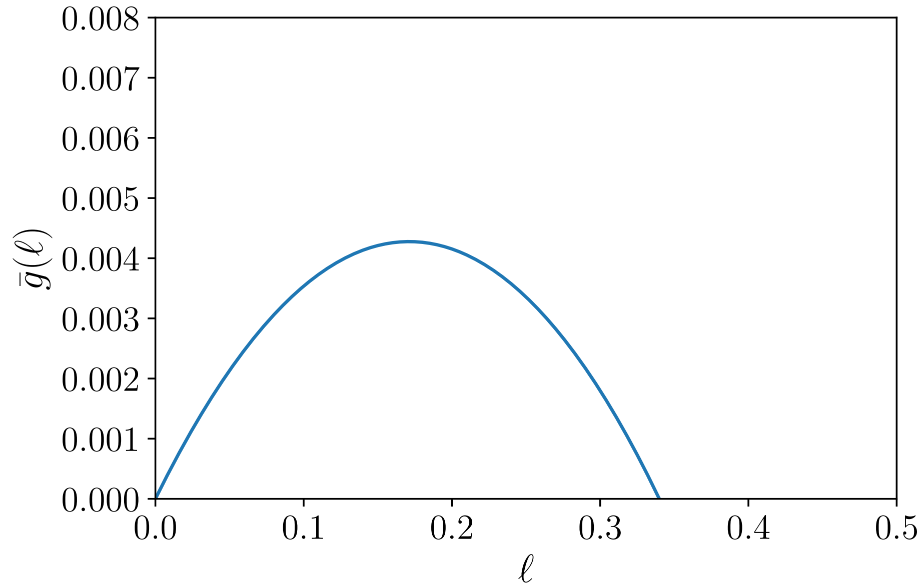


(a) Time-average growth rate as a function of the leverage.

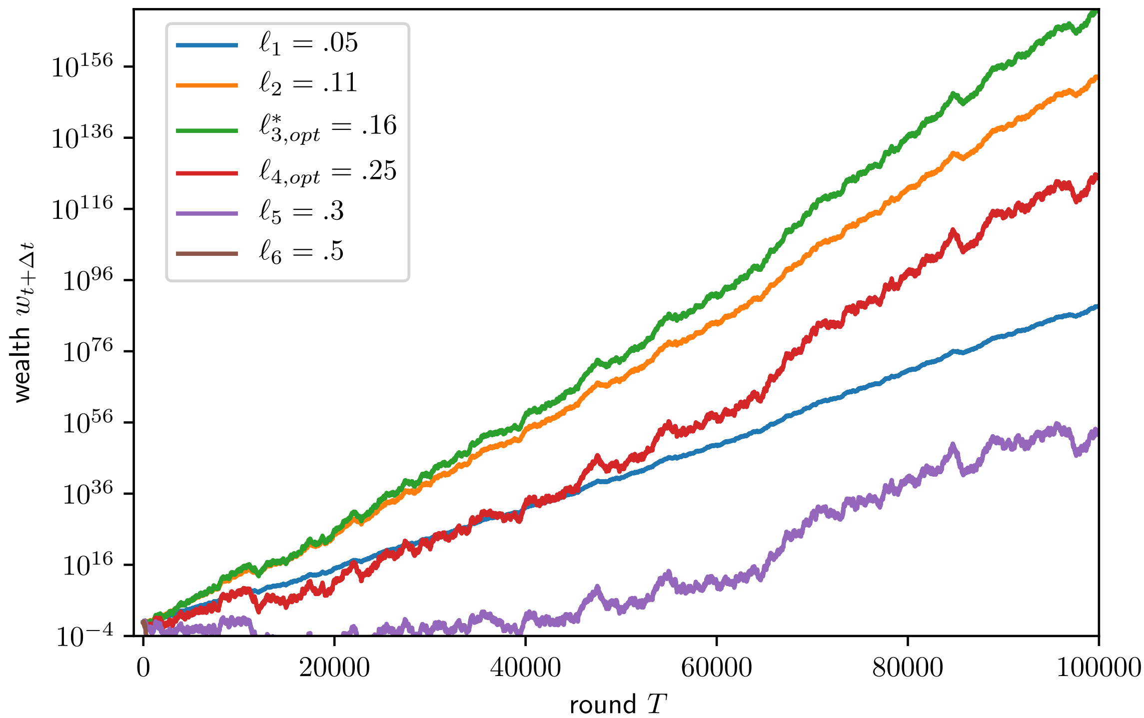


(b) Wealth trajectories for different leverages.

Figure 8.4: Wealth Trajectories for Different Leverages in the Unperturbed Gamble. **Fig. 8.4b** depicts wealth trajectories for different leverages in the context of the unperturbed gamble, whose time-average growth rates are given in **Fig. 8.4a**.



(a) Time-average growth rate as a function of the leverage for the perturbed gamble.



(b) Wealth trajectories for different leverages of the perturbed gamble.

Figure 8.5: Wealth Trajectories for Different Leverages in the Perturbed Gamble. **Fig. 8.5b** depicts wealth trajectories for different leverages in the context of the perturbed gamble, whose time-average growth rates are given in **Fig. 8.5a**.

Outlook

The numerical simulations we present in **Fig. 8.4** and **Fig. 8.5** are promising and show results consistent with *Ergodicity Economics*. The issue of overbetting and prudent underbetting is of course known in the literature: 'The reason is that using too large an [ℓ_{opt}] and overbetting is much more severely penalized than using too small an [ℓ_{opt}] and underbetting.'[723] However, the issue is not convincingly solved. Our main future objective will be to derive an analytical expression for the decrease in the optimal time-average growth rate of wealth depending on the amount of ignorance a gambler attaches to his estimation of the gamble parameters, $\overline{g}\,(\ell, \varepsilon)$. However, the formula involved are more complex, but entail a surprising result which we only insinuate. The full derivation of the results will require some more months.

[723] Thorp 2006, p. 18.

9 Conclusion

In my opinion we are in a dangerous age of overspecialization. To me the danger of this period is not primarily that we are studying very special problems that the development of science has forced us to go into, but rather that we are in great danger of finding our outlook so limited that we may fail to see the bearing of important ideas because they have been formulated in what our organization of science has decreed to be alien territory.[724]

Norbert Wiener

9.1 A Look in the Rear View Mirror

The identification of the time perspective by Peters (2011c,b) have opened the door to understand the economics of uncertainty in a novel way. It enables us to systematise the history of probability theory and mathematical economics along the ensemble perspective and the time perspective. We use this categorisation in **Fig. 9.1** to summarise the story that is told in this thesis.

It starts with the beginnings of probability theory which grew out of problems in commerce. Soon the concept of the expectation value was created. Equipped with a momentous connotation the concept stuck and with it the embedding of randomness in the ensemble. It took until the second half of the 19$^{\text{th}}$ century that in the works of Boltzmann this conceptualisation of randomness is questioned. Since then the development of tools suitable for the time perspective exploded in mathematics and physics. In economics, however, this process should not begin before the seminal publication of Peters (2011c), although many influential scientists could have become an intellectual bridge between the two perspectives. It is not unlikely that the pendulum could have swung in the other direction much earlier. However, this was not the case and mathematical economics developed in a direction very remote from the time

724 Wiener 1956b, p. 48

perspective. Many of the scholars which appear under the umbrella of the time perspective in **Fig. 9.1** lived a maverick life of an underappreciated scholar. We hope that this thesis does not follow in this footsteps.

Beginnings of Probability Theory

Fibonacci 1202
Pacioli 1494
Cardano 1526

Fermat & Pascal (de Méré) 1654
Huygens 1657

Ensemble Perspective

J. Bernoulli, de Moivre 1713, 1718
Maxwell 1867
Gibbs 1902

Mathematical Economics

C. Menger 1871
Walras 1874
Jevons 1888
Fisher 1892
K. Menger 1934
v. Neumann & Morgenstern 1944
Samuelson 1947
Markowitz 1952
Arrow & Debreu 1954
Black, Scholes & Merton 1973

D. Bernoulli 1738
Laplace 1812
Bachelier 1900
Einstein 1905; 1906
E. B. Wilson 1920s
Tinbergen 1939
v. Neumann 1932b

Euler ~1740
Whitworth 1870

Time Perspective

Boltzmann 1872; 1884
P. & T. Ehrenfest 1911
G. D. Birkhoff 1931b
Wiener 1923; 1939; 1964a
Kolmogorov 1933
Lévy 1937
Itô 1944
Shannon, Kelly 1956
Mandelbrot 1963; 1967
Thorp 1973
Cover 1991; 2006
Taleb 2010a; 2014

Ergodicity Economics

Peters 2011c; 2011b
Peters & Adamou 2011c; 2018b; 2018c
Peters & Gell-Mann 2016

Figure 9.1: History of thought of ensemble thinking and thinking in the time domain. We depict the intellectual lineage from the founders of probability theory, statistical mechanics and mathematical economics to the founders ergodicity economics. See also Tretyakov (2013, ch. 1).

To understand the process of formalisation and axiomatisation in economics after 1945 the relationship between the work of Paul Samuelson and Edwin Bidwell Wilson turns out to be crucial. The evidence of Wilson's influence on Samuelson was mentioned by Samuelson himself many times, and is also worked out by several historians of economics and economists close to him such as von Weizsäcker (1997, p. 92): 'Samuelson ist ganz wesentlich verantwortlich für diese methodische Wende im Fach. Er wurde in seinen jungen Jahren von bedeutenden Physikern beeinflußt und hat sich zum Beispiel in seinen *Foundations* die präzise Begrifflichkeit der Physik zum Vorbild genommen.' Apart from this and the historical fact that Samuelsonian economics became mainstream, there is much critical literature on the methodological and metaphorical loaning.[725]

The direction of the mathematisation of economics after 1945 (as described among others in Weintraub (2002)) is unthinkable without the tacit assumption of ergodicity underlying almost all parts of textbook economics. The realisation of the relevance of non-ergodicity as the most important model property of a form of economics, that includes the explanation of novelty and evolutionary change, challenges the dominant line of thought that mathematical economics follows since 1945. One reason for the adoption of this line of thought is the fact, that economic methodology draw unreflecting on statistical mechanics, In the introduction to his *Cybernetics* Wiener (1985, p. 4) reminds us, that cybernetic ideas 'were all very much in the spirit of the thought of the time, and I do not for a moment wish to claim anything like the sole responsibility for their introduction.' What is true for the advent of cybernetics is also valid for the use of methods from the natural sciences in economics, especially physics. The demand to put economics on more scientific footing lay in the zeitgeist. Still, Samuelson did ingeniously excel in this movement with his seminal dissertation *Foundations of Economic Analysis* and soon after it with his famous textbook *Economics*. Nevertheless, Samuelson (1968) was well aware of the crucial assumption of ergodicity underlying the whole movement of the increasing mathematisation of economics, from today's point of view, he may well have underestimated its criticality to the whole enterprise.[726] Regrettably, this knowledge got lost through the years, despite its paramount importance, to a large part because it was rarely mentioned at all and clouded by silent rhetoric, never thoroughly brought into light. This is what this thesis aims at.

In his Nobel prize lecture the physicist Eugene P. Wigner defines the objective of physics as the explanation of (inanimate) nature. By explanation Wigner thinks of an understanding of nature which produces no surprises. Over the course of this thesis it became clear, economics is the science of a domain which is fundamentally driven by surprises.

[725] First and foremost see the work of Philip Mirowski (Mirowski 1989, 1991, 2013).
[726] Or as Samuelson (1952) suggests, he is simply an honest believer.

9.2 Ergodicity Economics – Not Just Another Physics Analogy

> [O]nce Pandora's box was opened, there seemed to be no stopping the spread of probabilistic concepts in physics. Probabilistic gas theory led to statistical mechanics, which begat quantum mechanics, which begat probability waves, which begat nonergodic and weakly stable systems [...] all of which begat a seemingly accidental universe.[727]
>
> Philip Mirowski

The general spirit of the thesis at hand is strongly inspired by the beauty of mathematics in the sciences and especially in physics, but in full awareness and admiration of *e.g.* Philip Mirowski's famous critiques in *More Heat than Light. Economics as Social Physics, Physics as Nature's Economics* and *Machine Dreams. Economics Becomes a Cyborg Science.*[728] It is for the following two reasons, that we think it is worthwhile to study *Ergodicity Economics*, which is admittedly again a concept that originated from physics. First, Mirowski (1991) deals mainly with the energy analogy of utility in neoclassical economics and derives a critique of neoclassical micro- and macroeconomics. The term 'ergodic' surfaces only two times and marginally in Mirowski (1991) and only four times in Mirowski (2002). The specific aspects of non-ergodicity, time averages, temporal optimisation and the general time perspective which manifest in the embedding of randomness within time are not discussed, but are our main focus. Therefore, this thesis and *Ergodicity Economics* has a much wider scope and is not restricted to any particular ideological school of thought in economics. The approach of temporal optimisation applies to all economic problems because *Ergodicity Economics* incorporates the dynamic of the decision environment explicitly and is thus able to cope with all kinds of dynamics, see *e.g.* Peters and Adamou (2018c). Second, since the time of Mirowski's publications science gained a deeper understanding of the relation between ergodic theory and the theories of dynamical systems and stochastic processes. Besides that, there is a strong need for clarification of some so far misunderstood implications of the logarithm and non-ergodicity.

The seminal papers by Shannon (1948) and Jaynes (1957a,b) mark cornerstones in the foundation of information theory and revealed that the statistical approach and the concept of entropy are deeper and quite independent from statistical mechanics, where these concepts had been simply found and formulated for the first time:

[727] Mirowski 1991, p. 64.
[728] Mirowski 1991, 2002.

> “[S]tatistical mechanics need not be regarded as a physical theory dependent for its validity on the truth of additional assumptions not contained in the laws of mechanics (such as ergodicity, metric transitivity, equal a priori probabilities, *etc.*).
> [. . .]
> The principles and mathematical methods of statistical mechanics are seen to be of much more general applicability than conventional arguments would lead one to suppose. In the problem of prediction, the maximization of entropy is not an application of a law of physics, but merely a method of reasoning which ensures that no unconscious arbitrary assumptions have been introduced.[729]”

Jaynes draws connections between many subjects that are the focus of this thesis: the Bayesian approach to subjective inference, ergodicity, probabilities as a time-average quantity[730] and a subjective form of the H-theorem.[731] In particular, he is the major advocate for the MEP, which has a wide range of applicability because it constitutes a very general inferential procedure:

> “Mathematically, the maximum-entropy distribution has the important property that no possibility is ignored; it assigns positive weight to every situation that is not absolutely excluded by the given information. This is quite similar in effect to an ergodic property.”[732]

And he expands on the relation between the ergodicity problem and statistical inference,

> “This would not be the case, however, unless we also supplemented our prediction problem with new experimental data which provided us with some information as to the likely values of these new constants. Even if we had a dear proof that a system is not metrically transitive, we would still have no rational basis for excluding any region of phase space that is allowed by the information available to us. In its effect on our ultimate predictions, this fact is equivalent to an ergodic hypothesis, quite independently of whether physical systems are in fact ergodic.”[733]

> “The essential point in the arguments presented above is that we accept the von-Neumann-Shannon expression for entropy, very literally, as a measure of the amount of uncertainty represented by a probability distribution; thus entropy becomes the primitive concept with which we work, more fundamental even than energy. If in addition we reinterpret the prediction problem of statistical mechanics in the subjective sense, we can derive the usual relations in a very

729 Jaynes 1957a, p. 620, pp. 629.
730 Jaynes 1957b, part 15, pp. 183.
731 Jaynes 1957b, p. 179.
732 Jaynes 1957a, p. 623.
733 Jaynes 1957a, p. 624.

> elementary way without any consideration of ensembles or appeal to the usual arguments concerning ergodicity or equal a priori probabilities. The principles and mathematical methods of statistical mechanics are seen to be of much more general applicability than conventional arguments would lead one to suppose. In the problem of prediction, the maximization of entropy is not an application of a law of physics, but merely a method of reasoning which ensures that no unconscious arbitrary assumptions have been introduced."[734]

Let us summarise the line of argument of the preceding section in order to hammer home the important message: The much wider range of applicability of concepts that only originated in certain disciplines, but are by no means confined to them are the fruits the interdisciplinarily interested scientist hopes to harvest and show around. This is the case for the conceptual richness of (non-)ergodicity and information theory for economics,[735] that were simply discovered originally in statistical mechanics, but apply to much more than atoms and physics. A deeper understanding of their origin and their relevance *e.g.* for economics is what this thesis is striving for.

9.3 A Look Ahead

> Je suis persuadé que la seule épreuve décisive pour la fécundité d'idées ou d'une vision nouvelles est celle du temps. La fécondité se reconnait par la progéniture, et non par les honneurs.
>
> ALEXANDER GROTHENDIECK

We see this thesis in combination with the already existing results from *Ergodicity Economics* having the potential to motivate further fascinating research in a number of fields. First and foremost we provided an outlook on future research we will be occupied with ourselves. This follow-up project will further study the time perspective of fractional KELLY strategies and incorporate the gambler's ignorance about the parameters which define the gamble. The prospective results could lead to a move towards a BAYESIAN analysis. But already now the connection to puzzles in behavioural economics is looming and a further scrutinisation of (optimal) leverage seems promising. The contribution of empirical investigations just recently could become an important convincing argument.[736]

734 JAYNES 1957a, pp. 629.
735 PETERS and ADAMOU 2018b; ZHOU et al. 2013; YANG 2017.
736 BERMAN et al. 2016, 2017.

Furthermore, new experimental tests are just looming on the horizon. The new field of neuroeconomics tries to combine complementary insights from economics, psychology and neuroscience into a coherent whole. Expositions of it often start with a motivation based on the St. Petersburg paradox.[737] We could read a preprint of a study, which is testing the behaviour of individuals when confronted with changing dynamics done by a group of neuroscientists in Copenhagen.[738] The preliminary results seem to be consistent with *Ergodicity Economics* predictions and reject most other decision theories. Personally, the expansion of our skill set in the domain of stochastic calculus will be one of the next steps. The exposition in this thesis is mostly limited to discrete time. Interesting extensions can be made to the analysis of continuous stochastic processes and lead to the field of stochastic finance.[739]

Finally, the reader has reached the end of this monograph, so it is advisable to close on a lighter note with the words by Yakov G. Sinai one of the leaders in the field of ergodic theory:

> “I have a very good example for an ergodic system which I always explain to my students. Suppose you want to buy a pair of shoes and you live in a house that has a shoe store. There are two different strategies: one is that you go to the store in your house every day to check out the shoes and eventually you find the best pair; another is to take your car and to spend a whole day searching for footwear all over town to find a place where they have the best shoes and you buy them immediately. The system is ergodic if the result of these two strategies is the same.”[740]

[737] Glimcher and Rustichini 2004, p. 447.
[738] Meder et al. 2019; Peters 2018b.
[739] Fernholz 2002; Föllmer and Schied 2016.
[740] Raussen and Skau 2015, p. 154.

Bibliography

Aaronson, Jon (1997). *An Introduction to Infinite Ergodic Theory.* American Mathematical Society, Providence (77).

Aarseth, Espen J. (1997). *Cybertext. Perspectives on Ergodic Literature.* Johns Hopkins University Press (199).

Ainslie, George (1975). 'Specious reward: A behavioral theory of impulsiveness and impulse control.' In: *Psychological Bulletin* 82 (4), 463–496. doi: 10.1037/h0076860 (256).

Ait-Sahalia, Yacine and Lars P. Hansen, eds. (2010). *Handbook of Financial Econometrics. Tools and Techniques.* Vol. 1. North-Holland, Amsterdam. doi: 10.1016/C2009-0-16945-3 (175).

Aitchison, John and James A. C. Brown (1963). *The Lognormal Distribution. With Special Reference to Its Uses in Economics.* Cambridge University Press (145).

Alaoglu, Leonidas and George D. Birkhoff (1939). 'General Ergodic Theorems'. In: *Proceedings of the National Academy of Sciences* 25 (12), 628 (219).

Albeverio, Sergio, Volker Jentsch and Holger Kantz (2006). *Extreme Events in Nature and Society.* Springer, Berlin. doi: 10.1007/3-540-28611-X (90).

Alchian, Armen A. (1950). 'Uncertainty, Evolution, and Economic Theory'. In: *Journal of Political Economy* 58 (3), 211–221. JSTOR: 1827159 (179, 213, 214).

Algoet, Paul H. and Thomas M. Cover (1988). 'A Sandwich Proof of the Shannon-McMillan-Breiman Theorem'. In: *Annals of Probability* 16 (2), 899–909. doi: 10.1214/aop/1176991794 (78).

Al Khwarizmi, Muhammad ibn Musa (1915). *The Algebra of Al-Khowarizmi. Robert of Chester's Latin Translation of the Algebra of Al-Khowarizmi: With an Introduction, Critical Notes and an English Version.* Ed. by Louis C. Karpinski. Macmillan, London. url: https://archive.org/details/robertofchesters00khuw/page/n11 (83).

Allais, Maurice (1953). 'Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine'. In: *Econometrica* 21 (4), 503–546. doi: 10.2307/1907921 (255).

Altmann, Eduardo G. (2007). 'Intermittent Chaos in Hamiltonian Dynamical Systems'. PhD thesis. Bergische Universität Wuppertal. url: http://elpub.bib.uni-wuppertal.de/edocs/dokumente/fbc/physik/diss2007/altmann (74).

Amable, Bruno, Jérôme Henry, Frédéric Lordon and Richard Topol (1994). 'Strong hysteresis versus zero-root dynamics'. In: *Economics Letters* 44 (1-2), 43–47. doi: 10.1016/0165-1765(93)00300-d (198).

Ambrose, Warren, Paul R. Halmos and Shizuo Kakutani (1942). 'The Decomposition of Measures, II'. In: *Duke Mathematical

Journal 9 (1), 43–47. DOI: 10.1215/s0012-7094-42-00905-0 (78).

Amir, Rabah, Igor V. Evstigneev, Thorsten Hens and Klaus R. Schenk-Hoppé (2005). 'Market selection and survival of investment strategies'. In: *Journal of Mathematical Economics* 41 (1-2), 105–122. DOI: 10.1016/j.jmateco.2003.10.006 (179).

Amir, Rabah, Igor V. Evstigneev, Thorsten Hens and Le Xu (2011). 'Evolutionary finance and dynamic games'. In: *Mathematics and Financial Economics* 5 (3), 161–184. DOI: 10.1007/s11579-011-0053-2 (179).

Amir, Rabah, Igor V. Evstigneev and Klaus R. Schenk-Hoppé (2012). 'Asset market games of survival: a synthesis of evolutionary and dynamic games'. In: *Annals of Finance* 9 (2), 121–144. DOI: 10.1007/s10436-012-0210-5 (179).

Angrist, Joshua D. and Jörn-Steffen Pischke (2010). 'The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics'. In: *Journal of Economic Perspectives* 24 (2), 3–30. DOI: 10.1257/jep.24.2.3 (184, 187).

Anosov, Dmitry V. (1967). 'Geodesic flows on closed Riemannian manifolds of negative curvature'. In: *Proceedings of the Steklov Institute of Mathematics* 90, 1–235. URL: http://mi.mathnet.ru/eng/tm/v90/p3 (74, 75).

Anosov, Dmitry V. and Yakov G. Sinai (1967). 'Some smooth ergodic systems'. In: *Russian Mathematical Surveys* 22 (5), 103–167. DOI: 10.1070/rm1967v022n05abeh001228 (74, 75).

Araujo, Aloisio P. and Wilfredo L. Araujo Aloisio P.Maldonado (2000). 'Ergodic chaos, learning and sunspot equilibrium'. In: *Economic Theory* 15 (1), 163–184. DOI: 10.1007/s001990050005 (74).

Arkes, Hal R., Gerd Gigerenzer and Ralph Hertwig (2016). 'How bad is incoherence?' In: *Decision* 3 (1), 20–39. DOI: 10.1037/dec0000043 (248).

Arnold, Ludwig (1998). *Random Dynamical Systems*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-12878-7 (77, 78).

Arnol'd, Vladimir I. (1963a). 'Proof of a theorem of A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian'. (in Russian). In: *Russian Mathematical Surveys* 18 (5), 9–36. DOI: 10.1070/rm1963v018n05abeh004130 (75).

Arnol'd, Vladimir I. (1963b). 'Small denominators and problems of stability of motion in classical and celestial mechanics'. In: *Russian Mathematical Surveys* 18 (6), 85–191. DOI: 10.1070/RM1963v018n06ABEH001143 (75).

Arnol'd, Vladimir I. and André Avez (1968). *Ergodic Problems of Classical Mechanics*. W. A. Benjamin, New York (77).

Arrow, Kenneth J. (1951). 'Alternative Approaches to the Theory of Choice in Risk-Taking Situations'. In: *Econometrica* 19 (4), 404–437. DOI: 10.2307/1907465 (100, 139).

Arrow, Kenneth J. (1965). 'The Theory of Risk Aversion'. In: *Aspects of The Theory of Risk-Bearing*. Yrjö Jahnsson Lectures. Yrjö Jahnssonin Säätiö, Helsinki, 28–44 (133).

Arrow, Kenneth J. and Gérard Debreu (1954). 'Existence of an Equilibrium for a Competitive Economy'. In: *Econometrica* 22 (3), 265–290. JSTOR: 1907353 (180, 227, 228, 278).

Arrow, Kenneth J. and Frank Hahn (1971). *Competitive Equilibrium Analysis* (180).

Arthur, W. Brian (1989). 'Competing Technologies, Increasing Returns, and Lock-In by Historical Events'. In: *Economic Journal* 99 (394), 116–131. JSTOR: 2234208 (182).

Arthur, W. Brian (1999). 'Complexity and the economy'. In: *Science* 284 (5411), 107–109. doi: 10.1126/science.284.5411.107 (251).

Arthur, W. Brian (2010). 'Complexity, The Santa Fe Approach, and Non-Equilibrium Economics'. In: *History of Economic Ideas* 18 (2), 149–166. doi: 10.1400/148277 (251).

Arthur, W. Brian (2013). 'Complexity Economics: A Different Framework for Economic Thought'. In: *SFI Working Paper*. url: http://www.santafe.edu/research/working-papers/abstract/36df2f7d8ecd8941d8fab92ded2c4547/ (43, 251).

Arthur, W. Brian (2015). *Complexity and the Economy*. Oxford University Press (43, 251).

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1983). 'A generalized Urn problem and its applications'. Originally published in Kibernetika Issue 1 pp. 49–56, January–February 1983 in Russian). In: *Cybernetics* 19 (1), 61–71. doi: 10.1007/bf01070110 (182).

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1986). 'Strong laws for a class of path-dependent stochastic processes with applications'. In: *Stochastic Optimization. Proceedings of the International Conference, Kiev, 1984*. Ed. by Vadim I. Arkin, Albert N. Shiraev and Roger J.-B. Wets. Springer, Berlin, 287–300. doi: 10.1007/bfb0007105 (182).

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1987a). 'Limit Theorems for Proportions of Balls in a Generalized Urn Scheme'. In: *IIASA Working Paper* WP-87-111. url: http://pure.iiasa.ac.at/id/eprint/2941/ (182).

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1987b). 'Non-Linear Urn Processes: Asymptotic Behavior and Applications'. In: *IIASA Working Paper* WP-87-085. url: http://pure.iiasa.ac.at/id/eprint/2967/ (182).

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1987c). 'Path-dependent processes and the emergence of macro-structure'. In: *European Journal of Operational Research* 30 (3), 294–303. doi: 10.1016/0377-2217(87)90074-9 (182).

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1988). 'Nonlinear Adaptive Processes of Growth with General Increments: Attainable and Unattainable Components of Terminal Set'. In: *IIASA Working Paper* WP-88-086. url: http://pure.iiasa.ac.at/id/eprint/3119/ (182).

Ashley, Steven (2015). 'Core Concept: Ergodic theory plays a key role in multiple fields'. In: *Proceedings of the National Academy of Sciences* 112 (7), 1914. doi: 10.1073/pnas.1500429112 (219).

Aumann, Robert J. (1977). 'The St. Petersburg Paradox: A Discussion of Some Recent Comments'. In: *Journal of Economic Theory* 14 (2), 443–445. doi: 10.1016/0022-0531(77)90143-0 (100, 112, 139).

Austin, Tim, Tanja Eisner and Terence Tao (2011). 'Nonconventional ergodic averages and multiple recurrence for von Neumann dynamical systems'. In: *Pacific Journal of Mathematics* 250 (1), 1–60. doi: 10.2140/pjm.2011.250.1 (78).

Auyang, Sunny Y. (1998). *Foundations of Complex-system Theories: in Economics, Evolutionary Biology, and Statistical Physics*. Cambridge University Press, Cambridge UK (52).

Axelrod, Robert and William D. Hamilton (1981). 'The Evolution of Cooperation'. In: *Science* 211 (4489), 1390–1396. doi: 10.1126/science.7466396 (213).

Bachelier, Louis (1900). 'Théorie de la speculation'. In: *Annales scientifiques de l'École Normale Supérieure* Série 3. (17), 21–86. doi: 10.24033/asens.476 (90, 278).

Badino, Massimiliano (2006). 'The Foundational Role of Ergodic Theory'. In: *Foundations of Science* 11 (4), 323–347. doi: 10.1007/s10699-005-6283-0 (71).

Badino, Massimiliano (2011). 'Mechanistic slumber vs. statistical insomnia: the early history of Boltzmann's H-theorem (1868-1877)'. In: *European Physical Journal H: Historical Perspectives on Contemporary Physics* 36 (3), 353–378. doi: 10.1140/epjh/e2011-10048-5 (38, 41, 70).

Bagchi, Debarshee and Constantino Tsallis (2018). 'Fermi-Pasta-Ulam-Tsingou problems: Passage from Boltzmann to q-statistics'. In: *Physica A: Statistical Mechanics and its Applications* 491, 869–873. doi: 10.1016/j.physa.2017.09.098 (73).

Böhm, Volker and Carl Chiarella (2005). 'Mean variance preferences, expectations formations, and the dynamics of random asset prices'. In: *Mathematical Finance* 15.1, 61–97. doi: 10.1111/j.0960-1627.2005.00211.x (177).

Bais, F. Alexander and J. Doyne Farmer (2008). 'The Physics of Information'. In: *Philosophy of Information*. Ed. by Pieter Adriaans and Johan van Benthem. Handbook of the Philosophy of Science, Vol. 8, 617–691. doi: 10.1016/B978-0-444-51726-5.50020-0 (42, 75, 204).

Barenblatt, Grigory I. (2003). *Scaling*. Cambridge University Press (119, 160).

Barkai, Eli, Yuval Garini and Ralf Metzler (2012). 'Strange kinetics of single molecules in living cells'. In: *Physics Today* 65 (8), 29–35. doi: 10.1063/pt.3.1677 (208).

