Friedrich-Alexander-Universität
Naturwissenschaftliche Fakultät

Master's Thesis
in Physics

SUPERBALLISTICITY IN ANOMALOUS DIFFUSION

Lennard Kossmann

Supervisor: Prof. Dr. Michael Schmiedeberg

Soft Matter Theory Group
Lab for Emergent Phenomena

Institut für Theoretische Physik I
Friedrich-Alexander-Universität Erlangen-Nürnberg

Work done in the group of Prof. Vasily Zaburdaev

Immunophysics Division
Zaburdaev group

Max-Planck-Zentrum für Physik in der Medizin

Submission date: 12.05.2026

Abstract

Studying the motion of biological cells is an inherently ambiguous endeavor due to the random nature of cell motility. Comparing the different motility patterns of cells is mostly based on estimating the *displacement* $x_r - x_s$ between times $s < r$ of a *typical* cell trajectory via different methods. Historically, two *mean squared displacement (MSD)* quantifiers are used for this: The *ensemble-averaged MSD (EAMSD)* EA_t, which averages the single displacement $x_t - x_0$ over multiple trajectories, and the *time-averaged MSD (TAMSD)* TA_Δ, which averages the displacements $x_{t+\Delta} - x_t$ of a single trajectory over the starting time t. Both the EAMSD ($EA_t \sim t^{\alpha_{EA}}$) and TAMSD ($TA_\Delta \sim \Delta^{\alpha_{AT}}$) tend to *scale* locally as a *power-law* in both theory and experiments, and the value of the *scaling exponents* α_{EA}, α_{TA} can be detrimental to the biophysical properties of the system. Although both MSDs agree with each other for simple systems, the movement of active particles typically shows a different scaling in EA- and TAMSD. While the interpretation of the EAMSD scaling is very intuitive, the TAMSD scaling is poorely understood. This is especially concerning, as the TAMSD is easier to measure in experiments and widely used to study cell motility. This work investigates the edge case where the EAMSD scales **superballistic**, i.e. $\alpha_{EA} > 2$, which is based on motility experiments of NK cells in biological tissue. This type of scaling suggests an underlying acceleration mechanics in the cell motility, which indicates active propulsion or a complex force field in the surrounding matter. In contrast, both the experiment and my calculations showed, that the TAMSD scales at most *ballistic*, i.e. $\alpha_{TA} = 2$, in these cases. One important tool for understanding the scaling of these two MSDs is the *velocity autocorrelation function (VACF)* $C_v(r,s) = \langle v_r v_s \rangle$. I investigate the properties of the VACF from a more theoretical perspective and show how it influences the scaling of both EAMSD and TAMSD directly and indirectly. A special role is given to *stationary* VACFs $C_v(r,s) = C(r-s)$, where the time - dependence lies solely in the lagtime $r - s$. The EAMSD and TAMSD agree for such systems, but the possibility of superballistic scaling is very narrow. More general, non-stationary VACFs show a greater variability in their EAMSD scaling, but their TAMSD does not cross the superballistic threshold. The underlying reason for this seems too complex to be answered in full generality, but for the cases of this thesis I demonstrate that the TAMSD scaling is determined by the *stationary part* of the VACF. This resolves the contrast to the EAMSD scaling, which is mostly affected by the *non-stationary part*. This thesis and the explored dichotomy between the EAMSD and TAMSD scaling highlights the difficulty in describing cell motility and other forms of stochastic motion both qualitatively and quantitatively, but also hints on how the behavior of the TAMSD can be utilized to attain a deeper understanding of the underlying mechanics itself.

Contents

1 Introduction

1.1 Dynamics of large systems and diffusion theory

Understanding the statistical behavior of many - particle systems has been one of the most prominent research questions in theoretical physics over the last two centuries. From very small subatomic scales to the giganteous dimensions of the universe, we have been able to get a quite solid understanding of the principles that govern nature. Aside of the unification of General Relativity and Quantum Field Theory, the dynamics of single to few particle systems has been solved. But these theories can only give appropriate answers for small systems. As soon as the number of particles increases, the equations become impossible to solve or to calculate numerically. This curse of many - particle systems is not unique to one particular subdiscipline, as physicists from classical and quantum mechanics have equally claimed despair over this. Now that we have arrived into a world, where semiconductors reach atomic scales and spectroscopy allows us to resolve biological systems on subcellular levels, we are confronted with the world of many - particle systems more than ever.

Demonstrating the sharp rise of complexity for larger systems is best illustrated by Newtonian gravity, where a system of massive point particles $r_i(t)$ evolve under Netwon's law:

$$m_i \ddot{r}_i = -G m_i \sum_{j \neq i} \frac{m_j (r_i - r_j)}{|r_i - r_j|^3} \tag{1.1}$$

The solution to eq. 1.1 for two particles was already established by Newton, and the tedious derivation of the solution is a standard two lecture endeavor in every introductory course on theoretical mechanics. Newton proved Kepler's theory that planets seem to follow elliptic orbits and trabants sling - shot in parabolic trajectories. All the possible solution could be topologically described using such conic sections.
Already for three particles, the dreaded **3 - body problem**, do we have to capitulate. Many aspiring scientists tried to find the analytical solution to the 3 - body problem, until Henri Poincare showed the impossibility of solving it using a closed expression. Even before the impossibility of a general solution to the more general N - body problem was shown, some scientists started in the mid to late 1800s to describe the behavior of many - particle systems using a different paradigm:

Beginning with Boltzmann and Maxwell, first attempts were made to understand the dynamics of particles in larger systems using statistical or **averaged** quantities. The first famous result has been the **Maxwell - Boltzmann distribution**,

$$f(v) = \left(\frac{m}{2\pi k_b T} \right)^{\frac{3}{2}} e^{-\frac{mv^2}{2\pi k_b T}}, \tag{1.2}$$

which is the probability distribution of the **speed** v of a particle within an ideal gas. There is no way to accurately predict the actual dynamics and trajectory of a single particle, but one can compute statistical averages in some cases. The discovery of the Maxwell - Boltzmann distribution was the advent of **kinetic gas theory**, which was the first success of statistical physics. It has been responsible for verifying the phenomenological predictions of gaseous thermodynamics, based on first principles. Kinetic gas theory allowed the calculation of averaged particle energies, averaged velocities and much more, which can be related to the thermodynamic properties of the larger system. This made kinetic gas theory inherently an **equilibrium theory** - it presupposes the thermodynamic equilibrium of the larger system. Although this is no restriction for ideal gases, many interesting physical system are out of equilibrium. The kinetic theory of gases can not be applied to these.

The major breakthrough for non - equilibrium system came in the *annus mirabilis* 1905 by Albert Einstein and concurrently 1906 by Marian Smoluchowski with the theoretical description of **Brownian motion**. The idea of Einstein was to study a light particle in a larger fluid / gas, with which it interacts only via collisions. Written in modern notation, both scientists considered the following modified Netwon equation:

$$m\dot{v}_t = -\gamma v_t + \zeta_t. \tag{1.3}$$

The dissipative relaxation term $-\gamma v_t$ represents the viscous friction in the engulfing fluid, whereas ζ_t represents the collision forces with the fluid molecules. Since it is not possible to analytically calculate the trajectories of the fluid molecules, they modeled the collision force ζ **statistically**:

1. The fluid is isotropic and the collisions have no preferred direction, so the statistical average of the force vanishes: $\langle \zeta_t \rangle = 0$
2. No two collisions occur simultaneously, but collisions happen frequently. It is unlikely that two subsequent collisions occur with the same fluid molecule. Since the fluid molecules are themselves uncorrelated[1], the collisions at two different times are independent from each other and uncorrelated. The force is statistically singularly correlated:

$$\langle \zeta_t \zeta_{t'} \rangle = 2D\delta(t - t') \tag{1.4}$$

Eq. 1.3 with this force ζ_t is called the **Langevin equation**. Einstein and Smoluchowski calculated the mean square displacement in eq. 1.3 from this Langevin equation in the limit $m \to 0$. The solution of eq. 1.3 is not a proper function, but rather a **stochastic process / stochastic trajectory** itself, called **Ornstein - Uhlenbeck process**. Einstein and Smoluchowski considered the solution in the limit of $m \to 0$, in which the solution is called **Brownian motion**. While the solution of eq. 1.3 is in itself random and no predictions can be made about the trajectories, they realized that it is indeed possible to compute **averages** of the trajectory process, especially the **mean - square displacement (MSD)** $\langle (x_t - x_0)^2 \rangle$. In the limit $m \to 0$ for Brownian motion, the MSD becomes linear:

$$\langle (x_t - x_0)^2 \rangle = 2Dt \tag{1.5}$$

---

[1] They only interact by collisions themselves.

If a particle behaves as Brownian motion, the MSD suggests that on average the particle moves in a fixed time t a distance of

$$|x_t - x_0| \sim \sqrt{Dt}.$$

It may seen counterintuitive, but this **square - root** behavior of the displacement has been observed in many different settings. This makes Brownian motion **universal**. Kinetic gas theory previously predicted thermodynamically relevant quantities like particle speeds and related them to macroscopic properties like pressure. It could not give qualitative and quantitative predictions on the particle trajectories themselves. The work of Einstein and Smoluchowski revolutionized the field, because it allows us to statistically quantify how such particles **move / diffuse** within the larger environment. With the advent of this so called **diffusion theory**, the particle trajectories in many - particle systems became a new field of study.

The Langevin equation did not only predict the mean square displacement in eq. 1.3, it allowed Einstein to calculate the **probability distribution** of such particles:

$$\rho(x,t) = \frac{1}{\sqrt{4Dt}} e^{-\frac{x^2}{2Dt}} \tag{1.6}$$

I cannot overstate the importance of this result: While the previous theories were only able to calculate averages or statistical properties, the new diffusion theory calculated probabilities. Not only did it reproduce the average results, the probabilistic theory allows to make quantitative predictions about the **particle fluctuations**: How much does the average particle deviate from eq. 1.3? What is the probability that a particle moves faster than the average?

The modern theory of diffusion is less a theory of thermodynamic averages anymore, but a theory of probability distributions. Particles are modeled using probabilistic models and their properties studied.

1.2 Cell motility and particle tracking

The theory of Brownian motion by Einstein and Smoluchowski has been observed many times in nature, as its movement can be even seen by some light pollen motion in the air. This allows to verify the theory with experiments. The statistical properties of such systems are well - known by now. Things are different in case of biophysics:

We have gathered a very mature understanding of the biochemical mechanisms that occur in living nature, from small bacteria to eucaryotic tissues. We have been able to sequence the DNA of a plethora of different species now, and the enrollment of DNA modification in plant crops and mRNA - vaccines are a testament to our sophisticated knowledge of the biochemistry of life. The progress in biochemistry seems quite surprising when compared to our current understanding of the statistical physics of life:

Biological systems like immune cells or large enzyme complexes are too big to be described using simple molecular quantum physics, but far too small to be in the realm of macroscopic theories like thermodynamics. They are inherently **mesoscopic**, and we need a new theoretical apparatus to describe the behavior of these systems.[2] While it was previously very difficult to capture motion of living cells in vivo or even in vitro, the advent of **single-particle tracking** (see [23] for a modern introduction) and other optical methods in the 90s allowed us now to even study the statistics of sub - cellular systems.

Motility and crowded environments

On a fundamental level, biological cells move completely different to Brownian particles: Brownian particles are not self - propelled, they cannot generate movement on their own. Their movement patterns are dictated by the external environment, i.e. the fluid molecules colliding with them. If the fluid is an ideal fluid, then the internal structure is homogeneous and the behavior of the fluid is determined by classical hydrodynamics. If one studies the motion of amoeboid cells in eucaryotic tissues, for example, the type of motion is completely different:

Amoeboid cells are able to perform self - propelled motion by means of rotational organelles like **flagella**, which resembles a form of swimming. They can perform active and directed motion by fueling on its metabolic energy, which is called **cell motility**. Swimming is an advantageous form of movement in a liquid - like environment. But the tissue of living organisms is not homogeneous. While the cytoplasm of larger cells can behave in a fluid - like fashion, the **extracellular matrix (ECM)** outside of the cell membranes forms a maze - like crowded environment. Within ECMs, swimming is not the only type of motion for such cells. It is possible for the cell to attach partly to the matrix walls and pull itself in direction of the wall due to the motion of the cytosceleton and the hardness of the wall. These cells can perform different types of motility patterns, depending on the environment. This makes the statistical description of their motility, compared to Brownian motion, extremely complicated.

1.3 Experimental setting

In her master's thesis (see [24]) Johanna Schürlein collected experimental data on the movement of natural killer cells (NK cells) in an external collagen matrix (see [3] for more info on the experimental setup). The external collagen matrix has been created artificially, but simulated the ECM observed in human tissue. The goal has been to reproduce the motility pattern of NK cells in human ECM and study the statistical behavior.

The experiment has been conducted in the following way: A high number of NK cells have been prepared in an artificial collagen matrix, simulating the ECM of living tissue. The NK cells freely traverse the collagen matrix and their movement is filmed in 11 different frames. For every single picture of the movie, the center of mass (CoM) of each NK cell in the frames, where it appears, is computed. The center of mass is then tracked over time as the *trajectory* x_i of the NK cell in a given frame. The trajectories are numerically centered $x_i(0) = 0$ for computational purposes. The sample

---

[2] Already their macroscopic limit theories are not understood. A prime example for this is the rheology of blood cells: While behaving in a fluid - like fashion, their viscoelastic properties defy the nature of the Navier - Stokes equation and naive fluid dynamics.

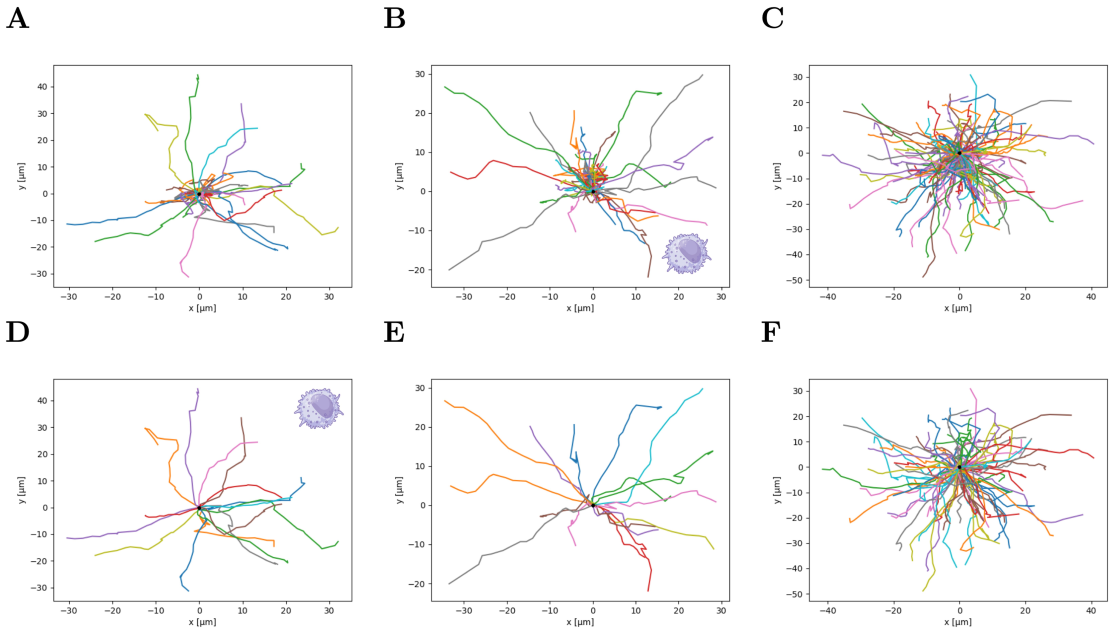

Figure 1.1: CoM trajectories for three frames (from [24]). In D - F, all inactive cells are removed.

CoM trajectories for three frames are shown in Fig. 1.1. The motility of the NK cells and their MSDs were studied using two different techniques:

For the **ensemble average MSD (EAMSD)** in a fixed frame, the squared displacement $(x_t - x_0)^2$ is averaged over all the CoM trajectories within the given frame:

$$\langle (x(t) - x(0))^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} (x_i(t) - x_i(0))^2 \, . \tag{1.7}$$

Conceptually, the EAMSD corresponds to the statistical average in the Langevin equation before and the stochastic average in later sections. The EAMSD for two arbitrary frames is shown in Fig. 1.2.

One characteristic of the EAMSDs of the different frames is the approximate power-law scaling $\langle (x_t - x_0)^2 \rangle \sim t^\alpha$ during two separate time regimes, which follows from the least-squares fit. The scaling exponent α for later times is roughly $\alpha \sim 2$, which is called **ballistic scaling**. A ballistic scaling suggests $|x_t - x_0| \sim t$, which corresponds to the displacement of a trajectory $x_t = x_0 + v_0 \cdot t$ with constant velocity.

The scaling exponent fit exceeds 2 for early times however, which is called **superballistic scaling**. Such a scaling suggests growing displacements $|x_t - x_0| \sim t^{\alpha/2}$ and an **increasing** velocity $|v_t| \sim t^{\alpha/2}$ when compared to the deterministic setting. Due to this comparison, it is often assumed that superballistic scaling involves some kind of acceleration mechanism.

There are some problems with the interpretation of the EAMSD here:

The ensemble average in the EAMSD relies on sampling over the different initial configurations of the NK cells. The EAMSD behaves differently if the initial velocities

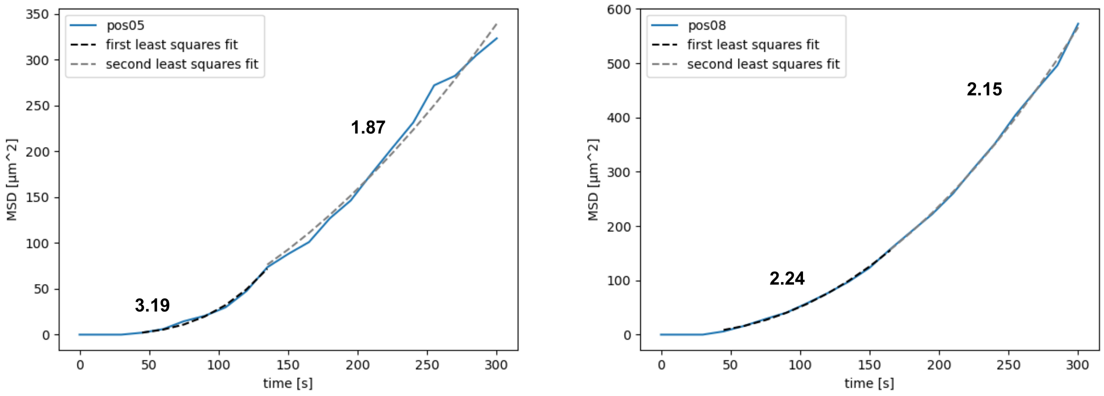

Figure 1.2: EAMSD for two sample arbitrary sample frames. The local power-law exponent is estimated using a least-square fit.

/ acceleration of NK cells vanish or are inherently finite. This may be the underlying reason, why the first least square fit leads to different scaling exponents in the superballistic regime. The interpretation of this α is therefore flawed.

Interpreting the EAMSD result is not always straightforward, especially when thinking about particle displacements. If the superballistic EAMSD scaling implies some underlying acceleration mechanism, then the differing scaling exponents between the frames suggest a stark inhomogeneity between the samples. Forming an average over wildly varying samples can be problematic in this case.

The second type of MSD, the **time - averaged MSD (TAMSD)**, overcomes this problem in parts. That is mainly because it is defined **pathwise**. If x is a given trajectory until the **observational time** T, then the TAMSD $\delta^2(\Delta, T)$ with **lagtime** Δ is defined as follows:

$$\delta^2(\Delta, T) = \frac{1}{T-\Delta}\int_0^{T-\Delta} (x_{t+\Delta} - x_t)^2 \, dt \tag{1.8}$$

Instead of averaging the displacement $x_\Delta - x_0$ over different particle samples, the TAMSD instead averages over the displacements of the trajectory $x_{t+\Delta} - x_t$ for all starting points $t = 0, ..., T - \Delta$. The TAMSD can be wildly varying between different samples, as it is itself a random object. But for many examples in nature or theoretical models, the TAMSD will have a similar scaling w.r.t. Δ in case of very large T. In case of Brownian motion it even holds that

$$\langle (x_\Delta - x_0)^2 \rangle = \lim_{T\to\infty} \delta^2(\Delta, T), \tag{1.9}$$

so the TAMSD can be used instead of the EAMSD.

This is a remarkable property of Brownian motion. Given that the TAMSD is computed by one (but long) trajectory, this equivalence allows to determine the properties of Brownian motion via a single trajectory instead of averaging over a large number of samples. Especially for experiments, where the sample size is often limited by the experimental setup and time of the Master's students, and computer simulations, where the sample size is limited by computational power, does this equivalence give a huge relief.

If the equivalence

$$\langle (x_\Delta - x_0)^2 \rangle = \lim_{T\to\infty} \delta^2(\Delta, T), \tag{1.10}$$

holds for a process x_t, it is historically called **(strongly) ergodic** (I condemn the convoluted usage of the word ergodicity in this context. and will use **large T convergence** instead).

Taking the TAMSD for the experimental data introduced a new problem, however: The superballistic scaling in the EAMSD **disappeared** in case of the TAMSD, the scaling became **ballistic**. This makes interpreting the TAMSD much more difficult, as only the EAMSD has a clear interpretation on the level of sample trajectories. While the superballistic EAMSD may stem from an acceleration mechanism of the NK cells themselves, this does not necessarily reflect in the TAMSD. If the acceleration is temporally confined to certain intervals, the effect of the acceleration on the TAMSD may be averaged out for large T (see [1] and [4] for a discussion on the TAMSD and EAMSD for anomalous diffusion).

The idea of both EAMSD and TAMSD is to estimate the *typical* displacement $x_t - x_s$ of a trajectory. Given the nature of the ensemble average, the EAMSD and its scaling exponent can give a very direct interpretation of this *typical* displacement. But for the TAMSD, this is not the case: The TAMSD averages $x_{t+\Delta} - x_t$ over $t = 0, ..., T - \Delta$ for one trajectory, so it is not possible to associate the behavior of $\delta^2(\Delta, T)$ with the behavior of some particular displacement. The qualitative and quantitative interpretation is not clear.

If a process is (strongly) ergodic / large T-convergent, this does not matter, since the TAMSD and EAMSD agree in the large T-limit. But many systems in biophysics are not ergodic. Since the TAMSD is quite easy to obtain in experiments, it is still widely used as a MSD quantifier. But there is a priori not direct connection with the EAMSD and EAMSD scaling. What does the TAMSD scaling measure, if it disagrees with the EAMSD? And why did it fail to reproduce the superballisticity of the EAMSD?

The overarching goal of this thesis has been to address these two question from a theoretical side.

1.4 Outline

The initial goal of this thesis has been to find stochastic models with superballistic scaling in **both** EAMSD and TAMSD, but which also model the run and tumble - like motion of the NK cells. After many different models I realized that this task may be impossible and went into a more theoretical direction, trying to analyze how the EAMSD and TAMSD scaling work qualitatively. Utilizing the underlying **velocity autocorrelation function (VACF)**

$$C_v(r, s) = \langle v_r v_s \rangle, \tag{1.11}$$

I have been able to make progress on different fronts and even gave partial answers as to why the (A)TAMSD scales so differently at times.

Ch. 2 introduces the necessary mathematical concepts from probability theory and stochastic calculus, which are needed for understanding the theoretical arguments later on. A focus is laid on exploring the EAMSD and TAMSD more conceptually and making the notion of **spreading** of stochastic processes more concrete.

Ch. 3 spends a considerable amount of pages on making the notion of power-law scaling $f(r) \sim r^{\beta}$ mathematically more precise. Since I will work mainly on theoretical grounds, I cannot heuristically define the scaling exponent using least square fits of MSD plots. Instead, different concepts and computational methods are introduced to give a clear definition of **scaling exponents**. I want to emphasize the concept of stable / transient scaling and in which way scaling exponents truly relate to a power-law nature of the underlying MSD.

Ch. 4 applies the introduced concepts to three classes of stochastic processes. These models can have a superballistic EAMSD scaling, but will show subballistic or ballistic (A)TAMSD scaling. A special emphasis is given on how to to compute the scaling exponents or estimate them heuristically.

Ch. 5 and Ch. 6 are two interlude chapters, which are used to introduce new theoretical results and methods. In Ch. 5, the **initial scaling exponent** $\alpha_f(0)$ of both EAMSD and (A)TAMSD is studied. Due to the equality with the lower index $\underline{\alpha}_f$, multiple theoretical estimates on $\alpha_f(0)$ will be analytically proven.

Ch. 6 studies the VACF $C_v(r, s) = \langle v_r v_s \rangle$ in more depth. The VACF can be used to compute the EAMSD and ATAMSD, it is therefore no surprise that the scaling of C_v itself dictates the EAMSD and ATAMSD scaling. I will show how the VACF relates to the higher derivatives of EAMSD and ATAMSD and provide a novel discretization scheme for the ATAMSD. The chapter concludes with a detailed study of stationary VACFs. Due to the direct relationship between the MSD and VACF for stationary VACFs, superballisticity can be ruled out for the initial and asymptotic regime. The only possibility for superballisticity in the stationary case are the multi-modal VACFs.

Ch. 7 studies the scaling a simple, but broad class of multi-modal VACFs. Although examples with a superballistic MSD are shown, these do constitute true VACFs due to a violation of the *positive-semidefiniteness* condition.

Ch. 8 is devowed to non-stationary VACFs. Since the EAMSD and ATAMSD do not coincide, the respective scaling behavior has to be studied separately. I will show that the ATAMSD scaling of a large class of non-stationary VACF is entirely dependent on the *stationary* part of VACF. The non-stationary part, which leads to superballisticity in the EAMSD case, influences the ATAMSD scaling w.r.t. the reduced observational time $T - \Delta$. The resulting ideas will be condensed into the **split scaling hypothesis**.

Since the theoretical part grew quite lengthy, owed to the large part of new results, I decided to exclude the simulated models entirely and only include computations for simple toy models. Still, the thesis became quite long in itself. I tried to make the chapters and sections inter - connected and weaving a red thread through the thesis, but even the interested reader may get occasionally lost in my Minoan labyrinth. This outline might help to understand how the current argument may fit in the overarching puzzle that is *scaling of stochastic processes*.

2 Stochastic primer: Basic stochastic calculus and spreading

Since probability theory / stochastic processes appear in almost every scientific discipline, many of the underlying concepts bear differing names, depending on the field in which it is introduced. This makes the conceptual field of stochastic motion / diffusion extremely vulnerable to confusing double - naming, which leaves any outsider in horror when comparing literature.

This chapter is meant to introduce the necessary conepts from probability theory.

2.1 Crash course on probability theory

This thesis relies heavily on probability theory on stochastic calculus. As these topics are not always discussed in depths during a standard physics programm, I decided to explain the basic notation and concepts for the reader. A more comprehensive, but mathematical reference can be found in [15].

Probability space

Anyone who studied the theory of Lebesgue measures knows the problem of defining *measure / area* for arbitrary, uncountable sets. In order to properly define the concept of probability, mathematicians use the concept of a probability space:

Definition 1. *A* ***probability space*** $W = (\Omega, \mathcal{A}, \mathbb{P})$ *is a tuple of three objects:*

1. *The underlying* ***sample space*** Ω
2. *The* ***event set*** $\mathcal{A} \subset \mathcal{P}(\Omega)$[1]
3. *The probability measure* $\mathbb{P} : \mathcal{A} \to [0, 1]$

In practice, only the probability measure $\mathbb{P}$ is relevant, while the sample space Ω and event set $\mathcal{A}$ are left implicit.

The following example illustrates the concept of probability spaces:

Assume you want to experiment with the statistics of a fair dice throw. Whatever you

---

[1] $\mathcal{A}$ consists of subsets of Ω, but has to be closed under certain set operations. The mathematical structure is called σ-**algebra** (see [9] for a more mathematical reference).

want to check, you will throw the dice exactly 10 times. These 10 trials (or **samples**) form the sample space:

$$\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Each number i corresponds to the unique i-th trial.

Since no trials are excluded, the event space consists of all subsets of Ω:

$$\mathcal{A} = \mathcal{P}(\Omega)$$

The probability measure $\mathbb{P} : \mathcal{A} \to [0, 1]$ assigns each trial a weight. Since the dice throws are assumed to be fair and unbiased, each trial is weighted the same. The probability of an individual trial i is therefore given by

$$\mathbb{P}(\{i\}) = \frac{1}{|\Omega|} = \frac{1}{10}.$$

The notion of probability space can be seen as the theoretical analogon of an experimental setup: It does not model the outcome of an experiment (e.g. the dice throw), but formalizes how the experiment can be conducted (e.g. exactly 10 fair dice are thrown and weighted equally).

The outcome of the experiment is modelled by a **random variable**:

Random variables

Definition 2. *Let $W = (\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A* ***(real valued) random variable*** *X is a measurable*[2] *function*

$$X : \Omega \to \mathbb{R}. \tag{2.1}$$

The probability space W models the experimental setup, but X models the concrete experiment:

Assume again the setup of the example, i.e. 10 fair dice throws. You conduct the experiment by simply throwing the dice 10 times, and for some miraculous reason you are able to throw 6 on each trial. If modeled as a random variable, this concrete experiment is given by

$$X(i) = 6 \tag{2.2}$$

for each $i \in \Omega$.

Each random variable has a **distribution / law**, which gives the probability of the outcome events:

Definition 3. *Let X be a real - valued random variable. The probability measure $\mathbb{P}_X$ defined by*

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) \tag{2.3}$$

for each open $B \subset \mathbb{R}$ is called the ***law / probability distribution*** *of X.*

If $\mathbb{P}_X$ is absolutely continuous w.r.t. the Lebesgue measure λ^1, the density function $p_X(x)$ is also called the ***distribution / density*** *of X.*

---

[2] X has to preserve the σ-algebra structure of $\mathcal{A}$.

If you were as lucky as in the example, the probability of throwing a 6 is 1. For each $B \subset \{1,2,3,4,5,6\}$ the law of X is given by

$$\mathbb{P}_X(B) = \begin{cases} 1, & \text{if } 6 \in B \\ 0, & \text{else} \end{cases}.$$

If you consider a generic result of this dice throw experiment, again modelled by a random variable X, the law of X is given by

$$\mathbb{P}_X(\{a\}) = \mathbb{P}(\{i \,:\, X(i) = a\}) = \frac{|\{i \,:\, X(i) = a\}|}{|\Omega|}$$

for each number $a \in \{1,2,3,4,5,6\}$. This is the fraction of throws i, for which a has been thrown.

In practice, the concrete probability space W and random variable X is not important, but rather the distribution $\mathbb{P}_X$.

Stochastic processes

Definition 4. *Let $T > 0$ be finite or $T = \infty$. Then a collection*

$$X = \{x_t \,:\, t \in [0,T]\} \tag{2.4}$$

of random variables x_t for each $t \in [0,T]$ is called a ***stochastic processes****. The process X itself will often be written just as x_t if no confusion arises.*

The index $t \in [0,T]$ is interpreted as time, so that x_t is the process X at time t. It is more reasonable to view X as a **random function / trajectory**. Stochastic processes are often written as $(x_t)_{t\in[0,T]}$ or simply x_t, if no confusion arises.

Whenever you will see the word stochastic process in this thesis, think of a random **trajectory**.

Moments of a random variable

From the applied perspective, only the distribution $\mathbb{P}_X$ of a random variable X is important. Most of the probabilistic quantities of interest[3] can be computed from $\mathbb{P}_X$ or the density $p_X(x)$ themselves.

But even if a density $p_X(x)$ exists, it can be too complex to be described analytically. It is therefore desirable to have a group of quantities, expressed using p_X, which can partially describe the behavior of the distributions. The simplies set of these are the **moments** of X:

Definition 5. *Let X be a random variable. The n-the moment μ_n of X is defined as the integral*

$$\mu_n = \mu_n^X = \int x^n \, d\mathbb{P}_X(x) = \int_{\mathbb{R}} x^n \rho_X(x) \, dx, \tag{2.5}$$

if it exists.

The first two moments are the most important: The **mean** and **second moment**.

---

[3] For example first-exit probabilities (see [5]).

Mean

Definition 6. *Let X be a random variable. The first moment*

$$\langle X\rangle = \mu_1 = \int_{\mathbb{R}} x\rho_X(x)\,dx \tag{2.6}$$

is called the ***mean*** *of X.*

The mean of X is usually denoted as $\langle X\rangle$ or $\mathbb{E}\{X\}$. Other words in the literature include **average** and **expectation value**.

If X is a fair dice throw, the probability for throwing number a is $\mathbb{P}(X = a) = 1/6$. The mean of X is therefore given by

$$\langle X\rangle = \sum_{a=1}^{6} a \cdot \mathbb{P}(X = a) = \frac{7}{2}.$$

I will not go into a deeper explanation of the mean, as the concept should be known.

Although the mean is a quite simple quantity, it is quite prominent due to its role in the **law of large numbers**:

Theorem 1. *Let $X_1, ..., X_n$ be a sequence of independent and identically distributed random variables.*[4] *If the identitcal mean $\langle X_i\rangle = \mu$ of the random variables is finite, the weighted sum*

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \tag{2.7}$$

converges to the average μ in probability.

The random variable $\overline{X}_n$ is often called the **empirical mean**. The law of large numbers assures that $\overline{X}_n$ will converge to $\langle X\rangle$ for large n, which means that $\overline{X}_n$ for finite n will have outcomes that become increasingly scattered around $\langle X\rangle$.

From an experimental perspective, the mean corresponds to the arithmetic mean of the outcomes if an experiment is repeated for sufficiently many trials. From a theoretical perspective, the mean can be seen as a value, around which the outcomes of $\overline{X}_n$ *scatter* or *concentrate*. This is due to the law of large numbers.

Although this interpretation is somewhat correct, it is dangerous to assume that also the outcomes of X itself scatter around $\langle X\rangle$. This is true for the Gaussian and other **super-exponential** distribution, since the probability distribution ρ_X tends to concentrate around the mean.

But for so called **sub-exponential** or **heavy-tailed**[5] distributions like the Cauchy distributions, this is not true. Either does the mean not exist due to integrability issues,

---

[4] X_i have the same probability distribution.

[5] The distribution of many *rare* events can be modeled using such distributions. Examples include the Pareto distribution for net income, strength of earthquakes on the Richter scale or insurance claims (see [20] for an introduction to extreme value theory).

or most of the mass of ρ_X is concentrated away from $\langle X \rangle$. Treating the mean $\langle X \rangle$ as a *typical value* of X in this case can lead to absurdly ridiculous conclusions.[6]

Aside of knowing the mean of a distribution, it is similarly ideal to measure how stark the outcomes of X deviate from the mean $\langle X \rangle$. This is done by the **variance**:

Variance

Definition 7. *Let X be a random variable with finite mean $\langle X \rangle$. The* ***variance*** *$Var(X)$ of X is defined by*

$$Var(X) := \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2 = \mu_2 - \mu_1^2. \tag{2.8}$$

It is quite simple to show that a vanishing variance $\mathrm{Var}(X) = 0$ implies $X \equiv \langle X \rangle$, i.e. the random variable X has always the mean $\langle X \rangle$ as outcome.

If $\mathrm{Var}(X) > 0$, the random variable X will have outcomes that deviate from the mean and the distribution is not concentrated at just $\langle X \rangle$. It is therefore the intuition that a larger variance implies a stronger deviance of the distribution away from the mean. The variance can be seen as a measure of how strongly the outcomes of X **spread** away from $\langle X \rangle$. Hence, the variance is normally interpreted as a quantifier of **spreading.**

Similar to the mean, the value of the variance itself cannot be compared between families of probability distributions. It only makes sense to compare results within a common class of distribution, e.g. within the class of Gaussian distributions.

2.2 Brownian motion and the Ornstein-Uhlenbeck process

Stochastic processes naturally appear in the context of **stochastic differential equations**. One simple, but quite fruitful example of these is the physical **Langevin equation** from the introduction:

$$m\dot{v}_t = -\lambda v_t + \eta_t, \tag{2.9}$$

v_t corresponds to the velocity process of a small particle inside some engulfing fluid at thermal equilibrium, with which it interacts via collision and friction. The first term $-\lambda v_t$ is the simple friction force with **friction / damping** coefficient λ and m is the mass of the particle. The external force η_t is called **white noise** and represents the forces on v induced by pairwise collisions with the surrounding gas molecules, which is a **stochastic force**. Due to the stochasticity of η_t, the velocity v_t itself will be a stochastic process.

From the isotropy of the surrounding gas molecules it should be clear that the average vanishes $\langle \eta_t \rangle = 0$ and the **(auto)correlation**

$$\langle \eta_t \eta_{t'} \rangle = D\delta(t - t') \tag{2.10}$$

---

[6] GDP per capita or net income per capita, a form of mean, is not a good quantifier of national prosperity on its own. In Germany in 2022, the richest 10% had 37.64 % of the income share before taxes, while the poorest 50% only had a whooping 18.78%. See [21] for the data.

is mutually singular. In physics, one chooses $D = 2\lambda k_B T$ for the diffusion coefficient, with k_B being the Boltzmann constant and T the gas temperature. Due to the more mathematical considerations in this thesis, I will work using a resized Langevin equation

$$\dot{v}_t = -\lambda v_t + \sqrt{2}\eta_t \tag{2.11}$$

with pairwise correlation

$$\langle \eta_t \eta_{t'} \rangle = \delta(t - t'). \tag{2.12}$$

The Langevin equation has one fatal flaw, which makes a rigorous treatment of this whole topic quite mathematical: The force η_t cannot be a function pathwise due to the mutually singular correlation. The delta distribution $\delta(t - t')$ in the pairwise correlation implies that η_t is a **Schwartz distribution** and not a genuine function at all. A priori, v_t should not be a function either, which is quite bad.

In the mathematical literature , the Langevin equation in eq. 2.11 is interpreted in a different way:

Every ODE can be turned equivalently into an integral equation. For the Langevin equation, the corresponding integral equation would look like

$$v_t = v_0 - \lambda \int_0^t v_s \, ds + \sqrt{2} \int_0^t \eta_s \, ds.$$

By defining $B_t := \int_0^t \eta_t \, dt$, the integral equation circumvents the white noise in favor of its integral. The correlation of B_t

$$\begin{aligned} \langle B_s B_t \rangle &= \left\langle \int_0^s \eta_{s'} \, ds' \int_0^t \eta_{t'} \, dt' \right\rangle = \int_0^t \int_0^s \langle \eta_{s'} \eta_{t'} \rangle ds' \, dt' \\ &= \int_0^t \int_0^s \delta(s' - t') \, ds' \, dt' = \min(s, t) \end{aligned}$$

is now a genuine function, so it is possible to give v_t a well-defined meaning if B_t is well-defined.

The integrated white noise process $B_t = \int_0^t \eta_s \, ds$ is historically called **Brownian motion** in the physics and **Wiener process** in the mathematics literature. Showing that B_t is well - defined is a little more cumbersome mathematically, so instead I will just give the modern definition of B_t:

Definition 8. *Let $B_t = (B_t)_{t \in [0,T]}$ be a stochastic process. Then B_t is called a **Browian motion** if*

1. $B_0 = 0$,
2. *The trajectories B_t are **continuous**,*
3. *The increments $B_t - B_s$ and $B_r - B_q$ are **independent** for $s \geq r$ and identically distribute,*
4. *The increment $B_t - B_s$ is **Gaussian** distributed with standard deviation $\sigma = \sqrt{t - s}$.*

*The trajectories of B_t have **Hölder regularity** γ for every $\gamma < \frac{1}{2}$ and are almost surely nowhere differentiable.*

There are multiple definitions of Brownian motion (and white noise), which all differ only by the normalization of the autocorrelation. The normalization condition I use disagrees with most of the physical papers. Conceptually, this differs from the physical convention in that I need to write the diffusion constant D externally into the evolution equation, instead of integrating it into the autocorrelation of the noise.

The modified integral Langevin equation

$$v_t = v_s - \lambda \int_s^t v_r \, dr + \sqrt{2} B_t \tag{2.13}$$

is called a **stochastic differential equation (SDE)** and written as

$$dv_t = -\lambda v_t \, dt + \sqrt{2} dB_t \tag{2.14}$$

in short form. The solution v_t is called **Ornstein-Uhlenbeck process**.

It is important for the short term behavior of v_t how the initial velocity v_0 is distributed. For most velocity distributions, the asymptotic probability distribution v_t will be the same, namely the **stationary** distribution π.

If $v_0 \sim \pi$, one can show $v_t \sim \pi$ for all future times $t > 0$. In this case the velocity distribution is **stationary** and the variance $\langle v_t^2 \rangle = 1$ remains constant.

2.3 Stochastic motion and spreading

The variance $\mathrm{Var}(X)$ of a random variable X can be used a quantifier on how much the outcomes of X deviate from the mean $\langle X \rangle$. This makes it a very simple, but quite effective measure on how much the outcomes of X **spread** away from $\langle X \rangle$. In case of stochastic processes, the notion of spreading is more ambiguous:

Although one can consider the **mean function** $\mu_t = \langle x_t \rangle$ of the stochastic process, many scientists like to work with the **centered** process $y_t = x_t - \mu_t$, since a bias in the measurement can lead to a spurious drift in μ_t. Naive quantities like the **second moment function** $SMF_t = \langle x_t^2 \rangle$ or **variance function** $\mathrm{Var}_t = \mathrm{Var}(x_t)$ do measure the spreading of x_t, but in terms of deviations from the mean function μ_t. This may be interesting in its own right, but in many cases this is the wrong kind of *spreading*:

The study of motility / movement of stochastic process in physics is more concerned with describing the distribution / behavior of the displacement from the starting position, $d_t = x_t - x_0$. The squared displacement $(x_t - x_0)^2$ quantifies how far x_t strays away from the origin x_0 and it is therefore more physical to study the distribution of the squared displacement as a way of describing the (spatial) spreading of x_t.

This dilemma of using x_t or $x_t - x_0$ as the stochastic object leads to the usage of **four** distinct quantities:

Second Moment Function (SMF)

Definition 9. *The* ***second moment function (SMF)*** *SMF_t is defined as*

$$SMF_t := \langle x_t^2 \rangle.$$

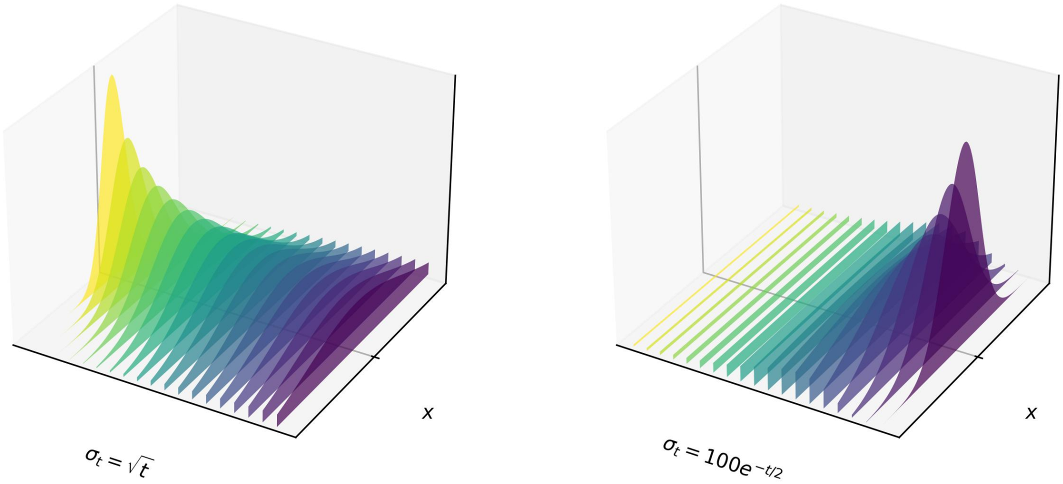

Figure 2.1: Probability distribution of $x_t \sim \mathcal{N}(0, \sigma_t)$ for $\sigma_t = \sqrt{t}$ and $\sigma_t = 100e^{-t/2}$ for $t \in [1, 20]$ with spacing $\Delta t = 0.5$.

It relates to the ***variance function (VF)*** *via*

$$Var(t) = \langle (x_t - \langle x_t \rangle)^2 \rangle = SMF_t - \mu_t^2. \tag{2.15}$$

Let x_t be a Gaussian process with $x_t \sim \mathcal{N}(0, \sigma_t)$ for each t. If $\sigma_t := e^{-\frac{t}{2}}$ is an exponentially decaying function, the SMF of the process will be

$$SMF_t = \langle x_t^2 \rangle = e^{-t} \tag{2.16}$$

a decaying exponential as well. What does SMF_t tell about the trajectories of x_t?

Since SMF_{t_0} gives a quantifier on how broadened the probability distribution of x_{t_0} really is, the exponentially decaying behavior implies that over time the probability distribution of x_t *shrinks*, i.e. contracts around the mean. But if the probability distribution starts to concentrate more and more around the mean of x_t, this implies that the process itself tends to concentrate / cluster around its mean position.

Brownian motion B_t shows the opposite behavior, since $B_t \sim \mathcal{N}(0, \sigma_t)$ with $\sigma_t = \sqrt{D \cdot t}$. The standard deviation increases as a square root $t^{\frac{1}{2}}$, so the SMF is given by

$$SMF_t = \langle B_t^2 \rangle = Dt, \tag{2.17}$$

which is linearly increasing. Since the SMF / standard deviation is temporally increasing, the probability distribution of B_t spreads out over time. This implies that B_t somehow diffuses further and further away over time from its mean, i.e. B_t **spreads out**.

The evolution of the distribution of x_t for both cases is shown in Fig. 2.1.

The SMF was used in both examples to quantify the spatial spreading (or shrinking) of the process trajectories. This interpretation was only valid because both processes were centered ($\langle x_t \rangle = \mu_t = 0$). A more direct notion version of (spatial) spreading can be acchieved when considering the displacement $x_t - x_0$ or squared displacement $(x_t - x_0)^2$ instead. This leads to the **mean squared displacement (MSD)** quantities:

Ensemble-Averaged MSD (EAMSD)

The most direct way to average over the squared displacement $d_t^2 = (x_t - x_0)^2$ is to consider the second moment of d_t:

Definition 10. *The* ***ensemble-averaged mean square displacement (EAMSD)*** *is defined as*

$$EA_t := \langle (x_t - x_0)^2 \rangle. \tag{2.18}$$

The EAMSD is connected to the SMF of x_t via the formula

$$\begin{aligned} EA_t &= \langle (x_t - x_0)^2 \rangle = \langle x_t^2 \rangle + \langle x_0^2 \rangle - 2\langle x_t x_0 \rangle \\ &= SMF_t + SMF_0 - 2C_x(t, 0), \end{aligned}$$

where $C_x(t, 0) = \langle x_t x_0 \rangle$ is the ***autocorrelation function*** *of x_t.*

The interpretation of the EAMSD is best explained if x_t is centered, i.e. $\langle x_t \rangle = 0$. In this case the average of the displacement d_t vanishes due to $\langle d_t \rangle = \langle x_t \rangle - \langle x_0 \rangle = 0$ and the EAMSD becomes variance of d_t. If the EAMSD grows, the variance of d_t increases. This can be interpreted as having a broadening in the distribution of d_t. If the distribution of d_t broadens, the it becomes more likely that x_t deviates stronger from the initial x_0. This suggests spatial spreading. The same logic also applies if the EAMSD is increasing, which suggest spatial contraction.

The EAMSD is the *gold standard* of spreading measurement for stochastic processes.

Time-Averaged MSD (TAMSD)

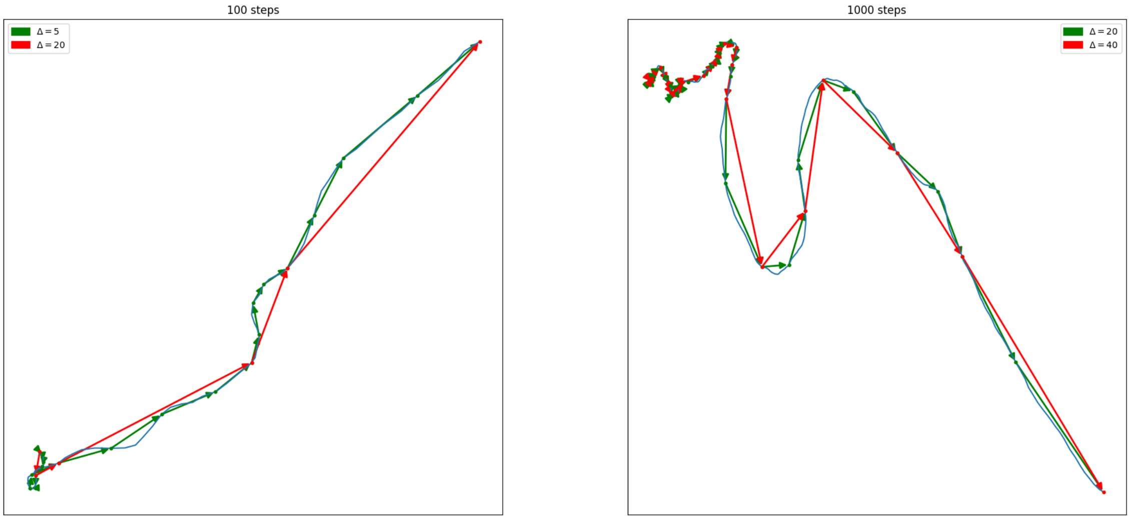

Figure 2.2: Plot of two sample trajectories with $T = 100$ and $T = 1000$. The green and red arrows show the lagtime increment $d_{\Delta,t} = x_{t+\Delta} - x_t$ for two different choices of lagtime Δ. With increasing lagtime Δ, the lagtime increments $d_{\Delta,t}$ tend to deviate stronger from the initial trajectory x_t.

