

Field theoretic formulation and empirical tracking of spatial processes

Saoirse Amarteifio

A PhD dissertation

Supervised by Dr. Gunnar Pruessner

Imperial College London
Department of Mathematics
Summer 2019

To my parents

*I remember how
you made me want
to learn*

Declaration

I declare that the work presented is entirely my own unless stated otherwise. Where new contributions are claimed by myself or in collaboration with colleagues, I endeavour to make clear the manner in which the work depends on the collaborations, the work of colleagues or the work of others.

Saoirse Amarteifio

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

To my supervisor, Gunnar, for introducing me to strange new topics in theoretical physics that I would otherwise not have encountered and for creating an environment of such openness and enthusiasm in which to do research. I put the greatest value in your advice and friendship. And to everyone in #GG, thank you guys for contributing to this wonderful environment too; particularly you non-equilibrium folks; Rosalba, Ben, Ignacio, Johannes, Nanxin - and to Brandon for taking great notes on your iPad in the early days of our reading group. I am deeply grateful to Ignacio for kindly offering to read a draft of the thesis, for comments and corrections and for offering a few words of encouragement when I needed them most. To my collaborators in biology and biophysics, particularly Giovanni, Todd, Delfin and Cristina, thank you for adding this other dimension to my research at Imperial. My thanks to Professor Henrik Jensen, who, in addition to first setting me on this path, is one of the nicest people I have had the pleasure to meet during my time in academia.

There are many people to whom I am grateful for discussions over the years, either in person or over email. Even brief conversations can have a profound impact on the direction one's research takes. I would like to thank Antal Jarai for a very useful discussion at Bath University in the early stages of the Abelian sandpile model project. I am grateful to certain researchers who have been so generous with their time after my emailing them out of the blue. Nicolas Behr has improved my understanding of combinatorics in chemical reaction networks and has been so incredibly responsive on the occasions that I enquired about aspects of his work. I am also grateful to Erik Panzer for discussing his PhD research with me and how ideas might transfer to non-equilibrium physics. Erik, who is surely one of a kind, has genuinely astounded me with the extent of his knowledge and his generosity in responding so insightfully to my silly questions.

My appreciation goes to the open-source and online community for sharing code and ideas and in particular, Manuel Kuehner for creating and sharing his wonderful LaTeX template!

In the last few months I have had the pleasure of spending some time visiting Humboldt Universität. I am truly grateful to Marcel Golz for helping me organise this visit and to Professor Dirk Kreimer for kindly inviting me. While visiting, I have had the privilege to discuss graph polynomials with Marcel and gain a deeper understanding of Hopf algebras and renormalization from Professor Kreimer. My thanks to everyone at Adlershof for your hospitality and to Nicolas for showing me the best of Berlin's bouldering halls!

Finally I would like to thank my family for moral support over the years, often bolstered by the sweetest lies. You lighten my heart. And thank you Chau, you have always done so incredibly much for me. My gratitude is beyond words.

A special thanks

To my external examiners
Guillaume Salbreux and Robert Endres
for their patient reading of, and helpful comments on, the manuscript.

Abstract

Spatial processes are attacked on two fronts. On the one hand, tools from theoretical and statistical physics can be used to understand behaviour in complex, spatially-extended multi-body systems. On the other hand, computer vision and statistical analysis can be used to study 4D microscopy data to observe and understand real spatial processes *in vivo*.

On the first of these fronts, analytical models are developed for abstract processes, which can be simulated on graphs and lattices before considering real-world applications in fields such as biology, epidemiology or ecology. In the *field theoretic formulation* of spatial processes, techniques originating in quantum field theory such as *canonical quantisation* and the *renormalization group* are applied to reaction-diffusion processes by analogy. These techniques are combined in the study of *critical phenomena* or *critical dynamics*. At this level, one is often interested in the scaling behaviour; how the correlation functions *scale* for different dimensions in geometric space. This can lead to a better understanding of how macroscopic patterns relate to microscopic interactions. In this vein, the *trace* of a *branching random walk* on various graphs is studied. In the thesis, a distinctly abstract approach is emphasised in order to support an *algorithmic* approach to parts of the formalism.

A model of *self-organised criticality*, the *Abelian sandpile model*, is also considered. By exploiting a bijection between *recurrent* configurations and spanning trees, an efficient Monte Carlo algorithm is developed to simulate sandpile processes on large lattices.

On the second front, two case studies are considered; migratory patterns of leukaemia cells and mitotic events in *Arabidopsis* roots. In the first case, tools from statistical physics are used to study the spatial dynamics of different leukaemia cell lineages before and after a treatment. One key result is that we can discriminate between migratory patterns in response to treatment, classifying cell motility in terms of sup/super/diffusive regimes. For the second case study, a novel algorithm is developed to processes a 4D light-sheet microscopy dataset. The combination of *transient* fluorescent markers and a poorly localised specimen in the *field of view* leads to a challenging tracking problem. A fuzzy registration-tracking algorithm is developed to track mitotic events so as to understand their spatiotemporal dynamics under normal conditions and after tissue damage.

Contents

Acknowledgements	7
Abstract	11
List of Figures	17
List of Tables	23
Introduction	25
1 Field theory of spatial Markov processes	31
1.1 Overview	31
1.2 Reaction-diffusion systems	32
1.2.1 Some history	33
1.2.2 Formulation	35
1.3 Heisenberg-Weyl semigroups	36
1.3.1 Markov semigroups	36
1.3.2 Heisenberg-Weyl algebra	39
1.3.3 Generating functions	40
1.4 Diffusion, propagators and responses	43
1.4.1 Introducing the propagator	44
1.5 Perturbative expansions and path integrals	45
1.5.1 Some history	45
1.5.2 Some key ideas	48
2 Graphs	55
2.1 Terminology	55
2.2 Representations	57
2.3 Spanning trees and graph polynomials	59
2.4 Abelian sandpiles: a useful bijection	62
2.4.1 Recurrent configurations to spanning trees	63
2.4.2 Spanning trees to recurrent configurations	64
2.5 The Feynman graph	65
2.5.1 Graph representations	65
2.5.2 Integral representations	72
2.5.3 Graph operations	79
2.5.3.1 A shuffle product	79
2.5.3.2 Contours, cover-ups and contractions	82
2.5.3.3 Parametric integration	86
2.6 Renormalization	87

3	Analysis of Scaling	93
3.1	Scaling symmetry and the renormalization group	94
3.2	Numerical studies	97
3.2.1	The Abelian sandpile	97
3.2.1.1	Algorithm validation results	100
3.2.2	The trace of a superprocess	101
3.2.2.1	What is a superprocess?	102
3.2.2.2	The relevant action	103
3.2.2.3	Minimal subtraction and RG functions	106
3.2.2.4	Scaling for superprocess on graphs	107
3.2.2.5	Numerics and graph generation algorithms	108
3.3	Results	112
4	Empirical studies with biological data	115
4.1	Overview of image processing	115
4.2	Lightsheet microscopy data	118
4.3	Tracking mitotic events	122
4.3.1	Tracking algorithm	123
4.3.2	Results	127
4.3.3	Testing against surrogate data	129
4.3.4	Comparing with manual tracking	130
4.4	Characterising cancer cell motility	134
4.4.1	Displacements	134
4.4.2	Distinct sites visited	136
4.4.3	Anomalous diffusion	137
4.4.4	Results	139
4.5	Summary	143
5	Discussion	145
5.1	Algebraic aspects of the Doi-Peliti/pQFT formalism	146
5.2	Parametric Feynman rules for out-of-equilibrium graphs	146
5.3	Efficient sandpile sampling	147
5.4	Scaling of superprocess trace on graphs	147
5.5	Fuzzy tracking of root cells	148
5.6	Quantifying cancer cell motility	148
	Bibliography	153
	Notes	168
A	Algorithms	173
B	Useful Results	175
B.1	The Gaussian integral	175
B.2	Gamma Functions	175
B.3	Combinatorial	176
C	Derivations, proofs and further examples	177
C.1	Homogeneous Poisson process	177
C.2	Lattice Laplacian operator	178
C.3	A note on random walks	179

C.4	Overview of diffusion	179
C.5	Hausdorff and spectral dimension	181
C.6	Chemical reaction networks out of equilibrium	181
C.7	All matrix tree theorem (proof)	183
C.8	Schwinger representation	184
C.9	Some graph matrices	185
C.10	On the Umbral calculus	186
C.11	Factorial moment generating functions	187
C.12	The Lévy α -stable spectrum	188
D	Diagrams for BWS	191

List of Figures

2.1	<i>Sample graphs and notation:</i> (a) An undirected graph G_a with $ \chi = 1, E = 5, V = 3, L = 2$. $\{e_0, e_1\}$ is an independent circuit and $\{e_3\}$ is a self-loop. v_1 has degree 3 and is incident to $\{e_0, e_1, e_2\}$. (b) A graph G_b with $ \chi = 2$ formed from the operation $G_a \setminus \{e_2\}$ (and a minor edge relabelling). Of the set of component graphs $\{g_1, g_2\} \in G_b$, $g_2 \sim g_1 // \{e_1\}$. (c) G_c is an arborescence, a directed acyclic graph rooted at v_0 . The vertex v_1 has in-degree 2 and out-degree 1. $\{v_2, v_3\}$ is the set of leaves in G_c . The graph G_a contains a bridge e_2 whose removal would increase $ \chi $ (the number of components) and is therefore not <i>one particle irreducible</i> (1PI).	57
2.2	<i>Wilson's algorithm</i> performs a random walk on the graph from each site, recording the most recent exit edge at each vertex (a). The numbers on each vertex correspond to edge directions. Loops are erased as the random walk re-exits sites (c). Random walks terminate when they hit the tree, which is initially a single chosen root vertex. A spanning tree is iteratively generated in this way (d).	61
2.3	<i>ASM dynamics:</i> Grey-scale cells denote the height in the sandpile. The darkest cell i.e. black has the highest value $q = 4$. If a new grain of sand is deposited on one of the black cells, there will be a toppling event as particles are redistributed to nearest neighbours. If the nearest neighbour happens to be black, there will be a second toppling event and so on...	62
2.4	<i>The burning algorithm on a lattice</i> is shown with the lattice boundary connected to v_∞ . Site heights $z(i)$ appear in larger font. The superscripts show the "burn time" based on the burn condition. Bootstrapping at time 0, v_∞ is burnt. Continuing in iterations, to burn a site, $z(i)$ must exceed the number of un-burnt neighbours $\bar{B}(i)$. As such, the top left non v_∞ site can be burnt at time 1 while the site to its right cannot be burnt until time 2.	64
2.5	<i>Half-edge pairing:</i> Pairing rules follow from the definition of graph isometry. In 1-vertex and 3-vertex graphs shown, how many ways can black half-edges on a be paired with black half-edges on d ? It would be wrong to enumerate all ways of pairing two edges on a with three edges on d . In fact there are only three ways; i. <i>both</i> edges on a are paired with b or ii. <i>only one</i> is iii. <i>none</i> are. Which edges are paired, on d say, does not matter as the choices lead to isometric (coloured) adjacency structures. However the graph produced by pairing $\{(a_1 \rightarrow c_1), (a_2 \rightarrow d_1)\}$ is not isomorphic to the graph produced by pairing $\{(a_1 \rightarrow b_1), (a_2 \rightarrow d_1)\}$. In total, black pairings include; no pairings, all ways of one of each on a going to only one of each of the others or both edges on a pairing with both edges on either c or d . Yellow edges can be paired in two ways in total; either no edge pairing occurs or a pairs with b	81

3.1	An annulus $[a, \sigma a]$ defines a scale range between a short-distance cutoff a e.g. lattice spacing and some scaling of it $\Lambda = \sigma \times a$. In order to remove short-range interaction effects one integrates out the degrees of freedom on the annulus and then rescales the system by a factor of inverse σ which creates an <i>effective</i> theory with a cutoff Λ . Short-distance behaviour describes a regime shorter than the correlation length but larger than the cut-off scale and this regime is governed by ultraviolet divergences. At the critical point, the short-distance behaviour extends to infinity and ultraviolet divergences are thus relevant to understand long-distance phenomena in many-body systems near the critical point	95
3.2	<i>Critical regimes for observables in branching processes:</i> Shown are profiles of the <i>logarithm of the ensemble average</i> of the number of particles observed as a function of <i>logarithmic time</i> for different parameter regimes in a branching process (or simple birth and death process) initialised with a single particle. The time-dependant solutions for the probability or moments of particle numbers appear with a factor $e^{(\beta-\varepsilon)t}$. For the critical regime with $\beta = \varepsilon$, any exponential factor drops out leaving a power law dependence on time. What this picture does not show is the divergence in the variance of the observable at the critical point. In general, dynamics of critical processes are strongly influenced by such fluctuations.	102
3.3	<i>Corollas (amputated vertices) of the Branching Wiener Sausage:</i> The propagators correspond to the free part of the Lagrangian and the interaction terms are used in a perturbative expansion.	105
3.4	<i>superprocess on Sierpinski carpet:</i> A realization of the superprocess on a Sierpinski carpet generated with parameters $b = 3, m = 8$. Visited sites are marked black.	109
3.5	<i>Some simple random trees:</i> Random trees are grown by a critical branching process. One way to think of the random tree is to start with a perfectly balanced binary tree with $2N + 1$ vertices. Visit each vertex and flip a fair coin; if the coin lands on heads, remove the sub-tree rooted at that vertex.	110
4.1	<i>Convolution</i> and some example kernels	117
4.2	<i>Variation in frame properties:</i> The tiles top(a) show four sample frames. Frame A is a "good" image. Frame B is a noisy image. Frame C is a noisy image with a low number of objects and some noise. Frame D is over-saturated and will be marked as degenerate. On the bottom(b), the same story is told quantitatively over 100 consecutive sample frames from one imaging session. The <i>signal</i> and <i>noise</i> levels are plotted with threshold values. The <i>narrow-band</i> plot shows the region between the 95 th and 99 th percentile ranges of the image histogram. Notably, these values fluctuate dramatically.	120
4.3	<i>Root Thresholding:</i> A rank-based adaptive threshold is used in conjunction with a Laplacian of a Gaussian to enhance large structures. Data intersecting the 95 th and 99 th percentile ranges are selected by band-pass filter.	121
4.4	<i>Isolating the Region of Interest (ROI):</i> A narrow-band filter (B) and aggressive Gaussian smoothing (C) is used to find the largest connected component in the narrow band of the data corresponding to the region of activity in the root tip (D).	121

4.5	<i>Root masking</i> : A Gaußian filter is first used to ensure the integrity of the components (first image). A binary mask is constructed (middle image). The labelled giant component is used to construct a 3D mask or <i>Region of Interest</i> , which is used to isolate the root pixels in the original 3D data for further processing.	121
4.6	<i>Blob detection</i> : Illustration of blob detection using a <i>Difference of Gaußians</i> . The 3D projection emphasises how thresholding might then be used to keep peaks in the image.	122
4.7	<i>Translation sampling with $k = 1$</i> : Panel (a) shows an arbitrary translation between a sample point in V and a sample point in U . Panel (b) shows the displacement of many such translations from the origin and includes a "consensus" translation i.e. the average of the top ranked translations. Panel (c) illustrates that in this case, the consensus translation is better than the one in panel (a) making better matches between circles and green dots. In this case, while not discernible from the image, the consensus translation turns out to be better than any of the other sampled translations.	124
4.8	<i>Constellation sampling</i> : Random permutations $\mathcal{P}(M, k)$ with $k = 3$ produce triangular constellations in each frame. Affine transformations dimensions mapping point triplets in one frame to point triplets in another frame are added to proposal set Γ	125
4.9	<i>Life matrix</i> : Rows (y-axis) correspond to detected individuals with time on the x-axis. The life matrix shows a generally consistent trend in the appearance and persistence of objects. At the beginning of the sequence, objects persist for an abnormally long time for biological/experimental reasons. . .	127
4.10	<i>Time series</i> : The population time series reflect the information shown in the life matrix. The high spike in the mean age is a result of one object surviving for a long time while others disappear. After this point more typical fluctuations in object activity are observed.	128
4.11	<i>Time-lapse plots</i> : Two different colourings are used to distinguish time(a) or differences between the object displacement vector and the global transformation vector(b). The ellipse represents the (Gaußian) statistical region of interest over many frames.	128
4.12	<i>Dealing with low information when ranking transformations</i> : The image sequence (left to right) demonstrates how identifiers can be incorrectly assigned when there are too few objects with white dots (at t) and red dots (at $t - 1$). The <i>identity transformation would</i> make the correct assignment in this example.	131
4.13	<i>Dis/agreement between lags</i> : Lag-1 identifiers for objects of interest (excluding single-frame detections) are compared with identifiers assigned by $\lambda_\tau : \tau \in \{2, 3\}$ transformations for objects that are older than τ . 734 objects appeared in the sample sequence. 42 $\approx 6\%$ showed lag1-2 disagreements and excluding these, 29 $\approx 4\%$ showed lag1-3 disagreements in specific frames. .	132

4.14	<i>Choosing an ε value:</i> The region 20 ± 5 , which is shaded in the plot, highlights the region where the reduction in single-frame detections begins to decrease before levelling out. This region can also be interpreted as the range of observed object radii. There is a trade-off in choosing an ε value. A value of $\varepsilon \ll 20$ results in excessive single frame detections as the tracker cannot accommodate small fluctuations in cell movement because the catchment region is too small, while values for $\varepsilon > 25$ merge lineages erroneously because the catchment region is too large.	132
4.15	<i>Anisotropic matching tolerance:</i> Differences in angles between the global transformation and the displacements between married objects are compared. White dots and red dots correspond to objects at time t and $t - 1$ respectively. Most red dots are to the left of white dots which corresponds to a global transformation vector of about 170° from the positive x-axis vector. Object 44 has a very different angle to the rest and in the image, the red dot appears to the right. Large displacements in conjunction with large angle differences for a number of consecutive frames suggests debris.	133
4.16	<i>Measuring distinct sites visited at different scales:</i> The illustration shows sites for $\sigma = 100$ and $\sigma = 50$. A volumetric mesh is constructed for each scale σ . The same track (tracked data points marked with dots) is shown passing through sites at two different scales. Four data points for a sample track are shown falling within sites for each scale. At scale $\sigma = 100$, 4 data points <i>fall in</i> 3 sites. At scale $\sigma = 50$, the same 4 data points fall in 1 site.	136
4.17	<i>Comparing linear and anomalous diffusion:</i> a Brownian motion on the left and a "typical" cell track on the right. Distribution of <i>jumps, angles, speed</i> etc. can be measured to understand how the particle explores space. The track on the right could be referred to as <i>Lévy-like</i> in that, qualitatively, it consists of more occurrences of large jumps than might be expected from Brownian motion.	137
4.18	<i>The displacement densities (normalised histograms) for each of the conditions:</i> These compare before (blue/lighter) and after (green/darker) treatment. We are interested primarily in detecting differences as a result of treatment. The Kullback-Liebler Divergence (KLD) is a measure of this difference (§4.4.4 for details). The impact of treatment is greatest for the T-ALL types as the KLD values are higher. For T-ALL, the KLD values are close to 0.3, for AML they are close to 0.1. The differences seem to correspond to results of other tests but are not striking <i>per se</i>	140
4.19	<i>Box plots of position-averaged displacements:</i> This shows a coarse-grained perspective of the displacements. These were computed by first taking an average within each mouse position and then showing the mean and standard deviation over the set of positions <i>per</i> condition. Consistent with the changes within displacement densities, we see most notable differences in T-ALL. p-values from ANOVA were (T-ALL no chemo, 0.004), (T-ALL post chemo, 0.001), (T-ALL post chemo, 0.004), (AML no chemo, 0.359), (AML post chemo, 0.980), confirming significance (95% confidence) for T-ALL but not for AML.	141

4.20	<i>Distinct sites visited as a function of the inverse scale parameter σ</i> : From left to right, fine-grained to coarse-grained, we see the expected decline in slope at coarse scales. Both the comparative rates of decline and how the distinct site visited <i>levels</i> change before and after treatment is consistent with other results, where we see the most dramatic changes in T-ALL after treatment.	141
4.21	The box plot above compares distinct sites visited at a fine-grained scale ($\sigma = 100$). The differences correspond to the gaps between the (start of) lines in the scaling behaviour shown in the previous figure.	142
4.22	<i>The log-log plot of time averaged mean squared displacement as a function of the lag (x axis) divided by the lag for each condition</i> . These trends show power law fits of the data. Each condition e.g. (T-ALL no chemo) takes a fixed colour, while PRE/POST pairs can be distinguished by line type; dashed for PRE and solid for POST. We can see the T-ALL <i>post chemo</i> POST stands out, with a slope that does not decline. T-ALL <i>post chemo</i> POST appears <i>superdiffusive</i> as opposed to <i>subdiffusive</i> (all other cases). . .	142

List of Tables

0.1	List of articles	26
3.1	<i>ASM avalanche area scaling:</i> For each moment, the mean and variance are shown for the scaling in the area (or trace) of avalanches.	99
3.2	<i>ASM avalanche duration:</i> For each moment, the mean and variance are shown for the scaling in the duration of avalanches.	99
3.3	<i>ASM avalanche size:</i> For each moment, the mean and variance are shown for the scaling in the size of avalanches. Size is defined as the number of sand grains involved in the event.	99
3.4	<i>Expected ASM heights:</i> Expected (asymptotic) site height distribution on recurrent ASM configuration of size $L = 10^6$	100
3.5	<i>Checking site heights against expected values:</i> Mean and standard deviation for 10^6 samples of recurrent configurations on a lattice of size $L = 10^3$ matching theoretically expected asymptomatic figures.	100
3.6	<i>Symmetry test without bias:</i> An <i>insignificant</i> differences between height counts top left and top right supports <i>uniform</i> sampling. For example observe $ 144228.9 - 144295.1 /144262.0 \approx 5 \times 10^{-4}$. Alternatively, the normalized deviation from the normalized mean (of the left and right values) should approach 0.5 and in this case it is 0.4998.	101
3.7	<i>Symmetry test with bias:</i> Significant differences between height counts, top left and top right, indicate a bias in configuration sampling. In this case, the difference is of the order 5×10^{-1} , which is significant. Alternatively, the normalized deviation from the normalized mean (of the left and right values) should approach 0.5 but in this case it is 0.2988.	101
3.8	<i>Scaling of visited sites in time by BRW on hypercubic lattice:</i> Scaling shown for lattices in dimension above and below the critical dimension $d_c = 4$. Comparisons are made between theory and numerical simulation.	112
3.9	<i>Scaling of visited sites in time by BRW on non-regular graphs:</i> Scaling is shown for the Sierpinski carpet, random tree and preferential attachment scale-free graph. The spectral dimensions for these graphs are, for Sierpinski carpet $d_s \approx 1.86$, for the random tree $d_s \approx 4/3$ and for the scale-free graph generated by the preferential attachment algorithm $d_s \approx 4$. Comparisons are made between theory and numerical simulation.	112
3.10	<i>Scaling of visited sites by a BRW as function of the system size on hypercubic lattices:</i> Scaling shown for lattices in dimension above and below the critical dimension $d_c = 4$. Comparisons are made between theory and numerical simulation.	113

3.11 *Scaling of visited sites by a BRW as function of the system size on non-regular graphs:* Scaling is shown for the Sierpinski carpet, random tree and preferential attachment scale-free graph. The spectral dimensions for these graphs are, for Sierpinski carpet $d_s \approx 1.86$, for the random tree $d_s \approx 4/3$ and for the scale-free graph generated by the preferential attachment algorithm $d_s \approx 4$. Comparisons are made between theory and numerical simulation. . 113

4.1 *Surrogate data parameters:* Surrogate data are generated respecting a minimum object separation ε . A random Euclidean transformation determined by parameters (r, s, u) is applied to a set of points U to produce a new set of points V . To simulate debris or particle births and deaths, p points are added at random locations to both U and V . Added points also respect the minimum object separation ε 129

4.2 *Surrogate data test results (loss):* With $n = 10$, $p = 3$, $u = 1.0$, results for average of 10 samples for each point in the (r, s) parameter space. These results illustrate that for small rotations (approx 5 degrees) or for small translations the tracker is robust and degrades gracefully for larger rotations (approx 15 degrees). 130

Introduction

An open challenge of our time is to understand the orchestration of complex spatially-extended processes as found in nature. We can see, right in front of us so to speak, complex and fascinating *macroscopic* patterns. Exploiting, for example, advances in experimental techniques such as time-lapse microscopy, we can increasingly see, right in front of us, *microscopic* behaviour. Yet it remains difficult to understand precisely how each of these scales influences the other. Of course *scale* itself is a relative concept and my choice of *macroscopic* and *microscopic* scales here is somewhat arbitrary. This in itself is a challenging notion. With conceptual challenges of this nature in mind, I explore some relevant methodologies;

1. Complex systems have long been modelled as stochastic process starting with zero-dimensional or mean field models. More recently, the geometry of space and also the role of fluctuations are included in more sophisticated models, which can be cross-validated with computer simulation. One approach is the *coherent state path integral formalism*. I describe this formalism and its relevance to spatial Markov processes in Chapter 1.
2. Graphs and other combinatorial structures surfaced numerous times in my research, often in surprising ways. I have collected these methods in their own chapter. Some novel graph-theoretic approaches to existing methods are discussed in Chapter 2.
3. Analysis of scaling behaviour employs universality and the reduction of relevant degrees of freedom in complex processes. This is a crucial perspective in understanding how macroscopic dynamics and patterns both arise from and influence microscopic dynamics and interactions. In Chapter 3 I discuss models and simulations exploring scaling behaviour in both self-organised critical systems and spatial Markov processes on geometric structures.
4. In Chapter 4 focus shifts to empirical methods. 4D microscopy can be used to monitor biological spatial processes *in vivo*. As the space-time resolution of machine vision datasets increases, large point cloud datasets can provide rich insight on spatiotemporal dynamics. In future this may offer a second method, beyond simulation, for the cross-validation of analytical models of multi-body systems. Today our efforts are more modest and we develop image processing and tracking methods for small point clouds. I consider the nature of conclusions which can be reached regarding the behaviour of cells using these methods. Case studies in cancer treatment and plant root regeneration are presented.

The four main chapters will be largely self-contained and independent. Due to the manner in which the work crosses disciplines it is infeasible to present it in a single linear structure. I try to use Appendix C to collect results and derivations, which may be standard to one

reader but unfamiliar to another, or in cases where the result plays a central role. I use *endnotes* (numeric subscripts) to keep track of tributaries and trivia which are not essential to follow the main thread. For *footnotes*, symbolic subscripts are used. I typically use *Examples* to indicate specific cases isolated from general cases and I use *Remarks* such as the following one to make auxiliary comments.