Barnett, William A. (2007). 'An Interview with Paul A. Samuelson. Conversations with Eminent Economists'. In: *Inside the Economist's Mind*. Ed. by Paul A. Samuelson and William A. Barnett. Blackwell Publishing, Malden. Chap. 7, 143–164. doi: 10.1017/S1365100504040039 (220).

Barreira, Luis and Yakov Pesin (2013). *Introduction to Smooth Ergodic Theory* (77).

Barron, Andrew R. (1985). 'The Strong Ergodic Theorem for Densities: Generalized Shannon-McMillan-Breiman Theorem'. In: *Annals of Probability* 13 (4), 1292–1303. doi: 10.1214/aop/1176992813 (78).

Basawa, Ishwar V. and David John Scott (1983). *Asymptotic Optimal Inference for Non-ergodic Models*. Springer New York. doi: 10.1007/978-1-4612-5505-5 (194).

Batterman, Robert W. (1998). 'Why Equilibrium Statistical Mechanics Works: Universality and the Renormalization Group'. In: *Philosophy of Science* 65 (2), 183–208. doi: 10.1086/392634 (71).

Bauer, Thomas K., Gerd Gigerenzer and Walter Krämer (2015). *Wahr oder wahrscheinlich? Über Risiken und Nebenwirkungen der Unstatistik*. Bundeszentrale für politische Bildung, Bonn (161).

Beaumont, Hubertus J. E., Jenna Gallie, Christian Kost, Gayle C. Ferguson and Paul B. Rainey (2009). 'Experimental evolution of bet hedging'. In: *Nature* 462 (7269), 90–93. doi: 10.1038/nature08504 (214).

Beck, Christian (2009). 'Generalised information and entropy measures in physics'. In: *Contemporary Physics* 50 (4), 495–510. doi: 10.1080/00107510902823517 (203).

Becker, Gary S. (1976). *The Economic Approach to Human Behavior*. University of Chicago Press (249).

BECKER, GARY S. and KEVIN M. MURPHY (1988). 'A Theory of Rational Addiction'. In: *Journal of Political Economy* 96 (4), 675–700. DOI: 10.1086/261558 (249).

BELKE, ANSGAR H., MATTHIAS GÖCKE and LAURA WERNER (2014). 'Hysteresis Effects in Economics Different Methods for Describing Economic Path-Dependence'. In: DOI: 10.2139/ssrn.2423282 (198).

BELL, ROBERT M. and THOMAS M. COVER (1980). 'Competitive Optimality of Logarithmic Investment'. In: *Mathematics of Operations Research* 5 (2), 161–166. DOI: 10.1287/moor.5.2.161 (164, 213).

BENEDICKS, MICHAEL and LENNART CARLESON (1985). 'On Iterations of $1 - ax^2$ on $(-1, 1)$'. In: *Annals of Mathematics* 122 (1), 1–25. DOI: 10.2307/1971367 (73).

BERG, NATHAN and GERD GIGERENZER (2010). 'As-If Behavioral Economics: Neoclassical Economics in Disguise?' In: *History of Economic Ideas* 18 (1), 133–166. DOI: 10.1400/140334 (252, 253).

BERGELSON, VITALY (2004). 'Some Historical Remarks and Modern Questions Around the Ergodic Theorem'. In: *Dynamics of Complex Systems*. URL: http://hdl.handle.net/2433/26078 (63).

BERGER, JAY M. and BENOÎT B. MANDELBROT (1963). 'A New Model for Error Clustering in Telephone Circuits'. ngerman. In: *IBM Journal of Research and Development* 7 (3), 224–236 (278).

BERMAN, YONATAN, OLE PETERS and ALEXANDER ADAMOU (2016). 'Far from equilibrium: Wealth reallocation in the United States'. In: arXiv: 1605.05631 (169, 234, 282).

BERMAN, YONATAN, OLE PETERS and ALEXANDER ADAMOU (2017). 'An Empirical Test of the Ergodic Hypothesis: Wealth Distributions in the United States'. In: URL: https://ssrn.com/abstract=2794830 (169, 234, 282).

BERNOULLI, DANIEL (1738). 'Specimen Theoriae Novae de Mensura Sortis'. In: *Commentarii Academiae Scientiarum Imperialis Petropolitanae Tomus V (Papers of the Imperial Academy of Sciences in Petersburg)* V (Ad Annos 1730 et 1731). German translation in BERNOULLI (1747, 1896), English translation in BERNOULLI (1954), 175–192. URL: bit.ly/MKBernoulliOriginal1738 (21, 85, 100, 114–116, 122, 124–126, 140, 168, 256, 278, 289).

BERNOULLI, DANIEL (1747). 'III. Auszug aus dem Versuch einer neuen Lehre, von dem Maaße der Glücksspiele. aus den Schriften der petersburgischen Akademie der Wissenschaften, 5 Band, 175 S. genommen'. In: *Hamburgisches Magazin oder gesammelte Schriften. zum Unterricht und Vergnügen, aus der Naturforschung und den angenehmen Wissenschaften überhaupt.* Vol. 1. This is the first published translation of the Latin original BERNOULLI (1738) into German. Chap. 5, 73–90. URL: bit.ly/MKBernoulli1747 (289).

BERNOULLI, DANIEL (1896). 'Versuch einer neuen Theorie der Wertbestimmung von Glücksfällen'. In: *Die Grundlage der modernen Wertlehre.* Ed. by ALFRED PRINGSHEIM. Brentano und Leser, Sammlung älterer und neuerer staatswissenschaftlicher Schriften des In- und Auslandes, Nr. 9. This is the second published translation of the Latin original BERNOULLI (1738) into German by ALFRED PRINGSHEIM with added remarks and explanations. Importantly, the figure on p. 32 is not contained in the original and is added for the first time in this version. Duncker & Humblot, Leipzig, 22–57. URL: bit.ly/MKBernoulli1896 (114, 115, 289, 290).

BERNOULLI, DANIEL (1954). 'Exposition of a New Theory on the Measurement of Risk'. In: *Eco-*

nometrica 22 (1). Translated by Louise Sommers with the support of Karl Menger. The figure from Bernoulli (1896, p. 32) is reproduced in this publication on p. 26., 23–36. doi: 10.2307/1909829 (84, 85, 114–116, 120, 123, 125, 129, 131, 289).

Bernoulli, Jakob (1713). *Ars Conjectandi.* Ed. by Nikolaus Bernoulli. Impensis Thurnisiorum fratrum, Basileae (85, 100, 104, 140, 278, 290).

Bernoulli, Jakob (1899). *Wahrscheinlichkeitsrechnung (Ars Conjectandi).* Ed. and trans. by Robert Haussner. Translation of Bernoulli (1713). Wilhelm von Engelmann, Leipzig (104).

Bernoulli, Jakob (1975). *Die Werke von Jakob Bernoulli. Bd. 3: Wahrscheinlichkeitsrechnung.* Ed. by Naturforschenden Gesellschaft in Basel and Bartel L. van der Waerden. Birkhäuser, Basel (84, 85, 108, 110, 131).

Billingsley, Patrick (1965). *Ergodic Theory and Information.* Wiley, New Academic Press, New York (77).

Billingsley, Patrick (1999). *Convergence of Probability Measures.* Wiley, New York (78).

Birkhoff, George D. (1927). *Dynamical Systems.* Vol. IX. Colloquium Publications. American Mathematical Society, Providence. url: https://archive.org/details/dynamicalsystems00birk (33).

Birkhoff, George D. (1928). 'Structure analysis of surface transformations'. In: *Journal de Mathématiques Pures et Appliquées* 7, 345–380 (70).

Birkhoff, George D. (1931a). 'Proof of a Recurrence Theorem for Strongly Transitive Systems'. In: *Proceedings of the National Academy of Sciences* 17 (12), 650–655. doi: 10.1073/pnas.17.12.650 (70, 219).

Birkhoff, George D. (1931b). 'Proof of the Ergodic Theorem'. In: *Proceedings of the National Academy of Sciences* 17 (12), 656–660. doi: 10.1073/pnas.17.2.656 (33, 62, 63, 220, 221, 278).

Birkhoff, George D. (1932). 'Probability and physical systems'. In: *Bulletin of the American Mathematical Society* 38 (6), 361–380. doi: 10.1090/s0002-9904-1932-05389-7 (70).

Birkhoff, George D. and Bernard O. Koopman (1932). 'Recent contributions to the ergodic theory'. In: *Proceedings of the National Academy of Sciences* 18 (3), 279–282. doi: 10.1073/pnas.18.3.279 (35, 63, 78, 219).

Bisin, Alberto, Ulrich Horst and Onur Özgür (2006). 'Rational expectations equilibria of economies with local interactions'. In: *Journal of Economic Theory* 127 (1), 74–116. doi: 10.1016/j.jet.2004.08.004 (171).

Black, Fischer and Myron S. Scholes (1973). 'The Pricing of Options and Corporate Liabilities'. In: *Journal of Political Economy* 81 (3), 637–654. doi: 10.1086/260062 (90, 278).

Blanchard, Olivier J. and Lawrence H. Summers (1986). 'Hysteresis and the European Unemployment Problem'. In: *NBER Macroeconomics Annual* 1, 15–78. doi: 10.1086/654013 (198).

Blanchard, Olivier J. and Mark W. Watson (July 1982). 'Bubbles, Rational Expectations and Financial Markets'. In: *NBER Working Paper* 945. doi: 10.3386/w0945 (248).

Blatt, John M. (1959). 'An Alternative Approach to the Ergodic Problem'. In: *Progress of Theoretical Physics* 22 (6), 745–756. doi: 10.1143/ptp.22.745 (71).

Blume, Lawrence E. (1979). 'The Ergodic Behavior of Stochastic Processes of Economic Equilibria'. In: *Econometrica* 47 (6), 1421–1432. doi: 10.2307/1914009 (181).

Blume, Lawrence E. (1993). 'The Statistical Mechanics of Strategic Interaction'. In: *Games* *and Economic Behavior* 5 (3), 387–424. DOI: 10.1006/game.1993.1023 (171).

Blume, Lawrence E. and David Easley (1992). 'Evolution and Market Behavior'. In: *Journal of Economic Theory* 58 (1), 9–40. DOI: 10.1016/0022-0531(92)90099-4 (174, 177).

Blume, Lawrence E. and David Easley (2002). 'Optimality and Natural Selection in Markets'. In: *Journal of Economic Theory* 107 (1), 95–135. DOI: 10.1006/jeth.2000.2764 (174, 177).

Blume, Lawrence E. and David Easley (2009). 'Market Selection and Asset Pricing'. In: *Handbook of Financial Markets. Dynamics and Evolution.* Ed. by Thorsten Hens and Klaus R. Schenk-Hoppé. North Holland, Amsterdam. Chap. 7, 403–437. DOI: 10.1016/B978-012374258-2.50011-7 (174).

Bogorad, Lawrence (1995). 'About the Proceedings'. In: *Proceedings of the National Academy of Sciences* 92 (1), i–ii. DOI: 10.1073/pnas.92.1.i (219).

Bolton, Gary E. and Axel Ockenfels (2000). 'ERC: A Theory of Equity, Reciprocity, and Competition'. In: *American Economic Review* 90 (1), 166–193. DOI: 10.1257/aer.90.1.166 (250, 256).

Boltzmann, Ludwig (1868). 'Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten'. In: *Wissenschaftliche Abhandlungen (1909). I. Band (1865-1874).* Ed. by Fritz Hasenöhrl. Verlag von Johann Ambrosius Barth, Leipzig. Chap. 5. 49–96 (37, 70).

Boltzmann, Ludwig (1871). 'Einige allgemeine Sätze über das Wärmegleichgewicht'. In: *Wissenschaftliche Abhandlungen (1909). I. Band (1865-1874).* Ed. by Fritz Hasenöhrl. Verlag von Johann Ambrosius Barth, Leipzig. Chap. 19. 259–287 (43, 44, 51).

Boltzmann, Ludwig (1872). 'Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen'. In: *Wissenschaftliche Abhandlungen (1909). I. Band (1865-1874).* Ed. by Fritz Hasenöhrl. Verlag von Johann Ambrosius Barth, Leipzig. Chap. 22. 316–402 (37, 40, 44, 60, 70, 278).

Boltzmann, Ludwig (1875). 'Über das Wärmegleichgewicht von Gasen, auf welche äußere, Kräfte wirken'. In: *Wissenschaftliche Abhandlungen (1909). II. Band (1875-1881).* Ed. by Fritz Hasenöhrl. Verlag von Johann Ambrosius Barth, Leipzig. Chap. 32, 1–30 (51).

Boltzmann, Ludwig (1877). 'Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respective den Sätzen über das Wärmegleichgewicht'. In: *Wissenschaftliche Abhandlungen (1909). II. Band (1875-1881).* Ed. by Fritz Hasenöhrl. Verlag von Johann Ambrosius Barth, Leipzig. Chap. 42. 164–223 (37, 70).

Boltzmann, Ludwig (1884). 'Über die Eigenschaften monozyklischer und damit verwandter Systeme'. In: *Wissenschaftliche Abhandlungen (1909). III. Band (1882-1905).* Ed. by Fritz Hasenöhrl. Verlag von Johann Ambrosius Barth, Leipzig. Chap. 73. 122–152 (42, 44, 45, 48–50, 221, 278).

Boltzmann, Ludwig (1887). 'Über die mechanischen Analogien des zweiten Hauptsatzes der Thermodynamik'. In: *Wissenschaftliche Abhandlungen (1909). III. Band (1882-1905).* Ed. by Fritz Hasenöhrl. Verlag von Johann Ambrosius Barth, Leipzig. Chap. 82. 258–271 (44, 45, 48).

Boltzmann, Ludwig (1895a). 'On Certain Questions of the Theory of Gases'. In: *Nature* 51 (1322), 413–415. DOI: 10.1038/051413b0 (46).

Boltzmann, Ludwig (1895b). 'On the Minimum Theorem in the Theory of Gases'. In: *Nature* 52 (1340), 221. doi: 10.1038/052221b0 (46).

Boltzmann, Ludwig (1895c). 'Reply to Culverwell's 'Professor Boltzmann's Letter on the Kinetic Theory of Gases''. In: *Nature* 51 (1329), 581. doi: 10.1038/051581a0 (40, 46).

Boltzmann, Ludwig (1898a). 'Ueber die sogenannte H-Curve'. In: *Mathematische Annalen* 50 (2-3), 325–332. doi: 10.1007/bf01448073 (40).

Boltzmann, Ludwig (1898b). *Vorlesungen über Gastheorie.* Verlag von Johann Ambrosius Barth, Leipzig (36).

Boltzmann, Ludwig (1964). *Lectures on Gas Theory.* Trans. by Stephen G. Brush. Dover Publications, New York (36, 49–51).

Borovkov, Aleksandr A. (1998). *Ergodicity and Stability of Stochastic Processes.* Wiley, New York (77).

Bottom, William P., Robert N. Bontempo and David Robert Holtgrave (1989). 'Experts, Novices, and the St. Petersburg paradox: Is one solution enough?' In: *Journal of Behavioral Decision Making* 2 (3), 139–147. doi: 10.1002/bdm.3960020302 (91).

Bouchaud, Jean-Philippe (1992). 'Weak ergodicity breaking and aging in disordered systems'. In: *Journal de Physique I* 2 (9), 1705–1713. doi: 10.1051/jp1:1992238 (207).

Bowen, Robert E. and David Ruelle (1975). 'The ergodic theory of Axiom A flows'. In: *Inventiones Mathematicae* 29 (3), 181–202. doi: 10.1007/bf01389848 (75).

Breiman, Leo (1957). 'The Individual Ergodic Theorem of Information Theory'. In: *Annals of Mathematical Statistics* 28 (3), 809–811. doi: 10.1214/aoms/1177706899 (78).

Breiman, Leo (1960a). 'Correction to "The Individual Ergodic Theorem of Information Theory"'. In: *Annals of Mathematical Statistics* 31 (3), 809–810. doi: 10.1214/aoms/1177705812 (78).

Breiman, Leo (1960b). 'Investment Policies for Expanding Businesses Optimal in a Long Run Sense'. In: *Naval Research Logistics Quarterly* 7 (4), 647–651 (265).

Breiman, Leo (1961). 'Optimal Gambling Systems for Favorable Games'. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability.* Ed. by Jerzy Neyman. Vol. 1: Contributions to the Theory of Statistics. Held at the Statistical Laboratory University of California June 20-July 30, 1960. University of California Press, 65–78. url: https://projecteuclid.org/euclid.bsmsp/1200512159 (170, 179, 265).

Breiman, Leo (1992). *Probability.* Society for Industrial and Applied Mathematics, Philadelphia (77).

Breiman, Leo (2001). 'Statistical Modeling: The Two Cultures'. In: *Statistical Science* 16 (3). (with comments and a rejoinder by the author), 199–231. doi: 10.1214/ss/1009213726 (185–187).

Brennan, Thomas J. and Andrew W. Lo (2011). 'The Origin of Behavior'. In: *Quarterly Journal of Finance* 1 (1), 55–108. doi: 10.1142/S201013921100002X (214).

Brito, Dagobert L. (1975). 'Becker's Theory of the Allocation of Time and the St. Petersburg Paradox'. In: *Journal of Economic Theory* 10 (1), 123–126. doi: 10.1016/0022-0531(75)90067-8 (112).

Brock, William A. and Steven N. Durlauf (2001). 'Discrete Choice with Social Interactions'. In: *Review of Economic Studies* 68 (2), 235–260. doi: 10.1111/1467-937x.00168 (171, 177).

Brock, William A. and Cars H. Hommes (1997). 'A Rational Route to Randomness'. In: *Econometrica* 65 (5), 1059–1095. doi: 10.2307/2171879 (178).

Brock, William A. and Cars H. Hommes (1998). 'Heterogeneous beliefs and routes to chaos in a simple asset pricing model'. In: *Journal of Economic Dynamics and Control* 22 (8-9), 1235–1274. doi: 10.1016/s0165-1889(98)00011-6 (178).

Brock, William A. and Leonard J. Mirman (1972). 'Optimal economic growth and uncertainty: The discounted case'. In: *Journal of Economic Theory* 4 (3), 479–513. doi: 10.1016/0022-0531(72)90135-4 (171, 172).

Broda, Engelbert (1955). *Ludwig Boltzmann. Mensch Physiker Philosoph.* Franz Deuticke, Wien (79).

Brown, Harvey R., Wayne Myrvold and Jos Uffink (2009). 'Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics'. In: *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 40 (2), 174–191. doi: 10.1016/j.shpsb.2009.03.003 (46, 70).

Brown, Robert (1828). *A Brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and anorganic bodies.* Richard Taylor, London (208).

Brush, Stephen G. (1967a). 'Evidence for Boltzmann's H as a capital eta'. In: *American Journal of Physics* 35 (9), 892. doi: 10.1119/1.1974281 (40).

Brush, Stephen G. (1967b). 'Foundations of statistical mechanics 1845-1915'. In: *Archive for History of Exact Sciences* 4 (3), 145–183. doi: 10.1007/bf00412958 (71).

Brush, Stephen G. (1971). 'Proof of the Impossibility of Ergodic Systems: The 1913 Papers of Rosenthal and Plancherel'. In: *Transport Theory and Statistical Physics* 1 (4), 287–298. doi: 10.1080/00411457108231454 (49, 70).

Bryan, George H. (1895). 'The Kinetic Theory of Gases'. In: *Nature* 51 (1318), 319–320. doi: 10.1038/051319b0 (46).

Buiter, Willem (6th Mar. 2009). 'The unfortunate uselessness of most 'state of the art' academic monetary economics'. In: *Financial Times Blog Maverecon.* url: https://voxeu.org/article/macroeconomics-crisis-irrelevance (6).

Bunimovich, Leonid A. (1974). 'On billiards close to dispersing'. In: *Mathematics of the USSR-Sbornik* 23 (1), 45–67. doi: 10.1070/sm1974v023n01abeh001713 (74).

Bunimovich, Leonid A. (1975). 'On ergodic properties of certain billiards'. In: *Functional Analysis and Its Applications* 8 (3), 254–255. doi: 10.1007/bf01075700 (74).

Bunimovich, Leonid A. (1979). 'On the ergodic properties of nowhere dispersing billiards'. In: *Communications in Mathematical Physics* 65 (3), 295–312. doi: 10.1007/bf01197884 (74).

Bunimovich, Leonid A. (1990). 'A theorem on ergodicity of two-dimensional hyperbolic billiards'. In: *Communications in Mathematical Physics* 130 (3), 599–621. doi: 10.1007/bf02096936 (74).

Bunimovich, Leonid A. (2001). 'Mushrooms and other billiards with divided phase space'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 11 (4), 802–808. doi: 10.1063/1.1418763 (74).

Bunimovich, Leonid A. (2007). 'Dynamical billiards'. In: *Scholarpedia* 2 (8). revision #91211, 1813. doi: 10.4249/scholarpedia.1813 (74).

Bunimovich, Leonid A. and Yakov G. Sinai (1973). 'On a fundamental theorem in the theory of dispersing billiards'. In: *Mathematics of the USSR-Sbornik* 19 (3), 407–423. doi: 10.1070/SM1973v019n03ABEH001786 (74).

Burbury, Samuel H. (1894). 'Boltzmann's Minimum Function'. In: *Nature* 51 (1308), 78. doi: 10.1038/051078a0 (40, 46).

Burbury, Samuel H. (1895a). 'Boltzmann's Minimum Function'. In: *Nature* 51 (1318), 320. doi: 10.1038/051320a0 (46).

Burbury, Samuel H. (1895b). 'Boltzmann's Minimum Function'. In: *Nature* 52 (1335), 104–105. doi: doi:10.1038/052104b0 (46).

Burov, Stas, Jae-Hyung Jeon, Ralf Metzler and Eli Barkai (2011). 'Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking'. In: *Physical Chemistry Chemical Physics* 13 (5), 1800. doi: 10.1039/c0cp01879a (207).

Byers, William (2011). *The Blind Spot. Science and the Crisis of Uncertainty.* Princeton University Press (2).

Callon, Michel (1998). 'Introduction: The embeddedness of economic markets in economics'. In: *The Laws of the Markets.* Ed. by Michel Callon. Blackwell Publishing, Oxford, 1–57 (238).

Camerer, Colin F. (1999). 'Behavioral economics: Reunifying psychology and economics'. In: *Proceedings of the National Academy of Sciences* 96 (19), 10575–10577. doi: 10.1073/pnas.96.19.10575 (254).

Campbell, David K., J. Doyne Farmer, Jim Crutchfield and Erica Jen (1985). 'Experimental Mathematics: The Role of Computation in Nonlinear Science'. In: *Communications of the ACM* 28 (4), 374–384. doi: 10.1145/3341.3345 (29).

Carathéodory, Constantin (1914). 'Über das lineare Maß von Punktmengen – eine Verallgemeinerung des Längenbegriffs'. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 404–426. url: http://www.digizeitschriften.de/dms/resolveppn/?PID=GDZPPN002504006 (48).

Carathéodory, Constantin (1919). 'Über den Wiederkehrsatz von Poincaré'. In: *Sitzungsberichte der Preußischen Akademie der Wissenschaften*, 580–584 (48).

Cardano, Girolamo (1633). *Liber de Ludo Aleae. Book on Games of Chance.* written around 1550 (81, 82).

Carleson, Lennart (1991). 'Stochastic behavior of deterministic systems'. In: *Journal of Economic Behavior & Organization* 16 (1-2), 85–92. doi: 10.1016/0167-2681(91)90044-x (73).

Carrión Álvarez, Miguel and Dirk Ehnts (2016). 'Samuelson and Davidson on ergodicity: A reformulation'. In: *Journal of Post Keynesian Economics* 39 (1), 1–16. doi: 10.1080/01603477.2016.1145062 (192, 219).

Cartwright, Mary L. and John E. Littlewood (1945). 'On Non-Linear Differential Equations of the Second Order: I. the Equation $\ddot{y} - k(1-y^2)\dot{y} + y = b\lambda k\cos(\lambda l + \alpha), k$ Large'. In: *Journal of the London Mathematical Society* s1-20 (3), 180–189. doi: 10.1112/jlms/s1-20.3.180 (39).

Caspari, Volker (2014). 'Keynes' Kritik an der „Tinbergenschen Ökonometrie". Ein Intermezzo in der Entwicklung der Makroökonomie?' In: *Studien zur Entwicklung der ökonomischen Theorie XXVIII. Die Ökonomik im Spannungsfeld zwischen Natur- und Geisteswissenschaften. Alte und neue Perspektiven im Licht des jüngsten Methodenstreits.* Ed. by Heinz D. Kurz.

Vol. 115. Schriften des Vereins für Socialpolitik. Duncker & Humblot, Berlin, 107–123 (182).

Cass, David (1965). 'Optimum Growth in an Aggregative Model of Capital Accumulation'. In: *Review of Economic Studies* 32 (3), 233–240. DOI: 10.2307/2295827 (172).

Chapman, Sydney (1937). 'Boltzmann's *H*-Theorem'. In: *Nature* 139 (3526), 931. DOI: 10.1038/139931a0 (40).

Chernoff, Herman and Lincoln E. Moses (1986). *Elementary Decision Theory*. Dover Publications, New York (88, 90).

Chernov, Nikolai I. (1991). 'New proof of Sinai's formula for the entropy of hyperbolic billiard systems. Application to Lorentz gases and Bunimovich stadiums'. In: *Functional Analysis and Its Applications* 25 (3), 204–219. DOI: 10.1007/BF01085490 (74).

Cherrier, Beatrice (2014). 'Toward a History of Economics at MIT, 1940-72'. In: *History of Political Economy* 46 (Supplement 1), 15–44. DOI: 10.1215/00182702-2716100 (222).

Cherstvy, Andrey G., Aleksei V. Chechkin and Ralf Metzler (2013). 'Anomalous diffusion and ergodicity breaking in heterogeneous diffusion processes'. In: *New Journal of Physics* 15 (8), 083039. DOI: 10.1088/1367-2630/15/8/083039 (206).

Chintschin, Alexander J. (1964). *Mathematische Grundlagen der statistischen Mechanik*. Bibliographisches Institut, Mannheim (42).

Choe, Geon Ho (2005). *Computational Ergodic Theory*. Springer, Berlin Heidelberg. DOI: 10.1007/b138894 (77).

Clausius, Rudolf (1857). 'Ueber die Art der Bewegung, welche wir Wärme nennen'. In: *Annalen der Physik und Chemie* 176 (3), 353–380. DOI: 10.1002/andp.18571760302 (35).

Clausius, Rudolf (1865). 'Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie'. In: *Annalen der Physik und Chemie* 201 (7), 353–400. DOI: 10.1002/andp.18652010702 (52).

Clausius, Rudolf (1875a). 'VI. On the theorem of the mean ergal, and its application to the molecular motions of gases'. In: *Philosophical Magazine* 50 (328), 26–46. DOI: 10.1080/14786447508641256 (51).

Clausius, Rudolf (1875b). 'XIV. On the theorem of the mean ergal, and its application to the molecular motions of gases'. In: *Philosophical Magazine* 50 (329), 101–107. DOI: 10.1080/14786447508641266 (51).

Clausius, Rudolf (1875c). 'XXIII. On the theorem of the mean ergal, and its application to the molecular motions of gases'. In: *Philosophical Magazine* 50 (330), 191–200. DOI: 10.1080/14786447508641279 (51).

Clements, Michael P. and David F. Hendry (1999). *Forecasting Non-Stationary Economic Time Series*. MIT Press (197).

Cohen, Dan (1966). 'Optimizing reproduction in a randomly varying environment'. In: *Journal of Theoretical Biology* 12 (1), 119–129. DOI: 10.1016/0022-5193(66)90188-3 (213).

Cohen, Joel E. (1976). 'Ergodicity of Age Structure in Populations with Markovian Vital Rates, I: Countable States'. In: *Journal of the American Statistical Association* 71 (354), 335–339. DOI: 10.1080/01621459.1976.10480343 (215).

Cohen, Joel E. (1977a). 'Ergodicity of age structure in populations with Markovian vital rates. II. General states'. In: *Advances in Applied Probability* 9 (1), 18–37. DOI: 10.2307/1425814 (215).

Cohen, Joel E. (1977b). 'Ergodicity of age structure in populations with Markovian vital rates, III: Finite-state moments and growth rate; an illustration'. In: *Advances in Applied Probability* 9 (3), 462–475. doi: 10.2307/1426109 (215).

Cohen, Joel E. (1979a). 'Ergodic theorems in demography'. In: *Bulletin of the American Mathematical Society* 1 (2), 275–295 (215).

Cohen, Joel E. (1979b). 'Ergodic Theorems of Population Dynamics'. In: *Rocky Mountain Journal of Mathematics* 9 (1), 27. doi: 10.1216/rmj-1979-9-1-27 (215).

Colander, David (2007). 'Conversations with James Tobin and Robert J. Shiller on the "Yale Tradition" in Macroeconomics. Conversations with Eminent Economists'. In: *Inside the Economist's Mind*. Ed. by Paul A. Samuelson and William A. Barnett. Blackwell Publishing, Malden. Chap. 16, 392–406 (219).

Colander, David, Peter Howitt, Alan P. Kirman, Axel Leijonhufvud and Perry Mehrling (2008). 'Beyond DSGE Models: Toward an Empirically Based Macroeconomics'. In: *American Economic Review Papers and Proceedings* 98, 2, 236–240. doi: 10.1257/aer.98.2.236 (251).

Cont, Rama, ed. (Feb. 2010). *Encyclopedia of Quantitative Finance*. John Wiley & Sons, Ltd. doi: 10.1002/9780470061602 (175).

Cornfeld, Isaak P., Sergei V. Fomin and Yakov G. Sinai (1982). *Ergodic Theory*. translated by A. B. Sossinskiĭ. Springer, New York. doi: 10.1007/978-1-4615-6927-5 (42, 77, 296).

Корнфелд, Исаак П., Яков Г. Синаи and Сергеи В. Фомин (1980). Эргодическая теория. Russian original of Cornfeld, Fomin and Sinai (1982). Наука, Москва (42).

Corradi, Valentina, Norman R. Swanson and Halbert White (2000). 'Testing for stationarity-ergodicity and for comovements between nonlinear discrete time Markov processes'. In: *Journal of Econometrics* 96 (1), 39–73. doi: 10.1016/s0304-4076(99)00050-0 (196).

Cournot, Antoine-Augustin (1843). *Exposition de la théorie des chances et des probabilités*. reprinted in 1984 as Volume I in Cournot (1973). Hachette, Paris (105, 107).

Cournot, Antoine-Augustin (1973). *Œvres complètes*. Ed. by Bernard Bru. Vrin, Paris (296).

Cover, Thomas M. (1991). 'Universal Portfolios'. In: *Mathematical Finance* 1 (1), 1–29. doi: 10.1111/j.1467-9965.1991.tb00002.x (213, 278).