The EAMSD is a good quantifier for spatial spreading and analytically tractable. But what about the experimentalists?

The EAMSD is defined via a stochastic / ensemble mean, so estimating the EAMSD and SMF in experiments requires a fairly large amount of samples, which need to be identically prepared. If the EAMSD is additionally increasing, the number of samples required for a given threshold of precision has to increase as well. This leads to a collapse of accuracy for long - time data, if the sample size is chosen to small. But one cannot easily increase the sample size in complex experiments: Either each sample / experimental trial takes a fairly long amount of time, which pushes the costs of additional experimental trials, or it is not possible to accurately reproduce the setup for many consequent trials.

The experiments on NK cell motility in the introduction (see [24]) required the tracking of NK cells in a collagen matrix. The collagen matrix can be damaged after many measurements, which can lead to an alteration in the motility pattern, or the NK cells themselves decay and the experimental setup needs to be renewed over time. Another drawback of the EAMSD is the distribution of the initial condition x_0:

Since the displacement distribution $d_t = x_t - x_0$ is highly sensitive to the initial condition, the EAMSD behavior changes drastically. Going back to the Ornstein-Uhlenbeck process v_t, which solved eq. 2.14

$$dv_t = -\lambda v_t\, dt + \sqrt{2} dB_t,$$

the EAMSD and SMF of v_t behave quite differently, depending on whether one starts with resting velocities $v_0 = 0$ are stationary distributed velocities $v_0 \sim \pi$. The difference in EAMSD scaling will be discussed in Sec. 4.3.2 and Sec. 4.3.3, which shows a great difference in the initial behavior of the EAMSD.

The EAMSD averages the displacement $d_t = x_t - x_0$ over different samples of the process. The different EAMSD scaling for the OU - process comes from the wildly different behavior of the initial distribution, because the initial position x_0 is different between the trajectories. It can also be a bad tactic to average the displacement over different samples of the process, if they dynamic itself varies between the trajectories themselves:

Take the example of

$$x_t = (t+1)^b,$$

where the exponent b is uniformly distributed between 1 and 2. The EAMSD of the process scales roughly as

$$EA_t \sim \frac{(t+1)^4}{2\log(1+t)}$$

for large t. This would suggest that the typical displacement $d_t = x_t - x_0$ behaves as

$$d_t \sim t^2,$$

although for each trajectory with exponent b, the displacement behaves as

$$d_t \sim t^b.$$

In this case it would be a wise choice to not average the displacement d_t over different samples, but rather a single trajectory. One possibility is to take the lagged displacements $x_{t+\Delta} - x_t$ of one fixed trajectory and average over the starting point t. This leads to the **time-averaged mean square displacement (TAMSD)**:

Definition 11. *Let x_t be a stochastic process. For each trajectory of x_t, the* ***time-averaged mean square displacement (TAMSD)*** *for* ***lagtime*** *Δ and observational time $T > \Delta$ is defined as*

$$\delta^2(\Delta, T) = \frac{1}{T-\Delta} \int_0^{T-\Delta} \left(x_{t+\Delta} - x_t\right)^2 \, dt. \tag{2.19}$$

Since x_t is a stochastic process, the TAMSD $\delta^2(\Delta, T)$ is itself ***random****.*

One of the surprising things about the TAMSD might be that it can often **agree** with the EAMSD, even though it is defined pathwise and therefore inherently random:

Take the example of Brownian motion $x_t = B_t$. The average of the TAMSD

$$\langle \delta^2(\Delta, T) \rangle = \frac{1}{T-\Delta} \int_0^{T-\Delta} \left\langle \left(x_{t+\Delta} - x_t\right)^2 \right\rangle \, dt \tag{2.20}$$

can be computed analytically, since the increments $x_{t+\Delta} - x_t \sim \mathcal{N}\left(0, \sqrt{t}\right)$ are Gaussian distributed. In this case the average of the TAMSD agrees with the EAMSD:

$$\begin{aligned} \langle \delta^2(\Delta, T) \rangle &= \frac{1}{T-\Delta} \int_0^{T-\Delta} \left\langle \left(B_{t+\Delta} - B_t\right)^2 \right\rangle \, dt \\ &= \frac{1}{T-\Delta} \int_0^{T-\Delta} \Delta \, dt = \Delta = EA_\Delta \end{aligned}$$

The averaged TAMSD agrees with the EAMSD of Brownian motion. This is not that interesting in itself, but it is possible to make a similar statement about the TAMSD itself:

One can show that the variance of $\delta^2(\Delta, T)$ decays as

$$\mathrm{Var}(\langle \delta^2(\Delta, T) \rangle) \sim \frac{\Delta^n}{T}$$

and **vanishes** in the large T limit:

$$\lim_{T\to\infty} \mathrm{Var}(\langle \delta^2(\Delta, T) \rangle) = 0 \tag{2.21}$$

This implies the pathwise convergence

$$\lim_{T\to\infty} \delta^2(\Delta, T) = \Delta = EA_\Delta. \tag{2.22}$$

The pathwise convergence is important, because it implies that the EAMSD of Brownian motion can be estimated by observing a single, but very long trajectory.

This phenomenon of convergence

$$\delta^2(\Delta, T) \overset{T\to\infty}{\to} EA_\Delta$$

is sometimes called **ergodicity** in the literature. I do not like this wording, as confusion may arise with ergodicity in the sense of classical thermodynamics or Markov processes. It may be a better choice to denote his phenomenon as **large T convergence**.

If large T-convergence holds, the distinction between EAMSD and TAMSD is ignored and the shared quantity is simply called **mean squared displacement (MSD)** MSD_t.

For a fair share of classical processes (e.g. processes with stationary velocity, see Sec. 6.5), the large T convergence works and the EAMSD can be reconstructed from a lucky trajectory. The problem becomes more pressing if the T convergence fails. In many cases the convergence cannot simply work due to the fact that the TAMSD $\delta^2(\Delta, T)$ retains an **explicit** T - dependence. Large T convergence cannot happen there, because the pathwise limit either diverges (see Sec. 4.1.1) or vanishes (see Sec. 8.6).

The fact that large T convergence happens seems to be more surprising, given how the **lagtime scaling** (scaling in Δ) of the TAMSD is interpreted:

The EAMSD $EA_t = \langle (x_t - x_0)^2 \rangle$ averages the displacement $d_t = x_t - x_0$ over a duration t from the starting point over multiple sample trajectories. The TAMSD averages over the lagtime increments

$$d_{\Delta,t} = x_{t+\Delta} - x_t$$

of a fixed trajectory instead. The outer integral in the TAMSD averages $d_{\Delta,t}$ over t from 0 to $T - \Delta$ and the formula

$$\frac{1}{T-\Delta} \int_0^{T-\Delta} (x_{t+\Delta} - x_t)^2 \, dt$$

is a weighted integral over the squared length of $d_{\Delta,t}$. Colloquially, it averages[7] over the squared length of the displacement vector $x_{t+\Delta} - x_t$.

I have included a visual representation of this in Fig. 2.2, where two sample trajectories and the lagtime increment $d_{\Delta,t}$ for two different choices of Δ are shown as green and red arrows.

For very small Δ, the displacement vectors seem match the underlying trajectory. But if Δ becomes progressively larger, the displacements vector tend to deviate more from the underlying trajectory. This occurs naturally if the trajectory **bends** over the interval Δ. You can see that the underlying trajectory performs its motion on very different timescales, as the dense agglomeration of arrows in the bottom left and top left corners transition into more isolated displacement vectors. As Δ remains constant for the red and green arrows respectively, a sparser accumulation of displacement vectors implies a faster progressing trajectory. Depending on how fast the trajectory progress and how strong it bends, the displacement vectors will deviate more from the trajectory for higher Δ values, which changes the value of $\delta^2(\Delta, T)$. The TAMSD tends to monotonically grow, but exactly how strong it grows depends on these underlying properties of the displacement.

The TAMSD still measures some sort of spreading of the trajectory displacements, but averaged over all possible displacements. The TAMSD around Δ gives information on how the lagtime increments $d_{\Delta,t}$ behave **on average**, but no direct conclusion on

---

[7] Average in the sense of the integral, i.e. a time integral average, not a stochastic average.

the initial displacement $x_\Delta - x_0$ can be drawn. This makes it conceptually difficult to relate the TAMSD and TAMSD scaling to the EAMSD, if large T convergence fails.

If large T convergence fails and δ^2 contains an explicit T-dependence, δ^2 depends on Δ and T simultaneously. The TAMSD scales not only in the lagtime Δ, but also in the observational time T. In the literature the lagtime Δ is often used for quantifying the spreading and **scaling** of the TAMSD, due to its agreement with the EAMSD scaling in case of large T convergence.

I will show in Ch. 8 for a model class of processes, that both the lagtime and observational time scaling can reveal information on the process. This will be condensed into the **split scaling hypothesis** in Sec. 8.4.

Averaged Time-Averaged MSD (ATAMSD)

The TAMSD is stochastic in nature, since it is computed pathwise. If analytic arguments are desired, the stochastic mean of the TAMSD is used instead:

Definition 12. *The* ***averaged time-averaged mean square displacement (ATAMSD)*** *is defined as the* ***(stochastic) mean*** *of the TAMSD:*

$$\langle \delta^2(\Delta, T) \rangle = \frac{1}{T - \Delta} \int_0^{T-\Delta} \left\langle (x_{t+\Delta} - x_t)^2 \right\rangle \, dt \tag{2.23}$$

It is often expected that the TAMSD converges to the ATAMSD for large T, namely

$$\delta^2(\Delta, T) \overset{T\to\infty}{\to} \langle \delta^2(\Delta, T) \rangle.$$

The mean

$$\left\langle (x_{t+\Delta} - x_t)^2 \right\rangle, \tag{2.24}$$

is called the **averaged lagtime increment (ALI)**. It is intimately related to the **velocity autocorrelation function (VACF)**

$$C_v(r, s) = \langle v_r v_s \rangle \tag{2.25}$$

in case of differentiable processes. This allows to estimate the scaling on the ATAMSD with the VACF.

3 Gambler's ruin: How random motion spreads and what scaling means

The previous chapter introduced the EAMSD and (A)TAMSD of a stochastic process as a way to measure the **spreading** of a process. In many cases the EAMSD and TAMSD **scale** approximately as a power - law

This chapter defines what an approximate **(power - law) scaling** means and how to describe it quantitatively. The previous chapter introduced the EAMSD and(A)TAMSD as quantities to measure the **spreading** of a stochastic process.

It is a heuristic observation, that both MSDs tend to grow as a power-law. In the case of the EAMSD, this often written as

$$\langle (x_t - x_0)^2 \rangle \sim t^{\alpha}.$$

This type of growth is called **power-law scaling**.

But power-law scaling is often only stated as a heuristic fact, and there is some ambiguity as to how one can deduce such a scaling and the corresponding **scaling exponent** β. This chapter gives a mathematically rigorous exposition to the concept of power-law scaling and how this relates to the spreading of stochastic processes.

3.1 Normal and anomalous diffusion

It is cumbersome to repeat the scaling concepts for each type of MSD, since the mathematical concept remains the same for all cases. For this reason, I decided to explain the scaling-related topics using the notion of **growth functions**:

Definition 13. *Let $f : [0, T] \to \mathbb{R}$ be a function. If $f(0) = 0$ and $f(r) \geq 0$ for any $r \geq 0$, f will be called a growth function.*

Aside of the SMF,[1] the EAMSD, TAMSD and ATAMSD all are growth functions. Since the symbol for time differs between the EAMSD (t) and the (A)TAMSD (Δ and T), I decided to use the neutral r and sometimes s for denoting time.

Normal diffusion

The first example of stochastic process has been Brownian motion B_t in Sec. 2.2. The EAMSD of B_t scales linearly

$$EA_t = t, \tag{3.1}$$

---

[1] For the discussion of scaling exponents, it does not matter if $f(0) \neq 0$. Most of the theorems are still valid for the SMF.

and due to large T-convergence, the same holds for the SMF, TAMSD and ATAMSD.

Going back to the growth function picture, processes x_t with

$$f(r) = D \cdot r \tag{3.2}$$

are said to **scale diffusively** in f.

The case of linear scaling (e.g. scaling exponent $\alpha = 1$) is special, since many classical *thermal* processes have an EAMSD that scales asymptotically linear. This type of process is called **normal diffusion**.

For normal diffusion, the standard technique to compare the *speed* of spreading between different processes by comparing the diffusion constant D.

Anomalous diffusion

Things become more interesting if the linear scaling is violated:

Definition 14. *The process x_t is said to be an* ***anomalous diffusion / scale anomalously*** *for a growth function f if the linear scaling is violated:*

$$f(r) \neq C \cdot r$$

Since f does not scale linear anymore, there is no direct way to define a diffusion constant and comparing the *speed of spreading* between processes becomes ambiguous.

One possible generalization for this kind of linear scaling would be the concept of **strict polynomial growth**:

Definition 15. *Let x_t be a process with growth function f. The process is said to scale* ***strictly polynomial*** *if*

$$f(r) = Cr^{\alpha} \tag{3.3}$$

for some ***scaling exponent*** *α.*

The behavior of the process is highly dependent on the scaling exponent α. Since different values of α lead to wildly different motility patterns, the ranges of α have been divided into five castes:

1. **Subdiffusivity** for $0 < \alpha < 1$:
   Superdiffusive processes appear naturally in the setting of stochastic motion within a confining / trap potential. Notable examples are Brownian motion in an external field / optical cavity, anti-persistent Brownian motion B_H with Hurst parameter $H < 1/2$ or the generic class of **fractional diffusions** (see [17]).

2. **Diffusivity** for $\alpha = 1$:
   Diffusive processes share many interesting properties with Brownian motion like the existence of a time-independent diffusivity constant. Notable examples are stochastic processes with an integrable VACF (see 6.5.3), Brownian motion and the Ornstein-Uhlenbeck process for large times.

3. **Superdiffusivity** for $1 < \alpha < 2$:
   Superdiffusive processes tend to have an angular decorrelation together with an acceleration mechanism. These two competing forces are somewhat balanced for superdiffusive processes, as those are slower than ballistic processes. Notable examples are Brownian motion in an repulsing / accelerating potential, persistent Brownian motion B_H with Hurst parameter $H > 1/2$ and Levy walks (see [28]).

4. **Ballisticity** for $\alpha = 2$:
   Ballistic processes are similar to superdiffusive processes, but for these the angular decorrelation and acceleration mechanism balance out. The corresponding spreading behaves similar to those of trajectories with constant velocity. One example I use in this thesis is the constant walker $x_t = v_0 t e_\theta$ presented in Sec. 4.1.1.

5. **Superballisticity** for $\alpha > 2$:
   Superballistic processes have an overwhelming acceleration mechanism. Examples in this thesis are the accelerated walker $x_t = t^2 e_\theta$ in Sec. 4.1.1 or the integrated Brownian motion[2] in Sec. 4.3.1.

A more complete list of example processes for each category can be found in [18].

I explained in the introduction that the goal of this thesis is to study the superballistic scaling. There are few detailed studies on the matter of superballistic scaling itself, which necessitates the more detailed analysis in this thesis.

The matter of strict scaling

The problem with the case of polynomial scaling is that most natural processes do not scale strictly as a power-law. From the examples I mentioned, only the accelerated walker $x_t = t^2 e_\theta$ and Brownian motion scale strictly.

For most natural processes, the power-law scaling is only local in nature and can change over time. This is illustrated best by the **integrated (equilibriated) OU process** x_t from Sec. 4.3.2:

The process x_t is defined as $x_t = \int_0^t v_t \, dt$, where the velocity v_t solves

$$dv_t = -v_t \, dt + \sqrt{2} \, dW_t \tag{3.4}$$

with the initial velocity $v_0 \sim \pi$ being stationary distributed. Since v_t is stationary, the process x_t is large T convergent and both EAMSD and (A)TAMSD agree.

The MSD of x_t can be computed as

$$MSD_t = 2\left(t + (e^{-t} - 1)\right), \tag{3.5}$$

which is clearly not a strict power-law. But for very small ($t \ll 1$) and very large ($t \gg 1$) times, the MSD has a **local** power-law nature:

---

[2] This is sometimes called the **random acceleration process**.

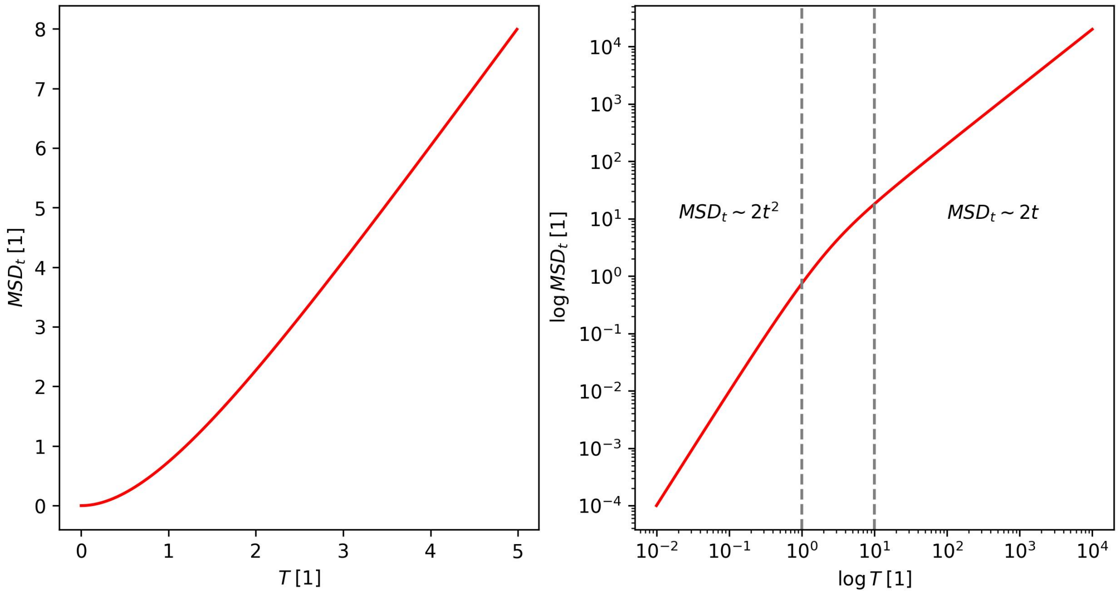

Figure 3.1: MSD of integrated OUP x_t as normal plot (left) and log-log plot (right).

A strictly polynomial scaling growth function $f(r) = C \cdot r^\alpha$ looks **linear** in the log-log plot

$$\log f(r) = \alpha \log r + \log C \tag{3.6}$$

and the scaling exponent appears as the **slope** of this linear function.[3] Fig. 3.1 shows the MSD in linear and log-log scaling. It is apparent, that the log-log plot is linear for the small time (**initial**) and and large time (**asymptotic**) regime. Estimating the slope in the log-log plot reveals, that the process scales quadratically (ballistic) at the beginning and linearly (diffusive) in the end.

Fitting to the log-log plot is numerically and experimentally very well justified, but from the prospects of theorists, it is desirable to compute the scaling exponents analytically. But can one define the slope of the plot *locally*?

3.2 Estimating the power-law exponent

Growth functions with strict polynomial scaling appear linear in the log-log plot. The offset $\log C$ is given by the normalization of the growth function, so after renormalization into

$$g(r) = \frac{f(r)}{C} = r^\alpha,$$

the log-log relation for g

$$\log g(r) = \alpha \log r$$

---

[3] Estimating this slope in the log-log plot is how one estimates the scaling exponent in experimental data.

becomes direct and the slope / scaling exponent can be computed as

$$\frac{\log g(r)}{\log r} = \alpha.$$

Since this quotient only involves g and r itself, this quantity can be defined for generic growth functions:

$$\lambda_r := \frac{\log f(r)}{\log r} \tag{3.7}$$

The index λ_r is called **crude scaling exponent**.[4]

If the growth function $f(r) = Cr^\alpha$ is not normalized, the crude scaling exponent can asymptotically recover α by means of the limit

$$\lim_{r\to\pm\infty} \lambda_r = \alpha,$$

but there is a divergence $\lim\limits_{r\to\pm 1} |\lambda_r| = \infty$ at $r = 1$.

Can the crude scaling exponent help in describing the local power-law scaling of the MSD of the integrated OU process x_t? For x_t, the crude scaling exponent is given by

$$\lambda_t = \frac{\log MSD_t}{\log t} = \frac{\log 2\,(t + (e^{-t} - 1))}{\log t} = 1 + \frac{\log\left(1 + \frac{e^{-t}-1}{t}\right)}{\log t} + \frac{\log 2}{\log t}$$

For $t \to 0$ the crude scaling exponent becomes $\lambda_0 = 2$, whereas for $t \to \infty$ it becomes $\lambda_\infty = 1$. The crude scaling exponent is therefore able to reproduce the initial and asymptotic slopes of the log-log plot in the limit. But there are issues with λ_r:

If the growth function vanishes $f(r) = 0$ for $r \neq 0$ or if $f(r) \neq 1$ for $r = 1$, the crude scaling exponent diverges $\lambda_r \to \pm\infty$. This is not a problem per se, as λ_r can still be used for estimating the *scaling behavior* at $r = 0$ or $r \to \infty$. But for processes more complex than the OUP, there can be a non-trivial scaling in between, which cannot be accurately quantified using λ_r, especially around the singularity at $r = 1$. Even if there is no intermediary divergence, the power-law exponents are only recovered in the limit $r \to \pm\infty$. If this limit converges poorly, the estimation can give a wrong results.

Due to the divergence properties of λ_r, the crude scaling exponent is seldom used (aside of analytic proofs) in favor of the **dynamical scaling exponent**:

The dynamical scaling exponent α_f

Definition 16. *Let $f(r)$ be a growth function, which is $\mathcal{C}^1$. The* ***(dynamical) scaling exponent*** *α_f is defined as the logarithmic derivative*

$$\alpha_f(r) := \frac{d\ \log f(r)}{d\ \log r} = r\partial_r \log f(r) = \frac{r}{f(r)}\partial_r f(r). \tag{3.8}$$

If the growth function f in α_f is clear from the context, $\alpha_f(r)$ will be written as α or $\alpha(r)$.

---

[4] At least by me.

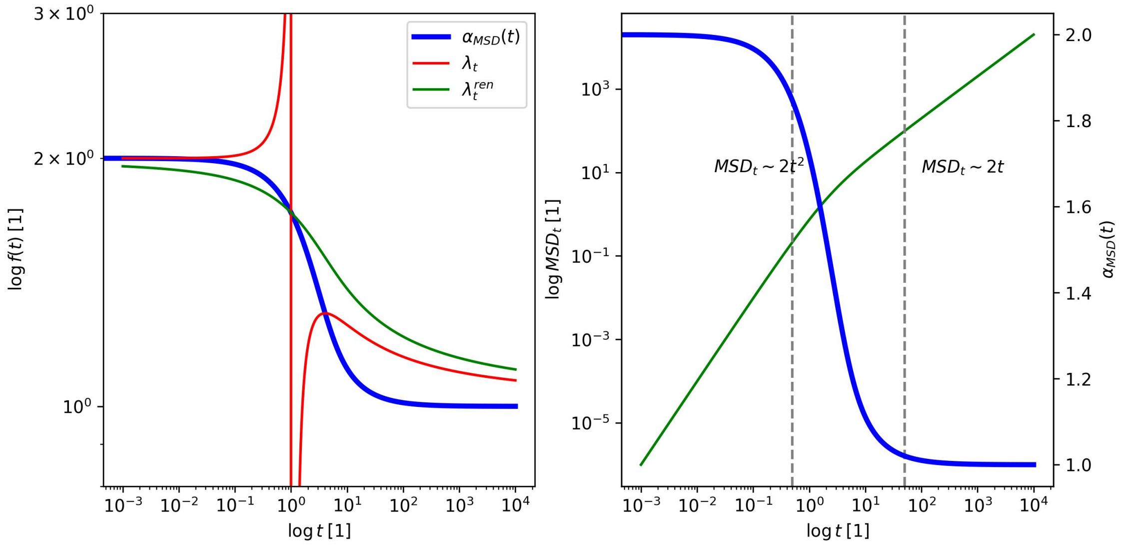

Figure 3.2: (Left) Comparison of the dynamical scaling exponent $\alpha_{MSD}(t)$, crude scaling exponent λ_t and renormalized crude scaling exponent λ_t^{ren} for the integrated OUP. For λ_t^{ren} the MSD has been renormalized such that $MSD_1^{ren} = 1$. (Right) MSD_t of the integrated OUP in log - log scaling and the dynamical scaling exponent $\alpha_{MSD}(t)$ in log time.

First of all, the dynamical scaling exponent can reproduce the power-law exponent in $f(r) = Cr^\alpha$:

$$\alpha_f = \frac{r}{Cr^\alpha} \alpha C r^{\alpha-1} = \alpha$$

But in comparison to the crude scaling exponent does α_f reproduce the exponent α at every time r, not only for the asymptotic $r \to \pm\infty$. The dynamical scaling exponent does not have a singularity at $r = 1$ in general, compared to λ_r.

The dynamical scaling exponent is well - defined for any finite r and the value of $\alpha_f(r)$ can be related to some power - law scaling of the growth function under suitable conditions. A comparison between λ_r and $\alpha_f(r)$ for the integrated OUP is included in Fig. 3.2. Although α_f is well-defined for every r, it does not mean that $\alpha_f(r)$ represents any factual local power-law exponent.

If the dynamical scaling exponent changes too much, then any interpretation of having a local power-law around r akin to

$$f(r) = r^{\alpha_f(r)}$$

is highly flawed. This issue will be addressed later on.

In the literature, only the (dynamical) scaling exponent is used to infer a power-law scaling. But if one is interested in the initial ($t \ll 1$) or asymptotic ($t \gg 1$) regimes, a different method is possible:

Upper and lower indices

The crude and dynamical scaling exponents utilize the fact, that a strict power-law becomes linear in the log-log plot. Another way to understand power-law scaling is to study the behavior of r^β at the asymptotic $r \to 0$ or $r \to \infty$:[5]

Let $f(r) = r^\beta$ be a strict power-law with $\beta > 0$. For any $\gamma > 0$ the growth function

$$f_\gamma(r) := \frac{f(r)}{r^\gamma} = r^{\beta-\gamma} \tag{3.9}$$

forms again a power-law. If $\gamma > \beta$, then the f_γ is an inverse power-law and the initial limit blows up:

$$\lim_{r\to 0} f_\gamma(r) = \infty$$

But if $\gamma < \beta$, then the exponent in f_γ is positive and the initial limit vanishes:

$$\lim_{r\to 0} f_\gamma(r) = 0$$

The exponent $\gamma = \beta$ is therefore the **smallest** exponent, for which the initial limit is finite. The important point is that f_γ can be defined for **arbitrary** growth functions as

$$f_\gamma(r) = \frac{f(r)}{r^\gamma}. \tag{3.10}$$

If the initial limit diverges, then it it clear that f grows stronger than r^γ in the initial regime. Likewise, the growth function grows weaker than r^γ in the initial regime if the initial limit converges.

The value of the initial limit of f_γ can be used to determine if f grows stronger or weaker than a power-law, so the behavior of the limits can be used to define an alternative characterization of **power-law scaling**. This is the intuition behind the **lower index**:

Definition 17. *Let f be a growth function. The* **lower index** $\underline{\alpha}_f$ *is defined as the infimum*

$$\underline{\alpha}_f := \inf\left\{\gamma \;:\; \lim_{r\to 0} f_\gamma(r) = \infty\right\}. \tag{3.11}$$

If $\underline{\alpha}_f$ is finite, it is guaranteed that f grows stronger than r^γ for $\gamma < \underline{\alpha}_f$ and weaker for $\gamma > \underline{\alpha}_f$. But the limit

$$\lim_{r\to 0} \frac{f(r)}{r^{\underline{\alpha}_f}} \tag{3.12}$$

is undefined in general, since the definition involved the infimum.

In any case, for a finite lower index it is always possible to decompose f into

$$f(r) = r^{\overline{\alpha}_f} h(r) \tag{3.13}$$

---

[5] The analysis and usage of upper and lower indices appears in the theory of **regular variations** (see [19] and [2]). But according to my knowledge, this connection has never been observed or used in the context of scaling exponents and stochastic processes.

with remainder h. The lower index of h is $\underline{\alpha}_h = 0$, so the limit

$$\lim_{r\to\infty} \frac{h(r)}{r^\gamma} = 0,$$

holds for any $\gamma > 0$. But the initial limit $\lim\limits_{r\to\infty} h(r)$ is in general undefined. The following example illustrates this:

The growth functions

$$f^1(r) = r^\beta, \quad f^2(r) = r^\beta \log(1+r), \quad f^3(r) = \frac{r^\beta}{\log(r+1)}$$

share the same lower index $\underline{\alpha}_f = \beta$, but the limit at $\gamma = \beta$ is constant, diverges or vanishes.

This power-law behavior of f_γ does not only work for $r \to 0$, but also for $r \to \infty$.Since

$$\lim_{r\to\infty} r^\gamma = \lim_{r\to 0} r^{-\gamma},$$

the ranges of convergence and divergences have to be swapped. The corresponding index is called the **upper index**:

Definition 18. *Let f be a growth function. The* ***upper index*** $\overline{\alpha}_f$ *is defined as the supremum*

$$\overline{\alpha}_f := \sup\left\{\gamma \ : \ \lim_{r\to\infty} f_\gamma(r) = \infty\right\}. \tag{3.14}$$

3.3 Sewing things together

Dynamical scaling exponents and indices

If the dynamical scaling exponent α_f is well-defined and continuous, it possible to define the **initial scaling exponent**

$$\alpha_f(0) := \lim_{r\to 0} \alpha_f(r) \tag{3.15}$$

and the **asymptotic scaling exponent**

$$\alpha_f(\infty) := \lim_{r\to\infty} \alpha_f(r). \tag{3.16}$$

If α_f converges for $r \to 0$ or $r \to \infty$, the slope of f in the log-log plot needs to converge as well. This limit slope can therefore be considered as a scaling exponent. It is a remarkable fact that the initial and asymptotic scaling exponents **equal** the lower and upper indices. This identity is not relevant for numerical or experimental analysis, as the upper and lower indices can not be easily calculated from real data.

But the connection is vital for theoretical arguments, as the indices can be used very neatly in theoretical estimates and scaling predictions, e.g. in Ch. 5.

The following proposition is crucial for establishing this relation:

Proposition 1. *Let the initial scaling exponent* $\alpha(0) = \beta$ *be well - defined. Then*

$$\lambda(0) := \lim_{r\to 0} \lambda_r = \alpha_f(0) = \beta. \tag{3.17}$$

Similarly, let the asymptotic scaling exponent $\alpha(\infty) = \eta$ *be well - defined. Then*

$$\lambda(\infty) := \lim_{r\to\infty} \lambda_r = \alpha_f(\infty) = \eta. \tag{3.18}$$

Proof. It suffices to show the proof for the initial scaling exponent, since the asymptotic scaling follows via a nearly identical argument.

If $\alpha(0) = \beta$ is well - defined, then

$$\alpha_r = \beta + g(r)$$

with remainder function g satisfying $\lim_{r\to 0} g(r) = 0$ and being continuous at 0. The scaling exponent formula

$$\alpha_f(r) = r\partial_r \log f(r)$$

can be rearranged and integrated into

$$\log f(r) - \log f(r_0) = \int_{r_0}^{r} \alpha_f(u)\frac{du}{u}$$

for any $0 < r_0 < r$. The remainder decomposition implies

$$\log f(r) - \log f(r_0) = \beta(\log r - \log r_0) + \int_{r_0}^{r} g(u)\frac{du}{u}.$$

The crude scaling exponent $\lambda_r = \frac{\log f(r)}{\log r}$ can be recovered from this formula as

$$\lambda_r = \beta + \underset{a)}{\frac{\log f(r_0) - \beta\log r_0}{\log r}} + \underset{b)}{\frac{1}{\log r}\int_{r_0}^{r} g(u)\frac{du}{u}}. \tag{3.19}$$

The denominator in $a)$ is constant and even finite for any r_0, such that $r_0 < 1$ and $0 < f(r_0) < 1$, hence

$$\lim_{r\to 0} \frac{\log f(r_0) - \beta\log r_0}{\log r} = 0.$$

The remainder g is continuous at 0, so for a fixed $\epsilon > 0$ I can choose a $\delta > 0$ such that for all $0 < r < \delta$ $|g(r)| < \epsilon$ and $|\frac{\log r - \log r_0}{\log r}| < 1$ hold true. The integral can be estimated from above by

$$\left|\frac{1}{\log r}\int_{r_0}^{r} g(u)\frac{du}{u}\right| d < \left|\frac{\epsilon}{\log r}\int_{r_0}^{r}\frac{du}{u}\right| = \epsilon\left|\frac{\log r - \log r_0}{\log r}\right| < \epsilon,$$

which proves the convergence

$$\lim_{r\to 0}\int_{r_0}^{r} \alpha_f(u)\frac{du}{u} = 0$$

and the claim

$$\lim_{r\to 0} \lambda_r = \beta.$$

□

The initial crude and dynamical scaling exponents agree. This is crucial for the following theorem, which links the lower index $\underline{\alpha}_f$ and the initial scaling exponent $\alpha_f(0)$:

Theorem 2. *Let f be a growth function with well-defined scaling exponent α_f. The initial scaling exponent and lower index agree:*[6]

$$\alpha_f(0) = \underline{\alpha}_f \tag{3.20}$$

Proof. The theorem is proven by showing an upper and lower estimate:

Let $0 < \gamma < \underline{\alpha}_f$, then

$$\lim_{r\to 0} \frac{f(r)}{r^\gamma} = 0.$$

The function f can be decomposed as $f(r) = r^\gamma h(r)$ with non-negative h. The remainder is continuous and satisfies $\lim_{r\to 0} h(r) = 0$ due to $f(0) = 0$. The crude scaling exponent decomposes into

$$\lambda_r = \gamma + \frac{\log h(r)}{\log r}.$$

Since h is continuous, there has to be a $0 < r_0 < 1$ such that $0 < h(r) < 1$ for all $r < r_0$. For those r the quotient $\frac{\log h(r)}{\log r} \geq 0$ has to be positive and consequently

$$\liminf_{r\to 0} \lambda_r = \gamma + \liminf_{r\to 0} \frac{\log h(r)}{\log r} \geq \gamma.$$

Since $\lim_{r\to 0} \lambda_r = \alpha_f(0)$, this implies $\lim_{r\to 0} \lambda_r = \alpha_f(0) \geq \gamma$ and $\alpha_f(0) \geq \underline{\alpha}_f$.

For the reverse estimate, assume that $\delta \geq \underline{\alpha}$, then the limit

$$\lim_{r\to 0} \frac{f(r)}{r^\delta} = \infty$$

diverges. The growth function f can still be decomposed as $f(r) = r^\gamma h(r)$ with the remainder being non - negative and continuous. But h has to diverge

$$\lim_{r\to 0} h(r) = \infty.$$

Due to the divergence, there has to be a $0 < r_0 < 1$ such that $h(r) > 1$ for all $r > r_0$. In this case $\log r \leq 0$ and $\log h(r) \geq 0$ and especially

$$\limsup_{r\to 0} \frac{\log h(r)}{\log r} \leq 0.$$

---

[6] There is no direct source of this proof that I could find, although the results seems to be known. A similar proof strategy for other theorems is employed in [19], so it is likely that a proof of this theorem already exists in the literature.

Then

$$\limsup_{r\to 0} \lambda_r \leq \delta$$

and together with the limit $\lim_{r\to 0} \lambda_r = \alpha_f(0)$ this implies $\lim_{r\to 0} \lambda_r = \alpha_f(0) \leq \delta$. This shows the reverse estimate

$$\alpha_f(0) \leq \underline{\alpha}.$$

□

The proof strategy relied on the initial growth of the remainder h. This proof tactic is easily adapted to the asymptotic regime:

Theorem 3. *Let f be a growth function with well-defined scaling exponent α_f. The asymptotic scaling exponent and upper index agree:*

$$\alpha_f(\infty) = \overline{\alpha}_f \tag{3.21}$$

What constitutes local power-law scaling?

The initial / asymptotic scaling exponents agree with the lower / upper indices. This strengthens the interpretation of $\alpha_f(0)$ and $\alpha_f(\infty)$ being the exponents of a local power-law, since the decompositions

$$f(r) = r^{\alpha_f(0)} h_1(r), \quad f(r) = r^{\alpha_f(\infty)} h_2(r)$$

are well-behaved and the remainders h_1 and h_2 have controlled growth.

For intermediary times, only the scaling exponent $\alpha_f(r)$ is available. I already mentioned that the value of $\alpha_f(r)$ does not need to constitute some local power-law exponent in the sense of

$$f(r) \sim r^{\alpha_f(r)}.$$

Take the example of the integrated OU process from before, where he MSD and scaling exponent $\alpha_{MSD}(t)$ is shown in Fig. 3.2:

α_{MSD} stays roughly constant until $t \sim 3 \cdot 10^{-1}$, during which it scales **ballistically** due to $\alpha_f \sim 2$, and again after $t \sim 10^2$, where it scales **diffusively** due to $\alpha_f \sim 1$ .But in between, where α_f changes between the ballistic and diffusive regime, the scaling exponent is highly varying. Does it make sense to assume an underlying power-law

$$f(r) \sim r^{\alpha_f(r)},$$

if $\alpha_f(r)$ changes rapidly? The scaling exponent ranges in $1 < \alpha_f(r) < 2$, but no rational human being would claim that the *OU process scales superdiffusively* in between. What assumptions need to be fulfilled for α_f to be seen as a local power-law exponent?

A persistent scaling $\alpha_f \sim \beta$ for some period $[t_1, t_2]$ corresponds to a roughly constant slope in the log-log plot. It should therefore be sufficient that α_f stays approximately constant:

Definition 19. *Let f be a growth function with dynamical scaling exponent α_f. Then f is said to* ***scale stable*** *around r_0 if there is some $\epsilon > 0$ such that*

$$\alpha_f(r) \sim \beta, \quad \partial_r \alpha_f(r) \sim 0 \tag{3.22}$$

for $r \in [r_0 - \epsilon, r_0 + \epsilon]$. The period $[r_0 - \epsilon, r_0 + \epsilon]$ is called a ***stable regime****. Otherwise f is said to scale* ***transiently*** *and the time interval a* ***transient regime****.*

The notion of stable and transient regime is not mathematically precise, and indeed it is only a heuristical definition.[7] It is broad enough to include important cases like the integrated OU process, while also allowing some theoretical leverage in the arguments.

The initial and asymptotic regimes are typically stable scaling regimes, although counterexamples like $f(r) = (e^r - 1)^2$ exist. For the integrated OU process these are the only stable regimes, but intermediary stable regimes are also possible for more complex processes.

Transient regimes may be interesting in their own right,[8] but the scaling exponent α_f cannot be interpreted as the exponent of an approximate power-law scaling. Therefore, α_f should only given meaning during **stable regimes**.

---

[7] A mathematician would probably define α_f to be δ-stable at r_0, if $|\partial_r \alpha_f(r)| < \delta$ for $r \in [r_0 - \epsilon, r_0 + \epsilon]$. But this gives us nothing in this case.

[8] The switching between stable scaling regimes suggest a qualitative change in the dynamics of the system.

4 The primitive ones: Deterministic walks, subordination and integration

This chapter explores three examples of stochastic processes and their respective EAMSD and (A)TAMSD scaling. Most of the examples have a superballistic EAMSD, but none were able to attain a superballistic ATAMSD scaling.

Although the example of the deterministic walkers and subordinated processes are unphysical as concepts, they are a nice demonstration on how to calculate the EAMSD / ATAMSD and their scaling exponents analytically. This chapter also demonstrates the different ways on how to infer the scaling from the analytic formulas.

4.1 Deterministic walkers

The deterministic walker model is conceptually quite intuitive and simple:

Definition 20. *Let $f(t)$ be a non - negative, differentiable function with $f(0) = 0$ and e_θ a random unit vector. The* ***deterministic walker*** *x_t with* ***position function*** *$f(t)$ is defined as*

$$x_t = f(t)e_\theta. \tag{4.1}$$

The ***velocity processes*** *is given by*

$$v_t = f'(t)e_\theta. \tag{4.2}$$

The velocity processes performs a pathwise straight forward motion in the direction of the random direction e_θ. Aside of e_θ, no stochasticity is involved in the process x_t. The major qualitative difference to *genuine* stochastic motion lies in the straight forward path: Many stochastic processes like Brownian motion, solution processes to SDEs or even more exotic objects like Levy walks tend to have **angular decorrelation** over time, i.e. the average function

$$A(r,s) := \left\langle \frac{x_r}{|x_r|} \frac{x_s}{|x_s|} \right\rangle \tag{4.3}$$

will decay in time from the initial value $A(r,r) = 1$.

In case of Brownian motion this can be explained physically: The *velocity* of B_t is formally given by the **white noise** process η_t.[1], which is mutually independent:

$$\langle \eta_t \eta_{t'} \rangle = \delta(t - t') \tag{4.4}$$

The *direction* of η changes rapidly over time and the direction of η and η' are mutually independent. The process increment

$$B_t - B_r = \int_r^t \eta_s \, ds \sim \sum_{i=1}^n \eta_{t_i}(\Delta t) \tag{4.5}$$

becomes increasingly independent and decorrelated from the past increments due to this mutual force independence. Due to this mutual independence, the increments of B_t decorrelate over time. This leads to angular decorrelation

Angular decorrelation is a natural phenomenon and appears whenever the underlying stochastic forces are decorrelating over time. Deterministic walkers stand in stark contrast to this, as they do **not** angularly decorrelate:

$$A(r,s) = \left\langle \frac{x_r}{|x_r|} \frac{x_s}{|x_s|} \right\rangle = \langle e_\theta^2 \rangle = 1$$

Since deterministic walkers do not angularly decorrelate, any change in the autocorrelation

$$C_x(r,s) = \langle x_r x_s \rangle$$

or SMF / EAMSD

$$\langle (x_t - x_0)^2 \rangle = \langle x_t^2 \rangle + \langle x_0^2 \rangle - 2\langle x_0 x_t \rangle$$

is purely attributed to the behavior of the position function f. This explains why tailoring the scaling of the EAMSD is quite trivial:

EAMSD

The position function $f(0) = 0$ was assumed to be centered. This implies that the EAMSD and SMF agree:

$$EA_t = \langle (x_t - x_0)^2 \rangle = \langle x_t^2 \rangle = SMF_t$$

The EAMSD can be computed fairly straightforward:

$$\begin{aligned} EA_t &= \langle (x_t - x_0)^2 \rangle = \langle f^2(t) e_\theta^2 \rangle \\ &= f^2(t) \langle e_\theta^2 \rangle = f^2(t). \end{aligned}$$

The positional function f is the root of the EAMSD $f(t) = \sqrt{EA_t}$. Given this direct relationship, virtually **any** EAMSD scaling behavior is possible given a suitable choice of f.

The simple choice of a monomial $f(r) = r^b$ for $b \geq 0$ leads to a scaling exponent of

$$\alpha_{EA}(t) = 2b, \tag{4.6}$$

---

[1] This is a heuristic argument. η_t is at most a distribution.

so anything from subdiffusion to severe superballisticity for $b \gg 1$ is possible.
A sufficient fine tuning of f allows intermediary scaling regimes, a changing scaling exponent and more. The EAMSD scaling is thus directly determined by f, making the question of EAMSD scaling a matter of finding a suitable f. The TAMSD, on the other hand, behaves quite differently.

TAMSD and ATAMSD

The random unit vector e_θ always satisfies $e_\theta^2 = 1$, so the increment

$$\begin{aligned}(x_{t+\Delta} - x_t)^2 &= (f(t+\Delta)e_\theta - f(t)e_\theta)^2 = \left(f(t+\Delta) - f(t)\right)^2 e_\theta^2 \\ &= \left(f(t+\Delta) - f(t)\right)^2\end{aligned}$$

is inherently deterministic. This implies that the TAMSD and ATAMSD agree with each other.

For general f, this increment depends on t and the integral in the (A)TAMSD cannot be solved further than

$$\delta^2(\Delta, T) = \frac{1}{T-\Delta}\int_0^{T-\Delta} (x_{t+\Delta} - x_t)^2\, dt = \frac{1}{T-\Delta}\int_0^{T-\Delta} \left(f(t+\Delta) - f(t)\right)^2\, dt.$$

I will consider two examples: The **polynomial walker** with $f(t) = t^b$ for integer b and the **exponential walker** with $f(t) = e^{\lambda t} - 1$.

4.1.1 $f(t) = t^b$ for integer b

TAMSD

The first step in computing the TAMSD is to simplify the lagtime difference

$$\left(f(t+\Delta) - f(t)\right)^2 = \left((t+\Delta)^b - t^b\right)^2.$$

In the case of $f(t) = t^b$ for integer b, the lagtime difference can be simplified by using the binomial expansion

$$(t+\Delta)^b = \sum_{i=0}^{b} \binom{b}{i} \Delta^{b-i} t^i. \tag{4.7}$$

The difference $(t+\Delta)^b - t^b$ becomes

$$(t+\Delta)^b - t^b = \sum_{i=0}^{b-1} \binom{b}{i} \Delta^{b-i} t^i$$

and the the lagtime difference can be expressed combinatorically:

$$\left((t+\Delta)^b - t^b\right)^2 = \sum_{i,j=0}^{b-1} \binom{b}{i}\binom{b}{j} \Delta^{2b-i-j} t^{i+j}$$

This combinatorial expansion makes the termwise computation of the TAMSD feasible:

$$\delta^2(\Delta,T) = \sum_{i,j=0}^{b-1} \binom{b}{i}\binom{b}{j}\Delta^{2b-i-j}\frac{1}{T-\Delta}\int_0^{T-\Delta} t^{i+j}\,dt = \sum_{i,j=0}^{b-1}\binom{b}{i}\binom{b}{j}\Delta^{2b-i-j}\frac{(T-\Delta)^{i+j}}{i+j+1}$$

Every term in $\delta^2(\Delta,T)$ is a monomial in Δ and $T-\Delta$ with combined order of $2b$, i.e. **homogeneous** of order $2b$. Only the combinatorial factors have an explicit i and j dependence, so it is possible to reparametrize the sum:

$$\begin{aligned}\delta^2(\Delta,T) &= \sum_{i,j=0}^{b-1}\binom{b}{i}\binom{b}{j}\Delta^{2b-i-j}\frac{(T-\Delta)^{i+j}}{i+j+1} = \sum_{i,j=1}^{b}\binom{b}{i-1}\binom{b}{j-1}\Delta^{2(b-1)-i-j}\frac{(T-\Delta)^{i+j-2}}{i+j-1}\\ &= \sum_{k=0}^{2(b-1)}\Delta^{2b-k}\frac{(T-\Delta)^k}{k+1}\sum_{i=0}^{k}\binom{b}{i}\binom{b}{k-i} = \sum_{k=0}^{2(b-1)}\binom{2b}{k}\Delta^{2b-k}\frac{(T-\Delta)^k}{k+1}\end{aligned}$$

It is only reasonable to study the TAMSD scaling if $T \gg 1$, which implies $(T - \Delta) \overset{T\gg\Delta}{\sim} T$ and the Δ in $T-\Delta$ does not contribute to the lagtime scaling. The only relevant contributions are the explicit Δ^{2b-k} terms.
The TAMSD is a **homogeneous** polynomial of order $2b$ in Δ and $T-\Delta$, since Δ has order $2b-k$ and $T-\Delta$ order k. This is expected, since the lagtime increment $(f(t+\Delta)-f(t))^2$ itself is a homogeneous polynomial of order b in t and Δ. The highest order in Δ is the $k=0$ term, for which Δ^{2b} would suggest a constant scaling exponent of $\alpha_{TA} = 2b$. But this apparent superballisticity is deceivingly false!