Remark 1. On the use of "we" and "I": I will use "I". When I use "we" I almost always mean "the reader and I". When referring specifically to collaborations I will use "we" to mean "my collaborators and I", which should be clear in context.

As elaborated upon below, the following will be addressed: (i) To better understand why and when the *field theoretic toolkit* applies to evolutionary Markov processes. (ii) To better understand scaling behaviour of Markov processes on general graphs. (iii) To evaluate a technique to simulate *sandpile models* on very large lattices so as to acquire better statistics about system scaling. (iv) To incorporate graph theoretic ideas into the evaluation of *Feynman integrals* applied to evolutionary Markov processes. (v) To explore the interface between theory and empirical data using computer vision and tracking methods for analysing spatiotemporal processes.

The emphasis in each chapter is quite different. In Chapter 1, I focus on restricting the class of stochastic processes to Markov processes, which have very nice properties and as such allow for a very direct approach to writing down the *path integral*. While in Chapter 3 I discuss the main results related to studies carried out in analysing scaling behaviour in both sandpile models and Markov processes, in Chapter 2 I focus on algebraic and combinatoric elements that are useful for abstracting and performing calculations. As such, Chapter 3 emphasise the physics and the results, while Chapter 2 emphasises mathematical abstractions. Chapter 4 is concerned with data analysis and computer vision methods. The thesis discusses work that appears in the following articles

Article	Status
Delfim Duarte, Saoirse Amarteifio, Heather Ang, Isabella Y Kong, Nicola Ruivo, Gunnar Pruessner, Edwin D Hawkins, and Cristina Lo Celso. <i>Defining the in vivo characteristics of acute myeloid leukemia cells behavior by intravital imaging</i> . Immunology and Cell Biology, 2018.	<i>published</i>
Saoirse Amarteifio, Todd Fallesen, Giovanni Sena, and Gunnar Pruessner. <i>Tracking transient fluorescent events in structured point clouds</i> . bioRxiv, page 419572, 2018.	<i>submitted</i>
Ignacio Bordeu, Saoirse Amarteifio, Rosalba Garcia-Millan, Benjamin Walter, Nanxin Wei, and Gunnar Pruessner. <i>Volume explored by a branching random walk on general graphs</i> , 2019	<i>submitted</i>
Saoirse Amarteifio and Gunnar Pruessner. <i>Exploiting Dhar’s spanning tree bijection and Wilson’s algorithm to compute abelian sandpile statistics efficiently</i> , 2019	<i>in preparation</i>

Table 0.1: List of articles

Motivating problems

The thesis crosses disciplines and methodologies. A number of different problems are considered over four chapters. Chapters 1 and 2 lead up to a discussion of the numerics supporting a renormalization group analysis of a spatial Markov process in Chapter 3. As explained in the following sections, the core problem is to apply field theoretic ideas to spatial Markov processes on arbitrary graphs. A challenge this thesis undertakes is how to abstract this procedure. In Chapter 1, field theory and chemical reaction networks on graphs i.e. spatial Markov processes are introduced. Some of the graph theoretic tools required to carry out the field theoretic procedure algebraically and symbolically are presented in Chapter 2. Chapter 3 then discusses the final analysis of these processes in the special case of a branching random walk, focusing on the Monte Carlo simulations used to validate the theory. Chapters 2 and 3 also considers the problem of efficient Monte Carlo simulation of the Abelian sand pile model and develops a new algorithm; the relevant graph theory and the Monte Carlo simulation find a home in Chapters 2 and 3 respectively. Chapter 4 changes perspective completely and deals with data analysis tools applied to real biological processes.

Algebraic aspects of the coherent state path integral formalism

I try to emphasise abstractions of the field theoretic treatment of spatial Markov processes. This culminates in an *algorithmic* treatment of the procedure, which has been partially prototyped in `Python` code. The field theoretic procedure can be cumbersome due to the possible combinatorial explosion of graphs (Feynman integrals) that can be generated for reaction-diffusion processes. Partly to make concrete some of the steps in the field theory procedure and partly to suggest a tool that could, for example, automatically generate graphs for arbitrary chemical reaction networks, the style of exposition in Chapter 1 and 2 leans towards the abstract. This is not standard in the field theory of critical phenomena literature. I briefly explore adopting methods such as parametric integral representations, which are used in field theories in high-energy physics. I recently visited Humboldt Universität to learn more about these methods from researchers who are developing them. In particular, some of the details outlined in Chapter 2 follow from recent discussions with Marcel Golz, who has developed parametric integral representations with a focus on Quantum Electrodynamics in his PhD research.

Efficient Monte Carlo sampling in sandpiles

The *Abelian sandpile model* (ASM) is a model of *self-organised criticality* [Dhar, 1990]. Self-organised criticality, as studied by statistical physicists, is thought to capture aspects of the dynamics of real-world (complex) systems such as neuronal firing patterns or earthquakes, which can be conceptualised as "avalanches". The sandpile model is typically studied on planar lattices for example an $L \times L$ regular lattice. Briefly, ASM dynamics involves depositing sandgrains on lattices according to a certain update rule, which induces avalanche events. The dynamics are inherently *deterministic*. By depositing sandgrains at random sites, randomness is introduced for Markov Chain Monte Carlo sampling and estimation of avalanche statistics. The configurations into which the ASM equilibrates

between avalanches may be *transient* or *recurrent*. The latter are the ones which produce useful statistics. In order to understand the scaling behaviour of this particular model of self-organised criticality, increasingly large lattices must be used in order to overcome *finite size effects*. However, it turns out, it takes a long time to first reach the recurrent epoch where meaningful statistics emerge.

There exists a bijection between *recurrent configurations* of sandpiles and *spanning trees*. Given any configuration, a so-called *burning test* can be applied. If having carried out the test a spanning tree is produced, then the configuration is *recurrent*. This suggests a more efficient way to sample ASM statistics. A spanning tree may be generated within the *cover time* of the lattice using an algorithm based on *loop erased random walks*. To efficiently compute ASM statistics for large lattices, the ASM system may be initialised with recurrent configurations mapped from uniform spanning trees, exploiting the mentioned bijection.

In Chapters 2 and 3, a novel algorithm that exploits this bijection is developed and evaluated. After collecting data from many [thousands] of hours of CPU time we are finalizing a manuscript that describes this work for lattices with up to $(32 \times 10^3)^2$ sites [Amarteifio and Pruessner, 2019].

Superprocess range on arbitrary graphs

Field-theoretic methods are applied to the study of reaction-diffusion processes. The objective in this case is to both develop and understand the theoretical framework i.e. the field theoretic method (by which I mean the coherent state path integral formalism and renormalization group) applied to critical reaction-diffusion processes and also to consider its role in tangible applications.

The particular reaction-diffusion process considered is a well-studied *superprocess*. A superprocess combines two sources of randomness; on the one hand particles (mass) are produced according to a *branching mechanism*. This in itself is independent of the geometry of space. A second source of randomness is added where particles are transported by some stochastic jump distribution. Brownian motion is a special case jump distribution. Such a (super)process is called a *super-Brownian motion* or in discrete space, a *branching random walk*. In applying the field theoretic method, studying the critical superprocess may not be, in and of itself, particularly interesting because as particles are produced by a branching mechanism that is independent of space, its critical features do not depend on the dimensionality (d) of space, which is where the field theory thrives. However, by introducing a second type of particle that tracks the range or trace (i.e. distinct sites visited) of the superprocess, an interesting d -dependant observable emerges. When emphasising the *volume mapped out* by the superprocess, the process is referred to as the *branching Wiener sausage*.

The first contribution of our work (in preparation [Bordeu et al., 2019]) is the field theoretic analysis of this process. Furthermore, by exploiting results about the *spectra* of general graphs, we can easily extend the results to derive the scaling behaviour of the time-dependant trace of the superprocess on arbitrary graphs. This is of practical interest, for example in understanding the spread of disease on social networks.

Problems in bioinformatics

The following two motivating problems, which are covered in Chapter 4, have their origin in biology. The biology problems beget data analysis problems, which will be the focus here. Microscopy forms a core part of the biologist's toolbox and today labs are producing enormous amounts of microscopy data. Microscopy imagery in itself can allow a human to observe *in vivo* processes but (automated) computational methods are required to extract meaningful *knowledge* from the raw images, particularly as the volume of such data grows. Commercial and open source software is available for many tasks. In some cases bespoke tools need to be developed to deal with image processing, computer vision and particle tracking problems in biology. Let me refer to these as challenges of the first kind. Having extracted knowledge such as, for example, cell morphology and trajectories by such means, these data then need to be analysed. Basic analysis includes for example calculating average cell speeds and displacements. However more sophisticated methods from fields such as stochastic process analysis, statistical physics and machine learning can provide an enriched interpretation of the data. Let me refer to these as problems of the second kind. Some of these techniques may be outside a biologist's core skill-set, leading to collaborations with computer scientists, physicists and mathematicians. The following two pieces of work have come about in this way.

Tracking mitotic events in a noisy microscopy data *The biology problem:* A single *Arabidopsis* primary root is grown and imaged on a custom-made light-sheet microscope setup. The root is hydroponically grown in a *perfusion chamber* maintained under constant light and constant temperature, with its liquid medium fully exchanged every 2 minutes. A full 3D scan (in fluorescence) of the root tip is composed of 60 optical sections $4\mu\text{m}$ apart, captured every 15 minutes for up to 7 days. To visualize mitotic events, *transient* fluorescent markers are used. In these plants, the fluorescent reporter accumulates in a cell transitioning between G2 and M phases of the cell cycle, and is quickly degraded after entering mitosis. Cells observed in 4D fluoresce during mitosis. By studying the 4D light-sheet image data, we wish to understand spatiotemporal dynamics of mitotic events under normal conditions and following root tissue damage.

The data analysis problem in this case is a problem of the first kind. From the perspective of the data processing, light-sheet microscopy presents some specific challenges. Imaging sessions typically last longer (hence there is more data to process) and the data may be more noisy and difficult to analyse because, for example, the subject may move around more in the *field of view*. In the setup described above, transient fluorescence markers are used. Tracking algorithms used in conjunction with fluorescent time-lapse microscopy data typically assume a continuous signal, where background and keypoints are permanently visible and objects of interest are permanently visible when not occluded. Such conditions allow for registration and identity management algorithms to track independently-moving objects. In the current context however, the need to track objects via transient fluorescent events arises. Under light-sheet microscopy conditions, the subject (root) moves around in the field of view. Cells are embedded in an almost rigid cell tissue structure, which acts to constrain independent object movement within the tissue but these objects are only visible for a short time and the global structure can move unpredictably. These conditions introduce distinct challenges.

An algorithm to deal with these data is described in Chapter 4. Two manuscripts, one emphasising the biology and one emphasising the novel tracking method are, respectively, in preparation and submitted. Details of the method are presented in the preprint [Amarteifio et al., 2018].

Quantifying the effect of cancer treatments on cell motility *The biology problem:* To understand the impact of a treatment on different leukaemia cell lineages observed before and after treatment (in mice) using 4D *fluorescent confocal microscopy*. We wish to characterize the *in vivo* migratory behaviour of acute myeloid leukemia (AML) cells. AML patients have a poor response to chemotherapy and survival of chemoresistant cells is thought to depend on leukemia-bone marrow microenvironment interactions, which are not well understood. This study aims to understand the migratory behaviour of AML and CXCR4 antagonism in the bone marrow and compare it to that of T-cell acute lymphoblastic leukemia (T-ALL) cells, which are highly motile in the bone marrow.

The data analysis problem: The problem considered here is a problem of the second kind. Having extracted tracks manually using a commercial software, 4D tracks (x, y, z, t) must be analysed to characterise differences in migratory behaviour across cell lines before and after treatment. Methods from statistical physics such as fractal measures and mean square displacement provide a more complete picture of spatiotemporal dynamics than do basic averages of speed and displacement.

In Chapter 4 results of this data analysis are presented. The work in collaboration with the Lo Celso lab at Imperial College London is presented in our recently published article [Duarte et al., 2018].

Declaration of contributions

Chapter 1 is background material developed from my own perspective. The perspective of generating functions and Markov semigroups and how these relate to quantisation of stochastic processes does not appear in this particular form elsewhere in the literature. Chapter 2 develops methods that borrow from other areas of quantum field theory applied in the context of non-equilibrium field theory and Markov processes. Parametric integrals and graph theoretic representations are used for the first time in this context. Chapter 2 also reviews an algorithm that I developed following a suggestion by my supervisor Gunnar Pruessner to exploit a certain bijection. Chapter 3 covers numeric methods developed by me and theory developed by my group. Notably, the data tabulated in Chapter 3 were prepared by Ignacio, Rosalba and Ben for our article. The RG analysis mentioned briefly in Chapter 3 and discussed in our article was carried out as a group. Chapter 4 discusses data analysis done by myself and an algorithm/code developed by myself and using data produced from experiments carried out by my collaborators.

1 Field theory of spatial Markov processes

The goal of this chapter is to map a *Markov process* onto a *field theory*. This is done by writing down a *path integral* representation of the process, which paves the way for using methods developed in perturbative quantum field theory. Yet it is not immediately obvious why one would wish to do this. Historically many useful results about the time evolution and statistics of Markov and birth-and-death-type processes have been developed in probability theory and the study of random processes. Initially such work has been carried out in a zero-dimensional setting. Starting from around the 1960s, the spatial case has been keenly studied. Handling space remains challenging and analytical methods are needed to determine spatial moments (correlation functions). Appropriate machinery for handling space and indeed fluctuations in critical spatial processes are hard to come by. It is here that a field theory may offer analytical insight, as suggested by physicists from around the 1970s. In the study of *dynamic critical phenomena* in particular, having found a path integral representation, one uses the *renormalisation group* to classify a spatial process in terms of its scaling exponents and *universality class*. Physically one might be interested in a process's phase transitions, which help in understanding how patterns emerge in spatial processes in $d \geq 2$ dimensions. The renormalization group and analysis of scaling are discussed in Chapter 3 and the workhorse of perturbation theory, the Feynman graph, is described in Chapter 2. For the moment Chapter 1 considers only the mapping of the spatial Markov process onto a path integral.

Tracing the classical treatment of birth and death processes, a common approach is to use the now textbook *generating function method* to map the differential form of the Chapman-Kolmogorov equation (or master equation) onto its *semigroup*. We will find actually, that much of the machinery that we need to map Markov processes to a path integral, bar some interpretation, already exists in the classical generating function method.* For the first half of this chapter, having introduced reaction-diffusion systems, I concentrate on the algebraic and combinatoric aspects of the mapping. This allows me to postpone the introduction of ideas that are specific to field theory until the later parts of the chapter. By then we should already have a good idea about how to write down a path integral. In the next section I quickly provide a birds-eye view before getting into the details.

1.1 Overview

Of central importance is the *compound Poisson process*. It can be described by a chemical reaction network, a collection of directed reactions. Consider for example the following set

*By extension, as the path integral is known as a generating functional we might begin by simply thinking of this path integral mapping as a *generating functional method for spatial Markov processes*.

of reactions known as the Gribov process.¹



Each reaction of the form $kA \xrightarrow{\lambda} lA$ transforms k particles of species A to l particles of species A at a rate λ . As written there is no concept of space so we embed this process on a graph such as a regular hypercubic lattice, where configurations $\mathcal{P}(n_A(\mathbf{x}), t); \mathbf{x} \in \mathbb{R}^d, t \geq 0, n_A < \infty$ describes the time evolution of finite configurations of a discrete spatial process. Particles are transported strictly by linear diffusion and the spatiotemporal operator takes the form $(\partial_t + \Delta)$ where $\nabla^2 = \Delta$ is the Laplacian operator. For each reaction we can write down a *factorial moment semigroup* with infinitesimal generator

$$\mathcal{Q}_\lambda[\partial_z, z] = \lambda((z+1)^l - (z+1)^k) \partial_z^k \quad (1.2)$$

Adding together all such generators from the chemical reaction network, the integrand of the Lagrangian density takes the form

$$\widehat{\mathcal{L}}[\partial_z, z] = \underbrace{z(\partial_t + \Delta)\partial_z}_{\mathcal{L}_0} - \underbrace{\sum \mathcal{Q}_j[\partial_z, z]}_{\mathcal{L}_I} \quad (1.3)$$

where I have distinguished between the *free* or *bilinear* part of the Lagrangian (corresponding to the spatiotemporal operator) and the *interaction* part (corresponding to the chemical reactions). Depending on one's perspective, the interactions can contribute bilinear terms to $\mathcal{L}_0$ but the expression as written is a convenient simplification for the time being. A semigroup from the generating function method can always be written

$$\Psi_t = e^{\mathcal{Q}t} \Psi_0 \quad (1.4)$$

and the *path integral* or *generating functional* can be written

$$Z = e^{\int -\mathcal{L}[\phi, \tilde{\phi}]} \quad (1.5)$$

having formally transitioned from operators to fields in the generators $\mathcal{Q}[\phi, \tilde{\phi}]$. Expectations of observables can be determined by a convolution with a *Lie-Kato-Trotter* approximation of the functional semigroup and can be abstractly written;

$$\langle \mathcal{O} \rangle = \langle \mathcal{O}, e^{\mathcal{L}t} \rangle \quad (1.6)$$

The goal of this chapter is to explore this mapping. Generally one says that having carried out this mapping, the bilinear part is solvable by Gaussian integration while the non-linear part can be dealt with *perturbatively* using Feynman diagrams, which will be discussed in later chapters.

1.2 Reaction-diffusion systems

Reaction-diffusion systems are formulated in terms of *microscopic interactions* where *particles* interact at a certain rate in a compound, changing the concentration of reactants or indeed *creating* and *destroying* reactants. The microscopic picture is quite useful in many

applications. The word "chemical" often appears for example in the description of *chemical reaction networks* or *chemical master equation*. While reaction-diffusion processes may well model true chemical reaction networks, often the word "chemical" is used metaphorically to describe systems that are in fact ecological, social, biological or physical. For any system where agents, particles, animals, people, companies and so on interact and one wishes to derive an equation of motion from the "bottom up", the perspective of (chemical) reaction-diffusion processes is useful. The word *particle* will be used to describe atoms of these processes while I retain in some cases the metaphor of *chemical reactions* to make links with the literature. The use of the word "chemical" is historically rooted to work in the first part of the twentieth century for example by Smoluchowski [1916] or Delbrück [1940]. In spatial contexts one considers not only the reactions but particle transport. Such models are referred to as *reaction-diffusion processes* or *interacting particle systems* [Liggett, 2012].

I will model reaction-diffusion processes as site-bound reactions on a graph, formulated as a compound Poisson process, juxtaposed with particle transport by linear, isotropic diffusion. In the reaction-diffusion formulation, one starts with the microscopic interactions in order to derive an equation of motion for a stochastic process. In the context of this thesis, a reaction-diffusion process is a spatial Markov Process. Markov processes will be defined below but the Markov property i.e. the memoryless property is a useful simplification to which we will adhere.

1.2.1 Some history

Using stochastic processes to model natural processes has a long history. Guttorp [1995] gives a brief account in his introduction describing the origin of the various textbook processes that are studied in introductory random processes. These basic processes feature throughout this thesis.

- The textbook example of a *Markov chain* (discrete time *Markov process*) is that modelling rain fall patterns from day to day as studied by Quetelet [1852].
- The simple *branching process* was used to compute the probability of noble family names to become extinct Bienaymé [1845]. Often this is described as the Galton-Watson process as they undertook the same study some time later Watson and Galton [1875].
- The *Poisson process* was developed by Rutherford et al. [1910].
- The *birth and death process* as an example of a compound Poisson process with non-linear interactions was introduced by McKendrick [1914, 1925] and developed in the context of the *generating function method* by Kendall [1948].
- Early examples of modelling chemical reactions include Smoluchowski [1916] and Delbrück [1940].
- The phenomenon of *Brownian motion* observed by the botanist Robert Brown [Brown, 1828] was used by Einstein [1905] to describe the motion of suspended

particles in a solution and by Bachelier [1900] to describe stock price movement. Furthermore one finds Brownian motion describing movement of random particles in *Fickian diffusion* [Fick, 1855].

After the 1960s the spatial context has been considered. Etheridge [2000] gives an introduction to *superprocesses*, which combine two sources of randomness; (i) production of particles by a branching mechanism such as Galton-Watson-Bienaymé branching processes with (ii) particle transport for which increments come from Lévy α -stable distributions. A particularly important special case, for which the branching process is the jump distribution, is a Brownian motion (or diffusion in the discrete case) is known as the Dawson-Watanabe process or *branching random walk* [Etheridge, 2000]. It is this process that we derive a field theory for in Chapter 3. As we pay particular attention to the *trace* (volume explored) of the superprocess, the description *branching Wiener sausage* [Donsker and Varadhan, 1975, Nekovar and Pruessner, 2016] is also used.

Leaving the classically studied mathematical objects and turning to applications, it becomes more difficult to give a concise historical account as different language and different methodologies might be used to study the same fundamental processes; applications to problems in epidemiology and ecology with and without the geometry of space can be described as *population models*. Dawson [2017] gives an excellent overview which builds on the classical processes introduced. Liggett [2012] presents many useful results in the framework of *interacting particle systems*. For a notably mathematical perspective on spatial Markov processes one can consult for example the thesis of Friesen [2016] and references therein. As discussed in detail below, Doi [1976a] was the first to make the link between statistical methods used for random "chemical" process and the framework of quantum mechanics and quantum field theory.² In essence Doi's methods is a reformulation of the generating function method as developed by Kendall [1948] in the language of quantum mechanics. It is a useful metaphor. As further clarified by Peliti [1985], this allows for a generating functional (path integral) representation of birth and death processes which paves the way for perturbative expansions and renormalization group analysis of such stochastic processes. Ultimately one wishes to study the *critical* dynamics of these systems [Vasil'ev, 2004, Täuber et al., 2005, Stollenwerk and Jansen, 2011].

A particular concern when dealing with population models in space is the calculation of spatial moments or correlation functions. This leads to a problem where moments need to be expressed in terms of higher order moments *ad infinitum* leading to various heuristic moment closure techniques [Van Kampen, 1992, Plank and Law, 2015]. A perturbative expansion in a small parameter offers an alternative approach to moment closure. This of course is the route taking in the field theory, where Feynman diagrams are used as the bookkeeping device in this expansion. I am aware of at least one example outside of the field theory where a perturbative perspective is taken in population models in the work of Ovaskainen et al. [2014]. Another example related to population models and chemical reaction networks, which does not develop the field theory program but does use Doi's formulation with a perturbative approach to chemical processes, can be found in the series of articles [Krishnamurthy and Smith, 2017, Smith and Krishnamurthy, 2017, 2018].

The reaction part of a reaction-diffusion processes is formulated as a chemical reaction network. In the strictly Markov case at least, this is a compound Poisson process. I make a few comments on the chemical reaction network picture. As mentioned, early example of

studying the mathematics of individual reactions goes back to at least Smoluchowski [1916] and Delbrück [1940]. A concept of central importance is that of *detailed balance*³ as worked out by Wegscheider [1911]. The mathematical theory of chemical reaction networks was developed in the 1960's and 1970s. For example *detailed balanced* was generalized by Horn and Jackson [1972] to the notion of *complex balanced* reaction networks while Feinberg [1972] developed a useful formulation for general chemical reaction networks. Recently, a nice review by Baez and Biamonte [2012] attempts to unify concepts and language coming from quantum field theory with that of probability theory in order to describe chemical reaction networks, where they discuss conditions under which complex-balanced reaction networks have solutions that reduce to a product of Poisson distributions.

Reaction-diffusion processes on non-regular graphs (something that is not a hypercubic lattice) are discussed in Chapter 3 and have enjoyed some attention in the literature [Lindenberg et al., 1991, Gallos and Argyrakis, 2004, Catanzaro et al., 2005, Baronchelli et al., 2008, Yun et al., 2009, Zhang et al., 2011]

1.2.2 Formulation

Definition 1. A *chemical reaction network* $\mathcal{C}$ is a compound Poisson process with sub-processes or reactions of the form $\sum_{i=0}^m k_i A_i \xrightarrow{\lambda} l_i A_i$, for m species $\in \{A, B, \dots\}$ with λ being a Poisson rate parameter. The systems we care about are *non-equilibrium systems* i.e. the system of elementary reactions do not typically satisfy *detailed balance* (cf. C.6).

Definition 2. A *discrete chemical reaction-diffusion system* $S(\mathcal{C}, G, W)$ is a chemical reaction network $\mathcal{C}$ living on a graph G . All reactions occur at sites $x \in G$. Particles hop continuously between sites according to a jump distribution W , which in the scope of this thesis will be linear diffusion (Brownian motion in the continuum). The state of the system is described by a *configuration* $\mathcal{P}(n_A(x); t) : t \geq 0, n_A(x) < \infty$.

Remark 2. Take for example G to be a regular hypercubic lattice with configurations $\mathcal{P}(n_A(\mathbf{x}); t) : \mathbf{x} \in \mathbb{R}^d$

Definition 3. The *standard particle proclivity* describes a process for which particles are identical and *bosonic* and reaction rates are proportional to particle density.

Remark 3. *bosonic* (as opposed to *fermionic*) means there is no exclusion principle at lattice sites i.e. particles can co-occupy sites.

Remark 4. The standard particle proclivity implies that, in such reactions, the combinatorics of particle sampling is described by the *permutations* ${}_nP_k$ i.e. sampling k particles from n particles without replacement.

A (compound) Poisson process is a *Markov process*, which as discussed in the next section admits a *semigroup*. Doi [1976a] gave a *second-quantized* interpretation of the semigroup mapping procedure. I describe this mapping primarily from the perspective of combinatorics and formal power series in the next few sections and provide an overview from the perspective of quantum field theory towards the end of this chapter. I add an endnote about terminology.⁴

1.3 Heisenberg-Weyl semigroups

I begin with general statements, which are to be developed in the next few subsections.

Definition 4. A *Heisenberg-Weyl semigroup* determines the time evolution of the system state Ψ_t and takes the form

$$\Psi_t = e^{\mathcal{Q}[\partial_z, z]} \Psi_0 \quad (1.7)$$

The *infinitesimal generator* $\mathcal{Q}$ of the semigroup is generated by a Heisenberg-Weyl algebra.

For every chemical reaction equation $kA \rightarrow lA$, a Poisson process within the *standard particle proclivity*, there is a corresponding *infinitesimal generator* with respect to a Markov *semigroup*. It turns out the interesting semigroup corresponds to the time evolution of *factorial moments* and its generator is

$$\mathcal{Q}[\partial_z, z] = \left((z+1)^l - (z+1)^k \right) \partial_z^k \quad (1.8)$$

where $[\partial_z, z]$ is the Lie bracket of the Heisenberg-Weyl algebra $\mathfrak{g}$. The aim of this section is to explain this statement. These $\mathcal{Q}$'s are the caterpillars that turn into certain colourful butterflies we call *Feynman diagrams*. The perspective I shall take here borrows loosely from the interpretation of formal power series in the sense of the *Umbral calculus* [Roman, 2005]. Furthermore I emphasise the particular importance of Poisson random variables and their generating functions in the formalism. In the next two subsections I give some facts about semigroups and the Heisenberg-Weyl algebra and then discuss their use in terms of generating functions.