Cover, Thomas M. (2011). 'On the St. Petersburg Paradox'. In: *IEEE International Symposium on Information Theory Proceedings*. July 31 - August 5, 2011 in St. Petersburg, Russia, 1758–1761. doi: 10.1109/isit.2011.6033850 (164).

Cover, Thomas M. (2012). 'On the St. Petersburg Paradox'. In: Talk given at the Information Theory Forum at the Information Systems Laboratory at Stanford. url: https://web.stanford.edu/group/isl/public/documents/st_pete_stanfordstats_talk_print.pdf (87).

Cover, Thomas M. and Joy A. Thomas (2006). *Elements of Information Theory*. 2nd ed. Wiley John & Sons, Hoboken (77, 164, 213, 278).

Cowen, Tyler and Jack High (1988). 'Time, bounded utility, and the St. Petersburg paradox'. In: *Theory and Decision* 25 (3), 219–223. doi: 10.1007/bf00133163 (111).

Cox, James C., Eike B. Kroll, Marcel Lichters, Vjollca Sadiraj and Bodo Vogt

(2018). 'The St. Petersburg paradox despite risk-seeking preferences: an experimental study'. In: *Business Research* (online first), 1–18. DOI: 10.1007/s40685-018-0078-y (91).

Cox, James C., Vjollca Sadiraj and Bodo Vogt (2009). 'On the empirical relevance of St. Petersburg lotteries'. In: *Economics Bulletin* 29 (1), 214–220. URL: https://ideas.repec.org/a/ebl/ecbull/eb-09-00013.html (91).

Crean, Angela J. and Dustin J. Marshall (2009). 'Coping with environmental uncertainty: dynamic bet hedging as a maternal effect'. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1520), 1087–1096. DOI: 10.1098/rstb.2008.0237 (214).

Cubitt, Robin P. and Robert Sugden (2001). 'On Money Pumps'. In: *Games and Economic Behavior* 37 (1), 121–160. DOI: 10.1006/game.2000.0834 (247).

Culverwell, Edward P. (1894a). 'Dr. Watson's Proof of Boltzmann's Theorem on Permanence of Distributions'. In: *Nature* 50 (1304), 617. DOI: 10.1038/050617c0 (46).

Culverwell, Edward P. (1894b). 'The Kinetic Theory of Gases'. In: *Nature* 51 (1308), 78–79. DOI: 10.1038/051078b0 (46).

Culverwell, Edward P. (1895). 'Professor Boltzmann's Letter on the Kinetic Theory of Gases'. In: *Nature* 51 (1329), 581. DOI: 10.1038/051581a0 (46).

Cynamon, Barry Z., Steven Fazzari and Mark Setterfield, eds. (2012). *After the Great Recession. The Struggle for Economic Recovery and Growth*. Cambridge University Press. DOI: 10.1017/cbo9781139059589 (189).

Czuber, Emanuel (1882). 'Das Petersburger Problem'. In: *Archiv der Mathematik und Physik* 67, 1–28. URL: https://archive.org/details/archivdermathem61unkngoog (100).

Daníelsson, Jón (2002). 'The emperor has no clothes: Limits to risk modelling'. In: *Journal of Banking & Finance* 26 (7), 1273–1296. DOI: 10.1016/S0378-4266(02)00263-7 (238).

Daníelsson, Jón (2008). 'Blame the models'. In: *Journal of Financial Stability* 4 (4), 321–328 (238).

Daníelsson, Jón, Paul Embrechts, Charles Goodhart, Con Keating, Felix Muennich, Olivier Renault and Hyun Song Shin (2001). 'An Academic Response to Basel II'. In: *Special Paper Series* LSE Financial Markets Group.130 (238).

Daston, Lorraine (1995). *Classical Probability in the Enlightenment*. Princeton University Press (78).

Dauxois, Thierry (2008). 'Fermi, Pasta, Ulam, and a mysterious lady'. In: *Physics Today* 61 (1), 55–57. DOI: 10.1063/1.2835154 (73).

David, Paul A. (1985). 'Clio and the Economics of QWERTY'. In: *American Economic Review Papers and Proceedings* 75 (2), 332–337. JSTOR: 1805621 (171, 182).

David, Paul A. (2001). 'Path dependence, its critics and the quest for 'historical economics''. In: *Evolution and Path Dependence in Economic Ideas: Past and Present*. Ed. by Pierre Garrouste and Stavros Ioannides. Edward Elgar, Cheltenham. Chap. 2 (182).

Davidson, James (1994). *Stochastic Limit Theory. An Introduction for Econometricians*. Oxford University Press. DOI: 10.1093/0198774036.001.0001 (194).

Davidson, Paul (1982). 'Rational Expectations: A Fallacious Foundation for Studying Crucial Decision-Making Processes'. In: *Journal of Post Keynesian Economics* 5 (2), 182–198. DOI: 10.1080/01603477.1982.11489355 (188, 189).

Davidson, Paul (1984). 'Reviving Keynes's revolution'. In: *Journal of Post Keynesian Eco-*

nomics 6 (4), 561–575. DOI: 10.1080/01603477.1984.11489467 (188).

DAVIDSON, PAUL (1991). 'Is Probability Theory Relevant for Uncertainty? A Post Keynesian Perspective'. In: *Journal of Economic Perspectives* 5 (1), 129–143. DOI: 10.1257/jep.5.1.129 (188, 190, 192, 218, 232).

DAVIDSON, PAUL (1992). 'Reponse to Bateman'. In: *Journal of Economic Perspectives* 6 (1), 206–208. DOI: 10.1257/jep.6.1.201. Correspondence to P. Davidson "Is Probability Theory Relevant for Uncertainty? A Post Keynesian Perspective" (1991) (188).

DAVIDSON, PAUL (1996). 'Reality and economic theory'. In: *Journal of Post Keynesian Economics* 18 (4), 479–508. DOI: 10.1080/01603477.1996.11490083 (188, 189, 192, 218).

DAVIDSON, PAUL (2002). *Financial Markets, Money and the Real World*. Edward Elgar, Cheltenham (189, 190).

DAVIDSON, PAUL (2003). 'Setting the record straight on *A history of Post Keynesian economics*'. In: *Journal of Post Keynesian Economics* 26 (2), 245–272. DOI: 10.1080/01603477.2003.11051396 (188, 189).

DAVIDSON, PAUL (2006). 'Samuelson and the Keynes/Post Keynesian Revolution'. In: *Samuelsonian Economics. And the Twenty-First Century*. Ed. by MICHAEL SZENBERG, LALL RAMRATTAN and ARON A. GOTTESMAN. Oxford University Press, 178–196 (188).

DAVIDSON, PAUL (2009a). 'Can future systemic financial risks be quantified? Ergodic vs nonergodic stochastic processes'. In: *Brazilian Journal of Political Economy (Revista de Economia Politica)* 29 (4), 324–340. DOI: 10.1590/S0101-31572009000400001 (188, 218, 232, 233).

DAVIDSON, PAUL (2009b). *The Keynes Solution. The Path to Global Economic Prosperity*. Palgrave Macmillan, New York (189).

DAVIDSON, PAUL (2010). 'Black swans and Knight's epistemological uncertainty: are these concepts also underlying behavioral and post-Walrasian theory?' In: *Journal of Post Keynesian Economics* 32 (4), 567–570. DOI: 10.2753/pke0160-3477320404 (189, 191).

DAVIDSON, PAUL (2012). 'Is economics a science? Should economics be rigorous?' In: *real-world economics review* (59), 58–66. URL: http://www.paecon.net/PAEReview/issue59/Davidson59.pdf (188, 190, 192, 218).

DAVIDSON, PAUL (2014). 'Biochemist to Economist'. In: *Eminent Economists II. Their Life and Work Philosophies*. Ed. by MICHAEL SZENBERG and LALL RAMRATTAN. Cambridge University Press, 72–83 (188, 192).

DAVIDSON, PAUL (2015). 'A rejoinder to O'Donnell's critique of the ergodic/nonergodic explanation of Keynes's concept of uncertainty'. In: *Journal of Post Keynesian Economics* 38 (1), 1–18. DOI: 10.1080/01603477.2015.1078701 (188, 192).

DAVIDSON, PAUL (2016). 'Rejoinder to Rosser, O'Donnell, and Carrión Álvarez and Ehnts on their criticisms of my ergodic/nonergodic formulation of Keynes's concept of an actuarial certain future vs. an uncertain future'. In: *Journal of Post Keynesian Economics* 39 (3), 308–333. DOI: 10.1080/01603477.2016.1203728 (192).

DAY, RICHARD H. and WAYNE SHAFER (1987). 'Ergodic Fluctuations in Deterministic Economic Models'. In: *Journal of Economic Behavior & Organization* 8 (3), 339–361. DOI: 10.1016/0167-2681(87)90049-7 (73, 181).

DEBREU, GÉRARD (1959). *Theory of Value. An Axiomatic Analysis of Economic Equilibrium*. Yale University Press (180).

Debreu, Gérard (1974). 'Excess demand functions'. In: *Journal of Mathematical Economics* 1 (1), 15–21. doi: 10.1016/0304-4068(74)90032-9 (180, 227).

Debreu, Gérard (1991). 'The Mathematization of Economic Theory'. In: *American Economic Review* 81 (1), 1–7. JSTOR: 2006785 (228).

Dehling, Herold G. (1997). 'Daniel Bernoulli and the St. Petersburg Paradox'. In: *Nieuw Archief voor Wiskunde* 6 (3), 223–227 (108).

Delli Gatti, Domenico, Saul Desiderio, Edoardo Gaffeo, Pasquale Cirillo and Mauro Gallegati (2011). *Macroeconomics from the Bottom-up*. Springer, Milan. doi: 10.1007/978-88-470-1971-3 (251).

Delli Gatti, Domenico, Giorgio Fagiolo, Mauro Gallegati, Matteo Richiardi and Alberto Russo, eds. (2018). *Agent-based Models in Economics. A Toolkit*. Cambridge University Press. doi: 10.1017/9781108227278 (196).

Delli Gatti, Domenico, Edoardo Gaffeo, Mauro Gallegati, Gianfranco Giulioni and Antonio Palestrini (2008). *Emergent Macroeconomics. An Agent-Based Approach to Business Fluctuations*. Springer, Milan. doi: 10.1007/978-88-470-0725-3 (251).

Derman, Cyrus (1954). 'Ergodic Property of the Brownian Motion Process'. In: *Proceedings of the National Academy of Sciences* 40 (12), 1155. url: http://www.pnas.org/content/40/12/1155.full.pdf (219).

Derman, Emanuel (2004). *My Life as a Quant. Reflections on Physics and Finance*. John Wiley & Sons, New York (228).

Derndinger, Roland, Rainer Nagel and Günther Palm (1987). 'Ergodic Theory in the Perspective of Functional Analysis'. unpublished lecture notes. url: http://www.math.uni-tuebingen.de/arbeitsbereiche/agfa/members/rana/ergthfuan.pdf (49, 77).

Derndinger, Roland, Rainer Nagel and Günther Palm (1998). *Ergodic Theory in the Perspective of Functional Analysis*. url: https://www.uni-ulm.de/fileadmin/website_uni_ulm/iui.inst.130/Mitarbeiter/palm/Ergodic.TEX.pdf (77).

Devlin, Keith (2009). *Pascal, Fermat und die Berechnung des Glücks. Eine Reise in die Geschichte der Mathematik*. C. H. Beck, München (83).

Devlin, Keith (2010). 'The Pascal-Fermat Correspondence: How Mathematics Is Really Done'. In: *Mathematics Teacher* 103 (8), 579–582 (83).

DeWitt, Bryce S. (1970). 'Quantum mechanics and reality'. In: *Physics Today* 23 (9), 155–167. doi: 10.1063/1.3022331 (212).

DeWitt, Bryce S. and Neill Graham, eds. (1973). *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton University Press (212).

Diba, Behzad T. and Herschel I. Grossman (1987). 'On the Inception of Rational Bubbles'. In: *Quarterly Journal of Economics* 102 (3), 697–700. doi: 10.2307/1884225 (248).

Diba, Behzad T. and Herschel I. Grossman (1988a). 'Rational inflationary bubbles'. In: *Journal of Monetary Economics* 21 (1), 35–46. doi: 10.1016/0304-3932(88)90044-x (248).

Diba, Behzad T. and Herschel I. Grossman (1988b). 'The Theory of Rational Bubbles in Stock Prices'. In: *Economic Journal* 98 (392), 746–754. doi: 10.2307/2233912 (248).

Dierker, Egbert (1974). *Topological Methods in Walrasian Economics*. Springer, Berlin Heidelberg. doi: 10.1007/978-3-642-65800-6 (181).

DIERKER, EGBERT, HILDEGARD DIERKER and WALTER TROCKEL (1980a). 'Continuous mean demand functions derived from non-convex preferences'. In: *Journal of Mathematical Economics* 7 (1), 27–33. DOI: 10.1016/0304-4068(80)90018-x (180).

DIERKER, EGBERT, HILDEGARD DIERKER and WALTER TROCKEL (1980b). 'Smoothing demand by aggregation with respect to wealth'. In: *Journal of Mathematical Economics* 7 (3), 227–247. DOI: 10.1016/0304-4068(80)90011-7 (180).

DIERKER, EGBERT, HILDEGARD DIERKER and WALTER TROCKEL (1984). 'Price-dispersed preferences and C^1 mean demand'. In: *Journal of Mathematical Economics* 13 (1), 11–42. DOI: 10.1016/0304-4068(84)90022-3 (180).

DOMOWITZ, IAN and MAHMOUD A. EL-GAMAL (1993). 'A Consistent Test of Stationary-Ergodicity'. In: *Econometric Theory* 9 (04), 589–601. DOI: 10.1017/s0266466600007994 (196).

DOMOWITZ, IAN and MAHMOUD A. EL-GAMAL (2001). 'A consistent nonparametric test of ergodicity for time series with applications'. In: *Journal of Econometrics* 102 (2), 365–398. DOI: 10.1016/s0304-4076(01)00058-6 (195, 196).

DONALDSON-MATASCI, MATINA C., CARL T. BERGSTROM and MICHAEL LACHMANN (2010). 'The fitness value of information'. In: *Oikos* 119 (2), 219–230. DOI: 10.1111/j.1600-0706.2009.17781.x (213).

DONALDSON-MATASCI, MATINA C., MICHAEL LACHMANN and CARL T. BERGSTROM (2008). 'Phenotypic diversity as an adaptation to environmental uncertainty'. In: *Evolutionary Ecology Research* 10, 493–515. URL: http://www.evolutionary-ecology.com/abstracts/v10/2266.html (213, 214).

DOSI, GIOVANNI, YURI M. ERMOLIEV and YURI M. KANIOVSKI (1994). 'Generalized urn schemes and technological dynamics'. In: *Journal of Mathematical Economics* 23 (1), 1–19. DOI: 10.1016/0304-4068(94)90032-9 (182).

DOWNAROWICZ, TOMASZ (2011). *Entropy in Dynamical Systems*. Cambridge University Press (77).

DUARTE QUEIROS, SILVIO M., CELIA ANTENEODO and CONSTANTINO TSALLIS (2005). 'Power-law distributions in economics: a nonextensive statistical approach (Invited Paper)'. In: *Noise and Fluctuations in Econophysics and Finance*. Ed. by DEREK ABBOTT, JEAN-PHILIPPE BOUCHAUD, XAVIER GABAIX and JOSEPH L. MCCAULEY. Vol. 5848. SPIE Third International Symposium on Fluctuations and Noise. Chap. Statistical Physics/Thermodynamics Viewpoint, 151–164. DOI: 10.1117/12.619469 (205).

DUFFIE, DARRELL (1987). 'Stochastic equilibria with incomplete financial markets'. In: *Journal of Economic Theory* 41 (2), 405–416. DOI: 10.1016/0022-0531(87)90027-5 (181).

DUFFIE, DARRELL, JOHN GEANAKOPLOS, ANDREU MAS-COLELL and ANDY MCLENNAN (1994). 'Stationary Markov Equilibria'. In: *Econometrica* 62 (4), 745–781. DOI: 10.2307/2951731 (181).

DUMAS, H. SCOTT (2014). *The KAM Story. A Friendly Introduction to the Content, History, and Significance of Classical Kolmogorov–Arnold–Moser Theory*. World Scientific, Singapore (66, 75, 77).

DURLAUF, STEVEN N. (1993). 'Nonergodic Economic Growth'. In: *Review of Economic Studies* 60 (2), 349–366. DOI: 10.2307/2298061 (171, 172).

DURLAUF, STEVEN N. (2001). 'A Framework For The Study of Individual Behavior and Social

Interactions'. In: *Sociological Methodology* 31 (1), 47–87. DOI: 10.1111/0081-1750.00089 (171).

Durlauf, Steven N., Paul A. Johnson and Jonathan R. W. Temple (2005). 'Growth Econometrics'. In: *Handbook of Economic Growth*. Ed. by Philippe Aghion and Steven N. Durlauf. Vol. 1.A. Elsevier. Chap. 8, 555–677. DOI: 10.1016/s1574-0684(05)01008-7 (175).

Dutka, Jacques (1988). 'On the St. Petersburg paradox'. In: *Archive for History of Exact Sciences* 39 (1), 13–39. DOI: 10.1007/bf00329984 (18, 92, 100, 107).

Dyson, Freeman (2004). 'A meeting with Enrico Fermi'. In: *Nature* 427 (6972), 297. DOI: 10.1038/427297a (250).

Earman, John and Miklós Rédei (1996). 'Why Ergodic Theory Does Not Explain the Success of Equilibrium Statistical Mechanics'. In: *British Journal for the Philosophy of Science* 47 (1), 63–78. DOI: 10.1093/bjps/47.1.63 (72).

Ebeling, Werner, Jan Freund and Frank Schweitzer (1998). *Komplexe Strukturen: Entropie und Information*. Teubner; Stuttgart, Leipzig (206).

Eckmann, Jean-Pierre and David Ruelle (1985). 'Ergodic theory of chaos and strange attractors'. In: *Reviews of Modern Physics* 57 (3), 617–656. DOI: 10.1103/revmodphys.57.617 (73, 75).

Economist, The (25th Jan. 2010a). 'Data, data everywhere'. In: *Economist* (185).

Economist, The (25th Jan. 2010b). 'Data deluge'. In: *Economist* (185).

Economist, The (20th Apr. 2013). 'The 90% question'. In: URL: http://www.economist.com/news/finance-and-economics/21576362-seminal-analysis-relationship-between-debt-and-growth-comes-under (223).

Edwards, Harold M. (1981). 'Read the Masters!' In: *Mathematics Tomorrow*. Ed. by Lynn Arthur Steen. Springer New York, 105–110. DOI: 10.1007/978-1-4613-8127-3_11 (35).

Ehrenfest, Paul and Tatiana A. Ehrenfest (1907). 'Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem'. In: *Physikalische Zeitschrift* 8, 311–314 (40, 48).

Ehrenfest, Paul and Tatiana A. Ehrenfest (1911). 'Begriffliche Grundlagen der statistischen Auffassung der Mechanik'. In: *Encyklopädie der mathematischen Wissenschaften. Mit Einschluss ihrer Anwendungen*. Ed. by Felix Klein and Conrad Müller. .Band IV, vierter Teilband, Heft 6, Art. 32. D. Mechanik der aus sehr zahlreichen diskreten Teilchen bestehenden Systeme. B. G. Teubner, Leipzig, 1–90. URL: bit.ly/MKEhrenfest1911 (40, 42, 47–49, 71, 184, 278).

Einsiedler, Manfred and Thomas Ward (2011). *Ergodic Theory. with a view towards Number Theory*. Springer, London. DOI: 10.1007/978-0-85729-021-2 (56, 77).

Einstein, Albert (1905). 'Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen'. In: *Annalen der Physik* 322 (8), 549–560 (208, 278).

Einstein, Albert (1906). 'Zur Theorie der Brownschen Bewegung'. In: *Annalen der Physik* 324 (2), 371–381. DOI: 10.1002/andp.19063240208 (278).

Eisner, Tanja, Bálint Farkas, Markus Haase and Rainer Nagel (2015). *Operator Theoretic Aspects of Ergodic Theory*. Springer, Heidelberg. DOI: 10.1007/978-3-319-16898-2 (66, 77).

ELLIS, RICHARD S. (2006). *Entropy, Large Deviations, and Statistical Mechanics*. Springer Berlin Heidelberg. DOI: 10.1007/3-540-29060-5 (77).

EREV, IDO, IRA GLOZMAN and RALPH HERTWIG (2008). 'What impacts the impact of rare events'. In: *Journal of Risk and Uncertainty* 36 (2), 153–177. DOI: 10.1007/s11166-008-9035-z (91).

ESPOSITO, ELENA (2007). *Die Fiktion der wahrscheinlichen Realität*. Suhrkamp Verlag, Frankfurt am Main (199).

ESPOSITO, ELENA (2010). *Die Zukunft der Futures. Die Zeit des Geldes in Finanzwelt und Gesellschaft*. Carl-Auer, Heidelberg (199).

EULER, LEONHARD (1862). 'The True Valuation of the Risk in Games'. In: *Opera Postuma*. Trans. by A. BERRA and RICHARD J. PULSKAMP. Vol. 1. E811, 315–318. URL: https://scholarlycommons.pacific.edu/euler-works/811/ (100, 162, 169).

EVERETT III, HUGH (1957). '"Relative State" Formulation of Quantum Mechanics'. In: *Reviews of modern physics* 29 (3), 454–462. DOI: 10.1103/RevModPhys.29.454 (212, 302).

EVERETT III, HUGH (1973). 'The Many-Worlds Interpretation of Quantum Mechanics'. In: *The Many-Worlds Interpretation of Quantum Mechanics*. Ed. by BRUCE S. DEWITT and NEILL GRAHAM. Longer published version of his PhD thesis EVERETT III (1957). Princeton University Press, 3–140 (212).

EVSTIGNEEV, IGOR V., THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ (2002). 'Market Selecetion of Financial Trading Strategies: Global Stability'. In: *Mathematical Finance* 12 (4), 329–339. DOI: 10.1111/j.1467-9965.2002.tb00127.x (179, 180).

EVSTIGNEEV, IGOR V., THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ (2006). 'Evolutionary stable stock markets'. In: *Economic Theory* 27 (2), 449–468. DOI: 10.1007/s00199-005-0607-8 (180).

EVSTIGNEEV, IGOR V., THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ (2008). 'Globally evolutionarily stable portfolio rules'. In: *Journal of Economic Theory* 140 (1), 197–228. DOI: 10.1016/j.jet.2007.09.005 (179).

EVSTIGNEEV, IGOR V., THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ (2009). 'Evolutionary Finance'. In: *Handbook of Financial Markets. Dynamics and Evolution*. Ed. by THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ. North Holland, Amsterdam. Chap. 9, 507–566. DOI: 10.1016/B978-012374258-2.50013-0 (173, 174, 179).

EVSTIGNEEV, IGOR V., THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ (2011a). 'Local stability analysis of a stochastic evolutionary financial market model with a risk-free asset'. In: *Mathematics and Financial Economics* 5 (3), 185–202. DOI: 10.1007/s11579-011-0056-z (179, 180).

EVSTIGNEEV, IGOR V., THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ (2011b). 'Survival and Evolutionary Stability of the Kelly Rule'. In: *The Kelly Capital Growth Investment Criterion. Theory and Practice*. Ed. by LEONARD C. MACLEAN, EDWARD O. THORP and WILLIAM T. ZIEMBA. .II: Classic Papers and Theories. World Scientific, New Jersey. Chap. 20, 273–284. DOI: 10.1142/9789814293501_0020 (179).

EVSTIGNEEV, IGOR V., THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ (2015). *Mathematical Financial Economics. A Basic Introduction*. Springer, Cham. DOI: 10.1007/978-3-319-16571-4 (179, 180).

FÖLLMER, HANS (1974). 'Random economies with many interacting agents'. In: *Journal of*

Mathematical Economics 1 (1), 51–62. DOI: 10.1016/0304-4068(74)90035-4 (171).

FÖLLMER, HANS, ULRICH HORST and ALAN P. KIRMAN (2005). 'Equilibria in financial markets with heterogeneous agents: a probabilistic perspective'. In: *Journal of Mathematical Economics* 41 (1-2), 123–155. DOI: 10.1016/j.jmateco.2004.08.001 (177, 178).

FÖLLMER, HANS and ALEXANDER SCHIED (2016). *Stochastic Finance. An Introduction in Discrete Time*. 4th ed. de Gruyter, Berlin (283).

FÖLLMER, HANS and MARTIN SCHWEIZER (1993). 'A Microeconomic Approach to Diffusion Models For Stock Prices'. In: *Mathematical Finance* 3 (1), 1–23. DOI: 10.1111/j.1467-9965.1993.tb00035.x (178).

FARJOUN, EMMANUEL and MOSHÉ MACHOVER (1983). *Laws of Chaos. A Probabilistic Approach to Political Economy*. Verso, London (214, 215).

FARQUHAR, IAN E. (1961). 'The Present State of Ergodic Theory∗'. In: *Nature* 190 (4770), 17–18. DOI: 10.1038/190017a0 (72).

FECHNER, GUSTAV T. (1869). *Elemente der Psychophysik. Erster Theil.* Breitkopf und Bartel, Leipzig (113, 129, 130).

FEHR, ERNST and KLAUS M. SCHMIDT (1999). 'A theory of fairness, competition, and cooperation'. In: *Quarterly Journal of Economics* 114 (3), 817–868 (250, 256).

FEIGENBAUM, MITCHELL J. (July 1978). 'Quantitative universality for a class of nonlinear transformations'. In: *Journal of Statistical Physics* 19 (1), 25–52. DOI: 10.1007/bf01020332 (39).

FELLER, WILLIAM (1945). 'Note on the Law of Large Numbers and "Fair" Games'. In: *Annals of Mathematical Statistics* 16 (3), 301–304. DOI: 10.1214/aoms/1177731094 (90, 100, 140–143).

FELLER, WILLIAM (1957). *An Introduction to Probability Theory and Its Applications*. 2nd ed. Vol. I. John Wiley & Sons, New York (77, 88, 89, 103, 140–143).

FELLER, WILLIAM (1971). *An Introduction to Probability Theory and Its Applications*. 2nd ed. Vol. II. John Wiley & Sons, New York (77).

FERMI, ENRICO (1923). 'Beweis, dass ein mechanisches Normalsystem im Allgemeinen quasi-ergodisch ist'. In: *Physikalische Zeitschrift* 24. Also published as chapter 11a in FERMI, AMALDI, ANDERSON, PERSICO, RASETTI, SMITH, WATTENBERG and SEGRÈ (1962, pp. 79–86), 261–265 (72).

FERMI, ENRICO (1924). 'Über die Existenz quasi-ergodischer Systeme'. In: *Physikalische Zeitschrift* 25. Also published as chapter 11b in FERMI, AMALDI, ANDERSON, PERSICO, RASETTI, SMITH, WATTENBERG and SEGRÈ (1962, pp. 87–88), 166–167 (72).

FERMI, ENRICO, E. AMALDI, H. L. ANDERSON, E. PERSICO, F. RASETTI, C. S. SMITH, A. WATTENBERG and EMILIO SEGRÈ, eds. (1962). *Collected papers. (Note e Memorie) Volume I Italy 1921-1938* (303).

FERMI, ENRICO, E. AMALDI, H. L. ANDERSON, E. PERSICO, F. RASETTI, C. S. SMITH, A. WATTENBERG and EMILIO SEGRÈ, eds. (1965). *Collected papers. (Note e Memorie) Volume II Italy 1939-1954* (73, 303).

FERMI, ENRICO, JOHN PASTA and STANISLAW M. ULAM (1955). 'Studies of Nonlinear Problems'. In: *Document Los Alamos-1940*. Also published as chapter 266 in FERMI, AMALDI, ANDERSON, PERSICO, RASETTI, SMITH, WATTENBERG and SEGRÈ (1965, pp. 977–988) (72, 73).

FERNHOLZ, E. ROBERT (2002). *Stochastic Portfolio Theory*. Springer New York. DOI: 10.1007/978-1-4757-3699-1 (283).

Feyerabend, Paul (1976). *Wider den Methodenzwang. Skizze einer anarchistischen Erkenntnistheorie*. Suhrkamp Verlag, Frankfurt am Main (230).

Fibonacci, Leonardo Pisano (1202). *Liber Abbaci*. url: bit.ly/MKFibonacci1202 (81, 278).

Fishburn, Peter C. (1988). *Nonlinear Preference and Utility Theory*. Wheatsheaf, Brighton (113, 255).

Fishburn, Peter C. (1991). 'Decision Theory: The Next 100 Years?' In: *Economic Journal* 101 (404), 27–32. doi: 10.2307/2233833 (255).

Fisher, Irving (1892). 'Mathematical Investigations in the Theory of Value and Prices'. In: *Transactions of the Connecticut Academy of Sciences and Arts* ix (July) (37, 221, 278).

Fisher, Irving (1906). *The Nature of Capital and Income*. Macmillan, London (37).

Foley, Duncan K. (2010). 'Mathematical Formalism and Political-Economic Content'. In: *Conference Paper*. INET Inaugural Conference at King's College, Cambridge UK, url: https://www.ineteconomics.org/uploads/papers/INET-C@K-Paper-Session-6-Foley.pdf (5).

Forestani, Lorenzo (1682). *Practica d'arithmetica e geometria*. Filippo Succhielli, Siena. Google Books: 2rNJAAAAMAAJ (81).

Fox, Craig R. and Russell A. Poldrack (2014). 'Prospect Theory and the Brain'. In: *Neuroeconomics. Decision Making and the Brain*. Ed. by Paul W. Glimcher and Ernst Fehr. 2nd ed. Elsevier, London, 533–567. doi: 10.1016/b978-0-12-416008-8.00042-5 (136).

Friedman, Daniel, R. Mark Isaac, Duncan James and Shyam Sunder (2014). *Risky Curves. On the empirical failure of expected utility*. Routledge, New York (100, 113, 254).

Friedman, K. S. (1976). 'A Partial Vindication of Ergodic Theory'. In: *Philosophy of Science* 43 (1), 151–162. doi: 10.1086/288675 (71).

Friedman, Milton (1953). 'The Methodology of Positive Economics'. In: *Essays in Positive Economics*. University of Chicago Press. Chap. Part I: Introduction, 3–43 (179, 214, 223, 225).

Frigg, Roman, Joseph Berkovitz and Fred Kronz (2014). 'The Ergodic Hierarchy'. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2014. url: http://plato.stanford.edu/archives/sum2014/entries/ergodic-hierarchy/ (76).

Frigg, Roman and Charlotte Werndl (2011). 'Explaining Thermodynamic-Like Behavior in Terms of Epsilon-Ergodicity∗'. In: *Philosophy of Science* 78 (4), 628–652. doi: 10.1086/661567 (71).

Frydman, Roman (1982). 'Towards an Understanding of Market Processes: Individual Expectations, Learning, and Convergence to Rational Expectations Equilibrium'. In: *American Economic Review* 72 (4), 652–668. JSTOR: 1810008 (198).