The TAMSD scales in Δ and $T-\Delta$ simultaneously and the scaling w.r.t. lagtime Δ can **change** for different observational times T. This is best observed when computing the TAMSD directly via $\alpha_{TA}(\Delta,T) = \frac{\Delta}{\delta^2}\partial_\Delta\delta^2$:

$$\begin{aligned}\Delta\partial_\Delta\delta^2 &= \sum_{k=0}^{2(b-1)}\binom{2b}{k}\frac{2b-k}{k+1}\Delta^{2b-k}(T-\Delta)^k - \sum_{k=1}^{2(b-1)}\binom{2b}{k}\frac{k}{k+1}\Delta^{2b-k+1}(T-\Delta)^{k-1}\\ &= 2b\Delta^2(T-\Delta)^{2(b-1)} + \sum_{k=0}^{2(b-1)-1}\binom{2b}{k+1}\left(1-\frac{k+1}{k+2}\right)\Delta^{2b-k}(T-\Delta)^k\end{aligned}$$

The scaling exponent is given by this petite formula:

$$\alpha_{TA}(\Delta,T) = \frac{2b\Delta^2(T-\Delta)^{2(b-1)} + \sum_{k=0}^{2(b-1)-1}\binom{2b}{k+1}\left(1-\frac{k+1}{k+2}\right)\Delta^{2b-k}(T-\Delta)^k}{b\Delta^2(T-\Delta)^{2(b-1)} + \sum_{k=0}^{2(b-1)-1}\binom{2b}{k}\Delta^{2b-k}\frac{(T-\Delta)^k}{k+1}} \tag{4.8}$$

For $T \to \Delta$, the scaling exponent becomes

$$\alpha_{TA}(\Delta,\Delta) = \frac{\binom{2b}{1}\left(1-\frac{1}{2}\right)\Delta^{2b}}{\binom{2b}{0}\Delta^{2b}} = b. \tag{4.9}$$

Only in the limit $T \to \infty$ or for large $T \gg 1$ does the TAMSD and its scaling matter. In this limit, the scaling becomes **ballistic**:

$$\lim_{T\to\infty} \alpha_{TA}(\Delta, T) = \frac{2b\Delta^2}{b\Delta^2} = 2 \tag{4.10}$$

The term in δ^2 with the dominant $T - \Delta$ scaling dominates the lagtime scaling in the large T limit. This is a consequence of Thm. 21, which will be proven later. On the current level of analysis, this can be reasoned with by canceling the highest $T - \Delta$-power in each term:

$$\alpha_{TA}(\Delta, T) = \frac{2b\Delta^2 + \sum_{k=0}^{2(b-1)-1} \binom{2b}{k+1}\left(1 - \frac{k+1}{k+2}\right)\Delta^{2b-k}(T-\Delta)^{k-2(b-1)}}{b\Delta^2 + \sum_{k=0}^{2(b-1)-1} \binom{2b}{k}\Delta^{2b-k}\frac{(T-\Delta)^{k-2(b-1)}}{k+1}}$$

With the exception of the Δ^2 term, every other term contains a non-trivial inverse power of $T - \Delta$, which vanishes in the $T \to \infty$ limit. Only the Δ^2-terms contributes to the limiting scaling exponent. A comparison of the EAMSD and TAMSD, together

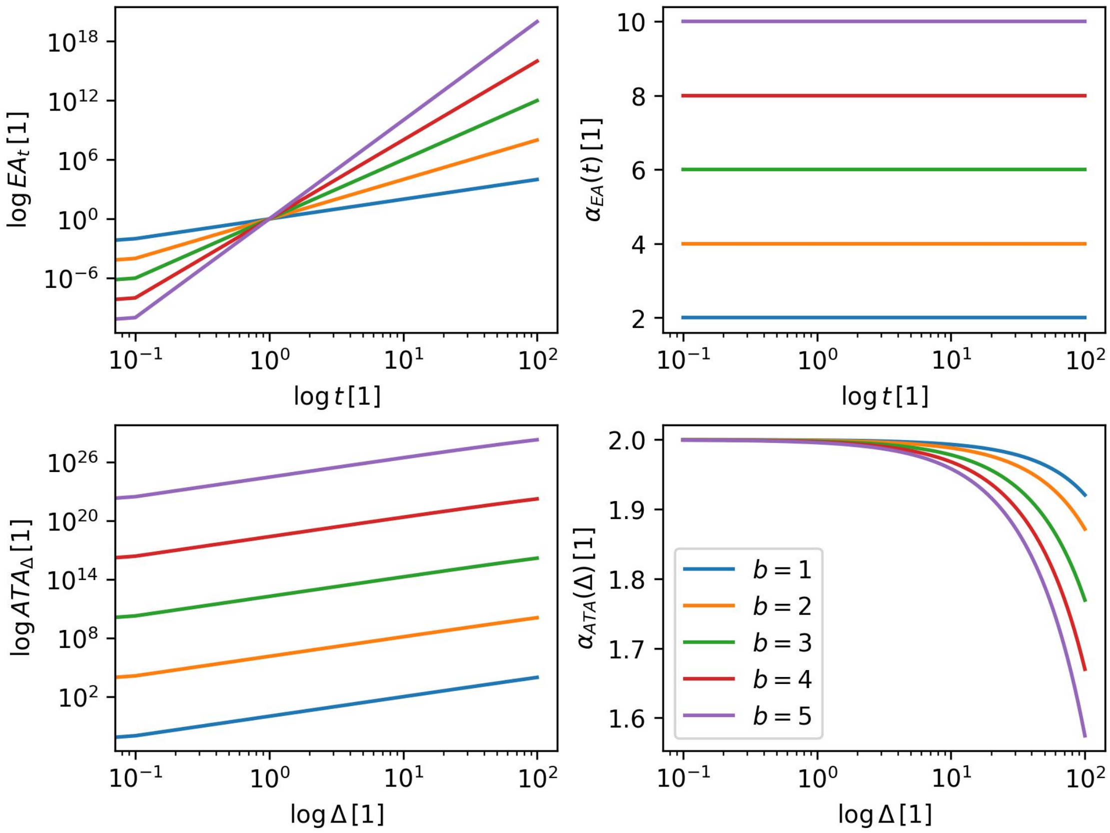

Figure 4.1: EAMSD / ATAMSD and scaling exponents for the polynomial walker $x_t = t^b e_\theta$. The ATAMSD scaling exponent is constant $\alpha_{ATA} = 2$ (later drop is due to numerical instabilities). Different choices of b only change the offset m in $\log ATA_\Delta = 2\log\Delta + m$, not the scaling exponent.

with their scaling exponents, is included in Fig. 4.1. while the EAMSD scaling $\alpha_{EA}(t) =$

$2b$ can be superballistic, depending on the value of b, the ATAMSD scaling is persistently ballistic $\alpha_{TA}(\Delta, T) \overset{T\gg 1}{\sim} 2$.

4.1.2 $f(t) = e^{\lambda t} - 1$

EAMSD

The EAMSD for this choice of f is $EA_t = (e^{\lambda t} - 1)^2$ and the scaling exponent becomes

$$\alpha_{EA}(t) = \frac{2\lambda t(e^{\lambda t} - 1)e^{\lambda t}}{(e^{\lambda t} - 1)^2} = \frac{2\lambda t}{1 - e^{-\lambda t}}. \tag{4.11}$$

For $t \ll 1$ the scaling exponent becomes ballistic $\alpha_{EA} \sim 2$, but the scaling exponent **diverges** asymptotically due to $\alpha_{EA}(t) \sim t$. The scaling exponent varies constantly and is not defined at $t = \infty$, so there is no **stable** scaling regime.

It should not be surprising that $f(r) = e^r - 1$ leads to no stable scaling, as exponential growth is inherently antithetical to power-law / algebraic growth. But even despite this conceptual difference, the initial and asymptotic scaling exponents still have a valid interpretation:

Using the Taylor series $e^{\lambda t} = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!}$, the EAMSD can be written as

$$\begin{aligned} EA_t = (e^{\lambda t} - 1)^2 &= \left(\sum_{n=1}^{\infty} \frac{(\lambda t)^n}{n!}\right)^2 \\ &= \sum_{n,m=1}^{\infty} \frac{\lambda^{n+m} t^{n+m}}{n!m!} = \lambda^2 t^2 \sum_{n,m=1}^{\infty} \frac{\lambda^n \lambda^m t^{n+m}}{(n+1)!(n+m)!} \end{aligned}$$

The remaining power series has a finite limit for $t \to 0$,

$$\lim_{t\to 0} \frac{(e^{\lambda t} - 1)^2}{t^2} = \lambda^2,$$

so the initial scaling has to be **ballistic** due to Thm. 2.

For large t, the EAMSD scales as $EA_t = e^{2\lambda t}$. The exponential grows stronger than any power of t, i.e.

$$\lim_{t\to\infty} \frac{EA_t}{t^\gamma} = \lim_{t\to\infty} \frac{e^{2\lambda t}}{t^\gamma} = \infty$$

for all $\gamma > 0$. The upper index is therefore $\overline{\alpha}_{EA} = \infty$ and this explains the diverging scaling exponent.

The EAMSD scaling, although transient, is heavily superballistic. Does this also persists in the TAMSD scaling?

TAMSD

The LI for $f(t) = e^{\lambda t} - 1$ is

$$(f(t + \Delta) - f(t))^2 = (e^{\lambda(t+\Delta)} - e^{\lambda t})^2 = e^{2\lambda t}(e^{\lambda\Delta} - 1)^2.$$

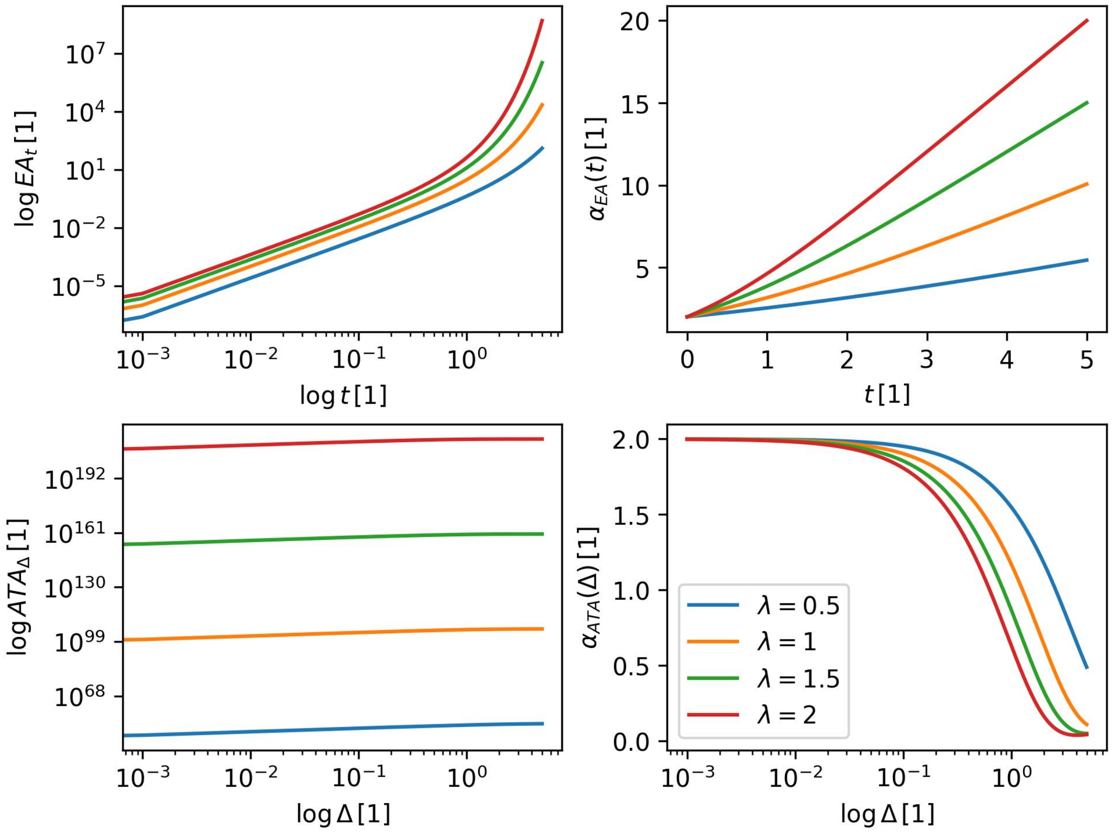

Figure 4.2: EAMSD / ATAMSD and scaling exponents for the polynomial walker $x_t = t^b e_\theta$. The EAMSD scaling exponent diverges linearly $\alpha_{EA}(t) \sim \lambda t$ for large times, whereas the ATAMSD scaling exponent vanishes $\alpha_{ATA}(\Delta) \sim 0$. Note the absurdly large offset of the ATAMSD.

It factorizes in Δ and t - dependent parts, so the TAMSD can be calculated without problems:

$$\begin{aligned}\delta^2(\Delta, T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} \left(f(t+\Delta) - f(t)\right)^2 \, dt \\ &= \frac{\left(e^{\lambda\Delta} - 1\right)^2}{T-\Delta}\int_0^{T-\Delta} e^{2\lambda t}\, dt = \left(e^{\lambda\Delta} - 1\right)^2 \frac{e^{2\lambda(T-\Delta)} - 1}{2\lambda(T-\Delta)}\end{aligned}$$

Using the logarithmic decomposition

$$\log \delta^2(\Delta, T) = 2\log\left(e^{\lambda\Delta} - 1\right) + \log\left(e^{2\lambda(T-\Delta)} - 1\right) - \log\left(2\lambda(T-\Delta)\right),$$

the scaling exponent is easy to calculate as well:

$$\alpha_{TA}(\Delta, T) = \Delta\partial_\Delta \log \delta^2(\Delta, T) = \frac{2\lambda\Delta}{1 - e^{-\lambda\Delta}} - \frac{2\lambda\Delta}{1 - e^{-2\lambda(T-\Delta)}} + \frac{2\lambda\Delta}{T-\Delta}$$

The large T limit is simple:

$$\alpha_{TA}(\Delta) = \lim_{T\to\infty} \alpha_{TA}(\Delta, T) = \frac{2\lambda\Delta}{e^{\lambda\Delta} - 1} \tag{4.12}$$

Both EAMSD and TAMSD, together with their scaling exponents, are shown in Fig. 4.2. The initial scaling of the TAMSD is ballistic due to $\alpha_{TA}(0) = 2$, which agrees with the EAMSD. But the intermediary and asymptotic regime differs drastically:

While the EAMSD scaling exponent increases linearly, the TAMSD scaling exponent **decays** to 0 over time due to the exponentially diverging denominator. But why? It does not seem that the TAMSD scaling exponent decays from the analytic formula alone, since there is an explicit $(e^{\lambda\Delta} - 1)^2$ term. But some guesses can be made: The TAMSD can be regrouped as

$$\delta^2(\Delta, T) = \left(1 - e^{-\lambda\Delta}\right)^2 \frac{e^{2\lambda T} - e^{2\lambda\Delta}}{2\lambda(T - \Delta)}. \tag{4.13}$$

The second factor is the arithmetic mean of $e^{2\lambda T}$ and $e^{\lambda\Delta}$. If $\Delta \ll T$, this term can be approximated as

$$\frac{e^{2\lambda T} - e^{2\lambda\Delta}}{2\lambda(T - \Delta)} \sim \frac{e^{2\lambda T}}{2\lambda T}$$

and the only Δ dependence remains in the first factor $\left(1 - e^{-\lambda\Delta}\right)^2$. But this saturates to 1 for $\Delta \gg \lambda$, leading to a constant plateau of δ^2. If the TAMSD plateaus as $\delta^2(\Delta, T) \sim$ const, the scaling exponent has to vanish $\alpha_{TA} \sim 0$.

This is one of the peculiarities that can appear for the scaling exponent if the underlying functions are inherently non-algebraic. The TAMSD suggest some *localization* of the trajectories due to the vanishing scaling exponent, but the process is not confined. It is vital to assure a stable scaling regime, because otherwise seemingly paradoxical results like the localized TAMSD are mistakenly interpreted.

4.2 Subordinated processes

Subordination is a procedure that distorts a given stochastic process x_t *in time*.

Definition 21. *Let x_t be a prescribed stochastic process and f be a differentiable, strictly increasing function with $f(0) = 0$. The process*

$$y_t := x_{f(t)} \tag{4.14}$$

*is called a **subordinated process**. The function f is called the **subordinator**.*

From the physical point of view, a subordinated process y_t can be interpreted as x_t viewed with an altered time perception. The observational time, during which x_t performs its motion, is being skewed by the subordinator f.

Subordination in itself is a pretty unphysical procedure: Although the concept of subordination is used by probability theorists and people in finance, the concept of altered time frames is not really convincing in the physical setting. Given that stochastic motions tend to be observed in inanimate objects, which do not have a perception of time scales, it is unreasonable that a given dynamics is being performed in a distorted timely manner.

From the theoretical side it is nonetheless interesting, as subordinated processes can be used as toy models for superballistic diffusions.

EAMSD and SMF

As the procedure of subordination alters solely the time dependence of the process, direct quantities like the EAMSD and SMF are unaffected. Writing EA_t^x and SMF_t^x for the EAMSD and SMF of the original process x_t, the corresponding quantities for the subordinated process follow easily:

$$SMF_t^y = \langle y_t^2 \rangle = \langle x_{f(t)}^2 \rangle = SMF_{f(t)}^y$$

$$EA_t^y = \langle (y_t - y_0)^2 \rangle = \langle (x_{f(t)} - x_{f(0)})^2 \rangle = \langle (x_{f(t)} - x_0)^2 \rangle = EA_{f(t)}^x$$

The EAMSDs and SMFs of y_t and the original x_t are only altered by a relative time change. The corresponding scaling exponents should thus be also related:

$$\begin{aligned} \alpha_{EA}^y(t) &= \frac{t}{EA_t^y} \partial_t (EA_t^y) = \frac{t}{EA_{f(t)}^x} \partial_t (EA_{f(t)}^x) = \frac{t}{EA_{f(t)}^x} f'(t) (\partial_t EA_{f(t)}^x) \\ &= \frac{t f'(t)}{f(t)} \frac{t f(t)}{EA_{f(t)}^x} (\partial_t EA_{f(t)}^x) = \frac{t f'(t)}{f(t)} \alpha_{EA}^x (f(t)) \end{aligned}$$

The EAMSD scaling exponents of x_t and y_t are again related by some time change, albeit with a new modulation term appearing. In case of general subordinators, this can destroy any stable scaling relation.[2] If the subordinator is chosen as a monomial $f(t) = t^b$ for $b \geq 0$, then the scaling exponent of y_t still admits a coherent form:

$$\alpha_{EA}^y(t) = b \cdot \alpha_{EA}^x(f(t)) \tag{4.15}$$

The subordinator shifts the scaling exponent multiplicatively. If x_t has a constant scaling exponent of $\alpha_{EA}^x = a$, the subordinator process will have a stable scaling of $\alpha_{EA}^y = a \cdot b$. The EAMSD scaling can be easily tailored to any superballistic regime, similarly to the case of the deterministic walker. Since a power - law subordinator $f(r) = r^b$ only modulates the scaling exponent multiplicatively, any stable scaling regime of x corresponds to a stable regime of y_t.

Take the example of subordinated Brownian motion (SBM) $x_t = B_{t^b}$ with the power - law subordinator $f(t) = t^b$ for $b > 0$:

Since $\langle B_t^2 \rangle = t$, the EAMSD of SBM is given by $EA_t = \langle B_{t^b}$ and the EAMSD scaling is persistent:

$$\alpha_{EA}(t) = b \tag{4.16}$$

b can be chosen arbitrarily, so the SBM allows for **any** stable scaling.

TAMSD and ATAMSD

The EAMSD of x_t and y_t are intimately related due to the time change involved. This simple time change becomes very problematic in case of the TAMSD and ATAMSD. Since the formulae for both involve a time - averaged integral,

$$\frac{1}{T - \Delta} \int_0^{T-\Delta} (x_{t+\Delta} - x_t)^2 \, dt,$$

---

[2] Take $f(t) = e^t$ and everything becomes mayhem.

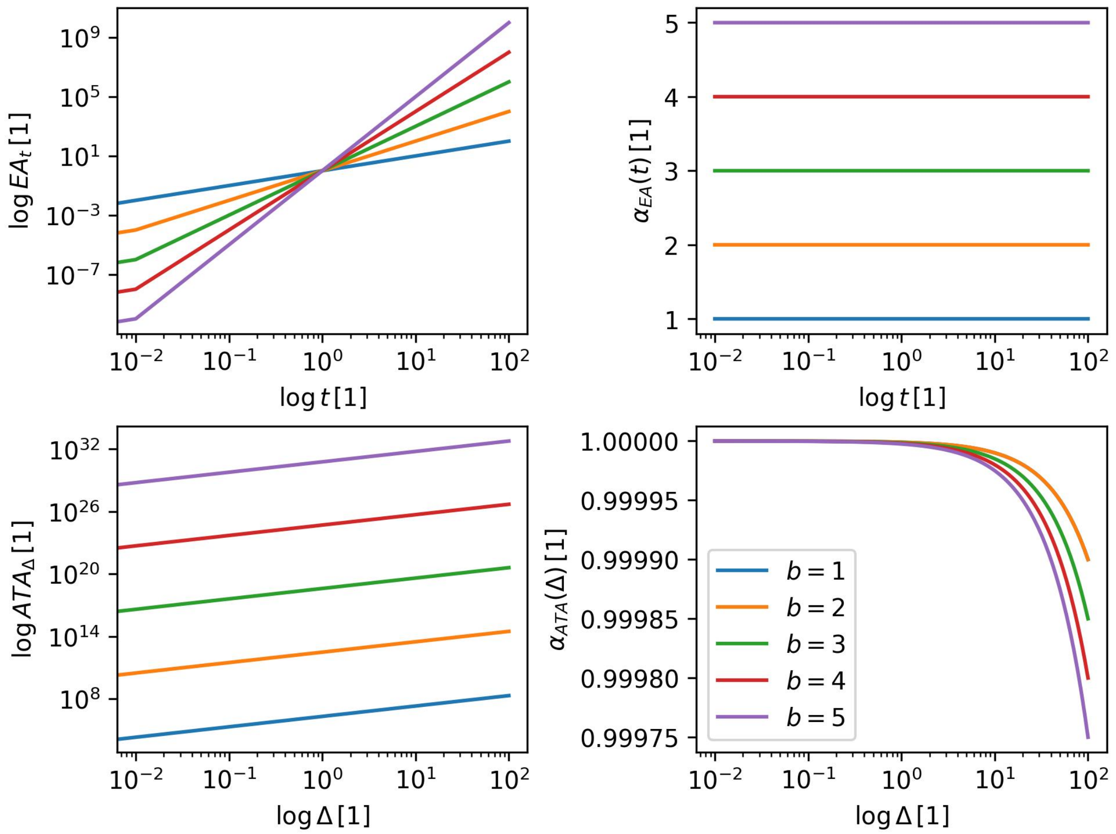

Figure 4.3: EAMSD / ATAMSD and scaling exponents for subordinated Brownian motion $x_t = B_{t^b}$. The ATAMSD scaling exponent is constant $\alpha_{ATA} = 1$ (later drop is due to numerical instabilities). Different choices of b only change the offset m in $\log ATA_\Delta = b \log \Delta + m$, not the scaling exponent. Note that stark similarity to Fig. 4.1.

a reparametrization destroys the structure of this integral. It is therefore not possible to relate the TAMSD and ATAMSD of x_t and y_t to each other.

If the subordination is done by a simple linear scaling $f(r) = cr$, then the TAMSD can be matched. This is due to the affine parametrization invariance of the time integral:

$$
\begin{aligned}
\delta_y^2(\Delta, T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} (y_{t+\Delta} - y_t)^2\,dt = \frac{1}{T-\Delta}\int_0^{T-\Delta} (x_{c(t+\Delta)} - x_{ct}^2\,dt \\
&= \frac{1}{T-\Delta}\int_0^{T-\Delta} (x_{ct+c\Delta)} - x_{ct})^2\,dt \overset{ct\to t}{=} \frac{1}{c(T-\Delta)}\int_0^{c(T-\Delta)} (x_{t+c\Delta} - x_t)^2\,dt \\
&= \delta_x^2(c\Delta, cT)
\end{aligned}
$$

Any non-linear subordinator distorts the lagtime difference, so the (A)TAMSD of y_t can behave quite different than the (A)TAMSD of x_t.

Subordinated Brownian motion (SBM)

If $x_t = B_{t^b}$ (with subordinator $f(t) = t^b$), the process is called **subordinated Brownian motion (SBM).**[3] For SBM, the ATAMSD can be computed analytically:

$$\begin{aligned}\delta_x^2(\Delta, T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} \langle (x_{t+\Delta} - x_t)^2 \rangle \, dt = \frac{1}{T-\Delta}\int_0^{T-\Delta} \left\langle \left(B_{(t+\Delta)^b} - B_{t^b}\right)^2 \right\rangle dt \\ &= \frac{1}{T-\Delta}\int_0^{T-\Delta} (t+\Delta)^b - t^b \, dt = \frac{1}{b+1}\frac{T^{b+1} - \Delta^{b+1} - (T-\Delta)^{b+1}}{T-\Delta} \\ &= \frac{1}{b+1}\sum_{k=1}^{b}\binom{b+1}{k}(T-\Delta)^k \Delta^{b+1-k}\end{aligned}$$

The ATAMSD of x_t has conceptionally the same form as the ATAMSD of the **polynomial walker** $y_t = f(t)e_\theta$ for $f(t) = t^b$ in Sec. 4.1.1. The lagtime scaling has been determined by the highest power in $T - \Delta$, which should also be expected here. For SBM, the highest term is given by

$$\delta^2(\Delta, T) \sim (T-\Delta)^b \Delta,$$

which suggests a **diffusive** scaling for all possible integer values of b, despite the **superballistic** EAMSD scaling. This is confirmed by the plots in Fig. 4.3.

4.3 Integrated processes

The last class of processes I will explore here are the **integrated processes**:

Definition 22. *Let x_t be a prescribed processes. The* ***integrated process*** *y_t is pathwise defined via integration:*

$$y_t = y_0 + \int_0^t x_s \, ds \tag{4.17}$$

The starting point y_0 can be arbitrarily chosen / distributed.

4.3.1 Integrated Brownian motion

The first example is **integrated Brownian motion**:[4]

$$x_t = \int_0^t B_s \, ds$$

The **averaged lagtime increment (ALI)** $\langle (x_{t+\Delta} - x_t)^2 \rangle$ is crucial in computing the EAMSD and ATAMSD. For integrated Brownian motion it can be calculated analyti-

---

[3] The TAMSD and EAMSD of subordinated Brownian motion are similar to those of **scaled Brownian motion** (see [14]).

[4] Another name is the **random acceleration process**.

cally:

$$\begin{aligned}\langle (x_{t+\Delta}-x_t)^2\rangle &= \left\langle \left(\int_t^{t+\Delta} B_s\,ds\right)^2\right\rangle = \int_t^{t+\Delta}\int_t^{t+\Delta}\langle B_rB_s\rangle\,dr\,ds\\ &= \int_t^{t+\Delta}\int_t^{t+\Delta} C_B(r,s)\,dr\,ds = \int_t^{t+\Delta}\int_t^{t+\Delta} min(r,s)\,dr\,ds\\ &= \frac{1}{3}\Delta^3 + t\Delta^2.\end{aligned}$$

EAMSD

The EAMSD follows directly from the averaged lagtime difference:

$$EA_t = \langle (x_t - x_0)^2\rangle = \frac{1}{3}t^3 \tag{4.18}$$

The EAMSD of integrated Brownian motion scales **superballistically** with scaling exponent $\alpha = 3$. This scaling is persistent over time, which can be expected from the persistent diffusive scaling of B_t itself. The exact value of $\alpha = 3$ can be explained heuristically:

The main diagonal of the ACF $C_B(r,r) = \langle B_r^2\rangle = r$ is the SMF of Brownian motion, which scales diffusively. For very small $t \ll 1$ the ACF $C_B(r,s) \sim C_B(t,t)$ can be assumed constant for $r,s \in [0,t]$ and the crude approximation

$$\begin{aligned}EA_t &= \int_0^t\int_0^t C_B(r,s)\,dr\,ds \sim \int_0^t\int_0^t C_B(t,t)\,dr\,ds\\ &= t\int_0^t\int_0^t dr\,ds = t^3\end{aligned}$$

works for the initial scaling regime. This heuristic argument suggest a cubic scaling for small times, which persists due to the persistent scaling of B_t itself.

I will prove a weaker variant of the scaling relation

$$\alpha_x(t) \sim \alpha_v(t) + 2$$

between the EAMSD scaling exponent of x_t and its velocity SMF scaling exponent for initial times in Thm. 8.

ATAMSD

The ATAMSD of x_t can be computed from

$$\langle \delta^2(\Delta,T)\rangle = \frac{1}{T-\Delta}\int_0^{T-\Delta}\langle (x_{t+\Delta}-x_t)^2\rangle\,dt$$

thanks to the analytic formula for the ALI:

$$\langle (x_{t+\Delta}-x_t)^2\rangle = \frac{1}{3}\Delta^3 + t\Delta^2$$

The ATAMSD of integrated Brownian motion becomes

$$\langle \delta^2(\Delta, T) \rangle = 2\Delta^2(T - \Delta) + \frac{\Delta^3}{3}. \tag{4.19}$$

The ATAMSD has a cubic term in Δ, just as the EAMSD. This is expected, since

$$\langle \delta^2(\Delta, \Delta) \rangle = EA_\Delta.$$

But there is a second term with quadratic Δ scaling, since $T - \Delta$ is effectively non - scaling in Δ for sufficiently large T. Indeed, the scaling exponent for the ATAMSD can be computed directly as

$$\alpha_{ATA}(\Delta, T) = \frac{\Delta \partial_\Delta \langle \delta^2 \rangle}{\langle \delta^2 \rangle} = \frac{4\Delta^2(T - \Delta) - \Delta^3}{2\Delta^2(T - \Delta) + \frac{\Delta^3}{3}}.$$

A naive guess for the scaling exponent at $T \sim \Delta$ would be $\alpha_{ATA}(\Delta, \Delta) = \alpha_{EA}(\Delta) = 3$, but the computation shows

$$\alpha_{ATA}(\Delta, \Delta) = -3.$$

The reason for this might be that the ATAMSD can have cusp - like singularities around $T = \Delta$, so the ATAMSD scaling exponent does not converge to the EAMSD scaling exponent for $\Delta \to T$. The scaling exponent only has a well - defined meaning if $T \gg \Delta$, so this is not a pressing problem.

In the large T-limit, the scaling exponent reduces to

$$\lim_{T \to \infty} \alpha_{ATA}(\Delta, T) = \lim_{T \to \infty} \frac{4\Delta^2 - \frac{\Delta^3}{(T-\Delta)}}{2\Delta^2 + \frac{\Delta^2}{3(T-\Delta)}} = 2 \tag{4.20}$$

and ATAMSD scales persistently **ballistic**. But why? From the mathematical side, this is simple:

The ATAMSD contains a cubic term in Δ, but the scaling exponent is computed for $\Delta \ll T$. In this case, the **effective** scaling is determined by the term with the **highest** $T - \Delta$ - dependence. For integrated Brownian motion, this term is $2\Delta^2(T - \Delta)$, hence the effective ballistic scaling. But this answer does not give any intuitive insight into why this happens on a physical level. I will address this concern more conceptually in Ch. 8.

4.3.2 Integrated equilibriated Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck process has been introduced in Sec. 2.2 as the solution process to the following Langevin SDE:

$$dv_t = -v_t \, dt + \sqrt{2} \, dB_t$$

If the initial velocity $v_0 \sim \pi$ is distributed according to the invariant distribution π, this process is called the **equilibriated Ornstein-Uhlenbeck (EOU) process**. The integrated process x_t will be called the **integrated equilibriated Ornstein-Uhlenbeck**

process. For quite obvious reasons I will call this object the integrated EOU process instead.

The EAMSD and ATAMSD of x_t can be computed analytically, since the ACF of the EOU process is explicitly known:

$$C_v(r,s) = \langle v_r v_s \rangle = e^{-|r-s|} \tag{4.21}$$

The ACF only depends on the difference $r-s$, it is therefore **stationary**. The important class of stationary ACFs will be dealt with in detail in 6.5.

The ALI of x_t is

$$\begin{aligned}\langle (x_{t+\Delta}-x_t)^2\rangle &= \int_t^{t+\Delta}\int_t^{t+\Delta} e^{-|r-s|}\,dr\,ds = \int_0^t\int_0^t e^{-|r-s|}\,dr\,ds\\ &= 2\int_0^t\int_s^t e^{-(r-s)}\,dr\,ds = 2\int_0^t 2(1-e^{-1})\,ds\\ &= 2(\Delta - 1 + e^{-\Delta}).\end{aligned}$$

There is no dependence on t in the final formula for the ALI, which is characteristic for stationary ACFs. This has as a consequence that the ATAMSD and EAMSD **agree**:

$$\begin{aligned}\langle \delta^2(\Delta,T)\rangle &= \frac{1}{T-\Delta}\int_0^{T-\Delta}\langle (x_{t+\Delta}-x_t)^2\rangle\,dt\\ &= \frac{1}{T-\Delta}\int_0^{T-\Delta}\langle (x_\Delta-x_0)^2\rangle\,dt\\ &= \langle (x_\Delta - x_0)^2\rangle = EA_\Delta\end{aligned}$$

In this case, both quantities are written as MSD_t.

MSD

Since the MSD is just the averaged lagtime increment, it is

$$MSD_t = \langle (x_t - x_0)^2\rangle = 2\,(t-1+e^{-t})\,. \tag{4.22}$$

The scaling exponent can be calculated directly:

$$\alpha_{MSD}(t) = \frac{t(1+e^{-t})}{t+1-e^{-t}} \tag{4.23}$$

The MSD scales **ballistically** for short times due to

$$\alpha_{MSD}(t) = \frac{t(1+e^{-t})}{t+1-e^{-t}} = \frac{1+e^{-t}}{1+\frac{1-e^{-t}}{t}} \to \frac{2}{1} = 2,$$

but becomes asymptotically diffusive:

$$\alpha_{MSD}(t) = \frac{t(1+e^{-t})}{t+1-e^{-t}} = \frac{1+e^{-t}}{1+\frac{1-e^{-t}}{t}} \overset{t\gg 1}{\sim} 1.$$

The integrated EOU process does **not** perform superballistic motion. The underlying reason for this comes from the stationary ACF of the EOU - process and the monotonic decay of

$$C_v(r,s) = e^{-|r-s|}.$$

Superballistic scaling for processes with a stationary VACF (velocity ACF) will be investigated in detail in Sec. 6.5.

It is, however, possible for the EAMSD to scale superballistically if the equilibrium condition is dropped relaxed:

4.3.3 Integrated non-equilibriated Ornstein-Uhlenbeck process

The equilibriated OU process has been defined by the SDE

$$dv_t = -v_t\,dt + \sqrt{2}\,dB_t$$

with $v_0 \sim \pi$ being invariantly distributed. If I consider the solution to this SDE with deterministic, vanishing initial velocity $v_0 \equiv 0$, a different process emerges. I will call this process the **non-equilibriated Ornstein-Uhlenbeck (NEOU) process**.

The substantial difference to the previous model is the **non-stationary** ACF:

$$C_v(r,s) = \langle v_r v_s \rangle = e^{-|r-s|} - e^{-(r+s)} \tag{4.24}$$

The ACF is the sum of the stationary ACF of the EOU process and the non-stationary kernel $e^{-(r+s)}$. This makes the calculated of the ALI quite trivial:

$$\begin{aligned}
\langle (x_{t+\Delta} - x_t)^2 \rangle &= 2\left(\Delta - 1 + e^{-\Delta}\right) - \int_t^{t+\Delta}\int_t^{t+\Delta} e^{-(r+s)}\,dr\,ds \\
&= 2\left(\Delta - 1 + e^{-\Delta}\right) - \left(\int_t^{t+\Delta} e^{-r}\,dr\right)^2 \\
&= 2\left(\Delta - 1 + e^{-\Delta}\right) - e^{-2t}\left(1 - e^{-\Delta}\right)^2
\end{aligned}$$

The averaged lagtime difference has an explicit t - dependence, so the EAMSD and ATAMSD will disagree. They need to be studied separately:

EAMSD

The EAMSD follows from the ALI:

$$\begin{aligned}
EA_t = \langle (x_t - x_0)^2 \rangle &= 2\,(t - 1 + e^{-t}) - (1 - e^{-t})^2 \\
&= 2t - 3 + 4e^{-t} - e^{-2t}
\end{aligned}$$

The EAMSD of the integrated NEOU process and the MSD of the integrated EOU process agree asymptotically:

$$\frac{2t - 3 + 4e^{-t} - e^{-2t}}{2\,(t + 1 - e^{-t})} \overset{t\to\infty}{\to} 1$$

Due to the asymptotically identical EAMSD, the integrated NEOU process scales asymptotically **diffusive**:

$$EA_t \overset{t\gg 1}{\sim} 2t \tag{4.25}$$

This can be deduced from the VACF of the process as well:

The non - stationary term $e^{-(r+s)}$ in the VACF corresponds to $-\left(1-e^{-t}\right)^2$ in the EAMSD, which becomes asymptotically constant:

$$-\left(1-e^{-t}\right)^2 \overset{t\to\infty}{\to} -1,$$

This has a weaker scaling than the stationary contribution $2\left(t-1+e^{-t}\right)$, thus the stationary diffusive part dominates in the long run.

The non - stationary contribution becomes important in the initial regime, as the modulation term *alters* the scaling behavior in comparison to the equilbriated OU process. It turns out that $\alpha_{EA}(0) = 3$, i.e. the process scales initially **cubically**:

One possibility for showing this is by analyzing the scaling exponent directly:

$$\alpha_{EA}(t) = \frac{2t\left(1+e^{-2t}-2e^{-t}\right)}{2t-3+4e^{-t}-e^{-2t}}$$

Both parts of the fraction vanish for $t \to 0$, and likewise their higher derivatives. The initial scaling exponent is only revealed after 3**(!)** iterated applications of l'Hospital:

$$\begin{aligned}\lim_{t\to 0}\alpha_{EA}(t) &= \lim_{t\to 0}\frac{2t\left(1+e^{-2t}-2e^{-t}\right)}{2t-3+4e^{-t}-e^{-2t}} = 1+\lim_{t\to 0}\frac{2t\left(e^{-t}-e^{-2t}\right)}{1+e^{-2t}-2e^{-t}}\\ &= 2+\lim_{t\to 0}\frac{2t\left(2e^{-2t}-e^{-t}\right)}{2e^{-t}-e^{-2t}} = 3+\lim_{t\to 0}\frac{t\left(e^{-t}-4e^{-2t}\right)}{2e^{-2t}-e^{-t}}\\ &= 3\end{aligned}$$

Calculating the scaling exponent directly has multiple downside: It is not feasible to calculate for complex MSDs, numerical approximations have to be used for the limit and the physical interpretation remains obscure.

Thm. 2 established the deep result that the initial scaling exponent is related to the lower index $\underline{\alpha}_{EA}$. Instead of calculating the fraction or log - derivative in $\alpha_{EA}(0)$ directly, it is possible to test the convergence of the limit

$$\lim_{t\to 0}\frac{EA_t}{t^\gamma}$$

for different values of γ instead. The EAMSD contains not only powers in t, but also the transcendental exponential e^{-t} and e^{-2t}. For these it is possible to use the **Taylor series** of e^t, which allows to expand both transcendental terms:

$$\begin{aligned}4e^{-t} &= 4-4t+2t^2-\frac{2}{3}t^3+4\sum_{n=4}^{\infty}\frac{(-t)^n}{n!}\\ -e^{-2t} &= -1+2t-2t^2+\frac{4}{3}t^3-\sum_{n=4}^{\infty}\frac{(2t)^n}{n!}\end{aligned}$$

This gives the EAMSD the following power series expansion:

$$EA_t = \frac{2}{3}t^3 + \sum_{n=4}^{\infty} \left((-1)^n 4 - 2^n\right) \frac{t^n}{n!}$$

The remainder of the power series contains only terms with order $n \geq 4$, so the cubic limit

$$\lim_{t\to 0} \frac{EA_t}{t^3} = \frac{2}{3} + \lim_{t\to 0} \sum_{n=0}^{\infty} \left((-1)^{n+4} 4 - 2^{n+4}\right) \frac{t^{n+1}}{(n+4)!} = \frac{2}{3} \tag{4.26}$$

is finite! Therefore $\underline{\alpha}_{EA} = 3$ and the initial scaling exponent is $\alpha_{EA}(0) = 3$.

This technique of computing the initial (and also asymptotic) scaling exponent via power series expansion will be used multiple times in this thesis in favor of the more direct computation of α. The initial scaling exponent can thus be computed by finding the lowest non-vanishing order in t.

There is also a third way to get to the cubic scaling, namely through a heuristical analysis of the SDE itself:

The NEOU process solves

$$dv_t = -v_t dt + \sqrt{2} dB_t$$

with initial condition $v_0 \equiv 0$, which is interpreted as the integral equation

$$v_t = v_0 + \int_0^t v_s \, ds + \sqrt{2} B_t. \tag{4.27}$$

Since the initial velocity $v_0 \equiv 0$ vanishes, the integral equation simplifies to

$$v_t = \int_0^t v_s \, ds + \sqrt{2} B_t.$$

A $v_0 \equiv 0$ and v_t is pathwise continuous, the velocity v_t can be assumed to be very small for $t \ll 1$ and therefore approximately $v_t \sim 0$.[5] But if $v_t \sim 0$ is negligible, the integral $\int_0^t v_s \, ds \sim 0$ becomes sufficiently small for $t \ll 1$, too. The SDE reduces to

$$v_t = \int_0^t v_s \, ds + \sqrt{2} B_t \overset{t \ll 1}{\sim} \sqrt{2} B_t$$

and the velocity is approximately **Brownian**:

$$v_t \overset{t \ll 1}{\sim} \sqrt{2} B_t \tag{4.28}$$

It follows that the velocities of the integrated NEOU process and integrated Brownian motion are roughly identical for small t,[6] therefore they share the same **initial** EAMSD scaling. This argumentation is also physically justified:

---

[5] The velocity of course does not vanish there. But for the following argument it is sufficient that the velocity is insignificant **compared** to Brownian motion.

[6] Aside of the different normalization $\sqrt{2}$.

The drift term $-v_t\,dt$ of the OU process represents a physical friction force, which is insignificant for sufficiently small velocities. In this case the velocity v_t is dominated by the Brownian noise and effectively behaves as a Brownian particle.

Does this scaling also hold for the TAMSD?

ATAMSD

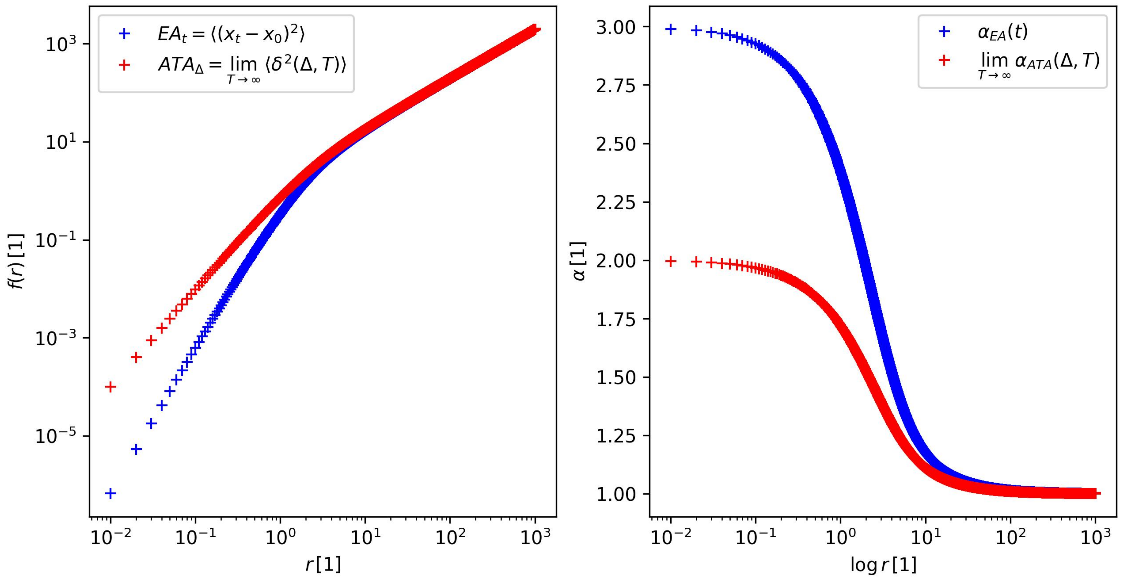

Figure 4.4: Comparison of EAMSD and ATAMSD scaling for the integrated non-equilibriated Ornstein-Uhlenbeck process.

The ALI

$$\langle (x_{t+\Delta} - x_t)^2 \rangle = 2\left(\Delta - 1 + e^{-\Delta}\right) - e^{-2t}\left(1 - e^{-\Delta}\right)^2$$

is already split into a stationary part (independent of t) and a t - dependent remainder, the ATAMSD can be computed directly:

$$\begin{aligned}
\langle \delta^2(\Delta, T) \rangle &= \frac{1}{T-\Delta} \int_0^{T-\Delta} \langle (x_{t+\Delta} - x_t)^2 \rangle \, dt \\
&= 2\left(\Delta - 1 + e^{-\Delta}\right) - \frac{(1-e^{-\Delta})^2}{T-\Delta} \int_0^{T-\Delta} e^{-2t}\, dt \\
&= 2\left(\Delta - 1 + e^{-\Delta}\right) - \frac{(1-e^{-\Delta})^2}{2} \frac{1-e^{-2(T-\Delta)}}{T-\Delta}
\end{aligned}$$

For $\Delta \to T$, the T-dependent part converges to

$$\lim_{\Delta\to T} \frac{1-e^{-2(T-\Delta)}}{2(T-\Delta)} = 1$$

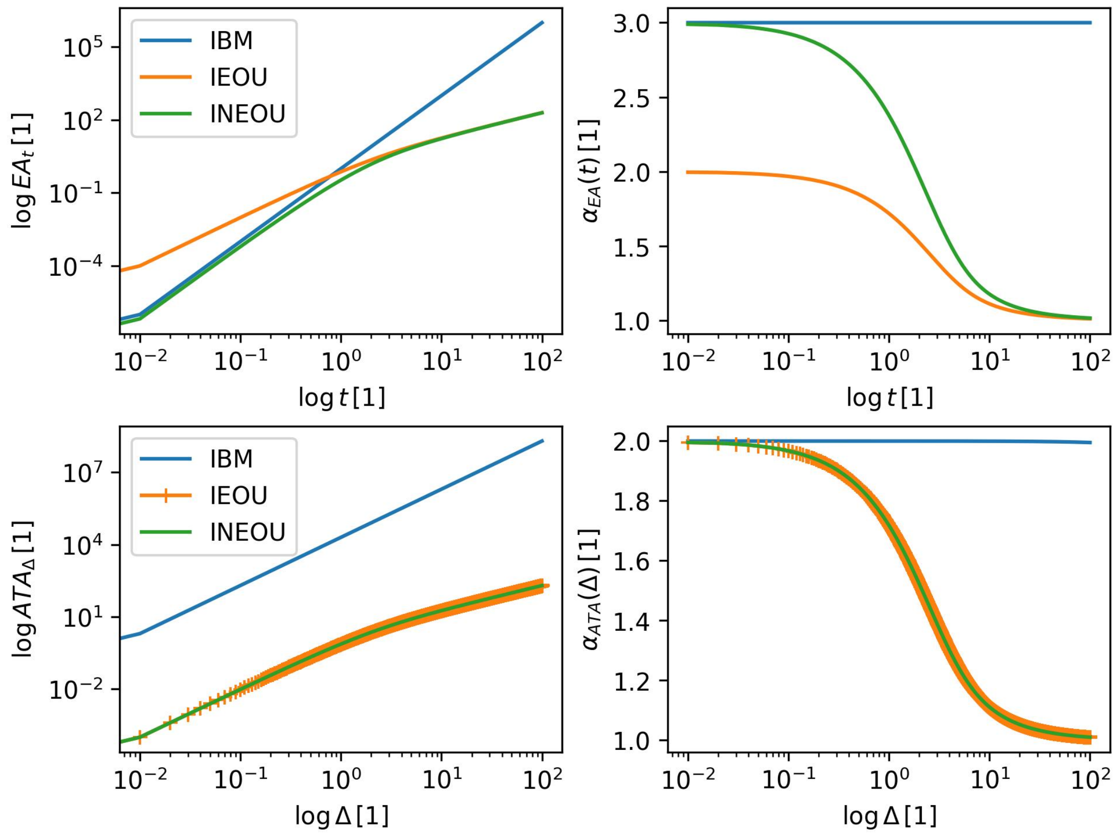

Figure 4.5: EAMSD / ATAMSD and scaling exponents for the integrated Brownian motion, integrated EOU process and integrated NEOU process. Notice that the EAMSD of the INEOU process is initially identical to IBM, but converges at later time with the MSD of the IEOU process. In the ATAMSD, both IEOU and INEOU processes are identical.

and the ATAMSD becomes just the EAMSD:

$$\langle \delta^2(\Delta, \Delta) \rangle = 2\left(\Delta - 1 + e^{-\Delta}\right) - \left(1 - e^{-\Delta}\right)^2$$

But if $\Delta \neq T$, the T-independent part is exactly the (ATA)MSD of the **integrated EOU process**! The initial superballisticity of the integrated NEOU process originated from the cancellation effects the $\left(1 - e^{-\Delta}\right)^2$-term introduced. But for $\Delta \ll T$ this cannot happen in the ATAMSD. Note that the T-dependent modulator term

$$\frac{1 - e^{-2(T-\Delta)}}{2(T - \Delta)} \tag{4.29}$$

is rapidly decreasing for large $T - \Delta$. The ATAMSD even has a finite T-limit

$$\lim_{T \to \infty} \langle \delta^2(\Delta, T) \rangle = 2\left(\Delta - 1 + e^{-\Delta}\right), \tag{4.30}$$

which is the MSD of the integrated EOU process.

The EAMSD and ATAMSD are shown in Fig. 4.4. Both MSDs have a stable initial scaling until roughly $r \sim 10^{-1}$. Between $r \sim 10^{-1}$ and $r \sim 5 \cdot 10^{1}$, an equilibriation period leads to a drastic decay of the scaling exponent, which saturates to the asymptotic diffusive $\alpha \sim 1$ afterwards. Note that even though both MSDs agree after $r \sim 10^{1}$ and their scaling as well, the initial scaling exponents are wildly differing. Th EAMSD of the INEOU process agrees initially with the EAMSD of integrated Brownian motion.

It is not easy to answer why they scale differently in the initial regime, given that they agree in the asymptotic regime. One thing to note however is the structure of the ACF of the NEOU process:

$$C_v(r,s) = e^{-|r-s|} - e^{-(r+s)} = C_v^{eq}(r,s) - e^{-r}e^{-r} \tag{4.31}$$

It can be decomposed in the the ACF of the EOU process, $C_v^{eq}(r,s) = e^{-|r-s|}$, and a **factorizable** non - stationary contribution

$$-e^{-r}e^{-s}.$$

Remember that the main diagonal of the ACF $C_v(r,r) = \langle v_r^2 \rangle$ forms the velocity SMF?