1.3.1 Markov semigroups

This subsection briefly reviews well-known properties of Markov processes [Stroock, 2013] and their semigroups [Engel and Nagel, 1999]. The key property of a Markov process is that it has no memory; a Markov process conditioned on the present is independent of the past. Examples are random walks (Brownian motion) where each increment is drawn independently or a Poisson process, which will feature heavily in this chapter.⁵ A Markov process can be written in the form of the Kolmogorov equations. For linear operator U ,

$$\partial_t \Psi = U \Psi \quad (1.9)$$

When the operator does not change with time the Markov process is called *time-homogeneous*.

An important equation is the *Chapman-Kolmogorov equation*. It says that for a Markov process transitioning from x at s to y at t via some intermediate space time points $(\mathbf{z}, \mathbf{u})$, it can be broken into two steps

$$(x, s) \rightarrow (\mathbf{z}, \mathbf{u}) \rightarrow (y, t) \quad (1.10)$$

here summing over the set of transitions between all intermediate time points. This implies that a Markov process can be described by its *semigroup*⁶ of operators, which are a one-parameter family of linear operators characterised by the *infinitesimal generator* U . The

semigroup property is a restatement of the Chapman-Kolmogorov condition for a discrete homogeneous Markov chain;

$$U(s+t) = U(s)U(t) \quad : \quad s, t \geq 0, \quad U(0) = \mathbb{1} \quad (1.11)$$

One distinguishes the perspectives of how the time evolution of the processes are studied. The *backward* equation with adjoint operator U^* considers the evolution of observables[†] while the *forward* equation considers the evolution through states[‡]. Both the backward and forward Kolmogorov equations can be written in the form of (1.9). The forward equation is also known as the *Fokker-Planck equation*. Both forms are formally solved by their respective semigroup;

$$\Psi_t = e^{U_t} \Psi_0 \quad (1.12)$$

This equation is written formally and is not necessarily exactly solvable. However, for a Markov process it is sufficient to study the semigroup of the process to determine the statistics of the process.

Example 1.3.1. Brownian motion is an example of a homogeneous Markov process. Its generator is $\frac{1}{2}\partial_x^2$ and its semigroup is $\Psi_t = e^{t \cdot \frac{1}{2}\partial_x^2}$

Example 1.3.2. the Schrödinger equation is a semigroup with U being a *unitary operator* having the semigroup property.

What is Ψ ? Certainly it describes something about the evolving state of the system. In the Doi picture this is interpreted as a (Bargmann-)Fock space. I think of this as merely a *formal power series* in the *Umbral* sense (§C.10). The formal power series is the generating function for the statistics of the processes.

The purpose of the generating function method is to rewrite the so-called chemical master equation in terms of its semigroup having identified U . The Markov operator or generator U determines the evolution of dynamical spatial systems, possibly with stochastic fluctuations. While the dynamical system may first be thought of as the time evolution of entire (particle) point configurations, we can focus on the evolution under a probability measure. Such measures describe point distributions in phase space. This measure corresponds to what we might refer to in the field theory context as *coherent states*.

Remark 5. There is perhaps a subtlety in changing between measures. Later we consider both the generating function of the configuration of occupation numbers and also the generating function of the factorial moments. Furthermore, at some point we make a transition from *operators* to *fields* (densities). The latter is done formally and if one blinks one might miss it. The intuition for *change of basis* (between formal power series) or *change of measure* (shift to fields) is described towards the end of the chapter.

Example 1.3.3. For continuous-time Markov processes, we can encode the transition between states in a time dependent matrix. For example, consider a simple Poisson process defined by $\mathbb{P}(N_t = k) \sim \frac{(\lambda t)^k}{k!} e^{-\lambda t}$.

[†]ala Heisenberg or Koopman

[‡]ala Schrödinger or Frobenius-Peron

We can write the transition semigroup

$$\mathbf{P}(t) = \begin{bmatrix} e^{-\lambda t} & \lambda t e^{-\lambda t} & \frac{\lambda^2}{2} t^2 e^{-\lambda t} & \dots \\ 0 & e^{-\lambda t} & \lambda t e^{-\lambda t} & \dots \\ 0 & 0 & e^{-\lambda t} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (1.13)$$

The matrix obeys the *semigroup property*

$$P(s+t) = P(t)P(s) \quad (1.14)$$

which is related in this context to the Chapman-Kolmogorov equations which describe a possibly infinite dimensional matrix multiplication. By defining the derivative

$$\mathbf{P}'(t) = \lim_{h \rightarrow 0} \frac{\mathbf{P}(t+h) - \mathbf{P}(t)}{h} \quad (1.15)$$

we can express this as $\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}$, where the matrix $\mathbf{Q}$ is the infinitesimal generator describing the probabilities of transitioning between states. This matrix is a *stochastic matrix*, which is to say, positive entries in each row sum to 1. The equations $P'(t) = QP(t)$ and $P'(t) = Q^*P(t)$ are the backward and forward Kolmogorov equations, formally solved via a matrix exponential, where Q^* is the adjoint and $P'(t)$ is the time derivative of the state. The matrix exponential in d dimensions is the series

$$e^{t\mathbf{Q}} = \mathbf{I}_d + \sum_{n=1}^{\infty} \frac{t^n}{n!} \mathbf{Q}^n \quad (1.16)$$

We can calculate for example the stationary distribution π by solving $\pi\mathbf{Q} = 0$ where the elements of π , being a distribution, of course sum to one •

Extended semigroups: Semigroups appear in the interpretation of the path integral and in particular the means by which time evolution is split into infinitesimal discrete intervals. In order to obtain a path integral representation one exploits the Lie-Kato-Trotter product formula [Trotter, 1959] in addition to properties of coherent states. In particular one writes the approximation

$$e^{-Qt} = \lim_{n \rightarrow \infty} \left(1 - \frac{Qt}{n}\right)^n = \lim_{\Delta t \rightarrow 0} (1 - Q\Delta t)^{\Delta t/t} \quad (1.17)$$

which is an extension of the *Lie product formula* for complex (or real) matrices

$$e^{A+B} = \lim_{N \rightarrow \infty} (e^{A/N} e^{B/N})^N \quad (1.18)$$

The approximation (1.17) is used when breaking up the (path integral) evolution operator into a sequence of *infinitesimal*, tractable steps. Armed with this, in the coherent state representation one can construct an eigenspace from a complete set of coherent states. The relation (1.19) below is the resolution of identity (I write with integration measure $\mathbb{D} = \pi^{-1} d\phi d\phi^*$)

$$\mathbb{1} = \int \mathbb{D} e^{-\phi^* \phi} e^{\phi z} |0\rangle \langle 0| e^{\phi^* \partial_z} \quad (1.19)$$

and the interested reader can find a proof of this being satisfied for the coherent state path integral in [Wiese, 2016] (§I). This identity is inserted into the *time-loop*, producing terms of the form below (after appropriate permuting and commuting of terms)

$$\langle 0 | 1 - \mathcal{Q} \Delta t | 0 \rangle e^{\phi^*(t') \phi(t)} \quad (1.20)$$

Further details in the context of coherent state path integrals can be found in [Cardy, 2006, Wiese, 2016, Weber and Frey, 2017] or [Negele and Orland, 1988] (ch. 2). My interest here extends only to motivating the perturbative expansion of $\mathcal{Q}$ in diagrammatic form as a realisation of the path integral.

1.3.2 Heisenberg-Weyl algebra

The infinitesimal generator is understood to be generated by a Heisenberg-Weyl algebra. This is a differential algebra $\mathfrak{g}$ with Lie bracket

$$[\partial_z, z] = \mathbb{1}, \quad \partial_z z - z \partial_z = \mathbb{1} \quad (1.21)$$

The definition extends to multiple species, i, j say, by use of the Kronecker delta

$$[\partial_{z_i}, z_j] = \delta_{ij} \quad (1.22)$$

Remark 6. In what I will refer to as *the Doi picture* [Doi, 1976a], to be described in more detail later⁷, the bracket is satisfied by the ladder operators $[a, a^\dagger]$; the creation operator $a^\dagger$ and the annihilation operator a . While the operators ∂_z, z can represent the familiar differential algebra (multiplication and partial derivative), more generally it is any pair of operators that satisfy the commutation relations.

Combinatorics of $\mathfrak{g}$ are given by its action on monomials, where monomials appear as $|n\rangle := z^n$ in Dirac's notation;

$$z |n\rangle = |n+1\rangle \quad (1.23a)$$

$$\partial_z^k |n\rangle = n_{(k)} |n-k\rangle \quad (1.23b)$$

$$z^k \partial_z^k |n\rangle = n_{(k)} |n\rangle \quad (1.23c)$$

I have introduced the falling factorial (*permutations*) notation $n_{(k)} = {}_n P_k$,

$$n_{(k)} = \frac{n!}{(n-k)!} \quad (1.24)$$

which is the number of unordered samples of k particles from n particles without replacement. This is particularly relevant for Poisson spatial statistics as shall be discussed. The terms $z^k \partial_z^k$ are *factorial moments* and a term $z \partial_z = \hat{n}$ is called the *number operator*. These moments are related by the Stirling numbers⁸

$$z^k \partial_z^k = \begin{bmatrix} k \\ \underline{l} \end{bmatrix} \hat{n}^l, \quad \hat{n}^l = \left\{ \begin{matrix} l \\ \underline{k} \end{matrix} \right\} z^k \partial_z^k \quad (1.25)$$

⁸My convention: an underlined variable is summed over e.g. in the first bracket, $\sum_{l=0}^k$

Monomials in the algebra can be expressed in terms of (tree-level) graphs or *corollas*. This is done by having as many legs entering from the left as there are ∂_z 's and as many legs exiting from the right as there are z 's. Later we will map operators to fields $z \rightarrow \tilde{\phi}$ and $\partial_z \rightarrow \phi$ and the representations will transfer naturally.

Example 1.3.4.

$$z^2 \partial_z = \text{diagram} \quad (1.26)$$

1.3.3 Generating functions

Recall the *ordinary generating function* for random variable N and distribution p_n ,

$$\Pi_N(z) = \mathbb{E}[z^N] = \sum_n p_n \cdot z^n \quad (1.27)$$

The *factorial (exponential) moment generating function* can be expressed as an ordinary moment generating function (see also §C.11). By definition of the exponential generating function with $a_n = n_{\{k\}}$ and using the binomial theorem (§B.10)

$$\Psi_N(z) = \sum_{k \leq n} n_{(k)} \frac{z^k}{k!} = \mathbb{E} \left[\sum_{k \leq n} \binom{n}{k} z^k \right] = \mathbb{E}[(1+z)^N] = \Pi_N(1+z) \quad (1.28)$$

The generating function method Starting from this microscopic equation, the time evolution of $\mathcal{P}_t$, we can apply the generating function method, which is to say use the Heisenberg-Weyl algebra and determine the semigroup.

Under the assumptions given in formulating the problem, a *chemical master equation*⁸ (i.e. the differential form of the Chapman Kolmogorov equations) for any reaction $kA \rightarrow lA$ can be written in terms of the falling factorials $n_{(k)}$ introduced above. The chemical master equation, the time evolution of the probability, for transitions between states n and n' is of this form

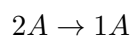
$$\partial_t P(n, t) = \underbrace{R_{n' \rightarrow n} P(n', t)}_{gain} - \underbrace{R_{n \rightarrow n'} P(n, t)}_{loss} \quad (1.29)$$

$R_{i \rightarrow j}$ denote the rate of moving between states i, j . Under the assumptions we have set in formulating reaction-diffusion processes, the chemical master equation takes a combinatoric form

$$\partial_t P(n, t) = (n + k - l)_{(k)} P(n, t) (n + k - l) - n_{(k)} P(n, t) \quad (1.30)$$

While this general form is well-known (cf. Behr et al. [2017]), it is useful to consider an example.

Example 1.3.5. The curious case of coagulation: Coagulation involves the reaction



A master equation generally takes the form $\partial_t P = P_{gain} - P_{loss}$. The gain state is the state for which two particles are just about to coagulate therefore there is one extra particle

with respect to the reference state, which we write as $n_x + 1$ for site x . I will omit the x subscript henceforth. For the two particles involved, I can sample from $n + 1$ particles *without replacement* and then from n remaining particles. The loss state is the flow out of the reference state. In this state there are n particles, any two of which might coagulate. For the two particles involved, I can sample from n particles *without replacement* and then from $n - 1$ remaining particles. Omitting the x which is understood to be summed over for all sites, I write this equation as

$$\partial_t P(n, t) = \chi [n(n + 1)P(n + 1, t) - n(n - 1)P(n, t)]$$

Using the definition for $n_{(k)}$ given in (1.24), this agrees with the general statement for the combinatorial master equation (1.30) •

Remark 7. Actually, the representation above is the zero dimensional one. On a lattice, we can simply rewrite this expression summing over all lattice sites, making the replacement $n \rightarrow n_x$.

The generating function method rewrites the time evolution of the probability as the time evolution of a moment generating function. In probability theory and combinatorics, this basis is a formal power series while in the Doi picture it is a Bargmann-Fock space. Doi [1976a] uses Dirac's concise notation to write the generating function. For multiple particle species,

$$|\Psi_t\rangle = \sum_{n,m} p_t(n, m) z^n z^m |0\rangle = \sum_{n,m} p_t(n, m) |n, m\rangle \quad (1.31)$$

The zero-ket represents the *vacuum*. By applying *raising* operators to it we recover the pure states $|n, m\rangle$. Now the general statement for writing down generators for the semigroup follows an example.

Example 1.3.6. The curious case of coagulation continued: The Heisenberg-Weyl algebra is used to find the generator of the semigroup. By the definition of the ordinary generating function and in particular its Dirac representation (1.31), map the time evolution of the configuration P to the time evolution of its statistics generating function Ψ . Plugging in our P for p in (1.31)

$$\partial_t |\Psi_t\rangle = \chi \sum_n [n(n + 1)P(n + 1, t) - n(n - 1)P(n, t)] |n\rangle \quad (1.32)$$

Using the Heisenberg-Weyl algebra this can be written as

$$\partial_t |\Psi_t\rangle / \chi = \sum_n z \partial_z^2 P(n + 1, t) |n + 1\rangle - z^2 \partial_z^2 P(n, t) |n\rangle \quad (1.33a)$$

$$= \sum_n z \partial_z^2 P(n, t) |n\rangle - z^2 \partial_z^2 P(n, t) |n\rangle \quad (1.33b)$$

$$= \sum_n [z \partial_z^2 - z^2 \partial_z^2] P(n, t) |n\rangle \quad (1.33c)$$

$$= \underbrace{[(z - z^2) \partial_z^2]}_{\mathcal{Q}} \Psi(t) \quad (1.33d)$$

The manipulations are easily carried out but just to be clear, we wish to express the master equation in terms of the generating function. Using the definition of the generating function and the algebra of the operators, we essentially "factor out" $\mathcal{Q}$. We do this by

applying the right combination of creation and annihilation operators to express the state as operators acting on a pure state $|n\rangle$ which *explain* the combinatorial rate terms such as $n(n+1)$. Take for example the first term in (1.33b). Now apply each of the operators $\partial_z, \partial_z, z$, in that order, to the state $|n+1\rangle$.

$$\begin{aligned}
 & n(n+1)P(n+1, t) |n\rangle \\
 & \implies z\partial_z\partial_z P(n+1, t) |n+1\rangle \quad (= z\partial_z\partial_z P(n, t) |n\rangle) \\
 & \implies (n+1)z\partial_z P(n+1, t) |n\rangle \\
 & \implies n(n+1)\mathbf{z}P(n+1, t) |n-1\rangle \\
 & \implies n(n+1)P(n+1, t) |n\rangle
 \end{aligned}$$

In the first step I write the generating function of the probability. On the far right in the second line I note the equivalence under relabelling of indices. The remaining steps just recover the first line from the second by applying the operators one by one. All terms in a chemical master equation including those with multiple species can be treated in a similar manner •

Remark 8. While this is a useful exercise and allows for any microscopic process to be written down first as a master equation and then as a semigroup, given the combinatorics of particle sampling in the standard particle proclivity, it is certainly more convenient to write any reaction $kA \rightarrow lA$ directly as either a master equation using (1.30) or as a semigroup using generators (1.35) or (1.36) below. With reference to (1.35) below, observe that in (1.33d) $k, l = 2, 1$

General forms... Using the Heisenberg-Weyl algebra, the infinitesimal generator $\hat{Q}$, corresponding to the semigroup acting on the ordinary generating function of the *particle occupation numbers*, can be written

$$\hat{Q}[\partial_z, z] = (z^l - z^k)\partial_z^k \quad (1.35)$$

Using the Heisenberg-Weyl algebra, the infinitesimal generator Q corresponding to the semigroup acting on the ordinary generating function for *factorial moments* can be written

$$Q[\partial_z, z] = ((z+1)^l - (z+1)^k)\partial_z^k \quad (1.36)$$

Remark 9. Factorial moments correspond to *correlation functions* under a Poisson ansatz; interactions occur in a particle bath described by Poisson point processes in $\mathbb{R}^d$ - complete spatial randomness. Such configurations correspond to *coherent states* [Glauber, 1963].

Example 1.3.7. By now we have made sufficiently many restrictions that we are in a position to simply plug in parameters to formulae. For the Gribov process (1.1) in particular, for the single species A , the stoichiometric matrix is

$$S_{kl} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 0 \end{bmatrix} \quad (1.37)$$

And the (factorial moment) infinitesimal generators are

$$\mathcal{Q}_\chi[\partial_z, z] = \chi((z+1)^1 - (z+1)^2) \partial_z^2 \quad (1.38a)$$

$$\mathcal{Q}_\beta[\partial_z, z] = \beta((z+1)^2 - (z+1)^1) \partial_z^1 \quad (1.38b)$$

$$\mathcal{Q}_\epsilon[\partial_z, z] = \epsilon((z+1)^0 - (z+1)^1) \partial_z^1 \quad (1.38c)$$

We can write down the action terms (integrand corresponding to Lagrangian density) by summing the generators (terms and their signs follow from expanding 1.38)

$$\mathcal{A}[z, \partial_z] = z(\partial_t + \Delta) \partial_z - (\epsilon - \beta) z \partial_z - \chi z^2 \partial_z^2 - \chi z \partial_z^2 + \beta z^2 \partial_z \quad (1.39)$$

Any chemical reaction network in the standard proclivity can be mapped to an action in this formulaic way •

Remark 10. Arbitrary reaction-diffusion processes can be mapped from first principles, a mechanical but tedious process, using a suitable algebra as shown in the examples.

Already we have achieved plenty. The Heisenberg-Weyl differential algebra produces operators that can be perturbatively expanded in diagrammatic notation. This is a combinatoric exercise discussed in Chapter 2. The semigroup in restricted cases, such as the simple birth and death process,[¶] can be solved in zero-dimensions using the *method of characteristics* (cf. [Allen, 2010]), which exploits the symmetry of the continuous one-parameter semigroup. In fact, in Chapter 3 we will exploit another symmetry, a scaling symmetry, in the one-parameter group that is the *renormalization group*. There too one can use the method of characteristics to solve the *renormalization group equations* [Zinn-Justin, 1996].

1.4 Diffusion, propagators and responses

In Chapter 4, Brownian motion is used as a yardstick to classify cancer cell motility regimes. This follows from Brownian motion's important role in physics (following seminal work of Einstein [1905]) to describe the free, random movement of particles in a medium. In reaction-diffusion processes as formulated in §1.2.2, diffusion describes the combined motions of a population of particles performing random walks on a lattice. The emphasis on restricting particle transport to linear diffusion earlier in this chapter may obscure the general importance of diffusion/Brownian motion in formulating any classical field theory. Indeed Itzykson and Drouffe [1991] (ch. 1) make a point of constructing Euclidean field theory from Brownian motion. In the context of *Euclidean* or *statistical*⁹ field theory *per se*, we might think of Brownian motion in terms of a *measure* relating to the volume of phase space.

Accounting for reentrant or intersection probabilities of diffusing particles can lead to an effective change in the interaction rates given by a mean-field theory (i.e. where the geometry of space is *not* taken into account). The Poisson interaction rates thus need to be *renormalized* when developing an *effective* field theory in order to accurately describe observables in reaction-diffusion processes. Consider for example intersections of Brownian

[¶]The simplest process that still contains an interesting non-linearity

paths in d dimensions (discussed for example by Lawler and Limic [2010] ch. 10) or indeed intersections of trajectories of an evolving population of Brownian particles i.e. a *superprocess* in d dimensions [Etheridge, 2000]. Random walk intersection probabilities influence the rate at which site-bound reactions occur by diminishing or replenishing local reactants and thus play a central role in Euclidean field theory and in the study of critical phenomena. Now where exactly does diffusion enter in field theory calculations? Diffusion becomes manifest in *the propagator*, which is introduced below.

Diffusion operator on graphs

Diffusion/Brownian motion and Poisson processes are examples of homogeneous Markov processes. It is possible to interpret the diffusion semigroup in the same manner as was done for the Poisson processes above. Introduce a generator for particles hopping between states (sites) i, j

$$\mathcal{Q}_D = D(z_j - z_i)\partial_{z_i} \quad (1.40)$$

or by superposition of symmetric processes, back and forth between sites

$$\mathcal{Q}_D = D(z_j - z_i)(\partial_{z_i} - \partial_{z_j}) \quad (1.41)$$

In the continuum limit this is

$$\mathcal{Q}_W = -D \int_x \nabla z(x) \nabla \partial_z(x) \quad (1.42)$$

Hence the Laplace operator $\nabla^2 = \Delta$ will be the general spatial operator used for reaction-diffusion processes where reactions occur at sites and particles are transported by linear diffusion (see §C.2).

Diffusion on hypercubic lattices is the most straightforward to understand but in Chapter 3, we will also consider the distinct sites visited by a superprocess on non-regular graphs. This requires techniques for dealing with diffusion on general graphs [Burioni and Cassi, 2005, Masuda et al., 2017], as we shall see.

1.4.1 Introducing the propagator

The short version of the story of defining a propagator for non-equilibrium systems is to consider the Fourier transform of the spatiotemporal operator $(\partial_t + D\Delta)$ that we introduced for the reaction-diffusion processes under consideration. This gives something of the form

$$\frac{1}{-i\omega + Dk^2 + m^2} \quad (1.43)$$

This object is analogous to the propagator in static Euclidean field theories where the *d'Alembert operator* $\square$ replaces $(\partial_t + D\Delta)$. In the case of critical dynamics, the interpretation of the squared mass is as a parameter that measures the distance from criticality in terms of the rate parameters of the compound Poisson processes. For example in a branching process, the critical point is $\beta = \varepsilon$ therefore $m^2 \leftarrow |\beta - \varepsilon|$. Making the connection with critical statics, this is analogous to a *temperature* parameter in ϕ^4 theory.

The longer version of the story, or at least one version of it, begins with Schwinger [1961], Keldysh [1964] and others (c.f. [Kamenev, 2002, 2011, Stefanucci and Van Leeuwen, 2013]) where one asks the question *how does one say anything at all about interacting systems out of equilibrium?* One of the postulates of critical dynamics is that it makes sense to apply machinery developed in the equilibrium context in the non-equilibrium context of critical dynamics [Vasil'ev, 2004].

Remark 11. In the case of coherent state path integrals one carries out this procedure with a Poisson ansatz (i.e. background noise distribution characterised by complete spatial randomness) [Glauber, 1963].

[Kamenev, 2002] gives a short account of the Keldysh technique which motivates a free (quadratic) action term independent of the particle hopping argument in the last section. The free part of the action reads

$$S_0[\tilde{\phi}, \phi] = \frac{1}{2} dt(\phi, \tilde{\phi}) \begin{pmatrix} 0 & \overbrace{-\partial_t + D\Delta}^A \\ \underbrace{\partial_t + D\Delta}_R & K \end{pmatrix} (\phi, \tilde{\phi})^T \quad (1.44)$$

where A, R, K relate to the *advanced*, *retarded* and *Keldysh* terms. For example the advanced propagator is

$$\langle \phi(t) \tilde{\phi}(t') \rangle = \frac{1}{-i\omega + D\Delta} \quad (1.45)$$

These correlation functions translate into statements such as *create a particle at time t' and then probe the system for a response at time t* , or *vice versa*. By now it should be clear that any conclusions we draw will be based on random walk and Poisson statistics.

1.5 Perturbative expansions and path integrals

1.5.1 Some history

Martin et al. [1973] may have been the first to consider the applications of quantum field theory machinery to classical systems and stochastic processes. I reproduce here the first paragraph of their 1973 article as it sets the scene nicely, providing some motivation:

Despite the deluge of papers that have been written over many years on the organization and calculation of the statistical properties of classical systems, there is not to our knowledge a satisfactory theory with the utility, generality, and precision of the quantum-field theories. In particular there is no parallel to the functional equations of Schwinger, or the equivalent diagrammatic techniques of Feynman for expressing the statistical and dynamical properties of a classical system, conservative or dissipative, in terms of closed albeit complicated equations involving the first few of the exact correlation functions for the system. As a consequence there is no renormalized perturbation theory for a classical system.

The Martin-Siggia-Rose program leads to *generating functionals* for describing the evolution of correlation generating functionals (as opposed to complete time evolution of configuration densities). Doi [1976a,b] implemented this program using a *second quantisation*¹⁰ representation of classical processes that leads conveniently to writing down a (Markovian) semigroup. As such his approach, which turns out to be analogous to the *generating function method* [Kendall, 1948] but for a basis of coherent states, bound the MSR formalism elegantly to reaction-diffusion Markov processes. While many ideas used by Doi already existed within the field of probability, his mapping, particularly after its exposition by Peliti [1985],¹¹ made the formalism popular with physicists studying the critical behaviour of reaction-diffusion processes (cf. [Täuber et al., 2005]).

Note on review and introductory articles: As of 1998 the best review article about the Doi-Peliti formalism, which I will generally refer to as the *coherent state path integral formalism*, is perhaps the one by Mattis and Glasser [1998] as it gives a good account of *the how* and *why* of the formalism. It also contains an excellent bibliography. Another good review is Täuber et al. [2005]. The most recent and complete account of stochastic path integrals that I am aware of is by Weber and Frey [2017], which also contains an excellent bibliography. The article by Wiese [2016] is also noteworthy and provides some unique perspectives. As tutorial introduction, [Cardy, 2006] is often cited. The particular flavour of field theory is characterised by studying the dynamic aspects of critical phenomena [De Dominicis et al., 1975]. Two books which emphasize the *dynamics* are [Vasil'ev, 2004] and [Täuber, 2014] and in a perhaps more general sense, [Zinn-Justin, 1996].