Frydman, Roman and Michael D. Goldberg (2003). 'Imperfect Knowledge Expectations, Uncertainty-Adjusted Uncovered Interest Rate Parity, and Exchange Rate Dynamics'. In: *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*. Ed. by Phillipe Aghion, Roman Frydman, Joseph Stiglitz and Michael Woodford. Princeton University Press, 145–182 (198).

Frydman, Roman and Michael D. Goldberg (2007). *Imperfect Knowledge Economics: Exchange Rates and Risk*. Princeton University Press (235).

Frydman, Roman and Michael D. Goldberg (2011). *Beyond Mechanical Markets: Asset price* *swings, risk, and the role of the state*. Princeton University Press (198, 235, 305).

Frydman, Roman and Michael D. Goldberg (2012). *Jenseits rationaler Märkte*. German translation of Frydman and Goldberg (2011). Wiley VCH, Weinheim (198).

Frydman, Roman and Michael D. Goldberg (2013). 'Change and expectations in macroeconomic models: recognizing the limits to knowability'. In: *Journal of Economic Methodology* 20 (2), 118–138. doi: 10.1080/1350178X.2013.804677 (198).

Frydman, Roman, Søren Johansen, Anders Rahbek and Morten Nyboe Tabor (2017). 'The Qualitative Expectations Hypothesis: Model Ambiguity, Consistent Representations of Market Forecasts, and Sentiment'. In: *INET Working Papers* 59. url: https://www.ineteconomics.org/uploads/papers/WP_59-Frydman-Qualitative-Expectation-Hypothesis.pdf (198).

Frydman, Roman and Edmund S. Phelps (1983). *Individual forecasting and aggregate outcomes. "Rational expectations" examined*. Cambridge University Press (198).

Frydman, Roman and Edmund S. Phelps, eds. (2013). *Rethinking Expectations. The Way Forward for Macroeconomics*. Princeton University Press. doi: 10.23943/princeton/9780691155234.001.0001 (198).

Furstenberg, Harry (1977). 'Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions'. In: *Journal d'Analyse Mathématique* 31 (1), 204–256. doi: 10.1007/bf02813304 (78).

Furstenberg, Harry and Yitzak Katznelson (1978). 'An ergodic Szemerédi theorem for commuting transformations'. In: *Journal d'Analyse Mathématique* 34 (1), 275–291. doi: 10.1007/bf02790016 (78).

Furstenberg, Hillel (2017). *Ergodic Theory and Fractal Geometry*. Vol. 120. First Indian Edition. American Mathematical Society, Providence (76).

Gödel, Kurt (1931a). 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I'. In: *Monatshefte für Mathematik* 38 (1), 173–198. doi: 10.1007/BF01700692 (225).

Gödel, Kurt (1931b). 'Diskussion zur Grundlegung der Mathematik'. In: *Erkenntnis* 2 (1). 2. Tagung für Erkenntnislehre der exakten Wissenschaften der Berliner Gesellschaft für Empirische Philosophie, 7.-9. September 1930 in Königsberg, Diskussion am 7. September 1930, 135–151. doi: 10.1007/BF02028146 (225).

Gallagher, Isabelle, Laure Saint-Raymond and Benjamin Texier (2014). *From Newton to Boltzmann: Hard Spheres and Short-range Potentials*. European Mathematical Societey, Zurich. doi: 10.4171/129 (38).

Gallavotti, Giovanni (1995). 'Ergodicity, ensembles, irreversibility in Boltzmann and beyond'. In: *Journal of Statistical Physics* 78 (5-6), 1571–1589. doi: 10.1007/BF02180143 (49).

Gallavotti, Giovanni (1999). *Statistical Mechanics. A Short Treatise*. Springer, Berlin (49).

Gallavotti, Giovanni (2008). 'Introduction to FPU'. In: *The Fermi-Pasta-Ulam Problem. A Status Report*. Ed. by Giovanni Gallavotti. Springer, Berlin Heidelberg, 1–19. doi: 10.1007/978-3-540-72995-2_1 (73).

Gallavotti, Giovanni, Federico Bonetto and Guido Gentile (2004). *Aspects of Ergodic, Qualitative and Statistical Theory of Motion*. Springer, Berlin Heidelberg. doi: 10.1007/978-3-662-05853-4 (49, 50).

Garson, James W. (2013). *Modal Logic for Philosophers*. Cambridge University Press (99).

Güth, Werner (1995). 'On ultimatum bargaining experiments – A personal review'. In: *Journal of Economic Behavior & Organization* 27 (3), 329–344. doi: 10.1016/0167-2681(94)00071-l (252).

Güth, Werner (2008). '(Non)Behavioral Economics'. In: *Journal of Psychology* 216 (4), 244–253. doi: 10.1027/0044-3409.216.4.244 (249, 252).

Güth, Werner and Hartmut Kliemt (2004). 'Perfect or Bounded Rationality?' In: *Analyse & Kritik* 26 (2), 364–381. doi: 10.1515/auk-2004-0203 (252).

Geisendorf, Sylvie (2009). 'The irrational foundations of neoclassical economics. How a simplified Newtonism shaped the perception of the economy'. In: *Der neoliberale Markt-Diskurs. Ursprünge, Geschichte, Wirkungen.* Ed. by Walter Ötsch and Claus Thomasberger. Metropolis-Verlag, Marburg, 151–166 (224, 250).

Geisendorf, Sylvie (2010). 'The irrational foundations of economics. Why neoclassical economists choose the wrong theory from physics'. In: *TRANS Internet-Zeitschrift für Kulturwissenschaften.* url: http://www.inst.at/trans/17Nr/2-9/2-9_geisendorf17.htm (224, 250).

Gell-Mann, Murray and Constantino Tsallis, eds. (2004). *Nonextensive Entropy – Interdisciplinary Applications.* Oxford University Press (204, 206).

Gelman, Andrew and Cosma R. Shalizi (2012). 'Philosophy and the practice of Bayesian statistics'. In: *British Journal of Mathematical and Statistical Psychology* 66 (1), 8–38. doi: 10.1111/j.2044-8317.2011.02037.x (185).

George, Donald A. R. (2011). 'Shifting Sands: Non-Linearity, Complexity and Randomness in Economics'. In: *Journal of Economic Surveys.* Special Issue on Nonlinearity, Randomness and Complexity 25 (3), 634–637. doi: 10.1111/j.1467-6419.2011.00687.x (6).

Ghirardato, Paolo, Fabio Maccheroni and Massimo Marinacci (2004). 'Differentiating ambiguity and ambiguity attitude'. In: *Journal of Economic Theory* 118 (2), 133–173. doi: 10.1016/j.jet.2003.12.004 (258).

Giannakouros, Panayotis and Lihua Chen (2015). 'A problem-solving approach to data analysis for economics'. In: *Forum for Social Economics* 47 (1), 87–114. doi: 10.1080/07360932.2015.1078737 (185, 186).

Gibbs, J. Willard (1892). *Thermodynamische Studien.* Trans. by Wilhem Ostwald. Engelmann Verlag, Leipzig (220).

Gibbs, J. Willard (1902). *Elementary Principles in Statistical Mechanics. Developed with Especial Reference to the Rational Foundation of Thermodynamics.* Charles Scribner's Sons, New York (36, 210, 211, 220–222, 248, 278).

Gigerenzer, Gerd (2000a). *Adaptive Thinking. Rationality in the real world.* Oxford University Press (251).

Gigerenzer, Gerd (2000b). 'How to Make Cognitive Illusions Disappear'. In: *Adaptive Thinking.* Oxford University Press, New York. Chap. 12, 241–266 (251).

Gigerenzer, Gerd, Ralph Hertwig and Thorsten Pachur, eds. (2011). *Heuristics. The Foundations of Adaptive Behavior.* Oxford University Press (248, 251).

Gigerenzer, Gerd and Reinhard Selten, eds. (2002a). *Bounded Rationality. The Adaptive Toolbox.* Berlin, March 14-19, 1999. Report of the 84th Dahlem Workshop. MIT Press (100, 251).

Gigerenzer, Gerd and Reinhard Selten (2002b). 'Rethinking Rationality'. In: *Bounded Rationality. The Adaptive Toolbox.* Ed. by Gerd Gigerenzer and Reinhard Selten.

Report of the 84th Dahlem Workshop. MIT Press, 1–12 (251, 252).

Gigerenzer, Gerd, Peter M. Todd and ABC Research Group, eds. (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press (251).

Gihman, Iosif I. and Anatoliĭ V. Skorokhod (2004). *The Theory of Stochastic Processes I*. Springer Berlin Heidelberg. doi: 10.1007/978-3-642-61943-4 (77).

Gilboa, Itzhak and David Schmeidler (1989). 'Maxmin expected utility with non-unique prior'. In: *Journal of Mathematical Economics* 18 (2), 141–153. doi: 10.1016/0304-4068(89)90018-9 (258).

Glimcher, Paul W. and Ernst Fehr, eds. (2014). *Neuroeconomics. Decision Making and the Brain*. 2nd ed. Elsevier, London. doi: 10.1016/C2011-0-05512-6 (136).

Glimcher, Paul W. and Aldo Rustichini (2004). 'Neuroeconomics: The Consilience of Brain and Decision'. In: *Science* 306 (5695), 447–452. doi: 10.1126/science.1102566 (283).

Gnedenko, Boris V. and Andrei N. Kolmogorov (1954). *Limit distributions for sums of independent random variables*. Trans. by Kai L. Chung. Addison-Wesley, Cambridge, MA (205).

Gode, Dhananjay K. and Shyam Sunder (1993). 'Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality'. In: *Journal of Political Economy* 101 (1), 119–137. doi: 10.1086/261868 (176, 224, 248).

Gode, Dhananjay K. and Shyam Sunder (1997). 'What Makes Markets Allocationally Efficient?' In: *Quarterly Journal of Economics* 112 (2), 603–630. doi: 10.1162/003355397555307 (176, 248).

Goldenfeld, Nigel (2005). *Lectures on Phase Transitions and the Renormalization Group*. Levant Books, Kolkata (234).

Goodhart, Charles A. E. (1975). 'Monetary relationships: a view from Threadneedle Street'. In: *Papers in Monetary Economics* 1 (236, 237).

Granger, Clive W. J. and Paul Newbold (1986). *Forecasting Economic Time Series*. Academic press, San Diego (194).

De Grauwe, Paul (2010). 'Top-Down versus Bottom-Up Macroeconomics'. In: *CESifo Economic Studies* 56 (4), 465–497. doi: 10.1093/cesifo/ifq014 (251).

De Grauwe, Paul (2012). *Lectures on Behavioral Macroeconomics*. Princeton University Press (251).

Graves, Timothy, Robert Gramacy, Nicholas Watkins and Christian Franzke (2017). 'A Brief History of Long Memory: Hurst, Mandelbrot and the Road to ARFIMA, 1951-1980'. In: *Entropy* 19 (9), 437. doi: 10.3390/e19090437 (200, 202).

Gray, Robert M. (2011a). *Entropy and Information Theory*. Springer, New York. doi: 10.1007/978-1-4419-7970-4 (78).

Gray, Robert M. (2011b). *Probability, Random Processes, and Ergodic Properties*. 2nd ed. Springer, New York. doi: 10.1007/978-1-4419-1090-5 (78).

Grazzini, Jakob (2012). 'Analysis of the Emergent Properties: Stationarity and Ergodicity'. In: *Journal of Artificial Societies and Social Simulation* 15 (2). doi: 10.18564/jasss.1929 (196).

Grazzini, Jakob and Matteo Richiardi (2015). 'Estimation of ergodic agent-based models by simulated minimum distance'. In: *Journal of Economic Dynamics and Control* 51, 148–165. doi: 10.1016/j.jedc.2014.10.006 (196).

GREEN, BENJAMIN and TERENCE TAO (2008). 'The primes contain arbitrarily long arithmetic progressions'. In: *Annals of Mathematics* 167 (2), 481–547. DOI: 10.4007/annals.2008.167.481 (78).

GRIMBERGEN, ARD JAN, JEROEN SIEBRING, ANA SOLOPOVA and OSCAR P. KUIPERS (2015). 'Microbial bet-hedging: the power of being different'. In: *Current Opinion in Microbiology* 25, 67–72. DOI: 10.1016/j.mib.2015.04.008 (214).

HAAVELMO, TRYGVE (1944). 'The Probability Approach in Econometrics'. In: *Econometrica* 12 (Supplement Jul.), iii–vi+1–115. DOI: 10.2307/1906935 (185).

HACKING, IAN (1980). 'Strange Expectations'. In: *Philosophy of Science* 47 (4), 562–567. DOI: 10.1086/288956 (100).

HÖFLICH, PAUL (1927). 'Wahrscheinlichkeitstheoretische Begründung der Ergodenhypothese'. In: *Zeitschrift für Physik A Hadrons and nuclei* 41 (8-9), 636–673. DOI: 10.1007/bf01390493 (69).

HAFNER, WOLFGANG and HEINZ ZIMMERMANN, eds. (2009). *Vinzenz Bronzin's Option Pricing Models. Exposition and Appraisal*. Springer, Berlin Heidelberg. DOI: 10.1007/978-3-540-85711-2 (90).

HAGAR, AMIT (2005). 'Discussion: The Foundations of Statistical Mechanics – Questions and Answers'. In: *Philosophy of Science* 72 (3), 468–478. DOI: 10.1086/498474 (71).

HAKANSSON, NILS H. (1971). 'On Optimal Myopic Portfolio Policies, With and Without Serial Correlation of Yields'. In: *Journal of Business* 44 (3), 324–334. JSTOR: 2351346 (178).

HALMOS, PAUL R. (1939). 'Invariants of Certain Stochastic Transformations: The Mathematical Theory of Gambling Systems'. In: *Duke Mathematical Journal* 5 (2), 461–478. DOI: 10.1215/s0012-7094-39-00538-7 (78).

HALMOS, PAUL R. (1941). 'The Decomposition of Measures'. In: *Duke Mathematical Journal* 8 (2), 386–392. DOI: 10.1215/s0012-7094-41-00830-x (78).

HALMOS, PAUL R. (1944). 'Approximation Theories for Measure Preserving Transformations'. In: *Transactions of the American Mathematical Society* 55 (1), 1. DOI: 10.2307/1990137 (78).

HALMOS, PAUL R. (1946). 'An Ergodic Theorem'. In: *Proceedings of the National Academy of Sciences* 32 (5), 156. URL: http://www.pnas.org/content/32/5/156.full.pdf (78, 219).

HALMOS, PAUL R. (1947). 'Invariant Measures'. In: *Annals of Mathematics* 48 (3), 735–754. DOI: 10.2307/1969138 (78).

HALMOS, PAUL R. (1949). 'Measurable transformations'. In: *Bulletin of the American Mathematical Society* 55 (11), 1015–1035. DOI: 10.1090/s0002-9904-1949-09305-9 (78).

HALMOS, PAUL R. (1956). *Lectures on Ergodic Theory*. Chelasea Publishing Company, New York (77).

HALMOS, PAUL R. (1958). 'Von Neumann on Measure and Ergodic Theory'. In: *Bulletin of the American Mathematical Society* 64 (3), 86–95. DOI: 10.1090/s0002-9904-1958-10203-7 (78).

HALMOS, PAUL R. (1961). 'Recent progress in ergodic theory'. In: *Bulletin of the American Mathematical Society* 67 (1), 70–81. DOI: 10.1090/s0002-9904-1961-10496-5 (78).

HAMILTON, JAMES D. (1994). *Time Series Analysis*. Princeton University Press (194).

HANEL, RUDOLF, STEFAN THURNER and MURRAY GELL-MANN (2011). 'Generalized entropies and the transformation group of superstatistics'. In: *Proceedings of the National Academy* *of Sciences* 108 (16), 6390–6394. DOI: 10.1073/pnas.1103539108 (205, 206).

Hanel, Rudolf, Stefan Thurner and Murray Gell-Mann (2014). 'How multiplicity determines entropy and the derivation of the maximum entropy principle for complex systems'. In: *Proceedings of the National Academy of Sciences* 111 (19), 6905–6910. DOI: 10.1073/pnas.1406071111 (205).

Hannan, Edward J. (1970). *Multiple Time Series*. John Wiley & Sons, New York. DOI: 10.1002/9780470316429 (194).

Harris, Theodore E. and Herbert Robbins (1953). 'Ergodic Theory of Markov Chains Admitting an Infinite Invariant Measure'. In: *Proceedings of the National Academy of Sciences* 39 (8), 860. URL: http://www.pnas.org/content/39/8/860.full.pdf (219).

Haug, Espen G. and Nassim N. Taleb (2011). 'Option traders use (very) sophisticated heuristics, never the Black–Scholes–Merton formula'. In: *Journal of Economic Behavior & Organization* 77 (2), 97–106. DOI: 10.1016/j.jebo.2010.09.013 (90).

Hayden, Benjamin Y. and Michael L. Platt (2009). 'The mean, the median, and the St. Petersburg paradox'. In: *Judgement and Decision Making* 4 (4), 256–272. URL: http://journal.sjdm.org/9226/jdm9226.html (91).

Von Hayek, Friedrich A. (1945). 'The Use of Knowledge in Society'. In: *American Economic Review* 35 (4), 519–530. JSTOR: 1809376 (198).

Von Hayek, Friedrich A. (1975). 'The Pretence of Knowledge'. In: *Swedish Journal of Economics* 77 (4). Nobel Memorial Lecture held December 11, 1974., 433. DOI: 10.2307/3439337 (198).

Von Hayek, Friedrich A. (1994). 'The Theory of Complex Phenomena'. In: *Readings in the Philosphy of Social Science*. Ed. by Michael Martin and Lee C. McIntyre. MIT Press, Cambridge. Chap. 4, 55–69 (229, 230).

Von Hayek, Friedrich A. (1989). 'The Pretence of Knowledge. Nobel Memorial Lecture, December 11, 1974'. In: *American Economic Review*. first published in 1974, 3–7. JSTOR: 1914347 (229).

He, Y., Stas Burov, Ralf Metzler and Eli Barkai (2008). 'Random Time-Scale Invariant Diffusion and Transport Coefficients'. In: *Physical Review Letters* 101 (5), 058101. DOI: 10.1103/physrevlett.101.058101 (208).

Hedlund, Gustav A. (1939). 'The Dynamics of Geodesic Flows'. In: *Bulletin of the American Mathematical Society* 45 (4), 241–261. DOI: 10.1090/s0002-9904-1939-06945-0 (74).

Heidegger, Martin (2002). *Sein und Zeit*. 11 unveränderte. Niemeyer, Tübingen (215).

Heise, Arne (2012). 'When the facts change, I change my mind ... Some developments in the economic scientific community and the situation in Germany'. In: *real world economics review* (62), 83–97. URL: www.paecon.net/PAEReview/issue62/Heise62.pdf (242).

Hempel, Carl G. and Paul Oppenheim (1948). 'Studies in the Logic of Explanation'. In: *Philosophy of Science*, 135–175. DOI: 10.1086/286983 (225).

Hendry, David F. (1980). 'Econometrics-Alchemy or Science?' In: *Economica* 47 (188), 387–406. DOI: 10.2307/2553385 (187).

Hendry, David F. (1995). *Dynamic Econometrics*. Oxford University Press. DOI: 10.1093/0198283164.001.0001 (196, 197).

Hendry, David F. (2000a). *Econometrics: Alchemy of Science? Essays in Econometric Methodology*. 2nd ed. Oxford University Press. DOI: 10.1093/0198293542.001.0001 (197).

HENDRY, DAVID F. (2000b). 'On detectable and non-detectable structural change'. In: *Structural Change and Economic Dynamics* 11 (1-2), 45–65. DOI: 10.1016/s0954-349x(00)00020-5 (197).

HENDRY, DAVID F. (2012). 'Mathematical Models and Economic Forecasting: Some Uses and Mis-Uses of Mathematics in Economics'. In: *Probabilities, Laws, and Structures*. Ed. by DENNIS DIEKS, WENCESLAO J. GONZÁLEZ, STEPHAN HARTMANN, MICHAEL STÖLTZNER and MARCEL WEBER. Springer, Dordrecht, 319–335. DOI: 10.1007/978-94-007-3030-4_23 (197).

HENDRY, DAVID F. and GRAYHAM E. MIZON (2011a). 'Econometric Modelling of Time Series with Outlying Observations'. In: *Journal of Time Series Econometrics* 3 (1), 1–24. DOI: 10.2202/1941-1928.1100 (197).

HENDRY, DAVID F. and GRAYHAM E. MIZON (2011b). 'What Needs Rethinking in Macroeconomics?' In: *Global Policy* 2 (2), 176–183. DOI: 10.1111/j.1758-5899.2011.00082.x (197).

HENDRY, DAVID F. and GRAYHAM E. MIZON (2014). 'Unpredictability in economic analysis, econometric modeling and forecasting'. In: *Journal of Econometrics*. Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions 182 (1), 186–195. DOI: 10.1016/j.jeconom.2014.04.017 (197).

HENDRY, DAVID F. and MARY S. MORGAN (1997). *The Foundations of Econometric Analysis*. Cambridge University Press (197).

HENDRY, DAVID F. and BENT NIELSEN (2007). *Econometric Modeling. A Likelihood Approach*. Princeton University Press (197).

HENS, THORSTEN and MARC OLIVER RIEGER (2016). *Financial Economics. A Concise Introduction to Classical and Behavioral Finance*. 2nd ed. Springer, Berlin Heidelberg. DOI: 10.1007/978-3-662-49688-6 (100, 179, 180).

HENS, THORSTEN and KLAUS R. SCHENK-HOPPÉ (2004). 'Survival of the Fittest on Wall Street'. In: *Institutioneller Wandel, Marktprozesse und dynamische Wirtschaftspolitik. Perspektiven der Evolutorischen Ökonomik*. Ed. by MARCO LEHMANN-WAFFENSCHMIDT, ALEXANDER EBNER and DIRK FORNAHL. Metropolis, Marburg, 393–367. URL: https://www.economics.ku.dk/research/publications/wp/2004/0403.pdf/ (179).

HENS, THORSTEN and KLAUS R. SCHENK-HOPPÉ (2005). 'Evolutionary stability of portfolio rules in incomplete markets'. In: *Journal of Mathematical Economics* 41 (1-2), 43–66. DOI: 10.1016/j.jmateco.2003.01.001 (179, 180).

HENS, THORSTEN and KLAUS R. SCHENK-HOPPÉ, eds. (2009). *Handbook of Financial Markets. Dynamics and Evolution*. North Holland, Amsterdam. DOI: 10.1016/s1568-4997(08)x0004-6 (173, 178).

HERNDON, THOMAS, MICHAEL ASH and ROBERT POLLIN (2014). 'Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff'. In: *Cambridge journal of economics* 38 (2), 257–279. DOI: 10.1093/cje/bet075 (223).

HERTWIG, RALPH, ULRICH HOFFRAGE and ABC RESEARCH GROUP, eds. (2013). *Simple Heuristics in a Social World* (251).

HESSE, JAN-OTMAR (2010). *Wirtschaft als Wissenschaft. Die Volkswirtschaftslehre in der frühen Bundesrepublik*. Campus Verlag, Frankfurt am Main (190).

HEUKELOM, FLORIS (2015). 'Prospect Theory'. In: *International Encyclopedia of the Social & Behavioral Sciences*. Ed. by JAMES D. WRIGHT. 2nd ed. Elsevier, 261–269. DOI: 10.1016/b978-0-08-097086-8.03193-7 (134).

HICKS, JOHN R. (1980). 'IS-LM: An Explanation'. In: *Journal of Post Keynesian Economics*

3 (2), 139–154. DOI: 10.1080/01603477.1980.11489209 (192).

HILBERT, DAVID (1900). 'Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900'. In: *Nachrichten der königlichen Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse* Heft 3, 253–297. URL: http://www.digizeitschriften.de/dms/resolveppn/?PPN=GDZPPN002498863 (25, 225).

HILBERT, MARTIN (2017). 'Complementary Variety: When Can Cooperation in Uncertain Environments Outperform Competitive Selection?' In: *Complexity* 2017, 1–15. DOI: 10.1155/2017/5052071 (214).

HILDENBRAND, WERNER (1971). 'Random preferences and equilibrium analysis'. In: *Journal of Economic Theory* 3 (4), 414–429. DOI: 10.1016/0022-0531(71)90039-1 (171).

HILDENBRAND, WERNER (1974). *Core and Equilibria of a Large Economy*. Princeton University Press (180).

HILDENBRAND, WERNER (1980). 'On the Uniqueness of Mean Demand for Dispersed Families of Preferences'. In: *Econometrica* 48 (7), 1703–1710. DOI: 10.2307/1911929 (180).

HILDENBRAND, WERNER (1994). *Market Demand. Theory and Empirical Evidence*. Princeton University Press (180).

HIRTE, KATRIN (2010). 'Performativity of Economics – ein tragfähiger Ansatz zur Analyse der Rolle der Ökonomen in der Ökonomie?' In: *Krise! Welche Krise? Zur Problematik aktueller Krisendebatten*. Ed. by WALTER O. ÖTSCH, KATRIN HIRTE and JÜRGEN NORDMANN. Metropolis-Verlag, Marburg, 49–75 (238).

HIRTE, KATRIN (2014). 'Performative Wissenschaft. Ökonomiekritik, Ökonomietheorien und die Verantwortung von ÖkonomInnen'. In: *Wissen! Welches Wissen? Zu Wahrheit, Theorien und Glauben sowie ökonomischen Theorien*. Ed. by KATRIN HIRTE, SEBASTIAN THIEME and WALTER O. ÖTSCH. Metropolis-Verlag, Marburg, 267–301 (238).

HIRTE, KATRIN and SEBASTIAN THIEME (2013). 'Mainstream, Orthodoxie und Heterodoxie. Zur Debatte um die Ausrichtung sowie einer Klassifizierung der Wirtschaftswissenschaften'. In: *Discussion Papers Institute for Comprehensive Analysis of Economy, Universität Linz* 16. URL: https://www.jku.at/fileadmin/gruppen/108/ICAE_Working_Papers/wp16.pdf (242).

HOPF, EBERHARD (1932). 'Complete Transitivity and the Ergodic Principle'. In: *Proceedings of the National Academy of Sciences* 18 (2), 204. URL: http://www.pnas.org/content/18/2/204.short (63, 70, 219).

HOPF, EBERHARD (1937). *Ergodentheorie*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-86630-2 (77).

HOPF, EBERHARD (1939). 'Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung'. In: *Berichte über die Verhandlungen der sächsichen Akademie der Wissenschaften zu Leipzig* 91, 261–304. DOI: 10.1007/978-3-642-86630-2 (74, 75).

HOPF, EBERHARD (1954). 'The General Temporally Discrete Markoff Process'. In: *Journal of Rational Mechanics and Analysis* 3, 13–45. JSTOR: 24900278 (56).

HOPF, EBERHARD (1971). 'Ergodic Theory and the Geodesic Flow on Surfaces of Constant Negative Curvature'. In: *Bulletin of the American Mathematical Society* 77 (6), 863–878. DOI: 10.1090/s0002-9904-1971-12799-4 (74, 75).

HORST, ULRICH (2005). 'Financial price fluctuations in a stock market model with many interacting agents'. In: *Economic Theory* 25 (4). DOI: 10.1007/s00199-004-0500-x (178).

Horst, Ulrich (2008). 'Ergodicity and Non-Ergodicity in Economics'. In: *The New Palgrave Dictionary of Economics*. Ed. by Steven N. Durlauf and Lawrence E. Blume. DOI: 10.1057/9780230226203.0499 (170, 171).

Horst, Ulrich and José A. Scheinkman (2006). 'Equilibria in systems of social interactions'. In: *Journal of Economic Theory* 130 (1), 44–77. DOI: 10.1016/j.jet.2005.02.012 (171, 177).

Horst, Ulrich and José A. Scheinkman (2009). 'A limit theorem for systems of social interactions'. In: *Journal of Mathematical Economics* 45 (9-10), 609–623. DOI: 10.1016/j.jmateco.2007.12.003 (171, 177).

Horst, Ulrich and Jan Wenzelburger (2007). 'On non-ergodic asset prices'. In: *Economic Theory* 34 (2), 207–234. DOI: 10.1007/s00199-006-0175-6 (173, 177, 178).

Hossenfelder, Sabine (2018). *Lost in Math. How Beauty Leads Physics Astray.* Basic Books, New York (227).

Host, Bernard and Bryna Kra (2005a). 'Convergence of polynomial ergodic averages'. In: *Israel Journal of Mathematics* 149 (1), 1–19. DOI: 10.1007/bf02772534 (78).

Host, Bernard and Bryna Kra (2005b). 'Nonconventional ergodic averages and nilmanifolds'. In: *Annals of Mathematics* 161 (1), 397–488. DOI: 10.4007/annals.2005.161.397 (78).

Hunsaker, Jerome and Saunders Mac Lane (1973). 'Edwin B. Wilson 1879–1964'. In: *National Academy of Sciences*, 285–320. URL: http://www.nasonline.org/member-directory/deceased-members/20001544.html (219).

Huschens, Stefan (2017). *Theorie und Methodik der Statistik.* Dresdner Beiträge zu Quantitativen Verfahren 70. Lecture Notes. URL: http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-222630 (175, 194).

Hutson, Matthew (2018). 'Artificial intelligence faces reproducibility crisis'. In: *Science* 359 (6377), 725–726. DOI: 10.1126/science.359.6377.725 (187).

Huygens, Christiaan (1657). *De ratiociniis in ludo aleae. On Reasoning in Games of Chance.* Ex officinia J. Elsevirii. Google Books: yS76XzFNcpMC (256, 278).

Irwin, Joseph O. (1967). 'William Allen Whitworth and a Hundred Years of Probability'. In: *Journal of the Royal Statistical Society. Series A (General)* 130 (2), 147–176. DOI: 10.2307/2343399 (162).

Itô, Kiyosi (1944). 'Stochastic integral'. In: *Proceedings of the Imperial Academy* 20 (8), 519–524. DOI: 10.3792/pia/1195572786 (278).

Jacobs, Konrad (1960). *Neuere Methoden und Ergebnisse der Ergodentheorie.* Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-94777-3 (74, 77).

Jacobs, Konrad (1965). 'Einige neuere Ergebnisse der Ergodentheorie.' In: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 67, 143–182. URL: http://eudml.org/doc/146529 (78).

Jakobson, M. V. (1981). 'Absolutely continuous invariant measures for one-parameter families of one-dimensional maps'. In: *Communications in Mathematical Physics* 81 (1), 39–88. DOI: 10.1007/bf01941800 (73).

Jarzynski, Christopher (2011). 'Single-molecule experiments: Out of equilibrium'. In: *Nature Physics* 7 (8), 591–592. DOI: 10.1038/nphys2054 (208).