For the ACF of the NEOU process, the non - stationary extra term leads to an effective cancellation at small times,

$$C_v(r,r) = e^{-|r-r|} - e^{-2r} = 1 - e^{-2r} \sim 2r$$

and an effective linear SMF scaling for the initial regime. This leads to the cubic scaling of the EAMSD. As the non - stationary term is decaying for larger times, it's impact on the EAMSD scaling becomes weaker over time. This leads to the asymptotically equivalent EAMSD scaling of the EOU and NEOU process.

The non - stationary $e^{-(r+s)}$ does **not** contribute to the ATAMSD scaling, however: The corresponding part of the ATAMSD,

$$\frac{(1-e^{-\Delta})^2}{2}\frac{1-e^{-2(T-\Delta)}}{T-\Delta},$$

is monotonically decreasing in T, therefore

$$\lim_{T\to\infty} \frac{(1-e^{-\Delta})^2}{2}\frac{1-e^{-2(T-\Delta)}}{T-\Delta} = 0 \tag{4.32}$$

and the the non - stationary part **doesn't** influence the ATAMSD scaling.

I will demonstrate in Ch. 8 that such a disparity between EAMSD and ATAMSD scaling appears for a wider range of velocity processes. While non - stationary terms in the ACF lead to a superballistic EAMSD scaling, the corresponding terms in the ATAMSD vanish or become insignificant up to leading order in the large T limit. This corresponds to a weaker lagtime scaling of the ATAMSD.

4.4 What to take away?

None of the examples with superballistic EAMSD scaling had an ATAMSD scaling exponent greater than 2. The same holds for many other processes in the literature (see

[18] for a comparison). There seems to be some *saturation* mechanism that restricts the ATAMSD to scale at most ballistically.

The remaining chapters of this thesis, beginning with Ch. 5, are devowed to explore this apparent *saturation* of the ATAMSD. This will culminate in the **split scaling hypothesis** in Sec. 8.4, which gives a partial answer to this dilemma.

5 Interlude I: Asymptotic analysis and initial scaling

This interlude is the first of two chapters purely devoted to the machinery and theoretical predictions on **scaling exponents**. Here I will explore what can be said about the **initial scaling** of the TA- and EAMSD scaling.

In the **first** part I investigate what can be said about the initial TAMSD scaling exponent $\alpha_{TA}(0)$. The value of $\alpha_{TA}(0)$ is determined by the **pathwise** regularity of the process trajectory x_t, e.g. differentiability or (limital) Hölder continuity. In the **second** part, the initial EAMSD scaling exponent $\alpha_{EA}(0)$ is explored. Although no unique relationship can be established there, the scaling exponent can be related to the initial scaling exponent $\alpha^{v}_{SMF}(0)$ of the velocity process SMF $SMF_t = \langle v_t^2 \rangle$.[1]

5.1 Initial TAMSD scaling exponent $\alpha_{TA}(0)$

The TAMSD

$$\delta^2(\Delta, T) = \frac{1}{T-\Delta} \int_0^{T-\Delta} (X_{t+\Delta} - X_t)^2 \, dt$$

is defined pathwise and is itself **stochastic** for most stochastic processes. If the TAMSD is path - dependent, the scaling should depend on the underlying properties of the path. For the initial TAMSD scaling exponent

$$\alpha_{TA}(0) := \lim_{\Delta \to 0} \alpha_{TA}(\Delta, T), \tag{5.1}$$

the pathwise property that matters is the **pathwise regularity**.

There are two scaling laws that I could establish analytically: First for jump processes and then for differentiable processes. This corresponds to piecewise - constant and C^1 - trajectories. A third partial result for Hölder continuous processes allows to estimate the scaling exponent from below.

---

[1] This chapter relies heavily on the equality $\alpha_f(0) = \underline{\alpha}_f$ between the initial scaling exponent and the lower index. The notation and theory behind those has been introduced in Sec. 3.3.

Jump processes

Theorem 4. *Let r_t be a piecewise - constant trajectory with* ***locally finite jumps***[2] *and representation*

$$x_t = \sum_{i=0}^{N_t} r_i, \tag{5.2}$$

where $r_i \in \mathbb{R}^n$ is the i-th jump and N_t the latest jump index at time t. Then the TAMSD scales initially ***diffusively***:

$$\alpha_{TA}(0) = 1 \tag{5.3}$$

Proof. Fix an observational time $T > 0$ and let

$$J := \{ t_j \ : \ j \in 0, ..., N_T \}$$

be the set of jumping times in the interval $[0, T]$. There are only finitely many jumps until T, so there exists a $\Delta_0 > 0$ such that

$$[t_i - \Delta_0, t_i + \Delta_0] \cap [t_j - \Delta_0, t_j + \Delta_0] = \emptyset$$

for any two jumping times t_i, t_j.

The increment vanishes if there is no jump in $(t, t+\Delta]$.[3] If there is a single jumping time $t_j \in (t, t + \Delta]$, the increment is given by

$$x_{t+\Delta} - x_t = |r_j|^2.$$

But if we choose $\Delta < \Delta_0$, there can be at most one jump in any interval $(t, t + \Delta]$ and $x_{t+\Delta} - x_t = |r_i|^2 \neq 0$ for a jump r_i if and only if $t \in [t_i - \Delta, t_i)$. This allows to explicitly evaluate the integral in the TAMSD:

$$\begin{aligned}\int_0^{T-\Delta} |x_{t+\Delta} - x_t|^2 \, dt &\overset{!}{=} \int_0^{T} |x_{t+\Delta} - x_t|^2 \, dt = \sum_{j\in J} \int_{t_j-\Delta}^{t_j} |x_{t+\Delta} - x_t|^2 \, dt \\ &= \sum_{j\in J} \int_{t_j-\Delta}^{t_j} |r_j|^2 \, dt = \Delta \cdot \sum_{j\in J} |r_j|^2\end{aligned}$$

Rearranging the TAMSD into

$$\frac{\delta^2(\Delta, T)}{\Delta} = \frac{1}{T-\Delta} \int_0^{T-\Delta} \frac{|x_{t+\Delta} - x_t|^2}{\Delta} \, dt = \frac{\sum_{j\in J} |r_j|^2}{T-\Delta}$$

gives the limit

$$\lim_{\Delta\to 0} \frac{\delta^2(\Delta, T)}{\Delta} = \frac{\sum_{j\in J} |r_j|^2}{T}, \tag{5.4}$$

which is finite and non-zero for at least one jump. The lower index has to be $\underline{\alpha}_{TA} = 1$ and the claim follows due to Thm. 2. is proven. $\square$

---

[2] Only finitely many jumps are possible in a finite time interval.

[3] Whenever N_t is constant on $]t, t + \Delta]$.

This result has been established for pathwise - constant trajectories with locally finite jumps. This entails the majority of studied jump processes, like ordinary CTRWs (see [16]) or heavy - tailed Levy walks (see [28]). It is independent of the jump length and waiting time distributions, so it is **universal** for the class of jump processes. This is in contrast to the initial EA-MSD scaling exponent, which can differ wildly: α_{EA} depends on the jump and waiting time distributions of the process.

Differentiable processes

Theorem 5. *Let $r \in \mathcal{C}^1([0,T^*],\mathbb{R}^n)$ be a continuously differentiable trajectory, which is not entirely constant on $[0,T]$ for some observational time $T > 0$. Then the TAMSD scales initially ballistic:*

$$\alpha_{TA}(0) = 2 \tag{5.5}$$

Proof. I will prove the following limit:

$$\lim_{\Delta\to 0}\frac{\delta^2(\Delta,T)}{\Delta^2} = \int_0^T |r'(t)|\, dt \tag{5.6}$$

Since the limit is finite if $r'(t) \neq 0$ for at least one time t, Thm. 2 proves the ballistic TA-MSD scaling.

Let $\Delta > 0$ and $T > T^* - \Delta$. The quotient of the TA-MSD is given by

$$\frac{\delta^2(\Delta,T)}{\Delta^2} = \underbrace{\int_0^T \left|\frac{r_{t+\Delta}-r_t}{\Delta}\right|^2 dt}_{a)} + \underbrace{\int_{T-\Delta}^T \left|\frac{r_{t+\Delta}-r_t}{\Delta}\right|^2 dt}_{b)}.$$

Define for each $\Delta > 0$ the following function w_Δ, which appears inside each integral:

$$w_\Delta(t) = \left|\frac{r_{t+\Delta}-r_t}{\Delta}\right|^2$$

For each t this collection of functions converges pointwise to the squared derivative

$$w_\Delta(t) \overset{\Delta\to 0}{\to} \left|r'_t\right|^2.$$

But since we integrate over these functions, pointwise convergence does not suffice. Note that that the first integral converges against the desired result and the integral in $b)$ should vanish, as the integration domain shrinks.

The first integral can be handled using the dominated convergence theorem. Since a pointwise limit has been established, it suffices to find an integrable dominating function, i.e. a function w such that $|w_\Delta(t)| \leq w(t)$ for each $\Delta > 0$ and time t.

Due to the mean value theorem the difference can be expressed as

$$\frac{r_{t+\Delta}-r_t}{\Delta} = r'(\zeta)$$

where $\zeta \in [0, t+\Delta]$. For any finite observation time $T > 0$, the bound

$$w_\Delta(t) = \left|\frac{r_{t+\Delta}-r_t}{\Delta}\right|^2 = |r'(\zeta)|^2 \leq \max_{s\in[0,T]} |r'_s|^2$$

is true and $w(t) := \max_{s\in[0,T]} |r'_s|^2$ is an integrable dominating function. Therefore

$$\lim_{\Delta\to 0} \int_0^T \left|\frac{r_{t+\Delta} - r_t}{\Delta}\right|^2 dt = \int_0^T |r'(t)|^2\, dt.$$

Likewise, the second integral can be bounded using the estimate via

$$\int_{T-\Delta}^T \left|\frac{r_{t+\Delta} - r_t}{\Delta}\right|^2 dt \leq \max_{s\in[0,T]} |r'_s|^2 \Delta \overset{\Delta\to 0}{\to} 0.$$

This proves the limit in eq. 5.6. Since this limit is finite and non - zero, Thm. 2 implies

$$\alpha_{TA}(0) = \underline{\alpha}_{TA} = 2.$$

□

The initial TAMSD scaling exponent of a differentiable process has to be $\alpha_{TA}(0) = 2$, so they necessary scale ballistic for small times. This explains the initial TAMSD scaling results obtained in Ch. 4, where always $\alpha_{TA}(0) = 2$ despite the initial EAMSD scaling being superballistic.

Although Thm. 8 will establish a constraint on the initial EAMSD scaling as well, this constraint is much more lenient and broad enough to allow a rich palate of initial EAMSD scaling exponents.

If the initial TAMSD scaling is always ballistic, a superballistic TAMSD scaling can only be possible during intermediary or asymptotic regimes.

Hölder continuous processes

The last class of processes I will consider here are **Hölder processes**.

Definition 23. *Let $f_t : [0, T] \to \mathbb{R}^n$ be a trajectory. Then f is called* ***Hölder continuous*** *with* ***Hölder exponent*** $0 < \beta < 1$ *($f \in \mathcal{C}^\beta$) if the norm*

$$||f_t||_\alpha := \sup_{s\neq t\in[0,T]} \frac{|f_t - f_s|}{|t-s|^\beta} < \infty \tag{5.7}$$

is finite. In this case there exists a $C > 0$ such that

$$|f(t) - f(s)| \leq C|t-s|^\beta \tag{5.8}$$

for all $r, s \in [0, T]$. The ***limit Hölder space*** *C^{β^-} is defined as all trajectories f, such that $f \in \mathcal{C}^\gamma$ is Hölder for $0 < \gamma < \beta$, but not Hölder $f \notin \mathcal{C}^\gamma$ for any $\gamma > \beta$.*

The definition of Hölder continuity can be seen as a weaker form of differentiability / Lipschitz continuity. The theory of Hölder functions shares many intersections with the theory of fractals and self-similar processes, most importantly via the **Kolmogorov - Chentsov theorem**.

A notable example is **fractional Brownian motion** B^H with **Hurst parameter** H, which is Hölder continuous for any $\beta < 2H$, but not for $\beta = 2H$ or higher. The increment is stationary distributed as

$$\langle (B_{t+\Delta} - B_t)^2 \rangle = \Delta^{2H}. \tag{5.9}$$

Theorem 6. *Let x_t be a stochastic process. If $x_t \in \mathcal{C}^H$ pathwise, i.e. x_t is* ***pathwise Hölder continuous*** *with*

$$|x_{t+s} - x_t| \leq Cs^H \tag{5.10}$$

for some $C > 0$, then

$$\alpha_{TA}(0) \geq 2H \tag{5.11}$$

Proof. With Thm. 2 it suffices to show

$$\lim_{\Delta\to 0} \frac{\delta^2(\Delta, T)}{\Delta^\beta} = 0$$

for all $\beta < 2H$.

The quotient can be bound using the Hölder condition

$$\begin{aligned} \frac{\delta^2}{\Delta^\beta} &= \frac{1}{T-\Delta}\int_0^{T-\Delta} \frac{(x_{t+\Delta}-x_t)^2}{\Delta^\beta}\, dt \\ &\leq \frac{1}{T-\Delta}\int_0^{T-\Delta} \frac{C^2\Delta^{2H}}{\Delta^\beta}\, dt \\ &= C^2\Delta^{2H-\beta}. \end{aligned}$$

Since δ^2 is non-negative and $2H - \beta > 0$ this implies

$$0 \leq \lim_{\Delta\to 0} \frac{\delta^2(\Delta, T)}{\Delta^\beta} \leq C^2 \lim_{\Delta\to 0} C^2\Delta^{2H-\beta} = 0 \tag{5.12}$$

and the claim is proven. ☐

For general Hölder continuous processes / trajectories, only a lower bound on the initial TAMSD scaling exponent can be established. The main obstacle in giving an upper bound comes from the Hölder condition itself:

$$|x_{t+\Delta} - x_t|^2 \leq C\Delta^{2H} \tag{5.13}$$

The integrand of δ^2 can bound from above by this, which helps proving convergence. But for divergence, a diverging lower bound akin to

$$|x_{t+\Delta} - x_t|^2 \geq C\Delta^{2H}$$

is required. Violating Hölder continuity implies the existence of some t_0 and a sequence $\Delta_n \to 0$ such that

$$\frac{|x_{t_0+\Delta_n} - x_{t_0}|}{\Delta^{2H}} \to \infty. \tag{5.14}$$

But this is not sufficient, as the divergence of the increment **only at** t_0 is guaranteed. For the integral in δ^2 this does not suffice and only the divergence on some larger set is sufficient. This is the assumption the reverse direction requires:

Theorem 7. *Let x_t be a stochastic process with $x_t \notin \mathfrak{C}^H$, i.e. x_t is not Hölder continuous with exponent H. Additionally, assume that there exists some $I > 0$ and $C > 0$ such that for each $\Delta < \Delta_0$ sufficiently small the set*

$$D_\Delta := \left\{ t \in [0,T] \ : \ |x_{t+\Delta} - x_t| > C\Delta^H \right\} \tag{5.15}$$

has measure $Vol(D_\Delta) = I$.
Then

$$\alpha_{TA}(0) \leq 2H \tag{5.16}$$

Proof. The fraction $\frac{\delta^2}{\Delta^\beta}$ can be bounded from below by

$$\begin{aligned} \frac{\delta^2(\Delta,T)}{\Delta^\beta} &= \frac{1}{T-\Delta}\int_{[0,T-\Delta]} \frac{(x_{t+\Delta}-x_t)^2}{\Delta^\beta}\,dt \\ &\geq \frac{1}{T-\Delta}\int_{D_\Delta} \frac{(x_{t+\Delta}-x_t)^2}{\Delta^\beta}\,dt \geq \frac{1}{T-\Delta}\int_{D_\Delta} C^2\Delta^{2H-\beta}\,dt \\ &= \frac{C^2 I}{T-\Delta}\Delta^{2H-\beta} \geq \frac{C^2 I}{T}\Delta^{2H-\beta} \end{aligned}$$

The limit diverges for $\beta > 2H$ due to

$$\frac{\delta^2(\Delta,T)}{\Delta^\beta} \geq \frac{C^2 I}{T}\lim_{\Delta\to 0}\Delta^{2H-\beta} = \infty$$

and the upper bound follows:

$$2H \geq \underline{\alpha}_{TA} = \alpha_{TA}(0). \tag{5.17}$$

□

5.2 Initial EAMSD scaling exponent $\alpha_{EA}(0)$

The initial EAMSD scaling of differentiable processes is heavily tied to the **velocity autocorrelation function (VACF)**. A detailed discussion of the VACF can be found in Ch. 6.

In contrast to the TAMSD, no general relationship for the initial EAMSD scaling exponent is known. The underlying reason for that might be that if x_t is a differentiable process with velocity v_t, there are two quantities that can quantify the spreading of v_t:

First we have the typical EAMSD $\langle (v_t - v_0)^2 \rangle$, but there is also the VACF $C_v(r,s) = \langle v_r v_s \rangle$. The main diagonal of this function plays an important role, the **second moment function (SMF)** $C_v(t,t) = \langle v_t^2 \rangle$. The SMF naturally appears in the VACF, which relates it to the EAMSD of x_t via the Green - Kubo relation. But the EAMSD of v_t only satisfies

$$\langle (v_t - v_0)^2 \rangle = C_v(t,t) + C_v(0,0) - 2C_v(0,t), \tag{5.18}$$

and there is no nice relation between the EAMSD of x_t and v_t.

Although no direct link between the initial EAMSD scaling $\alpha^x_{EA}(0)$ of x_t and the initial EAMSD scaling $\alpha^v_{EA}(0)$ of v_t can be established, there is a relation between $\alpha^x_{EA}(0)$ and the initial SMF scaling

$$\alpha^v_{SMF}(0) = \lim_{t\to 0} \frac{d \log\langle v_t^2\rangle}{d \log t} \tag{5.19}$$

of the velocity v_t, namely

$$\alpha^x_{EA}(0) \geq \alpha^v_{SMF}(0) + 2.$$

Heurististically, this relation appears naturally from Green-Kubo relation:[4]

$$\langle (x_t - x_0)^2\rangle = \int_0^t \int_0^t C_v(r,s)\, dr\, ds \tag{5.20}$$

The initial SMF of v_t will roughly scale as

$$\langle v_t^2\rangle \sim t^{\alpha_{SMF}(0)}$$

for $t \ll 1$. For such t, one can make the approximation $C_v(r,s) \sim C_v(t,t) = \langle v_t^2\rangle$ and the EAMSD reduces to

$$\begin{aligned}\langle (x_t - x_0)^2\rangle = \int_0^t \int_0^t C_v(r,s)\, ds\, dr \sim C_v(t,t) \int_0^t \int_0^t dr\, ds \\ \sim t^2 \langle v_t^2\rangle \sim t^{2+\alpha^v_{SMF}(0)}.\end{aligned}$$

This *suggests* $\alpha^x_{EA}(0) = \alpha^v_{SMF}(0) + 2$.

This argument is only heuristical and highly flawed, mathematically speaking. Cancellation effects in the EAMSD for small times can potentially lead to a violated relation

$$\alpha^x_{EA}(0) > \alpha^v_{SMF}(0) + 2.$$

But the lower bound can be proven under mild assumptions on v_t:

Theorem 8. *Let x_t be a differentiable stochastic process with velocity process v_t. Assume that the initial velocity SMF exponent $\alpha^v_{SMF}(0)$ is well-defined. Then*

$$\alpha^x_{EA}(0) \geq \alpha^v_{SMF}(0) + 2. \tag{5.21}$$

Proof. Using Theorem 2 it suffices to show that

$$\lim_{t\to 0} \frac{\partial_t \langle (x_t - x_0)^2\rangle}{t^{\gamma+1}} = 0$$

for every $\gamma < \alpha^v_{SMF}(0) + 2$.

Assume that the assumptions hold and let $\gamma < \alpha^v_{SMF}(0) + 2$. The derivative of the EAMSD is given by

$$\partial_t \langle (x_t - x_0)^2\rangle = 2 \int_0^t C_v(r,t)\, dt,$$

---

[4] The Green-Kubo relation is a classical relation in the field. Some references are [26] or [8].

but this is unfavorable for the moment, as the assumptions only concern the velocity SMF $C_v(r,r) = \langle v_r^2 \rangle$. Using the Cauchy - Schwartz inequality

$$C_v(r,s) \leq \sqrt{C_v(r,r)}\sqrt{C_v(s,s)} = \sqrt{\langle v_r^2 \rangle \langle v_s^2 \rangle},$$

the derivative can be bound from above via

$$\begin{aligned} \partial_t \langle (x_t - x_0)^2 \rangle &= 2\int_0^t C_v(r,t)\,dt \leq 2\sqrt{\langle v_t^2 \rangle}\int_0^t \sqrt{\langle v_r^2 \rangle}\,dr \\ &\leq 2t\sqrt{\langle v_t^2 \rangle}\max_{r\in[0,t]}\sqrt{\langle v_r^2 \rangle} \leq 2t\sqrt{\langle v_t^2 \rangle \max_{r\in[0,t]}\langle v_r^2 \rangle} \end{aligned}$$

The quotient can be bounded by

$$\frac{\partial_t \langle (x_t - x_0)^2 \rangle}{t^{\gamma+1}} \leq 2\sqrt{\frac{\langle v_t^2 \rangle}{t^\gamma}\frac{\max_{r\in[0,t]}\langle v_r^2 \rangle}{t^\gamma}}.$$

Since $\langle v_t^2 \rangle / t^\gamma \to 0$ for $t \to 0$ and v_t is continuous, the maximum also obeys $\max_{r\in[0,t]}\langle v_r^2 \rangle / t^\gamma \to 0$ and the limit follows. The initial quotient can be bound by

$$\left| \frac{\langle (x_t - x_0)^2 \rangle}{t^{2+\gamma}} \right| = \left| \frac{\int_0^t \partial_s \langle (x_s - x_0)^2 \rangle\,ds}{t^{2+\gamma}} \right| \leq \left| 2\sqrt{\frac{\langle v_t^2 \rangle}{t^\gamma}\frac{\max_{r\in[0,t]}\langle v_r^2 \rangle}{t^\gamma}} \right|$$

and the claim follows. □

This theorem relies crucially on the relationship between the VACF and EAMSD of x_t. There is no lower bound similar to the Cauchy - Schwartz inequality, which could help proving the divergence of the limit for $\gamma > \alpha^v_{SMF}(0) + 2$. Another approach could be to bound the VACF $C_v(r,s)$ in the integrand of $\partial_t \langle (x_t - x_0)^2 \rangle$ using higher derivatives for differentiable / analytical VACFs, but I could not write a valid proof in this case.

There is, however, another and direct proof possible for **stationary** VACFs:

Theorem 9. *Let x_t be a differentiable process with* ***stationary*** *velocity v_t, whose VACF is continuous and $\langle v_0^2 \rangle \neq 0$. Then $\alpha^v_{SMF}(t) = 0$ for all t and*

$$\alpha^x_{EA}(0) = 2 = \alpha^v_{SMF}(0) + 2 \tag{5.22}$$

Proof. First of, the EAMSD of x_t for stationary VACF $C(\tau)$ can be computed via the single integral

$$\langle (x_t - x_0)^2 \rangle = 2\int_0^t (t-\tau)C(\tau)\,d\tau$$

The first step is to prove that $\frac{\langle (x_t - x_0)^2 \rangle}{t^\gamma} \to 0$ for $\gamma < 2$:

A consequence of the CSI for stationary VACFs is the global bound $|C(\tau)| \leq C(0)$, which will be established in eq. 6.9. Using this inequality the quotient can be estimated as

$$
\begin{aligned}
0 \leq \left| \frac{\langle (x_t - x_0)^2 \rangle}{t^\gamma} \right| &= 2 \left| \int_0^t \frac{(t-\tau)}{t^\gamma} C(\tau)\, d\tau \right| \\
&\overset{|C(\tau)| \leq C(0)}{\leq} \frac{2C(0)}{t^\gamma} \int_0^t (t-\tau)\, d\tau \ = C(0) t^{2-\gamma}
\end{aligned}
$$

If $\gamma < 2$ the upper bound vanishes in the limit, establishing $\frac{\langle (x_t - x_0)^2 \rangle}{t^\gamma} \to 0$ for $\gamma < 2$.

Now let $\gamma = 2$. Since $\langle v_0^2 \rangle \neq 0$, $C(0) > 0$ and there exists an $\zeta \in (0, r)$ such that C is positive on $[0, \zeta]$. In this case one can bound

$$
\min_{r \in [0,\zeta]} C(r) \leq C(\tau).
$$

This gives the reverse bound

$$
\begin{aligned}
\min_{r \in [0,\zeta]} C(r) t^{2-\gamma} &= 2 \int_0^t \frac{(t-\tau)}{t^\gamma} \min_{r \in [0,\zeta]} C(r)\, d\tau \\
&\leq 2 \int_0^t \frac{(t-\tau)}{t^\gamma} C(\tau)\, d\tau = \frac{\langle (x_t - x_0)^2 \rangle}{t^\gamma}
\end{aligned}
$$

Now if $\gamma > 2$ the lower bound blows up and $\alpha_{EA}^x(0) = 2$. Since the velocity SMF $\langle v_t^2 \rangle = \langle v_0^2 \rangle = \langle v_0^2 \rangle t^0$ is constant, $\alpha_{SMF}^v(0) = 0$ and the relation holds. □

One consequence of the proof in Thm. 9 is the global bound

$$
\min_{r \in [0,t]} C(r) t^2 \leq \langle (x_t - x_0)^2 \rangle \leq C(0) t^2 \tag{5.23}
$$

for **positive** VACF $C(\tau) \geq 0$. This estimate has the consequence that a process x_t with stationary VACF **can not** scale asymptotically superballistically. This is discussed later on in Thm. 17.

6 Interlude II: Velocity autocorrelation functions and asymptotic scaling

The first interlude has been devoted to studying the initial scaling of both TAMSD and EAMSD. Constraints on the initial scaling exponents have been proven, which relied heavily on the relation between the initial scaling exponent $\alpha_f(0)$ and the lower index $\underline{\alpha}_f$ Theorem 2. This theorem allows to estimate the scaling exponents by determining the limit behavior of

$$\lim_{r\to 0} \frac{f(r)}{r^\gamma}$$

The initial EAMSD scaling exponent for differentiable x_t has been shown to obey

$$\alpha^x_{EA}(0) \geq \alpha^v_{SMF}(0) + 2.$$

The EAMSD scaling for small times is thus largely determined by the underlying velocity process in form of the velocity SMF $\langle v_t^2 \rangle$. In the asymptotic regime, the scaling is also determined by the velocity process, but the velocity SMF does not suffice for completely determining the scaling exponent. Instead the complete **velocity autocorrelation function (VACF)**

$$C_v(r,s) = \langle v_r v_s \rangle \tag{6.1}$$

needs to be utilized. This chapter will study how the VACF can be used to estimate the scaling exponent during **stable regimes**, mainly the asymptotic scaling exponent $\alpha_f(\infty)$.

6.1 Velocity Autocorrelation Function (VACF)

A more detailed study of the VACF is found in [22].

Definition 24. *Let x_t be a differentiable process with velocity v_t. The* ***velocity autocorrelation function (VACF)*** *$C_v(r,s)$ is the bivariate function defined by*

$$C_v(r,s) = \langle v_r \cdot v_s \rangle \tag{6.2}$$

In case that the VACF only depends on the time difference,

$$C_v(r,s) = C(r-s) \tag{6.3}$$

for some symmetric C, C_v will be called ***stationary****.*

The velocity SMF $\langle v_t^2 \rangle$ appears as the main diagonal of the VACF:

$$C_v(r,r) = \langle v_r^2 \rangle \tag{6.4}$$

In contrast to general bivariate functions, the VACF has to be **positive-semidefinite**:

Positive-Semidefiniteness

Let v_t be a fixed (velocity) process, $r_1, ..., r_n$ a collection of times and $\alpha_1, ..., \alpha_n$ arbitrary real numbers. Then the linear combination

$$w := \sum_{i=1}^{n} \alpha_i v_{r_i} \tag{6.5}$$

forms a random variable. Since v_t is assumed to have a finite SMF, i.e. $\langle v_t^2 \rangle < \infty$ for every $t \geq 0$, the resulting variable w has finite SM $\langle w^2 \rangle < \infty$. The SM $\langle w^2 \rangle \geq 0$ is always non - negative, and it can be expressed using the ACF of v_t:

$$\begin{aligned} 0 \leq \langle w^2 \rangle &= \left\langle \sum_{i=1}^{n} \alpha_i v_i, \sum_{i=1}^{n} \alpha_i v_i \right\rangle \\ &= \sum_{i,j} \alpha_i \alpha_j \langle v_{r_i}, v_{r_j} \rangle = \sum_{i,j} \alpha_i \alpha_j C_v(r_i, r_j) \\ &=: \sum_{i,j} \alpha_i \alpha_j C_{i,j} \end{aligned}$$

The linear combination $\sum_{i,j} \alpha_i \alpha_j C_{i,j} \geq 0$ has to be non-negative. But since w is arbitrary and r_i, α_i as well, this has to hold for any such choices. This condition is called **positive-semidefiniteness**.

Definition 25. *Let $K(r,s)$ be a bivariate kernel. Then K is called **positive-semidefinite (PSD)** iff for any collection of times $r_1, ..., r_n$ and real coefficients $\alpha_1, ..., \alpha_n$ the following inequality holds:*

$$\sum_{i,j=1}^{n} \alpha_i \alpha_j K(r_i, r_j) \geq 0 \tag{6.6}$$

For stationary VACF, the PSD condition can be stated in terms of $C(\tau)$:

Definition 26. *$C(\tau)$ is **positive - semidefinite** iff for any collection of times $r_1, ..., r_n$ and real coefficients $\alpha_1, ..., \alpha_n$ the inequality*

$$\sum_{i,j=1}^{n} \alpha_i \alpha_j C(r_i - r_j) \geq 0 \tag{6.7}$$

holds true.

The PSD condition severely restricts the class of possible VACFs, but it is crucial condition for many theorems in this thesis. Most of the results in Ch. 5 fail if PSD is violated.

Cauchy-Schwartz inequality

A consequence of the bilinearity of the inner product $\langle \cdot, \cdot \rangle$ and the PSD of $C_v(r,s) = \langle v_r, v_s \rangle$ is the **Cauchy-Schwartz inequality (CSI)**:

$$C_v(r,s) \leq \sqrt{C_v(r,r)}\sqrt{C_v(s,s)} \tag{6.8}$$

The CSI for stationary VACF $C_v(r,s) = C(r-s)$ leads to a global bound

$$|C(\tau)| \leq C(0) \tag{6.9}$$

due to the constant SMF $C_v(r,r) = C_v(s,s) = C(0)$. Stationary VACF are therefore **globally** bounded by the initial SM $C(0) = \langle v_0^2 \rangle$.

VACF and quadratic functionals

The process x_t can be expressed using its velocity via $x_t = x_0 + \int_0^t v_r \, ds$. Since any quadratic difference

$$(x_{t+\Delta} - x_t)^2 = \int_t^{t+\Delta} \int_t^{t+\Delta} v_r v_s \, ds$$

involves only the inner product of the velocity, the average can be computed using the VACF. The averaged increment

$$A_{t,\Delta} = \langle (x_{t+\Delta} - x_t)^2 \rangle \tag{6.10}$$

is called **averaged lagtime increment (ALI)**. The ALI can be expressed via the VACF as follows:

$$\begin{aligned} A_{t,\Delta} &= \langle (x_{t+\Delta} - x_t)^2 \rangle = \left\langle \left(\int_t^{t+\Delta} v_s \, ds \right)^2 \right\rangle \\ &= \int_t^{t+\Delta} \int_t^{t+\Delta} \langle v_s v_r \rangle \, ds \, dr = \int_t^{t+\Delta} \int_t^{t+\Delta} C_v(r,s) \, ds \, dr \\ &= 2 \int_t^{t+\Delta} \int_r^{t+\Delta} C_v(r,s) \, ds \, dr \end{aligned}$$

The ALI appears in both the EAMSD and ATAMSD, which allows to calculate both MSDs via the VACF alone. This does not work for the TAMSD, as it is define pathwise, so I will instead study the ATAMSD. Under mild conditions, all the results about the ATAMSD hold almost surely for the TAMSD as well in the large T limit.

6.2 Scaling via derivatives

One important theme in this thesis is the scaling exponent estimation of a growth function f via the scaling exponents of higher derivatives $f^{(n)}$. The heuristical idea is this:

Let $f(r) = r^\alpha$ be a strict power-law and the exponent $\alpha \notin \mathbb{N}_0$ not an integer. The higher derivatives are formal given by

$$f^{(n)}(r) = r^{\alpha-n} \prod_{i=0}^{n-1} \alpha - i, \tag{6.11}$$

which are power-laws themselves. The exponents of f and $f^{(n)}$ satisfy

$$\alpha_{f^{(n)}}(r) = \alpha - n = \alpha_f(r) - n. \tag{6.12}$$

Taking derivatives decreases the order of a monomial by one, so this scaling relation should not be surprising.[1]

If a growth function scales locally as a power-law $f(r) \sim r^\alpha$, one would expect a similar relation. But this is not always true: If f is a generic growth function

$$f(r) = r^{\alpha_f(r)}$$

with non-constant scaling exponent α_f, the first derivative of f is not a strict power-law:

$$f'(r) = \partial_r \alpha_f \log r r^{\alpha_f(r)} + \alpha r^{\alpha_f(r)-1}$$

This still looks manageable, but already the second derivative

$$\begin{aligned} f''(r) =& r^{\alpha_f(r)} \cdot \left[(\alpha_f(r) + \partial_r \alpha_f(r) \log r)\partial_r \alpha_f(r) \log r + \partial_r^2 \alpha_f(r) + \log r\right] \\ &+ r^{\alpha_f(r)-1} \cdot \partial_r \alpha_f(r)(1 + \alpha_f(r) \log r) \\ &+ r^{\alpha_f(r)-2} \alpha_f(r)(\alpha_f(r) - 1) \end{aligned}$$

is too convoluted and a local power-law nature seems unlikely. Given that f' and f'' are composed of multiple terms, the scaling exponents of f' and f'' should not necessarily be related to α_f. But it can work during **stable regimes**:

The derivatives f' and f'' will be highly fluctuating if f scales transiently. But for **stable scaling regimes**[2] the scaling exponent $\alpha_f(r)$ is roughly constant and the derivative $\partial_r \alpha_f(r) \sim 0$ **vanishes**. If this stable approximation is valid, the first derivative reduces to

$$f'(r) \sim \alpha r^{\alpha_f(r)-1}. \tag{6.13}$$

For stable scaling regimes, where f scales roughly as a power-law, the **first derivative** will scale as a power - law as well and the important relation

$$\alpha_{f'}(r) \sim \alpha_f(r) - 1 \tag{6.14}$$

remains valid. This is also mathematically justified:

---

[1] Remember that $\alpha \notin \mathbb{N}_0$. For integer α, the higher derivatives can vanish.

[2] Only the scaling exponents during stable regimes are useful to us, remember. So this is in fact not a real restriction here.

Theorem 10. *Let f be a twice-differentiable growth function, such that f' and f'' are well -defined. Let f and f' be non - zero and assume that for time r_0*

$$\partial_r \alpha_f(r_0) = 0. \tag{6.15}$$

Then

$$\alpha_f(r_0) = \alpha_{f'}(r_0) + 1. \tag{6.16}$$

Proof. The difference $d(r) = \alpha_f(r) - \alpha_{f'}(r)$ can be expressed as

$$\begin{aligned}\alpha_f - \alpha_{f'} &= \frac{rf'}{f} - \frac{rf''}{f'} = r\left(\frac{f'}{f} - \frac{f''}{f'}\right) \\ &= r\left(\frac{(f')^2 - ff''}{ff'}\right) = r\left(\frac{f'}{f} - \frac{f''}{f'}\right)\end{aligned}$$

Since

$$\partial_r \alpha_f = \frac{f'}{f} + r\left(\frac{f''}{f} - \frac{(f')^2}{f^2}\right)$$

and

$$\frac{\partial_r \alpha_f}{\alpha_f} = \frac{1}{r} + \left(\frac{f''}{f'} - \frac{f'}{f}\right),$$

the difference $d(r)$ can be written as

$$d(r) = r\left(\frac{f'}{f} - \frac{f''}{f'}\right) = 1 - \frac{r\partial_r \alpha_f}{\alpha_f} = 1 - r\partial_r \log \alpha_f. \tag{6.17}$$

Since $\partial_r \alpha_f(r_0) = 0$, the difference is $g(r_0) = 1$ and the claim

$$\alpha_f(r_0) = \alpha_{f'}(r_0) + g(r_0) = \alpha_{f'}(r_0) + 1$$

follows. □

If used with care, this heuristical argument allows to calculate scaling exponents of growth functions via their higher derivatives. This is advantageous in case of the EAMSD and ATAMSD, which can be expressed via integrals of the VACF.

The higher derivatives of both MSDs are conceptually more clearly related to the VACF, which makes scaling arguments on these derivatives quite easy. But this tactic does not always work:

Careful counter - example

Thm. 10 only holds at r_0 if $\partial_r \alpha_f(r_0) = 0$, but the trick of computing the scaling exponent of f' instead of f works in general as a heuristical argument. One has to be careful, however:

Take the example of $f(r) = r + (1+r)^{1-\alpha}$ for $0 < \alpha < 1$. The first derivative

$$f'(r) = 1 + (1-\alpha)(1+r)^{-\alpha} = r^0 + (1-\alpha)(1+r)^{-\alpha}$$

still predicts the correct asymptotic scaling, since $f'(r) \sim r^0 = r^{\alpha_f - 1}$. But the second derivative

$$f''(r) = -\alpha(1-\alpha)(1+r)^{-(\alpha+1)}$$

fails horribly, as

$$\alpha_{f''} + 2 = 1 - \alpha$$

would suggest a **sublinear growth**. Why does the scaling estimation fail?

The problem lies in f, as it contains more than one monomial: The linear r and the sublinear $(1+r)^{1-\alpha} \sim r^{1-\alpha}$.

The linear term $r = r^1$ has an **integer** power and will vanish for higher-order derivatives. The linear term disappears from the second derivative f'' onwards, but the sublinear term **still** remains.

This happens for diffusive growth functions at the second derivative, so one is advised to always check the first and second derivative. Otherwise, a wrong scaling behavior may be deduced.

6.3 EAMSD via VACF

The EAMSD

$$EA_t = \langle (x_t - x_0)^2 \rangle$$

is given by the ALI

$$\langle (x_t - x_0)^2 \rangle = A_{0,t}.$$

The ALI can be expanded using the VACF and the EAMSD expressed as a double integral:

$$EA_t = \langle (x_t - x_0)^2 \rangle = \int_0^t \int_0^t C_v(r,s)\, dr\, ds \tag{6.18}$$

The structure of this formula is quite simple, as the integration domain over C_v is just the plain square $[0,t]^2$ for each time t.

Before I investigate the properties of eq. 6.18 and its derivatives in more detail, a few words about the discretization in terms of the VACF is necessary.

Discretization

The integral formula

$$\langle (x_t - x_0)^2 \rangle = \int_0^t \int_0^t C_v(r,s)\, dr\, ds$$

can be naively discretized as

$$EA_n = dt^2 \sum_{i,j=0}^{n-1} C_{ij} \tag{6.19}$$

Here $\Omega = \{\, t_i = i \cdot dt \ : \ i = 0, \ldots, n \,\}$ is an equidistant grid of mesh size $dt = 1/n$ and $C_{ij} = C_v(r_i, r_j)$ the matrix coefficients on this grid.

This naive discretization requires the computation of n^2 matrix coefficients per step, which would result in a total cubic complexity $\mathcal{O}(n^3)$ for the full EAMSD from 0 to t. This not an advantageous complexity, especially when the asymptotic scaling of the EAMSD needs to be studied. The computational problem in eq. 6.19 comes from the repeated computation of the matrix coefficients C_{ij}:

The summation domain at step $n-1$ is the square $R_n = \{(i,j) \ : \ i,j = 0,...,n-2\}$, which has a recursive structure $R_n = R_{n-1} \cup D_n$, where

$$D_n = \{(n-1,n-1)\} \cup \{(n-1,i) \ : \ i = 0,...,n-2\} \cup \{(i,n-1) \ : \ i = 0,...,n-2\}.$$

This recursive splitting of R_n leads to a recursion relation for EA_n:

Theorem 11. *The discretized EAMSD EA_n satisfies the recursion relation*

$$EA_n = EA_{n-1} + \mathcal{Q}_n \tag{6.20}$$

with increment

$$\mathcal{Q}_n = 2(dt)^2\left(C_{n-1,n-1} + \sum_{i=0}^{n-2} C_{i,n-1}\right). \tag{6.21}$$

Proof. The sum $\sum_{i,j=0}^{n-1} C_{ij}$ can be regrouped into

$$\sum_{i,j=0}^{n-1} C_{ij} = C_{n-1,n-1} + 2\sum_{i=0}^{n-2} C_{i,n-1} + \sum_{i,j=0}^{n-2} C_{ij}.$$

The total recursion relation follows simply:

$$\begin{aligned} EA_n = dt^2 \sum_{i,j=0}^{n-1} C_{ij} &= (dt)^2\left(C_{n-1,n-1} + 2\sum_{i=0}^{n-2} C_{i,n-1} + \sum_{i,j=0}^{n-2} C_{ij}\right) \\ &= EA_{n-1} + (dt)^2 C_{n-1,n-1} + 2(dt)^2 \sum_{i=0}^{n-2} C_{i,n-1} = EA_{n-1} + \mathcal{Q}_n. \end{aligned}$$

□

The computational complexity of the recursion relation is determined by$\mathcal{Q}_i$. Since $\mathcal{Q}_i$ requires the computation of i matrix coefficients, the total complexity for computing EA_i for $i = 1,...,n$ is **quadratic**:

$$\sum_{i=0}^{n} i = \frac{n(n+1)}{2} \sim n^2.$$

6.3.1 Derivatives

The EAMSD in terms of the VACF is given as

$$\langle (x_t - x_0)^2 \rangle = A_{0,t} = \int_0^t \int_0^t C_v(r,s)\, ds\, dr = \int_0^t f(r,t)\, dr$$

with $f(r,t) = \int_0^t C_v(r,s)\, ds$. Since the EAMSD takes the form of a parametric integral w.r.t. t in this setting, the higher derivatives can be computed using the Leibnitz formula.

First derivative

The first derivative is simple. Applying Leibnitz once leads to

$$\partial_t\langle(x_t - x_0)^2\rangle = f(t,t) + \int_0^t \partial_t f(r,t)\,dr \tag{6.22}$$

$$= 2\int_0^t C_v(r,t)\,dr. \tag{6.23}$$

For **stationary** VACFs, this can be further reduced to

$$\partial_t\langle(x_t - x_0)^2\rangle = 2\int_0^t C(\tau)\,d\tau. \tag{6.24}$$

Second derivative

A second application of the Leibnitz formula yields the second derivative:

$$\partial_t^2\langle(x_t - x_0)^2\rangle = 2\langle v_t^2\rangle + 2\int_0^t \frac{\partial C_v(r,t)}{\partial t}\,dr \tag{6.25}$$

The integral can be simplified for stationary VACFs due to $C_v(t,t) = C(0)$ and $\frac{\partial C(r,t)}{\partial t} = C'(t-r)$:

$$\partial_t^2\langle(x_t - x_0)^2\rangle = 2C_v(t,t) + 2\int_0^t \frac{\partial C(r,t)}{\partial t}\,dr$$

$$= 2C(0) + 2\int_0^t C'(t-r)\,dr = 2C(t)$$

The second derivative of the EAMSD **equals** the VACF in the stationary case, which simplifies many scaling arguments.

6.3.2 EA-MSD scaling

Scaling for stationary VACF

The second derivative of the (EA)MSD for stationary VACF depends on $C(t)$ itself due to

$$\partial_t^2\langle(x_t - x_0)^2\rangle = 2C(t).$$

If the VACF scales as $C(t) \sim t^\alpha$, the scaling exponent during a stable regime should satisfy

$$\alpha_{EA}(t) \sim \alpha + 2.$$

Stationary VACF will be studied more in detail in Sec. 6.5.

Scaling for general VACF

The formula

$$\partial_t^2 \langle (x_t - x_0)^2 \rangle = 2C(t),$$

relates the VACF directly to the EAMSD scaling for **stationary** VACF. Since $C(t) = \langle v_t v_0 \rangle$ and $\langle v_t^2 \rangle = \langle v_0^2 \rangle$, the VACF measures the change in correlation (e.g. decorrelation) directly. Any power - law scaling $C(t) \sim t^\alpha$ thus represents directly how correlation changes over time. For non - stationary VACF $C_v(r,s) = \langle v_r v_s \rangle$ things become more complicated:

It is true that the behavior of $C_v(r, r+t) = \langle v_r v_{r+t} \rangle$ w.r.t. t measures the change of correlation from time r, similar to $C(t)$ in the stationary case, but what **correlation** means is more subtle.

The velocity SMF of stationary VACF $\langle v_t^2 \rangle = \langle v_0^2 \rangle = C(0)$ is temporally constant, whereas for non - stationary VACFs it can change over time. The issue of a non-trivial velocity SMF scaling can be best explained with an example:

Let

$$C_v(r,s) = r^n s^n e^{-|r-s|} \tag{6.26}$$

for $n \geq 1$.[3] The derivative of the VACF,

$$\partial_t C_v(r,t) = (nt^{n-1} - t^n) r^n e^{-|t-r|},$$

is positive for $t < n$, fixed r and $r < t$, but becomes negative for $t > n$. The VACF thus **increases** until $t = n$ and then **decreases** afterwards. It should be expected that $C_v \sim 0$ for large enough separations $|r-s| \gg 1$, but does the increase of C_v at $t < n$ imply an **increasing** correlation?

Comparing this to the **normalized VACF**

$$\rho(r,s) = \frac{C_v(r,s)}{\sqrt{C_v(r,r) C_v(s,s)}} = e^{-|r-s|}, \tag{6.27}$$

it becomes clear that $\rho(r,s)$ is always decreasing for fixed r and growing s. The increase of C_v for $t < n$ is therefore a consequence of the growing SMF $\langle v_t^2 \rangle = t^{2n}$.

Whether or not the behavior ρ or C_v represents the change of correlation better is a matter of hot debate. The important takeaway is that you cannot simply translate the interpretation of correlation / decorrelation for the stationary case to the non stationary case. The scaling behavior of the velocity SMF can heavily skew the behavior, depending on whether ρ or C_v is used as a measure of correlation.

This Janus - headed nature of C_v is present in the formula for the second EAMSD derivative:

$$\partial_t^2 \langle (x_t - x_0)^2 \rangle = 2\langle v_t^2 \rangle + 2 \int_0^t \frac{\partial C_v(r,t)}{\partial t} \, dr$$

The first term is the velocity SMF contribution, whereas the averaged integral is a mixed contribution. In case of the example $C_v(r,s) = r^n s^n e^{-|r-s|}$, the factorizable $r^n s^n$

---

[3] Such VACFs will be studied in Sec. 8.2

represents the SMF contribution, whereas $e^{-|r-s|}$ represents the underlying **decorrelation mechanism**. The SMF scaling can overshadow this decorrelation for early times. Indeed, the formula of the second derivative

$$\partial_t^2\langle(x_t - x_0)^2\rangle = 2C_v(t,t) + 2\int_0^t \partial_t C(r,t)\,dr$$

can be approximated for sufficiently small $t \ll 1$:

Since $C_v(r,t) \sim C_v(t,t) = \langle v_t^2\rangle$ for $|r-t| \ll 1$, the VACF derivative as can be approximated as

$$\partial_t C_v(r,t) \sim \partial_t\langle v_t^2\rangle.$$

The second derivative of the EAMSD simplifies to

$$\partial_t^2\langle(x_t - x_0)^2\rangle \sim 4\langle v_t^2\rangle$$

and the scaling of EAMSD is related to the velocity SMF scaling. This explains intuitively why the scaling relation

$$\alpha_{EA}^x(0) \sim \alpha_{SMF}^v(0) + 2$$

should hold. If t becomes larger, then an underlying decorrelation mechanism can destroy this approximation and the scaling relation is violated.

Understanding the EAMSD scaling thus requires a sufficient understanding of the velocity SMF scaling, especially in the initial scaling regime. But for larger times the correlation term can become more dominant. It is therefore detrimental to also understand the (de)correlation mechanism of the VACF itself. This part will play a major role in the behavior of the ATAMSD in Ch. 8.

Superballistic EAMSD

Given the scaling arguments made about the EASMD, there are two possible way to produce a superballistic scaling:

For a non-stationary VACF with velocity SMF $\langle v_t^2\rangle \sim t^\alpha$, an at least subdiffusive SMF scaling leads to a superballistic EAMSD scaling. This is guaranteed for the initial scaling regime due to

$$\alpha_{EA}^x(0) \geq \alpha_{SMF}^v(0) + 2.$$

I have shown examples for this in Sec. 4.3 about integrated processes.

It it also possible to have a superballistic scaling in the intermediary and asymptotic regime for such a velocity SMF, but this persists only under a moderate decorrelation mechanism. If cancellation effects are introduced, the scaling can become subballistic for later stable regimes. Such an atleast subdiffusive velocity SMF implies an underlying acceleration mechanism. But most of the examples I presented with such a velocity SMF are *artificial* in nature. Aside of the integration mechanism, all of the previous examples in Ch. 4 were artificial by design.

If the VACF is stationary, the derivative formula

$$\partial_t^2\langle(x_t - x_0)^2\rangle = 2C(t)$$

implies that EAMSD and VACF scaling are directly related. If the VACF scales stable as a power - law $C(\tau) \sim \tau^{\beta}$ for $\beta > 0$, the EAMSD will scale superballistically $\langle (x_t - x_0)^2 \rangle \sim t^{\beta+2}$. There are some theoretical objections, however, as the PSD condition of $C(\tau)$ does severely restrict the possible scaling of the VACF. I will show that the PSD condition rules out **asymptotic superballisticity**. Since the EAMSD scales ballistically initially and at most ballistically asymptotically, superballistically can only be achieved at intermediary scaling regimes. Stationary VACFs will be studied more in detail in Sec. 6.5 and Ch. 7.