The Poisson ansatz

My choice of the description *coherent state path integral* instead of *Doi-Peliti* is a response to something that is arguably obscured in the literature. In what follows, *coherent states* [Glauber, 1963] can effectively be taken as synonymous with *Poisson distributions*. In particular, one may attempt to use this formalism for arbitrary stochastic processes. However, reaction-diffusion processes have a particular structure for which a basis in terms of *factorial moments* makes sense. As discussed, the factorial moment generating function is intimately related to Poisson statistics and the particular assumptions that have been made about particle sampling is manifest in the choice of infinitesimal generator and formal power series basis. Doi has specifically taken advantage of this. Much of the literature fails to emphasise this important point with some exceptions. Wiese [2016] (§II C) comments on the combinatorics involved in reaction-diffusion processes, which warrant the Poisson ansatz. Weber and Frey [2017] point out that while a Poisson basis is suitable for reaction-diffusion processes, a different basis might be used in a random walk model, for example.

It is instructive to compare Doi's approach with the so-called Poisson representation of Gardiner and Chaturvedi [1977]. In fact this was done by Droz and McKane [1994] where the extension to Doi's method to path integrals by [Peliti, 1985] was compared with the extension of the Gardiner and Chaturvedi [1977] method by Elderfield [1985]. Just like Doi [1976a] does, Gardiner and Chaturvedi [1977] likened their method to the coherent-state-based *Glauber-Sudarshan P-representation* [Glauber, 1963]. They consider the expansion of the probability distribution over configurations as a superposition of Poisson distributions, introducing a "quasiprobability" distribution, that leads to a semigroup. They refer to it as

a *generalised Fokker-Planck equation*.^{||} They make a substitution; they map the probability density P in the master equation to one specifically represented by a Poisson distribution. Recall the generating function $\Psi(t, z) = \sum P_{n,t} z^n$ and use

$$P_{n,t} = \int d\alpha e^{-\alpha} \frac{\alpha^n}{n!} \cdot \hat{g}(\alpha, t) \quad (1.46)$$

which as discussed by Gardiner and Chaturvedi [1977], exists for some tempering *generalised function* $\hat{g}$ i.e. the *quasiprobability*¹². On substituting this into the generating function (or the master equation)

$$\begin{aligned} \Psi(t, z) &= \int d\alpha e^{-\alpha} \frac{1}{n!} (a^\alpha z^n) \cdot \hat{g}(\alpha, t) \\ &= \int d\alpha e^{\alpha(z-1)} \hat{g}(\alpha, t) \end{aligned}$$

This generating function generates factorial moments

$$n_{(k)}(\hat{g}) = \langle n(n-1) \dots (n-k+1) \rangle = \underbrace{\int d\alpha a^k \hat{g}}_{\partial_z^k G|_{z=1}} = \langle \alpha^k \rangle \quad (1.48)$$

The coherent states in the coherent state path integral formalism correspond to the generating functions of Poisson distributions. Coherent states do not have a fixed number of particles but rather the occupation number for some coherent state is Poisson distributed (cf. Negele and Orland [1988] pp. 25).

Statistical moments in coherent state path integral formalism

Coherent states can be defined

$$|\zeta\rangle = e^{\zeta z} |0\rangle = \partial_z |\zeta\rangle = \zeta |\zeta\rangle \quad (1.49)$$

Here ζ relate to factorial moments $n_{(k)}$ rather than simple particle number states and a Poisson density describes coherent states

$$p(n) = e^{-\zeta} \frac{\zeta^n}{n!} \quad (1.50)$$

There is the following known relations between moments using the language of normal ordered operators [Wiese, 2016]],

$$e^{m\hat{n}} = \mathcal{N} \left(e^{(e^m-1)\hat{n}} \right) \quad (1.51)$$

where I introduce the normal ordering operator $\mathcal{N}$ (Taylor expand and shift all z to the left of all ∂_z). Following Wiese [2016] and using that the generating function of *connected* moments is the logarithm of the function,

$$\langle e^{m\hat{n}} \rangle_\zeta^c = (e^m - 1)\zeta \quad (1.52)$$

The p^{th} connected moment of the number operator is expressed in terms of ζ .

^{||}A Fokker-Planck equation which, they emphasise, is different to the Fokker-Planck equation that arises in a *system size expansion* (cf. [Van Kampen, 1992]) as it is derived from microscopic interactions.

Example 1.5.1.

$$\zeta^2 = \langle \hat{n}(\hat{n} - 1) \rangle \quad (1.53)$$

Example 1.5.2.

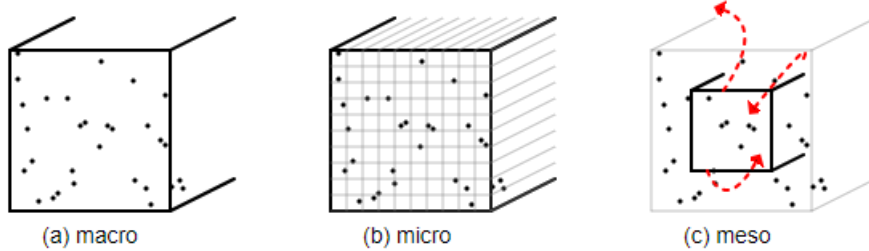
$$\langle \hat{n}^3 \rangle = \zeta(1 + 3\zeta + \zeta^2) \quad (1.54)$$

The relationship between moments was given in (1.25). The mechanics of normal ordered creation and annihilation operators are discussed in the coherent state path integral literature e.g. [Wiese, 2016] and in the combinatorics literature [Blasiak and Flajolet, 2011]. However I have tried to avoid reliance on these mechanics in favour of directly determining the form of the appropriate infinitesimal generator in context.

1.5.2 Some key ideas

I provide a brief overview of the main ideas used in field theories. I begin with some illustrations of basic concepts of path measures, continuum approximations of configurations, and perturbations used in modelling stochastic fluctuations.

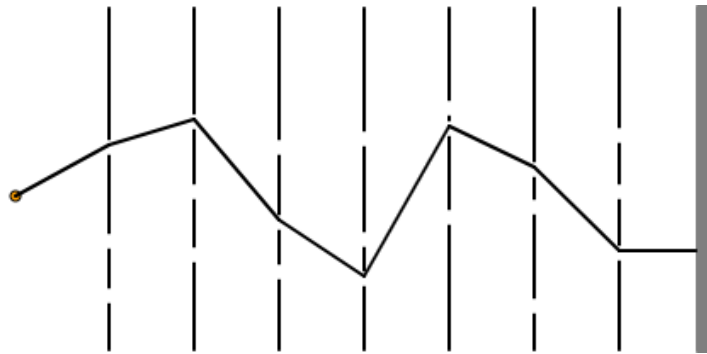
Limits and scales For (spatial) Markov processes one considers different limits and scales. Classically, random processes such as *the simple birth and death process* were studied without the geometry of space. One assumes a well-mixed system of particles as though all particles in the configuration can interact with all others and then determines the time evolution of particle numbers.



Such a *macroscopic* model shown in (a) is a mean-field model and does not take into account the impact of space and diffusion in either limiting or catalysing reactions by diminishing or replenishing particles at reaction sites. At another extreme, the *microscopic* scale (b), one can impose a mesh with d -dimensional volume elements of size a^d where interactions occur only at sites and particles diffuse to and fro. The choice of a influences what happens e.g. if a is very small, particles may never interact in higher dimensions while if a is very large, one returns to the mean-field. Another perspective sits somewhere in the middle and is the *mesoscopic* scale. In the context of birth and death processes, suppose we take a typical mesoscopic box of size b^d and treat the region as a locally well-mixed, mean-field system. Particles in the box may leave and particles outside the box may enter. If particles leave they may return or not return. These situations are described by the red arrows in (c). While this picture may be overly simplistic, it illustrates one of our objectives. If we start with a mean-field solution, we can then try to measure the deviation from the

mean-field when diffusion is taken into account.¹³ Linear diffusion is well understood in this regard and we can effectively combine first-passage probability considerations with mean-field models.

Feynman's path measure The use of path integrals in quantum field theory goes back to Feynman in the early 1940s building on ideas initially introduced by Dirac.¹⁴ Feynman exploits the semigroup property of the time evolution operator (fundamental composition Law) to slice the path into equal times slices. In the continuum limit the path integral is a product of integrals. Phenomenologically, Feynman considers the amplitudes or relative contributions of all possible paths that a particle might take between two spacetime points. In the continuum limit, aggregated paths converge to a Lagrangian system under a least action principle. In the first few chapters of Feynman's lectures [Feynman et al., 1965] the basic ideas of path measures are discussed. In particular, the following *Young-slits interference experiment is introduced*.



Particles emitted from a point source are detected on a surface having passed through slits in an intermediate screen(s). Possible trajectories are weighted by amplitudes for which the total probability of observing a particle at the sink is a sum over amplitudes. This is reminiscent of the Chapman-Kolmogorov conditions and indeed will lead to an evolutionary operator which has a semigroup property. This allows Feynman to consider slicing time into small intervals and then perform an integration over paths. One introduces the notion of a *path measure* to consider not just the mean trajectory but all possible paths of stochastic events (in principle). The so-called classical path is the one that makes the action stationary and a set of configurations traversed by this path is sometimes referred to as the *Landau configuration*. What we wish to study in fact, are fluctuations away from this configuration which motivates the use of perturbation theory. One postulate of critical phenomena is that the transition from a mean-field to a model incorporating fluctuations makes sense [Vasil'ev, 2004]. When considering fluctuations, a free Gaussian field can be used to model deviations from the mean field. In the context of Markov processes in field theory, Dynkin [1983] observes that Symanzik [1969] noted a *Euclidean field theory*¹⁵ can be formulated as a "gas" of Brownian paths.

I do not discuss the derivation of path integrals in great detail in this chapter as they are discussed in any textbook in quantum field theory such as [Itzykson and Zuber, 1980] (ch. 9), Le Bellac [1991] (ch. 8), Kleinert [2009] (ch. 2) or Zinn-Justin [1996].

Events and correlations in evolutionary systems A nice illustration of the interpretation of perturbative expansions in *Feynman diagrams*** relating to reaction-diffusion processes is given by Wiese [2016]. Consider the probability $p(t = \tau)$ of two particles meeting at times τ .

$$p(t = 2) \Rightarrow \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \begin{array}{c} | \\ | \\ | \\ | \\ | \\ | \\ | \end{array} \begin{array}{c} t \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \quad p(t = 3) \Rightarrow \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \begin{array}{c} | \\ | \\ | \\ | \\ | \\ | \\ | \end{array} \begin{array}{c} t \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \quad (1.55)$$

Furthermore one can consider the probability of particles meeting, dissipating and then meeting again at a later time

$$p(t = [2, 4]) \Rightarrow \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \begin{array}{c} | \\ | \\ | \\ | \\ | \\ | \\ | \end{array} \begin{array}{c} t \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \quad (1.56)$$

Now consider the *first passage*¹⁶ or first meeting probability $p_1(t = \tau)$. The probability of meeting *for the first time* at time τ is the probability of meeting at time τ minus the sum of probabilities of also meeting at some earlier time (with a double-counting correction).

$$p_1(t = 3) \Rightarrow \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \begin{array}{c} | \\ | \\ | \\ | \\ | \\ | \\ | \end{array} \begin{array}{c} t \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} - \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \begin{array}{c} | \\ | \\ | \\ | \\ | \\ | \\ | \end{array} \begin{array}{c} t \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} - \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \begin{array}{c} | \\ | \\ | \\ | \\ | \\ | \\ | \end{array} \begin{array}{c} t \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} + \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \begin{array}{c} | \\ | \\ | \\ | \\ | \\ | \\ | \end{array} \begin{array}{c} t \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \quad (1.57)$$

The Feynman diagrams in the theory correspond to these sorts of events. At *tree level*,^{††} coagulation and quartic vertices describe different "scattering" events. In the coagulation process discussed earlier in this chapter, was the generator

$$\mathcal{Q}_\chi[\partial_z, z] = \chi \left((z + 1)^1 - (z + 1)^2 \right) \partial_z^2 = \chi z \partial_z^2 + \chi z^2 \partial_z^2 \quad (1.58)$$

The last term is the quartic vertex below^{††}

$$z^2 \partial_z^2 := \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} \quad (1.59)$$

which Wiese [2016] notes is mapping a problem of two particles meeting at time τ to the problem of two particles meeting *for the first time* at τ .

Perturbations, the main idea: To understand the general structure, take a sample field theory. Introduce the Lagrangian density for a field¹⁷

$$\mathcal{L}[x, \phi(x), \partial_x \phi(x)] = \frac{1}{2} [\phi^* \Delta \phi] - J(x) \phi(x) - \frac{\lambda_1}{3!} \phi(x)^3 - \frac{\lambda_2}{4!} \phi(x)^4 \quad (1.60)$$

Many of the specific data are not the point of this example. Rather note the source or external field term $J(x)$ and monomials which are powers in the field with couplings λ_i .

**Unlike actual Feynman diagrams, time is ordered from left to right in these illustrations

††Diagrams without loops

††In diagram convention, unlike the first passage illustrations, time is ordered from right to left

Here Δ is in fact a placeholder for a general differential operator and may not be the Laplacian operator. The *Euler-Lagrange equations* can be written as the relation¹⁸

$$\Delta\phi(x) = J(x) + \frac{\lambda_1}{2!}\phi(x)^2 + \frac{\lambda_2}{3!}\phi(x)^3 \quad (1.61)$$

This is in general a non-linear differential equation that cannot be solved. But if the coupling constants λ_i are small, it can be solved perturbatively, regarding interacting terms as perturbations of free terms. A recursive relation can be written

$$\phi(x') = \int_{\mathbb{R}^d} d^d x' G_0(x' - x) \left[J(x) + \frac{\lambda_1}{2!}\phi(x)^2 + \frac{\lambda_2}{3!}\phi(x)^3 \right] \quad (1.62)$$

where G_0 is the free propagator or *Green's function* corresponding to the operator Δ . This propagator was encountered in the section on diffusion above. Solutions can be expressed as formal power series in powers of the coupling constant

$$\phi(x) = \sum_0^\infty \lambda^n \phi_n(x) \quad (1.63)$$

It is straight-forward but tedious to carry out the exercise of iteratively substituting the value of the function in place of $\phi(x)$ up to some $O(\lambda^p)$. Managing this tedium is precisely what Feynman diagrams are for. The terms $\phi_n(x)$ are finite sums of Feynman integrals.

We are generally interested in statistical moments referred to as the n -point Green's functions $G_n(\mathbf{q}) = \langle \phi(\mathbf{q}_1) \dots \phi(\mathbf{q}_n) \rangle$. A functional integral in principle extends (1.63) to n -point functions and the solution again admits a power series representation

$$G(x) = \sum_0^\infty \lambda^n G_n(x) \quad (1.64)$$

In the context of reaction-diffusion processes, these Green's functions describe certain observables. In an m -species *second quantized* process, there are $2m$ fields rather than the one field ϕ in the Lagrangian density (1.60) yet the principle is the same. In the previous section, the calculation of observables via the path integral was described as a "sifting" convolution (1.71), a series expansion in operators/diagrams. In perturbative quantum field theory the formal path integral is implemented by constructing a perturbative expansion in operators/diagrams acting on an appropriate basis such as one describing coherent states. While in principle an infinite expansion in diagrams might be considered, in practice it turns out that only the first few terms of the expansion are required to match theory with experiment. That this makes sense i.e. that going to higher orders does not qualitatively change the properties of the solution but only increases accuracy, is another postulate of critical phenomena [Vasil'ev, 2004]. In high energy physics, perturbative methods have been hugely successful in providing analytical results in agreement with experiment to high accuracy. In critical phenomena, this postulate is typically validated by matching analytical results with numerical simulation.¹⁹

A (very) high-level overview of field theory procedures

In this thesis, ideas originating in field theory are applied to classical stochastic processes. The fundamentals can be studied as either *quantum* or *statistical* field theory, where the

latter is formally regarded as a continuation of the former to imaginary time. There are different approaches to field theory - I consider only the *perturbative* kind.

Starting with the semigroup of the stochastic process, a *path integral* or *generating functional* is constructed and a *perturbative expansion* in Feynman diagrams is used to evaluate the generally non-linear part of the functional. At an overview level of quantum field theory, the following mappings are carried out;

$$S \xrightarrow{\mathbb{P}} e^{\int S} \xrightarrow{\mathbb{F}} Z \xrightarrow{\log} W \xrightarrow{\mathbb{L}} \Gamma \quad (1.65)$$

(i) The first maps the action S onto a path integral²⁰ (ii) The Fourier transform of the path integral gives the partition function, which is a sum over graphs. (iii) The logarithm of the partition function gives the normalized *free energy* W , which is the set of *connected* graphs only (iv) The Legendre transform produces the effective action Γ . In the QFT literature this is formally developed via functional methods [Zinn-Justin, 1996].

Remark 12. A perspective to perturbative quantum field theory that puts less emphasis on the functional aspects, is the combinational perspective as discussed for example by Jackson et al. [2018] or Borinsky [2018]. It would appear that all stages of the mapping, from the representation of the probability in terms of its generating function to the *renormalisation group functions*, can be treated in terms of formal power series. The entire game might be understood as finding suitable transformations between formal power series; a finite set of coefficients of the formal power series preserve the important data relating to the physics, regardless of whether or not the series is a convergent one.

In stochastic reaction-diffusion processes such as the superprocess discussed in Chapter 3, an *offspring mechanism* generates a population of interacting particles and we would like to study the time evolution of configurations and their correlation functions. The path integral in this context is called the *coherent state path integral*. Standard derivations of the coherent state path integral formalism applied to reaction-diffusion can be found for example in the original work of Peliti [1985], [Cardy, 2006] or [Wiese, 2016] and I recommend the exposition by Weber and Frey [2017]. In a more general context outside of the application to reaction-diffusion many texts on quantum field theory such as [Negele and Orland, 1988] discuss the path integral derivation. The path integral can be written as

$$\int \mathcal{D}\phi \mathcal{D}\tilde{\phi} e^{-S[\tilde{\phi}, \phi]} \Big|_{\phi_\alpha}^{\phi_\beta} \quad (1.66)$$

where the action S is

$$S[\tilde{\phi}, \phi] = \int_t^{t'} dt \tilde{\phi}(\partial_t + \Delta)\phi - \mathcal{Q}[\tilde{\phi}, \phi] \quad (1.67)$$

All the fields depend on t which I omit. Terms ϕ_α, ϕ_β symbolize boundary conditions to allow a definition in terms of a transition amplitude $\mathcal{A}(\phi_\alpha | \phi_\beta)$ whatever that means for a specific problem. Analogous to the forward and backward Kolmogorov equations both forward and backward time versions can be derived [Weber and Frey, 2017]. Above, $\mathcal{Q}$ is the sum of infinitesimal generators for elementary reactions in a chemical reaction network, each taking the form (1.2) and the Lagrangian is constructed from this sum as given by (1.3). This Lagrangian plays the role of the action (1.67) and enters the path integral (1.66), which was introduced symbolically as (1.5) in the overview at the start of the chapter.

Remark 13. As discussed in the section on *extended semigroups*, after making use of the Lie-Kato-Trotter formula to approximate the generator of a semigroup $\mathcal{Q}$ and breaking the trajectory into time chunks, one aims to show that the mapping to a Hilbert space makes sense, which requires demonstrating *resolution of identity*. The operator algebra generates a perturbation series which is represented in diagrams. In the formal derivation of the generating functional this is motivated by a *Wick expansion* of correlation functions [Negele and Orland, 1988]. Alternatively, this can be understood through the representation of formal power series, here a formal power series in the *species* of Feynman graphs. In the case of evolutionary Markov processes, the diagrams are very intuitive with respect to causal structure of actual stochastic events as illustrated in Figure 1.57.

In reaction-diffusion problems, we are interested in various statistical expectations i.e. *observables*. Observables are accessible via their generating functions,

$$\langle a \rangle = \sum_{n=0}^{\infty} a(n) p(t, n) \quad (1.68)$$

In the Poisson representation (i.e. in coherent state path integral) one has in fact a different observables. I define this,

$$\langle\langle a \rangle\rangle = \sum_{n=0}^{\infty} a(n) \frac{x^n e^{-x}}{n!} \quad (1.69)$$

The path integral representation of the observable can be written

$$\int \mathcal{D}\phi \mathcal{D}\tilde{\phi} e^{-S[\tilde{\phi}, \phi]} \langle\langle a \rangle\rangle \quad (1.70)$$

I use a short-hand

$$\langle \mathcal{O} \rangle = \langle \mathcal{O}, e^S \rangle \quad (1.71)$$

to describe a "convolution" with an observable and an operator expansion. This can be seen as a type of sifting operation where terms in the expansion that contribute to an observable are retained. This is best understood with the use of diagrams (Chapter 2).

Summary

The key observation in this chapter is that in principle, for an arbitrary chemical reaction network living on a graph (a compound Poisson process with reactions at sites on a hypercubic lattice and particle transport by linear diffusion), the semigroup can be formally written down by formulae. The generator of the semigroup is the infinitesimal generator acting on the generating function of the factorial moments, a formal power series. Using the factorial moment generator results in, what is referred to in the literature as, the *Doi-shifted* action. The generator of the semigroup is the Lagrangian density for the process. The semigroup cannot (in general) be solved so we apply perturbative methods through Feynman diagrams. The interaction part of the Lagrangian is simply a linear combination of the infinitesimal generators corresponding to each reaction in the compound process. The non-linear interaction part will be dealt with perturbatively in an expansion in diagrams. Whether or not the field theory can then be tackled meaningfully for an arbitrary process is another matter.

I have avoided describing in detail, certain formal, functional aspects of quantum field theory. Quantum field theory boasts sophisticated methods and a rich history and any adequate coverage would take us too far astray without (arguably) adding new insight in the given context. I attempted to take essential ingredients, remove the quantum field theory ballast and re-express the ideas using the basic tools of formal power series and their semigroups. The duality between an operator algebra and a complete set of states can be expressed in the language of *Umbral calculus*. In that picture, *series of binomial type* (cf. C.10) admit a semigroup property. Suitable formal power series represent a complete set of states and are the moment generating functions. For any such series, a dual operator algebra $\mathfrak{g}$ obeying the commutation rule is sought. A perturbative (operator) expansion for a given basis is generated by $\mathfrak{g}$. This expansion is represented diagrammatically in Chapter 2.

2 Graphs

In this chapter I discuss the role of graphs in two separate applications. The contributions of this chapter are algorithmic methods for problems in physics, which also provide insight on inherent structure in the underlying problems.

In the first case, a bijection between *spanning trees* and *recurrent configurations* of the *Abelian sandpile model* of self-organised criticality is exploited for efficient Monte Carlo sampling. Here I discuss a new algorithm that builds on the so-called *burning test* [Majumdar and Dhar, 1992] for checking if a configuration is *recurrent*. In the algorithm developed and evaluated in [Amarteifio and Pruessner, 2019], the inverse mapping from a *spanning tree* is used to produce recurrent configurations to seed Monte Carlo sampling of *avalanche* statistics.

In the second case, I discuss aspects of graph theory useful in working with *Feynman diagrams*. I discuss *algorithmically* generating Feynman diagrams directly from the chemical reaction network of Chapter 1. I also discuss the *parametric integral representation* expressed in terms of *graph polynomials*. In certain cases such as for one-loop Feynman diagrams, the *momentum integral representation* might be preferred and is the method most typically discussed in textbooks. Graph polynomials and the parametric integral representation have been exploited for symbolic and algorithmic methods [Heinrich, 2008, Panzer, 2015] and to better understand the mathematical structure of Feynman graphs [Speer and Westwater, 1971, Nakanishi, 1971, Brown, 2009]. Yet these methods have not been used in applications to *critical dynamics* and non-equilibrium field theories.²¹

The chapter ends with an overview of *renormalization* which, as will be discussed, constitutes a necessary "post processing" of Feynman integrals. This last section leads naturally to the topics discussed in Chapter 3, where renormalized Feynman integrals are used to determine critical exponents for a reaction-diffusion process.

2.1 Terminology

I introduce some notation and vocabulary with examples. This section serves to define objects that are used in applications in the remaining sections. The terminology is standard and can be found in introductory textbooks on graph theory such as [Harris et al., 2009] or [Diestel, 2010]. With particular relevance to Feynman diagrams, [Nakanishi, 1971] offers unique insights.

A graph $G(V, E)$ consists of a set of vertices V and a set of edges E . We say a vertex is *incident* to an edge if the edge is connected to the vertex. If two vertices are connected by

an edge they are *adjacent*. A graph may be *directed* or *undirected*. In a directed graph we distinguish edges exiting or entering a vertex. A directed edge is sometimes called an *arc*. The number of edges connected to a vertex is its *degree* or *valence*. In physics, the degree of vertices (sites) on a regular hypercubic lattice is also called the *coordination number*. In an undirected graph we may distinguish *in-degree* or *out-degree*. Strictly, an edge must be incident to two vertices. However, in physics applications one often finds *unactualized graphs*, graphs for which one or more edges are connected to only one vertex. Such edges are not true edges and are referred to as *half-edges*, *stubs* or *flags*. While they are useful animals in physics, unactualized graphs are not permitted for many graph operations and this is remedied by connecting all half-edges to a newly introduced vertex (or multiple vertices). The convention here will be to designate a single new vertex as v_∞ and connect all half-edges to it. I will refer to this map as the *actualization*^{*} of the graph. The inverse map is the *amputation* of v_∞ . In the particular case where an unactualized graph has only one vertex it is called an *amputated vertex* or a *corolla*. Given a vertex, the vertex plus the set of edges incident to it is called a *star*. The *cover time* of a graph is the expected time needed for a random walk on a finite graph to visit every vertex at least once.

A path $\pi(s, t)$ is a set of edges connecting vertices s and t . If $s = t$ the path is called a *circuit* and if the circuit does not repeat vertices it is called a *cycle*. A graph is formed of a number of connected components $|\chi| \geq 1$. A connected component $\chi \subseteq G$ is defined such that any pair of vertices $v_1, v_2 \in \chi$ are connected by a path $\pi(v_1, v_2)$. The first Betti number L gives the number of independent cycles in the graph;

$$L = |E| - |V| + |\chi| \quad (2.1)$$

In graph theory a *loop* is a self-edge i.e. an edge connecting a vertex to itself. But in our applications, we shall use *loop* synonymously with *cyclic circuits*, hence the use of L above. In any case, we will seldom encounter loops in the graph theoretic sense, except during some *contraction* operations including cases where we produce the so called *tadpole* diagrams.