Jaynes, Edwin T. (1957b). 'Information Theory and Statistical Mechanics. II'. In: *Physical Re-*

view 108 (2), 171–190. DOI: 10.1103/PhysRev.108.171 (254, 280, 281).

Jaynes, Edwin T. (1957a). 'Information Theory and Statistical Mechanics'. In: *Physical Review* 106 (4), 620–630. DOI: 10.1103/PhysRev.106.620 (254, 280–282).

Jaynes, Edwin T. (1971). 'Violation of Boltzmann's *H* Theorem in Real Gases'. In: *Physical Review A* 4 (2), 747–750. DOI: 10.1103/physreva.4.747 (71).

Jensen, Johan L. W. V. (1906). 'Sur les fonctions convexes et les inégalités entre les valeurs moyennes'. In: *Acta Mathematica* 30 (1), 175–193. DOI: 10.1007/bf02418571 (118).

Jevons, William S. (1888). *Theory of Political Economy*. 3rd ed. Macmillan & Co., London (278).

Johansen, Søren (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press. DOI: 10.1093/0198774508.001.0001 (198).

Johnson, Timothy C. (2015). 'Reciprocity as a Foundation of Financial Economics'. In: *Journal of Business Ethics* 131 (1), 43–67. DOI: 10.1007/s10551-014-2257-x (81, 82).

Johnson, Timothy C. (2016). 'The necessity of multi-disciplinary scholarship for finance: On Ayache and Roffe'. In: *Finance and Society* 2 (2), 189–204. DOI: 10.2218/finsoc.v2i2.1733 (14, 90, 105, 167).

Jorland, Gérard (1987). 'The Saint Petersburg Paradox 1713-1937'. In: *The Probabilistic Revolution*. Ed. by Lorenz Krüger, Lorraine J. Daston and Michael Heidelberger. Vol. 1: Ideas in History. MIT Press, Cambridge. Chap. 7, 157–190 (92, 100, 101, 103, 108).

Jost, Jürgen (2005). *Dynamical Systems*. Springer, Berlin Heidelberg. DOI: 10.1007/3-540-28889-9 (77).

Del Junco, Andrés (2009). 'Ergodic Theorems'. In: *Encyclopedia of Complexity and Systems Science*. Ed. by Robert A. Meyers. Springer, New York, 2933–2956. DOI: 10.1007/978-0-387-30440-3_176 (173).

Juselius, Katarina (2006). *The Cointegrated VAR Model*. Oxford University Press (198).

Kahneman, Daniel and Amos Tversky (1979). 'Prospect theory: An analysis of decision under risk'. In: *Econometrica* 47 (2), 263–291. DOI: 10.2307/1914185 (119, 126, 134, 135, 248, 256, 258).

Kakutani, Shizuo (1952). 'Ergodic Theory'. In: *Proceedings of the International Congress of Mathematicians. Cambridge, Massachusetts, U. S. A. 1950 Under Auspices of the American Mathematical Society*. Ed. by Lawrence M. Grayes, Einar Hille, Paul A. Smith and Oscar Zabiski. Vol. II. American Mathematical Society, Providence, 128–142. URL: https://www.mathunion.org/fileadmin/ICM/Proceedings/ICM1950.2/ICM1950.2.ocr.pdf (78).

Kallianpur, Gopinath and Herbert Robbins (1953). 'Ergodic Property of the Brownian Motion Process'. In: *Proceedings of the National Academy of Sciences* 39 (6), 525. URL: http://www.pnas.org/content/39/6/525.full.pdf (219).

Van Kampen, Nico G. (1998). 'Remarks on Non-Markov Processes'. In: *Brazilian Journal of Physics* 28 (2), 90–96. DOI: 10.1590/S0103-97331998000200003 (53, 54).

Van Kampen, Nico G. (2007). *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam. DOI: 10.1016/b978-0-444-52965-7.x5000-4 (61, 193).

Kantorovich, Leonid V. (1939). *Mathematical Methods in the Organization and Planning of Production*. (in Russian). English translation in Kantorovich (1960). Leningrad University (39).

Kantorovich, Leonid V. (1942). 'On the Translocation of Masses'. (in Russian). In: *Doklady Akademiĭ Nauk* 37 (7-8). English translation in Kantorovich (2006), 227–229 (39).

Kantorovich, Leonid V. (1960). 'Mathematical Methods of Organizing and Planning Production'. In: *Management Science* 6 (4). Translated by Robert W. Campbell and W. H. Marlow, 366–422. doi: 10.1287/mnsc.6.4.366 (314).

Kantorovich, Leonid V. (2006). 'On the Translocation of Masses'. In: *Journal of Mathematical Sciences* 133 (4). Translated by A. N. Sobolevskiĭ, 1381–1382. doi: 10.1007/s10958-006-0049-2 (314).

Kantz, Holger and Thomas Schreiber (2003). *Nonlinear Time Series Analysis*. Cambridge University Press. doi: 10.1017/cbo9780511755798 (194).

Kapeller, Jakob (2011). 'Was sind ökonomische Modelle?' In: *Philosophie und Wirtschaftswissenschaft*. Ed. by Volker Gadenne and Reinhard Neck. Mohr Siebeck, Tübingen, 29–50 (249).

Kapetanios, George and Yongcheol Shin (2011). 'Testing the Null Hypothesis of Nonstationary Long Memory Against the Alternative Hypothesis of a Nonlinear Ergodic Model'. In: *Econometric Reviews* 30 (6), 620–645. doi: 10.1080/07474938.2011.553568 (195, 196).

Körber, Hans-Günther, ed. (1961). *Aus dem wissenschaftlichen Briefwechsel Wilhelm Ostwalds. 1. Teil*. Briefwechsel mit Ludwig Boltzmann, Max Plack, Georg Helm und Josiah Willard Gibbs. Akademie-Verlag, Berlin (220, 221).

Kauffman, Stuart A. (1993). *The Origins of Order. Self Organization and Selection in Evolution*. Oxford University Press (16).

Kelly Jr., John L. (1956). 'A New Interpretation of Information Rate'. In: *Bell System Technical Journal* 35 (4), 917–926. doi: 10.1002/j.1538-7305.1956.tb03809.x (170, 174, 175, 178, 213, 259, 264, 266, 267, 278).

Keynes, John M. (1921). *A Treatise on Probability*. Macmillan & Co., London (84, 100, 103, 189).

Keynes, John M. (1924). 'Alfred Marshall, 1842-1924'. In: *Economic Journal* 34 (135), 311–372. doi: 10.2307/2222645 (229).

Keynes, John M. (1936). *The General Theory of Employment, Interest and Money*. Macmillan & Co., London (189, 192).

Keynes, John M. (1937). 'The General Theory of Employment'. In: *Quarterly Journal of Economics* 51 (2), 209–223. doi: 10.2307/1882087 (209, 210).

Keynes, John M. (1939). 'Professor Tinbergen's Method'. In: *Economic Journal* 49 (195), 558–568. JSTOR: 2224838 (182–184).

Keynes, John M. (1940). 'On a Method of Statistical Business-Cycle Research. A Comment'. In: *Economic Journal* 50 (197), 154–156. doi: 10.2307/2225764 (182).

Khinchin, Alexander I. (1949). *Mathematical Foundations of Statistical Mechanics*. Dover Publications, New York (3, 13, 30, 33, 40, 42, 62, 63, 71).

Khinchin, Alexander I. (1957). *Mathematical Foundations of Information Theory*. Two papers originally published in Uspekhi Matematicheskikh Nauk translated by R.A. Silverman and M.D. Friedman. Dover Publications (41, 77).

KHINTCHINE, ALEXANDER YA. (1933). 'Zu Birkhoffs Lösung des Ergodenproblems'. In: *Mathematische Annalen* 107 (1), 485–488. DOI: 10.1007/bf01448905 (33, 42, 63).

KIM, YOUNG CHIN (1973). 'Choice in the Lottery-Insurance Situation Augmented-Income Approach'. In: *Quarterly Journal of Economics* 87 (1), 148–156. DOI: 10.2307/1885231 (111, 112).

KIRCHGÄSSNER, GEBHARD, JÜRGEN WOLTERS and UWE HASSLER (2013). *Introduction to Modern Time Series Analysis*. 2nd ed. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-33436-8 (7, 8, 187, 194).

KIRMAN, ALAN P. (2006). 'Demand Theory and General Equilibrium: From Explanation to Introspection, a Journey down the Wrong Road'. In: *History of Political Economy* 38 (Suppl 1), 246. DOI: 10.1215/00182702-2005-025 (36, 227).

KIRSTEIN, BERND (5th Feb. 2007). *Grundlagen der Maß- und Integrationstheorie*. unpublished lecture notes, Teil VIII. Chap. 17 Ergodentheorie (i).

KIRSTEIN, MARK (2012). 'Stochastische Finanzmarktmodelle der Preisevolution und Ursachen für die Instabilität von Finanzmärkten'. Diplomarbeit. Universität Leipzig (iii, 9, 205).

KIRSTEIN, MARK (2016). 'Von der Ergodenhypothese der Physik zum Ergodenaxiom in der Ökonomik. Genealogie einer Idee und ihre methodologischen Implikationen'. In: *Ökonomie! Welche Ökonomie? Stand und Status der Wirtschaftswissenschaften*. Ed. by JAKOB KAPELLER, STEPHAN PÜHRINGER, KATRIN HIRTE and WALTER O. ÖTSCH. Metropolis-Verlag, Marburg, 71–104. URL: bit.ly/FromEH2EA (23, 233).

KITCHENS, BRUCE (2007). 'Computational Ergodic Theory by GEON HO CHOE'. In: *Bulletin of the American Mathematical Society* 44 (10), 147–155. URL: http://www.ams.org/journals/bull/2007-44-01/S0273-0979-06-01120-7/S0273-0979-06-01120-7.pdf (77).

KLAFTER, JOSEPH and IGOR M. SOKOLOV (2011). *First Steps in Random Walks. From Tools to Applications*. Oxford University Press (208).

KLEIN, TONY (2017). 'Conditional Variance Modeling of Financial Time Series. Essays on Advances and Applications to Commodity and Foreign Exchange Markets'. PhD thesis. TU Dresden, Economics Department. DOI: 10.13140/RG.2.2.28851.12324 (202).

KLENKE, ACHIM (2014). *Probability Theory. A Comprehensive Course*. 2nd ed. Springer London. DOI: 10.1007/978-1-4471-5361-0 (77).

KOLESNIKOV, ALEXEI V. and THORSTEN RÜHL (2010). 'Ergodicity of financial indices'. In: *Europhysics Letters* 90 (3), 30004. DOI: 10.1209/0295-5075/90/30004 (214).

KOLMOGOROFF, ANDREI N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin Heidelberg. DOI: 10.1007/978-3-642-49888-6 (14, 140, 278).

KOLMOGOROFF, ANDREI N. (1937). 'Ein vereinfachter Beweis des Birkhoff-Khintchineschen Ergodensatzes'. In: *Matematicheskii Sbornik* 2 (44), 367–368. URL: http://eudml.org/doc/64907 (64).

KOLMOGOROV, ANDREI N. (1954). 'On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian'. (in Russian). In: *Doklady Akademiĭ Nauk USSR* 98 (527), 2–3 (75).

KOOPMANS, TJALLING C. (1965). 'On the Concept of Optimal Economic Growth'. In: *The Econometric Approach to Development Planning*. Study Week, 7-13 October 1963 (Pontifical Academy of Sciences). Scripta Varia 28. Part I. Academia Vaticana. Chap. 4, 225–287.

URL: http://www.pas.va/content/dam/accademia/pdf/sv28apas.pdf (172).

Korinek, Anton (1st Oct. 2015). 'Thoughts on DSGE Macroeconomics'. In: DOI: 10.2139/ssrn.3022009. URL: http://www.korinek.com/download/DSGEMacroEssay.pdf (171).

Korzybski, Alfred H. S. (1933). 'A Non-Aristotelian System and its Necessity for Rigour in Mathematics and Physics'. In: *Science and Sanity*, 747–761 (237).

Kra, Bryna (2005). 'The Green-Tao Theorem on arithmetic progressions in the primes: an ergodic point of view'. In: *Bulletin of the American Mathematical Society* 43 (1), 3–24. DOI: 10.1090/s0273-0979-05-01086-4 (78).

Krüger, Lorenz, Lorraine J. Daston and Michael Heidelberger, eds. (1987a). *The Probabilistic Revolution*. Vol. 1: Ideas in History. MIT Press (16, 78).

Krüger, Lorenz, Gerd Gigerenzer and Mary S. Morgan, eds. (1987b). *The Probabilistic Revolution*. Vol. 2: Ideas in the Sciences. MIT Press (16, 78).

Krasnosel'skii, Mark A. and Aleksei V. Pokrovskii (1989). *Systems with Hysteresis*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-61302-9 (198).

Krengel, Ulrich (1978). 'On the speed of convergence in the ergodic theorem'. In: *Monatshefte für Mathematik* 86 (1), 3–6. DOI: 10.1007/bf01300052 (65).

Krengel, Ulrich (1985). *Ergodic Theorems. With a Supplement by Antoine Brunel*. Walter de Gruyter, Berlin. DOI: 10.1515/9783110844641 (65, 77).

Krugman, Paul (1993). 'How I Work'. In: *American Economist* 37 (2), 25–31. DOI: 10.1177/056943459303700204 (19).

Kullback, Solomon and Richard A. Leibler (1951). 'On Information and Sufficiency'. In: *Annals of Mathematical Statistics* 22 (1), 79–86. JSTOR: 2236703 (206).

Kurz, Heinz D., ed. (2014). *Studien zur Entwicklung der ökonomischen Theorie XXVIII. Die Ökonomik im Spannungsfeld zwischen Natur- und Geisteswissenschaften. Alte und neue Perspektiven im Licht des jüngsten Methodenstreits*. Vol. 115. Schriften des Vereins für Socialpolitik. Duncker & Humblot, Berlin (226).

Kurz, Mordecai (2009). 'Rational Diverse Beliefs and Market Volatility'. In: *Handbook of Financial Markets. Dynamics and Evolution*. Ed. by Thorsten Hens and Klaus R. Schenk-Hoppé. North Holland, Amsterdam. Chap. 8, 439–506. DOI: 10.1016/b978-012374258-2.50012-9 (174).

Landau, Lev D. and Evgeny M. Lifshitz (1980). *Statistical Physics. Part I*. Volume 5 of Course of Theoretical Physics. 3rd ed. Pergamon Press, Oxford (71, 210).

Lanford III., Oscar E. (1973). 'Entropy and Equilibrium States in Classical Statistical Mechanics'. In: *Statistical Mechanics and Mathematical Problems*. Ed. by Andrew Lenard. Springer, Berlin Heidelberg, 1–113. DOI: 10.1007/bfb0112756 (74).

Lange, Kenneth (1979). 'On Cohen's stochastic generalization of the strong ergodic theorem of demography'. In: *Journal of Applied Probability* 16 (03), 496–504. DOI: 10.2307/3213079 (215).

Laplace, Pierre-Simon (1812). *Théorie Aanalytique des Probabilités*. Courcier, Paris. URL: bit.ly/MKLalace1812 (125, 140, 256, 278).

Laplace, Pierre-Simon (1902). *A Philosophical Essay on Probabilities*. John Wiley & Sons, London (12).

Lasota, Andrzej and Michael C. Mackey (1994). *Chaos, Fractals, and Noise. Stochastic*

Aspects of Dynamics. Springer New York. DOI: 10.1007/978-1-4612-4286-4 (77).

Latané, Henry A. (1959). 'Criteria for Choice Among Risky Ventures'. In: *Journal of Political Economy* 67 (2), 144–155. DOI: 10.1086/258157 (170, 265).

Lütkepohl, Helmut (2005). *New Introduction to Multiple Time Series Analysis*. Springer, Berlin Heidelberg. DOI: 10.1007/978-3-540-27752-1 (194).

Lauer, Thomas (1996). *Die Dynamik von Konsumgütermärkten. Eine evolutorische Sicht*. Physica, Heidelberg. DOI: 10.1007/978-3-662-01589-6 (181).

Lavis, David (1977). 'The Role of Statistical Mechanics in Classical Physics'. In: *British Journal for the Philosophy of Science* 28 (3), 255–279. DOI: 10.1093/bjps/28.3.255 (71).

Lavoie, Marc (2011). *Post Keynesian Economics. New Foundations*. Edward Elgar, Cheltenham (189).

Lévy, Paul (1937). *Théorie de l'Addition des Variables Aléatoires*. Gauthier-Villars, Paris (205, 278).

Lawson, Tony (1988). 'Probability and uncertainty in economic analysis'. In: *Journal of Post Keynesian Economics* 11 (1), 38–65. DOI: 10.1080/01603477.1988.11489724 (184).

Lawson, Tony (2009). 'Contemporary economics and the crisis'. In: *real-world economics review* (50), 122–131. URL: http://www.paecon.net/PAEReview/issue50/Lawson50.pdf (184).

Leamer, Edward E. (1983). 'Let's Take the Con Out of Econometrics'. In: *American Economic Review* 73 (1), 31–43. JSTOR: 1803924 (184, 187).

Leamer, Edward E. (2010). 'Tantalus on the Road to Asymptopia'. In: *Journal of Economic Perspectives* 24 (2), 31–46. DOI: 10.1257/jep.24.2.31 (184, 187).

Lebowitz, Joel L. and Elliott W. Montroll, eds. (1983). *Nonequilibrium Phenomena I The Boltzmann Equation*. Studies in Statistical Mechanics Volume X. North-Holland, Amsterdam (38, 39).

Lebowitz, Joel L. and Oliver Penrose (1973). 'Modern ergodic theory'. In: *Physics Today* 26 (2), 23–29. DOI: 10.1063/1.3127948 (220).

Leeds, Stephen (1989). 'Malament and Zabell on Gibbs Phase Averaging'. In: *Philosophy of Science* 56 (2), 325–340. DOI: 10.1086/289491 (71).

Leeson, Robert (1998). '"The Ghosts I Called I Can't Get Rid of Now": The Keynes-Tinbergen-Friedman-Phillips Critique of Keynesian Macroeconometrics'. In: *History of Political Economy* 30 (1), 51–94. DOI: 10.1215/00182702-30-1-51 (182).

Lehmann-Waffenschmidt, Marco (1990). 'Predictability of Economic Processes and the Morgenstern Paradox'. In: *Swiss Journal of Economics and Statistics* 126 (II), 147–161. URL: http://www.sjes.ch/papers/1990-II-4.pdf (238).

Lehmann-Waffenschmidt, Marco (2010). 'Contingency and Causality in Economic Processes – Conceptualizations, Formalizations, and Applications in Counterfactual Analysis'. In: *European Review* 18 (4), 481–505. DOI: 10.1017/S106279871000027X (15, 99, 234).

Leijonhufvud, Axel (1973). 'Life Among the Econ'. In: *Economic Inquiry* 11 (3), 327–337. DOI: 10.1111/j.1465-7295.1973.tb01065.x (199).

Lengwiler, Yvan (2009). 'The Origins of Expected Utility Theory'. In: *Vinzenz Bronzin's Option Pricing Models. Exposition and Appraisal*. Ed. by Wolfgang Hafner and Heinz Zim-

MERMANN. .Part F Finance, Economics and Actuarial Science. Springer, Berlin Heidelberg. Chap. 20, 535–545. DOI: 10.1007/978-3-540-85711-2_26 (112).

LENSBERG, TERJE and KLAUS R. SCHENK-HOPPÉ (2007). 'On the Evolution of Investment Strategies and the Kelly Rule—A Darwinian Approach∗'. In: *Review of Finance* 11 (1), 25–50. DOI: 10.1093/rof/rfm003 (179).

LEWONTIN, RICHARD C. and DAN COHEN (1969). 'On Population Growth in a Randomly Varying Environment'. In: *Proceedings of the National Academy of Sciences* 62 (4), 1056–1060. DOI: 10.1073/pnas.62.4.1056 (145, 213).

LI, TIEN-YIEN and JAMES A. YORKE (1975). 'Period Three Implies Chaos'. In: *American Mathematical Monthly* 82.10, 985–992. DOI: 10.1080/00029890.1975.11994008 (39).

LINDGREN, KRISTIAN (1992). 'Evolutionary Phenomena in Simple Dynamics'. In: *Artificial Life II*. Ed. by CHRISTOPHER G. LANGTON, CHARLES TAYLOR, J. DOYNE FARMER and STEEN RASMUSSEN. Vol. X. Santa Fe Institute Studies in the Sciences of Complexity, 295–312 (213).

VAN LITH, JANNEKE (2001). 'Ergodic Theory, Interpretations of Probability and the Foundations of Statistical Mechanics'. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32 (4), 581–594. DOI: 10.1016/s1355-2198(01)00027-2 (71).

LIU, CHUANG (2003). 'Spontaneous Symmetry Breaking and Chance in a Classical World'. In: *Philosophy of Science* 70 (3), 590–608. DOI: 10.1086/376786 (234).

LO, ANDREW W., H. ALLEN ORR and RUIXUN ZHANG (2018). 'The growth of relative wealth and the Kelly criterion'. In: *Journal of Bioeconomics* 20 (1), 49–67. DOI: 10.1007/s10818-017-9253-z (179, 214).

DE LONG, J. BRADFORD, ANDREI SHLEIFER, LAWRENCE H. SUMMERS and ROBERT J. WALDMANN (1990). 'Noise Trader Risk in Financial Markets'. In: *Journal of Political Economy* 98 (4), 703–738. DOI: 10.1086/261703 (176, 248).

LOPES, LOLA L. (1991). 'The Rhetoric of Irrationality'. In: *Theory & Psychology* 1 (1), 65–82. DOI: 10.1177/0959354391011005 (253).

LORENZ, EDWARD N. (1963). 'Deterministic Nonperiodic Flow'. In: *Journal of the Atmospheric Sciences* 20 (2), 130–141. DOI: 10.1175/1520-0469(1963)020<0130:dnf>2.0.co;2 (39).

LOSCHMIDT, JOSEF (1876a). 'Über den Zustand des Wärmegleichgewichtes eines Systemes von Körpern mit Rücksicht auf die Schwerkraft I'. In: *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien, mathematisch-naturwissenschaftliche Classe* 73, 128–142. URL: https://loschmidt.chemi.muni.cz/biography/ (48).

LOSCHMIDT, JOSEF (1876b). 'Über den Zustand des Wärmegleichgewichtes eines Systemes von Körpern mit Rücksicht auf die Schwerkraft II'. In: *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien, mathematisch-naturwissenschaftliche Classe* 73, 366–372. URL: https://loschmidt.chemi.muni.cz/biography/ (48).

LOSCHMIDT, JOSEF (1877). 'Über den Zustand des Wärmegleichgewichtes eines Systemes von Körpern mit Rücksicht auf die Schwerkraft III'. In: *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien, mathematisch-naturwissenschaftliche Classe* 75, 287–298. URL: https://loschmidt.chemi.muni.cz/biography/ (48).

LOSCHMIDT, JOSEF (1878). 'Über den Zustand des Wärmegleichgewichtes eines Systemes von Körpern mit Rücksicht auf die Schwerkraft IV'. In: *Sitzungsberichte der Kaiserlichen Akademie* *der Wissenschaften zu Wien, mathematisch-naturwissenschaftliche Classe* 76, 209–225. URL: https://loschmidt.chemi.muni.cz/biography/ (48).

Lotka, Alfred J. (1925). *Elements of Physical Biology*. Dover Publications (220).

Louçã, Francisco (1999). 'The econometric challenge to Keynes: arguments and contradictions in the early debates about a late issue'. In: *European Journal of the History of Economic Thought* 6 (3), 404–438. DOI: 10.1080/10427719900000074 (182).

Lucas, Robert E. Jr (1976). 'Econometric policy evaluation: A critique'. In: *Carnegie-Rochester Conference Series on Public Policy* 1 (1), 19–46. DOI: 10.1016/S0167-2231(76)80003-6 (188, 236, 237).

Lux, Thomas and Michele Marchesi (1999). 'Scaling and criticality in a stochastic multi-agent model of a financial market'. In: *Nature* 397 (6719), 498–500. DOI: 10.1038/17290 (176).

Lux, Thomas and Michele Marchesi (2000). 'Volatility Clustering in Financial Markets: A Microsimulation of Interacting Agents'. In: *International Journal of Theoretical and Applied Finance* 03 (04), 675–702. DOI: 10.1142/s0219024900000826 (176).

Lux, Thomas and Didier Sornette (2002). 'On Rational Bubbles and Fat Tails'. In: *Journal of Money, Credit and Banking* 34 (2), 589–610 (248).

Lv, Yingdong and Bernhard K. Meister (2010). 'Implication of the Kelly Criterion for Multi-Dimensional Processes'. In: *International Journal of Theoretical and Applied Finance* 13 (1), 93–112. DOI: 10.1142/s0219024910005693 (179).

MacKenzie, Donald (2003). 'An Equation and its Worlds. Bricolage, Exemplars, Disunity and Performativity in Financial Economics'. In: *Social Studies of Science* 33 (6), 831–868. DOI: 10.1177/0306312703336002 (238).

Mackey, George W. (1974). 'Ergodic Theory and Its Significance for Statistical Mechanics and Probability Theory'. In: *Advances in Mathematics* 12 (2), 178–268. DOI: 10.1016/s0001-8708(74)80003-4 (71).

MacLean, Leonard C., Edward O. Thorp and William T. Ziemba (2011a). 'Introduction to the Classic Papers and Theories'. In: *The Kelly Capital Growth Investment Criterion. Theory and Practice*. Ed. by Leonard C. MacLean, Edward O. Thorp and William T. Ziemba. .II: Classic Papers and Theories. World Scientific, New Jersey. Chap. 11, 143–146. DOI: 10.1142/9789814293501_0045 (175).

MacLean, Leonard C., Edward O. Thorp and William T. Ziemba (2011b). 'Introduction to the Early Ideas and Contributions'. In: *The Kelly Capital Growth Investment Criterion. Theory and Practice*. Ed. by Leonard C. MacLean, Edward O. Thorp and William T. Ziemba. .I: The Early Ideas and Contributions. World Scientific, New Jersey. Chap. 1, 3–10. DOI: 10.1142/9789814293501_0001 (175, 265).

MacLean, Leonard C., Edward O. Thorp and William T. Ziemba (2011c). 'Introduction to the Evidence of the Use of Kelly Type Strategies by the Great Investors and Others'. In: *The Kelly Capital Growth Investment Criterion. Theory and Practice*. Ed. by Leonard C. MacLean, Edward O. Thorp and William T. Ziemba. .VI: Evidence of the use of Kelly type strategies by the great investors and others. World Scientific, New Jersey. Chap. 45, 657–662. DOI: 10.1142/9789814293501_0045 (175).

MacLean, Leonard C., Edward O. Thorp and William T. Ziemba (2011d). 'Introduction to the Good and Bad Properties of Kelly'.

In: *The Kelly Capital Growth Investment Criterion. Theory and Practice*. Ed. by Leonard C. MacLean, Edward O. Thorp and William T. Ziemba. .IV: Critics and Assessing the Good and Bad Properties of Kelly. World Scientific, New Jersey. Chap. 30, 459–464. doi: 10.1142/9789814293501_0030 (175).

MacLean, Leonard C., Edward O. Thorp and William T. Ziemba (2011e). 'Introduction to the Relationship of Kelly Optimization to Asset Allocation'. In: *The Kelly Capital Growth Investment Criterion. Theory and Practice*. Ed. by Leonard C. MacLean, Edward O. Thorp and William T. Ziemba. .III: The Relationship of Kelly Optimization to Asset Allocation. World Scientific, New Jersey. Chap. 22, 301–304. doi: 10.1142/9789814293501_0022 (175).

MacLean, Leonard C., Edward O. Thorp and William T. Ziemba (2011f). 'Introduction to the Utility Foundations of Kelly'. In: *The Kelly Capital Growth Investment Criterion. Theory and Practice*. Ed. by Leonard C. MacLean, Edward O. Thorp and William T. Ziemba. .V: Utility Foundations. World Scientific, New Jersey. Chap. 40, 575–576. doi: 10.1142/9789814293501_0040 (175).

MacLean, Leonard C., Edward O. Thorp and William T. Ziemba, eds. (2011g). *The Kelly Capital Growth Investment Criterion. Theory and Practice*. World Scientific, New Jersey. doi: 10.1142/9789814293501 (175).

Majorana, Ettore (2005). 'The value of statistical laws in physics and social sciences'. In: *Quantitative Finance* 5 (2). Third translation of Majorana's article 'Il valore delle leggi statistiche nella fisica e nelle scienze sociali' published in 1942 after Majorana's disappearance by Giovanni Gentile Jr in *Scientia*, 1942, Vol. 36, pp. 55–66, 135–140. doi: 10.1080/14697680500148174 (16, 17).

Majumdar, Mukul and Tapan Mitra (1994). 'Robust ergodic chaos in discounted dynamic optimization models'. In: *Economic Theory* 4 (5), 677–688. doi: 10.1007/bf01212024 (74).

Malament, David B. and Sandy L. Zabell (1980). 'Why Gibbs Phase Averages Work–The Role of Ergodic Theory'. In: *Philosophy of Science* 47 (3), 339–349. doi: 10.1086/288941 (72).

Mañé, Ricardo (1987). *Ergodic Theory and Differentiable Dynamics*. Springer Berlin Heidelberg. doi: 10.1007/978-3-642-70335-5 (77).

Mandelbrot, Benoît B. (1963). 'The Variation of Certain Speculative Prices'. In: *Journal of Business* 36 (4), 394–419 (200).

Mandelbrot, Benoît B. (1965a). 'Self-Similar Error Clusters in Communication Systems and the Concept of Conditional Stationarity'. In: *IEEE Transactions on Communications* 13 (1), 71–90. doi: 10.1109/tcom.1965.1089090 (200).

Mandelbrot, Benoît B. (1965b). 'Time varying channels, 1/f noises, and the infrared catastrophe: or why does the low frequency energy sometimes seem infinite?' In: *IEEE Communication Convention* (200).

Mandelbrot, Benoît B. (1967). 'How long is the coast of Britain? Statistical self-similarity and fractional dimension'. In: *Science* 156 (3775), 636–638. doi: 10.1126/science.156.3775.636 (200, 278).

Mandelbrot, Benoît B. (1987). *Die fraktale Geometrie der Natur*. Akademie-Verlag, Berlin (39, 91, 206).

Mandelbrot, Benoît B. and Richard L. Hudson (2004). *Fraktale und Finanzen*. Piper Verlag, München (91).

Mandelbrot, Benoît B. and John W. Van Ness (1968). 'Fractional Brownian Motions, Fractional Noises and Applications'. In: *SIAM Review* 10 (4), 422–437. DOI: 10.1137/1010093 (200).

Mandelbrot, Benoît B. and James R. Wallis (1969). 'Computer Experiments with Fractional Gaussian Noises: Part 2, Rescaled Ranges and Spectra'. In: *Water Resources Research* 5 (1), 242–259. DOI: 10.1029/wr005i001p00242 (201).