6.4 ATAMSD via VACF

The ATAMSD in terms of the ALI is given by

$$\langle \delta^2(\Delta, T) \rangle = \int_0^{T-\Delta} \int_0^{T-\Delta} \langle (x_{t+\Delta} - x_t)^2 \rangle \, dt. \tag{6.28}$$

and the VACF-resolved formula is

$$\langle \delta^2(\Delta, T) \rangle = \frac{1}{T-\Delta} \int_0^{T-\Delta} \int_t^{t+\Delta} \int_t^{t+\Delta} C_v(r,s) \, dr \, ds \, dt. \tag{6.29}$$

The ATAMSD

$$\langle \delta^2(\Delta, T) \rangle = \frac{1}{T-\Delta} \int_0^{T-\Delta} A_{t,\Delta} \, dt,$$

involves the ALI with lagtime Δ, but is averaged over all increment starting points. This stands in contrast to the EAMSD, which is given by the initial ALI:

$$\langle (x_t - x_0)^2 \rangle = A_{0,t}$$

From this structural difference alone it should be clear that the ATAMSD is substantially more complicated than the EAMSD. Using the ATAMSD derivatives for scaling arguments and discretizing the ATAMSD is thus a more subtle issue.

Before I go into the matter of ATAMSD derivatives and scaling arguments, I want to shed some light on the discretization of the ATAMSD.

6.4.1 Discretization

The naive discretization of the ATAMSD is

$$\delta_{n,m} = \delta(t_n, T) = \frac{(\Delta t)^2}{m-n} \sum_{k=0}^{m-n-1} \sum_{i,j=k}^{k+n-1} C_{ij}, \tag{6.30}$$

with $\Omega = \{ t_n = n \cdot dt \; : \; n = 0, ..., m \}$ being an equidistant mesh for $T > 0$, mesh size $dt = T/m$ and matrix coefficients $C_{ij} := C_V(t_i, t_j)$. This formula is horrible in more than one way:

First, there are 3 sums in $\delta_{n,m}$, leading to $n^2(m-n)$ terms per step. If one computes $\delta_{n,m}$ for all $n = 0, ..., m-1$, the complexity becomes $\mathcal{O}(m^4)$.[4] Additionally, most of

---

[4] Even the symmetry of C_{ij} does not bring much relieve to the complexity.

the matrix coefficients are computed multiple times. This has been a problem for the EAMSD discretization as well.

I could not find any paper that investigated the efficient discretization of the ATAMSD in detail. In [12], the discrete ATAMSD is presented as a quadratic form of the trajectory vector $x \in \mathbb{R}^m$ with a triangular matrix $M \in \mathbb{R}^{m\times m}$, but this hasn't been utilized for numerical purposes. Therefore, I propose a new discretization algorithm.

The algorithm will not compute the discretized ATAMSD $\delta_{n,m}$ directly, but the unnormalized ATAMSD

$$\zeta_{n,m} = \sum_{k=0}^{m-n-1} \sum_{i,j=k}^{k+n-1} C_{ij} \tag{6.31}$$

instead.[5]

Theorem 12. *The unnormalized ATAMSD $\zeta_{n,m}$ obeys the recursion relation*

$$\zeta_{n,m} = \zeta_{n-1,m} + \mathcal{I}_n \tag{6.32}$$

with ***(first) increment***

$$\mathcal{I}_n = D_{n,m} + \sum_{a=1}^{n-1} I_{a,n,m}. \tag{6.33}$$

Here

$$I_{a,n,m} = 2 \sum_{i=n-1-a}^{m-n-1} C_{i,i+a} \tag{6.34}$$

is the a-th off diagonal and

$$D_{n,m} = \sum_{i=n-1}^{m-n-1} C_{i,i} \tag{6.35}$$

the main diagonal. The first increment $\mathcal{I}_n$ obeys the recursion relation

$$\mathcal{I}_n = \mathcal{I}_{n-1} + \mathcal{T}_n \tag{6.36}$$

with ***second increment***

$$\mathcal{T}_n = I_{n-1,n,m} - \left(C_{n-2,n-2} + C_{m-n,m-n} + 2\sum_{a=1}^{n-2} C_{n-2-a,n-2} + C_{m-n,m-n+a} \right). \tag{6.37}$$

Proof. The recursion relation can be established by extracting the terms for $\zeta_{n-1,m}$ out of the triple sum in $\zeta_{n,m}$. The element $\zeta_{n-1,m}$ is given by

$$\zeta_{n-1,m} = \sum_{k=0}^{m-n} \sum_{i,j=k}^{k+n-2} C_{ij}$$

---

[5] The first terms are given by $\zeta_{0,m} = 0$ and $\zeta_{1,m} = \sum_{k=0}^{m-2} C_{kk}$ for convention.

and we start by rearranging $\zeta_{n,m}$ in a way that $\zeta_{n-1,m}$ will appear:

$$\begin{aligned}
\zeta_{n,m} &= \sum_{k=0}^{m-n-1} \sum_{i,j=k}^{k+n-1} C_{ij} = -\sum_{i,j=m-n}^{m-1} C_{ij} + \sum_{k=0}^{m-n} \sum_{i,j=k}^{k+n-1} C_{ij} \\
&= -\sum_{i,j=m-n}^{m-1} C_{ij} + \sum_{k=0}^{m-n} \sum_{i,j=k}^{k+n-2} C_{ij} + \sum_{k=0}^{m-n} C_{k+n-1,k+n-1} + \sum_{k=0}^{m-n} \sum_{i=k}^{k+n-2} C_{i,k+n-1} + C_{k+n-1,i} \\
&= \zeta_{n-1,m} - \underset{a)}{\sum_{i,j=m-n}^{m-1} C_{ij}} + \underset{b)}{\sum_{k=0}^{m-n} C_{k+n-1,k+n-1}} + \underset{c)}{\sum_{k=0}^{m-n} \sum_{i=k}^{k+n-2} C_{i,k+n-1} + C_{k+n-1,i}}.
\end{aligned}$$

The recursion relation is already there, but the terms $a), b), c)$ need to be simplified into $\mathcal{I}_n$ yet.

Starting with c), I propose the following Ansatz:

$$c) = \sum_{a=0}^{m-1} \sum_{i=n-1-a}^{m-1-a} C_{i,i+a} + C_{i,i+a}$$

This can be proven by index change and the principle of exhaustion, but the written argument is too lengthy. The left figure in Fig. 6.1 shows the elements of C involved in sum $c)$. The matrix elements are naturally ordered in off - diagonals $C_{i,i+a}$, which is expressed in the Ansatz.

The sum $\sum_{i,j=m-n}^{m-1} C_{ij}$ in $a)$ can be split:

$$\sum_{i,j=m-n}^{m-1} C_{ij} = \sum_{k=0}^{m-1} C_{kk} + \sum_{i,j=m-n,j\geq i}^{m-1} C_{ij} + C_{ji}$$

The diagonal part can be combined with the diagonal term $b)$ into $\sum_{k=n-1}^{m-n-1} C_{kk}$, which gives the intermediary

$$\zeta_{n,m} = \zeta_{n-1,m} + \sum_{k=n-1}^{m-n-1} C_{kk} + \sum_{a=1}^{n-1} \sum_{i=n-1-a}^{m-1-a} C_{i,i+a} + C_{i+a,i} - \underbrace{\sum_{i=m-n,j>i}^{m-1} C_{ij} + C_{ji}}_{d)} .$$

The summed indices in the remaining negative term $d)$ are plotted in the middle figure of Fig. 6.1. These indices can again be sorted into off - diagonals and an index exhaustion argument gives

$$\sum_{a=0}^{m-1} \sum_{i=m-n}^{m-1-a} C_{i,i+a} + C_{i,i+a}.$$

The terms c) and d) can be recombined, leading to the sum

$$\sum_{a=0}^{n-1} \sum_{i=n-1-a}^{m-n-1} C_{i,i+a} + C_{i+a,i}.$$

The corresponding index set is shown in the right figure in Fig. 6.1. This proves the recursion relation and the form of the first increment

$$\mathcal{J}_n = D_{n,m} + \sum_{a=1}^{n-1} I_{a,n,m}.$$

Due to the decomposition of $\mathcal{J}_n$ into diagonal $D_{n,m}$ and off-diagonal terms $I_{a,n,m}$, it suffices to show the recursion relation for these terms separately:

The main diagonal satisfies

$$D_{n,m} = D_{n-1,m} - C_{n-2,n-2} - C_{m-n,m-n}$$

and the off-diagonals

$$I_{a,n,m} = I_{a,n-1,m} - 2 \cdot \left(C_{n-2-a,n-2} + C_{m-n,m-n+a}\right),$$

which proves the second increment and the second recursion relation. □

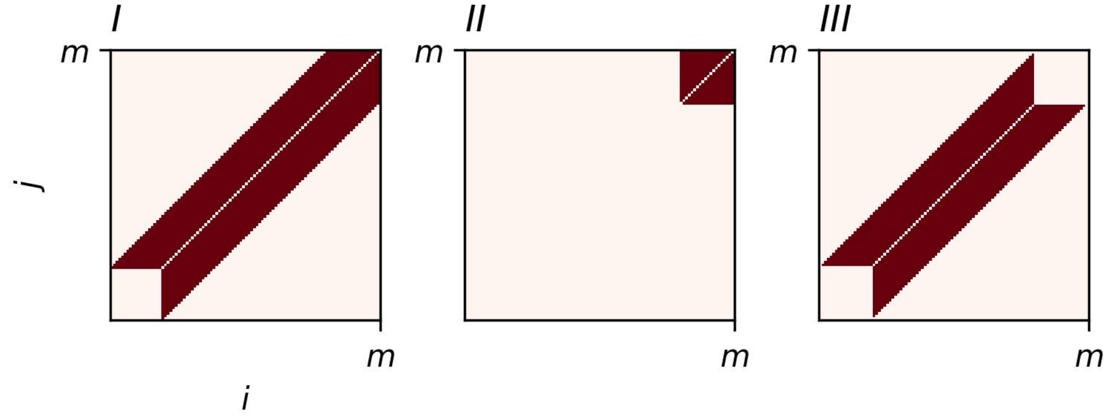

Figure 6.1: Matrix elements of C_{ij} appearing in the proof of Thm. 12.

6.4.2 Derivatives

The unnormalized ATAMSD satisfies the discrete recursion relation

$$\delta_{n,m} = \delta_{n-1,m} + \mathcal{J}_n$$

and the increment the recursion

$$\mathcal{J}_n = \mathcal{J}_{n-1} + \mathcal{T}_n.$$

Since recursion relations are the discrete counterpart to derivatives, the derivative of the ATAMSD will given by a continuous analogon of the first $\mathcal{J}_n$ and second increment $\mathcal{T}_n$.

Algorithm 1 Computation of $\zeta_{n,m}$ for $n = 0, ..., n_{max} - 1$ for $n_{max} < \frac{m}{2}$ (General case)

```
procedure $\zeta(n_{max}, m)$
    Set $D_c \leftarrow \sum_{k=0}^{m-2} C_{kk}$
    Allocate array $I_c$ with $n_{max} - 1$ entries
    Allocate array $\zeta$ with $n_{max}$ entries
    Set $\zeta[0] \leftarrow 0$
    Set $\zeta[0] \leftarrow D_{1,m}$
    Set $D_c \leftarrow 0$ and $I_c \leftarrow 0$
    for $n = 2, ..., n_{max} - 1$ do
        Set $D_c \leftarrow D_c - C_{n-2,n-2} - C_{m-n,m-n}$
        Set $\zeta[n] \leftarrow \zeta[n-1] + D_c$
        for $a = 1, ..., n - 1$ do
            if $a = n - 1$ then
                Set $I_c[a-1] \leftarrow \sum_{i=0}^{m-a-2} C_{i,i+a}$
            else
                Set $I_c[a-1] \leftarrow I_c[a-1] - 2\left(C_{n-2-a,n-2} + C_{m-n,m-n+a}\right)$
            Set $\zeta[n] \leftarrow \zeta[n] + I_c[a-1]$
```

First derivative

The derivative can be split up into 3 terms using Leibnitz's formula:

$$\begin{aligned}\partial_\Delta \langle \delta^2 \rangle &= \partial_\Delta \frac{1}{T-\Delta} \int_0^{T-\Delta} A_{t,\Delta}\, dt \\ &= -\frac{\langle \delta^2 \rangle}{(T-\Delta)^2} + \frac{1}{T-\Delta} \left(\int_0^{T-\Delta} \partial_\Delta A_{t,\Delta}\, dt - A_{T-\Delta,\Delta} \right)\end{aligned}$$

The first term $\frac{\langle \delta^2 \rangle}{(T-\Delta)^2}$ becomes insignificant for the scaling in the large T limit.

The derivative $\partial_\Delta A_{t,\Delta}$ of the ALI follows the same steps as the calculation of $\partial_t \langle (x_t - x_0)^2 \rangle$:

$$\partial_\Delta A_{t,\Delta} = 2 \int_t^{t+\Delta} C_v(r, t+\Delta)\, dr$$

In the stationary case this integral would reduce to

$$\int_0^{T-\Delta} \partial_\Delta A_{t,\Delta}\, dt = (T-\Delta) \int_0^{\Delta} C(r)\, dr, \tag{6.38}$$

but I want to derive a formula for arbitrary VACF, so this is not sufficient. The double integral

$$2 \int_0^{T-\Delta} \int_t^{t+\Delta} C_v(r, t+\Delta)\, dr\, dt$$

is a 2D - integral

$$\int_\Delta^T 2 \int_{t-\Delta}^{t} C_v(r,t)\, dr\, dt = \int_{B_1} C_v(r,t) d\lambda_2(r,t)$$

with integration domain

$$B_1 = \{(r,t) \ : \ \Delta \le t \le T, \, t-\Delta \le r \le t\}.$$

The integration domain can be decomposed into off - diagonals, which has been done for the discrete case. The transformed integral is

$$\int_\Delta^T \int_{t-\Delta}^t C_v(r,t)\, dr\, dt = \int_0^\Delta \int_\Delta^T C_v(t, t-s)\, dt\, ds.$$

Similarly, the remainder $A_{T-\Delta,\Delta}$ can be simplified into off - diagonal integrals:

$$\begin{aligned} A_{T-\Delta,\Delta} &= \int_{T-\Delta}^T \int_{T-\Delta}^T C_v(r,s)\, dr\, ds = 2\int_{T-\Delta}^T \int_r^T C_v(r,s)\, dr\, ds \\ &= 2\int_{B_2} C_v(r,s)\, d\lambda_2(r,s) = 2\int_0^\Delta \int_{T-\Delta+s}^T C_v(t,t-s)\, dt\, ds \end{aligned}$$

Finally, the bracket simplifies to

$$\begin{aligned} \frac{1}{T-\Delta}\left(\int_0^{T-\Delta} \partial_\Delta A_{t,\Delta}\, dt - A_{T-\Delta,\Delta}\right) &= 2\int_0^\Delta \frac{1}{T-\Delta}\int_\Delta^{T-\Delta+s} C_v(t,t-s)\, dt\, ds \\ &= 2\int_0^\Delta I(s,\Delta,T)\, ds. \end{aligned}$$

The term

$$I(s,\Delta,T) = \frac{1}{T-\Delta}\int_\Delta^{T-\Delta+s} C_v(t,t-s)\, dt \tag{6.39}$$

is called **averaged off-diagonal(AOD)** with lag s.[6] The first derivative can be written and approximated as

$$\partial_\Delta\langle\delta^2\rangle = 2\int_0^\Delta I(s,\Delta,T)\, ds - \frac{\langle\delta^2\rangle}{(T-\Delta)^2} \overset{T\gg\Delta}{\sim} 2\int_0^\Delta I(s,\Delta,T)\, ds. \tag{6.40}$$

Second derivative

The formula for the second derivative can be divided into three parts:

$$\partial_\Delta^2\langle\delta^2\rangle = \underset{a)}{2I(\Delta,\Delta,T)} + \underset{b)}{2\int_0^\Delta \partial_\Delta I(s,\Delta,T)\, ds} + \underset{c)}{\partial_\Delta \frac{\langle\delta^2\rangle}{(T-\Delta)^2}} \tag{6.41}$$

Resolving the first integrals gives

$$b) = -\int_0^\Delta C_v(T-\Delta, T-\Delta-s) + C_v(\Delta, \Delta-s)\, ds + \int_0^\Delta \frac{I}{T-\Delta}\, ds.$$

---

[6] Note the discrete analogon $I_{a,n,m} \sim I(s,\Delta,T)$.

The part in c) is straightforward

$$c) = \frac{\partial_\Delta \langle \delta^2 \rangle}{(T-\Delta)^2} + 2\frac{\langle \delta^2 \rangle}{(T-\Delta)^3}$$

and the second derivative becomes:

$$\begin{aligned}\partial^2_\Delta \langle \delta^2 \rangle =& 2\left(I(\Delta,\Delta,T) - \int_0^\Delta C_v(T-\Delta, T-\Delta-s) + C_v(\Delta,\Delta-s)\, ds \right) \\ &+ \frac{2}{T-\Delta}\int_0^\Delta I(s,\Delta,T)\, ds + \frac{2\langle \delta^2 \rangle}{(T-\Delta)^3} + \frac{\partial_\Delta \langle \delta^2 \rangle}{(T-\Delta)^2} \\ &\overset{T \gg \Delta}{\sim} 2 \cdot I(\Delta,\Delta,T)\end{aligned}$$

If no confusion arises, I will call $I(\Delta, T) = I(\Delta,\Delta,T)$ **AOD** as well.

6.4.3 Scaling of ATAMSD

The ATAMSD is phenomenologically more complex than the EAMSD. Not only does it contain more terms,[7] but there is no direct scaling relation to the VACF scaling, as even the second derivative still contains the VACF in integrated form.
Another problematic part is the simultaneous dependence on Δ and T. Since

$$\langle \delta^2(\Delta,\Delta) \rangle = \langle (x_\Delta - x_0)^2 \rangle,$$

it should be expected that in cases that large T convergence fails, the scaling w.r.t. Δ should change for different observational times T. I will therefore first discuss the case of stationary VACF, before I tackle the problem of non - stationary VACF afterwards in Ch. 8.

6.5 Stationary VACFs

The mathematically more consistent definition of **stationary VACFs** involves the concept of **weak stationarity**:

Definition 27. *Let v_t be a process, which is* ***second order***[8]*. Assume with out loss of generality that $\langle v_t \rangle = 0$. Then v_t is called* ***weakly - stationary / moment stationary*** *iff*

$$\langle v_t \rangle = \langle v_0 \rangle, \quad Cov(r,s) = Cov(|r-s|, 0) \tag{6.42}$$

For the purpose of this thesis it suffices that the VACF

$$C_v(r,s) = \langle v_r v_s \rangle = \langle v_{|r-s|} v_0 \rangle = C(r-s) \tag{6.43}$$

depends only on the time difference $r - s$.

Classical references for the topic of stationary VACFs in the wider context of SDEs and stochastic processes are [22], [7]. A more direct link with statistical mechanics and molecular simulations is given in [13] and [11].

---

[7] Most of these are not relevant of the scaling, however, due to their weaker T - scaling.

[8] Second order means that the mean $\langle v_t \rangle$ and SMF $\langle v_t^2 \rangle$ exist.

6.5.1 Ergodicity for stationary VACF

From a theoretical standpoint, stationary VACFs have the great advantage that the integrated process x_t is **large T convergent**:

Theorem 13. *Let x_t be a differentiable process with stationary VACF. Then*

$$\langle \delta^2(\Delta, T) \rangle = \langle (x_\Delta - x_0)^2 \rangle. \tag{6.44}$$

This holds for any observational time $T > \Delta$.

The theorem follows from the simpler proposition that the ALI does not depend on the initial time for stationary VACF:

Proposition 2. *Let x_t be a differentiable process with stationary VACF. Then*

$$A_{t,\Delta} = \int_0^\Delta (\Delta - \tau) C(\tau)\, d\tau. \tag{6.45}$$

Consequently $A_{t,\Delta} = A_{0,\Delta}$.

Proof. The double integral of the ALI can be rearranged as

$$\begin{aligned} A_{t,\Delta} &= \int_t^{t+\Delta} \int_t^{t+\Delta} C_v(r, s)\, ds\, dr = \int_0^\Delta \int_0^\Delta C_v(t + r, t + s)\, ds\, dr \\ &= 2 \int_0^\Delta \int_r^\Delta C_v(t + r, t + s)\, ds\, dr \end{aligned}$$

due to the symmetry of the VACF. Since $C_v(t+r, t+s) = C(s-r)$, a simple substitution leads to

$$\begin{aligned} A_{t,\Delta} &= 2 \int_0^\Delta \int_r^\Delta C_v(t + r, t + s)\, ds\, dr = 2 \int_0^\Delta \int_r^\Delta C(s - r)\, ds\, dr \\ &= 2 \int_0^\Delta \int_0^{\Delta - r} C(\tau)\, d\tau\, dr. \end{aligned}$$

Changing the integration order using Fubini's theorem leads to

$$\begin{aligned} A_{t,\Delta} &= 2 \int_0^\Delta \int_0^{\Delta - r} C(\tau)\, d\tau\, dr = 2 \int_0^\Delta \int_0^{\Delta - \tau} C(\tau)\, dr\, d\tau \\ &= 2 \int_0^\Delta C(\tau) \int_0^{\Delta - \tau} dr\, d\tau = 2 \int_0^\Delta (\Delta - \tau) C(\tau)\, d\tau. \end{aligned}$$

□

6.5.2 MSD scaling

The EAMSD and ATAMSD agree for processes with stationary VACFs. The combined growth function is called **MSD** MSD_t.

The MSD for stationary VACFs is

$$MSD_t = 2\int_0^t (t-\tau)C(\tau)\,d\tau. \tag{6.46}$$

The first derivative is given by

$$\partial_t MSD_t = 2\int_0^t C(\tau)\,d\tau. \tag{6.47}$$

and second derivative directly by the VACF

$$\partial_t^2 MSD_t = 2C(t). \tag{6.48}$$

Thm. 5 implied that the (A)TAMSD of any differentiable processes scales initially ballistically, so the MSD of processes with stationary VACF scales initially ballistic:

$$\alpha_{MSD}(0) = 2$$

This was a consequence of the pathwise regularity. The scaling of the intermediary and asymptotic regimes are not restricted by this, so the scaling exponent is not necessarily constrained to be ballistic.

The second derivative of the MSD is directly proportional to the VACF **itself**. This implies that any stable scaling of the VACF akin to $C(t) \sim t^\beta$ implies a stable scaling of the MSD as $MSD_t \sim t^{2+\beta}$. If it is possible to construct a VACF with a stable scaling for $\beta > 0$, then this direct relation implies a **stable superballistic scaling** of the MSD.

Before I address the possibility of such a scaling, I want to explore this relation in case that $\beta \leq 0$.

Constant walker and constant VACF

The **constant walker model**[9] is a variation of the deterministic walker form Sec. 4.1.1,

$$x_t = f(t)e_\theta, \tag{6.49}$$

where $f(t)$ is a continuously differentiable function and e_θ a random unit vector.

The constant walker is given by the linear position function $f(t) = v_0 \cdot t$. The respective velocity process $v_t = v_0 e_\theta$ is pathwise constant and the VACF is given by

$$C(\tau) = \langle v_\tau v_0\rangle = v_0^2\langle e_\theta^2\rangle = v_0^2. \tag{6.50}$$

This is a very specific setting, since the VACF remains **constant** over time. In fact, every such constant VACF implies a pathwise constant velocity:

---

[9] Constant here refers to the constant velocity.

Theorem 14. *Let v_t be a stochastic process with stationary ACF $C(t)$, i.e. $C(t) = C(0)$ for all t. Then*

$$v_t = v_0 \tag{6.51}$$

almost surely.

Proof. Due to the constant VACF, the EAMSD of v_t has to vanish:

$$\begin{aligned}\langle (v_t - v_0)^2 \rangle &= \langle v_t^2 \rangle + \langle v_0^2 \rangle - 2\langle v_t v_0 \rangle \\ &= 2\left(C(0) - C(t)\right) \overset{!}{=} 0\end{aligned}$$

This implies $v_t \equiv v_0$ almost surely. □

The stationary VACF $C(t)$ has to be globally bounded $|C(t)| \leq C(0)$ by the initial correlation. If the stationary velocity is **not** pathwise constant, equality cannot hold and the velocity has to **decorrelate** eventually at some point.

The persistent ballisticity of the constant walker can be derived from the second derivative

$$\partial_t^2 MSD_t = 2C(t) = 2v_0^2 = 2v_0^2 t^0,$$

which gives a ballistic $MSD_t = 2v_0^2 t^2$.

But this logic does also work if the VACF $C(t)$ becomes asymptotically constant, i.e. if the limit $\lim\limits_{t\to\infty} C(t) = C_\infty \neq 0$ if finite and non zero. The scaling estimation via the second derivative would imply that the MSD has to scale ballistically as

$$MSD_t \overset{t\gg 1}{\sim} C_\infty t^2. \tag{6.52}$$

This can be proven analytically:

Theorem 15. *Let x_t be a differentiable stochastic process with continuous, stationary velocity v_t. If the VACF $C(t)$ has a non - zero asymptotic limit, i.e.*

$$\lim_{t\to\infty} C(t) = C_\infty \neq 0, \tag{6.53}$$

*then the MSD scales asymptotically ballistically with **effective squared velocity***

$$v_{eff}^2 := \lim_{t\to\infty} \frac{MSD_t}{t^2} = C_\infty. \tag{6.54}$$

Proof. Since $\lim\limits_{t\to\infty} C(t) = C_\infty \neq 0$ and v_t is continuous, there exists a continuous function h with $\lim\limits_{t\to\infty} h(t) = 0$, such that

$$C(t) = C_\infty + h(t).$$

Proving asymptotic ballisticity follows from showing that the limit in v_{eff}^2 is finite:

The quotient of the MSD can be divided into

$$\frac{MSD_t}{t^2} = C_\infty + 2\frac{\int_0^t (t-\tau)h(\tau)d\tau}{t^2}.$$

The second term can be bounded by

$$\left|\frac{\int_0^t (t-\tau)h(\tau)d\tau}{t^2}\right| \le \left|\frac{\int_0^t h(\tau)d\tau}{t}\right| \to 0$$

due to Thm. 25. This gives the upper bound

$$\lim_{t\to\infty}\left|\frac{MSD_t}{t^2}\right| \le C_\infty.$$

The lower bound follows from eq. 5.23

$$C_\infty + \min_{r\in[0,t]} h(r) \le \frac{MSD_t}{t^2}$$

and

$$\lim_{t\to\infty}\min_{r\in[0,t]} h(r) = 0.$$

□

Constant VACFs or those with a non-zero asymptotic limit have to scale ballistically in the asymptotic regime. A heuristical derivation via derivatives is also possible:

$$\partial_t^2 MSD_t = 2C(t) \overset{t\gg 1}{\sim} 2C_\infty = 2C_\infty t^0.$$

This argument also works for intermediary regimes, so locally constant VACF $C(t) \sim$ const is associated with a ballistic MSD scaling.

Scaling exponent for LRD processes

The asymptotically constant VACFs $C(t) \sim C_\infty \cdot t^0$ are special in the sense that they do not asymptotically decorrelate, i.e.

$$\lim_{t\to\infty} C(t) \neq 0.$$

Most stochastic models have an underlying decorrelation mechanism and do asymptotically decorrelate. There are stark differences, however, in the *way* they decorrelate. The **LRD processes** forms a special subclass of those.

Definition 28. *The velocity v_t is said to have* ***long-range dependence (LRD)*** *if the VACF decays asymptotically algebraically*

$$C(t) \overset{t\to\infty}{\sim} t^{-\alpha} \tag{6.55}$$

for $0 < \alpha < 1$. v_t is then called a ***LRD process*** *or said to be* ***heavy tailed****.*

The VACFs of LRD velocities vanish asymptotically, but the are **not integrable**:

$$\lim_{t\to\infty}\int_0^t C(\tau)\,d\tau = \infty$$

It is the non - intractability that makes LRD processes more exotic. Notable examples include Levy walks (see [28]), fractional Brownian motion with Hurst index $H > \frac{1}{2}$ (see [6], Sec. 3.2) and Dyson Brownian motion in random matrix theory (see [10], Ch. 12).

Although examples of LRD processes became more prominent in the last decades, especially due to a rise in the study of *heavy tailed phenomena* (see [19] for a self-contained exposition), the examples above are strictly non-stationary. I will therefore use the following conceptual VACF:[10]

$$C(t) = (1+t)^{-\alpha} \tag{6.56}$$

The MSD of a process x_t with such a VACF can be analytically calculated:

$$MSD_t = \frac{2}{1-\alpha}\left(\frac{(1+t)^{2-\alpha}}{2-\alpha} - t - \frac{1}{2-\alpha}\right) \tag{6.57}$$

There is no quadratic term in the MSD that suggests an initial ballistic scaling, but looking at the individual terms alone is misleading. Initial ballisticity arises due to initial cancellation effects:

The first term has a Taylor series

$$\frac{(1+t)^{2-\alpha}}{2-\alpha} = \frac{1}{2-\alpha} + t + \frac{(1-\alpha)}{2}t^2 + \sum_{n=3}^{\infty} c_n t^n$$

around $t = 0$. After canceling with the remaining terms, The MSD simplifies to

$$MSD_t = t^2 + \sum_{n=3}^{\infty} c_n t^n$$

and t^2 becomes the lowest effective power of this power series. The ballistic scaling follows from the finite limit

$$\lim_{t\to 0} \frac{MSD_t}{t^2} = 1.$$

Although there is a linear term in the MSD, the term $(1+t)^{2-\alpha}$ has as stronger scaling due to $2-\alpha > 1$. The asymptotic superdiffusivity can be established by proving the convergence of

$$\frac{MSD_t}{t^{2-\alpha}} = \frac{2}{1-\alpha}\left(\frac{1}{2-\alpha}\frac{(1+t)^{2-\alpha}}{t^{2-\alpha}} - t^{\alpha-1} - \frac{1}{(2-\alpha)t^{2-\alpha}}\right).$$

Since $\alpha - 1 < 0$, the middle term vanishes and similarly the last term. Only the first term has a finite limit and the whole MSD quotient converges:

$$\lim_{t\to\infty} \frac{MSD_t}{t^{2-\alpha}} = \frac{2}{(1-\alpha)(2-\alpha)}$$

---

[10] Although the VACF is not integrable, PSD can still be proven using a version of Bochner's theorem via spectral measures.

The MSD scales asymptotically superdiffusively

$$MSD_t \overset{t\to\infty}{\sim} \frac{2}{(1-\alpha)(2-\alpha)} t^{2-\alpha}$$

with $\alpha_{MSD}(\infty) = 2 - \alpha$. This scaling can also be shown using the derivative analysis:

The first derivative of the MSD is given by

$$\partial_t MSD_t = \frac{2}{1-\alpha}\left((1+t)^{1-\alpha} - 1\right)$$

and since $1 - \alpha > 0$, the dominant term has power $1 - \alpha$. This suggest an asymptotic scaling exponent of

$$\alpha_{MSD}(\infty) = 1 - \alpha + 1 = 2 - \alpha,$$

which has been verified in the direct MSD computation.

For more general LRD processes, the expression

$$C(t) \overset{t\to\infty}{\sim} t^{-\alpha}$$

is interpreted via asymptotic analysis as the limit

$$\lim_{t\to\infty} \frac{C(t)}{t^{-\alpha}} = C \neq 0.$$

In this case the second derivative really satisfies

$$\lim_{t\to\infty} \frac{\partial_t^2 MSD_t}{t^{-\alpha}} = \lim_{t\to\infty} \frac{2C(t)}{t^{-\alpha}} = 2C \neq 0$$

with a finite limit and

$$\alpha_{\partial_t^2 MSD}(\infty) = -\alpha.$$

Therefore one can assume

$$\alpha_{MSD}(\infty) = \alpha_{\partial_t^2 MSD}(\infty) + 2 = 2 - \alpha$$

and that such heavy tailed processes do in fact scale asymptotically superdiffusively.[11]

The cautious tale of diffusive MSDs

I have warned previously that the scaling exponent estimation using derivatives can lead to erroneous results if not handled with care. Aside of the requirement of stable scaling, it is very important that the scaling exponent is not an **integer**, or at least not diffusive in case of using the second derivative. The problem for diffusive MSDs can best be seen by analyzing the following VACF:

$$C(\tau) = (1+\tau)^{-\alpha} \tag{6.58}$$

---

[11] Superdiffusively, but also strictly subballistically.

The functional form is identical to the LRD example, but now the exponent is $\alpha > 1$.

The second derivative

$$\partial_t^2 MSD = (1+t)^{-\alpha} \sim t^{-\alpha}$$

becomes asymptotically a power - law and one would assume $\alpha_{\partial_t^2 MSD}(\infty) = -\alpha$. Following the previous examples, I expect that the relation

$$\alpha_{MSD}(\infty) = \alpha_{\partial_t^2 MSD}(\infty) + 2 = 2 - \alpha$$

remains true and for $1 < \alpha < 2$ the MSD should scale subdiffusive in the asymptotic regime. But this is **not** the case:

The MSD is structurally the same as for the LRD example:

$$MSD_t = \frac{2}{\alpha - 1}\left(\frac{1}{2-\alpha} + t - \frac{(1+t)^{2-\alpha}}{2-\alpha}\right).$$

But $(1+t)^{2-\alpha}$ grows **weaker** than t due to $2 - \alpha < 1$. In fact

$$\lim_{t\to\infty} \frac{MSD_t}{t} = \frac{2}{\alpha - 1}$$

and the MSD scales **diffusively** in the asymptotic regime. But why does the second derivative *fail us*?

The first derivative is still fine:

$$\partial_t MSD_t = \frac{2}{\alpha - 1}\left(1 - (1+t)^{1-\alpha}\right)$$

Since $1 - \alpha < 0$, the constant term dominates as t^0 and the scaling relation

$$\alpha_{MSD}(\infty) = \alpha_{\partial_t MSD}(\infty) + 1 = 1$$

remains true. The first derivative accurately predicts the scaling exponent, but the leading term is **constant**. The linear term $\partial_t^2 t = 0$ **vanishes** in the second derivative, so it cannot influence the scaling in $\partial_t^2 MSD_t$.

I mentioned in Sec. 6.2 that computing the scaling exponent during stable regimes via higher derivatives does not necessarily work for integer exponents. This is due to the fact that derivatives of simple monomials t^n vanish after the n-th derivative. The third derivative can become an unreliable estimator for ballistic MSDs, and the second derivative unreliable for diffusive MSDs.

This is especially true for **integrable** VACFs:

6.5.3 Integrable VACF

The MSD formula

$$MSD_t = 2\int_0^t (t-\tau)C(\tau)\,d\tau$$

can be divided into two parts:

$$MSD_t = 2\int_0^t (t-\tau)C(\tau)\,d\tau = 2t\int_0^t C(\tau)\,d\tau - 2\int_0^t \tau C(\tau)\,d\tau$$

Here

$$D(t) := \int_0^t C(\tau)\,d\tau \tag{6.59}$$

is the **(instantaneous) diffusivity constant** and $R(t)$ the **correction term**.

Under normal circumstances this distinction of terms would not matter, because $D(t)$ and $R(t)$ do not have to converge and can scale arbitrarily. But if the VACF is **integrable** in the sense that the **effective diffusivity constant**

$$D := \lim_{t\to\infty} D(t) = \lim_{t\to\infty}\int_0^t C(\tau)\,d\tau \neq 0 \tag{6.60}$$

exists, the distinctions matters. The MSD will scale asymptotically diffusive:

Theorem 16. *Let x_t be a differentiable process with stationary VACF $C(\tau)$. Assume that $C(\tau)$ is integrable, i.e.*

$$\int_0^\infty C(\tau)\,d\tau < \infty,$$

then

$$\lim_{t\to\infty}\frac{MSD_t}{t} = 2D \tag{6.61}$$

with effective diffusivity constant $\int_0^\infty C(\tau)\,d\tau$. The process x_t has to scale asymptotically diffusive.

Proof. The starting point is

$$\frac{MSD_t}{t} = 2\int_0^t C(\tau)\,d\tau - \frac{2}{t}\int_0^t \tau C(\tau)\,dt,$$

which has the limit

$$\lim_{t\to\infty}\frac{MSD_t}{t} = 2D - 2\lim_{t\to\infty}\frac{\int_0^t \tau C(\tau)\,dt}{t}.$$

To prove the theorem it suffices to show that the second limit vanishes:

First of, if the integral $\int_0^\infty \tau C(\tau)\,dt < \infty$ converges, the theorem is already proven. Now assume that $\int_0^\infty \tau C(\tau)\,dt = \infty$ strictly diverges. In this case, using l'Hospital transforms the limit into

$$\lim_{t\to\infty}\frac{\int_0^t \tau C(\tau)\,dt}{t} = \lim_{t\to\infty} tC(t).$$

The theorem is proven if the limit $\lim\limits_{t\to\infty} tC(t) < \infty$ holds:

Assume that $\lim_{t\to\infty} tC(t) = \infty$. In this case there exists a sequence t_n with $t_{n+1} \geq 2t_n$ such that

$$|C(t_n)| > \frac{\epsilon}{t_n}.$$

The integral

$$\int_{t_n}^{2t_n} |C(\tau)|\, d\tau > \frac{\epsilon}{t_n}(2t_n - t_n) = 2\epsilon$$

can be bounded from below by 2ϵ. Since the intervals $[t_n, 2t_n]$ are disjoint by construction, the integral $\int_0^\infty |C(\tau)|\, d\tau)$ can be bounded from below by

$$\int_0^\infty |C(\tau)|\, d\tau \geq \sum_n \int_{t_n}^{2t_n} |C(\tau)|\, d\tau \geq 2\sum_n \epsilon = \infty.$$

This contradicts the integrability of $C(\tau)$, therefore $\lim_{t\to\infty} tC(t) < \infty$. □

In contrast to general asymptotically diffusive processes, those with integrable VACF have a finite limit

$$\lim_{t\to\infty} \frac{MSD_t}{t} = 2D.$$

This implies the existence of some continuous h such that $MSD_t = (2D + h)\,t$ and $\lim_{t\to\infty} t \cdot h(t) = 0$. The scaling exponent of the remainder h has to be $\alpha_h(\infty) = \beta \leq 0$. It can happen that $\alpha_h(\infty) < 0$, and exactly this situation can spoil the derivative argument:

The first derivative

$$\partial_t MSD_t = 2D + \partial_t h + t\partial_t h$$

retains the dominant constant term $2D$, so $\alpha_{\partial_t MSD}(\infty) = 0$. In this case the scaling relation still holds. But in the second derivative

$$\partial_t^2 MSD_t = 2\partial_t h + t\partial_t^2 h,$$

the linear term of the MSD vanishes entirely, giving h the prime spot in prescribing the scaling exponent of $\partial_t^2 MSD_t$. Depending on the remainder scaling exponent $\alpha_h(\infty) := \beta$, the second derivative has an asymptotic scaling exponent

$$\alpha_{\partial_t^2 MSD}(\infty) = \beta - 1.$$

Relating the scaling exponent of $\partial_t^2 MSD_t$ to MSD_t can lead to an erroneous

$$\alpha_{MSD}(\infty) \overset{?}{\sim} 1 + \beta \leq 1$$

subdiffusive scaling exponent if $\beta \neq 0$. This does not necessarily happen only for integrable stationary VACF, but most of the examples are integrable.

6.5.4 Asymptotic scaling for stationary VACF

The major takeaway from Sec. 6.5.2 is the validity of the scaling argument

$$C(t) \overset{t\to\infty}{\sim} t^{\alpha} \Rightarrow MSD_t \overset{t\to\infty}{\sim} t^{2+\alpha}$$

for non - diffusive MSDs in the asymptotic regime. This should likewise hold for intermediary stable regimes.

If a stationary VACF is found, that **increases** asymptotically as $C(t) \sim t^{\beta}$ for some $\beta > 0$, then the corresponding trajectory process x_t scales asymptotically superballistically with $\alpha_{MSD}(\infty) = 2 + \beta$. But sadly this is not possible:

I have shown in Thm. 5 that any differentiable process scales initially ballistic in the (TA)MSD due to the pathwise differentiability, i.e.

$$\alpha_{MSD}(0) = 2.$$

For sufficiently small times t, such that C is non-negative, the VACF has been shown to satisfy the upper and lower bound

$$\min_{r\in[0,t]} C(r) \leq \frac{MSD_t}{t^2} \leq C(0).$$

These bounds can be used to rule out *asymptotic superballisticity*:

Theorem 17. *Let x_t be a stochastic process with stationary VACF. Then the asymptotic scaling exponent is bounded from above:*

$$\alpha_{MSD}(\infty) \leq 2 \tag{6.62}$$

*x_t **cannot** scale asymptotically superballistically.*

Proof. With Thm. 3 it suffices to show the bound $\overline{\alpha}_{MSD} \leq 2$ for the upper index. Since

$$\lim_{t\to\infty} \frac{MSD_t}{t^2} \leq \lim_{t\to\infty} \frac{C(0)t^2}{t^2} = C(0) < \infty,$$

the bound $\overline{\alpha}_{MSD} \leq 2$ holds true. □

The core argument in this proof is the global bound

$$C(t) \leq |C(t)| \leq C(0) < \infty.$$

Asymptotic superballisticity requires an asymptotically **increasing** VACF of the type $C(t) \sim t^{\beta}$ with $\beta > 0$ for t large enough. But this contradicts the global boundedness of C. The VACF can at most saturate at some stable value $\lim_{t\to\infty} C(t) = C_\infty \neq 0$, for which only *asymptotic ballisticity* is achieved. Asymptotic superballisticity is not possible for stationary VACF.

6.5.5 Rigid scaling for stationary VACF

At this point the properties of superballistic MSDs and stationary VACFs do not seem to get along. The increasing nature of VACFs of superballistic MSDs is in kahoots with the global boundedness and PSD of the VACF. There are three important takeaways:

1. The MSD scales initially ballistically: $\alpha_{MSD}(0) = 2$
2. The MSD scales asymptotically at most ballistically: $\alpha_{MSD}(\infty) \leq 2$
3. The value of α_{MSD} in any stable regime depends on the stable scaling behavior $C(t) \sim t^{\beta}$

Things may look bleak for superballisticity at this point. But hope is not lost yet!

These objections restricts $\alpha_{MSD}(0)$ and $\alpha_{MSD}(\infty)$, so the initial and asymptotic regime can sustain at most a ballistic scaling. But nothing constrains **intermediary** scaling regimes. It is of course possible to have an increasing VACF akin to $C(t) \sim t^{\beta}$ for $\beta > 0$ on some intermediary time scale!

Intermediary scaling regimes are the only option in case of stationary VACFs. And this is a problem in its own right:

None of the previously presented VACFs had **any** stable intermediary scaling regimes. The problem is that these VACFs have been **monotonically decaying** VACFs, i.e.

$$C(t') \leq C(t)$$

for any $t \leq t'$. Monotonically decaying VACFs do not increase at all and the respective MSD scaling cannot exceed ballisticity. Moreover, no stable intermediary scaling has been observed. Examples like

$$C(t) = (1 + t)^{-\alpha}$$

start ballistically, but after a small time the VACF starts to show the asymptotic tail behavior $C(t) \sim t^{-\alpha}$. This passage from the initial to the asymptotic regimes is only accompanied by an intermediary transient regime. No stable intermediary scaling is possible. These VACF are too **rigid**.

A superballistic stable scaling requires $C(t) \sim t^{\beta}$ for $\beta > 0$ during some intermediary period. For this to happen, C has to decay for a while away from $t = 0$,[12] increase as a power t^{β} during intermediary times and either asymptotically decay at $t \to \infty$ or saturate.

The previous VACFs could not do this, since they were *mono - modal*: They had just **one** maximum, the initial $C(0)$. From the description it should be clear that C needs to have a local minimum before the onset of the superballistic intermediary regime, and a local maximum during or slightly after. The VACF has to be **multi - modal**.

The next chapter explores the possibility of multi-modal VACFs. There will be problems on the way, as usual.

---

[12] This is due to the global bound $|C(t)| \leq C(0)$.

7 Taking balance: Stationary VACF and multi-modality

The previous Chapter 6 established that a process x_t with stationary velocity can only have a stable superballistic MSD if the VACF grows as $C(t) \sim t^{\beta}$ for $\beta > 0$. Such a behavior is only possible if the VACF $C(t)$ is **multi - modal**. It is not difficult to find or construct multi-modal functions with such an increasing power - law scaling. The problem for VACFs, however, lies in the PSD condition. In fact, the PSD condition is the only obstruction in constructing valid multi - modal VACFs.

The later part of this chapter is devoted to studying concrete examples of multi - modal VACFs. It is detrimental to verify the PSD condition in these cases. Therefore, the first part is spent to introduce the tools necessary to verify the PSD condition in concrete examples.

7.1 PSD for stationary VACF

The concept of **positive-semidefiniteness** for a bivariate kernel $K(r, s)$ is explained in Sec. 6.1. In general, there is no direct way to verify PSD aside of numerical estimations or direct proofs. In case of stationary kernels, a different route is possible via **Bochner's theorem**:

7.1.1 Bochner's theorem

Theorem 18 (Bochner's theorem - *see [22], Sec. 1.2*). *Let $C(\tau)$ be a continuous positive-semidefinite function. Then there exists a unique nonnegative measure ρ such that $\rho(\mathbb{R}) = C(0)$ and*

$$C(\tau) = \int_{\mathbb{R}} e^{i\omega\tau} \rho(d\omega) \;\; \forall t \in \mathbb{R}. \tag{7.1}$$

Conversely, every such C is positive - semidefinite.

If the measure ρ has a density $S(\omega)$ w.r.t. the Lebesgue measure λ_1 in the sense of

$$\int_A \rho(d\omega) = \int_A S(\omega) d\lambda_1(\omega) = \int_A S(\omega) d\omega,$$

eq. 7.1 can be rewritten as

$$C(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\,\omega. \tag{7.2}$$

The density $S(\omega)$ is the *Fourier transform* of the VACF

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} C(\tau)\, d\tau \tag{7.3}$$

and is called the **spectral density** of C.

Since the spectral measure ρ has to be non-negative, the spectral density S has to be non-negative for real frequencies $\omega \in \mathbb{R}$ as well. This connection makes it possible to verify the PSD condition via the Fourier transform of C:

Theorem 19. *Let $C(\tau)$ be a continuous function. Then the $C(\tau)$ is PSD if the spectral density is non-negative for real frequencies:*

$$S(\omega) \geq 0 \;\; \forall \omega \in \mathbb{R} \tag{7.4}$$

7.1.2 PSD and characteristic functions

If a stationary VACF is normed, $C(0) = 1$, there is a direct connection to random variables via the characteristic function

$$\varphi_X(t) = \langle e^{itX} \rangle. \tag{7.5}$$

Using the probability measure ρ of X, the characteristic function can be written as

$$\varphi_X(t) = \int_{\mathbb{R}} e^{i\omega t} \rho(d\omega). \tag{7.6}$$

Any stationary VACF with $C(0) = 1$ therefore corresponds to a unique probability distribution.

7.2 Exponentially weighted polynomials

The conceptually easiest class, for which PSD can be analytically verified and multi-modality is achieved, is the group of the **exponentially weighted polynomials:**

$$C(\tau) = p_n(|\tau|) e^{-|\tau|} \tag{7.7}$$

The PSD condition is going to be verified using Bochner's theorem. Since the VACF is the product of a polynomial and the decaying exponential, the Fourier transform of C can be analytically computed for arbitrary polynomials. Verifying the non-negativity of S is combinatorically difficult in general and numerical estimates have to be used.