If a graph contains no cycles it is called *acyclic*. An acyclic graph is called a *forest* or a *k-tree* where k denotes the number of components. If a forest has one component we refer it as a *1-tree* or simply a *tree*. The *leaves* of a tree are its 1-degree vertices. For any tree, we can pick a distinguished vertex called the *root* to which any other vertex is connected by a unique path. If we wish to emphasise the directed edges along these paths culminating at the root, such a tree is referred to as an *arborescence*. A *bridge* is any edge whose removal *disconnects* the graph i.e. increases $|\chi|$. This is relevant when discussing the so-called *one particle irreducible* diagrams i.e. Feynman diagrams that do not contain a bridge. The deletion of an edge from a set of edges is denoted $E \setminus e$ and the contraction of an edge is denoted $E // e$. Some authors use (respectively) $E - e$ and E / e . A contraction *identifies* vertices at the end point of an edge as shown in Figure 2.1.

An isomorphism between two graphs is a bijection between their vertex sets that preserves adjacency. An *isometry class* is the set of all graphs that are isomorphic to each other modulo internal symmetries. In some instances we may be interested in the size of this class, which can be described as a *symmetry factor* $\mathfrak{s}(G)$. If two graphs are isomorphic, the relationship will be written $G \sim F$.

^{*}The word *actualization* is more cumbersome than say, "completion", however *complete graphs* have another meaning in graph theory, which might cause confusion.

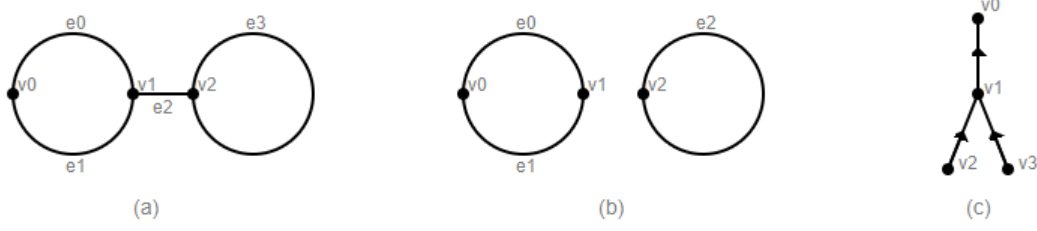


Figure 2.1: *Sample graphs and notation:* (a) An undirected graph G_a with $|\chi| = 1, |E| = 5, |V| = 3, L = 2$. $\{e_0, e_1\}$ is an independent circuit and $\{e_3\}$ is a self-loop. v_1 has degree 3 and is incident to $\{e_0, e_1, e_2\}$. (b) A graph G_b with $|\chi| = 2$ formed from the operation $G_a \setminus \{e_2\}$ (and a minor edge relabelling). Of the set of component graphs $\{g_1, g_2\} \in G_b$, $g_2 \sim g_1 // \{e_1\}$. (c) G_c is an arborescence, a directed acyclic graph rooted at v_0 . The vertex v_1 has in-degree 2 and out-degree 1. $\{v_2, v_3\}$ is the set of leaves in G_c . The graph G_a contains a bridge e_2 whose removal would increase $|\chi|$ (the number of components) and is therefore not *one particle irreducible* (1PI).

2.2 Representations

The matrices used to represent graphs will always appear **boldface**. They will be indexed by subscripts. For example $\mathbf{E}_{ij}$ denotes the matrix cell that is the i^{th} column and j^{th} row. The convention in this thesis is to zero-index. Later an indexing notation will be introduced to describe a restricted graph i.e. deletions of rows and or columns. In this case a bracket subscript will be introduced. For example $\mathbf{E}_{[ij]}$ is the matrix produced by *deleting* the i^{th} column and j^{th} row. In fact we can delete entire sets of rows and or columns with the notation $\mathbf{E}_{[IJ]}$ where the upper case specifically denotes row or column *sets*.

Definition 5. The *adjacency matrix* $\mathbf{A}$ of size $|V| \times |V|$ is

$$\mathbf{A}_{ij} = \begin{cases} \text{degree}(v_i) & \text{if } i = j \\ -1 & \text{if } v_i \text{ adjacent to } v_j \\ 0 & \text{otherwise} \end{cases} \quad (2.2)$$

Definition 6. The *diagonal matrix* $\mathbf{D}$ of size $|V| \times |V|$ is given by

$$\mathbf{D}_{ij} = \begin{cases} \text{degree}(i) & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

Definition 7. The *incidence matrix* is a matrix $\mathbf{E}$ of size $|E| \times |V|$, that is, by convention I place edges on columns and vertices on rows.

$$\mathbf{E}_{ev} = \begin{cases} -1 & \text{if } e \text{ exits } v \\ 1 & \text{if } e \text{ enters } v \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

Example 2.2.1. This example shows the incidence matrix for a Feynman graph called *the sunset*. All three edges enter v_0 as shown by the signs in the first row of $\mathbf{E}$.

$$G = \begin{array}{c} \text{v}_0 \text{---} \text{v}_1 \\ \text{e}_1 \quad \text{e}_2 \quad \text{e}_3 \end{array}, \quad \mathbf{E} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix} \quad (2.5)$$

Definition 8. The *Laplacian matrix* is $\mathbf{L} = \mathbf{D} - \mathbf{A} = \mathbf{E}\mathbf{E}^T$

On graph residues and graph actualization

The set of external edges $R = \mathfrak{R}(G)$ will be referred to as the *residue* of the graph. In each of the applications discussed in this chapter, the graphs need to be *actualized*. The Feynman graphs we will care about are expected to have no bridges.

Feynman diagrams are represented such that their external vertices are half-edges. By convention, the external half-edges are connected to the vertex v_∞ in order to actualize the graph.

Example 2.2.2. Considering diagrams encountered in reaction-diffusions, we have already seen that monomials in the Heisenberg-Weyl algebra are depicted diagrammatically.

$$z^n \partial_z^m = n \times m \quad (2.6)$$

The graph is actualized by connecting the external edges to the vertex v_∞ . This tree-level or "vertex diagram" can be represented by an incidence matrix using the conventions already discussed. Following conventions, the residue edges q_i are connected to v_∞ , which always features as the last row in the incidence matrix.

$$\begin{array}{c} v_\infty \bullet \\ \vdots \\ \text{---} \bullet \end{array} \Rightarrow \begin{bmatrix} q_0 & -q_1 & -q_2 \\ -q_0 & q_1 & q_2 \end{bmatrix} \quad (2.7)$$

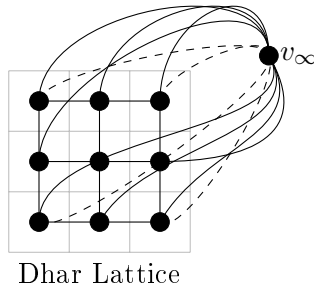
In this example $R = \{q_0, q_1, q_2\} = E$

Example 2.2.3. The one-loop graph below has a different overall structure but the same residue as in the previous example

$$\begin{array}{c} v_\infty \bullet \\ \vdots \\ \text{---} \bullet \end{array} \Rightarrow \begin{bmatrix} p_0 & p_1 & 0 & -q_1 & -q_2 \\ -p_0 & -p_1 & q_0 & 0 & 0 \\ 0 & 0 & -q_0 & q_1 & q_2 \end{bmatrix} \quad (2.8)$$

Remark 14. In perturbative expansions it will be necessary to group diagrams according to their residue.

Example 2.2.4. In the case of lattice simulations on open boundaries, the open boundary is another example where the half-edges need to be connected to v_∞ . This 2D lattice has coordination number $q = 4$. The vertices at the corners are connected to v_∞ by two edges.

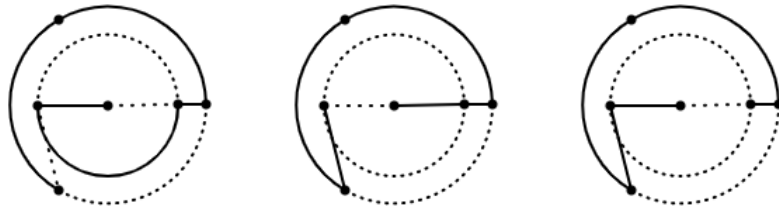


2.3 Spanning trees and graph polynomials

Definition 9. Given a graph $G(V, E)$, a *spanning tree* $\Gamma(G) \subset G$ is an acyclic graph that contains all the vertices of G .

Remark 15. A spanning tree is a minimally (simply) connected graph. We delete edges in any cycles until there are no cycles, leaving the graph minimally connected.

Example 2.3.1. The example below shows a graph that contains 4 independent cycles. Three different spanning trees (of a total of 46 for this graph) are shown, each of which deletes any one edge in each independent cycle. Deleted edges are represented as dotted lines. $4 = L$ edges have been removed such that the graphs are minimally connected.



Adding back any of the edges marked by dotted lines creates a cycle and the resulting graph is not a spanning tree •

In order to systematically enumerate (or generate) spanning k-trees there is one very important theorem due to Kirchhoff [1847]. First some terminology. Given incidence matrix $\mathbf{E}$ of size $|E| \times |V|$, the *sub matrix* $\mathbf{E}_{[ij]}$ is the matrix formed by deleting single row $i \in V$ and single column $j \in E$ from G .

Given incidence matrix $\mathbf{E}$ of size $|E| \times |V|$, the *sub matrix* $\mathbf{E}_{[IJ]}$ is the matrix formed by deleting a row subset $I \subset V$ and a column subset $J \subset E$ from G .

Remark 16. For square subgraph $\mathbf{E}_{[IJ]}$, $|E \setminus J| = |V \setminus I|$

Occasionally definitions require us to delete only the v^{th} row or the e^{th} column in which case the notation $E_{[v]}$ or $E_{[e]}$ is required.

Definition 10. The IJ -cofactor of $\mathbf{E}$ is $(-1)^{I+J} \det \mathbf{E}_{[IJ]}$

The *Cauchy-Binet formula* which is useful for noting variations in definitions appears below. Let $\hat{\mathbf{E}} = \mathbf{E}_{[v]}$ and S be permutations of index sets of sizes $|V| - 1$ then

$$\det \hat{\mathbf{L}} = \mathbf{L}_{[EV]} = \sum_S \det \hat{\mathbf{E}}_S \det \hat{\mathbf{E}}_S^T = \sum_S \det(\hat{\mathbf{E}}_S)^2 \quad (2.9)$$

This relates the determinant of the *reduced Laplacian* on the left hand side with the determinant of square of the reduced incidence matrix on the right hand side.

Finally we come to the matrix tree theorems. The most general statement, generalising Kirchhoff's matrix tree theorem which follows as a corollary, is known as the *all minors Matrix-tree theorem* (cf. Chaiken [1982]).

Theorem 2.1. (*All minors matrix-tree theorem*) For square matrices $\mathbf{E}_{[IJ]}$,

$$\det(\mathbf{E}_{[IJ]}) = \begin{cases} \pm 1 & \text{for the } |\chi| \text{-tree } E \setminus I \text{ with } |\chi| \text{ connected components,} \\ & \text{each tree having one vertex in component } \chi. \\ 0 & \text{otherwise} \end{cases} \quad (2.10)$$

Proof. The proof is given in (§C.7) □

I state one specific case as a corollary, *Kirchhoff's matrix tree theorem* for spanning 1-trees:

Corollary 2.1.1. Any ij -cofactor, $\det \mathbf{L}_{[ij]}$ of the *Laplacian* matrix of the graph G , enumerates the spanning trees (i.e. 1-trees) of G .

Remark 17. The proof of this particular case is discussed for example in [Harris et al., 2009] (pp. 48). I give explicit examples below, which will hopefully be equally illuminating.

Example 2.3.2. In this example I take a graph based on the sunset diagram given in the last example. Attach the external stubs that could be attached to each of the external vertices v_0, v_1 in the true Feynman graph (as will be discussed later in this chapter). Connect the stubs to v_∞ thus *actualizing the graph*. Internal edges will be denoted p_i and q_j will denote external edges. By convention, the last row of the incidence matrix will always correspond to the vertex v_∞ . The *symbolic*[†] incidence matrix for this example, with symbols corresponding to edge labels is

$$\mathbf{E} = \begin{bmatrix} p_0 & p_1 & p_2 & q_0 & 0 \\ -p_0 & -p_1 & -p_2 & 0 & q_1 \\ 0 & 0 & 0 & -q_0 & -q_1 \end{bmatrix}$$

The Laplacian matrix can be computed as $\mathbf{E}\mathbf{E}^T$. For reasons that will become clear I change variables for external edges $q_i \rightarrow z$. These edges will be severed in the cofactor. Thus we have for the *symbolic* Laplacian matrix and its ij -cofactor,

$$\mathbf{L} = \begin{bmatrix} p_0 + p_1 + p_2 + z & -p_0 - p_1 - p_2 & -z \\ -p_0 - p_1 - p_2 & p_0 + p_1 + p_2 + z & -z \\ -z & -z & 2z \end{bmatrix}$$

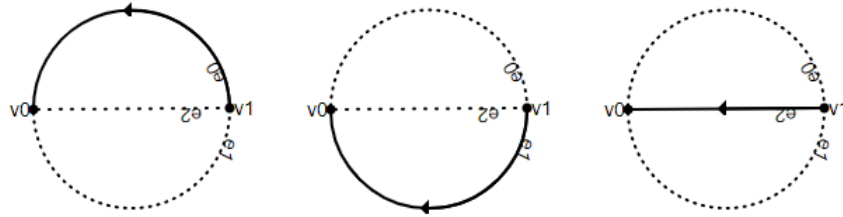
$$\mathbf{L}_{[2,2]} = \begin{bmatrix} p_0 + p_1 + p_2 + z & -p_0 - p_1 - p_2 \\ -p_0 - p_1 - p_2 & p_0 + p_1 + p_2 + z \end{bmatrix}$$

The symbolic determinant of $\mathbf{L}_{[2,2]}$ gives a polynomial in z called the *Kirchhoff polynomial*. Terms of order k in the polynomial give the k -trees!

$$\det \mathbf{L}_{[2,2]} = z^2 + z(p_0 + p_1 + p_2) \quad (2.12)$$

In this case the 2-tree which consists of two isolated vertices, contains no edges. The three spanning 1-trees are denoted by one of the edges $p_0, p_1, \dots$ remaining in that spanning tree. These spanning trees are shown below.

[†]A *symbolic* matrix is simply the matrix in $\mathbb{Z}$ multiplied by a suitable diagonal matrix which attributes symbolic labels e.g. $\mathbf{L}_{sym} = \mathbf{E}\mathbf{D}_e\mathbf{E}^T$ where $\mathbf{D}_e = \text{diag}(p_1, p_2, \dots, q_1, q_2)$



Randomly sampling spanning trees

The matrix tree theorem is one way to enumerate spanning trees of graphs. When graphs are small this method is transparent and convenient. Later in this chapter, various methods for working with Feynman graphs will focus on graph polynomials and the matrix tree theorem. Up until about 1990, algorithms used to generate spanning trees were based on the matrix tree theorem [Kelner and Madry, 2009]. Then a number of stochastic methods emerged. Two noteworthy and fast methods for generating spanning trees are a method based on random walks discovered independently by Aldous [1990] and Broder [1989] and a method based on *loop erased random walks* discovered by Wilson [1996]. In the next section we learn about an interesting bijection between the *recurrent configurations* of the *Abelian sandpile model* and spanning trees. In order to exploit this bijection in an efficient Monte Carlo sampling algorithm, it is important to have a fast method for uniformly sampling spanning trees on large graphs. Wilson's algorithm is the method of choice and I introduce it here.

Wilson's algorithm Wilson [1996] gives an algorithm for generating uniform spanning trees on arbitrary directed or undirected graphs more quickly than the cover time.

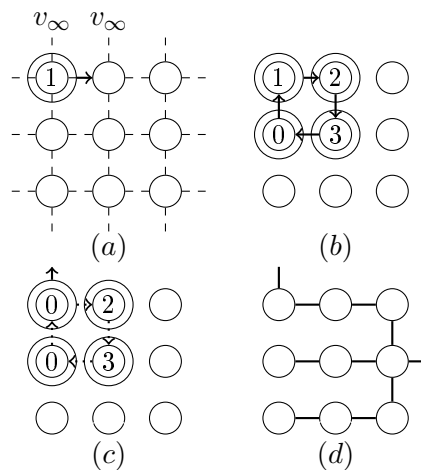


Figure 2.2: *Wilson's algorithm* performs a random walk on the graph from each site, recording the most recent exit edge at each vertex (a). The numbers on each vertex correspond to edge directions. Loops are erased as the random walk re-exits sites (c). Random walks terminate when they hit the tree, which is initially a single chosen root vertex. A spanning tree is iteratively generated in this way (d).

The Wilson algorithm is illustrated in Figure 2.2. The underlying graph G is shown connected at the borders to the vertex v_∞ for the *Dhar lattice* (Example 2.2.4). In general any vertex can be chosen as the root to bootstrap Wilson's algorithm. However in this implementation, v_∞ is always chosen as the root of the undirected spanning tree $\Gamma(G)$. While this would produce a bias when sampling *directed* spanning trees, here the *arborescence* rooted at v_∞ is only used for convenience to designate a convenient and fixed starting point for the breadth first traverse in the reverse mapping. Having designated a root vertex to represent the initial tree $\Gamma(G)_0$, the Wilson algorithm, starting from lattice site $v_0 \in G \setminus v_\infty$, performs a loop-erased random walk to connect v_0 to $\Gamma(G)$. The numbers shown on vertices in Figure 2.2 correspond to edge choices stored on a stack e.g. $0 \leftarrow N, 1 \leftarrow E, 2 \leftarrow S, 3 \leftarrow W$. In (b) a loop has been created. However in (c) when a new random choice is made at v_0 , the stack value is replaced and the loop erased. The dotted arrows show the loop erasure. The edge exiting v_0 becomes connected to v_∞ and hence $\Gamma(G)$. As such the edge set $\mathcal{V}_0$ of vertices in the loop-erased random walk started at v_0 are added to $\Gamma(G)$. The spanning tree shown in (d) is completed in this way, as the edge chain $\mathcal{V}_i$ from each vertex $v_i \notin \Gamma(G)$ connects to a vertex $v_j \in \Gamma(G)$. In practice, as shown in Algorithm 2 (§A), a second while loop requires a retracing walk from v_i to $\Gamma(G)$ to "commit" the path to the tree. For this algorithm, $\Gamma(G)$ grows slowly at first then quickly.

2.4 Abelian sandpiles: a useful bijection

The *Abelian sandpile model's* *raison d'être* is discussed in the next chapter. For the moment I introduce it as a dynamical system on a lattice without any motivation.

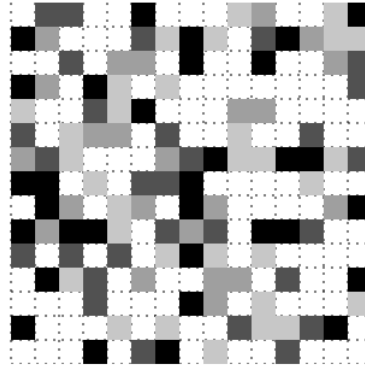


Figure 2.3: *ASM dynamics:* Grey-scale cells denote the height in the sandpile. The darkest cell i.e. black has the highest value $q = 4$. If a new grain of sand is deposited on one of the black cells, there will be a toppling event as particles are redistributed to nearest neighbours. If the nearest neighbour happens to be black, there will be a second toppling event and so on...

Consider a $2D$ lattice with side length L and where $q = 4$ is the coordination number. Starting from an initial (empty) d -dimensional lattice, a *random* lattice site i is *driven* (alias *charged*) by dropping one particle so as to increase its height $z_i \rightarrow z_i + 1$. While $z_i < q$ the configuration remains *stable* otherwise a toppling occurs, distributing particles

to neighbouring sites according to $z_j \rightarrow z_j - \Delta_{ij}$ with the discrete lattice Laplacian

$$\Delta_{ij} = \begin{cases} q & i = j \\ -1 & |i - j| = 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.13)$$

Remark 18. The Abelian Sandpile Model is driven by depositing particles, which may induce an avalanche. The Abelian property means the final state after topping is independent of the order in which sites are driven.

The dynamics are inherently *deterministic*. By driving the system from random sites, randomness is introduced for Markov Chain Monte Carlo sampling and estimation of avalanche statistics. The configurations (i.e. lattice states) into which the ASM equilibrates between avalanches may be *transient* or *recurrent*. Transient configurations ψ are passed through once during equilibration. Recurrent configurations ϕ on the other hand, once entered will be revisited surely within a finite time (assuming the system runs for long enough). The configurations the system passes through on return (called an *orbit*) are necessarily recurrent

$$\underbrace{\psi_1 \rightarrow \psi_2 \rightarrow \dots \rightarrow \psi_p}_{\text{equilibration}} \rightarrow \overbrace{\phi_1 \rightarrow \phi_2 \rightarrow \dots \rightarrow \phi_n \rightarrow \dots \rightarrow \phi_1}^{\text{recurrent orbit}}$$

From any recurrent configuration ϕ there are a set of orbits $\{\omega_i(\phi) \in \Omega(\phi)\}$ returning to ϕ . The recurrent epoch is ergodic in that any ϕ_i can be reached by any ϕ_j in a finite number of charges. [Majumdar and Dhar, 1992] introduced a bijection between *recurrent configurations* of sandpiles and spanning trees. Given any configuration a so-called *burning test* can be applied. If having carried out the test a spanning tree is produced, then the configuration is *recurrent*.

2.4.1 Recurrent configurations to spanning trees

For a given candidate configuration, the burning test proceeds in iterations, comparing each lattice site v_i of height $z(i)$ with $\bar{\mathcal{B}}(i)$, the number of (nearest) neighbours of site i *not* already *burnt* at some time t . I use a *complement* notation on the lattice with coordination number q such that $\bar{\mathcal{B}}(i) = q - \mathcal{B}(i)$. The procedure as illustrated in Figure 2.4 is as follows:

1. Bootstrap by *burning* the vertex v_∞
2. For each lattice site, if $z(i) > \bar{\mathcal{B}}(i)$, *burn* site i
3. Iterate (2) until all sites are burnt or it is no longer possible to burn another site.
4. To each site burnt at t , add a directional edge connecting it to a site burnt at $t - 1$.
5. A configuration is *recurrent* if every site is burnable (results in a spanning tree).

When adding tree edges (step 4 above), if there is more than one candidate direction $d \in \mathcal{D}$, an edge precedence $\mathcal{P}$ is employed to ensure an injective mapping $f(v_i, \mathcal{D}, \mathcal{P}) \rightarrow d$. In Figure 2.4, 0-indexing non v_∞ sites from top left, we may have for site v_9 i.e. $\boxed{3^{(3)}}$ the direction choice;

$$f(v_9, \{W, N\}, [N > E > S > W]) \rightarrow W$$

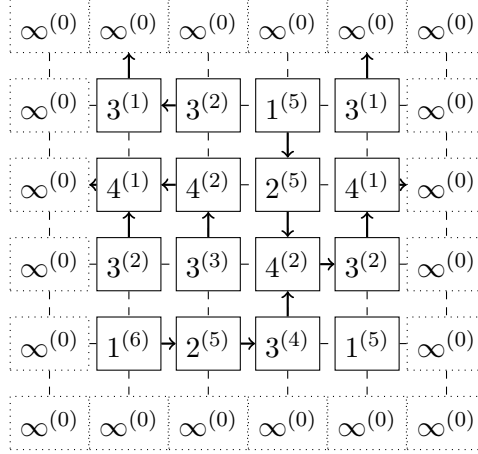


Figure 2.4: The burning algorithm on a lattice is shown with the lattice boundary connected to v_∞ . Site heights $z(i)$ appear in larger font. The superscripts show the “burn time” based on the burn condition. Bootstrapping at time 0, v_∞ is burnt. Continuing in iterations, to burn a site, $z(i)$ must exceed the number of un-burnt neighbours $\bar{\mathcal{B}}(i)$. As such, the top left non v_∞ site can be burnt at time 1 while the site to its right cannot be burnt until time 2.

Constraints on heights In the burning test procedure, when mapping from a recurrent configuration to a spanning tree, the site height is constrained by

$$\bar{\mathcal{B}}_i < z_i \leq \bar{\mathcal{B}}'_i \quad (2.14)$$

where $\bar{\mathcal{B}}_i, \bar{\mathcal{B}}'_i$ are respectively the number of neighbours (of site i) burnt *at or after* time t or *at or after* $t - 1$. This follows directly from the burning condition; the left-hand side of the inequality is precisely the burn condition that dictates when site i burns. The right-hand-side follows from the fact that site i could not have burned by $t - 1$ as the burn condition was not satisfied. The number of neighbouring sites $\mathcal{D}_i$ that were burnt precisely at $t - 1$ is $\bar{\mathcal{B}}'_i - \bar{\mathcal{B}}_i$. Any *ambiguity* in inequality (2.14) is resolved by employing the edge precedence to choose a specific d . As such we have [Majumdar and Dhar, 1992, Járai, 2005],

$$z_i = \bar{\mathcal{B}}_i + d \quad (2.15)$$

2.4.2 Spanning trees to recurrent configurations

Given a vertex v_i of a spanning tree $\Gamma(G)$, the configuration height z_i can be resolved via an inverse mapping with edge precedence $\mathcal{P}$ of form

$$f^{-1}(v_i, \Gamma(G)_i, \mathcal{P}) \rightarrow z_i \quad (2.16)$$

The breadth-first traverse algorithm (§A, Algorithm 1) is similar to other inverse mappings, such as work by Cori and Le Borgne [2003] and the depth-first algorithm described by Perkinson et al. [2013]. The breadth-first traverse has relatively low additional stack (memory) requirements, dominated by the storage of $\bar{\mathcal{B}}_i(t), \bar{\mathcal{B}}'_i(t)$ of the order of tree width. The bijection provides an injective function that locally and unambiguously maps a vertex of depth $\tau_i \in \Gamma(G)$, to the site height z_i within the recurrent configuration. The choice

of variables is motivated by what is known during a breadth-first traverse of $\Gamma(G)$ while maintaining a low memory footprint. The tree depth corresponds directly to burn-time as can be seen in Figure 2.4 where each burn-time is also the tree-distance between v_i and v_∞ . The algorithm can compute $\mathcal{B}'$, vertices burnt *strictly* before $t - 1$, without storing their *burn time*. It then computes $\mathcal{D} = \mathcal{B} - \mathcal{B}'$ i.e. vertices burnt *at* $t - 1$. Given $\mathcal{B}$ it is possible to use (2.15) to compute z_i . $\bar{\mathcal{B}}_i$ is resolved as $q - \mathcal{B}_i$ and d follows from using the index of the edge exiting $v_i \in \Gamma(G)$ among the set of candidate edges $\mathcal{D}$ as in the previous section. In Algorithm 1 (§A), d becomes *parentIndex*.