Mandelbrot, Benoit B. (1997). *Fractals and Scaling in Finance. Discontinuity, Concentration, Risk.* Vol. Selecta Volume E. Springer New York. DOI: 10.1007/978-1-4757-2763-0 (91).

Mangabeira Unger, Roberto and Lee Smolin (2015). *The Singular Universe and the Reality of Time. A Proposal in Natural Philosophy.* Cambridge University Press (100, 216, 232).

Mantel, Rolf R. (1974). 'On the characterization of aggregate excess demand'. In: *Journal of Economic Theory* 7 (3), 348–353. DOI: 10.1016/0022-0531(74)90100-8 (180, 227).

Markowitz, Harry M. (1952). 'Portfolio selection'. In: *Journal of Finance* 7 (1), 77–91 (177, 278).

Markowitz, Harry M. (1959). *Portfolio Selection.* John Wiley & Sons, New York (153, 170).

Markowitz, Harry M. (1976). 'Investment for the Long Run New Evidence for an Old Rule'. In: *The Journal of Finance* 31 (5), 1273–1286. DOI: 10.1111/j.1540-6261.1976.tb03213.x (139).

Marx, Robert (2018). 'Die Wirkung von Extremereignissen auf die Solvenz von Finanzintermediären. Eine empirische und simulationsgestützte Analyse vor dem Hintergrund der Finanzkrise 2007-2009'. PhD thesis. TU Dresden (65, 229).

Mas-Colell, Andreu, Michael Dennis Whinston and Jerry R. Green (1995). *Microeconomic theory.* Oxford University Press (100).

Masani, Pesi R. (1990). *Norbert Wiener 1894-1964.* Birkhäuser, Basel. DOI: 10.1007/978-3-0348-9252-0 (13, 30, 78, 232).

Mathieu, Martin (1988). 'On the origin of the notion 'Ergodic Theory''. In: *Expositiones Mathematicae* 6, 373–377 (49, 50).

Maxwell, J. Clerk (1860a). 'II. Illustrations of the dynamical theory of gases'. In: *Philosophical Magazine* 20 (130), 21–37. DOI: 10.1080/14786446008642902 (35).

Maxwell, J. Clerk (1860b). 'V. Illustrations of the dynamical theory of gases.—Part I. On the motions and collisions of perfectly elastic spheres'. In: *Philosophical Magazine* 19 (124), 19–32. DOI: 10.1080/14786446008642818 (35).

Maxwell, J. Clerk (1867). 'On the Dynamical Theory of Gases'. In: *Philosophical Transactions of the Royal Society of London* 157, 49–88. DOI: 10.1098/rstl.1867.0004 (37, 278).

Maxwell, J. Clerk (1890). 'On Boltzmann's Theorem on the average distribution of energy in a system of material points'. In: *The Scientific Papers of James Clerk Maxwell.* Ed. by William D. Niven. Vol. II. originally published 1879. Dover Publications, New York, 713–741 (44).

Mayer, Jürgen, Khaled Khairy and Jonathon Howard (2010). 'Drawing an elephant with four complex parameters'. In: *American Journal of Physics* 78 (6), 648–649. DOI: 10.1119/1.3254017 (250).

Mayergoyz, Isaak D. (1991). *Mathematical Models of Hysteresis.* Springer New York. DOI: 10.1007/978-1-4612-3028-1 (198).

Mazo, Robert M. (2002). *Brownian Motion. Fluctuations, Dynamics, and Applications.* Oxford University Press. DOI: 10.1093/acprof:oso/9780199556441.001.0001 (61).

McCauley, Joseph L. (2008). 'Time vs. ensemble averages for nonstationary time series'. In: *Physica A* 387 (22), 5518–5522. DOI: 10.1016/j.physa.2008.05.057 (196).

McCauley, Joseph L. (2009). *Dynamics of Markets. The New Financial Economics.* 2nd ed. Cambridge University Press (194).

McCloskey, Deirdre N. (2002). 'Samuelsonian Economics'. In: *Eastern Economic Journal*, 425–430. JSTOR: 40326119 (222).

McCloskey, Donald N. (1983). 'The Rhetoric of Economics'. In: *Journal of Economic Literature* 21 (2), 481–517. JSTOR: 2724987 (23, 169, 187, 225, 228, 240–243).

McKenzie, Lionel W. (1959). 'On the Existence of General Equilibrium for a Competitive Market'. In: *Econometrica* 27 (1), 54. DOI: 10.2307/1907777 (227).

McMillan, Brockway (1953). 'The Basic Theorems of Information Theory'. In: *Annals of Mathematical Statistics* 24 (2), 196–219. DOI: 10.1214/aoms/1177729028 (78).

McTaggart, J. Ellis (1908). 'The Unreality Of Time'. In: *Mind* 17 (68), 457–474. JSTOR: 2248314 (232).

Meder, David, Finn Rabe, Tobias Morville, Kristoffer Madsen, Hartwig R. Siebner and Oliver J. Hulme (2019). 'Ergodicity-breaking reveals time optimal economic behavior in humans'. In: arXiv: 1906.04652 (283).

Mehra, Rajnish and Edward C. Prescott (1985). 'The equity premium: A puzzle'. In: *Journal of Monetary Economics* 15 (2), 145–161. DOI: 10.1016/0304-3932(85)90061-3 (169).

Mehrling, Perry (2008). 'Book Review: Geoffrey Poitras (ed.) Pioneers of Financial Economics, Vol. I: Contributions Prior to Irving Fisher; Geoffrey Poitras and Franck Jovanovic (eds.) Pioneers of Financial Economics, Vol. II: Twentieth Century Contributions'. In: *Journal of the History of Economic Thought* 30 (3), 422–425. DOI: 10.1017/s1053837208000412 (190, 191).

Meitz, Mika and Pentti Saikkonen (2008). 'Ergodicity, mixing, and existence of moments of a class of Markov models with applications to GARCH and ACD models'. In: *Econometric Theory* 24 (5), 1291–1320. DOI: 10.1017/s0266466608080511 (195, 196).

Menger, Carl (1871). *Grundsätze der Volkswirthschaftslehre.* Braumüller, Wien (278).

Menger, Karl (1934). 'Das Unsicherheitsmoment in der Wertlehre. Betrachtungen im Anschluß an das sogenannte Petersburger Spiel'. In: *Zeitschrift für Nationalökonomie/Journal of Economics* 5 (4). English translations in Menger (1967) and Menger (1979), 459–485. DOI: 10.1007/BF01311578 (100, 112, 133, 137–140, 168, 278).

Menger, Karl (1967). 'The Role of Uncertainty in Economics'. In: *Essays in Mathematical Economics. In Honor of Oskar Morgenstern.* Ed. by Martin Shubik. Trans. by Wolfgang Schoellkopf and W. Giles Mellon. Princeton University Press, 211–231. DOI: 10.1515/9781400877386-018 (322).

Menger, Karl (1979). 'The Role of Uncertainty in Economics'. In: *Selected Papers in Logic and Foundations, Didactics, Economics.* D. Reidel, Dordrecht, 259–278. DOI: 10.1007/978-94-009-9347-1_25 (322).

Menger, Karl (1998). *Ergebnisse eines Mathematischen Kolloquiums.* Ed. by Egbert Dierker and Karl Sigmund. Springer, Wien. DOI: 10.1007/978-3-7091-6470-9 (325).

Merton, Robert C. (1973). 'Theory of Rational Option Pricing'. In: *Bell Journal of Economics* *and Management Science* 4 (1), 141–183. DOI: 10.2307/3003143 (90).

Merton, Robert C. and Paul A. Samuelson (1974). 'Fallacy of the log-normal approximation to optimal portfolio decision-making over many periods'. In: *Journal of Financial Economics* 1 (1), 67–94. DOI: 10.1016/0304-405x(74)90009-9 (170).

Metzler, Ralf (2015). 'Forever ageing'. In: *Nature Physics* 12.2, 113–114. DOI: 10.1038/nphys3585 (207).

Metzler, Ralf, Jae-Hyung Jeon, Andrey G. Cherstvy and Eli Barkai (2014). 'Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking'. In: *Physical Chemistry Chemical Physics* 16 (44), 24128–24164. DOI: 10.1039/c4cp03465a (208).

Metzler, Ralf, Vincent Tejedor, Jae-Hyung Jeon, Yong He, W. H. Deng, Stas Burov and Eli Barkai (2009). 'Analysis of single particle trajectories: From normal to anomalous diffusion'. In: *Acta Physica Polonica B* 40 (5), 1315–1331. URL: http://www.actaphys.uj.edu.pl/fulltext?series=Reg&vol=40&page=1315 (208).

Meyers, Robert A., ed. (2009). *Encyclopedia of Complexity and Systems Science*. Springer, New York. DOI: 10.1007/978-0-387-30440-3 (173).

Meyers, Robert A., ed. (2011). *Complex Systems in Finance and Econometrics*. Springer, New York. DOI: 10.1007/978-1-4419-7701-4 (173).

Mills, Terence C. (2015). *Time Series Econometrics*. Palgrave Macmillan, Houndmills (194).

Mills, Terence C. and Raphael N. Markellos (2008). *The Econometric Modelling of Financial Time Series*. Cambridge University Press. DOI: 10.1017/cbo9780511817380 (194).

Mirowski, Philip (1989). 'How not to do things with metaphors: Paul Samuelson and the science of neoclassical economics'. In: *Studies in History and Philosophy of Science Part A* 20 (2), 175–191. DOI: 10.1016/0039-3681(89)90002-2 (220, 242, 279).

Mirowski, Philip (1991). *More Heat than Light. Economics as Social Physics, Physics as Nature's Economics*. Cambridge University Press (279, 280).

Mirowski, Philip (2002). *Machine Dreams. Economics Becomes a Cyborg Science*. Cambridge University Press (280).

Mirowski, Philip (2007). 'Markets come to bits: Evolution, computation and markomata in economic science'. In: *Journal of Economic Behavior & Organization* 63 (2), 209–242. DOI: 10.1016/j.jebo.2005.03.015 (36).

Mirowski, Philip (16th Feb. 2013). 'Die Linke hat auch keine Antwort'. In: *Frankfurter Allgemeine Zeitung* (279).

Von Mises, Richard (1920a). 'Ausschaltung der Ergodenhypothese in der physikalischen Statistik'. In: *Physikalische Zeitschrift* 21 (9), 225–232. URL: https://babel.hathitrust.org/cgi/pt?id=mdp.39015086723080 (69).

Von Mises, Richard (1920b). 'Ausschaltung der Ergodenhypothese in der physikalischen Statistik II'. In: *Physikalische Zeitschrift* 21 (9), 256–262. URL: https://babel.hathitrust.org/cgi/pt?id=mdp.39015086723080 (69).

de Moivre, Abraham (1718). *The Doctrine of Chances. Or, A Method of Calculating the Probability of Events in Play*. W. Pearson, London. URL: https://archive.org/details/doctrinechances00moivgoog (visited on 13/04/2017) (83, 278).

MOLCHANOV, STANISLAV A. (1997). 'Wiener-Kolmogorov Conception of the Stochastic Organization of Nature'. In: *Proceedings of the Norbert Wiener Centenary Congress, 1994. Michigan State University, November 27-December 3, 1994*. Ed. by VIDYADHAR MANDREKAR and PESI R. MASANI. Vol. 52. American Mathematical Society, Providence, 1–36 (13–15).

MOLENAAR, PETER C. M. (2004). 'A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever'. In: *Measurement: Interdisciplinary Research & Perspective* 2 (4), 201–218. DOI: 10.1207/s15366359mea0204_1 (215).

MOLENAAR, PETER C. M. (2007). 'On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intra-individual variation'. In: *Developmental Psychobiology* 50 (1), 60–69. DOI: 10.1002/dev.20262 (215).

MONOD, JACQUES (1971). *Chance and Necessity*. Alfred A. Knopf, New York (15).

DE MONTMORT, PIERRE R. (1713). *Essay d'analyse sur les jeux de hazard*. Reprinted by the American Mathematical Society 2006. Jacque Quillau, Paris. Google Books: q6AzGMfjJPkC (84).

MOORE, CALVIN C. (2015). 'Ergodic theorem, ergodic theory, and statistical mechanics'. In: *Proceedings of the National Academy of Sciences* 112 (7), 1907–1911. DOI: 10.1073/pnas.1421798112 (71, 219).

MORGAN, MARY S. (1990). *The history of econometric ideas*. Cambridge University Press (187).

MORGENSTERN, OSKAR (1935). 'Vollkommene Voraussicht und wirtschaftliches Gleichgewicht'. In: *Journal of Economics* 6 (3), 337–357. DOI: 10.1007/BF01311642 (238).

MORGENSTERN, OSKAR (1972). 'Descriptive, Predictive, and Normative Theory'. In: *Kyklos* 25 (4), 699–714. DOI: 10.1111/j.1467-6435.1972.tb01077.x (13).

MOSER, JÜRGEN (1962). 'On invariant curves of area-preserving mappings of an annulus'. In: *Nachrichten der Akademie der Wissenschaften Göttingen Mathematisch-Physikalische Klasse II*, 1–20 (75).

MOSER, JÜRGEN (1966a). 'A rapidly convergent iteration method and non-linear differential equations. II'. In: *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze* 20 (3), 499–535. URL: http://eudml.org/doc/83390 (75).

MOSER, JÜRGEN (1966b). 'A rapidly convergent iteration method and non-linear partial differential equations - I'. In: *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze* 20 (2), 265–315. URL: http://eudml.org/doc/83381 (75).

MUMFORD, DAVID (2000). 'The Dawning of the Age of Stochasticity'. In: *Mathematics: Frontiers and Perspectives*. Ed. by VLADIMIR I. ARNOL'D, MICHAEL F. ATIYAH, PETER D. LAX and BARRY MAZUR. American Mathematical Society, Providence, 197–218 (14).

MUTH, JOHN F. (1961). 'Rational Expectations and the Theory of Price Movements'. In: *Econometrica* 29 (3), 315–335. DOI: 10.2307/1909635 (176, 198).

NADKARNI, M. G. (2013). *Basic Ergodic Theory*. 3rd ed. Hindustan Book Agency, New Delhi. DOI: 10.1007/978-93-86279-53-8 (77).

NASIR, MUHAMMAD A. and JAMIE MORGAN (2018). 'The unit root problem: Affinities between ergodicity and stationarity, its practical contradictions for central bank policy, and some consideration of alternatives'. In: *Journal of Post Keynesian Economics* 41 (3), 339–363. DOI: 10.1080/01603477.2017.1387060 (192, 195, 233).

NATURE (2007). 'Editorial – Parallel worlds galore'. In: *Nature* 448 (7149), 1. DOI: 10.1038/448001a (212).

NAVARRO, LUIS (1998). 'Gibbs, Einstein and the Foundations of Statistical Mechanics'. In: *Archive for History of Exact Sciences* 53 (2), 147–180. DOI: 10.1007/s004070050025 (71).

NEUGEBAUER, TIBOR (2010). 'Moral Impossibility in the Petersburg Paradox: A Literature Survey and Experimental Evidence'. In: *Luxembourg School of Finance Working Paper Series*. URL: https://EconPapers.repec.org/RePEc:crf:wpaper:10-14 (91).

VON NEUMANN, JOHN (1932a). *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin. DOI: 10.1007/978-3-642-61409-5 (46).

VON NEUMANN, JOHN (1932b). 'Proof of the Quasi-Ergodic Hypothesis'. In: *Proceedings of the National Academy of Sciences* 18 (1), 70–82. DOI: 10.1073/pnas.18.1.70 (63, 64, 221, 278).

VON NEUMANN, JOHN (1929). 'Beweis des Ergodensatzes und des *H*-Theorems in der neuen Mechanik'. In: *Zeitschrift für Physik* 57 (1-2), 30–70. DOI: 10.1007/bf01339852 (63, 211).

VON NEUMANN, JOHN (1932c). 'Physical Applications of the Ergodic Hypothesis'. In: *Proceedings of the National Academy of Sciences* 18 (3), 263–266. DOI: 10.1073/pnas.18.3.263 (64, 219).

VON NEUMANN, JOHN (1937). 'Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes'. In: *Ergebnisse eines mathematischen Kolloquiums*. Ed. by KARL MENGER. English translation in VON NEUMANN (1945). The German original is reprinted in MENGER (1998, pp. 453–463). Wien, 73–83. DOI: 10.1007/978-3-7091-6470-9 (172).

VON NEUMANN, JOHN (1945). 'A Model of General Economic Equilibrium'. In: *Review of Economic Studies* 13 (1), 1–9. DOI: 10.2307/2296111 (325).

VON NEUMANN, JOHN and OSKAR MORGENSTERN (1967). *Spieltheorie und wirtschaftliches Verhalten*. Ed. by FRIEDRICH SOMMER. 2nd ed. Physica-Verlag, Würzburg (18).

VON NEUMANN, JOHN and OSKAR MORGENSTERN (1955). *Theory of Games and Economic Behavior*. 3rd ed. 1st edition 1944. Princeton University Press (18, 100, 103, 226, 228, 246, 255, 258).

NEUSSER, KLAUS (2016). *Time Series Econometrics*. Springer, Cham. DOI: 10.1007/978-3-319-32862-1 (194).

NICKEL, GREGOR (2000). 'Epilogue on Determinism: Scenes from the Interplay Between Metaphysics and Mathematics'. In: *One-Parameter Semigroups for Linear Evolution Equations*. Ed. by KLAUS-JOCHEN ENGEL and RAINER NAGEL. Springer, New York, 531–554. DOI: 10.1007/b97696 (11).

NICOL, MATTHEW and KARL PETERSEN (2009). 'Ergodic Theory: Basic Examples and Constructions'. In: *Encyclopedia of Complexity and Systems Science*. Ed. by ROBERT A. MEYERS. Springer, New York, 2956–2980. DOI: 10.1007/978-0-387-30440-3_177 (173).

NIEHANS, JÜRG (1948). 'Zur Preisbildungen bei ungewissen Erwartungen'. In: *Swiss Journal of Economics and Statistics* 84.V, 433–456. URL: http://www.sjes.ch/papers/1948-V-3.pdf (258).

NIEHANS, JÜRG (1997). 'Samuelsons »Foundations«: Ein nichtmathematischer Leitfaden'. In: *Paul A. Samuelsons »Foundations of Economic Analysis«*. Ed. by JÜRG NIEHANS, PAUL SAMUELSON and C. CHRISTIAN VON WEIZSÄCKER. Vademecum zu einem Klassiker der

Gegenwart. Verlag Wirtschaft und Finanzen, Düsseldorf, 55–87 (231).

Nishimura, Kazuo, Gerhard Sorger and Makoto Yano (1994). 'Ergodic chaos in optimal growth models with low discount rates'. In: *Economic Theory* 4 (5), 705–717. doi: 10.1007/bf01212026 (74).

Nolan, John P. (30th Jan. 2018). *Stable Distributions. Models for Heavy Tailed Data*. In progress, Chapter 1 online. url: http://fs2.american.edu/jpnolan/www/stable/chap1.pdf (91).

North, Douglass C. (2005). *Understanding the Process of Economic Change*. Princeton University Press (192).

Nowak, Martin A. (2006). 'Five Rules for the Evolution of Cooperation'. In: *Science* 314 (5805), 1560–1563. doi: 10.1126/science.1133755 (145).

Nualart, David (2007). 'Fractional Brownian Motion stochastic calculus and applications'. In: *Proceedings of the International Congress of Mathematicians. Madrid, August 22–30, 2006*. Ed. by Marta Sanz-Solé, Javier Soria, Juan L. Varona and Joan Verdera. Vol. 3. European Mathematical Society, Zurich, 1541–1562. url: http://www.icm2006.org/proceedings/Vol_III/contents/ICM_Vol_3_74.pdf (201).

NZZ (2002). 'Daniel Bernoulli und seine schwierige Familie'. In: *Neue Zürcher Zeitung*. 17.03.2002. url: https://www.nzz.ch/article80ZZW-1.378426 (85).

O'Donnell, Rod (2014). 'A critique of the ergodic/nonergodic approach to uncertainty'. In: *Journal of Post Keynesian Economics* 37 (2), 187–209. doi: 10.2753/PKE0160-3477370201 (192).

O'Donnell, Rod (2016a). 'Second contribution to the ENE critique: Reply to Davidson, part 1'. In: *Journal of Post Keynesian Economics* 39 (1), 17–43. doi: 10.1080/01603477.2016.1148617 (192).

O'Donnell, Rod (2016b). 'Third contribution to the ergodic/nonergodic critique: Reply to Davidson, part 2'. In: *Journal of Post Keynesian Economics* 39 (2), 145–171. doi: 10.1080/01603477.2016.1169431 (192).

Olofsson, Helen, Jörgen Ripa and Niclas Jonzén (2009). 'Bet-hedging as an evolutionary game: the trade-off between egg size and number'. In: *Proceedings of the Royal Society B: Biological Sciences* 276 (1669), 2963–2969. doi: 10.1098/rspb.2009.0500 (214).

Oppenheim, Irwin, Kurt E. Shuler and George H. Weiss (1977). *The Master Equation*. MIT Press, Cambridge (38).

Ornes, Stephen (2017). 'Core Concept: How nonequilibrium thermodynamics speaks to the mystery of life'. In: *Proceedings of the National Academy of Sciences* 114 (3), 423–424. doi: 10.1073/pnas.1620001114 (208).

Orr, H. Allen (2009). 'Fitness and its role in evolutionary genetics'. In: *Nature Reviews Genetics* 10 (8), 531–539. doi: 10.1038/nrg2603 (214).

Orr, H. Allen (2018). 'Evolution, finance, and the population genetics of relative wealth'. In: *Journal of Bioeconomics* 20 (1), 29–48. doi: 10.1007/s10818-017-9254-y (214).

Oseledets, Valery I. (1968). 'Multiplicative ergodic theorem: Characteristic Lyapunov exponents of dynamical systems'. (in Russian). In: *Trudy MMO* 19, 179–210 (78).

Oseledets, Valery I. (2008). 'Oseledets theorem'. In: *Scholarpedia* 3 (1). revision #142085, 1846. doi: 10.4249/scholarpedia.1846 (78).

Ostwald, Wilhelm (1909). *Energetische Grundlagen der Kulturwissenschaft*. Dr. Werner Klinkhardt Verlag, Leipzig (220, 221).

Oxtoby, John C. and Stanislav M. Ulam (1941). 'Measure-Preserving Homeomorphisms and Metrical Transitivity'. In: *Annals of Mathematics* 42 (4), 874–920. DOI: 10.2307/1968772 (70).

Pacioli, Luca (1494). *Summa de arithmetica, geometria, proportioni et proportionalità*. published 1523. Paganino de Paganini. Google Books: iqgPe49fhrsC (81–83, 278).

Paine, Alasdair D. M. (1985). ''Ergodic' reasoning in geomorphology: time for a review of the term?' In: *Progress in Physical Geography* 9 (1), 1–15. DOI: 10.1177/030913338500900101 (215).

Palmer, Richard G. (1989). 'Broken Ergodicity'. In: *Lectures in the Sciences of Complexity. The Proceedings of the 1988 Complex Systems Summer School held June-July, 1988 in Santa Fe, New Mexico*. Ed. by Daniel L. Stein. Addison-Wesley, Redwood City CA, 275–300 (41, 65, 234).

Pareto, Vilfredo F. (1964). *Cours d'Economie Politique*. Ed. by G.-H. Bousquet and G. Busino. Oeuvres complètes Tome I. first ed. 1896–1897. Librairie Droz, Genève (91).

Patrascioiu, Adrian (1987). 'The Ergodic Hypothesis. A Complicated Problem in Mathematics and Physics'. In: *Los Alamos Science* (15), 263–279. URL: https://la-science.lanl.gov/lascience15.shtml (30, 73).

Pearson, Karl (1905a). 'The Problem of the Random Walk'. In: *Nature* 72 (1865), 294. DOI: 10.1038/072294b0 (208).

Pearson, Karl (1905b). 'The Problem of the Random Walk'. In: *Nature* 72 (1867), 342. DOI: 10.1038/072342a0 (208).

Pei-Dong, Liu; and Min Qian (1995). *Smooth Ergodic Theory of Random Dynamical Systems*. Springer Berlin Heidelberg. DOI: 10.1007/bfb0094308 (77).

Peters, Ole (2009). 'On Time and Risk'. In: *Santa Fe Institute Bulletin* 24 (1), 36–41. URL: http://www.santafe.edu/media/bulletin_article_pdf/TimeRisk.pdf (iii).

Peters, Ole (2011a). 'Menger 1934 revisited'. In: arXiv: 1110.1578v1 [q-fin.RM] (22, 140, 168).

Peters, Ole (2011b). 'Optimal leverage from non-ergodicity'. In: *Quantitative Finance* 11 (11), 1593–1602. DOI: 10.1080/14697688.2010.513338 (168, 260, 277, 278).

Peters, Ole (2011c). 'The time resolution of the St Petersburg paradox'. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 369 (1956), 4913–4931. DOI: 10.1098/rsta.2011.0065 (3, 5, 20, 21, 26, 88, 97, 104, 112, 113, 145, 149, 168, 169, 171, 193, 249, 277, 278).

Peters, Ole (2014). 'A conversation on utility theory'. In: *mimeo* (22, 114, 123).

Peters, Ole (18th July 2017). 'Doing a Laplace'. In: *Ergodicity Economics*. URL: https://ergodicityeconomics.com/2017/07/18/doing-a-laplace/ (114).

Peters, Ole (2018a). 'Comment on D. Bernoulli (1738)'. In: URL: https://www.researchers.one/article/2019-05-1 (22, 114, 121, 123).

Peters, Ole (29th May 2018b). 'The Copenhagen experiment'. In: *Ergodicity Economics*. URL: https://ergodicityeconomics.com/2018/05/29/the-copenhagen-experiment/ (283).

Peters, Ole (16th Feb. 2018c). 'The trouble with Bernoulli 1738'. In: *Ergodicity Economics*. URL: https://ergodicityeconomics.com/2018/02/16/the-trouble-with-bernoulli-1738/ (114, 122, 123, 168).

Peters, Ole and Alexander Adamou (2013). 'Stochastic Market Efficiency'. In: URL: https://www.santafe.edu/research/

results / working - papers / stochastic - market-efficiency (169).

PETERS, OLE and ALEXANDER ADAMOU (2017a). 'Insurance makes wealth grow faster'. In: arXiv: 1507.04655 (112, 169).

PETERS, OLE and ALEXANDER ADAMOU (2017b). 'Leverage Efficiency'. In: arXiv: 1101.4548v2 (169).

PETERS, OLE and ALEXANDER ADAMOU (2018a). 'An evolutionary advantage of cooperation'. In: arXiv: 1506.03414v2 (145, 169).

PETERS, OLE and ALEXANDER ADAMOU (2018b). *Ergodicity Economics. Lecture Notes.* 5.0, 2018/06/30. URL: https://ergodicityeconomics.com/lecture-notes/ (5, 67, 68, 112, 113, 119, 145, 160, 169, 180, 193, 256, 278, 282).

PETERS, OLE and ALEXANDER ADAMOU (2018c). 'The time interpretation of expected utility theory'. In: arXiv: 1801.03680 (22, 159, 160, 168, 175, 278, 280).

PETERS, OLE and ALEXANDER ADAMOU (2019). 'An evolutionary advantage of cooperation'. In: URL: https://www.researchers.one/article/2019-03-4 (214).

PETERS, OLE and MURRAY GELL-MANN (2016). 'Evaluating gambles using dynamics'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 26 (2), 023103. DOI: 10.1063/1.4940236 (114, 139, 140, 159, 160, 168, 278).

PETERS, OLE and WILLIAM KLEIN (2013). 'Ergodicity Breaking in Geometric Brownian Motion'. In: *Physical Review Letters* 110 (10), 100603. DOI: 10.1103/PhysRevLett.110.100603 (234).

PETERS, OLE and MICHAEL J. MAUBOUSSIN (2012). 'Shaking the Foundation – Revisiting Basic Assumptions about Risk, Reward, and Optimal Portfolios – An Interview with Ole Peters'. In: URL: http://trendfollowing.com/whitepaper/Mauboussin2.pdf (26, 231).

PETERS, OLE and MAXIMILIAN WERNER (2017). 'A recipe for irreproducible results'. In: arXiv: 1706.07773 (20).

PETERSEN, KARL E. (1989). *Ergodic theory.* Cambridge University Press. DOI: 10.1017/cbo9780511608728 (55, 65, 77, 193).

PETERSEN, KARL (1996). 'Ergodic Theorems and the Basis of Science'. In: *Synthese* 108 (2), 171–183. DOI: 10.1007/bf00413496 (20, 64, 235).

PHELPS, EDMUND S. (1961). 'The Golden Rule of Accumulation: A Fable for Growthmen'. In: *American Economic Review* 51 (4), 638–643. JSTOR: 1812790 (172).

PHELPS, EDMUND S. (1965). 'Second Essay on the Golden Rule of Accumulation'. In: *American Economic Review Papers and Proceedings* 55 (4), 793–814. JSTOR: 1823937 (172).

PHELPS, EDMUND S. (1966). 'Models of Technical Progress and the Golden Rule of Research'. In: *Review of Economic Studies* 33 (2), 133–145. DOI: 10.2307/2974437 (172).

PHELPS, EDMUND S. (1967). *Golden Rules of Economic Growth. Studies of Efficient and Optimal Investment.* W. W. Norton, New York (172).

PHELPS, EDMUND S., ed. (1970). *Microeconomic Foundations of Employment and Inflation Theory.* Norton, New York (198).

PHILIPPI, TOM and JON SEGER (1989). 'Hedging one's evolutionary bets, revisited'. In: *Trends in Ecology & Evolution* 4 (2), 41–44. DOI: 10.1016/0169-5347(89)90138-9 (214).

PISCITELLI, LAURA, ROD CROSS, MICHAEL GRINFELD and HARBIR LAMBA (2000). 'A Test for Strong Hysteresis'. In: *Computational Economics* 15 (1/2), 59–78. DOI: 10.1023/a:1008638827537 (198).

PLANCHEREL, MICHEL (1913). 'Beweis der Unmöglichkeit ergodischer mechanischer Systeme'.

In: *Annalen der Physik* 347 (15), 1061–1063. DOI: 10.1002/andp.19133471509 (70).

Von Plato, Jan (1987). 'Probabilistic Physics the Classical Way'. In: *The Probabilistic Revolution*. Ed. by Lorenz Krüger, Gerd Gigerenzer and Mary S. Morgan. Vol. 2: Ideas in the Sciences. MIT Press. Chap. 14, 379–407 (26, 211, 212).

Von Plato, Jan (1994). *Creating Modern Probability*. Cambridge University Press (70, 71, 78).

Von Plato, Jan (1982). 'Probability and Determinism'. In: *Philosophy of Science* 49 (1), 51–66. DOI: 10.1086/289034 (71).