7.2.1 Positive - semidefiniteness via Bochner's theorem

Due to the linearity of the Fourier transform, it suffices to compute the Fourier transform of the weighted monomials $g_n(x) = |x|^n e^{-|x|}$ individually:

The function g_n is even, so the Fourier transform can be rewritten as

$$\begin{aligned}
S_{g_n}(\omega) &= \int g_n(x)e^{-i\omega x}\,dx = \int_{-\infty}^{0} g_n(x)e^{-i\omega x}\,dx + \int_0^\infty g_n(x)e^{-i\omega x}\,dx \\
&= \int_0^\infty g_n(x)e^{i\omega x}\,dx + \int_0^\infty g_n(x)e^{-i\omega x}\,dx = 2\int_0^\infty g_n(x)cos(\omega x)\,dx \\
&= 2Re\left\{\int_0^\infty g_n(x)e^{i\omega x}\,dx\right\}.
\end{aligned}$$

Due to the Gamma function identity $\int_0^\infty x^n e^{-\lambda x}\,dx = \frac{n!}{\lambda^{n+1}}$ for complex λ, the monomial can be simplified into

$$2Re\left\{\int_0^\infty x^n e^{-x} e^{i\omega x}\,dx\right\} = 2Re\left\{\int_0^\infty x^n e^{-(1-i\omega)x}\,dx\right\} = 2Re\left\{\frac{n!}{(1-i\omega)^{n+1}}\right\}$$

The real part can be resolved using a binomial expansion:

$$S_{g_n}(\omega) = \frac{2}{(1+\omega^2)^{n+1}} n! Re\left\{(1+i\omega)^{n+1}\right\} = \frac{2}{(1+\omega^2)^{n+1}} n! \sum_{j=0}^{n+1} \binom{n+1}{j} \omega^j Re\left\{i^j\right\}$$

Despite its combinatorial nature, the Fourier transform of the weighted monomial g_n can be compactly written as

$$S_{g_n}(\omega) = \frac{2P_n(\omega)}{(1+\omega^2)^{n+1}}.$$

P_n is the **even** polynomial

$$P_n(\omega) = n! \sum_{j=0}^{n+1} \binom{n+1}{j} \omega^j Re\{i^j.\} \tag{7.8}$$

Verifying the non-negativity of S_{g_n} is equivalent to verifying the non-negativity of P_n for real ω. The case of general polynomials is more subtle, as can be seen by the simple example of $p(x) = g_n(x) + g_{n+1}(x)$:

$$\begin{aligned}
S_p(\omega) &= \frac{2P_n(\omega)}{(1+\omega^2)^{n+1}} + \frac{2P_{n+1}(\omega)}{(1+\omega^2)^{n+2}} = \frac{2P_n(\omega)(1+\omega^2)}{(1+\omega^2)^{n+2}} + \frac{2P_{n+1}(\omega)}{(1+\omega^2)^{n+2}} \\
&= \frac{2\left(P_n(\omega)(1+\omega^2) + P_{n+1}(\omega)\right)}{(1+\omega^2)^{n+2}}
\end{aligned}$$

Since the Fourier transforms of g_n and g_{n+1} involve **different** inverse powers, one cannot simply add up the polynomials P_n and verify the non-negativity of this summed polynomial. If the weighted polynomial is given by $C(\tau) = p_n(|\tau|)e^{-|\tau|}$ with

$$p_n(\tau) = \sum_{k=0}^{n} c_k \tau^k,$$

it is still possible to write the Fourier transform of C using a single fraction

$$S_C(\omega) = \frac{2\tilde{P}_n(\omega)}{(1+\omega^2)^{n+1}}, \tag{7.9}$$

albeit with the weighted polynomial

$$\tilde{P}(\omega)_n = \sum_{k=0}^{n} c_k P_k(\omega)(1+\omega^2)^{n-k}. \tag{7.10}$$

The closed form of $\tilde{P}_n$ is also known:

$$\tilde{P}_n(\omega) = \sum_{k=0}^{n} k! c_k \sum_{l=0}^{n-k} \sum_{j=0}^{n+1} \binom{n-k}{l}\binom{k+1}{j} \omega^{2l+j} Re\{i^j\} \tag{7.11}$$

7.2.2 Cubic polynomials

Constant polynomials $C(t) = ae^{-|t|}$ are ACFs of the the scaled OU process, which are monotone. Linear polynomials

$$C(t) = (c_1|t| + c_0)e^{-|t|}$$

cannot have more than one extremum, so they also do not work. Quadratic polynomials are in principle sufficient to create a multi - modal kernel, with a distinct minimum and maximum. But the PSD condition leads to degenerate parabolas and monotone VACFs. The first interesting class to study are the cubic polynomials:

$$C(t) = (|t|^3 + c_2|t|^2 + c_1|t| + c_0)e^{-|t|} \tag{7.12}$$

The Fourier transform of the relevant monomials $g_n(t) = |t|^n e^{-|t|}$ are the following:

$$\begin{aligned}
S_{g_0}(\omega) &= \frac{2}{(1+\omega^2)^4}\left(1 + 3\omega^2 + 3\omega^4 + \omega^6\right) \\
S_{g_1}(\omega) &= \frac{2}{(1+\omega^2)^4}\left(1 + \omega^2 - \omega^4 - \omega^6\right) \\
S_{g_2}(\omega) &= \frac{2}{(1+\omega^2)^4}\left(2 - 4\omega^2 - 6\omega^4\right) \\
S_{g_3}(\omega) &= \frac{2}{(1+\omega^2)^4}\left(6 - 36\omega^2 + 6\omega^4\right)
\end{aligned}$$

7.2.3 Simple cubics: $C(t) = (|t|^3 + a)e^{-|t|}$

I will start with the case of simple cubics $C(t) = (|t|^3 + a)e^{-|t|}$:

The Fourier transform of the monomial $|x|^3 e^{-|x|}$ is

$$S_{g_3}(\omega) = \frac{2\left(6\omega^4 - 36\omega^2 + 6\right)}{(1+\omega^2)^4}$$

and for the constant term $ae^{-|x|}$

$$S_{g_0}(\omega) = \frac{2a}{(1+\omega^2)} = \frac{2a\left(\omega^6 + 3\omega^4 + 3\omega^2 + 1\right)}{(1+\omega^2)^4}.$$

The total Fourier transform thus becomes

$$S_C(\omega) = \frac{2}{(1+\omega^2)^4}\left(a\omega^6 + (6+3a)\omega^4 + (3a-36)\omega^2 + (6+a)\right) = \frac{2P(\omega^2)}{(1+\omega^2)^4}, \quad (7.13)$$

where

$$P(\zeta) = a\zeta^3 + (6+3a)\zeta^2 + (3a-36)\zeta + (6+a) \quad (7.14)$$

is the reduced polynomial of the denominator with $\zeta = \omega^2$. The PSD of C is equivalent to P being non-negative for $\zeta \geq 0$, so it is important to investigate the values of a, for which this condition is satisfied.

If all coefficients of P are strictly non-negative, then P is non-negative for all ζ as well. Thus for $a \geq 6$ the PSD condition is always satisfied. But this threshold is not optimal. Although for $a \sim 6$ the multimodality is still there, the extremal difference

$$C(t_+) - C(t_-)$$

between the local minimum $C(t_-)$ and maximum $C(t_+)$ of C becomes too small. The corresponding VACF does not show any superballisticity.

If $a < 0$, the leading coefficient is negative and the polynomial will become asymptotically negative due to $P(\zeta) \to -\infty$. The Fourier transform S becomes negative at some point, ruling out PSD. What happens for $a \in [0, 6]$?

If $a \neq 0$, the polynomial P is truly cubic with derivative

$$P'(\zeta) = 3a\zeta^2 + 6(a+2)\zeta + (3a-36), \quad (7.15)$$

whose roots are

$$\zeta_\pm = 2\left(\frac{-6(a+2) \pm \sqrt{4a+1}}{6(a+2)}\right) - 1 = \frac{2}{a}\left(-(a+2) \pm 2\sqrt{4a+1}\right). \quad (7.16)$$

Now if $a \geq -\frac{1}{4}$ the derivative has two roots and since $P'(0) = 3(a-12) < 0$ for $a < 12$, ζ_- will be a maximum and ζ_+ a minimum. Both extrema ζ_-, ζ_+ are strictly positive for $a \geq 0$ and due to $P(0) > 0$ for $a > -6$, ζ_- will be the global minimum on $\{\zeta > 0\}$. It suffices to determine ζ_- and $P(\zeta_-)$: When $P(\zeta_-)$ is non - negative, the kernel C is PSD.

There is no closed analytical expression for the roots of $P(\zeta_+)$ with respect to a, which would allow to calculate the limit threshold for a. But the numerical estimate is $a_0 \sim 3.3124$. Whenever $a > a_0$, the kernel C is PSD.

A plot of the kernel C and the corresponding ATAMSD scaling exponent is included in Fig. 7.1. The plot of C and α_{MSD} supports my argument that an increasing regime of C drives α_{MSD} into the superballistic regime. The increase of C does not persist for a sufficiently long time and the power - law scaling $C(t) \sim t^\beta$ is not fulfilled during this regime, so the scaling of the MSD is not stable.

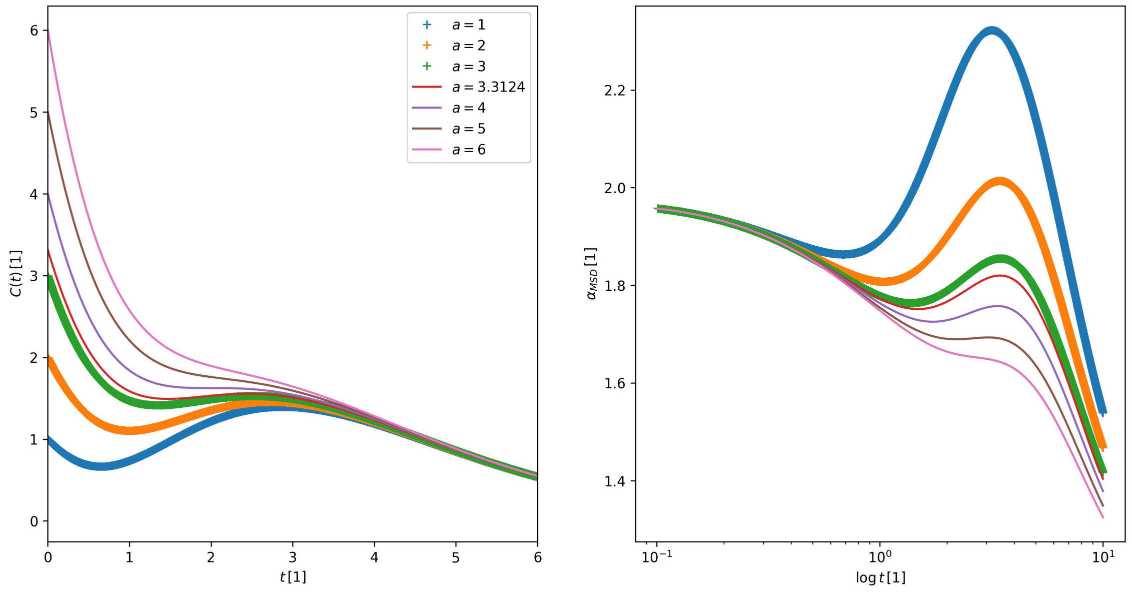

Figure 7.1: Plot of $C(t) = (|t|^3 + a)e^{-|t|}$ and the (hypothetical) MSD scaling exponent α_{MSD}. The functions, for which PSD is violated, are drawn with the + marker.

Leaving the issue of transient scaling aside, the numerical simulation suggests that the MSD does indeed scale **superballistic** for low values of a. But superballisticity is only achieved for the values $a = 1$ and $a = 2$, which all lie below the threshold a_0. The kernel C is not PSD for these values and C can not be a valid stationary VACF.

One reason for the failure could be that ζ_- and ζ_+ lie too close to each other:

Since all the superballistic simple cubics had a large extremal difference

$$C(t_+) - C(t_-),$$

the kernel C has to *bend* very strongly inbetween the extrema, due to their proximity. This seems to be a problem for the PSD property:

If the distance between the extrema could be increased, while maintaining a similar large extremal difference, a more stable scaling could be established. In this case C is still increasing, which allows for a superballistic scaling.

The simple cubics have demonstrated that a superballistic scaling is possible. The next step is to try more general cubic polynomials.

7.2.4 General cubics: $C(t) = (|t|^3 + c_2|t|^2 + c_1|t| + c_0)e^{-|t|}$

For general cubics

$$C(t) = (|t|^3 + c_2|t|^2 + c_1|t| + c_0)e^{-|t|},$$

three parameters can be tuned. The Fourier transform of C is given by

$$S_C = \frac{2}{(1+\omega^2)^4} P(\omega^2) = \frac{2}{(1+\zeta)^4} P(\zeta)$$

with polynomial

$$\begin{aligned} P(\zeta) = &(c_0 + c_1 + 2c_2 + 6c_3) \\ &+ (3c_0 + c_1 - 4c_2 - 36c_3)\,\zeta \\ &+ (3c_0 - c_1 - 6c_2 + 6c_3)\,\zeta^2 \\ &+ (c_0 - c_1)\,\zeta^3. \end{aligned}$$

Verifying the PSD of C is equivalent to showing $P(\zeta) \geq 0$ for $\zeta \geq 0$. Similarly to the previous case, one can compute the minimum ζ_- of $P(\zeta)$ via the derivative $P'(\zeta)$,

$$\begin{aligned} P'(\zeta) = &(3c_0 + c_1 - 4c_2 - 36c_3) \\ &+ (6c_0 - 2c_1 - 12c_2 + 12c_3)\,\zeta \\ &+ (3c_0 - 3c_1)\,\zeta^2, \end{aligned}$$

and check if $P(\zeta_-) \geq 0$.

The parameter space is three - dimensional (c_0, c_1, c_2), so constraining the possible values for which $P(\zeta_-) \geq 0$ holds, is analytically not feasible. The PSD condition has to be checked numerically.

Fig. 7.3 includes the curves C and the corresponding α_{MSD} for three different variations of the initial $c_0 = 4, c_1 = 0, c_2 = 0$. In the first row c_1 is varied, in the second c_2 and in the third row c_1 and c_2 are varied simultaneously.

Varying c_1 alone affects the overall shape of C. For increasing $c_1 \geq 0$, the multimodality of C is lost, whereas negative c_1 leads to a stronger extremal difference Ext $= C(t_+) - C(t_-)$. Varying c_2 does not strongly affect the extremal difference Ext. Instead, a variation in c_2 shifts the position of the extrema to the left ($c_2 < 0$) or to the right ($c_2 > 0$). Larger c_2 - values lead to stronger separation of t_- and t_+, but also changes Ext slightly. This combined effect preserves partly the shape of the *bumpwise increase* in α_{MSD}. A combined variation of c_1 and c_2 allows to change the shape of the *bumpwise increase* in α_{MSD}, while also changing the position slightly.

In all three cases, the curves C with superballlistic MSD were not PSD, so they cannot be a valid VACF. On the other hand, the qualitative picture is the same as for the simple cubics:

The curve C has at most a local minimum and maximum, which allows for an intermediary increase and the possibility of a *bumpwise increase* in α_{MSD}. In the case of general cubics, this bump can be moved (via c_2) and shapewise altered (via c_1). Especially no stable scaling is possible.

One additional possibility, instead of moving to higher - order polynomials, is to look at **rescaled cubics**:

7.2.5 Rescaled cubics: $C_s(t) = \left(|\frac{t}{s}|^3 + c_2|\frac{t}{s}|^2 + c_1|\frac{t}{s}| + c_0\right)e^{-|\frac{t}{s}|}$

One of the drawbacks for cubic polynomials has been the rigidity of the **extremal difference** Ext $= C(t_+) - C(t_-)$ and the **extremal distance** Dis $= t_+ - t_-$. A stable and possibly superballistic scaling presupposes a large domain of C, over which the curve increases. This requires a large extremal distance and difference.

Changing the parameters c_1 and c_2 for cubics only changed the extremal difference, but not the extremal distance. The corresponding scaling became more superballistic, but highly transient.

A major drawback of weighted polynomials is the exponential weighting factor e^{-t}, as this leads to an asymptotic exponential tail. Since e^{-t} decays pretty fast, the curves C for different polynomials do converge pretty quickly. All the curves in Fig. 7.3 converged to the exponential tail at around $t \sim 10$. This exponential convergence makes increasing the extremal distance Dis very difficult and obstructs a stable scaling regime.

One possibility is to increase the order of the polynomial, e.g. starting with quartic polynomials. Another possibility is to look at **rescaled polynomials**:

If

$$C(t) = (|t|^3 + c_2|t|^2 + c_1|t| + c_0)e^{-|t|} = p_3(|t|)e^{-|t|}$$

is some generic cubic polynomial, then the **rescaled polynomial** is defined as

$$C_s(t) = C\left(\frac{t}{s}\right) = p_3\left(\left|\frac{t}{s}\right|\right)e^{-\left|\frac{t}{s}\right|} \tag{7.17}$$

with **scale factor** $s > 0$.

The idea behind the scaling is simple: If C has a given extremal difference and distance, then the rescaled polynomial C_s has a greater extremal distance for $s > 1$, since the exponential damping term $e^{-|t/s|}$ gets suppressed.

This logic is indeed valid, i.e. the extremal distance increases with increasing scale factor $s > 1$. The problem is that this rescaling also changes the extremal difference, and for increasing s the extremal difference is **reduced**. Both effects seem to cancel, as the **shape** of the MSD scaling exponent does not change for different choices of s. A comparison of this is included in Fig. 7.3.

It does not seem that rescaling alters any transient or stable scaling regime. The only effect is a temporal rescaling of the scaling exponent itself. This can be proven analytically:

Theorem 20. *Let C be a VACF and MSD_t the MSD of a corresponding trajectory process x_t. Let $C_s(x) := C(x/s)$ for $s \neq 0$ and MSD_t^s the corresponding MSD. Then*

$$MSD_t^s = s^2 MSD_{t/s} \tag{7.18}$$

and

$$\alpha_{MSD^s}(t) = \alpha_{MSD}(t/s). \tag{7.19}$$

Proof. The identity for the MSDs is straightforward:

$$\begin{aligned} MSD_t^s &= \int_0^t (t-\tau)C(\tau)\,d\tau = \int_0^{t/s} (t-us)C(u)s\,du \\ &= s^2 \int_0^{t/s} (t/s-u)\,C(u)\,du = s^2 MSD_{t/s} \end{aligned}$$

The scaling exponents follow likewise:

$$\begin{aligned}\alpha_{MSD^s}(t) &= \frac{t}{MSD^s_t}\partial_t(MSD^s_t) = \frac{t/s}{MSD_{t/s}}\frac{1}{s}\partial_t(MSD_{t/s})\\ &= \frac{t/s}{MSD_{t/s}}(\partial_t MSD)_{t/s} = \alpha_{MSD}(t/s)\end{aligned}$$

□

Since rescaling can only rescale the scaling exponent temporally, a stable scaling can only be achieved when going to sufficiently higher polynomials.

7.3 Conclusion

I have introduced the weighted polynomials $C(t) = p_n(|t|)e^{-|t|}$ as an introductory example for multi - modal VACFs. The possibility of superballistic scaling has been demonstrated.

Although superballistic scaling has been observed for the MSD in some cases, the scaling was not stable and the corresponding VACFs not PSD. No stable superballistic scaling for valid VACFs has been observed, and it can be true that no superballistic scaling is possible for PSD functions. A definite answer to the possibility of intermediary superballisticity of the MSD for stationary VACFs is therefore not given.

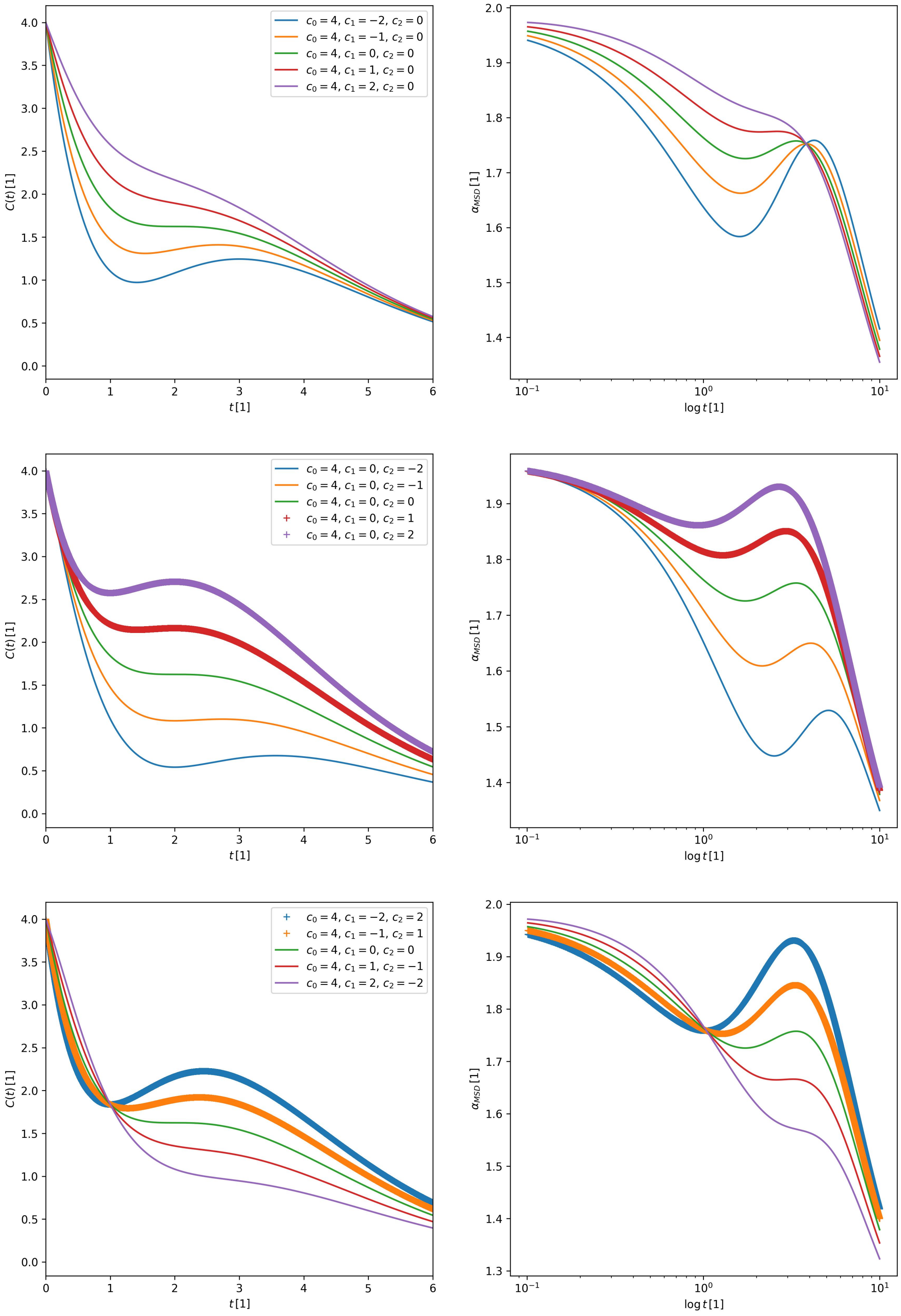

Figure 7.2: Plot of general cubics and the (hypothetical) MSD scaling exponent α_{MSD} for different variations of c_1 and c_2. The functions, for which PSD is violated, are drawn with the + marker.

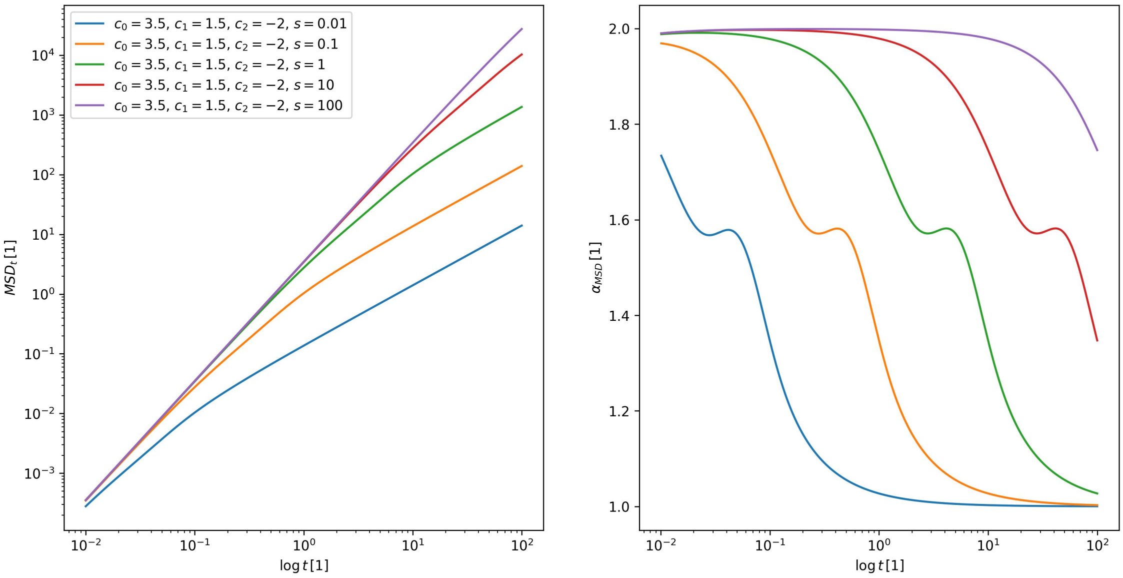

Figure 7.3: MSD and scaling exponent plots for rescaled cubics with different scale factors s.

8 The wild West: Non-stationary VACFs

The EAMSD and ATAMSD scaling of stationary VACFs has been covered in the previous chapter. Since both MSDs are identical, their scaling agrees. The second derivative of the MSD is given by the VACF directly, which allows a very direct characterization of the scaling.

For non-stationary VACFs, on the other hand, the EAMSD and ATAMSD are distinct. Understanding the EAMSD scaling is more complex, but the scaling can still be readily derived from the VACF directly. The ATAMSD, however, is more cryptic. The formulae for the derivatives in Ch. 6 are not given by the VACF, but by off - diagonal averages of the VACF. This obscures any attempts to link the scaling of the VACF directly to the ATAMSD. Another difference is the scaling of ATAMSD in lagtime Δ **and** reduced observational time $T - \Delta$. This *double scaling* can be seen as an annoyance, but it leads to a new understanding of the ATAMSD scaling. I will establish for the model class of **stationary modulated VACFs**, that the scaling w.r.t Δ / $T - \Delta$ is only determined by the *stationary* / *non-stationary* parts of the VACF respectively. This leads to the **split scaling hypothesis**.

8.1 Positive-semidefiniteness for non-stationary VACFs

The verification of the PSD condition in the stationary case can be done analytically, thanks to Bochner's theorem (see Thm. 18). Bochner's theorem only works for 1D functions and there is no generalization to the case of general bivariate kernels $C_v(r, s)$.

There are non-stationary VACFs for which PSD has been confirmed using direct analytically proofs. A well-known example is the ACF of *fractional Brownian motion*,

$$C_v(r, s) = \frac{1}{2}\left(|r|^{2H} + |s|^{2H} - |r - s|^{2H}\right),$$

for Hurst parameter $0 < H < 1$ (see [6], Sec. 4.3).

But coming up with an arbitrary kernel $K(r, s)$ and verifying the PSD by mere luck is not a good basis, upon which one should construct stochastic models. Instead, a well-practiced approach to constructing new PSD kernels (alias VACFs) is to rely on operations that preserve the PSD condition. If $K_i = K_i(r, s)$ are valid PSD kernels, then the following operations preserve the PSD condition and lead to a valid VACF:

1. Positive linear combinations $\sum_i \lambda_i K_i$ with $\lambda_i \geq 0$
2. Pointwise products $K(r, s) := K_i(r, s) \cdot K_j(r, s)$
3. Pointwise limits $K(r, s) := \lim_{n \to \infty} K_n(r, s)$

4. Pointwise modulations $K(r,s) := f(r)f(s)K_i(r,s)$ with $f(0) \geq 0$ and f monotonically increasing

I will illustrate the operation of pointwise modulation:

8.2 Polynomially modulated exponentials: $C_v(r,s) = r^b s^b e^{-|r-s|}$

A simple, but quite illustrative example class for pointwise modulation are the **polynomially modulated exponential** VACFs:

$$C_v(r,s) = r^b s^b e^{-|r-s|} \tag{8.1}$$

8.2.1 Stationary modulated VACFs

Polynomially modulated exponential VACFs are a special case of **stationary modulated** VACFs:

$$C_v(r,s) = \langle v_r v_s \rangle = f(r)f(s)C(r-s) \tag{8.2}$$

$C(t)$ is a stationary VACF, which represents the stationary decorrelation part of the VACF and is commonly associated with the angular (decorrelation) dynamics of the velocity. $f(r)$ is a monotone-driving function, which modulates the stationary dynamics. Indeed, the velocity SMF

$$SMF_t = \langle v_t^2 \rangle = f(t)^2 \tag{8.3}$$

is determined by the driving function f. It represents some persistent, *drift-like*[1] dynamics of the velocity, which determines the behavior of the SMF and variance. The VACF in eq. 8.2 and the corresponding velocity v_t has a very intuitive interpretation:

If u_t is a velocity process with ACF given by $C(\tau)$, then the modulated velocity $v_t := f(t)u_t$ has the ACF of eq. 8.2:

$$C_v(r,s) = \langle v_r v_s \rangle = f(r)f(s)\langle u_r u_s \rangle = f(r)f(s)C(r-s)$$

The velocity u_t has a stationary ACF, which is associated with *equilibriated dynamics*. Think of the EOU process, which represents the noise relaxation of a Brownian particle in a surrounding fluid. The distribution of the EOU process does not change in time and the system as a whole is interpreted being in thermal equilibrium. This interpretation can also be justified for more general velocities: Stationary velocities correspond to the notion of *equilibrium dynamics* in the stochastic setting.

The velocity $v_t = f(t)u_t$ is modulated by the driving function f. The modulated velocity u_t still can be seen as the equilibriated dynamics of v_t, but with a non-stationary modulation. Since f destroys the stationarity of v_t, the modulation of v_t corresponds to the *non-equilibrium dynamics* of v_t. This correspondence with the umbrella term "non-equilibrium" is justified by the velocity SMF $SMF_t = f(t)^2$ itself: In stationary

---

[1] The usage of drift-like should not be confused with the actual drift of a stochastic process. Drift should resemble more the *genetic drift* in evolutionary biology.

/ equilibrium dynamics, the second moment cannot change, since the distribution is invariant.[2] The non - trivial SMF of v_t drives the dynamics out of equilibrium, as the distribution of v_t will change over time, depending on the behavior of f itself.

Stationary modulated VACFs therefore separate into two conceptual different **parts**:

$$C_v(r,s) = \langle v_r v_s \rangle = \underbrace{f(r)f(s)}_{non-stationary} \; \underbrace{C(r-s)}_{stationary} \tag{8.4}$$

This idea of conceptually dividing C_v into a stationary and non-stationary part is important for the second part of this chapter, where the ATAMSD scaling of general non-stationary VACFs is investigated.

The rest of this section explores the EAMSD and ATAMSD scaling analytically for the polynomially modulated exponential VACF in case of $b = 1$.

ALI

Utilizing the integral identities

$$\int_a^b xe^{-x}\,dx = e^{-a}(1+a) - e^{-b}(1+b), \quad \int_a^b xe^{x}\,dx = e^{b}(b-1) - e^{a}(a-1),$$

the averaged lagtime increment (ALI) can be calculated analytically:

$$\begin{aligned}
A_{t,\Delta} &= 2\int_t^{t+\Delta}\int_r^{t+\Delta} rse^{-|r-s|}\,ds\,dr = 2\int_t^{t+\Delta} re^r \int_r^{t+\Delta} se^{-s}\,ds\,dr \\
&= 2\int_t^{t+\Delta} re^r\left(e^{-r}(1+r) - e^{-(t+\Delta)}(1+t+\Delta)\right)dr = 2\int_t^{t+\Delta} r^2 + r - e^{-(t+\Delta)}(1+t+\Delta)re^r\,dr \\
&= \frac{2}{3}\left((t+\Delta)^3 - t^3\right) + \left((t+\Delta)^2 - t^2\right) - 2\int_t^{t+\Delta} e^{-(t+\Delta)}(1+t+\Delta)re^r\,dr \\
&= \frac{2}{3}\left((t+\Delta)^3 - t^3\right) + \left((t+\Delta)^2 - t^2\right) - 2e^{-(t+\Delta)}(1+t+\Delta)\left(e^{t+\Delta}(t+\Delta-1) - e^t(t-1)\right) \\
&= \frac{2}{3}\left((t+\Delta)^3 - t^3\right) - \left((t+\Delta)^2 + t^2\right) + 2\left(1 + e^{-\Delta}(t(t+\Delta) - \Delta - 1)\right)
\end{aligned}$$

It is possible to compute the ALI analytically for other integer b, but I refrain from presenting this here due to the analytical complexity.

EAMSD

The EAMSD $EA_t = A_{0,t}$ consists of three terms:

$$\langle (x_t - x_0)^2 \rangle = \frac{2}{3}t^3 - t^2 + 2\left(1 - e^{-t}(1+t)\right). \tag{8.5}$$

The leading cubic term dictates the asymptotic scaling, since the limit

$$\frac{EA_t}{t^3} = \frac{2}{3} - \frac{1}{t} + \frac{2\left(1 - e^{-t}(1+t)\right)}{t^3} \overset{t\to\infty}{\to} \frac{2}{3}$$

---

[2] The system has already equilibriated.

is finite. The asymptotic scaling exponent is therefore $\alpha_{EA}(\infty) = 3$ and the EAMSD scales asymptotically **cubic**, i.e. superballistic.

The velocity SMF $SMF_t = r^2$ scales ballistically, so the scaling relation in Thm. 8 suggest an initial **quartic** scaling $\alpha_{EA}(0) = 4$. Even though no term in EA_t has an explicit quartic dependence, the initial quartic scaling appears *implicitly* due to a termwise cancellation between the polynomial parts and the transcendental part $e^{-t}(1+t)$:

The remainder term $2\,(1 - e^{-t}(1+t))$ has a series expansion

$$2\,(1 - e^{-t}(t+1)) = \frac{2}{3}t^4 - \frac{2}{3}t^4 + t^2 + \mathcal{O}(t^5),$$

and the leading $\frac{2}{3}t^3$ and subleading t^2 terms cancel the lowest terms in this expansion. The limit

$$\lim_{t\to 0} \frac{EA_t}{t^4} = \frac{2}{3}$$

is finite, which proves the initial quartic scaling $\alpha_{EA}(0) = 4$.

Both of these scaling regimes and scaling exponents can also be obtained via the derivative analysis of

$$\partial_t \langle (x_t - x_0)^2 \rangle = 2t\,(t - 1 + e^{-t})\,.$$

The initial scaling is dominated by the velocity SMF r^2, which leads to a quartic scaling. The stationary exponential $e^{-|r-s|}$ leads to a decorrelation, which begins with the intermediary transient regime. The reduction of the quartic to the cubic scaling for large times is a direct consequence of the exponential decorrelation. For more general integer b, the same picture holds:

The initial scaling exponent is $\alpha_{EA}(0) = 2b + 2$, due to velocity SMF r^{2b}. After a transient regime, which occurs due to the onset of the decorrelation mechanism, the final scaling becomes $\alpha_{EA}(\infty) = 2b + 1$. The stationary decorrelation reduces the scaling exponent by one, i.e.

$$\alpha_{EA}(0) - \alpha_{EA}(\infty) = 1 \tag{8.6}$$

holds for all values of b.

ATAMSD for $b = 1$

The ALI $A_{t,\Delta}$ for $b = 1$ has been computed as

$$A_{t,\Delta} = \frac{2}{3}\left((t+\Delta)^3 - t^3\right) - \left((t+\Delta)^2 + t^2\right) + 2\left(1 + e^{-\Delta}(t(t+\Delta) - \Delta - 1)\right).$$

The ATAMSD

$$\langle \delta^2(\Delta, T) \rangle = \frac{1}{T-\Delta} \int_0^{T-\Delta} A_{t,\Delta}\, dt$$

requires only the resolution of the t-integral. I spare you the lengthy computations, but you have to believe me that the ATAMSD has the following form:

$$\begin{aligned}\langle \delta^2(\Delta, T) \rangle =& \frac{1}{6}\frac{T^4 - \Delta^4 - (T-\Delta)^4}{T-\Delta} - \frac{1}{3}\frac{T^3 - \Delta^3 + (T-\Delta)^3}{T-\Delta} \\ &+ 2\left(1 - e^{-\Delta}(1+\Delta)\right) + e^{-\Delta}\left(\frac{2}{3}(T-\Delta)^2 + (T-\Delta)\Delta\right).\end{aligned}$$

The ATAMSD consists of four terms with different Δ and T scaling. It is not directly obvious in this form, which term will dominate the overall lagtime scaling in the large T limit. The first step is to reduce any explicit T dependence, so that only an implicit $T - \Delta$ occurs. This only affects the first two terms, which are of the generic form $T^n - \Delta^n - (T-\Delta)^n$. The explicit T-dependence can be resolved via $T^n = ((T-\Delta)+\Delta)^n$ and a binomial expansion. The implicit forms are

$$\frac{1}{6}\frac{T^4 - \Delta^4 - (T-\Delta)^4}{T-\Delta} = \frac{2}{3}(T-\Delta)^2\Delta + (T-\Delta)\Delta^2 + \frac{2}{3}\Delta^2$$

and

$$\frac{1}{3}\frac{T^3 - \Delta^3 + (T-\Delta)^3}{T-\Delta} = \frac{2}{3}(T-\Delta)^2 + (T-\Delta)\Delta + \Delta^2$$

respectively. Since only $T - \Delta$-dependencies are involved, the ATAMSD can be regrouped into successive powers of $T - \Delta$:

$$\begin{aligned}\langle\delta^2(\Delta,T)\rangle = &\frac{2}{3}(T-\Delta)^2\left(e^{-\Delta} + \Delta - 1\right)\\ &+(T-\Delta)\left(\Delta e^{-\Delta} + \Delta(\Delta-1)\right)\\ &+\left(2 + \frac{2}{3}\Delta^3 - \Delta^2 - 2(\Delta+1)e^{-\Delta}\right)\end{aligned}$$

You can try to compute the scaling exponent analytically, which is possible, but this only obscures the scaling. Especially, since the fraction appearing α_{ATA} becomes impossible to write down nicely in one line. Instead, one can reason on the scaling using the decomposition in terms of $T - \Delta$ powers:

The last term, which has no $T - \Delta$ dependence, equals the EAMSD. This is just a consequence of $\langle\delta(\Delta,\Delta)\rangle = \langle(x_\Delta - x_0)^2\rangle$. It will not contribute to the final scaling, as the other terms are increasing w.r.t. to $T - \Delta$. Similarly, the second term does not influence the scaling for large T, as it has a linear scaling in $T - \Delta$. The highest power, i.e. the first term with quadratic $T - \Delta$ dependence, dictates the scaling. This has been heuristically justified in Ch. 4, but a theoretical justification can be found later on in Thm. 21. If the quadratic term dominates the scaling, the ATAMSD is effectively given as

$$\langle\delta^2(\Delta,T)\rangle \overset{T\gg\Delta}{\sim} (T-\Delta)^2\left(e^{-\Delta} + \Delta - 1\right). \tag{8.7}$$

But this is just the **MSD** of the integrated EOU process in eq. 4.22!

The VACF of the integrated EOU process is given by the exponential

$$C_v(r,s) = e^{-|r-s|},$$

which is the stationary part of the modulated VACF in eq. 8.1. The ATAMSD scaling w.r.t. Δ is somehow only affected by the stationary kernel, whereas the non - stationary part $f(r)f(s) = rs$ does not contribute anything.

This may only be a coincidence for the case of $b = 1$, but numerical simulations suggest this for *arbitrary* b.

ATAMSD scaling for general b

If the ALI is computed, the ATAMSD can be easily calculated as well for general (integer) b. The problem becomes again that the integrals involved become computationally intractable. In case of $b = 1$, where the formula is analytically known, the ATAMSD for large T is approximately given as

$$\langle \delta^2(\Delta, T) \rangle \sim (T-\Delta)^2 \left(e^{-\Delta} + \Delta - 1 \right).$$

Although there is an explicit $T-\Delta$ dependence, this does not affect the lagtime scaling, since

$$\alpha_{T-\Delta}(\Delta, T) = \frac{\Delta \partial_\Delta (T-\Delta)}{T-\Delta} = \frac{1}{1-T/\Delta} \overset{T \gg \Delta}{\sim} 0 \,.$$

A naive guess would be that for general b, the lagtime scaling is still only given by the stationary part. The *power* with which $T-\Delta$ scales in highest power should only be affected by the non-stationary part. The ATAMSD for large T should be approximately

$$\langle \delta^2(\Delta, T) \rangle \sim (T-\Delta)^{2b} \left(e^{-\Delta} + \Delta - 1 \right). \tag{8.8}$$

A comparative plot of the ATAMSD and scaling exponent computations for different

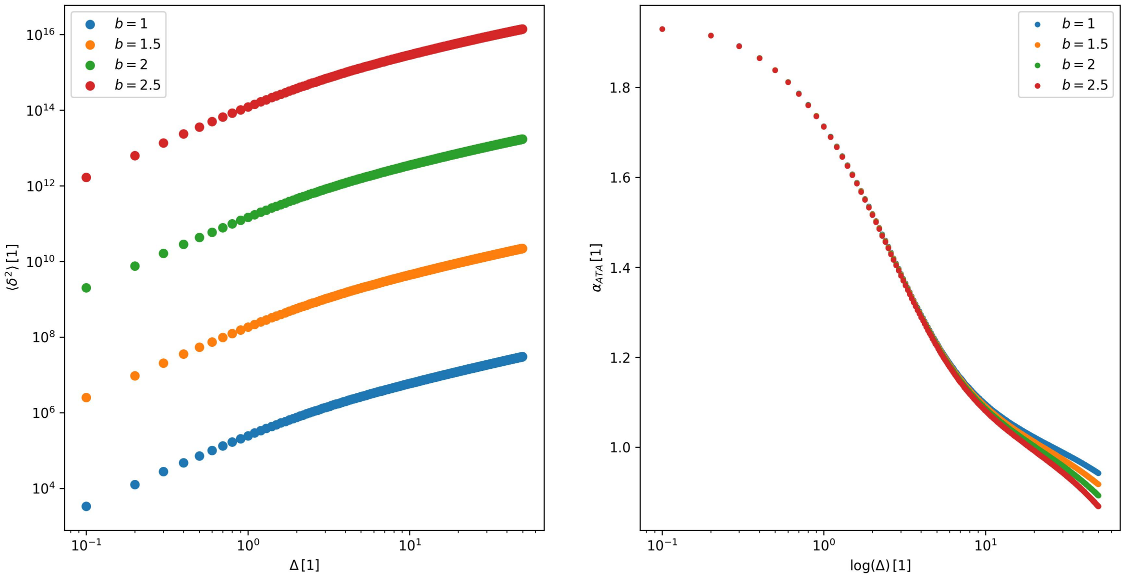

Figure 8.1: Plot of $\langle \Delta^2(\Delta, T) \rangle$ and α_{ATA} for $C_v(r,s) = r^b s^b e^{-|r-s|}$ with $T = 1000$. Although the ATAMSDs differ by a multiplicative constant, the scaling exponents α_{ATA} agree for the different values of b. The difference after $\Delta \sim 10$ is a numerical artifact.

b values is shown in Fig. 8.1. Although the ATAMSDs seem to have different offsets m in the logarithmic plot, the scaling exponent suggest that they indeed have an **identical** lagtime scaling. Although only four values of b have been sampled, I can confidently claim this for all b: *The lagtime scaling is purely dependent on the stationary VACF component $e^{-|r-s|}$*. This explains why the lagtime scaling matches the MSD scaling of the IEOU process in eq. 4.22.

Beside the claim, that the lagtime scaling is only affected by the stationary part, which is numerically supported by Fig. 8.1, also the scaling w.r.t. $T-\Delta$ can be justified:

Indeed, the heuristical approximation

$$\langle \delta^2(\Delta, T)\rangle \sim (T-\Delta)^{2b}\left(e^{-\Delta}+\Delta-1\right)$$

suggests that the scaling of the ATAMSD w.r.t. $T-\Delta$ should be $2b$, i.e.

$$\alpha_{ATA}^{T-\Delta}(\Delta, T) = 2b. \tag{8.9}$$

The offset m of the ATAMSD in the log-log plot of Fig. 8.1 for different values of T is

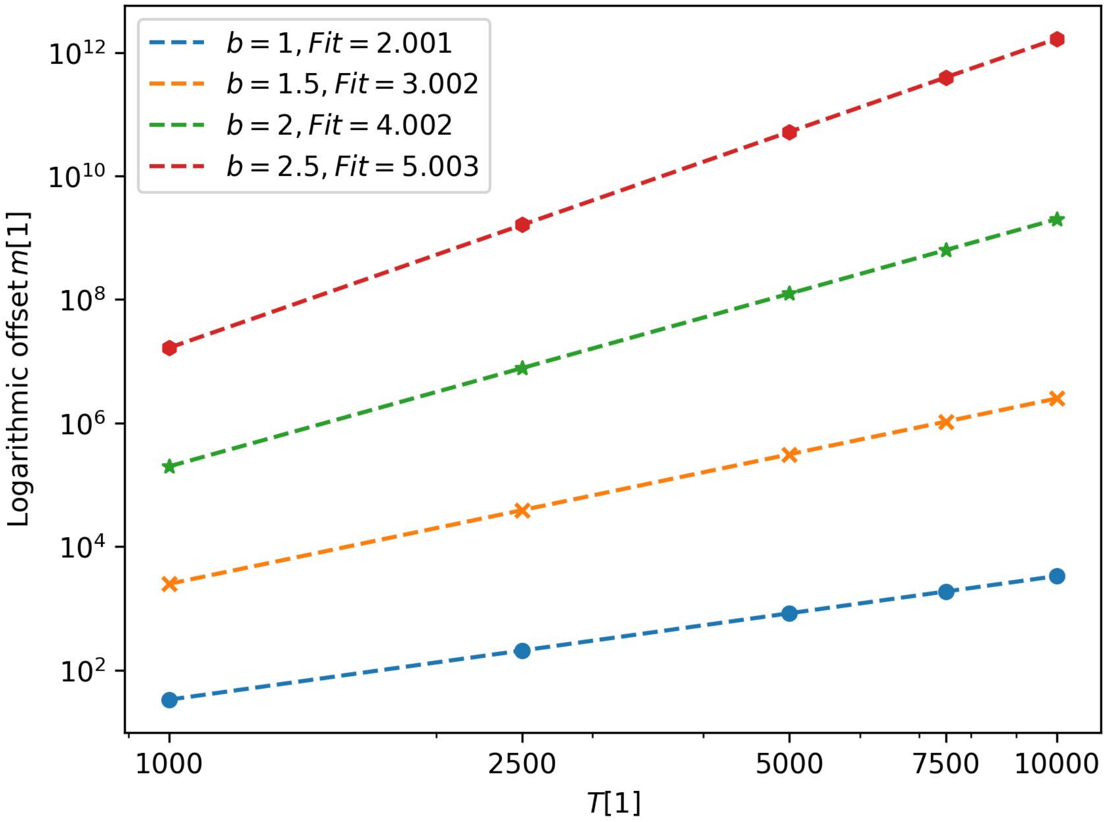

Figure 8.2: Plot of discrete ATAMSD $\delta_{1,m}$ with varying discrete observational time m for $C_v(r,s) = r^b s^b e^{-|r-s|}$. For every b the slope of the log - log plot is exactly $2b$.

shown in Fig. 8.2. Since the curves are linear in the log-log plot, the ATAMSD scales as a power-law w.r.t. $T-\Delta$ for large T. The linear fit of each curve has a slope of roughly $2b$ for each value of b, which suggests that the ATAMSD scales as $(T-\Delta)^{2b} \sim T^{2b}$ w.r.t. T for large observational times.

It appears that the non-stationary VACF part $r^b s^b$ is irrelevant for the lagtime scaling, but entirely determines the scaling w.r.t. the observational time T (or $T-\Delta$). As no effect of the stationary part $e^{-|r-s|}$ changes[3] the scaling w.r.t. T, this can be attributed

---

[3] This may be too vague, but one can repeat the whole numerical simulations for VACFs $C_v(r,s) = f(r)f(s)C(r-s)$ with other stationary parts, e.g. $C(t) = (1+t)^{-\alpha}$. Irrespective of α, the scaling of the ATAMSD offset in the log-log plot w.r.t. T will be the same.

as a purely non-stationary (non-equilibrium) effect.

8.2.2 Scaling hypothesis

The analysis for the non-stationary VACF $C_v(r,s) = r^b s^b e^{-|r-s|}$ suggest the following hypothesis, which I call the **split scaling hypothesis**:

If the VACF can be factorized into a stationary and (purely) non-stationary part, then the ATAMSD scaling also splits. The non-stationary part affects the scaling w.r.t. T and if this scaling is stable, the lagtime scaling depends entirely on the stationary part.

There is a problem in this hypothesis, namely that a decomposition akin to eq. 8.4 is not generally possible. I will therefore confine myself to the class of stationary modulated VACFs first, because the hypothesis can be **proven** under mild assumptions.

Even though such a decomposition is not always possible, many physical VACFs allow for a heuristical separation into **stationary** dynamics, which correspond to some equilibrium behavior and *fast* dynamics of the velocity, and strictly **non-stationary** dynamics, which correspond to some out-of-equilibrium behavior and **slow**[4] dynamics of the velocity. What this means for the scaling of the VACF will be briefly glimpsed upon at the end of this chapter.

8.3 Commensurated powers and slowly-varying remainders

The core idea in proving that the stationary part of the VACF determines the lagtime scaling of the ATAMSD relies on the limit

$$\alpha_{\partial^2_\Delta ATA}(\Delta) = \lim_{T\to\infty} \alpha_{\partial^2_\Delta ATA}(\Delta, T) = \alpha_{C_{stat}}(\Delta), \tag{8.10}$$

where $C_{stat}(\Delta)$ corresponds to the stationary component of the VACF.

Proving this for the comparatively simple case of stationary modulated VACFs is no easy task in itself, and the argument relies on two crucial results: The **principle of commensurated powers** and the **slowly-varyiedness** of the remainder in the power-law decomposition of the driving function f.

8.3.1 Principle of commensurated powers

The core idea in the heuristical derivations of the lagtime scaling has been that the highest power in $T-\Delta$ determines the lagtime scaling in the large T-limit. As always, this is not a rigorous argument. Luckily, it can be made precise. This is the **principle of commensurated powers**:

---

[4] Fast and slow does not necessarily correspond to the usage of fast and slow degrees of freedom in the Mori-Zwanziger theory or similar projection formalisms. The Mori-Zwanziger formalism is deterministic, in the sense that it divides the dynamics (degrees of freedom) of a system into different regimes. The split into stationary / strictly non-stationary parts of the velocity is not based on the dynamics, but on the VACF. Since the VACF measures correlation, but not direct dynamics, one should be careful to not conflate the usage of fast / slow in both cases.