2.5 The Feynman graph

2.5.1 Graph representations

The incidence matrix $\mathbf{E}$ is *almost* enough to describe a Feynman graph. We may have multiple particle *species*, which will determine the *kinematics* (masses and momenta) carried on graph edges or the combinatorics of combining graphs. For this reason, the graph should be *decorated*. A vector $\mathcal{C}$ of length $|E|$ gives the edge colouring so as to map an undecorated graph to a decorated one.

Definition 11. A Feynman graph $\mathcal{G}(\mathbf{E}, \mathcal{C})$ is a directed, decorated graph described by its incidence matrix $\mathbf{E}$ and an edge colouring $\mathcal{C}$.

Remark 19. The incidence matrix contains all the data for a regular graph $G(V, E)$. By convention the last vertex in the incidence matrix is v_∞ which allows partitioning of edges into a set of external and internal edges. External edges are those edges incident to v_∞ .

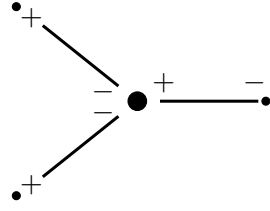
There are two types of edges on the diagrammatic representation of monomials in the Heisenberg-Weyl algebra; those entering a vertex and those exiting a vertex. We distinguish between edges that enter any vertex in the graph, corresponding to operators z (or fields $\tilde{\phi}$), and edges that exit the vertex, corresponding to operators ∂_z (or fields ϕ). Recall that the incidence matrix already carries information about whether an edge enters (+1) or exits (−1) a vertex. L -loop graphs are constructed by combining primitive tree-level graphs. As will be shown in examples throughout the chapter, the composite incidence matrix necessarily preserves the edge direction for each vertex, thus respecting the underlying algebra.

To see how the edge direction features in graphs, we need to *zoom in* on the *stars* of the graph i.e. primitive vertices and their edges. First, recall the propagator and assign a *polarity* based on the fields (below, spending the delta function, the resolved superscripts on k^+, ω^+ are to be dropped in the denominator as they are irrelevant);

$$\frac{\delta(\omega^+ + \omega^-) \delta^{(d)}(k^+ + k^-)}{-i\omega^+ + D(k^+)^2 + m^2} = \frac{\phi}{\tilde{\phi}} = \frac{+}{-} \quad (2.17)$$

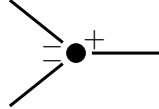
I have chosen the polarity so that it matches the convention of the incidence matrix i.e. any edge that exits a vertex will have a negative sign (at the vertex end). For example the

branching vertex with momenta flowing from right to left is,



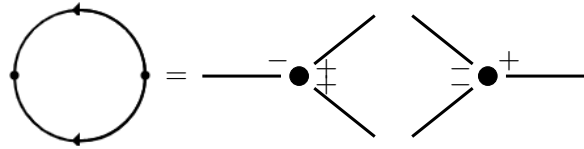
$$(2.18)$$

The so-called amputated vertex or corolla includes *only the half-edges* and the vertex;



$$(2.19)$$

When combining amputated vertices, we can only join positive to negative ends i.e. we must pair up creation with annihilation fields. We can see below that the flow always moves in a *positive* direction via the vertices as will be reflected in the incidence matrix of any *composite* graph.

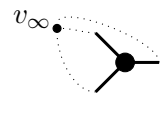


$$(2.20)$$

Later in the chapter I define a *graph product*, which combines graphs to form new graphs to a specified loop order. In the composite graphs, at the level of the incidence matrix, the edge directions corresponding to momentum flow are preserved under the graph product operation.[‡]

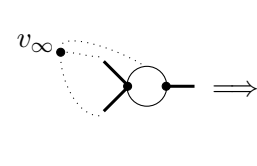
Remark 20. As will be seen when defining the graph product, the edge colouring i.e. the species type of the fields will naturally constrain how graphs are combined.

Example 2.5.1. Incidence matrix for primitive vertex with a clockwise edge arrangement and a row in $\mathbf{E}$ for each vertex including v_∞ :



$$\Rightarrow \begin{bmatrix} q_0 & -q_1 & -q_2 \\ -q_0 & q_1 & q_2 \end{bmatrix} \quad (2.21)$$

Example 2.5.2. Incidence matrix for one-loop graph with the same residue as the previous example: This graph is constructed as shown in (2.20). The 3 vertices including v_∞ and 5 edges (with preserved edge directionality) are represented in the incidence matrix.



$$\Rightarrow \begin{bmatrix} p_0 & p_1 & 0 & -q_1 & -q_2 \\ -p_0 & -p_1 & q_0 & 0 & 0 \\ 0 & 0 & -q_0 & q_1 & q_2 \end{bmatrix} \quad (2.22)$$

[‡]The correct flow is crucial when solving systems of conserved momentum equations to determine the kinematics flowing on edges!

Remark 21. A subtle but important point is that flow in the composite diagram is unambiguous modulo diagram symmetry. The polarity of edges at vertices is not always evident from looking directly at a composite graph.

We may wish to know if two graphs are isomorphic or have common sub-graphs. Two graphs G, F are *loop isomorphic* if $G \setminus \mathfrak{R}(G)$ is isomorphic to $F \setminus \mathfrak{R}(F)$. That is, we ignore the graph residues and compare the structure of the directed loop. As for defining isomorphism on all graphs, this is done using a suitable *graph canonisation* i.e. we compute a map $C : G \rightarrow G_c$ to determine the canonical form for $G \setminus \mathfrak{R}(G)$, which gives a canonical vertex labelling.

Symmetry factors

There are a number of different conventions for defining symmetry factors. I give a simple definition on $\mathcal{G}(\mathbf{E}, \mathcal{C})$ data that will suffice for the graphs we'll encounter. The symmetry factor is determined by enumerating stars in the graph and taking the factorial of the number of ϕ_s fields of a given species. Denote the number of the *outgoing edges* (at least one of which is internal) of species j on the i^{th} star $|\phi_{ij}| \geq 0$ for $v_i \neq v_\infty$ and the *incoming edges* of species j on the i^{th} star $|\phi_{ij}| \geq 0$ for $v_i = v_\infty$ then the symmetry factor, with $0! = 1$, is

$$\mathfrak{s}(G) = \prod_j \prod_i^{s_v |V|} |\phi_{ij}|! \quad (2.23)$$

Remark 22. Choosing outgoing fields, each time an edge of a certain colour (species) is formed, there are as many ways to merge two vertices as there are outgoing edges of that colour. Once the edge is chosen, the remaining edges can be fixed in the same way. It is possible to distinguish internal and external symmetry factors by dividing by the symmetry factor of the residue R , which is defined to correspond to the set of *incoming edges* at v_∞

Example 2.5.3. In this tree diagram with a single species there are 4 edges incoming to v_∞ and there are two *internal* vertices with internal edge pairings, where there is in both cases a choice of two edges.

$$\mathfrak{s}(\text{---} \begin{array}{c} \nearrow \nwarrow \\ \nearrow \nwarrow \\ \nearrow \nwarrow \end{array} \text{---}) = 4! \times 2! \times 2! \quad (2.24)$$

Example 2.5.4. In this loop diagram with a single species there is 1 external edge incoming to v_∞ and there is one *internal* vertex with an internal edge pairing, where there is a choice of two edges.

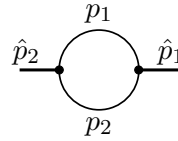
$$\mathfrak{s}(\text{---} \begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \text{---}) = 1! \times 2! \quad (2.25)$$

Remark 23. Strictly speaking, the definition for $\mathfrak{s}$ should be modified to divide by a possible overall *internal* graph symmetry that might be double-counted at the residue. This will not be a problem in the graphs we encounter.

Kinematic flows and conservation

The theory of network flows (e.g. electric circuits) is used to determine the kinematic data found on internal edges. The external half-edge structure dictates how external momentum flows on internal edges. Once this is resolved and we add the correct external momentum data onto the internal edges, we can ignore the external edges entirely! Therefore the size of graph matrices that we deal with will depend on whether we have resolved momentum flow or not.

Note on graph anatomy Consider the example graph



$$(2.26)$$

I introduce decorations to distinguish *loop* kinematics $\bar{\kappa}$, *external* kinematics $\hat{\kappa}$ and the linear combinations of loop and external momenta, which will be the (*unadorned*) κ_e momenta we actually care about on edges. For $|E|$ internal edges, L loops and $|R|$ external half-edges

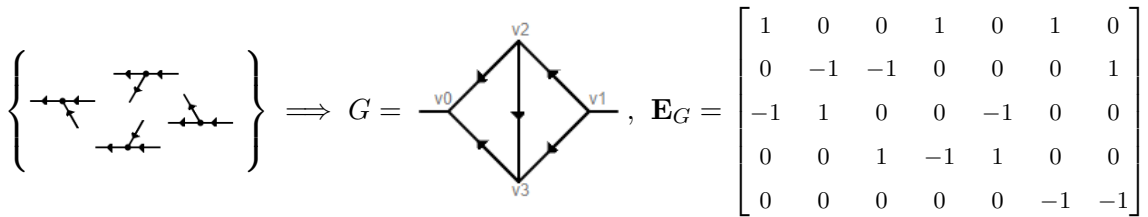
$$\kappa_e = \sum_j^L \iota_{ej} \bar{\kappa}_j + \sum_j^{|R|} \eta_{ej} \hat{\kappa}_j, \quad \iota_{ej}, \eta_{ej} \in \{0, -1, 1\} \quad (2.27)$$

Here ι_{ej}, η_{ej} are coefficients which determine if certain momenta are carried on a given edge. For Feynman graphs used in non-equilibrium field theory, the momenta contain space but also time components and a diffusion constant also appears in the propagator; $p := -i\omega + Dk^2$. As per (2.27), ω, k are linear combinations of loop and external momenta. Again for the external case I write $\hat{p}$ and for the non external case I write p (it is, after all, polite to take one's hat off inside).

A *constraint system* can be solved to determine a matrix $\mathcal{E}$ of size $|E| \times |E|$ which I refer to as the *edge basis*. The constraint equations are determined by using conservation of momentum at every vertex including v_∞ . The edge basis $\mathcal{E}$ can be expressed concisely in terms of the incidence matrix $\mathbf{E}$ restricted to the set of edges in a spanning tree T of G [Golz, 2018],

$$\mathcal{E} = \mathbf{E}_{[T]}^{-1} \mathbf{E}_{[G \setminus T]} \quad (2.28)$$

Example 2.5.5. The example graph below shows a suitable flow. There are two loops and two designated loop edges $\beta_0 = \alpha_2, \beta_1 = \alpha_4$ say. This example graph is constructed by a certain combination of 2 branching and 2 coagulation vertices - the internal edges connecting vertices $[v_i, v_j]$ by the first 5 columns in $\mathbf{E}_G$, are $[2, 0], [1, 2], [1, 3], [3, 0], [2, 3]$



$$\left\{ \begin{array}{c} \text{Four vertices with flow arrows} \end{array} \right\} \Rightarrow G = \begin{array}{c} \text{Graph with vertices } v_0, v_1, v_2, v_3 \end{array}, \quad \mathbf{E}_G = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{bmatrix}$$

Momentum is conserved at each vertex which produces a system of equations. Solving this system gives a possible assignment of kinematic variables on edges. From (2.28) one finds

$$\mathcal{E}_f = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.29)$$

Each edge is represented above in the *full* "loop basis" of size $|E| \times |E|$. When dealing with graph polynomials and Feynman diagrams generally, we will restrict attention to *internal* edges after the appropriate kinematics have been assigned to edges. In this solution matrix, looking at the rows, e_0 carries imaginary loop momentum $-i\omega_0$ and e_1 carries $-i\omega_0 + i\omega_1$ where the ω_j indices follow from the designated loop edge maps $\alpha_i \rightarrow \beta_j$. Later in this chapter we will integrate the omega variables by expressing the omega integrations as delta functions (cf. §2.5.3.2). These delta functions are determined by $\mathcal{E}_f$; looking at columns, the two delta functions (one for each loop) are $\delta(\alpha_2 - \alpha_0 + \alpha_1)$ and $\delta(\alpha_4 - \alpha_1 - \alpha_3)$ •

There is a harmless ambiguity in the choice of edge labels but there is a *dangerous* ambiguity in the choice of initial edge flows. Using the *shuffle product* (§2.5.3.1) to generate all graphs from primitive graphs guarantees the correct flow. For many Feynman diagrams, particularly at one loop, one finds a suitable planar mapping of the graph (2D drawing) and then assigns momentum flow always in one direction (e.g. left) on edges. This is a valid flow. However this is ambiguous for either very complicated graphs or working directly (algorithmically) from the incidence matrix. In this chapter the approach is to either determine the flow by inspection using a convention for simple graphs or use a *constructive* method to determine flow in general. In the latter case, each graph is generated by a product operation on primitive corollas which have directionality corresponding to the types of fields (incoming or outgoing). These directions are preserved in the composite graphs so that the signs in the incidence matrix correspond directly to the field types in the underlying theory. The constructive method differs from attempting to assign a flow to an existing graph having lost information about which corollas may have been used in generating the graph.²²

k-trees and the Symanzik polynomials

The *Symanzik polynomials* play an important role in *parametric* representations of Feynman integrals used later in this chapter. As these will be used to organise kinematic variables in the propagators, allow me to fix some notation and conventions.

The first and second Symanzik polynomials Ψ and Φ are closely related to the Kirchhoff polynomial $\mathcal{K}$ introduced above. In my notation, $\mathcal{K}$ is a polynomial in z where terms of order k give the k -trees. If the Kirchhoff polynomial gives the edges *in* the k -tree, the Symanzik polynomials focus instead on the *edge cutset*. As such, the terms on the Symanzik polynomials are derived simply from the *edge complements* of the terms found

in $\mathcal{K}$, which we know how to compute from the matrix tree theorem. Denote the coefficients of z^p as $\mathcal{K}_i^{(p)}$ and define $\mathcal{S}_i^{(p)}$ as their edge complements. For example for a spanning tree $\{e_0, e_1, e_2\} \setminus \{e_0\}$, $\mathcal{K}_0^{(1)} = \{e_1, e_2\}$ and $\mathcal{S}_0^{(1)} = \{e_0\}$. Using an intermediate polynomial φ , the first and second Symanzik polynomials are

$$\Psi = \sum_i \mathcal{S}_i^{(1)} = \sum_{T \in \Gamma_{\{1\}}} \prod_{e \notin T} \alpha_e \quad (2.30a)$$

$$\varphi = \sum_j \hat{p}_j^2 \mathcal{S}_j^{(2)} = - \sum_{F \in \Gamma_{\{2\}}} \hat{p}_1 \cdot \hat{p}_2 \prod_{e \notin F} \alpha_e \quad (2.30b)$$

$$\Phi = \Psi \sum_e \alpha_e m_e + \varphi \quad (2.30c)$$

$\hat{p}_j^2$ is the square of the momentum flowing across the cutset of the j^{th} 2-tree. For 2-trees that isolate internal vertices (those *not* connected to v_∞) from external vertices (those connected to v_∞), $\hat{p}_j^2$ will vanish as there is a zero total external momentum on the cutset. $\hat{p}_j^2$ is shown as a dot product of the momentum flowing *in* to the two trees in the 2-tree. Due to momentum conservation, $-\hat{p}_1 \cdot (-\hat{p}_2) \rightarrow \hat{p}_1^2 = \hat{p}_2^2$. From the definition of $\mathcal{S}^{(k)}$, the α parameters in these definitions are understood to be a product of all edge parameters *not in* the respective k trees. Below I focus on some of the matrix representations of the polynomials for use in symbolic computing applications.

Example 2.5.6. The spanning trees for the amputated sunset diagram were shown at the end of Example 2.2.1. Connecting each vertex to v_∞ , the graph polynomials are easy to see as they are formed of the cut-sets of the k -trees. For the three spanning trees shown in Example 2.2.1, the dotted lines in each spanning tree form a cutset and the first polynomial is a sum of the cutsets in each tree. The single 2-forest cuts all edges and disconnects the only two vertices from each other. The momentum flowing on the cutset $\{\alpha_0, \alpha_1, \alpha_2\} = E$ is the momentum either in/out of one of the edges connected to v_∞ . We have,

$$\Psi = \alpha_1 \alpha_2 + \alpha_0 \alpha_2 + \alpha_0 \alpha_1 \quad (2.31a)$$

$$\varphi = q^2 \alpha_0 \alpha_1 \alpha_2 \quad (2.31b)$$

The polynomial Φ is easily constructed from these •

Remark 24. The matrix representations that follow can be a little tedious unless there is a motivation for algorithmic methods as in this chapter. I provide a set of matrix examples in §C.9. Further details and proofs of the matrix representations are discussed in [Panzer, 2015] or [Bogner and Weinzierl, 2010].[§]

Matrix representations of graph polynomials For some suitable matrices $\mathbf{M}, \mathbf{Q}, \mathbf{K}$ the propagators can be expressed as

$$\sum_e \alpha_e (-\hat{p}_e + m_e^2) = \sum_{i,j=1} k_i \mathbf{M}_{ij} k_j + \sum_{i=1} 2k_i \cdot \mathbf{Q}_i + \mathbf{K}_j \quad (2.32)$$

and the Symanzik polynomials which will appear in parametric representations can be written [Bogner and Weinzierl, 2010]

$$\Psi = \det(\mathbf{M}), \quad \Phi = \Psi \cdot (\mathbf{Q}\mathbf{M}^{-1}\mathbf{Q} + \mathbf{K}) \quad (2.33)$$

[§]When comparing derivations and proofs care should be taken in conventions for example I use an incidence matrix with edges on columns and vertices on rows which may not be the case for other authors.

The determinant of the matrix $\mathbf{M}$, by linearity of the determinant can be expressed as a sum over determinants

$$\det(\mathbf{M}) = \sum_{S \in \Gamma_{\{1\}}} \prod_e \alpha_e \det \begin{pmatrix} 0 & \mathbf{E}_{[S]}^T \\ -\mathbf{E}_{[S]} & 0 \end{pmatrix} \quad (2.34)$$

The most convenient forms (for symbolic computing applications) are those that can be expressed directly in terms of the incidence matrix or its minors or the *Laplacian matrix* or its minors. Recall that the Laplacian matrix is given by $\mathbf{L} = \mathbf{E}\mathbf{E}^T$ and the *reduced Laplacian* $\mathbf{L}_R$ is the Laplacian of the cofactor matrix formed by deleting suitable rows and columns. To express graph polynomials in terms of alpha parameters, multiply the reduced Laplacian by the diagonal matrix in the alpha parameters to produce the *reduced symbolic Laplacian matrix*

$$\mathbf{L}_{RS} = \mathbf{E}_{[\Gamma]} \mathbf{D}_\alpha \mathbf{E}_{[\Gamma]}^T, \quad \mathbf{D}_\alpha = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_{|E|}) \quad (2.35)$$

I define also the dual matrix

$$\tilde{\mathbf{L}}_{RS} = \mathbf{E}_{[\Gamma]} \mathbf{D}_\alpha^{-1} \mathbf{E}_{[\Gamma]}^T \quad (2.36)$$

In what follows I drop the subscripts on the reduced symbolic Laplacian (because we will only ever care about the reduced matrices) and also the diagonal matrix, but their structure is understood.

Example 2.5.7. A spanning tree for the diagram in Example 2.5.5 is given below

$$\Gamma \left(\begin{array}{c} \text{Diagram of a graph with vertices } v_0, v_1, v_2, v_3 \text{ and edges } e_0, e_1, e_2, e_3, e_4. \end{array} \right) = \dots \quad (2.37)$$

The square matrix $\mathbf{E}_{[\Gamma]}$ below, is the (column) restriction of the (amputated) incidence matrix to the spanning tree edges $\{e_1, e_3, e_4\}$. Deleting any one designated vertex (row 0) from the internal vertex set and deleting the edges *not* in the spanning tree (columns $\{e_0, e_2\}$) gives the (square) minor

$$\mathbf{E}_{[\Gamma]} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 1 \end{bmatrix} \quad (2.38)$$

The reduced Laplacian matrix can now be calculated given the minor $\mathbf{E}_{[\Gamma]}$ and the diagonal symbol matrix $\mathbf{D}_\alpha$. These matrices and their determinants appear in definitions below. Carrying out the linear algebra, example matrices appear in §C.9 for reference •

The Symanzik polynomials can be written with external momenta vector $\hat{p}(F)$ representing the momentum flowing across the cutset of a given 2-tree F as [Panzer, 2015]

$$\Psi = \det \mathbf{D} \cdot \det \tilde{\mathbf{L}} = \sum_{T \in \Gamma_{\{1\}}} \prod_{e \notin T} \alpha_e \quad (2.39a)$$

$$\Phi = \Psi \cdot \left(\sum_e \alpha_e m_e + \mathbf{p}^T \tilde{\mathbf{L}}^{-1} \mathbf{p} \right) = \Psi \sum_E \alpha_e m_e + \sum_{F \in \Gamma_{\{2\}}} \hat{p}(F) \prod_{e \notin F} \alpha_e \quad (2.39b)$$

Remark 25. The exponentiated parametric representation (2.48) will take, as part of the exponent, the term

$$\frac{\Phi}{\Psi} = \sum_e \alpha_e m_e + \mathbf{p}^T \tilde{\mathbf{L}}^{-1} \mathbf{p} = \sum_E \alpha_e m_e + \frac{1}{\Psi} \sum_{F \in T_{\{2\}}} \hat{p}(F) \prod_{e \notin F} \alpha_e \quad (2.40)$$

In the alternate Feynman parametrisation (2.47), the term appears

$$\frac{\Psi^{\eta-d/2}}{\Phi^\eta} = \frac{\Psi^{\eta-d/2}}{\Phi^\eta} = (\Phi/\Psi)^{-\eta} \Psi^{-d/2} \quad (2.41)$$

For graphs carrying no external momenta, the form simplifies as the right-most term in (2.39b) vanishes $\square$

Remark 26. It is useful to switch between graph polynomials and their matrix representations, particularly representations based upon the symbolic Laplacian, which is easy to compute. The notation may appear cumbersome. However, the graph polynomials are easy to write down for simple graphs by determining their k -trees by inspection of the diagrams, while for more complex graphs or for algorithmic methods, polynomials are easy to compute in matrix form $\square$

Properties of Symanzik polynomials In the derivation of the parametric integral representation using graph polynomials it is important to observe some properties of the graph polynomials

1. The first polynomial Ψ is homogenous in the alpha parameters of degree L
2. The second polynomial Φ is homogenous in the alpha parameters of degree $L + 1$
3. The first polynomial is independent of kinematic data (masses and momenta) and contains only alpha parameters. All kinematic data are contained in the second polynomial
4. Φ is linear in the alpha parameters for zero mass and otherwise quadratic.

2.5.2 Integral representations

I summarise some of the main forms of integrals. Many of the representations are variations of each other.[¶] These results can be found in QFT textbooks such as Itzykson and Zuber [1980] or books about Feynman integration [Smirnov, 2005, 2006]. In what follows I list various formula and derivations of, in particular, the parametric representation.

It is desirable to be able to write arbitrary products of propagators in the form of the right-hand side below.

$$\frac{1}{\prod_e P_e} = \Gamma(\nu) \int_0^\infty \prod \alpha_e \delta(1 - \sum_E \alpha_e) \int d\bar{k} [k^2 + 2k \cdot Q + \sum_e \alpha_e p_e]^{-\nu} \quad (2.42)$$

[¶]This is nowhere more clearly seen than differences in language between *Feynman parametrisations* and *Schwinger parametrisations* which amount to the same things from two different perspectives, from two great physicists.

The Feynman parametrisation can be used in one of the following well-known forms. A general form of the Feynman parametrisation for propagators without raising propagators to powers is

$$\frac{1}{\prod_i A_i} = \Gamma(N) \int_0^1 d\bar{x}_{N-1} \frac{1}{[A_1 + \dots + x_{N-1}(A_N - A_{N-1})]^N} \quad (2.43)$$

The simplest case known as *Feynman's identity* can be written in terms of the *Beta function* (2.52). Considering powers of propagators,

$$\frac{1}{A^m B^n} = B(m, n) \int_0^1 d\alpha \frac{\alpha^{m-1} (1-\alpha)^{n-1}}{[A\alpha + B(1-\alpha)]^{n+m}} \quad (2.44)$$

with the special case $m, n = 1$

$$\frac{1}{AB} = \int_0^1 d\alpha [A\alpha + B(1-\alpha)]^{-2} = \int_0^1 d\alpha [\alpha(A-B) + B]^{-2} \quad (2.45)$$

For general powers in propagators (2.42) can be written

$$\frac{1}{\prod_e A_e^{\nu_e}} = \frac{\Gamma(\nu)}{\prod_e \Gamma(\nu_e)} \int_0^\infty \prod_e d\alpha_e \alpha_e^{\nu_e-1} \frac{\delta(1 - \sum_e \alpha_e)}{[\sum_e \alpha_e A_e]^\nu} \quad (2.46)$$

As shown in the previous section, we can collect the kinematics into two polynomials, the Symanzik polynomials. This makes it possible to express the parametric representation in terms of the graph polynomials,

$$\frac{\Gamma(\eta)}{\prod_e \Gamma(\nu_e)} \int_0^\infty \prod_e d\alpha_e \alpha_e^{\nu_e-1} \delta(1 - \sum_e \alpha_e) \frac{\Psi^{\eta-d/2}}{\Phi^\eta} \quad (2.47)$$

where $\eta = \nu - Ld/2$. Below I use the Schwinger representation of the propagator. Below we will derived a parametric representation from the momentum representation, exploiting the Euler Gamma function. This leads to the rather compact exponential form of the parametric integral,

$$\mathcal{I}_p(G) = \prod_{e \in E} \int_0^\infty \frac{\alpha_e^{\alpha_e-1} d\alpha_e}{\Gamma(\alpha_e)} \cdot \frac{e^{-\Phi/\Psi}}{\Psi^{d/2}} \quad (2.48)$$

Here I introduce the notation $\mathcal{I}_p(G)$ to refer to "the (parametric) integral representation" of the graph G . Spending the delta functions in mapping (2.47) $\rightarrow$ (2.48) uses a scaling identity with $\alpha_e = \nu_e/\nu$

$$\prod_e d\nu_i = \nu^{n-1} d\nu \prod_e d\alpha_e \delta(1 - \sum_e \alpha_e) \quad (2.49)$$

Illustrative calculations Before describing the parametric integral representation, I give some explicit calculations for one loop graphs in momentum space for use as reference. In momentum space we generally find an integral representation that is a product of *scalar propagators*

$$\int_\omega \int_k \prod_{i=1}^M \frac{1}{P_i^{n_i}}, \quad \int_\omega := \int \frac{d\omega}{2\pi}, \quad \int_k := \int \frac{d^d k}{(2\pi)^d} \quad (2.50)$$

We will cover various examples but for now I focus on this abstract form. The word *scalar* denotes the case where there are no vector terms in the numerator. For us this will be general enough but throughout the literature this form is often emphasised because *tensor* integrals can be reduced to (sums of) scalar ones [Tarasov, 1996].