Von Plato, Jan (1991). 'Boltzmann's Ergodic Hypothesis'. In: *Archive for History of Exact Sciences* 42 (1), 71–89. DOI: 10.1007/BF00384333 (71).

Poincaré, Henri (1890). 'Sur le problème des trois corps et les équations de la dynamique'. In: *Acta Mathematica* 13 (1), 3–270. DOI: 10.1007/BF02392506 (39).

Poincaré, Henri (1899). *Les Méthodes Nouvelles de la Mécanique Céleste*. Vol. III. Gauthier-Villars, Paris. URL: https://archive.org/details/lesmthodesnouv003poin (39, 48).

Poincaré, Henri (1914). *Wissenschaft und Methode*. B. G. Teubner, Leipzig (12).

Poitras, Geoffrey (2018). 'The pre-history of econophysics and the history of economics: Boltzmann versus the marginalists'. In: *Physica A* 507, 89–98. DOI: 10.1016/j.physa.2018.05.058 (193, 214).

Poitras, Geoffrey and John Heaney (2015). 'Classical Ergodicity and Modern Portfolio Theory'. In: *Chinese Journal of Mathematics* 2015, 1–17. DOI: 10.1155/2015/737905 (214).

Popper, Karl R. (1935). *Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft*. Springer, Wien (225).

Poundstone, William (2005). *Fortune's Formula. The untold story of the scientific betting system that beat the casinos and Wall Street*. Hill and Wang, New York (210).

Pratt, John W. (1964). 'Risk Aversion in the Small and in the Large'. In: *Econometrica* 32 (1/2), 122–136. DOI: 10.2307/1913738 (133).

Prelec, Dražen (1998). 'The Probability Weighting Function'. In: *Econometrica* 66 (3), 497–527. DOI: 10.2307/2998573 (136).

Prelec, Dražen (2000). 'Compound invariant weighting functions in prospect theory'. In: *Choices, Values, and Frames*. Ed. by Daniel Kahneman and Amos Tversky. Cambridge University Press New York, 67–92 (136).

Prigogine, Ilya (1980). *From being to becoming: time and complexity in the physical sciences*. Freeman & Co., San Francisco (186, 187, 215).

Prigogine, Ilya, Alkis P. Grecos and Claude George (1976). 'Kinetic Theory and Ergodic Properties'. In: *Proceedings of the National Academy of Sciences* 73 (6), 1802–1805. URL: http://www.pnas.org/content/73/6/1802.short (219).

Prigogine, Ilya and Isabelle Stengers (1984). *Order Out of Chaos. Man's New Dialogue With Nature*. Bantam Books, New York (16).

Quay S. J., Paul M. (1978). 'A Philosophical Explanation of the Explanatory Functions of Ergodic Theory'. In: *Philosophy of Science* 45 (1), 47–59. DOI: 10.1086/288778 (71).

Quiggin, John (1982). 'A theory of anticipated utility'. In: *Journal of Economic Behavior & Organization* 3 (4), 323–343. DOI: 10.1016/0167-2681(82)90008-7 (256, 258).

QUINN, FRANK (2012). 'A Revolution in Mathematics? What Really Happened a Century Ago and Why It Matters Today'. In: *Notices of the American Mathematical Society* 59 (1), 31–37. DOI: 10.1090/noti787 (229).

RABIN, MATTHEW (1993). 'Incorporating fairness into game theory and economics'. In: *American Economic Review*, 1281–1302. JSTOR: 2117561 (250).

RACHEV, SVETLOZAR T., ed. (2003). *Handbook of Heavy Tailed Distributions in Finance*. North-Holland, Amsterdam. DOI: 10.1016/b978-0-444-50896-6.x5000-6 (175).

RAMSEY, FRANK P. (1928). 'A Mathematical Theory of Saving'. In: *Economic Journal* 38 (152), 543–559. DOI: 10.2307/2224098 (172).

RAMSEY, FRANK P. (1950). *The Foundations of Mathematics. and other Logical Essays*. Ed. by RICHARD B. BRAITHWAITE. 2nd ed. first publ. 1931. Routledge, London (256, 258).

RÉNYI, ALFRÉD (1961). 'On Measures of Entropy and Information'. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Held at the Statistical Laboratory University of California, June 20-July 30, 1960). Ed. by JERZY NEYMAN. Vol 1: Contributions to the Theory of Statistics. University of California Press, 547–561. URL: https://projecteuclid.org/euclid.bsmsp/1200512181 (203).

RÜSCHENDORF, LUDGER (1998). 'Wasserstein metric'. URL: https://www.stochastik.uni-freiburg.de/emeriti/rueschendorf/inhalte/1998-artikel-wasserstein-metric (39).

RAUSSEN, MARTIN and CHRISTIAN SKAU (2015). 'Interview with Yakov Sinai'. In: *Notices of the American Mathematical Society* 62 (2), 152–160. DOI: 10.1090/noti1210 (283).

RAYLEIGH, JOHN WILLIAM STRUTT (1905). 'The Problem of the Random Walk'. In: *Nature* 72 (1866), 318. DOI: 10.1038/072318a0 (208).

REDNER, SIDNEY (1990). 'Random multiplicative processes: An elementary tutorial'. In: *American Journal of Physics* 58 (3), 267–273. DOI: 10.1119/1.16497 (145).

REES, MARK, C. JESSICA E. METCALF and DYLAN Z. CHILDS (2010). 'Comment: Bet-hedging as an evolutionary game: the trade-off between egg size and number'. In: *Proceedings of the Royal Society B: Biological Sciences* 277 (1685), 1149–1151. DOI: 10.1098/rspb.2009.1541 (214).

REINHART, CARMEN M. and KENNETH S. ROGOFF (2010). *Dieses Mal ist alles anders*. FinanzBuch Verlag, München (223).

RENN, JÜRGEN (2008). 'Boltzmann and the end of the mechanistic worldview'. In: *Boltzmann's Legacy*. Ed. by GIOVANNI GALLAVOTTI, WOLFGANG REITER and JAKOB YNGVASON. European Mathematical Society, Zurich, 7–25. DOI: 10.4171/057 (70).

RESCHER, NICHOLAS (1962). 'The Revolt Against Process'. In: *Journal of Philosophy* 59 (15), 410–417. DOI: 10.2307/2022958 (216).

RESCHER, NICHOLAS (1996). *Process Metaphysics*. State University of New York Press, Albany (216).

RESCHER, NICHOLAS (2000). *Process Philosophy. A Survey of Basic Issues*. University of Pittsburgh Press (216).

RIEDL, RUPERT (2004). *Meine Sicht der Welt*. 2nd ed. Seifert Verlag, Wien (223).

RIGATELLI, TOTI (1985). 'Il "Problema della parti" in manoscritti del XIV e XV seclo'. In: *Mathemata. Festschrift für Helmuth Gericke*. Ed. by MENSO FOLKERTS and UTA LINDGREN. Franz Steiner Verlag, Stuttgart, 229–236 (82).

Ripa, Jörgen, Helen Olofsson and Niclas Jonzén (2010). 'What is bet-hedging, really?' In: *Proceedings of the Royal Society B: Biological Sciences* 277 (1685), 1153–1154. doi: 10.1098/rspb.2009.2023 (214).

Risken, Hannes (1996). *The Fokker-Planck Equation*. 2nd ed. Springer, Berlin Heidelberg. doi: 10.1007/978-3-642-61544-3 (38).

Røed, Knut (1997). 'Hysteresis in Unemployment'. In: *Journal of Economic Surveys* 11 (4), 389–418. doi: 10.1111/1467-6419.00040 (198).

Rosenblueth, Arturo and Norbert Wiener (1945). 'The Role of Models in Science'. In: *Philosophy of Science* 12 (4), 316–321. doi: 10.1086/286874 (239).

Rosenthal, Artur (1913). 'Beweis der Unmöglichkeit ergodischer Gassysteme'. In: *Annalen der Physik* 347 (14), 796–806. doi: 10.1002/andp.19133471407 (70).

Rosser Jr., J. Barkley (2015). 'Reconsidering ergodicity and fundamental uncertainty'. In: *Journal of Post Keynesian Economics* 38 (3), 331–354. doi: 10.1080/01603477.2015.1070271 (192).

Rothstein, Jerome (1974). 'Loschmidt's and Zermelo's paradoxes do not exist'. In: *Foundations of Physics* 4 (1), 83–89. doi: 10.1007/bf00708557 (193).

Rovelli, Carlo (2011). 'Forget time'. In: *Foundations of Physics* 41 (9), 1475–1490. doi: 10.1007/s10701-011-9561-4 (216).

Rovelli, Carlo (2018a). 'Physics Needs Philosophy. Philosophy Needs Physics'. In: *Foundations of Physics* 48 (5), 481–491. doi: 10.1007/s10701-018-0167-y (216).

Rovelli, Carlo (2018b). *The Order of Time*. Allen Lane (216).

Rubinstein, Ariel (2013). '10 Q&A: Experienced Advice for "Lost" Graduate Students in Economics'. In: *Journal of Economic Education* 44 (3), 193–196. doi: 10.1080/00220485.2013.795448 (2).

Ruelle, David (1973). 'Ergodic Theory'. In: *The Boltzmann Equation. Theory and Applications*. Ed. by Ezechiel G. D. Cohen and Walter Thirring. Springer, Vienna, 609–617. doi: 10.1007/978-3-7091-8336-6_24 (71).

Samuelson, Paul A. (1947). *Foundations of Economic Analysis*. 1st ed. Harvard University Press (278).

Samuelson, Paul A. (1948). *Economics*. McGraw Hill, New York (221, 279).

Samuelson, Paul A. (1952). 'Economic Theory and Mathematics – An Appraisal'. In: *American Economic Review Papers and Proceedings* 42 (2), 56–66. JSTOR: 1910585 (279).

Samuelson, Paul A. (1960). 'The St. Petersburg paradox as a divergent double limit'. In: *International Economic Review* 1 (1), 31–37. doi: 10.2307/2525406 (100).

Samuelson, Paul A. (1962). 'Economists and the History of Ideas'. In: *American Economic Review* 52 (1), 1–18. JSTOR: 1823476 (219).

Samuelson, Paul A. (1965). 'Some Notions on Causality and Teleology in Economics'. In: *Cause and Effect. The Hayden Colloquium on Scientific Method and Concept*. Ed. by David Lerner. Collier-Macmillan, London, 99–143 (10, 254).

Samuelson, Paul A. (1968). 'What Classical and Neoclassical Monetary Theory Really Was'. In: *Canadian Journal of Economics* 1 (1), 1–15. doi: 10.2307/133458 (2, 218, 242, 279).

Samuelson, Paul A. (1969). 'Lifetime Portfolio Selection By Dynamic Stochastic Programming'. In: *Review of Economics and Statistics* 51 (3), 239–246. doi: 10.2307/1926559 (170).

Samuelson, Paul A. (1971a). 'Stochastic speculative price'. In: *Proceedings of the National Academy of Sciences* 68 (2), 335–337. URL: http://www.pnas.org/content/68/2/335 (219).

Samuelson, Paul A. (1971b). 'The "Fallacy" of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling'. In: *Proceedings of the National Academy of Sciences* 68 (10), 2493–2496. DOI: 10.1073/pnas.68.10.2493 (170, 219).

Samuelson, Paul A. (1972). 'Maximum Principles in Analytical Economics'. In: *American Economic Review* 62 (3), 249–262. JSTOR: 1803374 (254).

Samuelson, Paul A. (1973). 'Proof That Properly Discounted Present Values of Assets Vibrate Randomly'. In: *Bell Journal of Economics and Management Science* 4 (2), 369–374. DOI: 10.2307/3003046 (219).

Samuelson, Paul A. (1976a). 'Is Real-World Price a Tale Told by the Idiot of Chance?' In: *Review of Economics and Statistics* 58 (1), 120–123. DOI: 10.2307/1936018 (219).

Samuelson, Paul A. (1976b). 'Optimality of Sluggish Predictors under Ergodic Probabilities'. In: *International Economic Review* 17 (1), 1–7. JSTOR: 2526061 (181, 219).

Samuelson, Paul A. (1977). 'St. Petersburg Paradoxes: Defanged, Dissected, and Historically Described'. In: *Journal of Economic Literature* 15 (1), 24–55. JSTOR: 2722712 (100, 138, 139).

Samuelson, Paul A. (1979a). *Foundations of Economic Analysis*. 9th ed. first ed. 1947. Atheneum, New York (169, 217, 279).

Samuelson, Paul A. (1979b). 'Why we should not make mean log of wealth big though years to act are long'. In: *Journal of Banking & Finance* 3 (4), 305–307. DOI: 10.1016/0378-4266(79)90023-2 (170).

Samuelson, Paul A. (1983). 'My Life Philosophy'. In: *American Economist* 27 (2), 5–12. JSTOR: 25603574 (220).

Samuelson, Paul A. (1990). 'Gibbs in Economics'. In: *Proceedings of the Gibbs Symposium. Yale University May 15-17, 1989*. Ed. by Daniel G. Caldi and George D. Mostow. American Mathematical Society, Providence, 255–267 (220, 221, 242, 243).

Samuelson, Paul A. (1997). 'Some Memories of Norbert Wiener'. In: *The Legacy of Norbert Wiener: A Centennial Symposium. In Honor of the 100th Anniversary of Norbert Wiener's Birth, October 8-14, 1994, MIT*. Ed. by David Jerison, Isadore M. Singer and Daniel Stroock. Vol. 60. American Mathematical Society, Providence, 37–42 (222).

Samuelson, Paul A. (1998). 'How *Foundations* came to be'. In: *Journal of Economic Literature* 36 (3), 1375–1386. JSTOR: 2564803 (219, 220).

Samuelson, Paul A. and William A. Barnett, eds. (2007). *Inside the Economists's Mind. Conversations with Eminent Economists*. Blackwell Publishing, Malden.

Sargent, Thomas J. (1995). *Bounded Rationality in Macroeconomics*. Oxford University Press (198, 250).

Sauermann, Heinz and Reinhard Selten (1962). 'Anspruchsanpassungstheorie der Unternehmung'. In: *Zeitschrift für die gesamte Staatswissenschaft*, 577–597. JSTOR: 40748622. URL: http://www.digizeitschriften.de/dms/resolveppn/?PPN=GDZPPN001778587 (251).

Savage, Leonard J. (1951). 'The Theory of Statistical Decision'. In: *Journal of the Amer-*

ican Statistical Association 46 (253), 55–67. DOI: 10.1080/01621459.1951.10500768 (258).

Savage, Leonard J. (1972). *The Foundations of Statistics*. 2nd ed. first publ. 1954. Dover Publications, New York (255, 256, 258).

Schachermayer, Walter and Josef Teichmann (2008). 'How close are the option pricing formulas of Bachelier and Black–Merton–Scholes?' In: *Mathematical Finance* 18 (1), 155–170. DOI: 10.1111/j.1467-9965.2007.00326.x (90).

Schachermeyer, Walter (2017). *Asymptotic Theory of Transaction Costs*. European Mathematical Society Publishing House, Zurich. DOI: 10.4171/173 (178).

Schäfer, Benjamin, Christian Beck, Kazuyuki Aihara, Dirk Witthaut and Marc Timme (2018). 'Non-Gaussian power grid frequency fluctuations characterized by Lévy-stable laws and superstatistics'. In: *Nature Energy* 3 (2), 119–126. DOI: 10.1038/s41560-017-0058-z (205).

Schmeidler, David (1989). 'Subjective Probability and Expected Utility without Additivity'. In: *Econometrica* 57 (3), 571–587. DOI: 10.2307/1911053 (258).

Schmeling, Jörg (2009). 'Ergodic Theory: Fractal Geometry'. In: *Encyclopedia of Complexity and Systems Science*. Ed. by Robert A. Meyers. Springer, New York, 3015–3029. DOI: 10.1007/978-0-387-30440-3_179 (76).

Schneider, Ivo (1985). 'Luca Pacioli und das Teilungsproblem: Hintergrund und Lösungsversuche'. In: *Mathemata. Festschrift für Helmuth Gericke*. Ed. by Menso Folkerts and Uta Lindgren. Franz Steiner Verlag, Stuttgart, 237–246 (82, 83).

Schoemaker, Paul J. H. (1982). 'The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations'. In: *Journal of Economic Literature* 20 (2), 529–563. JSTOR: 2724488 (254).

Schrödinger, Erwin (1944). *What is Life?* reprinted 1967. Cambridge University Press (208).

Schrödinger, Erwin (1948). *Statistical Thermodynamics. A Course of Seminar Lectures delivered in January-March 1944, at the School of Theoretical Physics, Dublin Institute for Advanced Studies*. Cambridge University Press, UK (211).

Schuster, Heinz G. and Wolfram Just (2005). *Deterministic Chaos. An Introduction*. Wiley, Advanced Materials (Weinheim, Germany) (77).

Schwartz, Jacob T. (2008). 'The Pernicious Influence of Mathematics On Science'. In: *Discrete Thoughts. Essays on Mathematics, Science and Philosophy*. Ed. by Marc Kac, Gian-Carlo Rota and Jacob T. Schwartz. Revised and Corrected 1992 Edition. Birkhäuser, Boston, 19–25. DOI: 10.1007/978-0-8176-4775-9_3 (71).

Seger, Jon and H. Jane Brockmann (1987). 'What is bet-hedging?' In: *Oxford Surveys in Evolutionary Biology* 4, 182–211 (214).

Seidl, Christian (2013). 'The St. Petersburg Paradox at 300'. In: *Journal of Risk and Uncertainty* 46 (3), 247–264. DOI: 10.1007/s11166-013-9165-9 (130, 139).

Selten, Reinhard (1991). 'Evolution, Learning, and Economic Behavior'. In: *Games and Economic Behavior* 3 (1), 3–24. DOI: 10.1016/0899-8256(91)90003-w (213, 251, 252).

Selten, Reinhard (1998). 'Aspiration adaptation theory'. In: *Journal of Mathematical Psychology* 42 (2-3), 191–214. DOI: 10.1006/jmps.1997.1205 (251).

Selten, Reinhard (2001). 'Die konzeptionellen Grundlagen der Spieltheorie einst und jetzt'. In: *Bonn Econ Discussion Papers* 2, 1–16. URL: http://hdl.handle.net/10419/78400 (18).

Selten, Reinhard (2002). 'What Is Bounded Rationality?' In: *Bounded Rationality*. Ed. by Gerd Gigerenzer and Reinhard Selten. Report of the 84th Dahlem Workshop. MIT Press, 13–36 (248, 251, 252).

Sen, Amartya K. (1977). 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory'. In: *Philosophy & Public Affairs* 6 (4), 317–344. JSTOR: 2264946 (248).

Setterfield, Mark (1997). *Rapid Growth and Relative Decline. Modelling Macroeconomic Dynamics with Hysteresis*. Palgrave Macmillan. doi: 10.1057/9780230375871 (198).

Shafer, Glenn (2006). 'Why did Cournot's principle disappear?' In: url: http://www.glennshafer.com/assets/downloads/disappear.pdf (108).

Shafer, Glenn (2010). 'A betting interpretation for probabilities and Dempster-Shafer degrees of belief'. In: arXiv: 1001.1653v1 [math.ST] (105, 107).

Shafer, Glenn and Vladimir Vovk (2006). 'The Sources of Kolmogorov's *Grundbegriffe*'. In: *Statistical Science* 21 (1), 70–98. doi: 10.1214/088342305000000467 (105, 107).

Shalizi, Cosma R. (2017). *Advanced Data Analysis from an Elementary Point of View*. Manuscript version from 30th Jan, 2017. url: http://www.stat.cmu.edu/~cshalizi/ADAfaEPoV/ (visited on 09/06/2018) (51).

Shalizi, Cosma R. and Aryeh Kontorovich (2010). *Almost None of the Theory of Stochastic Processes. A Course on Random Processes, for Students of Measure-Theoretic Probability, with a View to Applications in Dynamics and Statistics*. url: https://www.stat.cmu.edu/~cshalizi/almost-none/ (77).

Shannon, Claude E. (1948). 'A Mathematical Theory of Communication'. In: *Bell System Technical Journal* 27 (3), 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x (40, 41, 77, 78, 280).

Shapley, Lloyd S. (1977a). 'Lotteries and menus: A comment on unbounded utilities'. In: *Journal of Economic Theory* 14 (2), 446–453. doi: 10.1016/0022-0531(77)90144-2 (111, 139).

Shapley, Lloyd S. (1977b). 'The St. Petersburg Paradox: A Con Game?' In: *Journal of Economic Theory* 14 (2), 439–442. doi: 10.1016/0022-0531(77)90142-9 (100, 111, 112, 139).

Shields, Paul C. (1996). *The Ergodic Theory of Discrete Sample Paths*. American Mathematical Society, Providence (78).

Shields, Paul C. (1998). 'The interactions between ergodic theory and information theory'. In: *IEEE Transactions on Information Theory* 44 (6), 2079–2093. doi: 10.1109/18.720532 (78, 213).

Shiryaev, Albert N. (2000). 'Andreĭ Nikolaevich Kolmogorov. (April 25, 1903 to October 20, 1987) A biographical Sketch of His Life and Creative Paths'. In: *Kolmogorov. In Perspective*. Vol. 20. History of Mathematics. American Mathematical Society and London Mathematical Society, 1–88 (76).

Shleifer, Andrei and Lawrence H. Summers (1990). 'The Noise Trader Approach to Finance'. In: *Journal of Economic Perspectives* 4 (2), 19–33. doi: 10.1257/jep.4.2.19 (176, 248).

Shlesinger, Michael F. (2000). 'Exploring phase space'. In: *Nature* 405 (6783), 135–137. doi: 10.1038/35012197 (71).

Sigler, Laurence (2002). *Fibonacci's Liber Abaci*. Springer, New York. doi: 10.1007/978-1-4613-0079-3 (81).

Silva, Cesar E. (2007). *Invitation to Ergodic Theory*. American Mathematical Society, Providence (77).

Silva, Cesar E. (2016). 'WHAT IS ... an Ergodic Transformation?' In: *Notices of the American Mathematical Society* 63 (1), 26–27. doi: 10.1090/noti1318 (91).

Simon, Herbert A. (1947). *Administrative Behavior. A Study of Decision-Making Processes in Administrative Organizations.* Macmillan, New York (251).

Simon, Herbert A. (1955). 'A Behavioral Model of Rational Choice'. In: *Quarterly Journal of Economics* 69 (1), 99–118. JSTOR: 1884852 (251).

Simon, Herbert A. (1956). 'Rational choice and the structure of the environment.' In: *Psychological Review* 63 (2), 129–138. doi: 10.1037/h0042769 (251).

Simon, Herbert A. (1957). *Models of Man. Social and Rational. Mathematical Essays on Rational Human Behavior in a Social Setting.* Wiley, Oxford (251).

Simons, Andrew M. (2009). 'Fluctuating natural selection accounts for the evolution of diversification bet hedging'. In: *Proceedings of the Royal Society B: Biological Sciences* 276 (1664), 1987–1992. doi: 10.1098/rspb.2008.1920 (214).

Simons, Andrew M. (2011). 'Modes of response to environmental change and the elusive empirical evidence for bet hedging'. In: *Proceedings of the Royal Society B: Biological Sciences* 278 (1712), 1601–1609. doi: 10.1098/rspb.2011.0176 (214).

Sims, Christopher A. (1980). 'Macroeconomics and Reality'. In: *Econometrica* 48 (1), 1–48. doi: 10.2307/1912017 (198).

Sims, Christopher A. (2002). 'Structural VAR's'. In: url: http://www.uh.edu/~bsorense/SimsVARS.pdf (198).

Sims, Christopher A. (2003). 'Implications of rational inattention'. In: *Journal of Monetary Economics* 50 (3), 665–690. doi: 10.1016/s0304-3932(03)00029-1 (248).

Sims, Christopher A. (2006). 'Rational Inattention: A Research Agenda'. In: url: http://sims.princeton.edu/yftp/RIplus/RatInattPlus.pdf (248).

Sims, Christopher A. (2010). 'But Economics Is Not an Experimental Science'. In: *Journal of Economic Perspectives* 24 (2), 59–68. doi: 10.1257/jep.24.2.59 (184, 187).

Sinai, Yakov G. (1963). 'On the foundations of the ergodic hypothesis for a dynamical system of statistical mechanics'. In: *Doklady Akademiĭ Nauk USSR* 153 (6). (in Russian), 1261–1264. url: http://mi.mathnet.ru/eng/dan28929 (74, 75).

Sinai, Yakov G. (1970). 'Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards'. In: *Russian Mathematical Surveys* 25 (2), 137–189. doi: 10.1070/rm1970v025n02abeh003794 (74, 76).

Sinai, Yakov G. (1972). 'Gibbs Measures in Ergodic Theory'. In: *Russian Mathematical Surveys* 27 (4), 21–69. doi: 10.1070/rm1972v027n04abeh001383 (75).

Sinai, Yakov G. (1977). *Introduction to Ergodic Theory.* Princeton University Press (77).

Sinai, Yakov G. (1989). 'Kolmogorov's work on ergodic theory'. In: *Annals of Probability* 17 (3). doi: 10.1214/aop/1176991249 (78).

Sinai, Yakov G. and Nikolai I. Chernov (1987). 'Ergodic properties of certain systems of two-dimensional discs and three-dimensional balls'. In: *Russian Mathematical Surveys* 42 (3), 181–207. doi: 10.1070/rm1987v042n03abeh001421 (74).

Singh, Navinder (2011). 'How and why does statistical mechanics work'. In: arXiv: 1103.4003v2 (50, 71).

Sinn, Hans-Werner (1980). *Ökonomische Entscheidung bei Ungewißheit*. Zugleich Diss. 1977 Universität Mannheim. Mohr Siebeck, Tübingen (336).

Sinn, Hans-Werner (1989). *Economic Decisions Under Uncertainty*. 2nd ed. English translation of Sinn (1980). Physica-Verlag, Heidelberg (129, 130, 170, 255).

Sinn, Hans-Werner (2003). 'Weber's Law and the Biological Evolution of Risk Preferences: The Selective Dominance of the Logarithmic Utility Function, 2002 Geneva Risk Lecture'. In: *Geneva Papers on Risk and Insurance Theory* 28 (2), 87–100. doi: 10.1023/a:1026384519480 (129, 130, 170).

Sinn, Hans-Werner and Alfons J. Weichenrieder (1993). 'Biological Selection of Risk Preferences'. In: *Risk is a Construct. Perceptions of Risk Perceptions*. Ed. by Bayerische Rück. Knesebeck, München, 67–83 (129, 130, 170).

Sklar, Lawrence (1973). 'Statistical Explanation and Ergodic Theory'. In: *Philosophy of Science* 40 (2), 194–212. doi: 10.1086/288515 (71).

Sklar, Lawrence (1993). *Physics and Chance. Philosophical issues in the foundations of statistical mechanics*. Cambridge University Press (30, 69, 70, 75, 78).

Smale, Stephen J. (1967). 'Differentiable dynamical systems'. In: *Bulletin of the American Mathematical Society* 73 (6), 747–818. doi: 10.1090/s0002-9904-1967-11798-1 (75, 181).

Smale, Stephen J. (1973). 'Global Analysis and Economics I. Pareto Optimum and a Generalization of Morse Theory'. In: *Dynamical Systems. Proceedings of a Symposium held at the University of Bahia, Salvador, Brasil, July 26–August 14, 1971*. Ed. by Mauricio M. Peixoto. Academic Press, New York, 531–544. doi: 10.1016/b978-0-12-550350-1.50044-8 (181).

Smale, Stephen J. (1974a). 'Global analysis and economics IIA. Extension of a theorem of Debreu'. In: *Journal of Mathematical Economics* 1 (1), 1–14. doi: 10.1016/0304-4068(74)90031-7 (181).

Smale, Stephen J. (1974b). 'Global analysis and economics III. Pareto Optima and price equilibria'. In: *Journal of Mathematical Economics* 1 (2), 107–117. doi: 10.1016/0304-4068(74)90002-0 (181).

Smale, Stephen J. (1974c). 'Global analysis and economics IV. Finiteness and stability of equilibria with general consumption sets and production'. In: *Journal of Mathematical Economics* 1 (2), 119–127. doi: 10.1016/0304-4068(74)90003-2 (181).

Smale, Stephen J. (1974d). 'Global analysis and economics V. Pareto theory with constraints'. In: *Journal of Mathematical Economics* 1 (3), 213–221. doi: 10.1016/0304-4068(74)90013-5 (181).

Smale, Stephen J. (1976). 'Dynamics in General Equilibrium Theory'. In: *American Economic Review Papers and Proceedings* 66 (2), 288–294. JSTOR: 1817235 (181).

Smale, Stephen J. (1980a). 'On the Problem of Reviving the Ergodic Hypothesis of Boltzmann and Birkhoff'. In: *Annals of the New York Academy of Sciences* 357 (1), 260–266. doi: 10.1111/j.1749-6632.1980.tb29691.x (181).

Smale, Stephen J. (1980b). *The Mathematics of Time. Essays on Dynamical Systems, Economic Processes, and Related Topics*. Springer, New York. doi: 10.1007/978-1-4613-8101-3 (181).

Smale, Stephen J. (June 2000). *The Collected Papers of Stephen Smale*. Ed. by Felipe

Cucker and Roderick S. C. Wong. Vol. 1. World Scientific. doi: 10.1142/4424-vol1 (181).

Smith, Adam (1853). *The Theory of Moral Sentiments. An Essay towards an Analysis of the Principles by which Men naturally judge concerning the Conduct and Character, first of their Neighbours, and afterwards of themselves.* To which is added, A Dissertation on the Origins of Languages. New Edition. With a biographical and critical Memoir of the Author. Ed. by Dugald Stewart. first published in 1759. Henry G. Bohn, London. url: https://oll.libertyfund.org/titles/2620 (246).

Smith, Adam (1904). *An Inquiry into the Nature and Causes of the Wealth of Nations. edited with an Introduction, Notes, Marginal Summary and an enlarged Index.* Ed. by Edwin Cannan. Vol. I. first published in 1776. Methuen, London. url: http://oll.libertyfund.org/titles/237 (246).

Smith, Vernon L. (2003). 'Constructivist and Ecological Rationality in Economics'. In: *American Economic Review.* Lecture Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel, 465–508. doi: 10.1257/000282803322156954 (251).

Smith, Vernon L. (2007). *Rationality in Economics. Constructivist and Ecological Forms.* Cambridge University Press (251).

Smolin, Lee (1992). 'Did the Universe evolve?' In: *Classical and Quantum Gravity* 9 (1), 173–191. doi: 10.1088/0264-9381/9/1/016 (100).

Smolin, Lee (1997). *The Life of the Cosmos.* Oxford University Press (100).

Smolin, Lee (2004). 'Cosmological natural selection as the explanation for the complexity of the universe'. In: *Physica A: Statistical Mechanics and its Applications* 340 (4), 705–713. doi: 10.1016/j.physa.2004.05.021 (100).