Theorem 21 (Principle of commensurated powers). *Let $f(r,x) = f_1(r)g_1(x)+f_2(r)g_2(x)$ be a bivariate growth function. If $\alpha^x_{g_1}(0) < \alpha^x_{g_2}(0)$, then*

$$\lim_{x\to 0} \alpha^r_f(r,x) = \alpha_{f_1}(r). \tag{8.11}$$

Conversely if $\alpha^x_{g_1}(\infty) < \alpha^x_{g_2}(\infty)$ we have

$$\lim_{x\to \infty} \alpha^r_f(r,x) = \alpha_{f_2}(r). \tag{8.12}$$

Proof. The r - scaling exponent of f is

$$\alpha^r_f = \frac{r(\partial_r f_1)g_1 + r(\partial_r f_2)g_2}{f_1 g_1 + f_2 g_2}.$$

Let $\gamma \in \big(\alpha^x_{g_1}(0), \alpha^x_{g_2}(0)\big)$. The initial limit

$$\lim_{x\to 0} \frac{g_2}{g_1} = \lim_{x\to 0} \frac{g_2}{r^\gamma}\frac{r^\gamma}{g_1} = 0$$

follows from $\lim\limits_{x\to 0} \frac{g_2}{r^\gamma} = 0$ and $\lim\limits_{x\to 0} \frac{g_1}{r^\gamma} = \infty$. The limit of α^r_f follows:

$$\begin{aligned} \alpha^r_f(r,x) &= \frac{r(\partial_r f_1)g_1 + r(\partial_r f_2)g_2}{f_1 g_1 + f_2 g_2} = \frac{r(\partial_r f_1) + r(\partial_r f_2)\frac{g_2}{g_1}}{f_1 + f_2\frac{g_2}{g_1}} \\ &\overset{x\to\infty}{\to} \frac{r\partial_r f_1}{f_1} = \alpha_{f_1}(r). \end{aligned}$$

The proof for $\lim\limits_{x\to\infty} \alpha^r(r,x)$ goes similarly. ☐

The principle of commensurated powers assures that the *weakest x - scaling initially wins*, but *asymptotically only the strongest x - scaling survives.*

Take the example of the ATAMSD

$$\langle \delta^2(\Delta,T)\rangle = (T-\Delta)^2\Delta + (T-\Delta)\Delta^2. \tag{8.13}$$

Thm. 21 assures that the linear term $(T-\Delta)\Delta^2$ dominates the scaling for the $T \to \Delta$ limit, leading to a ballistic lagtime scaling. But the ATAMSD is only meaningful if $T \gg 1$, so the quadratic term $(T-\Delta)^2\Delta$ dictates the scaling. This results in a **diffusive** lagtime scaling of the ATAMSD.

The principle of commensurated powers also works partly if the ATAMSD is given as a power series. If the ATAMSD is given by a series

$$\langle \delta^2(\Delta,T)\rangle = \sum_{n=0}^{\infty} c_n f_n(\Delta)(T-\Delta)^{b-n}$$

with decaying powers in $T-\Delta$, the principle assures that the lagtime scaling in the large T limit is determined by the initial term $c_0 f_0(\Delta)(T-\Delta)^b$. But what happens if $\alpha^x_{g_1}(\infty) \neq \alpha^x_{g_1}(\infty)$? In this case it seems that both terms have an equal power-law scaling, which gives no definite answer. Although the asymptotic scaling exponent can be used infer which function grows stronger asymptotic in terms of power-law behavior, it only suffices to know how the limit

$$\lim_{x\to\infty} \frac{g_1}{g_2}$$

behaves:

Theorem 22. *Let* $f(r,x) = f_1(r)g_1(x) + f_2(r)g_2(x)$ *be a bivariate growth function. If the asymptotic scaling exponents of* g_1 *and* g_2 *agree,*

$$\alpha^x_{g_1}(\infty) = \alpha^x_{g_1}(\infty),$$

and the limit $\lim\limits_{x\to\infty} \frac{g_1}{g_2} \to 0$ *holds, then*

$$\lim_{x\to\infty} \alpha^r_f(r,x) = \alpha_{f_2}(r). \tag{8.14}$$

Proof. The r - scaling exponent of f can be rewritten as

$$\alpha^r_f(r,x) = \frac{r(\partial_r f_1)g_1 + r(\partial_r f_2)g_2}{f_1 g_1 + f_2 g_2} = \frac{r(\partial_r f_1)\frac{g_1}{g_2} + r(\partial_r f_2)}{f_1\frac{g_1}{g_2} + f_2}$$

and the theorem follows from the vanishing quotient $\frac{g_1}{g_2} \to 0$. □

8.3.2 Slowly-varying remainder

If a growth function $f(r)$ has an asymptotic scaling exponent $\alpha_f(\infty) = b$, it admits a factorization $f(r) = r^b h(r)$ in terms of the power-law. This is not spectacular in its own right, as $h(r) = r^b/f(r)$. The import argument is that the remainder has an asymptotic scaling exponent $\alpha_h(\infty) = 0$, which implies

$$\lim_{r\to\infty} \frac{h(r)}{r^\gamma} = 0$$

for any $\gamma > 0$. The remainder has to grow weaker than any polynomial, but the value of the limit $\lim\limits_{r\to\infty} h(r)$ itself is undetermined. It can happen that h converges to a finite value, but the limit can also vanish or diverge. There is no growth estimate on h in the asymptotic regime in general.

For this reason, it is important that the asymptotic scaling exponent $\alpha_f(\infty)$ is actually well-defined. If only the upper index $\overline{\alpha}_f$ exists, the regularity of the remainder h is a priori pretty bad. But the fact that the log-log derivative

$$\alpha_f(r) = \frac{d\log f(r)}{d\log r} \tag{8.15}$$

exists gives enough regularity to the remainder. h has to be **slowly - varying**:[5]

---

[5] The concept of slowly-varying function stems from the theory of **regular variations**. See [19] for a nice introduction.

Theorem 23. *Let f be a differentiable growth function with asymptotic scaling exponent $\alpha_f(\infty) = b$. Then the remainder h in the decomposition*

$$f(r) = r^b h(r) \tag{8.16}$$

is ***slowly-varying****:*

$$\lim_{r\to\infty} \frac{h(cr)}{h(r)} = 1 \tag{8.17}$$

for every $\Delta > 0$.

Proof. The decomposition is well-defined due to $\alpha_f(\infty) = b$ being finite. The logarithm of the factorization gives

$$\log f = b \log r + \log h(r),$$

and the scaling exponent α_f satisfies

$$\alpha_f(r) = b + \frac{d \log h(r)}{d \log r} = b + \alpha_h(r)$$

with $\alpha_h(\infty) = 0$. h itself is non-negative and differentiable, so h and α_h obey the following relation for $c > 0$:

$$\begin{aligned} \log \frac{h(cr)}{h(r)} &= \log h(cr) - \log h(r) = \int_r^{cr} \partial_u \log h(u)\, du \\ &= \int_r^{cr} \frac{u\partial_u \log h(u)}{h(u)}\, du = \int_r^{cr} \frac{\alpha_h(u)}{u}\, du \end{aligned}$$

Since $\alpha_h(\infty) = 0$, there exists a $\delta > 0$ such that for all $r > \delta$

$$|\alpha_h(u)| < \frac{\epsilon}{|\log(c)|}$$

holds. The estimate

$$\begin{aligned} \left|\log \frac{h(cr)}{h(r)}\right| &\le \int_r^{cr} \left|\frac{\alpha_h(u)}{u}\right| du < \frac{\epsilon}{|\log(c)|} \int_r^{cr} \frac{1}{u}\, du \\ &= \frac{\epsilon}{|\log(c)|} (\log(cr) - \log r) = \epsilon \end{aligned}$$

implies $\log \frac{h(cr)}{h(r)} \overset{r\to\infty}{\to} 0$ and likewise $\frac{h(cr)}{h(r)} \overset{r\to\infty}{\to} 1$. □

The **uniform convergence theorem (UCT)** is a consequence of h being slowly-varying:

Theorem 24 (Uniform convergence theorem *(see [2], Thm. 1.2.1)*). *Let h be slowly-varying, then*

$$\lim_{r\to\infty} \sup_{\Delta\in K} \frac{h(r+\Delta)}{h(r)} = 1 \tag{8.18}$$

for every compact set $K \subset \mathbb{R}$.

The UCT implies the uniform convergence of the quotient on compact intervals. This is very important, as most theorems require a locally uniform bound on h.

One consequence of the UCT, which constrains the local growth behavior of h, are the **Potter** bounds:

Proposition 3 (Potter bounds *(see [2], Thm. 1.5.6)*). *Let h be a slowly-varying function. Then for every $\delta > 0$ and $A_\delta > 1$ there exists some $R > 0$, such that for all $r \geq R$ the following bound holds:*

$$A_\delta^{-1}\left(1+\frac{\Delta}{r}\right)^{-\delta} \leq \frac{h(r+\Delta)}{h(r)} \leq A_\delta\left(1+\frac{\Delta}{r}\right)^{\delta} \tag{8.19}$$

Since Prop. 3 gives an upper and lower bound to the growth of h, both estimates can be combined into a singular bound:

Proposition 4. *Let h be a slowly varying function. Then for every $\delta > 0$ and $A_\delta > 1$ there exists a $R > 0$ and $g(r)$, such that for all $r > R$.*

$$\left|\frac{h(r+\Delta)}{h(r)} - 1\right| \leq A_\delta - 1 + \frac{A_\delta \delta \Delta}{r} + g(r). \tag{8.20}$$

The remainder g satisfies

$$|g(r)| \leq \frac{C_3(\delta)}{r^2}. \tag{8.21}$$

Proof. The bound in Prop. 3 implies

$$\left|\frac{h(r+\Delta)}{h(r)} - 1\right| \leq max\left(A_\delta\left(1+\frac{\Delta}{r}\right)^{\delta} - 1, 1 - A_\delta^{-1}\left(1+\frac{\Delta}{r}\right)^{-\delta}\right)$$

Since $\delta > 0$, both terms in the maximum can be expanded via Taylor's theorem in

$$A_\delta\left(1+\frac{\Delta}{r}\right)^{\delta} - 1 = A_\delta - 1 + \frac{A_\delta \delta \Delta}{r} + f_1(r)$$

and

$$1 - A_\delta^{-1}\left(1+\frac{\Delta}{r}\right)^{-\delta} = 1 - A_\delta^{-1} + \frac{A_\delta^{-1}\delta\Delta}{r} + f_2(r)$$

with $|f_1(r)| \leq \left|\frac{C_1(\delta)}{r^2}\right|$ and $|f_2(r)| \leq \left|\frac{C_2(\delta)}{r^2}\right|$. The claimed upper bound follows from $g(r) = max\,(f_1(r), f_2(r))$. □

Another, very important consequence of the UCT is the perseverance of the scaling exponent after integration:

Theorem 25 (Asymptotic scaling exponent of integrals). *Let f be a growth function with asymptotic scaling exponent $\alpha_f(\infty) = b > -1$. Then the integrated*

$$F(r) = \int_{r_0}^{r} f(u)\,du \tag{8.22}$$

for some $r_0 > 0$ has the asymptotic scaling exponent $\alpha_F(\infty) = b + 1$.

Proof. It follows from Prop. 5 that $\alpha_F(\infty) \leq b + 1$. For the reverse bound let $\gamma < b$. For any fixed $r > 0$ and compact K, the UCT implies

$$\sup_{\Delta \in K} \left| \frac{h(r+\Delta)}{h(r)} - 1 \right| = 0.$$

Now let r_K be sufficiently high, such that for all $r > r_K$

$$1 - \epsilon \leq \frac{h(r+\Delta)}{h(r)} - 1 \leq 1 + \epsilon,$$

then the uniform bound becomes

$$h(r+\Delta) \geq h(r)(1-\epsilon)$$

uniformly for $\Delta \in K$ and $r > r_K$. Since K was arbitrary, the choice $K = [-\Delta, 0]$ leads to

$$h(u) \geq h(r)(1-\epsilon)$$

for all $u \in [r - \Delta, r]$. The integral F can therefore be estimated as

$$\begin{aligned} F(r) = \int_{r_0}^{r} u^b h(u)\, du &\geq \int_{r-\Delta}^{r} u^b h(u)\, du \\ &\geq h(r)(1-\epsilon) \int_{r-\Delta}^{r} u^b\, du = h(r)(1-\epsilon)\frac{r^{b+1} - (r-\Delta)^{b+1}}{b+1} \end{aligned}$$

for $r > r_K$. The integral can be bound from below (since $b > -1$) by

$$\frac{F(r)}{r^{\gamma+1}} \geq \frac{h(r)(1-\epsilon)}{b+1} \frac{r^{b+1} - (r-\Delta)^{b+1}}{r^{\gamma+1}} \overset{r\to\infty}{\to} \infty,$$

and the divergence implies $\alpha_F(\infty) \geq b + 1$. ☐

Thm. 25 fails horribly if $b \leq -1$: If $b < -1$, then f is **integrable** and

$$\lim_{r\to\infty} F(r) = \int_{r_0}^{\infty} f(r)\, dr < \infty. \tag{8.23}$$

Since F has a finite limit, the asymptotic scaling exponent of F is simply $\alpha_F(\infty) = 0$.

If $b = -1$, then the integrability of F depends on the behavior of h for large r. It is nonetheless possible to show $\alpha_F(\infty) = 0$. Thm. 25 remains thus true for

$$\alpha_F(\infty) = \max b + 1, 0. \tag{8.24}$$

8.4 Split scaling hypothesis for simple VACFs

This section proves the *split scaling hypothesis* of the ATAMSD in case of stationary modulated VACFs:

$$C_v(r,s) = f(r)f(s)C(r-s)$$

If the driving function f has an asymptotic scaling exponent $\alpha_f(\infty) = b$, the split scaling hypothesis is expressed via the second derivative as the following asymptotic scaling:

$$\partial_\Delta^2 \langle \delta^2(\Delta, T) \rangle \overset{T\gg 1}{\sim} (T-\Delta)^{2b} C(\Delta) \tag{8.25}$$

Heuristical argument

The heuristical argument starts with the AOD:

$$I(\Delta, T) = \frac{1}{T-\Delta}\int_0^{T-\Delta} C_v(r, r+\Delta)\,dr = \frac{C(\Delta)}{T-\Delta}\int_0^{T-\Delta} f(r)f(r+\Delta)\,dr$$

The overlap integral $\int_0^{T-\Delta} f(r)f(r+\Delta)\,dr$ has a priori an explicit dependence on Δ. If the asymptotic scaling exponent $\alpha_f(\infty) = b$ is well-defined, the modulator is given as a power-law $f(r) \sim r^b$ for large r. Since $(r+\Delta)^b \sim r^b$ for large r, the normalized overlap integral becomes

$$\begin{aligned} \frac{1}{T-\Delta}\int_0^{T-\Delta} f(r)f(r+\Delta)\,dr &\sim \frac{1}{T-\Delta}\int_0^{T-\Delta} r^b(r+\Delta)^b\,dr \\ &\sim \frac{1}{T-\Delta}\int_0^{T-\Delta} r^{2b}\,dr = \frac{(T-\Delta)^{2b}}{2b+1} \end{aligned}$$

in leading order of T. This gives the T-scaling. The lagtime scaling follows from the stationary C.

This argument is possible because power-laws have the property that

$$\frac{(r+\Delta)^b}{r^b} \to 1,$$

i.e. the **size** of $r+\Delta$ and r become comparable for large r. This leads to a loss of the Δ-dependence in the leading order for large r. The driving function f does not need to be a strict power-law, so this logic is not exact.

It is crucial that the asymptotic scaling exponent $\alpha_f(\infty) = b$ is well-defined, because this gives the meaningful factorization of f into

$$f(r) = r^b h(r)$$

The pressing problem for the proofs is to handle the remainder h in the integral, which requires a tighter analysis. Since h is **slowly-varying**, the logic of this heuristic argument still guides the proof strategy.

The split scaling heuristics in eq. 8.25 is not mathematically exact, so one has to be careful how to phrase this approximation in exact terms. I have decided to interpret the split scaling hypothesis as two separate postulates:

The **lagtime hypothesis**

$$\alpha_{\partial^2_\Delta ATA}(\Delta) = \lim_{T\to\infty} \alpha_{\partial^2_\Delta ATA}(\Delta, T) = \alpha_C(\Delta) \tag{8.26}$$

and the **observational time hypothesis**

$$\alpha^T_{\partial^2_\Delta ATA}(\infty) = \lim_{T\to\infty} \alpha^T_{\partial^2_\Delta ATA}(\Delta, T) = \lim_{T\to\infty} \frac{d\log \partial^2_\Delta\langle\delta^2(\Delta, T)\rangle}{d\log T} = 2b. \tag{8.27}$$

The problem in translating the results from $\alpha_{\partial^2_\Delta ATA}$ to α_{ATA} stems from the T-limit. Even if one assumes stable scaling or constant $\alpha_{\partial^2_\Delta ATA}(\Delta) = \alpha$, the analysis becomes too convoluted. But given the level of rigor in most of the surveys on this topic, I believe that these two postulates imply the corresponding scaling of $\alpha_{ATA}(\Delta)$ during stable regimes.

Proving the scaling hypothesis is quite technical and it will be done in three levels of increasing difficulty:

1. Monomial rank 1 VACFs: $C_v(r,s) = r^b s^b$
2. General rank 1 VACFs: $C_v(r,s) = f(r)f(s)$
3. Stationary modulated VACFs: $C_v(r,s) = f(r)f(s)C(r-s)$

Is $\alpha_f(\infty) = b$ necessary?

Both the heuristical and analytical version of the split scaling hypothesis assumed a well-defined asymptotic scaling exponent $\alpha_f(\infty) = b$. In the analytical version this is the case because of the estimates within the proofs, which require the identity between $\alpha_f(\infty)$ and $\overline{\alpha}_f$. On a heuristic level, it is the fact that

$$\frac{f(r+\Delta)}{f(r)} \to 1$$

for large r holds. This makes the explicit Δ-dependence in the overlap integral vanish in leading order of $T-\Delta$. The fact that $f(r+\Delta) \sim f(r)$ holds is a consequence of the well-defined $\alpha_f(\infty) = b$, and for diverging α_f the split scaling hypothesis fails:

Taking the example of $C_v(r,s) = (e^r - 1)(e^s - 1)$, the split scaling would suggest

$$\partial^2_\Delta ATA_{\Delta,T} \overset{T\gg 1}{\sim} (T-\Delta)^{2b}$$

for some exponent b and no lagtime dependence, i.e. a persistently **ballistic** ATAMSD scaling. But I showed in Sec. 4.1.2 that the scaling exponent is given by

$$\alpha_{ATA}(\Delta, T) \overset{T\gg 1}{\sim} \frac{2\lambda\Delta}{e^{\lambda\Delta} - 1}$$

in the large T limit, which contradicts the scaling hypothesis. But this is only possible because $\alpha_f(\infty) = \infty$, so the asymptotic scaling exponent is not well-defined. And it should not be, because $f(r) \sim e^r$ asymptotically. The exponential growth is stronger than any power-law scaling and the assumptions of the split scaling hypothesis is violated. Going through the argument in Sec. 4.1.2, the culprit is found in the overlap integral:

$$\int_0^{T-\Delta} e^r e^{r+\Delta}\, dr = e^\Delta \int_0^{T-\Delta} e^{2r}\, dr$$

The overlap integral does not loose its lagtime dependence, and the lagtime hypothesis fails. The assumption $\frac{f(r+\Delta)}{f(r)} \sim 1$ for large r is violated, since

$$\frac{f(r+\Delta)}{f(r)} = \frac{er+\Delta}{e^r} = e^\Delta \neq 1.$$

$f(r+\Delta) = e^r e^\Delta$ is always a factor e^Δ larger than $f(r) = e^r$ and the whole argument breaks down.

8.4.1 Deterministic walkers and Rank 1 VACFs

Definition 29. *A VACF $C_v(r,s)$ is called* ***rank 1*** *if there exists some function f such that C_v is factorizable:*

$$C_v(r,s) = f(r)f(s) \tag{8.28}$$

Rank 1 VACFs are a subclass of the more general stationary modulated VACFs, where the stationary component $C(t) = 1$ is constant. In view of the lagtime hypothesis eq. 8.26, the second derivative should scale as

$$\partial_\Delta^2 \langle \delta^2(\Delta, T) \rangle \overset{T \gg 1}{\sim} (T - \Delta)^{2b}$$

and the ATAMSD persistently **ballistic** for all Δ.

Rank 1 VACFs are not something new in this thesis, as the *deterministic walkers* from Sec. 4.1 had such a VACF:

Let $x_t = f(t)e_\theta$ be the deterministic walker with velocity process $v_t = f'(t)e_\theta$. Then the VACF is just

$$C_v(r,s) = \langle v_r v_s \rangle = f'(r)f'(s)\langle e_\theta^2 \rangle = f'(r)f'(s).$$

The numerical calculations of the ATAMSD in Fig. 4.1 suggested a persistent **ballistic scaling**:[6]

$$\alpha_{ATA}(\Delta, T) \overset{T \gg \Delta}{\sim} 2,$$

which agrees with the claim of the lagtime hypothesis. The persistent ballistic scaling of the ATAMSD, despite vastly different choices of f, has been a source of endless agony for me. Understanding the persistently ballistic ATAMSD in face of a superballistic EAMSD has been the main driving force for the theoretical odyssey from Ch. 5 onwards.

8.4.2 Monomials $C_v(r,s) = r^b s^b$

The monomials $C_v(r,s) = r^b s^b$ are the easiest case, for which the split scaling hypothesis can be verified. But do not get fooled - the argument is still quite technical. The proof strategy for this and the general case is the following:

The second derivative of the ATAMSD is given as the AOD,

$$\partial_\Delta^2 \langle \delta^2(\Delta, T) \rangle \overset{T \gg 1}{\sim} 2 \cdot I(\Delta, T),$$

in highest order in T.[7] For the monomial VACFs, the AOD reduces to a normalized overlap integral:

$$I(\Delta, T) = \frac{1}{T - \Delta} \int_0^{T-\Delta} (r + \Delta)^b r^b \, dr \tag{8.29}$$

---

[6] Only if f itself has an asymptotic stable scaling. Take $f(r) = (e^r - 1)^2$ as a classical counterexample.

[7] This may seem as a new approximation, but it can be made mathematically precise by the interested reader. The other subleading terms can be neglected.

The split scaling hypothesis then follows from the estimate

$$I(\Delta, T) \sim (T-\Delta)^{2b}, \tag{8.30}$$

which relies on estimating the integrand:

Theorem 26 (Split scaling for $C_v(r,s) = r^b s^b$)**.** *Let $C_v(r,s) = r^b s^b$ be the VACF of the process x_t. Then the ATAMSD of x_t fulfills the split scaling hypothesis:*

$$\lim_{T\to\infty} \alpha_{\partial^2_\Delta ATA}(\Delta) = 0 \tag{8.31}$$

and

$$\alpha^T_{\partial^2_\Delta ATA}(\infty) = 2b \tag{8.32}$$

Proof. The overlap integral in the AOD can be decomposed into two parts:

$$\begin{aligned}
I(\Delta, T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} (r+\Delta)^b r^b \, dr \\
&= \frac{1}{T-\Delta}\int_\Delta^{T-\Delta} (r+\Delta)^b r^b \, dr + \frac{1}{T-\Delta}\int_0^{\Delta} (r+\Delta)^b r^b \, dt \\
&= \frac{I_a}{T-\Delta} + \frac{I_R}{T-\Delta}
\end{aligned}$$

The remainder I_R has no explicit $T-\Delta$ dependence. Since I_a has an increasing dependence in $T-\Delta$, the I_R-term becomes negligible due to Thm. 21.

The split into I_a and I_R is quite helpful as $r > \Delta$ for all r in I_a. In this case the infinite binomial expansion

$$(r+\Delta)^b = r^b\left(1+\frac{\Delta}{r}\right)^b = r^b \sum_{k=0}^{\infty} \binom{b}{k} \Delta^k r^{-k}$$

is convergent. Since $(r+\Delta)^b r^b$ is analytical, the sum and integral can be interchanged:

$$\begin{aligned}
I_a = \int_\Delta^{T-\Delta} (r+\Delta)^b r^b \, dr &= \int_\Delta^{T-\Delta} \sum_{k=0}^{\infty} \binom{b}{k} \Delta^k r^{2b-k} \, dr \\
&= \sum_{k=0}^{\infty} \binom{b}{k} \Delta^k \int_\Delta^{T-\Delta} r^{2b-k} \, dr \\
&= \sum_{k=0}^{\infty} \binom{b}{k} \Delta^k \left(\int_0^{T-\Delta} r^{2b-k} \, dr - \int_0^{\Delta} r^{2b-k} \, dr \right) \\
&= \sum_{k=0}^{\infty} \binom{b}{k} \Delta^k \left(I_k - I_{k,R} \right)
\end{aligned}$$

The integrals $I_{k,R}$ have no explicit $T-\Delta$ dependence and I_k have the asymptotic scaling exponent $\alpha_{I_k}(\infty) = 2b + 1 - k$. The AOD is split into three parts:

$$I(\Delta, T) = \underbrace{\sum_{k=0}^{\infty} \binom{b}{k} \Delta^k \frac{I_k}{T-\Delta}}_{a)} - \underbrace{\sum_{k=0}^{\infty} \binom{b}{k} \Delta^k \frac{I_{k,R}}{T-\Delta}}_{b)} - \underbrace{\frac{I_R}{T-\Delta}}_{c)}$$

The rest b) and c) vanish in the large T limit, so they can be ignored. Term a) is given by an infinite series, but each individual term has the asymptotic scaling exponent $\alpha_{I_k}(\infty) = 2b + 1 - k$. The term with $k = 0$ dominates the scaling due to Thm. 21. But

$$I_0 = \Delta^0 \frac{I_0}{T - \Delta} = \frac{(T - \Delta)^{2b}}{2b + 1}$$

has **no** explicit Δ-dependence. The claims in eq. 8.31 and 8.32 follow. □

8.4.3 Stable $C_v(r, s) = f(r)f(s)$

The next, more general case is for stable rank 1 VACFs $C_v(r, s) = f(r)f(s)$. Here stable assumes $\alpha_f(\infty) = b$, which guarantees that f can be meaningfully factorized as

$$f(r) = r^b h(r). \tag{8.33}$$

The general proof idea is to reduce the argument to the previous case, namely by showing that the overlap integral

$$\int_0^{T-\Delta} (r + \Delta)^b r^b \, dr$$

dominates the scaling. But with the remainder, the overlap integral is a priori given by

$$\int_0^{T-\Delta} (r + \Delta)^b r^b h(r)h(r + \Delta) \, dr.$$

The basic argument is to split up the integral and use the asymptotic growth properties of h to get the estimate.

First I demonstrate this for the case that the limit $h_\infty = \lim_{r\to\infty} h(r)$ is finite and then for general remainders.

Converging remainder

Theorem 27 (Split scaling for $C_v(r, s) = f(r)f(s)$ for convergent h.). *Let $C_v(r, s) = f(r)f(s)$ be the VACF of the process x_t with $\alpha_f(\infty) = b > -1$. If the remainder h in*

$$f(r) = r^b h(r) \tag{8.34}$$

has a finite limit $h_\infty = \lim_{r\to\infty} h(r)$, the ATAMSD of x_t fulfills the split scaling hypothesis:

$$\alpha_{\partial^2_\Delta ATA}(\Delta) = 0 \tag{8.35}$$

and

$$\alpha^T_{\partial^2_\Delta ATA}(\infty) = 2b \tag{8.36}$$

The proof of Thm. 27 requires the following proposition:

Proposition 5. *Let $f(r)$ be a continuous function with*

$$\lim_{r\to\infty} \frac{f(r)}{r^\gamma} = 0.$$

Then

$$\lim_{r\to\infty} \frac{\int_0^r f(s)\,ds}{r^{\gamma+1}} = 0. \tag{8.37}$$

Proof of Prop. 5. Since $\frac{f(r)}{r^\gamma} \to 0$, for a fixed ϵ there is some $\delta_0 > 0$ such that for all $r > \delta_0$ the estimate

$$|f(r)| < \frac{\epsilon}{2} r^\gamma$$

holds. The integral can be estimated by

$$\frac{|\int_0^r f(s)\,ds|}{r^{\gamma+1}} < \frac{\epsilon}{2} + \frac{|\int_0^{\delta_0} f(s)\,ds|}{r^{\gamma+1}}.$$

The only problem is the remaining integral. Since f is continuous and locally integrable, there exists some $\delta > \delta_0$, such that

$$|\int_0^{\delta_0} f(s)\,ds| < \frac{\epsilon}{2}\delta^{\gamma+1}.$$

Then for every $r > \delta$ the second integral can be bounded as well:

$$\frac{|\int_0^{\delta_0} f(s)\,ds|}{r^{\gamma+1}} < \frac{\epsilon}{2}\left(\frac{\delta}{r}\right)^{\gamma+1} < \frac{\epsilon}{2}$$

The claim follows. □

Proof of Thm. 27. The AOD can be decomposed into two parts:

$$\begin{aligned} I(\Delta,T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} f(r+\Delta)f(r)\,dr \\ &= \frac{1}{T-\Delta}\int_0^{T-\Delta} (r+\Delta)^b r^b h(r+\Delta)h(r)\,dr \\ &= \underbrace{\frac{h_\infty^2}{T-\Delta}\int_0^{T-\Delta} (r+\Delta)^b r^b\,dr}_{=:V_a} + \underbrace{\frac{1}{T-\Delta}\int_0^{T-\Delta} (r+\Delta)^b r^b \left(h(r+\Delta)h(r) - h_\infty^2\right)\,dr}_{=:V_b} \end{aligned}$$

The integral V_a is the AOD of $f(r) = r^b$ form Thm. 26. Since $\lim_{T\to\infty} \frac{V_a}{(T-\Delta)^{2b}} = C \neq 0$, it suffices to show

$$\lim_{T\to\infty} \frac{V_b}{(T-\Delta)^{2b}} = 0$$

due to Thm. 21. Since the limit is of the form

$$\lim_{T\to\infty} \frac{\int_0^{T-\Delta}(r+\Delta)^b r^b \left(h(r+\Delta)h(r) - h_\infty^2\right)\,dr}{(T-\Delta)^{2b+1}} = 0,$$

Prop. 5 proves the argument if

$$\lim_{r\to\infty} \frac{(r+\Delta)^b r^b \left(h(r+\Delta)h(r) - h_\infty^2\right)}{r^{2b}} = 0 \tag{8.38}$$

holds. The individual limits $\lim\limits_{r\to\infty} \frac{(r+\Delta)^b r^b}{r^{2b}} = 1$ and $\lim\limits_{r\to\infty} h(r+\Delta)h(r) = h_\infty^2$ hold, so the limit in eq. 8.38 follows. □

Slowly-varying remainder

The proof in Thm. 27 relies crucially on the decomposition of the AOD into V_a and V_b. Although this decomposition is always possible, the limit

$$\lim_{T\to\infty} \frac{V_b}{(T-\Delta)^{2b}} = 0$$

is only given if $\lim\limits_{r\to\infty} h(r) = h_\infty$. Since this is not necessarily true, the proof strategy has to be revisited for more general remainders. Fortunately, Potter's bounds and the special version Prop. 4 suffices for estimating the remainder:

Theorem 28 (Split scaling for $C_v(r,s) = f(r)f(s)$ for general h.)**.** *Let $C_v(r,s) = f(r)f(s)$ be the VACF of the process x_t with $\alpha_f(\infty) = b > -1$. Then the ATAMSD of x_t fulfills the split scaling hypothesis:*

$$\alpha_{\partial_\Delta^2 ATA}(\Delta) = 0 \tag{8.39}$$

and

$$\alpha^T_{\partial_\Delta^2 ATA}(\infty) = 2b \tag{8.40}$$

Proof of Thm. 28. The AOD for slowly varying remainder can be decomposed into two parts:

$$\begin{aligned} I(\Delta,T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} (r+\Delta)^b r^b h(r)h(r+\Delta)\,dr = \frac{1}{T-\Delta}\int_0^{T-\Delta} r^b(r+\Delta)^b h^2(r)\frac{h(r+\Delta)}{h(r)}\,dr \\ &= \underbrace{\frac{1}{T-\Delta}\int_0^{T-\Delta} r^b(r+\Delta)^b h^2(r)\,dr}_{=:P_a} + \underbrace{\frac{1}{T-\Delta}\int_0^{T-\Delta} r^b(r+\Delta)^b h^2(r)\left(\frac{h(r+\Delta)}{h(r)} - 1\right)dr}_{=:P_b} \end{aligned}$$

The first part P_a can be decomposed similarly to I_a in the case of $C_v(r,s) = r^b s^b$:

$$\begin{aligned} P_a &= \frac{1}{T-\Delta}\int_0^{T-\Delta} r^b(r+\Delta)^b h^2(r)\,dr = \frac{1}{T-\Delta}\int_0^{T-\Delta} r^{2b}h^2(r)\left(1+\frac{\Delta}{r}\right)^b dr \\ &= \frac{1}{T-\Delta}\int_\Delta^{T-\Delta} r^{2b}h^2(r)\left(1+\frac{\Delta}{r}\right)^b dr - \frac{1}{T-\Delta}\int_0^{\Delta} r^{2b}h^2(r)\left(1+\frac{\Delta}{r}\right)^b dr \\ &= \sum_{k=0}^{\infty}\binom{b}{k}\Delta^k \frac{1}{T-\Delta}\int_\Delta^{T-\Delta} r^{2b-k}h^2(r)\,dr - \frac{P_a, R}{T-\Delta} = \sum_{k=0}^{\infty}\binom{b}{k}\Delta^k P_{a,k} - \frac{P_{a,R}}{T-\Delta} \end{aligned}$$

$P_{a,R}$ has no explicit $T-\Delta$ dependence, so the last term vanishes in the large T limit. The individual terms

$$P_{a,k} = \frac{1}{T-\Delta} \int_{\Delta}^{T-\Delta} r^{2b-k} h^2(r)\, dr \tag{8.41}$$

have an asymptotic scaling exponent $\alpha_{P_{a,k}}(\infty) = 2b-k$ due to Thm. 25 . The scaling of P_a is therefore determined by $P_{a,0}$. But this does not determine the scaling of $I(\Delta, T)$ yet, as P_b has a priori the same scaling exponents:

$$\begin{aligned}
P_b &= \frac{1}{T-\Delta} \int_{0}^{T-\Delta} r^{2b} h^2(r) \left(1+\frac{\Delta}{r}\right)^b \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr \\
&= \frac{1}{T-\Delta} \int_{\Delta}^{T-\Delta} r^{2b} h^2(r) \left(1+\frac{\Delta}{r}\right)^b \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr \\
&- \frac{1}{T-\Delta} \int_{0}^{\Delta} r^{2b} h^2(r) \left(1+\frac{\Delta}{r}\right)^b \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr \\
&= \frac{1}{T-\Delta} \int_{\Delta}^{T-\Delta} r^{2b} h^2(r) \left(1+\frac{\Delta}{r}\right)^b \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr - \frac{P_{b,R}}{T-\Delta} \\
&= \sum_{k=0}^{\infty} \binom{b}{k} \Delta^k \frac{1}{T-\Delta} \int_{\Delta}^{T-\Delta} r^{2b-k} h^2(r) \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr - \frac{P_{b,R}}{T-\Delta} \\
&= \sum_{k=0}^{\infty} \binom{b}{k} \Delta^k P_{b,k} - \frac{P_{b,R}}{T-\Delta}
\end{aligned}$$

The scaling exponent of $t(r) = \frac{h(r+\Delta)}{h(r)} - 1$ is in principle $\alpha_t(\infty) \leq 0$, but this inequality is not necessarily strict. The scaling exponent $\alpha_{P_{b,k}} \leq 2b-k$ of $P_{b,k}$ can match those of $P_{a,k}$. In this case it is necessary to show

$$\lim_{T\to\infty} \frac{P_{b,k}}{P_{a,k}} = 0. \tag{8.42}$$

For fixed $\delta > 0, A_\delta > 1$, the estimate in Prop. 4 splits $P_{b,k}$ into four terms:

$$\begin{aligned}
\left|P_{b,k}\right| &= \left| \int_{\Delta}^{T-\Delta} r^{2b-k} h^2(r) \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr \right| \\
&\leq \left| \int_{\Delta}^{R} r^{2b-k} h^2(r) \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr \right| + \left| \int_{R}^{T-\Delta} r^{2b-k} h^2(r) \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr \right| \\
&\leq \left| \int_{\Delta}^{R} r^{2b-k} h^2(r) \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr \right| + C_1(\delta) \int_{\Delta}^{R} r^{2b-k} h^2(r)\, dr \\
&+ C_2(\delta) \Delta \int_{\Delta}^{R} r^{2b-(k+1)} h^2(r)\, dr + \int_{\Delta}^{R} r^{2b} h^2(r) g(r)\, dr
\end{aligned}$$

The quotient likewise splits in four terms:

$$\left|\frac{P_{b,k}}{P_{a,k}}\right| = \frac{\left|\int_\Delta^R r^{2b-k} h^2(r) \left(\frac{h(r+\Delta)}{h(r)} - 1\right) dr\right|}{\int_\Delta^{T-\Delta} r^{2b-k} h^2(r)\, dr} + C_1(\delta) \underbrace{\frac{\int_R^{T-\Delta} r^{2b-k} h^2(r)\, dr}{\int_\Delta^{T-\Delta} r^{2b-k} h^2(r)\, dr}}_{a)}$$

$$+ C_2(\delta)\Delta \underbrace{\frac{\int_R^{T-\Delta} r^{2b-(k+1)} h^2(r)\, dr}{\int_\Delta^{T-\Delta} r^{2b-k} h^2(r)\, dr}}_{b)} + \underbrace{\frac{\int_R^{T-\Delta} r^{2b} h^2(r) g(r)\, dr}{\int_\Delta^{T-\Delta} r^{2b-k} h^2(r)\, dr}}_{c)}$$

The first term has a vanishing T-limit, as the numerator is independent of T. Since $g(r) \leq \frac{C_3(\delta)}{r^2}$ for $r \geq R$, the terms $b)$ and $c)$ vanish in the limsup. The same holds for the first term. Only term $a)$ survives as

$$\limsup_T \left|\frac{P_{b,k}}{P_{a,k}}\right| \leq C_1(\delta) \limsup_T \int_\Delta^{T-\Delta} r^{2b-k} h^2(r)\, dr. \tag{8.43}$$

The limsup on the right side is independent of δ and A_δ, whereas the constant $C_1(\delta) = A_\delta - 1$ can be made arbitrarily small. This implies

$$\limsup_T \left|\frac{P_{b,k}}{P_{a,k}}\right| = 0$$

and the limit

$$\lim_T \left|\frac{P_{b,k}}{P_{a,k}}\right| = 0.$$

Thm. 21 implies that $P_{a,0}$ dictates the scaling of the AOD and the claim follows. □

8.4.4 Subordinated processes and Stationary modulated VACFs

Definition 30. *A VACF $C_v(r,s)$ is called a* ***stationary modulated VACF*** *if*

$$C_v(r,s) = f(r) f(s) C(r-s) \tag{8.44}$$

for non-negative f and a stationary kernel C with $C(0) = 1$.

The concept of stationary modulated VACFs has been already introduced in Sec. 8.2.1. This class contains the rank 1 VACFs for the special choice of $C(\tau) = 1$, i.e. when no stationary (de)correlation mechanism is present in the underlying velocity dynamics. But just like the rank 1 VACFs entailed the case of deterministic walkers, the stationary modulated VACFs can partly explain the ATAMSD scaling of **subordinated processes**:[8]

Given a process x_t, the subordinated process y_t is defined by

$$y_t := x_{f(t)},$$

---

[8] The concept and scaling of subordinated processes has been investigated in Sec. 4.2.

where f is a strictly monotone, differentiable function with $f(0) = 0$. The process y_t has the same dynamics as x_t, but viewed using a different time scale, which is being altered by the subordinator $f(t)$. The velocity of x_t, v_t, and the velocity y_t, u_t, are directly related:

$$u_t = \frac{d}{dt}y_t = \frac{d}{dt}x_{f(t)} = \frac{dx}{dt}_f(t)\partial f = (\partial_t f)_t v_{f(t)}.$$

This relation between u_t and v_t likewise relates the VACF of x_t and y_t:

$$\begin{aligned} C_u(r,s) &= \langle u_r u_s \rangle = \langle (\partial_t f)_r v_{f(r)} (\partial_t f)_s v_{f(s)} \rangle \\ &= f'(r)f'(s)\langle v_{f(r)} v_{f(s)} \rangle = f'(r)f'(s)C_v(f(r), f(s)). \end{aligned}$$

Upon a time rescaling,

$$C_u(f^{-1}(r), f^{-1}(s)) = f'(f^{-1}(r))f'(f^{-1}(s))C_v(r,s),$$

the VACF looks strikingly like a stationary modulated VACF. The conceptual problem here is that the VACF $C_u(f^{-1}(r), f^{-1}(s))$ is itself given with rescaled time.

8.4.5 Stationary modulated VACFs $C_v(r,s) = f(r)f(s)e^{-|r-s|}$

Although the stationary modulated VACFs seem more complex than generic rank 1 VACFs, the proof of the split scaling hypothesis follows directly from Thm. 28:

Theorem 29 (Split scaling for $C_v(r,s) = f(r)f(s)C(r-s)$.). *Let $C_v(r,s) = f(r)f(s)C(r-s)$ be the VACF of the process x_t with $\alpha_f(\infty) = b^{\geq} - 1$. Then the ATAMSD of x_t fulfills the split scaling hypothesis:*

$$\alpha_{\partial^2_\Delta ATA}(\Delta) = \alpha_C(\Delta) \tag{8.45}$$

with $\alpha_C(\Delta)$ being the scaling exponent of the stationary C and

$$\alpha^T_{\partial^2_\Delta ATA}(\infty) = 2b. \tag{8.46}$$

Proof. The AOD of C_v can be factorized as

$$\begin{aligned} I(\Delta, T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} C_v(r+\Delta, r)\,dr = \frac{1}{T-\Delta}\int_0^{T-\Delta} f(r)f(r+\Delta)C(\Delta)\,dr \\ &= \frac{C(\Delta)}{T-\Delta}\int_0^{T-\Delta} f(r)f(r+\Delta)\,dr = C(\Delta)I_f(\Delta, T) \end{aligned}$$

Here I_f is **identical** to the AOD for the underlying VACF $C_v(r,s) = f(r)f(s)$. Since the lagtime scaling exponent of the AOD is given by

$$\alpha_I(\Delta, T) = \alpha_C(\Delta)\alpha_{I_f}(\Delta, T), \tag{8.47}$$

the lagtime hypothesis is a consequence from Thm. 28:

$$\alpha_{\partial^2_\Delta ATA}(\Delta) = \lim_{T\to\infty} \alpha_{\partial^2_\Delta ATA}(\Delta, T) = \alpha_C$$

$C(\Delta)$ has no T-dependence and the observational time hypothesis follows. ☐

8.4.6 Higher VACFs: $C_v(r,s) = \sum_{i=1}^n f_i(r) f_i(s) C_i(r-s)$

Stationary modulated VACFs have been considered as *toy model VACFs*, since they admit a nice factorization into stationary / non-stationary parts:

$$C_v(r,s) = \langle v_r v_s \rangle = \underbrace{f(r)f(s)}_{non-stationary}\ \underbrace{C(r-s)}_{stationary}$$

In general, such a clean split is not possible. There is no unique non-stationary or stationary dynamics for general VACFs, and the presence of multiple non-stationary makes the matter of scaling more complex. The next larger class that can show such a behavior are the **higher stationary modulated VACFs**:

$$C_v(r,s) = \sum_{i=1}^{n} f_i(r) f_i(s) C_i(r-s) = \sum_{i=1} D_i \tag{8.48}$$

Higher modulated VACFs are nothing more than sums of individual stationary modulated VACFs D_i. Since the modulators f_i are allowed to have widely different scaling behavior, the clear split in stationary / non-stationary dynamics is washed away. But this is not a problem in itself. Although the EAMSD scaling of a corresponding trajectory process x_t is indeed influenced by each D_i individually, the ATAMSD scales (in accordance with Thm. 21) via the *"winner takes it all"*-principle - only the strongest non-stationary part influences the scaling the large T limit:

Theorem 30 (Split scaling for rank n stationary modulated VACFs.). *Let*

$$C_v(r,s) \sum_{i=1}^{n} f_i(r) f_i(s) C_i(r-s)$$

be the VACF of the process x_t with $\alpha_{f_i}(\infty) = b_i$. Assume there exists a ***dominating*** *modulator f_j with $\alpha_{f_j}(\infty) \geq -1$, such that $b_i < b_j$ or*

$$\lim_{r\to\infty} \frac{f_i(r)}{f_j(r)} = 0$$

holds for all $i \neq j$. Then the ATAMSD of x_t fulfills the split scaling hypothesis:

$$\alpha_{\partial_\Delta^2 ATA}(\Delta) = \alpha_C(\Delta) \tag{8.49}$$

with $\alpha_{C_j}(\Delta)$ being the scaling exponent of the stationary part C_j and

$$\alpha^T_{\partial_\Delta^2 ATA}(\infty) = 2b_j. \tag{8.50}$$

Proof. The VACF $C_v(r,s) = \sum_{i=1}^n f_i(r) f_i(s) C_i(r-s) = \sum_{i=1}^n D_i(r,s)$ is naturally decomposed into individual VACFs

$$D_i(r,s) = f_i(r) f_i(s) C_i(r-s).$$

Since the AOD is linear in C_v, it likewise splits into the AODs of the individual VACFS:

$$
\begin{aligned}
I(\Delta, T) &= \frac{1}{T-\Delta} \int_0^{T-\Delta} C_v(r, r+\Delta)\, dr = \sum_{i=1}^{n} \frac{1}{T-\Delta} \int_0^{T-\Delta} D_i(r, r+\Delta)\, dr \\
&= \sum_{i=1}^{n} I_{D_i}(\Delta, T)
\end{aligned}
$$

Thm. 21 implies that I_{D_j} dominates the scaling and the lagtime / observational time hypotheses follow from Thm. 29. □

8.5 General split scaling hypothesis

Motivated by the split scaling hypothesis for the rather special class of higher modulated VACFs, I do believe that the following hypothesis can explain the ATAMSD lagtime and failure of large T convergence observed in the literature:

The second derivative of the ATAMSD is given in highest order as

$$\partial_\Delta^2 ATA_{\Delta,T} \overset{T \gg 1}{\sim} N(T-\Delta) S(\Delta). \tag{8.51}$$

N represents the leading non-stationary contribution of the VACF and dominates the scaling w.r.t. T. The non-stationary part N has a counterpart in the EAMSD, so it can be used to estimate the non-stationary contribution of the EAMSD scaling,in terms of the ATAMSD scaling w.r.t. T.

S is the stationary contribution connected to the non-stationary N and forms a stationary VACF. The lagtime scaling of $\partial_\Delta^2 ATA_{\Delta,T}$ (and of ATA during stable regimes) is only dependent on the stationary S, which represents the stationary (de)correlation mechanism accompanying the non-stationary contribution N.

8.6 The issue of subdiffusion

Why do I emphasize the split scaling hypothesis so much? Isn't superballistic motion in itself rather unphysical, so the scope of this hypothesis should not affect the regimes of diffusion normally seen in experiments?

This is valid objection, but note that the split scaling hypothesis does not restrict to the superballistic regime. It can shed light on the **subdiffusive regime** as well:

Consider the rank 1 VACF

$$C_v(r, s) = (1+r)^{-b}(1+s)^{-b} \tag{8.52}$$

for $b > 0$. This corresponds to a deterministic walker x_t with algebraically decreasing velocity

$$v_t = (1+t)^{-b} e_\theta. \tag{8.53}$$

Although the VACF is non-stationary, the second derivative of the EAMSD is given by

$$\partial_t^2 EA_t \sim t^{-2b}$$

This suggests an EAMSD scaling $EA_t \sim t^{2(1-b)}$ for large t. This would imply a subdiffusive scaling for $1/2 < b < 1$, but an asymptotically **constant** EAMSD for $b = 1$ and asymptotically **shrinking** EAMSD for $b > 1$.

The subdiffusive scaling for $1/2 < b < 1$ sounds plausible, but the constant EAMSD for $b = -1$ is strange. The velocity is not integrable

$$\int_0^t |v_r|^2 \, dr = \log(1+t) \to \infty,$$

so one would expect x_t to wander to infinity for large t and not remain confined.

Likewise, the shrinking EAMSD for $b < -1$ is unphysical. The velocity is integrable, which forbids each x_t from wandering to infinity. But since each velocity v_t grows away from 0, the EAMSD should become at least **constant**. A contracting EAMSD would suggest a sort of contraction of x_t, such that the realizations of x_t progressively move back to the origin x_0. The velocity v_t does not change sign / direction, so each trajectory **cannot** move back. This contradicts the shrinking of EA_t.

But why does this happen? Let's analyze this carefully:

EAMSD for $b = -1$

The ALI for $b = -1$ is given by

$$\langle (x_{t+\Delta} - x_t)^2 \rangle = \left(\int_t^{t+\Delta} (1+r)^{-1} \, dr \right)^2 = \log^2 \left(1 + \frac{\Delta}{1+t} \right), \tag{8.54}$$

and the EAMSD as

$$\langle (x_t - x_0)^2 \rangle = \log^2 (1+t) , \tag{8.55}$$

which supports the asymptotic divergence. The second derivative suggested

$$\partial_t^2 EA_t \sim t^0,$$

but this is only heuristical. It tells you the leading power in t, but neglects subpolynomial growth factors, like the logarithm in this case.