Complex graphs with many propagator edges may be difficult to integrate. The parametric representation allows products of propagators to be expressed in terms of a sum over propagator terms multiplied by alpha parameters. In the next sections, the parametric forms will be discussed including dealing with the *omega* integration. For now, I suppose that by encountering a sufficiently simple graph or by applying results such as *Feynman's identity*, which is a simple case of the parametric representation, one may arrive at an integral of this form

$$\frac{1}{C^\nu} \int_k \frac{(k^2)^m}{(k^2 + \frac{\mathcal{M}k^0}{C})^\nu} \quad (2.51)$$

for some constant C . Here I note that for a polynomial in k , linear terms are removed and the terms of order 0 are referred to here as a *quasimass*.^{||} The special integral i.e. the Euler-Beta function is

$$B(a, b) = \int_0^\infty dz \frac{z^{a-1}}{(z+1)^{a+b}} = \int_0^1 dy y^{a-1} (1-y)^{b-1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad (2.52)$$

This is almost enough to see the well-known general formula for Feynman integration i.e

$$\int_k \frac{(k^2)^n}{[k^2 + \mathcal{M}]^\nu} = \iota \frac{\Gamma(\alpha)\Gamma(\nu-\alpha)}{\Gamma(d/2)\Gamma(\nu)} [\mathcal{M}]^{\alpha-\nu} \quad (2.53)$$

where I have introduced $\alpha = n + d/2$, $\iota = (-1)^{\nu+n}/(4\pi)^{d/2}$. This formula can be written with $\mathcal{M} = m^2$ but I express it here with the generalised quasimass, which can contain any physical parameters with positive real part that can be treated as constants of integration in k . I emphasise that some terms arise from a change to polar coordinates as explained below. For scalar propagators i.e. where $n = 0$, this simplifies to

$$\int_k \frac{1}{[k^2 + \mathcal{M}]^\nu} = \iota \frac{\Gamma(\nu-d/2)}{\Gamma(\nu)} [\mathcal{M}]^{d/2-\nu} \quad (2.54)$$

In fact, the denominator term $\Gamma(a+b) = \Gamma(\nu)$ should cancel with the one in the parametric formula so that the integrals often take the form

$$\int_k \frac{1}{[k^2 + \mathcal{M}]^\nu} \implies \iota \Gamma(\nu-d/2) [\mathcal{M}]^{d/2-\nu} \quad (2.55)$$

Note that UV poles will generally appear in the Gamma function.

Example 2.5.8. Taking $\nu = 2$ for some $d = 4 - \varepsilon$ say,

$$\Gamma(2-d/2) = \Gamma(\varepsilon/2) = \frac{2}{\varepsilon} - \gamma_E + \mathcal{O}(\varepsilon) \quad (2.56)$$

where I have used the *Euler-Mascheroni constant*

$$\gamma_E = \lim_{n \rightarrow \infty} \left(\sum_{j=1}^n \frac{1}{j} - \ln n \right) = 0.5772 \quad (2.57)$$

Remark 27. A good integration strategy is to reduce any graph to terms of the form (2.55), or sums of such terms, such that denominator contains a linearly reduced polynomial in k (possibly also containing alpha parameters). In the *branching random walk* process studied in the next chapter, integrals can typically be written in this form modulo constant pre-factors.

^{||}The name *quasimass* alludes to the fact that an expression containing external momenta and alpha parameters can be treated as a mass in this integration formula. Later $\mathcal{M}$ will be used as a scale-related test function corresponding to the normalisation point. Given the freedom in fixing the normalisation point, $\mathcal{M}$ can be treated as a single scale parameter.

The *quasimass* $\mathcal{M}$ is a constant of k -integration and may contain alpha parameters e.g. if using the Feynman identity or generalised parametric representation. Alpha parameters contained in $\mathcal{M}$ may need to be integrated out. In some instances there may be *cross-terms*, terms linear in k , in the propagator data. This is often easily dealt with by completing the squares and leads to external momentum data in the quasimass.**

The formulae above can be used to integrate momentum integrals. However, generally, the parametric integral representation can be understood as replacing integrations in k_l by integrations in the alpha parameters α_e . Many examples to follow. In the *non-equilibrium Feynman graph*, one must also integrate ω_l , loop frequencies. I show below that the omega integration can always be written as delta functions in the alpha parameters and we can replace integrations over k and ω with a *parametric integral*. Following this reasoning, rather than focus on Feynman rules to write down *momentum integrals*, which can then be integrated by the formulae above, the focus in this chapter will be to directly write down the *parametric integrals*.

Polar coordinates: Hidden in the integration $\int_k$ is an integral over d dimension. From this, a standard trick is to express this integral in terms of a d-dependant *angular factor*. We make the transformation using the volume of a d-dimensional hypersphere $\mathcal{V}_d$

$$\frac{d^d k}{(2\pi)^d} \rightarrow \frac{d\kappa}{(2\pi)^d} \mathcal{V}_d \kappa^{d-1}, \quad \mathcal{V}_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} \quad (2.58)$$

This can be written as

$$\frac{1}{(2\pi)^d} \frac{2\pi^{d/2}}{\Gamma(d/2)} \kappa^{d-1} \quad (2.59)$$

In fact to use this formula in the integral, we are interested in making the measure substitution corresponding to substitution $k^2 \rightarrow \kappa$ i.e

$$\int_0^\infty dk k^{d-1} \frac{1}{P(k^2)} = \frac{1}{2} \int_0^\infty d\kappa \kappa^{d/2-1} \frac{1}{P(\kappa)} \quad (2.60)$$

So with $2dk = d\kappa$

$$\frac{1}{2} \frac{1}{(2\pi)^d} \frac{2\pi^{d/2}}{\Gamma(d/2)} \kappa^{d/2-1} = \frac{1}{(4\pi)^{d/2}} \frac{\kappa^{d/2-1}}{\Gamma(d/2)} \quad (2.61)$$

This clarifies the additional terms in the general formula (2.53).

Parametric integral

*Feynman rules*²³ map Feynman graphs to integrals. I focus here on the parametric representation in terms of graph polynomials. The original book by Nakanishi [1971] covers graph polynomials in great detail and includes historical notes about their development on pp 44. Lighter coverage can be found in Itzykson and Zuber [1980]. For modern references, the theses by Bogner [2009] or Panzer [2015] are recommended. In his *sector decomposition* review, Heinrich [2008] also gives a nice overview of graph polynomials and their usefulness in symbolic computing. Some examples of parametric representations have already

**See also for example the comments and formula (8.12) in [Kleinert and Schulte-Frohlinde, 2001].

been given above starting from the simple case of Feynman's identity (2.44). Below, the parametric representation is derived using the Schwinger representation of the propagator and graph polynomials.

The Schwinger representation of the general propagator is

$$\frac{1}{P^n} = \frac{1}{\Gamma(n)} \int_0^\infty d\alpha \alpha^{n-1} e^{-\alpha P}, \quad \Re(\alpha) > 0 \quad (2.62)$$

where the parameters α are treated as test variables over edges in the graph - there are as many α -parameters as there are edges in the graph.

Using the Schwinger representation of the propagator, the parametric integral is written in terms of an exponential, which has some advantages. These advantages manifest in an exchange between integrals in position or momentum variables for integrals in alpha parameters, which may or many not seem like progress. Generally; (i) integrations in momentum and space variables (x, k) become Gaussian integrals. (ii) Integrations in the $i\omega$ variables become delta functions in the alpha parameters (iii) Then, all alpha parameters by which the integrals are now defined can be tidied up into graph polynomials. When moving between the exponentiated and non-exponentiated forms of the parametric representation, Euler's Gamma function is useful,

$$\int_0^\infty d\alpha \alpha^{n-1} e^{-p\alpha} = \frac{1}{p^n} \Gamma(n), \quad \Gamma(n) = \int_0^\infty \alpha^{n-1} e^{-\alpha} = (n-1)! \quad (2.63)$$

The Schwinger representation allows integrals with ultraviolet divergences (in terms of momentum in $\mathbb{R}^d$) to be written as convergent Gaussian integrals, which contain infrared divergences in the parametric space $\mathbb{R}_+^\nu$ for ν propagators/alpha parameters. Using the Symanzik polynomials it can be shown that the parametric representation can be written for d the space-time dimension,

$$\mathcal{I}_p(G) = \prod_{e \in E} \int_0^\infty \frac{\alpha_e^{\alpha_e-1} d\alpha_e}{\Gamma(\alpha_e)} \cdot \frac{e^{-\Phi/\Psi}}{\Psi^{d/2}} \quad (2.64)$$

To show this, we use the properties of the Symanzik polynomials discussed above in addition to the Schwinger representation, which in turn exploits Gaussian integrals. I begin with a basic example before showing the general form

Example 2.5.9. It is instructive to work through an example for the simplest one loop graph

$$G = \begin{array}{c} p_1 \\ \circlearrowleft \\ \hat{p}_2 \bullet \quad \bullet \hat{p}_1 \\ \circlearrowright \\ p_2 \end{array}, \quad \delta(p_1 - p_2 - \hat{p}_1) \quad (2.65)$$

The single delta function constraint is determined by momentum conservation for this symmetric graph. Using the Schwinger representation with $p := (Dk^2 - i\omega)$ the exponent in the Schwinger representation is

$$\sum_e \alpha_e (\hat{p}_e + m_e) := \alpha_1 (p_1 + m_1) + \alpha_2 (p_2 + m_2) \quad (2.66a)$$

I focus on the equilibrium physics case first by setting $D = 1, i\omega = 0$ because as mentioned, the omega variables will ultimately produce delta functions and I wish to focus on the

structure determined by the remaining variables. Using the delta function which for this case gives $k_2 = \hat{k} - k_1$, we then have for the exponent

$$\alpha_1 k_1^2 + \alpha_2 (\hat{k} - k_1)^2 + \alpha_1 m_1 + \alpha_2 m_2 \quad (2.67a)$$

$$= \underbrace{\alpha_1 + \alpha_2}_{\Psi} \underbrace{\left(k_1 - \hat{k} \frac{\alpha_2}{\alpha_1 + \alpha_2} \right)^2}_k + \underbrace{\hat{k}^2 \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} + \alpha_1 m_1^2 + \alpha_2 m_2^2}_{\Phi/\Psi} \quad (2.67b)$$

In the last step a completion of the squares and a momentum relabelling is applied and I identify the graph polynomials. The integral over the k part of the exponent is a multivariate Gaußian integral in spacetime dimension d (B.1-B.3). After performing the Gaußian integral only the terms on the right i.e. Φ/Ψ appear in the exponent part below;

$$\frac{1}{(2\pi)^d} (2\pi)^{d/2} \int_0^\infty \frac{\prod_e d\alpha_e}{\Psi^{d/2}} \cdot e^{-\Phi/\Psi} \quad (2.68)$$

The exponential is dealt with using the Euler Gamma integral with a scaling of the alpha parameters. If we consider the well-known Feynman parametrisation (2.47) we notice a delta function in the alpha parameters and the emergence of an index $\sigma_d(G) = -|E| + \frac{1}{2}d \cdot L$. Here one exploits what is known as the *Cheng-Wu theorem* (cf. [Smirnov, 2006, Panzer, 2015]) to choose hyperplane to explain the delta function.

$$1 = \int_0^\infty d\lambda \delta(\lambda - H(\alpha)) = \frac{d\lambda}{\lambda} \delta(1 - H(\alpha)) \quad (2.69)$$

Using (2.63) produces a Gamma function which, when using *dimensional regularization* (later in this chapter), is a function of ε ;

$$\mathcal{I}(G) = \int_0^\infty \frac{\prod d\alpha_e \times \delta(1 - H(\alpha))}{\Psi^{d/2}} \cdot \int_0^\infty d\lambda e^{-\lambda(\Phi/\Psi)} \lambda^{-\sigma_d(G)-1} \quad (2.70a)$$

$$= \Gamma(-\sigma_d(G)) \int_{[\alpha]} \frac{1}{\Psi^{d/2}} \left(\frac{\Psi}{\Phi} \right)^{-\sigma_d(G)} \quad (2.70b)$$

This gives the simplest parametric integral representation valid for an equilibrium Feynman graphs with *simple propagators* (raised to power 1) •

Equation (2.70b) is a very compact integral representation. The alpha integration $\int_{[\alpha]}$ as defined allows for setting one of the alpha parameters to 1 in practice. I will discuss practical aspects in examples. Equation (2.70b) is expressed only in terms of the graph polynomials in the alpha parameters and the *superficial degree of divergence* of the graph $\sigma_d(G) = -|E| + \frac{1}{2}d \cdot L$. The function σ_d is explained in more detail below in a more general context. For now I simply use it as some function of properties of the graph i.e. the number of edges, the first Betti number L (i.e. the number of loops in this case 1) and the spacetime dimension d . It should seem at first glance a little suspicious that we have a function of the *first Betti number*. Recall however that the first Betti number is contained in the *degree* of the graph polynomials; Ψ is homogeneous of degree L and Φ is homogeneous of degree $L+1$. Hence when we scale the *alpha* parameters, Φ/Ψ is always of degree 1 in the alpha parameters and λ^1 thus appears as a pre-factor in (2.70a). Similarly, a power $L \cdot d/2$ emerges from scaling *alpha* parameters in $\Psi^{d/2}$. The factor $|E|$ is perhaps less surprising and comes from the scaling of the alpha parameters, where $\prod_e d\alpha = \lambda^{|E|-1} d\lambda$ by construction. For simple graphs, the graph polynomials can be easily written down by inspection and for more complicated graphs they can be computed from the incidence matrix of the graph as discussed below.

The general form The derivation closely follows Panzer [2015] but extended to the out-of-equilibrium context.^{††} I introduce the frequency variables $i\omega_l$ distinguishing each independent loop l . It turns out that these lead to a product of delta functions (one for each loop), and this product enters the parametric integral for the non-equilibrium case. I also scale the graph polynomials according to the diffusion constants in the propagators. The example above was derived from a graph containing only *simple* propagators (those appearing with power 1). In the Schwinger representation (2.62) generally we have powers $n \geq 1$. This adds a term $\prod_e \alpha_e^{n_e-1}/\Gamma(n_e)$ to (2.70b). Starting in momentum space,

$$\mathcal{I}(G) = \prod_e \int_{-\infty}^{\infty} \frac{d^d k_e}{(2\pi)^{d/2}} \frac{d\omega_e}{2\pi} \frac{1}{-i\omega_e + D_e k_e^2 + m_e^2} \cdot \underbrace{\prod_{v \setminus v_\infty} \pi^{d/2} \delta^{(d)}(k_v) \pi^{1/2} \delta(\omega_v)}_{\Delta_c} \quad (2.71)$$

The delta functions apply constraints to determine the momentum flow on each edge in the edge basis as per (2.28). The Schwinger representation produces

$$\mathcal{I}(G) = \int_{-\infty}^{\infty} \frac{d^d k_e}{(2\pi)^{d/2}} \int_{-\infty}^{\infty} \frac{d\omega_e}{2\pi} \int_0^\infty \prod_e d\alpha_e e^{-\alpha_e(-i\omega_e + D_e k_e^2 + m_e^2)} \cdot \Delta_c \quad (2.72a)$$

where the exponent $\mathcal{P}$ after applying the delta function constraints for momentum conservation at vertices is shown in matrix form in (2.73). There is a subtlety to the application of the delta function, the enforcer of Kirchoff's laws in this context, and how it leads to a transition from an integral basis on independent *edges* to one of independent *circuits*. In the matrix representation one writes the analogue of (2.67b)

$$\mathcal{P} = -\tilde{\mathbf{d}}\mathbf{W}^T - \mathbf{K}\tilde{\mathbf{D}}\mathbf{K}^T - \mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T - \underbrace{\left(\sum_e \alpha_e m_e + \mathbf{p}^T \tilde{\mathbf{L}}^{-1} \mathbf{p} \right)}_{\frac{\Phi}{\Psi}} \quad (2.73)$$

Having completed the squares and using suitably defined vectors $\bar{k}$, $\bar{\omega}$ on the edge basis and $\bar{x}$ on the vertex basis, the matrices in the equation are of the form

$$\mathbf{K} = (\bar{k} - \frac{i}{2}\tilde{\mathbf{D}}^{-1}\tilde{\mathbf{E}}\bar{x}), \quad \mathbf{X} = (\bar{x}/2 - i\mathbf{L}^{-1}p), \quad \mathbf{W} = (i\bar{\omega}) \quad (2.74)$$

In analogy to Example 2.5.9 one performs the Gaussian integrals over position and *loop* momentum matrices $\mathbf{K}, \mathbf{X}$. The alpha symbol matrix is scaled by diffusion constants on each edge and thus appears $\tilde{\mathbf{D}}$. Additionally, as illustrated in §2.5.3.2, using the integral representation of the Dirac delta function i.e.

$$\delta(\Sigma) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega \cdot \Sigma} \quad (2.75)$$

the omega variables in $\mathbf{W}$ produce a delta function for each circuit. The argument of the delta function Σ is a polynomial in the alpha parameters on circuit l . Identifying the graph polynomials in $\mathcal{P}$, the derivation of the generalised parametric integral representation proceeds as in Example 2.5.9.

The general form of the parametric integral extended to the *non-equilibrium* case is

$$\mathcal{I}(G) = \Omega_d \cdot \Gamma(-\sigma_d(G')) \int_{[\alpha]} \frac{\prod_e \alpha_e^{n_e-1}}{\Gamma(n_e)} \frac{1}{(\Psi')^{d/2}} \left(\frac{\Psi'}{\Phi'} \right)^{-\sigma_d(G')} \cdot \prod^L \delta(\Sigma_l) \quad (2.76)$$

^{††}This section relates to nascent work with Marcel Golz on out-of-equilibrium parametric integrals. At present I provide a basic perspective, which will be developed in future work.

where Σ_l is a polynomial in edge variables per circuit for which coefficients $\{\pm 1\}$ define a hyperplane which determines the domain of integration. This polynomial is determined directly from the edge basis matrix (2.28). Ω_d is the angular factor which can be understood either via the Gaussian integrations or as explained in the *calculations-polar coordinates* section. Scaling in diffusion constants $D_e \geq 0$ corresponding to each propagator has been applied to the polynomials and alpha parameters as denoted by the dashes. One observes that Ψ' and Φ' , which are scaled by diffusion constants, emerge during the completion of the squares with respect to the k variable on each loop - for example in (C.43) and (C.44) using definitions of the graph polynomials (2.30). The degree of divergence is computed for the graph G' after edge contractions due to application of $\delta(\Sigma_l)$. Examples of parametric integration can be found in §2.5.3.3.

Remark 28. While for simple integrals, at one loop say, it may be more direct to tackle the integral in momentum space, for algorithmic methods or for more complex integrals, the parametric representation can be convenient. This representation is also very useful for understanding the singularities in Feynman graphs i.e. for *desingularisation* as exploited for example in the sector decomposition algorithm [Heinrich, 2008] and for exploring general mathematical properties of Feynman graphs [Brown, 2009, Panzer, 2015]. The parametric representation also supports an interesting approach to renormalization [Brown and Kreimer, 2013] in which certain renormalized (finite) integrals can be written directly in parametric form.

2.5.3 Graph operations

2.5.3.1 A shuffle product

I define a systematic method to combine graphs in all possible ways using the matrix representations. The basic idea is straight-forward but requires good book-keeping. How many ways can we combine these diagrams?

$$\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \times \begin{array}{c} \diagup \\ \diagdown \end{array} = ? \quad (2.77)$$

Or with multiple species?

$$\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \times \begin{array}{c} \diagup \\ \diagdown \end{array} = ? \quad (2.78)$$

Is the product commutative?

$$\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \times \begin{array}{c} \diagup \\ \diagdown \end{array} \stackrel{?}{=} \begin{array}{c} \diagup \\ \diagdown \end{array} \times \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \quad (2.79)$$

And how do we deal with loops?

$$\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \times \begin{array}{c} \diagup \\ \diagdown \end{array} \stackrel{?}{=} \begin{array}{c} \diagup \\ \diagdown \end{array} \times \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \quad (2.80)$$

One way to define a product would be to allow *grafting* one diagram onto the other, which would require inserting external graph edge pairs into existing edges. This has certain desirable algebraic properties²⁴ but is not feasible in this context. For example there is no way to graft a transmutation vertex  into an existing edge.

I define a product that operates only on the *full combinatorial residue* of the graph. Recall the residue of a graph is the set of external edges with all internal edges contracted to a point. With the last correspondence I introduce a tensor object to show the cardinality of incoming and outgoing edges for each species.

$$\mathfrak{R}(\text{Diagram}) = \text{Diagram} \hat{=} \begin{bmatrix} 2 & 1 \end{bmatrix} \quad (2.81)$$

The full combinatorial residue distinguishes vertices in the graph and separates the residue of each vertex. This definition becomes necessary for enumerating all graphs.²⁵

$$\mathfrak{R}_f(\text{Diagram}) = \text{Diagram} \hat{=} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad (2.82)$$

Remark 29. The full combinatorial residue of a tree is the tree itself (distinguished vertices).

The rule for constructing graphs is easily stated as combine primitive graphs *in all possible ways* to generate new graphs up to some desired order in loop number. To define *all possible ways* I use the graph residues. While the full residue as exemplified in (2.82) which produces a $2 \times s \times |V|$ tensor is used in practice it is much easier to describe and visualise using the $2 \times s$ matrix for the simpler residue in (2.81). Notation to describe the possible edge pairings is demonstrated in the following example

Example 2.5.10. Here the input edges of the *left-hand-side* are compared with the output edges of the *right-hand-side* for each species, as shown in boldface. The *minimum* values are preserved and then mapped to a set of pairings for an arbitrary set of species $A, B, C \dots$

$$\text{Diagram 1} \times \text{Diagram 2} \quad (2.83a)$$

$$= \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \{s_A^{(1)}\} \quad (2.83b)$$

where $s_A^{(1)}$ corresponds to choosing 1 (of 1 possible) pairings for species A •

All potential pairings (see also Figure 2.5) are written $\Omega = \{s_A^{(0)}, s_A^{(1)}, \dots, s_A^{(n)}, s_B^{(0)}, s_B^{(1)}, \dots, s_B^{(m)}\}$. The relevance of this, a set of size N , is that it gives a maximal matching of two diagrams before considering set permutations ${}_N P_k$ for $k = 1..N - 1$. Given all possible permutations of half-edge pairings of a graph, I define multiplicative maps $\otimes_\mu$ for pairing μ , which can be enumerated in an algorithm to iteratively generate all graphs from a primitive set of corollas.

Definition 12. Let Ω be the maximal half-edge pairing between the full residues of graphs G' and G'' . Denote the union of all permutations of Ω as $\mathfrak{M}$ with items $\mu \in \mathfrak{M}$. There is a unique graph product $\otimes_\mu$ that connects diagrams over a filtration of possible half-edges. The *shuffle product* generates diagrams by applying all such products;

$$G' * G'' = \sum_{\mathfrak{M}} \otimes_\mu(G'G'') \quad (2.84)$$

Remark 30. By construction, there is a bijection between $\cup_i \otimes_{\mu_i}(V_1, V_2)$ over vertex graphs V_1, V_2 and the resulting composite graph G . The set of multiplications $\otimes_{\mu_i}$ with μ_i being a permutation of decorated half-edge sets, yields a unique G and given G the particular set of multiplications $\cup_i \otimes_{\mu_i}(V_1, V_2)$ can be recovered. The set $\cup_i \otimes_{\mu_i}(V_1, V_2)$ which is easily encoded is used as a *hash function* for the loops in the algorithmic context.

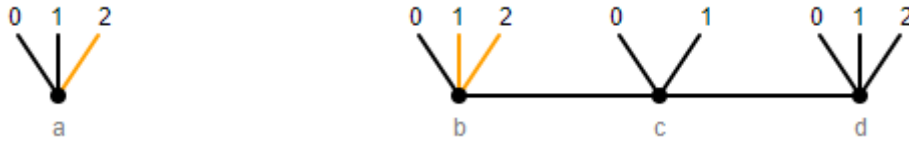


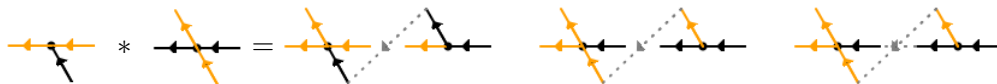
Figure 2.5: *Half-edge pairing:* Pairing rules follow from the definition of graph isometry. In 1-vertex and 3-vertex graphs shown, how many ways can black half-edges on a be paired with black half-edges on d ? It would be wrong to enumerate all ways of pairing two edges on a with three edges on d . In fact there are only three ways; i. *both* edges on a are paired with b or ii. *only one* is iii. *none* are. Which edges are paired, on d say, does not matter as the choices lead to isometric (coloured) adjacency structures. However the graph produced by pairing $\{(a_1 \rightarrow c_1), (a_2 \rightarrow d_1)\}$ is not isomorphic to the graph produced by pairing $\{(a_1 \rightarrow b_1), (a_2 \rightarrow d_1)\}$. In total, black pairings include; no pairings, all ways of one of each on a going to only one of each of the others or both edges on a pairing with both edges on either c or d . Yellow edges can be paired in two ways in total; either no edge pairing occurs or a pairs with b .

Graph generation algorithm

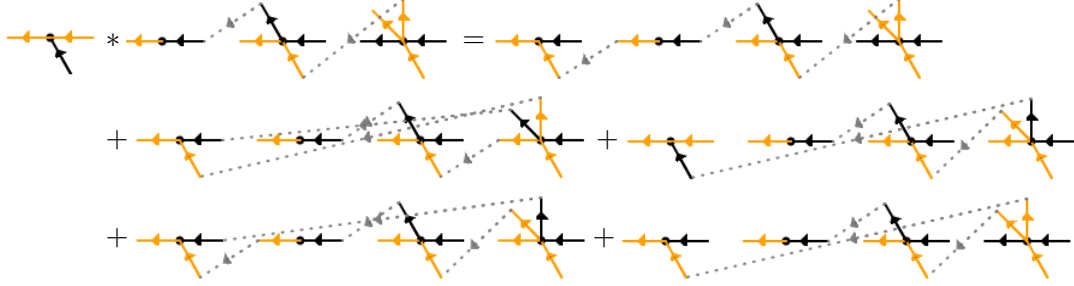
Starting from the primitive corollas, the graph generation algorithm generates L loop graphs up to some l_{max} . Any graphs that are generated with $L = 0$ are referred to as *radicals* in that they are retained in order to generate new graphs even if they are not otherwise required. Trees (radicals) may be grown to some maximal vertex count v_{max} .