Smolin, Lee (2006). *The Trouble with Physics.* Houghton Mifflin, New York (100).

Smolin, Lee (2009). 'Time and symmetry in models of economic markets'. In: arXiv: 0902.4274 (216, 232).

Smolin, Lee (2013). *Time Reborn. From the Crisis of Physics to the Future of the Universe.* Allen Lane, London (100, 216, 232).

Smolin, Lee (2014). 'Time, laws, and the future of cosmology'. In: *Physics Today* 67 (3), 38–43. doi: 10.1063/PT.3.2310 (216, 232).

Smolin, Lee (2015). 'Temporal Naturalism'. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 52 Part A, 86–102. doi: 10.1016/j.shpsb.2015.03.005 (216, 232).

Sokolov, Igor M. (2008). 'Statistics and the single molecule'. In: *Physics* 1, 1–3. doi: 10.1103/physics.1.8 (208).

Sonnenschein, Hugo (1972). 'Market Excess Demand Functions'. In: *Econometrica* 40 (3), 549–563. doi: 10.2307/1913184 (180, 227).

Sonnenschein, Hugo (1973). 'Do Walras' identity and continuity characterize the class of community excess demand functions?' In: *Journal of Economic Theory* 6 (4), 345–354. doi: 10.1016/0022-0531(73)90066-5 (180, 227).

Soros, George (1994). *The Alchemy of Finance. Reading the Mind of the Market.* John Wiley & Sons, New York (238).

Soros, George (2013). 'Fallibility, reflexivity, and the human uncertainty principle'. In: *Journal of Economic Methodology* 20 (4), 309–329. doi: 10.1080/1350178X.2013.859415 (236).

Starrfelt, Jostein and Hanna Kokko (2012). 'Bet-hedging-a triple trade-off between means, variances and correlations'. In: *Biological Re-*

views 87 (3), 742–755. DOI: 10.1111/j.1469-185x.2012.00225.x (214).

Stefani, Fernando D., Jacob P. Hoogenboom and Eli Barkai (2009). 'Beyond quantum jumps: Blinking nanoscale light emitters'. In: *Physics Today* 62 (2), 34–39. DOI: 10.1063/1.3086100 (208).

Stigler, George J. (1950). 'The Development of Utility Theory. II'. In: *Journal of Political Economy* 58 (5), 373–396. DOI: 10.1086/256980 (129).

Stoop, Ruedi and Willi-Hans Steeb (2006). *Berechenbares Chaos in dynamischen Systemen*. Birkhäuser, Basel. DOI: 10.1007/3-7643-7551-5 (77).

Stoppard, Tom (2003). *Rosencrantz & Guildenstern Are Dead*. ebook (18).

Strevens, Michael (2003). *Bigger than Chaos. Understanding Complexity through Probability*. Harvard University Press (74).

Strotz, Robert H. (1955). 'Myopia and Inconsistency in Dynamic Utility Maximization'. In: *Review of Economic Studies* 23 (3), 165–180. DOI: 10.2307/2295722 (256).

Sunder, Shyam (2006). 'Economic Theory: Structural Abstraction or Behavioral Reduction?' In: *History of Political Economy* 38 (Suppl 1), 322–342. DOI: 10.1215/00182702-2005-028 (176, 248).

Syll, Lars P. (2015). *On the use and misuse of theories and models in mainstream economics*. World Economics Association (183).

Syll, Lars P. (2018). 'The main reason why almost all econometric models are wrong'. In: *World Economics Association Newsletter* 8 (3), 5–10. URL: https://www.worldeconomicsassociation.org/newsletterarticles/econometric-models-wrong/ (183, 184).

Sylla, Edith D. (2003). 'Business Ethics, Commercial Mathematics, and the Origins of Mathematical Probability'. In: *History of Political Economy* 35 (Suppl 1), 309–337. DOI: 10.1215/00182702-35-suppl_1-309 (83, 87).

Szász, Domokos (2000a). 'Boltzmann's Ergodic Hypothesis, a Conjecture for Centuries?' In: *Hard Ball Systems and the Lorentz Gas*. Ed. by Domokos Szász. Springer, Berlin Heidelberg, 421–446. DOI: 10.1007/978-3-662-04062-1_14 (78).

Szász, Domokos, ed. (2000b). *Hard Ball Systems and the Lorentz Gas*. Springer, Berlin Physica-Verlag, Heidelberg. DOI: 10.1007/978-3-662-04062-1 (74).

Szemerédi, Endre (1969). 'On sets of integers containing no four elements in arithmetic progression'. In: *Acta Mathematica Academiae Scientiarum Hungaricae* 20 (1-2), 89–104. DOI: 10.1007/BF01894569 (78).

Szemerédi, Endre (1975). 'On sets of integers containing k elements in arithmetic progression'. In: *Acta Arithmetica* 27 (1), 199–245. URL: http://eudml.org/doc/205339 (78).

Szenberg, Michael, Lall Ramrattan and Aron A. Gottesman, eds. (2006). *Samuelsonian Economics. and the Twenty-First Century*. Oxford University Press (222).

Tabor, Morten N. (2014). 'Essays on Imperfect Knowledge Economics, Structural Change, and Persistence in the Cointegrated VAR Model'. PhD thesis. University of Copenhagen. URL: http://research.ku.dk/search/?pure=en/publications/essays-on-imperfect-knowledge-economics-structural-change-and-persistence-in-the-cointegrated-var-model(333d72bf-ddc9-4873-a648-adce898470d4).html (198).

Takahashi, Taiki (2007). 'A probabilistic choice model based on Tsallis' statistics'. In: *Physica*

A 386 (1), 335–338. DOI: 10.1016/j.physa.2007.07.005 (205).

TALEB, NASSIM N. (2005). *Narren des Zufalls*. 2nd ed. Wiley, Weinheim (90, 236).

TALEB, NASSIM N. (2010b). *Der Schwarze Schwan - Konsequenzen aus der Krise*. Carl Hanser Verlag, München (90, 237).

TALEB, NASSIM N. (2010a). *Der schwarze Schwan*. Deutscher Taschenbuch Verlag, München (90, 237, 278).

TALEB, NASSIM N. (14th July 2014). 'Ergodicity'. In: *Facebook Post*. URL: bit.ly/MKTalebIYIErgodicity (2, 278).

TANAKA, KATSUTO (2017). *Time Series Analysis. Nonstationary and Noninvertible Distribution Theory*. 2nd ed. Wiley, Chennai (194).

TAO, TERENCE (2006). 'A Quantitative Ergodic Theory Proof of Szemerédi's Theorem'. In: *Electronic Journal of Combinatorics*, R99. DOI: 10.37236/1125. URL: https://www.combinatorics.org/ojs/index.php/eljc/article/view/v13i1r99 (78).

TAO, TERENCE (2007). 'The dichotomy between structure and randomness, arithmetic progressions, and the primes'. In: *Proceedings of the International Congress of Mathematicians. Madrid, August 22–30, 2006*. Ed. by MARTA SANZ-SOLÉ, JAVIER SORIA, JUAN L. VARONA and JOAN VERDERA. Vol. I. European Mathematical Society, Zurich, 581–608. URL: http://www.icm2006.org/proceedings/Vol_I/26.pdf (78).

TAO, TERENCE (2008). 'Norm convergence of multiple ergodic averages for commuting transformations'. In: *Ergodic Theory and Dynamical Systems* 28 (2), 657–688. DOI: 10.1017/s0143385708000011 (78).

TAO, TERENCE (2015). 'Failure of the L^1 Pointwise and Maximal Ergodic Theorems for the Free Group'. In: *Forum of Mathematics, Sigma* 3, e27. DOI: 10.1017/fms.2015.28 (78).

TARTAGLIA, NICOLO (1556). *La prima parte del general trattato di numeri, et misure*. Curtio Troiano dei Navò, Venedig. DOI: 10.3931/e-rara-19205 (81).

TEGMARK, MAX (1998). 'The Interpretation of Quantum Mechanics: Many Worlds or Many Words?' In: *Fortschritte der Physik* 46 (6-8), 855–862. DOI: 10.1002/(SICI)1521-3978(199811)46:6/8<855::AID-PROP855>3.0.CO;2-Q (213).

TEGMARK, MAX (2003). 'Parallel Universes'. In: *Scientific American* 288 (5), 41–51. DOI: 10.1038/scientificamerican0503-40 (212).

TEGMARK, MAX (2007). 'Many lives in many worlds'. In: *Nature* 448 (7149), 23–24. DOI: 10.1038/448023a (212).

THALER, RICHARD H. (1981). 'Some empirical evidence on dynamic inconsistency'. In: *Economics Letters* 8 (3), 201–207. DOI: 10.1016/0165-1765(81)90067-7 (256).

THOMSON, WILLIANM (1874). 'Kinetic Theory of the Dissipation of Energy'. In: *Nature* 9 (232), 441–444. DOI: 10.1038/009441c0 (48).

THORP, EDWARD O. (1969). 'Optimal Gambling Systems for Favorable Games'. In: *Review of the International Statistical Institute* 37 (3), 273–293. DOI: 10.2307/1402118 (90, 265).

THORP, EDWARD O. (1973). 'A corrected derivation of the Black-Scholes option model. Presented at the CRSP meeting 1973'. In: *CRSP Proceedings 1976* (90, 265, 278).

THORP, EDWARD O. (2006). 'The Kelly Criterion in Blackjack, Sports Betting and the Stock Market'. In: *Handbook of Asset and Liability Management*. Ed. by STAVROS A. ZENIOS and WILLIAM T. ZIEMBA. Vol. 1. Elsevier, North-Holland Amsterdam. Chap. 9, 385–428.

DOI: 10.1016/B978-044453248-0.50015-0 (259, 268, 275).

Thorp, Edward O. (2017). *A Man for All Markets. From Las Vegas to Wall Street*. Random House, New York (90).

Thurner, Stefan and Rudolf Hanel (2012). 'The Entropy of Non-Ergodic Complex Systems – A Derivation from First Principles'. In: *International Journal of Modern Physics: Conference Series* 16, 105–115. DOI: 10.1142/s2010194512007817 (205).

Tinbergen, Jan (1939). *Statistical Testing of Business Cycle Theories. 1 A Method and Its Application to Investment Activity*. League of Nations, Geneva (182, 278).

Tinbergen, Jan (1940). 'On a Method of Statistical Business-Cycle Research. A Reply'. In: *Economic Journal* 50 (197), 141–154. DOI: 10.2307/2225763 (182).

Todd, Peter M., Gerd Gigerenzer and the ABC Research Group, eds. (2012). *Ecological Rationality. Intelligence in the World*. ABC stands for the research group at the Center for Adaptive Behavior and Cognition at the Max-Planck-Institute for Human Development in Berlin, Germany. Oxford University Press (251).

Todhunter, Isaac (1865). *A History of the Mathematical Theory of Probability. From the Time of Pascal to that of Laplace*. Macmillan & Co., Cambridge (100, 103, 124).

Tolman, Richard C. (1938). *The Principles of Statistical Mechanics*. Clarendon Press, Oxford (30, 71).

Tretyakov, Michael V. (2013). *Introductory Course on Financial Mathematics*. Imperial College Press, London (278).

Trockel, Walter (1983). 'Uniqueness of Mean Maximizers via an Ergodic Theorem'. In: *Mathematische Operationsforschung und Statistik. Series Optimization* 14 (3), 411–419. DOI: 10.1080/02331938308842873 (180).

Trockel, Walter (1984). *Market Demand. An Analysis of Large Economies with Non-Convex Preferences*. Springer, Berlin Heidelberg. DOI: 10.1007/978-3-642-46488-1 (180, 181).

Tsallis, Constantino (1988). 'Possible generalization of Boltzmann-Gibbs statistics'. In: *Journal of Statistical Physics* 52 (1-2), 479–487. DOI: 10.1007/bf01016429 (203, 204).

Tsallis, Constantino (2009). *Introduction to Nonextensive Statistical Mechanics. Approaching a Complex World*. Springer, New York. DOI: 10.1007/978-0-387-85359-8 (205, 206).

Tsallis, Constantino (2014). 'An introduction to nonadditive entropies and a thermostatistical approach to inanimate and living matter'. In: *Contemporary Physics* 55 (3), 179–197. DOI: 10.1080/00107514.2014.900977 (203).

Tsallis, Constantino, Celia Anteneodo, Lisa Borland and Roberto Osorio (2003). 'Nonextensive statistical mechanics and economics'. In: *Physica A* 324 (1-2), 89–100. DOI: 10.1016/s0378-4371(03)00042-6 (202, 203).

Tsay, Ruey T. (2002). *Analysis of Financial Time Series. Financial Econometrics*. Wiley, New York (194).

Tsay, Ruey T. and Rong Chen (2019). *Nonlinear Time Series Analysis*. Wiley, Hoboken (194).

Tversky, Amos (1975). 'A critique of expected utility theory: Descriptive and normative considerations'. In: *Erkenntnis* 9 (2), 163–173. DOI: 10.1007/bf00226380 (134).

Tversky, Amos and Daniel Kahneman (1986). 'Rational choice and the framing of decisions'. In: *Journal of Business*, S251–S278. JSTOR: 2352759 (134).

TVERSKY, AMOS and DANIEL KAHNEMAN (1992). 'Advances in Prospect Theory: Cumulative Representation of Uncertainty'. In: *Journal of Risk and Uncertainty* 5 (4), 297–323. DOI: 10.1007/BF00122574 (126, 134, 135, 248, 256, 258).

UFFINK, JOS (2007). 'Compendium of the Foundations of Classical Statistical Physics'. In: *Philosophy of Physics. Part B*. Ed. by JEREMY BUTTERFIELD and JOHN EARMAN. Nort Holland, Amsterdam, 923–1074. URL: http://philsci-archive.pitt.edu/2691/ (71).

ULAM, STANISLAW M. (1976). *Adventures of a Mathematician*. Charles Scribner's Sons, New York (63, 73, 248).

VARIAN, HAL R. (2010). *Intermediate Microeconomics. A Modern Approach*. W.W. Norton & Company, New Academic Press, New York (119).

VEBLEN, THORSTEIN (1898). 'Why is Economics not an Evolutionary Science?' In: *Quarterly Journal of Economics* 12 (4), 373–397. JSTOR: 1882952 (213).

VERSHIK, ANATOLY M. (2006). 'Kantorovich Metric: Initial History and Little-Known Applications'. In: *Journal of Mathematical Sciences* 133 (4), 1410–1417. DOI: 10.1007/s10958-006-0056-3 (39).

VERSHIK, ANATOLY M. (2013). 'Long History of the Monge-Kantorovich Transportation Problem'. In: *Mathematical Intelligencer* 35 (4), 1–9. DOI: 10.1007/s00283-013-9380-x (39).

VIANA, MARCELO and KRERLEY OLIVEIRA (2016). *Foundations of Ergodic Theory*. Cambridge University Press. DOI: 10.1017/cbo9781316422601 (56, 75, 77).

VILLANI, CÉDRIC (2008). '*H*-theorem and beyond: Boltzmann's entropy in today's mathematics'. In: *Boltzmann's Legacy*. Ed. by GIOVANNI GALLAVOTTI, WOLFGANG REITER and JAKOB YNGVASON. European Mathematical Society, Zurich, 129–143. DOI: 10.4171/057 (42).

VILLANI, CÉDRIC (2009). *Optimal Transport. Old and New*. Springer, Berlin Heidelberg. DOI: 10.1007/978-3-540-71050-9 (39).

VOGL, JOSEPH (2011). *Das Gespenst des Kapitals*. 5th ed. (199).

VRANAS, PETER B. M. (1998). 'Epsilon-Ergodicity and the Success of Equilibrium Statistical Mechanics'. In: *Philosophy of Science* 65 (4), 688–708. DOI: 10.1086/392667 (72).

WALD, ABRAHAM (1939). 'Contributions to the Theory of Statistical Estimation and Testing Hypotheses'. In: *Annals of Mathematical Statistics* 10 (4), 299–326. DOI: 10.1214/aoms/1177732144 (257).

WALD, ABRAHAM (1945). 'Statistical Decision Functions Which Minimize the Maximum Risk'. In: *Annals of Mathematics* 46 (2), 265. DOI: 10.2307/1969022 (256, 257).

WALD, ABRAHAM (1950). *Statistical Decision Functions*. Wiley, New York (256).

WALRAS, LÉON (1874). *Éléments d'économie politique pure ou théorie de la richesse sociale*. Corbaz, Lausanne (278).

WALTERS, PETER (1982). *An Introduction to Ergodic Theory*. Springer, New York (64, 77).

WALTHER, THOMAS (2017). 'Essays on Financial Econometrics. with Applications to Commodity, Equity, and Foreign Exchange Markets'. PhD thesis. TU Dresden, Economics Department. DOI: 10.13140/RG.2.2.27680.48641 (202).

WARD, TOM (2009). 'Ergodic Theory: Interactions with Combinatorics and Number Theory'. In: *Encyclopedia of Complexity and Systems Science*. Ed. by ROBERT A. MEYERS. Springer, New York, 3040–3053. DOI: 10.1007/978-0-387-30440-3_181 (173).

WATKINS, NICHOLAS W. (2017). 'On the continuing relevance of Mandelbrot's non-ergodic fractional renewal models of 1963 to 1967'. In: *European Physical Journal B* 90 (12). DOI: 10.1140/epjb/e2017-80357-3 (200, 202).

WATSON, HENRY W. and EDWARD P. CULVERWELL (1894). 'Boltzmann's Minimum Function'. In: *Nature* 51 (1309), 105. DOI: 10.1038/051105b0 (46).

WEAVER, WARREN (1948). 'Science and Complexity'. In: *American Scientist* 36 (4), 536–544. JSTOR: 27826254 (205).

WEBER, ERNST HEINRICH (1834). *De Pulsu, Resorptione, Auditu et Tactu. Annotationes Anatomicae et Physiologicae.* C. F. Koehler, Leipzig. Google Books: 5WtFAQAAIAAJ (113, 129).

WEBER, ERNST HEINRICH (1851). *Die Lehre vom Tastsinne und Gemeingefühle. auf Versuche gegründet.* Friedrich Vieweg und Sohn, Braunschweig. Google Books: NVlVAAAAMAAJ (129).

WEIDLICH, WOLFGANG (1991). 'Physics and social science - the approach of synergetics'. In: *Physics Reports* 204 (1), 1–163. DOI: 10.1016/0370-1573(91)90024-G (38).

WEIGEL, A. V., B. SIMON, M. M. TAMKUN and D. KRAPF (2011). 'Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking'. In: *Proceedings of the National Academy of Sciences* 108 (16), 6438–6443. DOI: 10.1073/pnas.1016325108 (208).

WEINTRAUB, E. ROY (1991). *Stabilizing Dynamics. Constructing Eonomic Knowledge.* Cambridge University Press (220, 222).

WEINTRAUB, E. ROY (2002). *How economics became a mathematical science.* Duke University Press Books (279).

WEISSERT, THOMAS P. (1997). *The Genesis of Simulation in Dynamics. Pursuing the Fermi-Pasta-Ulam Problem.* Springer, New York. DOI: 10.1007/978-1-4612-1956-9 (73).

VON WEIZSÄCKER, CARL FRIEDRICH (1939). 'Der zweite Hauptsatz und der Unterschied von Vergangenheit und Zukunft'. In: *Annalen der Physik* 428 (3-4), 275–283. DOI: 10.1002/andp.19394280309 (232).

VON WEIZSÄCKER, C. CHRISTIAN (1997). 'Paul Samuelsons wissenschaftliche Wirkung'. In: *Paul A. Samuelsons »Foundations of Economic Analysis«.* Ed. by JÜRG NIEHANS, PAUL SAMUELSON and C. CHRISTIAN VON WEIZSÄCKER. Vademecum zu einem Klassiker der Gegenwart. Verlag Wirtschaft und Finanzen, Düsseldorf, 89–105 (278).

WENZELBURGER, JAN (2009). 'Perfect Forecasting, Behavioral Heterogeneities, and Asset Prices'. In: *Handbook of Financial Markets. Dynamics and Evolution.* Ed. by THORSTEN HENS and KLAUS R. SCHENK-HOPPÉ. Nort Holland, Amsterdam. Chap. 6, 345–401. DOI: 10.1016/b978-012374258-2.50010-5 (173, 174, 176, 178).

WERNDL, CHARLOTTE (2013). 'Justifying typicality measures of Boltzmannian statistical mechanics and dynamical systems'. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 44 (4), 470–479. DOI: 10.1016/j.shpsb.2013.08.006 (71).

WERNDL, CHARLOTTE and ROMAN FRIGG (2015). 'Reconceptualising equilibrium in Boltzmannian statistical mechanics and characterising its existence'. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 49, 19–31. DOI: 10.1016/j.shpsb.2014.12.002 (71).

WERNDL, CHARLOTTE and ROMAN FRIGG (2017). 'Mind the Gap: Boltzmannian versus Gibbsian Equilibrium'. In: *Philosophy of Science* 84 (5), 1289–1302. DOI: 10.1086/694088 (71).

Weyl, Hermann (1916). 'Über die Gleichverteilung von Zahlen mod. Eins'. In: *Mathematische Annalen* 77 (3), 313–352. doi: 10.1007/bf01475864 (69).

White, Halbert (1984). *Asymptotic Theory for Econometricians*. Academic Press, Orlando (194).

Whitehead, Alfred N. (1978). *Process and Reality*. The Free Press, New York (215).

Whitehead, Alfred N. and Bertrand Russell (1910). *Principia Mathematica*. Vol. I. Cambridge University Press (225).

Whitehead, Alfred N. and Bertrand Russell (1927a). *Principia Mathematica*. Vol. II. Cambridge University Press (225).

Whitehead, Alfred N. and Bertrand Russell (1927b). *Principia Mathematica*. Vol. III. Cambridge University Press (225).

Whitworth, William Allen (1870). *Choice and Chance*. second edition, enlarged. Deighton Bell, Cambridge. url: bit.ly/MKWhitworth1870 (162, 169, 278).

Whitworth, William Allen (1886). *Choice and Chance*. Fourth edition, enlarged. Deighton Bell, Cambridge. url: bit.ly/MKWhitworth1886 (86, 162).

Whitworth, William Allen (1901). *Choice and Chance. With 1000 Exercises*. fifth edition much enlarged. Deighton Bell, Cambridge. url: bit.ly/MKWhitworth1901 (162–164).

Wiener, Norbert (1920). 'The Mean of a Functional of Arbitrary Elements'. In: *The Annals of Mathematics* 22 (2), 66. doi: 10.2307/1967855 (14).

Wiener, Norbert (1921b). 'The Average of an Analytic Functional and the Brownian Movement'. In: *Proceedings of the National Academy of Sciences* 7 (10), 294–298. doi: 10.1073/pnas.7.10.294 (14).

Wiener, Norbert (1921a). 'The Average of an Analytic Functional'. In: *Proceedings of the National Academy of Sciences* 7 (9), 253–260. doi: 10.1073/pnas.7.9.253 (14).

Wiener, Norbert (1923). 'Differential-Space'. In: *Journal of Mathematical Physics* 2, 131–174 (14, 278).

Wiener, Norbert (1939). 'The Ergodic Theorem'. In: *Duke Mathematical Journal* 5 (1), 1–18. doi: 10.1215/S0012-7094-39-00501-6 (30, 56, 222, 278).

Wiener, Norbert (1956a). 'Nonlinear Prediction and Dynamics'. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. Ed. by Jerzy Neyman. Vol. 3: Contributions to Astronomy and Physics. December 1954 and July-August 1955 Statistical Laboratory of the University of California, Berkeley. University of California Press, 247–252. url: https://projecteuclid.org/euclid.bsmsp/1200510046 (13, 222).

Wiener, Norbert (1956b). 'What is Information Theory?' In: *IRE Transactions on Information Theory* 2 (2), 48. doi: 10.1109/TIT.1956.1056796 (277).

Wiener, Norbert (1964a). *I Am a Mathematician. The Later Life of a Prodigy*. MIT Press (15, 278).

Wiener, Norbert (1964b). *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*. MIT Press (239, 240).

Wiener, Norbert (1985). *Cybernetics. Or control and communication in the animal and the machine*. 2nd ed. [1948]. MIT Press, Cambridge Massachusetts (14, 236, 238, 239, 279).

Wiener, Norbert and Aurel Wintner (1941a). 'Harmonic Analysis and Ergodic Theory'. In: *American Journal of Mathematics* 63 (2), 415. doi: 10.2307/2371534 (222).

WIENER, NORBERT and AUREL WINTNER (1941b). 'On the Ergodic Dynamics of Almost Periodic Systems'. In: *American Journal of Mathematics* 63 (4), 794–824. DOI: 10.2307/2371623 (222).

WIENER, NORBERT and AUREL WINTNER (1943). 'The Discrete Chaos'. In: *American Journal of Mathematics* 65 (2), 279. DOI: 10.2307/2371816 (222).

WIENER, NORBERT and AUREL WINTNER (1958). 'Random Time'. In: *Nature* 181 (4608), 561–562. DOI: 10.1038/181561a0 (222).

WIGHTMAN, ARTHUR S. (1985). 'Introduction to the Problems'. In: *Regular and Chaotic Motions in Dynamic Systems*. Ed. by GIORGIO VELO and ARTHUR S. WIGHTMAN. Plenum Press, New York, 1–26 (73).

WILKINSON, AMIE (2009). 'Smooth Ergodic Theory'. In: *Encyclopedia of Complexity and Systems Science*. Ed. by ROBERT A. MEYERS. Springer, New York, 8168–8183. DOI: 10.1007/978-0-387-30440-3_484 (173).

WILSON, EDWIN B. (1945). 'Obituary George David Birkhoff'. In: *Science* 102 (2658), 578–580. DOI: 10.1126/science.102.2658.578 (66).

WINDRUM, PAUL, GIORGIO FAGIOLO and ALESSIO MONETA (2007). 'Empirical Validation of Agent-Based Models: Alternatives and Prospects'. In: *Journal of Artificial Societies and Social Simulation* 10 (2). URL: http://jasss.soc.surrey.ac.uk/10/2/8.html (196).

WINTNER, AUREL (1932). 'Remarks on the Ergodic Theory of Birkhoff'. In: *Proceedings of the National Academy of Sciences* 18 (3), 248. URL: www.pnas.org/content/18/3/248.full.pdf (219).

WITT, ULRICH (2003). *The Evolving Economy: Essays on the Evolutionary Approach to Economics*. Edward Elgar, Cheltenham (235).

WOLD, HERMAN O. (1969). 'Econometrics as Pioneering in Nonexperimental Model Building'. In: *Econometrica* 37 (3), 369–381. DOI: 10.2307/1912787 (193).

WOLD, HERMAN and LARS JURÉEN (1953). *Demand Analysis. A study in econometrics*. John Wiley & Sons, New York (195).

WOLFE, GARY (2007). 'Surfing the multiverse'. In: *Nature* 448 (7149), 25–26. DOI: 10.1038/448025a (212).

YANG, JANGHO (2017). 'Information Theoretic Approaches in Economics'. In: *Journal of Economic Surveys* 32 (3), 940–960. DOI: 10.1111/joes.12226 (282).

YOSHIMURA, JIN, HIROMU ITO, DONALD G. MILLER III and KEI-ICHI TAINAKA (2013a). 'Dynamic decision-making in uncertain environments I. The principle of dynamic utility'. In: *Journal of Ethology* 31 (2), 101–105. DOI: 10.1007/s10164-013-0362-4 (214).

YOSHIMURA, JIN, HIROMU ITO, DONALD G. MILLER III and KEI-ICHI TAINAKA (2013b). 'Dynamic decision-making in uncertain environments II. Allais paradox in human behavior'. In: *Journal of Ethology* 31 (2), 107–113. DOI: 10.1007/s10164-013-0359-z (214).

YOSHIMURA, JIN, YUMI TANAKA, TATSUYA TOGASHI, SHIGEHIDE IWATA and KEI-ICHI TAINAKA (2009). 'Mathematical equivalence of geometric mean fitness with probabilistic optimization under environmental uncertainty'. In: *Ecological Modelling* 220 (20). See also the errata https://doi.org/10.1016/j.ecolmodel.2010.11.001, 2611–2617. DOI: 10.1016/j.ecolmodel.2009.06.046 (214).

YOSIDA, KÔSAKU and SHIZUO KAKUTANI (1939). 'Birkhoff's ergodic theorem and the maximal ergodic theorem'. In: *Proceedings of the Imperial Academy of Japan* 15 (6), 165–168. DOI: 10.3792/pia/1195579375 (56, 64).

Zapart, Christopher A. (2015). 'Econophysics: A challenge to econometricians'. In: *Physica A* 419, 318–327. DOI: 10.1016/j.physa.2014.10.013 (214).

Zermelo, Ernst (1896a). 'Ueber einen Satz der Dynamik und die mechanische Wärmetheorie'. In: *Annalen der Physik* 293 (3), 485–494. DOI: 10.1002/andp.18962930314 (48).

Zermelo, Ernst (1896b). 'Ueber mechanische Erklärungen irreversibler Vorgänge. Eine Antwort auf Hrn. Boltzmann's „Entgegnung"'. In: *Annalen der Physik* 295 (12), 793–801. DOI: 10.1002/andp.18962951216 (48).

Zhang, Ruixun, Thomas J. Brennan and Andrew W. Lo (2014a). 'Group Selection as Behavioral Adaptation to Systematic Risk'. In: *PLoS ONE* 9 (10), e110848. DOI: 10.1371/journal.pone.0110848 (214).

Zhang, Ruixun, Thomas J. Brennan and Andrew W. Lo (2014b). 'The origin of risk aversion'. In: *Proceedings of the National Academy of Sciences* 111 (50), 17777–17782. DOI: 10.1073/pnas.1406755111 (214).

Zhou, Rongxi, Ru Cai and Guanqun Tong (2013). 'Applications of Entropy in Finance: A Review'. In: *Entropy* 15 (12), 4909–4931. DOI: 10.3390/e15114909 (282).

Ziemba, Rachel E. S. and William T. Ziemba (2008). *Scenarios for Risk Management and Global Investment Strategies*. John Wiley & Sons, Chichester (176).

Ziemba, William T. and Raymond G. Vickson (2006). *Stochastic Optimization. Models in Finance*. 2nd ed. World Scientific, Singapore. DOI: 10.1016/c2013-0-11739-4 (176).

Ziliak, Stephen T. (2010). 'The Validus Medicus and a new gold standard'. In: *Lancet* 376 (9738), 324–325. DOI: 10.1016/S0140-6736(10)61174-9 (187).

Zund, Joseph D. (2002). 'George David Birkhoff and John von Neumann: A Question of Priority and the Ergodic Theorems, 1931–1932'. In: *Historia Mathematica* 29 (2), 138–156. DOI: 10.1006/hmat.2001.2338 (63).