EAMSD for $b \neq -1$

For $b \neq -1$, the ALI can be easily computed:

$$\begin{aligned} langle\, (x_{t+\Delta} - x_t)^2 \rangle &= \left(\int_t^{t+\Delta} (1+r)^{-b} \, dr \right)^2 = \left(\frac{(1+t+\Delta)^{1-b} - (1+t)^{1-b}}{1-b} \right)^2 \\ &= \frac{1}{(b-1)^2} \left((1+t+\Delta)^{1-b} - (1+t)^{1-b} \right)^2 \\ &= \frac{1}{(b-1)^2} \left((1+t+\Delta)^{2(1-b)} - 2(1+t)^{1-b}(1+t+\Delta)^{1-b} + (1+t)^{2(1-b)} \right) \end{aligned}$$

The EAMSD is then just

$$\langle (x_t - x_0)^2 \rangle = \frac{1}{(b-1)^2} \left((1+t)^{2(1-b)} - 2(1+t)^{1-b} + 1 \right) . \tag{8.56}$$

If $b < 1$, the power of the first term $2(1-b)$ is the highest and the EAMSD is given as

$$EA_t \overset{t\gg 1}{\sim} t^{2(1-b)}$$

for large t. This is in correspondence with the derivative argument.

But if $b > 1$, the first and second term have **negative** powers due to $1-b < 0$. Since the third term $1 = t^0$ has a higher power, it dominates the asymptotic scaling. The EAMSD becomes asymptotically constant (or **localized**):

$$\lim_{t\to\infty} EA_t = \frac{1}{(b-1)^2} \tag{8.57}$$

Th subdiffusive behavior for $b < 1$ is related to **trapping**, i.e. anomalous diffusion in crowded environments. An example for this kind of motion are mRNAs diffusing in the eucaryotic cells: The cytpolasm within cells is not a homogeneous fluid, so the movement cannot be modeled using naive Brownian motion. Due to collisions of the mRNA with molecules of comparable size (e.g. enzymes or vesicles) or organelles, the motion of the mRNA is suppressed. Characterizing this subdiffusivity is very crucial for understanding cell behavior, since the EAMSD scaling exponents α of different constituent molecules influences chemical reaction rates within the cell.

The toy model VACF produced an EAMSD

$$EA_t \overset{t\gg 1}{\sim} \begin{cases} t^{2(1-b)} & \text{for } b < 1 \\ \log^2 t & \text{for } b = 1 \\ \frac{1}{(b-1)^2} & \text{for } b > 1 \end{cases} \tag{8.58}$$

with quite different asymptotic scaling, depending on b. How does this compare to the ATAMSD?

ATAMSD

The split scaling hypothesis suggested an ATAMSD scaling

$$ATA_{\Delta,T} \overset{T\gg 1}{\sim} (T-\Delta)^{2b}\Delta^2$$

for a rank 1 VACF $C_v(r,s) = f(r)f(s)$ with $\alpha_f(\infty) = b \geq -1$. This does not include the case of

$$C_v(r,s) = (1+r)^{-b}(1+s)^{-b},$$

since the power-law exponent is negative.

It is possible to adapt the proof strategy of Thm. 26, as r^b and $(1+r)^{-b}$ are both asymptotic power-laws. The problem lies in the range of b and the estimate of the dominating term:

In the proof of Thm. 26, the scaling argument for the T-scaling relied on estimating the scaling of

$$I_k = \int_0^{T-\Delta} r^{2b-k}\, dr. \tag{8.59}$$

In the proof the scaling exponent of I_k was $\alpha_{I_k}(\infty) = 2b + 1 - k$ and this allowed to compare the different contributions. The scaling exponent can be derived using Thm. 25, but this is only true for $b > -1$.

For the modulator $(1+r)^{-b}$, the integral becomes

$$I_k \overset{T\gg 1}{\sim} \int_{r_0}^{T-\Delta} (1+r)^{-2b-k}\, dr.$$

Even for $k = 0$, the integrand scales as r^{-2b} and the proof strategy does not work for $b > \frac{1}{2}$. What changes here?

In case of $b \neq 1/2$ (and $b \neq -1$), the ALI is approximately given by

$$\begin{aligned}
\langle (x_{t+\Delta} - x_t)^2 \rangle &= \frac{1}{(b-1)^2} \left((1+t+\Delta)^{1-b} - (1+t)^{1-b} \right)^2 \\
&= \frac{(b-1)^{-2}}{(1+t)^{2b}} \left((1+t+\Delta) \left(\frac{1+t}{1+t+\Delta} \right)^b - (1+t) \right) \\
&\overset{t\gg\Delta 1}{\sim} \frac{(b-1)^{-2}}{(1+t)^{2b}} (1+t+\Delta-1+t) \\
&= \frac{\Delta^2}{(b-1)^2(1+t)^{2b}}
\end{aligned}$$

The approximation of the ALI can be integrated analytically, which suggests

$$\begin{aligned}
\delta^2(\Delta, T) &= \frac{1}{T-\Delta} \int_0^{T-\Delta} \langle (x_{t+\Delta} - x_t)^2 \rangle \, dt \\
&\sim \frac{\Delta^2}{(b-1)^2} \frac{1}{T-\Delta} \int_0^{T-\Delta} (1+t)^{-2b}\, dt \\
&= \frac{\Delta^2}{(1-2b)(b-1)^2} \left((T-\Delta)^{-2b} - (T-\Delta)^{-1} \right).
\end{aligned}$$

For $b < 1/2$, the scaling exponent $-2b$ is greater than -1, so the T-scaling is dominated by $(T-\Delta)^{-2b}$. The ATAMSD scales in accordance with the split scaling hypothesis as

$$\langle \delta^2(\Delta, T) \rangle \sim (T-\Delta)^{-2b} \Delta^2.$$

Although one could expect the same to be true for $b > 1/2$, $(T-\Delta)^{-2b}$ is not the dominating term anymore due to $-2b < -1$. In this case $(T-\Delta)^{-1}$ dominates the scaling, **regardless** of the specific value of b. Another way to deduce this is to track the signs of $(T-\Delta)^{-2b}$ and $(T-\Delta)^{-1}$, which change for $b > 1/2$. Since $(T-\Delta)^{-1}$ is the positive term here, it has to dominate the scaling. Otherwise the ATAMSD would become asymptotically negative.

If $b = 1/2$ the ALI can still be approximated as

$$\langle (x_{t+\Delta} - x_t)^2 \rangle \sim \frac{\Delta^2}{(b-1)^2(1+t)},$$

but the ATAMSD gets a logarithmic correction:

$$\begin{aligned}\delta^2(\Delta,T) &= \frac{1}{T-\Delta}\int_0^{T-\Delta} \langle (x_{t+\Delta}-x_t)^2\rangle \, dt \\ &\sim \frac{\Delta^2}{(b-1)^2}\int_0^{T-\Delta}\int_0^{T-\Delta} (1+t)^{-1}\, dt \\ &= \frac{\Delta^2}{(b-1)^2}\frac{\log(T-\Delta+1)}{T-\Delta} \sim \frac{\Delta^2}{(b-1)^2}\frac{\log(T-\Delta)}{T-\Delta}\end{aligned}$$

Localization and trapping

I explained that a subdiffusive $EA_t \sim t^{2(1-b)}$ often occurs in crowded environments in biology, where the different particle masses can be varying over multiple length scales. This kind of motion is sometimes termed **trapping** and plays a significant in biophysics. Indeed, experiments suggest that in vitro tracer particles with Brownian diffusion can act subdiffusive in vivo, i.e. in live tissue.

Likewise, an asymptotically constant EAMSD $EA_t \sim$ const relates to some **localization** of the trajectories. Localization can occur if the diffusion is being performed in a confined space, e.g. a cell with non-permeable membranes. If the particle cannot escape through the cell boundaries, the motion is confined to the cell and the EAMSD saturates at the cell boundary.

Trapping and localization are therefore two very important patterns that emerge for anomalous diffusion in living tissue. Understanding their behavior is important, but also the experimental verification for these two phenomena.

Since the EAMSD is expensive to estimate in experiments, the TAMSD of individual tracer particles is studied instead. But in this case there is a conundrum:

Even though the (A)TAMSD lagtime scaling disagrees with the EAMSD scaling for superballistic processes, the T-scaling can be utilized to roughly estimate the EAMSD scaling. In case of $C_v(r,s) = r^b s^b$ this was possible due to

$$EAMSD_t = t^{2b}$$

and

$$ATA_{\Delta,T} \overset{T\gg\Delta}{\sim} (T-\Delta)^{2b}\Delta^2.$$

But this is not as straightforward here. The EAMSD and ATAMSD scaling is approximately the following:

$$EA_t \overset{t\gg 1}{\sim} \begin{cases} t^{2(1-b)} & \text{for } b<1 \\ \log^2 t & \text{for } b=1 \\ \frac{1}{(b-1)^2} & \text{for } b>1 \end{cases} \tag{8.60}$$

$$ATA_{\Delta,T} \overset{T\gg\Delta}{\sim} \begin{cases} \frac{\Delta^2}{(T-\Delta)^{2b}} & \text{for } b<1/2 \\ \frac{\Delta^2\log(T-\Delta)}{(T-\Delta)} & \text{for } b=1/2 \\ \frac{\Delta^2}{(T-\Delta)} & \text{for } b>1/2 \end{cases} \tag{8.61}$$

For $b < 1/2$ there is a correspondence, since the exponents of Δ and $(T - \Delta)$ add up to $2(1 - b) = 2 - 2b$, which is the EAMSD scaling exponent. This superdiffusive case satisfies the split scaling hypothesis.

But for $b > 1/2$, the EAMSD scales subdiffusive. Although the EAMSD scaling exponent $2(1-b)$ changes with b, the ATAMSD saturates and scales as Δ/T for $T \gg 1$. The ALI is integrable and the integral

$$\int_0^{T-\Delta} \langle (x_{t+\Delta} - x_t)^2 \rangle \, dt \overset{T\to\infty}{\to} C$$

saturates to a fixed value C. Since the ATAMSD is the weighted integral

$$\frac{1}{T-\Delta} \int_0^{T-\Delta} \langle (x_{t+\Delta} - x_t)^2 \rangle \, dt,$$

the ATAMSD decays as T^{-1} for all $b > 1/2$. The approximated ATAMSD

$$\langle \delta^2(\Delta, T) \rangle \sim \frac{\Delta^2}{(1-2b)(b-1)^2} \left((T-\Delta)^{-2b} - (T-\Delta)^{-1} \right) \tag{8.62}$$

has T^{-1} as the dominant scaling due to $-2b < -1$.

The idea that the scaling is dominated by the leading term in the large T limit is valid in case of $f(r) = (1+r)^{-b}$, just like it applied to the case of $f(r) = r^b$ before. The pressing difference here is that $f(r) = r^b$ leads to a **diverging** ATAMSD

$$\lim_{T\to\infty} \langle \delta^2(\Delta, T) \rangle = \infty$$

for each $\Delta > 0$, whereas in case of $f(r) = (1+r)^{-b}$ the ATAMSD is **asymptotically vanishing**:

$$\lim_{T\to\infty} \langle \delta^2(\Delta, T) \rangle = 0$$

For the diverging ATAMSD, most of the terms are positive powers T^a in T. Even if the ATAMSD is given by

$$\langle \delta^2(\Delta, T) \rangle = f_1(\Delta) T^a + f_2(\Delta) T^b,$$

Thm. 21 assures that the scaling is dominated by the first term if $a > b$. The value of T, after which the first term becomes dominates, is entirely dependent on the values of the exponents a and b. But if both are positive, the convergence is very rapid. This allows to estimate the ATAMSD at lagtime Δ accurately for values of $T \sim 10 \cdot \Delta$.

If the ATAMSD is asymptotically vanishing, the scaling in T is given by inverse powers T^{-a} like in the case of eq. 8.62:

$$\langle \delta^2(\Delta, T) \rangle \sim \frac{\Delta^2}{(1-2b)(b-1)^2} \left((T-\Delta)^{-2b} - (T-\Delta)^{-1} \right)$$

Although T^{-2b} or T^{-1} will eventually determine the scaling due to Thm. 21, the convergence is pretty slow for inverse powers. This introduces numerical instabilities when computing the discretizations. These instabilities can render the lagtime scaling completely unusable: In Fig. 8.3 the numerical ATAMSD and scaling exponent

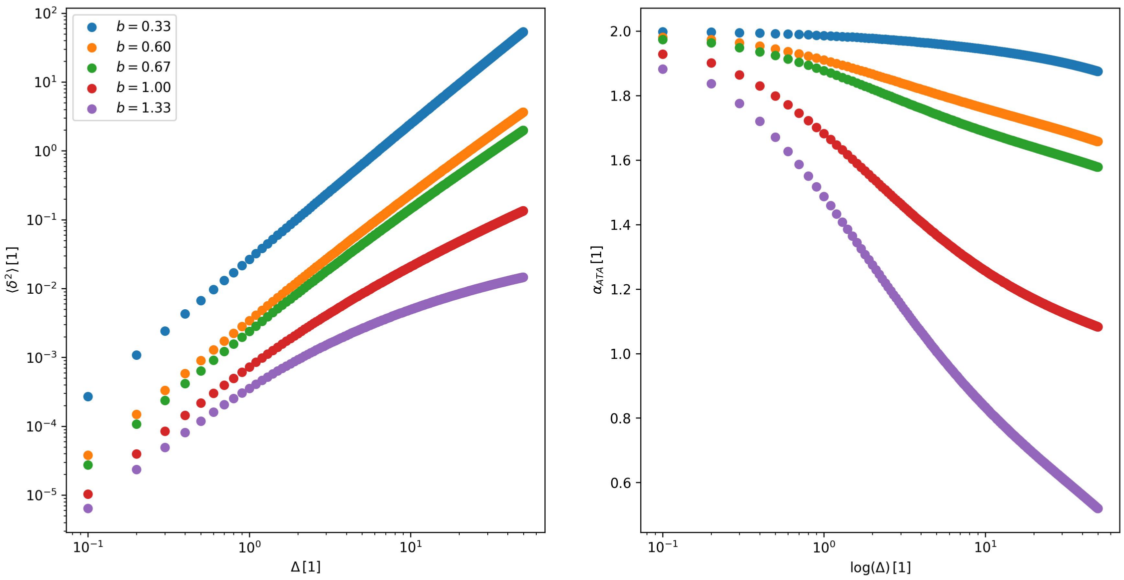

Figure 8.3: $ATA_{\Delta,T}$ and α_{ATA} for VACF $C_v(r,s) = (1+r)^{-b}(1+s)^{-b}$. The discretization has been run with $dt = 0.1$ and $T = 10000$.

are shown for time step $dt = 0.1$ and observational time $T = 10000$ for different values of b. Eq. 8.61 implies a persistent ballistic scaling

$$\langle \delta^2(\Delta, T)\rangle \overset{T\gg 1}{\sim} \Delta^2$$

for large T. This scaling is only valid for large observational times, although the experience from previous models suggests that $T \sim 10 \cdot \Delta$ can be sufficient for computing $ATA_{\Delta,T}$ accurately. The scaling exponents Fig. 8.3 suggest a **decorrelation** of the trajectories, since the scaling exponent decays away from 2. But this scaling is fictitious!

Fig. 8.4 shows the ATAMSD and scaling exponent for the particular choice of $C_v(r,s) = (1+r)^{-3/4}(1+s)^{-3/4}$ with differing values of the step size dt. The decay of the scaling exponent away from ballisticity seems to correlate with dt, which suggest numerical instabilities as the underlying reason for the subballisticity. Low values for T and high values for dt leads to a decaying ATAMSD.

This can be put aside as a numerical instability, but it also affects experimental data:

The qualitative behavior of the lagtime scaling of the (A)TAMSD is often used instead of the EAMSD in experiments due to small sample sizes. But if numerical instabilities affect the lagtime scaling, time resolution and length of the data set can severely alter the lagtime scaling. In case of Fig. 8.3, the integrable VACFs with $b > 1/2$ seem to show some plateauing in the ATAMSD, i.e. a vanishing lagtime scaling. Comparing the ATAMSD behavior with the EAMSD would suggest **localization** since $ATA_{\Delta,T} \sim$ const, although the corresponding EAMSD is subdiffusive.

Despite its obscure nature, the lagtime scaling itself is highly volatile to changes in the time resolution and measurement length. This is in contrast to the T-scaling of

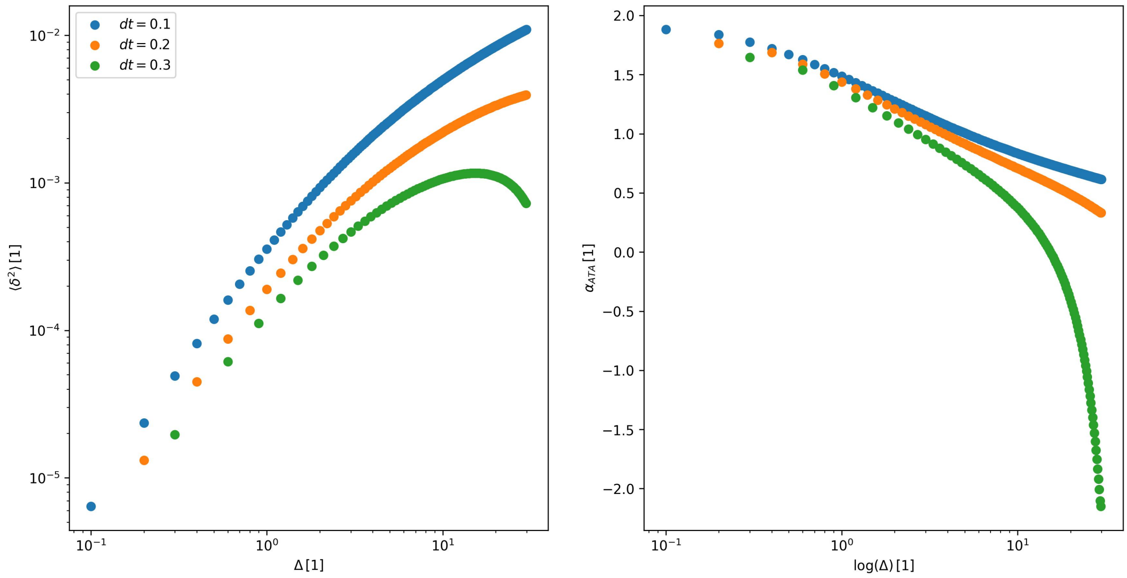

Figure 8.4: $ATA_{\Delta,T}$ and α_{ATA} for VACF $C_v(r,s) = (1+r)^{-3/4}(1+s)^{-3/4}$. The discretization has been run with $T = 10000$ and varying stepsize dt.

the ATAMSD, which is pretty stable. Although

$$\langle \delta^2(\Delta, T) \rangle \sim \frac{\Delta^2}{(1-2b)(b-1)^2} \left((T-\Delta)^{-2b} - (T-\Delta)^{-1} \right)$$

suggests at least two stable regimes due to $(T-\Delta)^{-2b}$ and $(T-\Delta)^{-1}$, the T-scaling is quite resilient:

Fig. 8.5 shows that despite low T-values, the T-scaling is numerically stable. It can be expected that the same holds for experimental data, i.e. the T-scaling is robust against finite resolution effects.

T-scaling

The T-scaling of the ATAMSD appear more robust in the case of $C_v(r,s) = (1+r)^{-b}(1+r)^{-s}$. Such a VACFs correspond to a subdiffusive or localized EAMSD and is therefore more connected to real biological motility patterns than the superballistic VACFs previously. Since the T-scaling agrees with the split scaling hypothesis for $b < 1/2$, the whole interpretation with non-stationary VACF components still holds. The T-scaling exponent of the ATAMSD is a better estimator for α_{EA} than the lagtime scaling exponent α_{ATA}.

A detailed study of the (A)TAMSD in the theoretical and experimental setting is therefore only meaningful when lagtime and T-scaling are studied simultaneously.

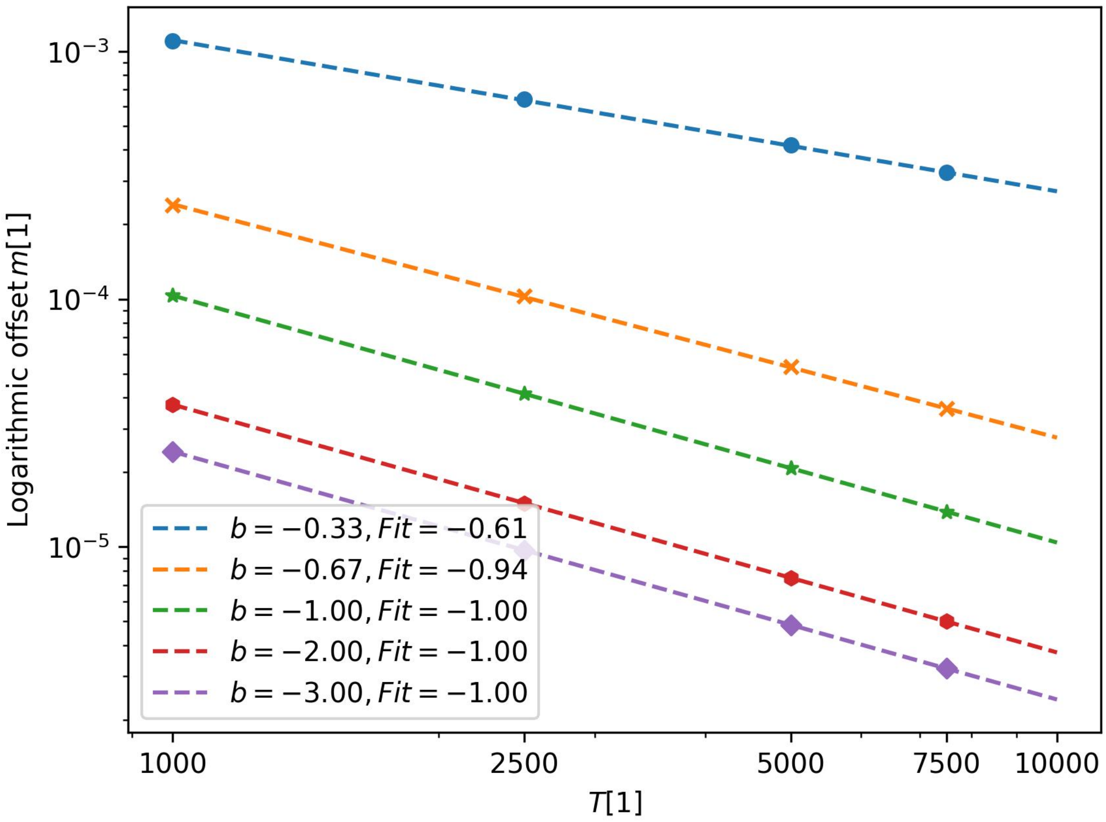

Figure 8.5: Offset m in $\log ATA_{\Delta,T} = m \cdot \log \Delta + c$ for varying observational time T. The linear slope corresponds to γ in $ATA_{\Delta,T} \sim (T - \Delta)^{\gamma} \Delta^{\alpha_{ATA}}$.

9 Summary and Outlook

Summary

The goal of this thesis has been to study the behavior of the EAMSD and (A)TAMSD for (superballistic) stochastic processes, with an emphasis on analytical results. The overarching theme has been *power-law scaling* $f(r) \sim r^{\beta}$. A considerable amount of the thesis has been spent on developing the analytical framework, in which these MSDs can be adequately studied.

In Ch. 2 and Ch. 3 the theoretical groundwork is laid, upon which a rigorous notion of power-law scaling $f(r) \sim r^{\beta}$ can be built. The usage of *scaling exponent* and *power-law behavior* is intuitively understood, but ambiguously defined. Hence, I tried to define and explain my usage of these concepts in a self - contained manner. The introduction of the upper and lower indices $\overline{\alpha}_f, \underline{\alpha}_f$ and their relation to the (dynamical) scaling exponent α_f was already heuristically known, but the analytical proof and conceptual framework within this thesis has not been made explicit in the literature. It is this analytical link that is used throughout the consequent work to establish theoretical results, beyond the heuristical level of rigor commonly used in the literature. The whole analysis has been made for the class of *growth functions* f, which extends the area of application of the results beyond the EAMSD and ATAMSD.

Ch. 4 stood out from the rest of the thesis, as it focused more on concrete models than the overarching principle and theory. The three classes of processes (deterministic walkers, subordinated processes and integrated processes) have been used to demonstrate the variability of scaling for the EAMSD. The contrast of EAMSD and (A)TAMSD scaling for the examples demonstrated, why the difference in scaling is not as easy as it seems. Differing (A)TAMSD scaling exponents and the concept of *ergodicity breaking* is addressed in the current literature, but the severity is often understated. This is not surprising, since the scaling exponents of both EAMSD and ATAMSD tend to disagree mildly.[1] The rest of the thesis has been devowed to understanding this scaling disagreement in terms of analytical arguments.

The example chapter is followed by another pair of twin chapters: Interlude I in Ch. 5 analyzed the scaling exponents in the small time limit for both EAMSD and (A)TAMSD in form of the initial scaling exponent $\alpha(0)$. For differentiable processes, the (A)TAMSD has to scale ballistically for small times. This follows from the existence of a pathwise velocity. The EAMSD in contrast, scales differently, depending on the distribution of the initial velocity v_0. For stationary velocities, both initial scaling exponent agree. Due to the emphasis on stationary velocities in the literature, the *er-*

---

[1] Most classical surveys consider subballistic processes, so both exponents are restricted to $0 \leq \alpha_{EA}, \alpha_{ATA} \leq 2$. Given the split scaling hypothesis, both exponent agree in many cases.

godicity breaking for small times is rarely observed. But it should be more surprising that both exponent agree, rather than disagree. These results have shown that the scaling of the (A)TAMSD and the EAMSD already should disagree for small times. This is often overlooked in literature survey (see [18]) due to the focus on the asymptotic scaling.

The *velocity autocorrelation function (VACF)* $C_v(r,s) = \langle v_r v_s \rangle$, an important property of differentiable processes, is introduced and studied in Interlude II in Ch. 6. Both EAMSD and ATAMSD can be expressed analytically via iterated integrals of the VACF. Since their formulas are conceptually very different, I spent most of this chapter on studying them separately. I showed that the higher derivatives of the both MSDs can be related more directly to the VACF, which allows to translate scaling arguments of the VACF partly to the MSDs themselves. The (more or less) heuristical relation of the MSD scaling exponent with the scaling exponents of the higher derivatives has been introduced and formed an important theme throughout the rest of the thesis. The formulas for the EAMSD have already been known and extensively studied in the stationary case, but to my knowledge, there have not been any major attempts to study the ATAMSD scaling in terms of the VACF from the analytical side. Especially the discretization scheme of the ATAMSD in terms of two recursions appears to have been unknown in the diffusive literature on the subject. The techniques are applied to the case of stationary velocities, for which both EAMSD and ATAMSD agree. Using the properties of the VACF, I was able to prove that a superballistic scaling is impossible for the initial ($t \ll 1$) and asymptotic ($t \gg 1$) regimes.

Whether or not a superballlistic scaling can be possible for stationary velocities is explored in Ch. 7. The only possibility are *multi-modal* VACFs, those with multiple maxima. Although the numerical examples for the study class of *exponentially weighted polynomials* suggested the possibility, the underlying nature of superballistic scaling seems to be counterproductive to the *positive-semidefiniteness* condition of the VACF itself. No definite analytical answer has been made, but the numerical examples together with the PSD property suggest, that a superballistic scaling is implausible for stationary velocities. [2]

The final chapter Ch. 8 concluded the thesis with a final tour de force: Starting with an example for non-stationary VACFs, I investigated the different EAMSD and ATAMSD scaling for the sample class of *polynomially modulated VACFs* $C_v(r,s) = r^b s^b e^{-|r-s|}$. Although a superballistic EAMSD is easily possible, the ATAMSD has again been shown to stay subballistic. I have made the very important observation, that the lagtime scaling of the ATAMSD is identical to the MSD scaling of the integrated equilibriated Ornstein-Uhlenbeck process in Sec. 4.3.2. Both processes share the *stationary* VACF component $e^{-|r-s|}$, and the total ATAMSD for different values of b only differed in their scaling w.r.t. the observational time T. This inspired the *split scaling hypothesis* for stationary modulated VACF $C_v(r,s) = f(r)f(s)C(r-s)$, which explains the scaling of the ATAMSD w.r.t. lagtime Δ and observational time T. The stationary component $C(r-s)$ determined the lagtime scaling and the non-stationary $f(r)f(s)$ the T-scaling. The split scaling hypothesis has been proven analytically, albeit via lengthy arguments. The analysis of the thesis concluded with a glimpse on how to phrase the split scaling hypothesis for general VACFs, why this helps in ex-

---

[2] At least when requiring stable scaling.

plaining the difference in (superballistic) EAMSD and (subballistic) (A)TAMSD scaling throughout the literature examples and how this can shed new light on the scaling of subdiffusive processes.

Own investigations

The main theme of this thesis has been to understand why the lagtime scaling of the ATAMSD in many examples and real-life experiments does not appear superballistic, even though the corresponding EAMSD does. Since superballisticity in itself is a rather niche topic in the research direction of anomalous diffusion, there are few reference papers on this topic.

The arguments on the ATAMSD scaling relied very heavily on deeper results on scaling itself, namely the scaling of higher derivatives and a rigorous definition of scaling exponents. The notion of the dynamical scaling exponent is known in the physical literature, just as the the upper and lower indices from the theory of regular variations (see [19]) are known in the mathematical literature. I have not found a reference that links these two topics in the context of diffusion processes and the analysis in Ch. 3 and Ch. 5 is my own, although parts of the proof ideas can be found in [2] in a different context.

The examples in Ch. 4 are partially known in the literature[3], but most of the detailed analysis of the scaling in terms of the scaling exponent has not been done before for these examples.

Ch. 6 introduced the VACF and discussed its relation with the EAMSD, ATAMSD and the special case of stationary VACFs. The concept of VACFs and the connection between the EAMSD, ATAMSD and VACF with formulas is widely known (see [17] and [16]), but most of the analysis in the literature has been done for the stationary case. The analysis via derivative scaling exponents seems to be not known or used. The same also holds with the analysis of the EAMSD in the general non-stationary case, although the formulas are known in the literature. Regarding the ATAMSD in Ch. 6, I can claim the results as my own. I could not find a thorough reference on the matter of the higher ATAMSD derivatives, the discretization in terms of recursion relations and the analysis of ATAMSD scaling in terms of Δ and $T-\Delta$. A fairly recent survey on this matter can be found in [27], but the analysis is focused on the ALI instead of the VACF. Some of the themes of my ATAMSD analysis are mirrored there, albeit on specific example processes only.

The results on multi-modal VACFs and the concrete examples in Ch. 7 are my own. Most of the arguments and the general concept of the split scaling hypothesis in Ch. 8 are from me, but parts of the mathematical analysis is again borrowed and adapted from [2].

Outlook

Two topics in this thesis have not been adequately answered: The issue of multi-modal VACFs and the ATAMSD scaling for more general VACFs. These represent new fruitful directions of research:

---

[3] like E.g. deterministic processes and the integrated Ornstein-Uhlenbeck processes.

Multi-modal VACFs and random walks in correlated environments

The analysis on multi-modal VACFs in Ch. 7 has been reduced in length due to the total volume of this thesis. I explained on how the concrete examples of multi-modal VACFs can produce intermediary superdiffusive and even superballlistic MSDs, if the PSD condition is relaxed. It is known that every true VACF, i.e. those being PSD, can be realized as a velocity v_t for some trajectory process x_t. The most direct approach to this is the Kolmogorov - Chentsov theorem for Gaussian processes, but the underlying process has no direct physical interpretation. The problem for multi-modality lies in the additional maxima:

If the VACF has a second [4] distinct maximum at x_0, then C has to decrease and increase between $x = 0$ and $x = x_0$. Since the velocity is stationary, the increase in C implies some increase in correlation within the velocity processes. Coming up with a model admitting such a multi-modal VACF is not an easy task, as an intermediary increase in correlation suggest some underlying *memory effect* of the velocity. One possibility would be to adjust the dynamics of the velocity process itself, e.g. by introducing an angular bias. This is a simple way to introduce correlation, but is inherently artificial. Another possibility would be **random walks in correlated environments** - instead of introducing the correlation via altering the dynamics, one can imprint the correlation via the underlying geometry itself: This can be achieved if the underlying geometry is modeled stochastically, such that the sites of the geometry (e.g. lattice sets, surfaces, cell boundaries) are given by a stochastic process. One example of this is the **Voronoi liquid** x_β, a cell growth model (see [25]):

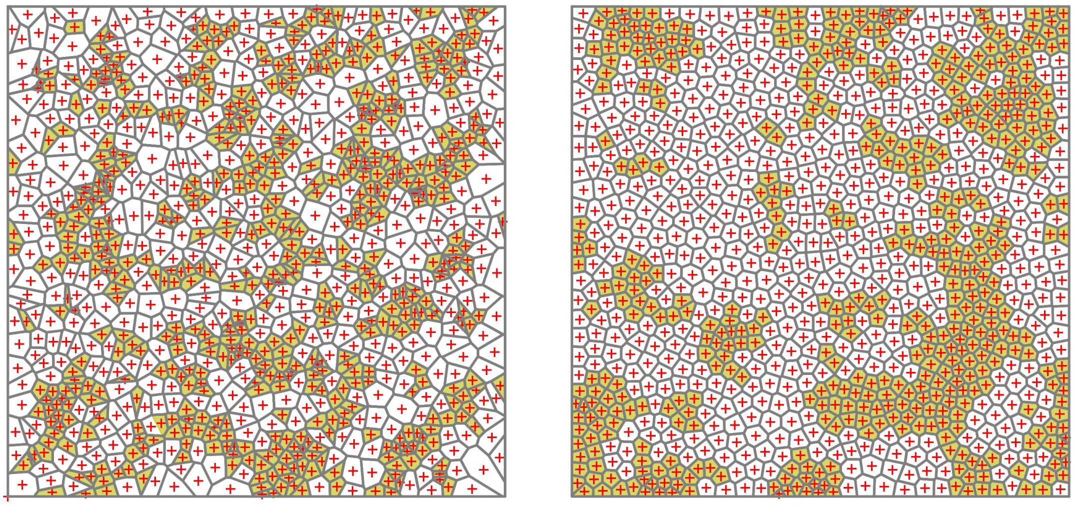

Figure 9.1: Voronoi tessellation of the Voronoi liquid x_β on the unit square with $N = 1000$ points at $T = 100$. The left is given by $\beta = 1$ and the right for $\beta = \infty$.

The process $x_\beta = (x_1, ..., x_N)$ consists of N points in the unit square $D = [0, 1]^2$,

---

[4] $C(0)$ is always the global maximum.

whose joint probability distribution $\rho_\beta(x)$ is given by the Gibbs distribution

$$\rho_\beta(x) = \frac{1}{Z} e^{-\beta V(x)} \tag{9.1}$$

with **Voronoi potential**

$$V(x) = \frac{1}{2} \sum_{i=1}^{N} \int_{Vor(x_i)} \|y - x_i\|^2 \, dy. \tag{9.2}$$

Two sample paths of x_β for $\beta = 1$ and $\beta = \infty$ with $N = 1000$ are shown in Fig. 9.1. The point process x_β has been simulated using a Langevin Monte Carlo method and the faces have been computed using Voronoi tessellations. In both cases, the cells are colored **yellow** if their area is smaller than the mean area $A = \frac{1}{N}$.

The right process is called **centroidal Voronoi tesselation (CVT)**, because each cell germ x_i coincides with the **cell centroid**

$$\mu_i = \frac{\int_{\mathrm{Vor}(x_i)} (y - x_i) \, dy}{|\mathrm{Vor}(x_i)|}.$$

The yellow cells are clustered together, in the sense that they form small, globular domains themselves.

The Voronoi liquid for $\beta = 1$ on the left is less regular and the yellow cells do not cluster into globular domains, but into fractal-like elongated stripes.

Since the distribution of the yellow cells is not independent, but dependent on β, their local distribution is spatially correlated. It is this correlation that can be used to create a multi-modal VACFs.

One simple realization would be to construct a Markov chain on such a grid, but that is forbidden to traverse the yellow cells. In case of the left figure ($\beta = 1$), the tubular wall-like distribution of the yellow cells can lead to channels, in which only a few cells are traversable in each direction. This increases the probability to traverse such a *channel* in a guided direction, leading to locally enhanced diffusion.

Whether or not such a model can generate a multi-modal VACF has to be checked in detail. There are some objections, that spoil this argument. If one averages over all realization of the cell grid, then the spatial correlation may be averaged out and multi-modality is lost. Likewise, a uniformly or invariantly distributed initial position can again average out the multi-modality, if the global structure of the grid is too regular.

Studying these **correlated geometries** and simulating the behavior of random walks across them remains a fruitful, yet open research question.

ATAMSD scaling for generic VACFs

The split scaling hypothesis has been shown for the example class of stationary modulated VACFs $C_v(r, s) = f(r)f(s)C(r - s)$. Despite the technical arguments, the split scaling hypothesis can be easily explained for these VACFs. This is mainly due to the split in stationary ($C(r - s)$ and non-stationary ($f(r)f(s)$) components, allowing to separate both contributions into the lagtime and observational time hypothesis.

For more general VACFs, this conceptual divide is not explicit. Either because the VACF is a linear combination of individual VACFs, or the split in stationary and non-stationary contributions is not possible. Take the following example:

$$C_v(r,s) = \frac{1}{1+\left(\frac{r+s}{2}\right)^2} e^{-|r-s|}$$

The ALI

$$\langle (x_{t+\Delta} - x_t)^2 \rangle = \int_t^{t+\Delta} \int_t^{t+\Delta} \frac{1}{1+\left(\frac{r+s}{2}\right)^2} e^{-|r-s|}\, dr\, ds \tag{9.3}$$

cannot be solved analytically, but it can be approximated

$$\langle (x_{t+\Delta} - x_t)^2 \rangle \sim 2\left(1 - e^{-\Delta}\right)(\arctan t + \Delta - \arctan t)$$

and the ATAMSD approximately solved as

$$\begin{aligned}\langle \delta^2(\Delta, T) \rangle &\sim \frac{2\left(1-e^{-\Delta}\right)}{T-\Delta} \int_0^{T-\Delta} \arctan t + \Delta - \arctan t\, dt \\ &= \frac{2\left(1-e^{-\Delta}\right)}{T-\Delta} \int_{T-\Delta}^{T} \arctan t\, dt - \int_0^{\Delta} \arctan t\, dt\end{aligned}$$

The first part $2\left(1-e^{-\Delta}\right)$ comes from the stationary kernel. The split scaling hypothesis would suggest that the lagtime scaling agrees with the MSD scaling of the integrated equilibrium Ornstein - Uhlenbeck process. But this is not true:

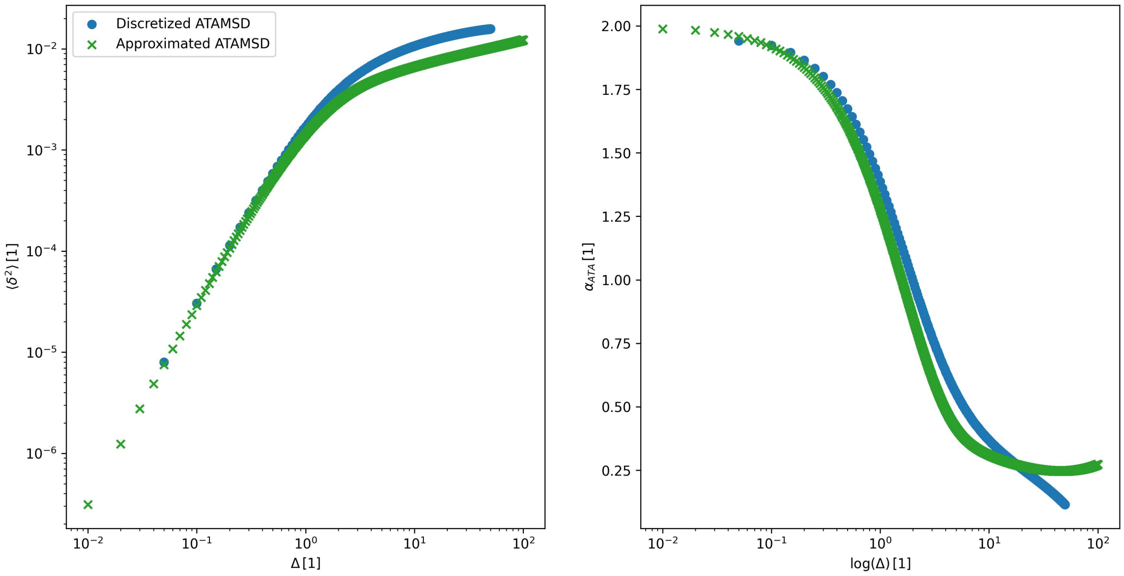

Figure 9.2: ATAMSD and scaling exponent for $C_v(r,s) = \frac{1}{1+\left(\frac{r+s}{2}\right)^2} e^{-|r-s|}$. The discretization used the recursion relation of the VACF, whereas the approximated ATAMSD is a simple numerical integration of the approximated formula.

The arcus tangens terms

$$\int_{T-\Delta}^{T} \arctan t \, dt - \int_{0}^{\Delta} \arctan t \, dt$$

have a **finite** T-limit. This makes the split scaling hypothesis fail, because the dominant term is **constant** in T. Since the constant term can have a non-trivial lagtime scaling, the lagtime scaling hypothesis breaks down. This is supported by the numerical results in Fig. 9.2, where the lagtime scaling clearly deviates from the expected scaling of the IEOUP in Sec. 4.3.2.

The breakdown of the split scaling hypothesis comes from the fact that the VACF is *integrable*. Take the AOD:

$$\begin{aligned} I(\Delta, T) &= \int_{0}^{T-\Delta} C_v(r, r+\Delta) \, dr = C(\Delta) \int_{0}^{T-\Delta} \frac{1}{1+\left(\frac{2r+\Delta}{2}\right)^2} \, dr \\ &= C(\Delta)\left(\arctan\left(T - \frac{\Delta}{2}\right) - \arctan\left(\frac{\Delta}{2}\right)\right) \end{aligned}$$

The split scaling hypothesis would suggest that $C(\Delta)$ determines the lagtime scaling, because the second term looses its explicit lagtime scaling in the large T limit. In the proof of the stationary modulated VACFs, this was true, because the remainder diverged in the large T limit and the leading term could not have a non-trivial lagtime scaling.

But here the remainder **converges**, since $\arctan(x) \to \pi/2$. More explicitly, the AOD has the following limit:

$$\lim_{T\to\infty} I(\Delta, T) = C(\Delta)\left(\frac{\pi}{2} - \arctan \arctan\left(\frac{\Delta}{2}\right)\right) = C(\Delta)D(\Delta) \tag{9.4}$$

which differs from the expected

$$\lim_{T\to\infty} I(\Delta, T) = C(\Delta).$$

$C(\Delta)$ linked the lagtime scaling to the EAMSD scaling by means of the stationary / non-stationary decomposition. This gave the ATAMSD some merit, as it really did reflect some underlying properties of the EAMSD or trajectory process. In this case, there is no direct link to the EAMSD anymore, and it is not clear what property of the process the lagtime scaling describes. But a quantity without a clear meaning or interpretation is pointless.

Closing remark

It remains a fascinating question, what the lagtime scaling of the ATAMSD implies for the EAMSD and trajectory process, not only for generic VACFs, but also for non-differentiable processes. My goal in this thesis was to come a step closer to answering the problem, but despite its length, many details remain vague and incomplete. Despite its simple appearance, one can not easily reason about the scaling of the

(A)TAMSD, although its vast usage in the literature and experimental data analysis falsely suggests otherwise. Now that single particle tracking and other optical methods have pushed the barrier of experimental solution to the cellular level, it is surprising that the (A)TAMSD and its scaling is not better understood. Hopefully, this work can help in improving our collective understanding of this simple appearing, but highly complex quantity.

Bibliography

[1] Eli Barkai, Yuval Garini, and Ralf Metzler. “Strange kinetics of single molecules in living cells”. In: *Physics Today* 65.8 (Aug. 2012), pp. 29–35. DOI: `10.1063/PT.3.1677`.

[2] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1987.

[3] David Böhringer et al. “Dynamic traction force measurements of migrating immune cells in 3D biopolymer matrices”. In: *Nature Physics* 20.11 (2024), pp. 1816–1823. DOI: `10.1038/s41567-024-02632-8`.

[4] Stas Burov et al. “Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking”. In: *Physical Chemistry Chemical Physics* 13.5 (2011), p. 1800. ISSN: 1463-9084. DOI: `10.1039/c0cp01879a`.

[5] Tom Chou and Maria R. D’Orsogna. *First Passage Problems in Biology*. 2014. arXiv: `1408.4518 [cond-mat.stat-mech]`.

[6] Serge Cohen and Jacques Istas. *Fractional Fields and Applications*. Mathématiques et Applications. volume 76. Springer, 2013, p. 270. DOI: `10.1007/978-3-642-36739-7`.

[7] Leticia F. Cugliandolo. *Introduction to Langevin Stochastic Processes*. 2025. arXiv: `2504.11994 [cond-mat.stat-mech]`.

[8] A. Dechant et al. “Scaling Green-Kubo Relation and Application to Three Aging Systems”. In: *Physical Review X* 4.1 (Feb. 2014). ISSN: 2160-3308. DOI: `10.1103/physrevx.4.011022`.

[9] Jürgen Elstrodt. *Maß- und Integrationstheorie*. 6th ed. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. ISBN: 9783540897286.

[10] L. Erdős and H.T. Yau. *A Dynamical Approach to Random Matrix Theory*. Courant Lecture Notes. Courant Institute of Mathematical Sciences, New York University, 2017. ISBN: 9781470436483.

[11] Daan Frenkel and Berend Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. 2nd ed. San Diego: Academic Press, 2002. ISBN: 9780122673511. DOI: `10.1016/B978-0-12-267351-1.X5000-7`.

[12] Denis S. Grebenkov. “Probability distribution of the time-averaged mean-square displacement of a Gaussian process”. In: *Phys. Rev. E* 84 (3 2011), p. 031124. DOI: `10.1103/PhysRevE.84.031124`.

[13] Jean-Pierre Hansen and I. R. McDonald. *Theory of Simple Liquids: With Applications to Soft Matter*. 4th ed. Oxford: Academic Press, 2013. ISBN: 9780123870322. DOI: `10.1016/C2010-0-68308-6`.

[14] Jae-Hyung Jeon, Aleksei V. Chechkin, and Ralf Metzler. "Scaled Brownian motion: a paradoxical process with a time dependent diffusivity for the description of anomalous diffusion". In: *Phys. Chem. Chem. Phys.* 16.30 (2014), 15811–15817. ISSN: 1463-9084. DOI: `10.1039/c4cp02019g`.

[15] Olav Kallenberg. *Foundations of Modern Probability*. 3rd ed. Probability Theory and Stochastic Modelling. Cham: Springer, 2021. ISBN: 978-3-030-61871-1. DOI: `10.1007/978-3-030-61871-1`.

[16] Joseph Klafter and Igor M. Sokolov. *First Steps in Random Walks: From Tools to Applications*. Oxford: Oxford University Press, 2011. ISBN: 9780199234868.

[17] Ralf Metzler and Joseph Klafter. "The random walk's guide to anomalous diffusion: a fractional dynamics approach". In: *Physics Reports* 339.1 (2000), pp. 1–77. ISSN: 0370-1573. DOI: `https://doi.org/10.1016/S0370-1573(00)00070-3`.

[18] Ralf Metzler et al. "Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking". In: *Phys. Chem. Chem. Phys.* 16 (44 2014), pp. 24128–24164. DOI: `10.1039/C4CP03465A`.

[19] Jayakrishnan Nair, Adam Wierman, and Bert Zwart. *The Fundamentals of Heavy Tails: Properties, Emergence, and Estimation*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2022.

[20] Philippe Naveau and Johan Segers. *Multivariate extreme value theory*. 2024. arXiv: `2412.18477 [math.ST]`.

[21] Our World in Data. *Economic Inequality*. Accessed: 2026-05-06. 2024.

[22] Grigorios A. Pavliotis. *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. Vol. 60. Texts in Applied Mathematics. Springer New York, 2014. ISBN: 9781493913220. DOI: `10.1007/978-1-4939-1323-7`.

[23] Michael J. Saxton. "Single-particle tracking: connecting the dots". In: *Nature Methods* 5.8 (2008), pp. 671–672. DOI: `10.1038/nmeth0808-671`.

[24] Johanna Schürlein. "Statistic analysis of natural killer cell's motility in a collagen matrix". MA thesis. Erlangen: Friedrich-Alexander-Universität Erlangen-Nürnberg, Dec. 2021.

[25] Anand Srinivasan, Steph S. M. H. Hohn, and Raymond E. Goldstein. *Stochastic Voronoi Tessellations as Models for Cellular Neighborhoods in Simple Multicellular Organisms*. 2024. arXiv: `2311.11939 [cond-mat.soft]`.

[26] Michael Theves et al. "How the motility pattern of bacteria affects their dispersal and chemotaxis". In: *PLoS ONE* 8.12 (2013), e81936. DOI: `10.1371/journal.pone.0081936`.

[27] Wei Wang et al. *Genuine and spurious (non-)ergodicity in single particle tracking*. 2026. arXiv: `2603.22989 [cond-mat.stat-mech]`.

[28] V. Zaburdaev, S. Denisov, and J. Klafter. “Lévy walks”. In: *Reviews of Modern Physics* 87.2 (June 2015), 483–530. ISSN: 1539-0756. DOI: `10.1103/revmodphys.87.483`.

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbstständig und ohne unerlaubte Hilfe verfasst habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel verwendet und alle wörtlich oder sinngemäß übernommenen Inhalte als solche kenntlich gemacht. Weiterhin versichere ich, dass diese Arbeit weder in gleicher noch in ähnlicher Form bereits Bestandteil eines anderen Prüfungsverfahrens war.

Erlangen, May 10, 2026

Lennard Kossmann