In iterations, primitives with $|V| = 1$ are combined via the shuffle product with radicals having vertex count p for $p = 1, 2, \dots, v_{max}$. Adding a vertex to a radical of size v_p will either create a loop or a radical of size v_{p+1} . If the maximum loop order l_{max} is reached or the graph contains a bridge, such a graph is not used in further graph product operations.

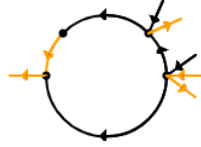
Example 2.5.11. The following example combines two radicals of size $p = |V| = 1$. This is the simplest case because both the left and right graph have simple residues where multiple vertices need not be distinguished. There are multiple possible edge pairings and a set of graphs are produced with loop numbers 0, 0, 1.



Example 2.5.12. In a more complicated example, now a radical of size $p = 1$ on the left is combined with a radical of size $p = 3$. By convention the algorithm always combines radicals of size $p = 1$ (on the left) with radicals of size $p \geq 1$ (right). The full combinatorial residue of the right-hand side determines symmetries of vertex pairings *for each species and for each vertex* before projecting onto specific edges



This representation emphasises the edge pairings but the loops are not particularly easy to make out when they exist. The fourth diagram in the result, for example, is this one-loop bridgeless graph,



It is a little tedious to consider combinations of all possible graphs but having defined a suitable graph product the iterative algorithm will simply output loop graphs of interest •

2.5.3.2 Contours, cover-ups and contractions

The non-equilibrium field theory is distinguished by the following integration measure in momentum space

$$\int \frac{d^d k}{(2\pi)^d} \frac{d\omega}{2\pi} \quad (2.86)$$

The additional ω integration not present in equilibrium field theories is the *causal fingerprint* of the evolutionary Markovian system. This corresponds to a temporal symmetry of the semigroup and conservation of energy. In practice, particularly at one loop, the ω (i.e. loop frequency) integration is trivial and seldom given much thought. However, in order to determine Feynman rules for directly writing down a parametric integral representation of the graph, it is worth thinking about the ω integration. In particular, using the *Schwinger representation*^{††} of the graph's propagators, an integration operator $\int_{\delta(\omega)}$ that uses only delta functions can be seen to be equivalent to the contour integral operator $\oint$, which uses the residue formula for rational functions.

The operator $\int_{\delta(\omega)}$ has the additional benefit of elucidating how the integration operators act on the graph. We observe (i) contractions of propagators, (ii) expansions in graphs by a combinatorial pre-factor that depends on the graph topology and (iii) broadcasting of kinematics among propagators at any loop order.

^{††}I am grateful to Erik Panzer for pointing out to me the merits of Schwinger's representation.

To begin with, it is easy to see by example that

$$\oint G = \int_{\delta(\omega)} G \quad (2.87)$$

The Schwinger representation of the propagators is

$$\frac{1}{P^n} = \frac{1}{\Gamma(n)} \int_0^\infty d\alpha \alpha^{n-1} e^{-\alpha P} \quad (2.88)$$

Here $\Gamma(n) = (n-1)!$ is the Gamma function.

Remark 31. The parameters α are called *alpha parameters*, *Schwinger parameters* or *Feynman parameters*. They should be treated as *test variables* associated with graph edges and play an important role in graph polynomials and parametric integral representations.

We recall the integral representation of the delta function

$$2\pi\delta(x) = \int_{-\infty}^{\infty} dt e^{-it \cdot x} \quad (2.89)$$

Cauchy tells us that $\oint G = 2\pi i \sum \text{Res}(G)$. The residue formula for rational functions $P = (\mp i\omega_l + \bullet)^{-\nu}$ can be written in this context as

$$\frac{1}{\Gamma(\nu)} (\mp i)^\nu (G//P)^{(\nu-1)} (\pm \bullet), \quad \Re(\bullet) > 0 \quad (2.90)$$

where I take the $(\nu-1)^{th}$ derivative of the momentum representation of the graph with the propagator edge(s) containing the pole of order ν contracted.

Example 2.5.13. As illustration, consider the simple one-loop diagram and its momentum and Schwinger representations. The momentum *omega* integral is

$$I(\text{one-loop}) = \int \frac{d\omega}{2\pi} \int_k \frac{1}{(-i(\omega + \hat{\omega}) + D(k + \hat{k})^2 + m)} \frac{1}{(i\omega + Dk^2 + m)} \quad (2.91a)$$

$$= \int_k \frac{1}{(D(k + \hat{k})^2 + Dk^2 - i\hat{\omega} + 2m)} \quad (2.91b)$$

$$= \frac{1}{2D} \int_k \frac{1}{k^2 + \underbrace{\frac{1}{2} \left(\frac{\hat{k}^2}{2} + \frac{2m - i\hat{\omega}}{D} \right)}_{\mathcal{M}}} = I'(\text{one-loop}) \quad (2.91c)$$

In the second line, a contour in an imaginary half plane is chosen and formula (2.90) is applied. I do not perform the k integration, which is not the point of the example. Alternatively using the Schwinger representation (2.88) the integration yields;

$$\begin{aligned} I(\text{one-loop}) &= \int \frac{d\omega}{2\pi} \int_k \int_0^\infty d\alpha_1 d\alpha_2 e^{-\alpha_1[-i(\omega + \hat{\omega}) + D(k + \hat{k})^2 + m]} e^{-\alpha_2(i\omega + Dk^2 + m)} \\ &= \int_k \int_0^\infty d\alpha_1 e^{-\alpha_1(-i\hat{\omega} + D(k + \hat{k})^2 + m)} e^{-\alpha_1(Dk^2 + m)} \\ &= \int_k e^{-\alpha_1(-i\hat{\omega} + D(k + \hat{k})^2 + m)} e^{-\alpha_1(Dk^2 + m)} \cdot \delta(\alpha_1 - \alpha_2) \\ &= \int_k \frac{1}{(D(k + \hat{k})^2 + Dk^2 - i\hat{\omega} + 2m)} = I'(\text{one-loop}) \end{aligned}$$

In the last line I *invert* the Schwinger representation to recover (2.91b), demonstrating the equivalence of the integration methods for this scalar Feynman graph •

The operator $\int_{\delta(\omega)}$ provides a convenient means to integrate frequency variables in an algorithmic context or for complex graphs. Determining the correct delta function requires proper resolution of the signs (flow) of $\pm i\omega_l$ variables. In general the L delta functions for an L -loop graph are determined unambiguously from the matrix (2.28). In simple cases it can be determined by inspection if there is no ambiguity about the signs that the omega variables can take up to graph symmetry. If, in a planar graph, it is possible to unambiguously choose a source and a sink external vertex on each independent circuit l (necessarily connected by two paths in that circuit) such that all positive loop frequencies $\pm i\omega_l$ flow on one path and all negative frequencies flow on the other path, then the alpha parameters in the delta functions will take the signs of the loop frequencies $\pm i\omega_l$ flowing on that edge. The delta function describes the integration region where a hyperplane should be chosen that partitions negative alpha parameters from positive ones.

Remark 32. I use the following language to describe the topological effect of integration in graph G . Either because of the delta function substitution or by the so-called *cover up rule*, when computing residues, the propagator P that carries the pole is removed. Topologically we can describe this removal as a *contraction* $G//P$. Furthermore, by evaluating $G//P$ at the pole or by applying the delta function constraint for the loop, all kinematics carried in the pole term are *broadcast* from terms $\pm i\omega_l$ in one imaginary half-plane to terms $\mp i\omega_l$ in the other when integrating loop momentum ω_l . The following example further illustrates broadcasting of kinematics after contracting edges.

Example 2.5.14. If all the propagators are of *order* 1 and the integration region is evident from the frequency signs, the integration remains straightforward for example in this two loop case

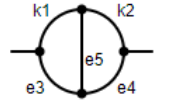
$$\mathcal{I}(G) = \frac{1}{p_0 p_1 p_2 p_3 p_4}, \quad \begin{cases} p_0 = (Dk_0^2 + i\omega_0) \\ p_1 = (Dk_1^2 + i\omega_1) \\ p_2 = (D(k'_1 + k_0)^2 + i\omega'_1 - i\omega_0) \\ p_3 = (D(k_0 + k_1)^2 + i\omega_0 - i\omega_1) \\ p_4 = (D(k'_1 + k_0 + k_1)^2 + i\omega'_1 - i\omega_0 - i\omega_1) \end{cases} \quad (2.93)$$

When integrated, two propagators i.e. p_0, p_1 carrying only (positive) loop momenta are *contracted*, and the kinematics from those propagators are broadcast to the other propagators on the respective loops.

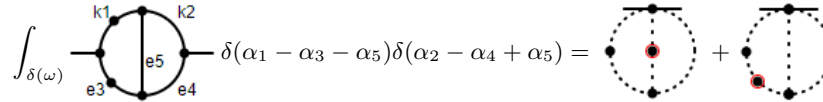
$$\oint \mathcal{I}(G) = \frac{1}{\hat{p}_2 \hat{p}_3 \hat{p}_4}, \quad \begin{cases} \hat{p}_2 = (Dk_0^2 + D(k'_1 + k_0)^2 + i\omega'_1) \\ \hat{p}_3 = (Dk_1^2 + D(k'_1 + k_0)^2 + D(k_0 + k_1)^2 + i\omega'_1) \\ \hat{p}_4 = (D(k'_1)^2 + 2Dk'_1 k_0 - 2Dk'_1 k_1 + 2Dk_0^2 + i\omega'_1) \end{cases} \quad (2.94)$$

For very simple graphs, the delta function determines the broadcasting of kinematics, which is the *only* impact of integration. This was the case in Example 2.5.13. To provide some intuition for the more general case of how the delta function acts on the Schwinger representation of the propagator (and in turn the parametric integral representation of the general graph), I use a graphical approach.

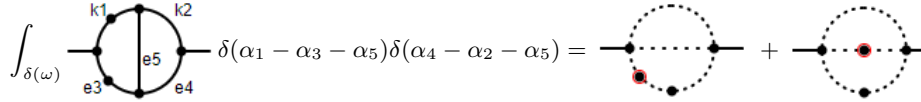
Comment on a graphical perspective of ω -integration The action of the contour integral on the graph has the following pattern. Consider the graph's edge labelling below


(2.95)

There are 2 external vertices, 2 internal vertices and 5 edges, 2 of which are designated loop edges k_1, k_2 . Increasing the power (splitting) of one of the loop edges, k_1 and one other edge e_3 I evaluate the integral graphically, taking delta functions from (2.28). Ordering delta functions starting with the loop variables, one might be tempted to write this as


(2.96)

However, similarly to how it is required to chose a positive *or* negative half plane in performing contour integration, here the correct hyperplane must be chosen for the delta function (note the different topology in the result due to contracting different edges!)


(2.97)

I make some remarks:

1. The *contracted* graph (isolated below) no longer contains edges k_1, e_4 after integrating out loop frequencies. I use dotted lines to emphasise that the loop frequencies have been integrated, leaving propagators that do not depend on $i\omega_l$ variables.



$$= \frac{1}{\hat{\rho}_3 \hat{\rho}_4 \hat{\rho}_5}$$
(2.98)

Integrated graphs are sums over such *basis* graphs with various powers of the basic propagators and some linear combination of kinematic data determined by broadcasting of kinematics from contracted edges (resembling the previous examples).

2. Because the loop edge k_1 in the Schwinger representation is raised to the power 2, which is represented graphically by adding an internal vertex on k_1 , the application of the delta function produces a pre-factor $\alpha_1 \rightarrow \alpha_3 + \alpha_5$. This produces the two graphs in (2.97). If propagators corresponding to both k_1 and k_2 are raised to power 2 then there would be a pre-factor $(\alpha_3 + \alpha_5)(\alpha_4 - \alpha_5)$ producing 4 terms.
3. Because the edge e_3 is raised to the power 2, which can be seen by the internal vertex splitting edge e_3 , this power is preserved in both of the new graphs - the edge remains split.
4. Additionally, circled in red, an extra internal vertex is added on each edge on the loop corresponding to the delta functions i.e. terms such as α_3, α_5 are substituted such that if the Schwinger representation were inverted, the power of the given propagator is increased by one.
5. The overall action can be appreciated as a sort of *Leibniz rule*.²⁶

2.5.3.3 Parametric integration

The general form of the parametric integral (2.76) derived above is reproduced here

$$\mathcal{I}(G) = \Omega_d \cdot \Gamma(-\sigma_d(G')) \int_{[\alpha]} \frac{\prod_e \alpha_e^{n_e-1}}{\Gamma(n_e)} \frac{1}{(\Psi')^{d/2}} \left(\frac{\Psi'}{\Phi'} \right)^{-\sigma_d(G')} \cdot \prod^L \delta(\Sigma_l) \quad (2.99)$$

and will be used for a number of examples. Generally, the angular factor is

$$\Omega_d = \frac{2\pi^{d/2}}{(2\pi)^d} \quad (2.100)$$

where the numerator comes from the d -dimensional Gaussian integral and the denominator from the integration measure $d^d k / (2\pi)^d$. I start with two examples of one loop integrals each with 2 edges for which

$$\sigma_d(G) = |E| - \frac{1}{2}d \cdot L = 2 - d/2 \quad (2.101)$$

and $n_e = 1, \forall e$. After applying the delta functions for omega integration, c of those edges are contracted and I denote this $\sigma_d(G')$ or $\sigma_d(G^{(c)})$ whenever it should be emphasised.

Example 2.5.15. Consider the graph with two propagators of species A with diffusion constant D_A ;

$$G = \text{---} \bullet \text{---} \bigcirc \text{---} \bullet \text{---}, \quad \begin{cases} \Psi' = D_A \alpha_1 + D_A \alpha_2 \\ \frac{\Phi'}{\Psi'}|_{\hat{k}, \hat{\omega}=0} = \alpha_1(m_A - i\omega) + \alpha_2(m_A + i\omega) \end{cases} \quad (2.102)$$

After applying the delta function with $\Sigma_l = \alpha_2 - \alpha_1$ and using the Cheng-wu theorem, one finds

$$\mathcal{I}(G) = \frac{\Gamma(-\sigma_d(G'))}{(2\pi)^{d/2}} \int d\alpha_2 \frac{1}{(2\alpha_2 D_A)^{d/2}} [\alpha_2 2m_A]^{\sigma_d(G')} \Big|_{\alpha_2=1} \quad (2.103a)$$

$$= \frac{\Gamma(1-d/2)}{(2\pi)^{d/2}} \frac{1}{(2D)^{d/2}} (2m)^{2/d-1} \quad (2.103b)$$

$$= \mathcal{A}_d \cdot \Gamma(1-d/2) [2m]^{2/d-1}, \quad \mathcal{A}_d = (4\pi D)^{-d/2} \quad (2.103c)$$

This example could be easily read off from the graph topology and data because after applying the delta function and exploiting the Cheng-wu theorem there are no *alpha* integrations required •

Example 2.5.16. Consider an example with the same topology but two different species such that propagator data are different

$$G = \text{---} \bullet \text{---} \bigcirc \text{---} \bullet \text{---}, \quad \begin{cases} \Psi' = D_A \alpha_1 + D_B \alpha_2, \quad D_B = 0 \\ \frac{\Phi'}{\Psi'}|_{\hat{k}, \hat{\omega}=0} = \alpha_1(m_A - i\omega) + \alpha_2(m_B + i\omega) \end{cases} \quad (2.104)$$

One finds

$$\mathcal{A}_d \cdot \Gamma(1-d/2) [m_A + m_B]^{2/d-1}, \quad \mathcal{A}_d = \frac{1}{2} (4\pi D)^{-d/2} \quad (2.105)$$

The argument of the Gamma function and the exponent of the *quasimass* follow from the graph topology and are unchanged while the data in the *quasimass* reflect the aggregated masses on the propagators of different species. The angular factor is modified due to the fact that species B does not diffuse •

Example 2.5.17. The last example considers a three-vertex graph

$$G = \text{Diagram}, \quad \begin{cases} \Psi' = D_A \alpha_1 + D_A \alpha_2 + D_B \alpha_3, & D_B = 0 \\ \frac{\Phi'}{\Psi'}|_{\hat{k}, \hat{\omega}=0} = \alpha_1(m_A + i\omega) + \alpha_2(m_A - i\omega) + \alpha_3(m_B - i\omega) \end{cases} \quad (2.106)$$

After applying the delta function with $\Sigma_l = \alpha_1 - \alpha_2 - \alpha_3$,

$$\begin{aligned} \mathcal{I}(G) &= \frac{\Gamma(-\sigma_d(G'))}{(2\pi)^{d/2}} \int d\alpha_2 d\alpha_3 \frac{1}{[(\alpha_2 + \alpha_3)D_A]^{d/2}} [\alpha_2 2m_A + \alpha_3 2m_B]^{\sigma_d(G')} \Big|_{\alpha_2=1} \\ &= \frac{\Gamma(2-d/2)}{(2\pi)^{d/2}} \int d\alpha_3 \frac{1}{[(1+\alpha_3)D_A]^{d/2}} [\alpha_3(2m_B) + 2m_A - 1]^{2/d-2} \\ &= \mathcal{A}_d \cdot \Gamma(2-d/2) [\widetilde{\mathcal{M}}]^{2/d-2}, \quad \mathcal{A}_d = (4\pi D)^{-d/2} \end{aligned}$$

All of these examples can be evaluated for $d = d_c - \varepsilon$. For the theory as discussed in Chapter 3, $d_c = 4$. This diagram has a symmetry factor $\mathfrak{s}(G) = 2$. I write the amplitude;

$$\frac{\mathcal{A}_d}{\mathfrak{s}(G)} \cdot \Gamma(\varepsilon/2) [1/\widetilde{\mathcal{M}}]^{\varepsilon/2} \quad (2.108)$$

This result will be revisited in Chapter 3 •

2.6 Renormalization

Renormalization has many faces. The idea of renormalization in physics may be traced back to Green [1835]²⁷ and an interesting discussion about the history of the concept in physics can be found in Dresden [1993].

In this section I introduce renormalization but I emphasise the abstractions that are, as much as possible, independent of physical theories. This is the appropriate partial perspective in the context of graphs. I will complete the story in Chapter 3 where one encounters the perspective due to Wilson [1971], where an intuitive picture based on Kadanoff's block-spin picture [Kadanoff, 1966] emerges. Wilson's intuitive idea, which was developed in the study of critical phenomena, can today be reinterpreted in the light of a more general idea of renormalization group theory, largely due to efforts by the "Saclay group" (cf. Vasil'ev [2004] ch. 1). One of many definitions of renormalization appears at the level of the Lagrangian density, which we have already met in the previous chapter.

Definition 13. Renormalization: A redefinition of the Lagrangian in terms of physical rather than bare (purely model-related) parameters. This is usually achieved by adding suitable counterterms to the Lagrangian.

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I + \mathcal{L}_{CT} \quad (2.109)$$

The *interaction* terms in $\mathcal{L}_I$ lead to a perturbative expansion in Feynman graphs mapped to integrals via *Feynman rules*. These integrals are not well-defined, which means series contributing to $\mathcal{L}_I$ are not well-defined. We can *temper* the objects (and thus the series) to make them well defined. At the level of the Lagrangian we can add suitable counterterms $\mathcal{L}_{CT}$ to arrive at finite, physically meaningful results. These counterterms are themselves evaluations of diagrams for certain choices of parameters i.e. normalisation points.

Remark 33. A theory is said to be *renormalizable* if the number of monomials (in terms of fields and their derivatives) that must be added to the Lagrangian remains finite. The counterterms are physically linked to physical observables hence the need for their finiteness. That the counterterms of a Lagrangian are polynomials in fields and their derivatives is referred to as *locality*.

It is possible to distinguish two different approaches to renormalization. The first method, the *counterterm method*, is direct and efficient at one loop for example, and it is the one used in the application in Chapter 3. This follows from the definition above where to each order in a coupling g , divergent counterterms are added to the bare action to produce a renormalised action. The second method is the BPHZ scheme introduced by Bogoliubov and Parasiuk [1957] and is perhaps more at home in this chapter as it proceeds graph by graph. This second route is more relevant at higher loop orders as it provides a recursive, combinatoric description of the renormalization procedure at higher loop order to manage what might otherwise be unwieldy calculations. This scheme has been reinterpreted in terms of Hopf algebras [Connes and Kreimer, 1999, 2000, 2001]. These two perspectives, the counter term and the BPHZ-based scheme, are equivalent *in effect*.

There are two procedures that are required to *temper* Feynman integrals; *regularization* and *renormalisation*. In a tongue and cheek sense, the first is to appease the mathematician i.e. to make analytical sense of the integral, while the second is to appease the physicist i.e. to make some physical sense of the result. At times the differences between *regularization* and *renormlsation* may be a little blurred. In the next chapter I discuss the combination of *dimensional regularization* and *minimal subtraction* due to Hooft and Veltman [1972], Hooft [1973], which is a common approach to carry out the combined program. This is particularly useful in one loop graphs. For more complex graphs, graphs may contain as well as divergences, nested sub-divergences. The Connes-Kreimer theory follows in the BPHZ tradition to systemically render all graphs finite by systematically dealing with sub-divergences. In fact unless there are additional practical reasons to regularize, BPHZ-like renormalization schemes can be applied at the level of the integrand without requiring an additional regularization method to arrive at a physically meaningful, finite result.

UV and IR divergences To understand the need to regularize/renormalize one begins by studying the pathologies of Feynman integrals. Of particular interest in this chapter are properties that can be determined from the graph topology.

The Dyson-Weinberg "power counting" theorem proved by Weinberg [1960] gives the superficial degree of (ultraviolet) divergence of 1PI subgraph γ . Here as illustration I take the definition as given by Kleinert and Schulte-Frohlinde [2001] in considering scalar graphs in ϕ^4 theory

$$\sigma_d(\gamma) = d \cdot L - 2|E| \quad (2.110a)$$

with the spatial dimension d and L the first Betti number. For $\sigma_d(\gamma) < 0$, γ is absolutely convergent. For $\sigma_d(\gamma) = 0$, $\sigma_d(\gamma) = 1$, the graph γ is respectively *logarithmically* and *quadratically* divergent.

Example 2.6.1. In $d = 4$, with $L = 1$ and $|E| = 2$

$$\sigma_d(\text{loop}) = 0 \quad (2.111)$$

is logarithmically UV divergent. Looking a little more closely, the graph has the following integral

$$\mathcal{I}(\text{loop}) = \int_{k^4} \frac{1}{k^2 + m^2} \frac{1}{(k - p)^2 + m^2} \quad (2.112)$$

where $\int_{k^4} = \int \frac{d^4 k}{(2\pi)^4}$ and p is the external momentum variable. This behaves like

$$\int_{\Lambda}^{\infty} d|k|^4 \frac{1}{|k|^4} \quad (2.113)$$

which has an *ultraviolet* divergence for $k \rightarrow \infty$ and an *infrared* divergence for $k \rightarrow 0$. Here Λ is some regularising scale boundary.

Definition 14. An *ultraviolet divergence* arises in the short-wavelength/large momenta region of the integral.

Definition 15. An *infrared divergence* occurs at small momenta and at the zero-mass limit.

Divergences and graph polynomials: Again in the context of dimensional regularization singularities can appear in the following ways in parametric representation of integrals; (i) single poles can appear depending on the argument of the Gamma functions being close to zero or negative (ii) if Ψ , which appears with a power in terms of ε vanishes at the boundary of the integration region, this leads to ultraviolet divergences (iii) if Φ , which contains the kinematics and appears with a power in terms of ε vanishes at the boundary of the integration region, this leads to infrared divergences. Infrared divergences appear in the parametric integrals for $\alpha_e \rightarrow \infty$ [Bogner and Weinzierl, 2010].

Dimensional Regularization and minimal subtraction: There are a number of ways to regularize Feynman integrals²⁸ and here only *dimensional regularization* is considered. In practice, the method is easy to state; Feynman diagrams are infinite in finite dimensions and the idea is to analytically continue the integral to dimension $d = 4 - \varepsilon$. Examples of Feynman integrals have been shown where poles appeared in the Gamma function with arguments in terms of ε . These poles contain all the information required to determine the critical exponents in critical dynamics. In minimal subtraction one defines an operator that extracts the pole part of the result. Under this scheme one can remove UV divergences for finite small ε at the critical point.

In understanding the (Bogoliubov-Parasiuk-Hepp-Zimmermann) BPHZ scheme, which is due initially to Bogoliubov and Parasiuk [Bogoliubov and Parasiuk, 1957], given a correct proof by Hepp [1966] and simplified in terms of the forest formula by Zimmermann [1968, 1969], one exploits the following insight; counterterms used to renormalize a Feynman graph γ are polynomials of degree $\sigma_d(\gamma)$. An operator can be applied to a divergent graph

to produce a convergent graph. Bogoliubov-Parasiuk referred to this as the R -operation. Applying a derivative operator with respect to momentum variables sufficiently many times to an integrand and subtracting this integrand from the original, will render a divergent graph convergent. The appropriate renormalization operator is a subtraction of a Taylor expansion up to order $\sigma_d(\gamma)$ evaluated for external momentum $\mathbf{p} = 0$.

Example 2.6.2. Suppose we take the integrand of the example above

$$\widehat{\mathcal{I}}(\text{loop}) = \frac{1}{k^2 + m^2} \frac{1}{(p - k)^2 + m^2} \quad (2.114)$$

Applying the derivative operator with respect to external momentum variable p and evaluating at vanishing external momentum

$$\partial_p \widehat{\mathcal{I}}(G)|_{p=0} = \frac{1}{(k^2 + m^2)^2} \quad (2.115)$$

If we subtract this term from the original integrand

$$\widehat{\mathcal{I}}(G) - \partial_p \widehat{\mathcal{I}}(G)|_{p=0} = \frac{2kp - p^2}{(k^2 + m^2)^2[(p - k)^2 + m^2]} \sim \frac{1}{|k|^5} \quad (2.116)$$

In $d = 4$, this differential form integrated with measure $\frac{d^4 k}{(2\pi)^d}$ is convergent.

Renormalized graph Suppose we wish to determine the contribution of a given graph, where that contribution is referred to in physics as the *amplitude*. The integral may not be well-defined therefore the following amplitude, which is also a function of a fiducial kinematic scale κ (e.g. relating to external momenta or masses), does not in general exist

$$\mathcal{A}(G, \kappa) = \int_k \widehat{\mathcal{I}}(G) \quad (2.117)$$

The proof of renormalizability due to Bogoliubov and Parasiuk [1957] and refined by Hepp [1966] says that there is a subtraction that produces a finite, renormalized amplitude and it is given by the subtraction operation

$$\mathcal{A}_R(G, \kappa) = \int_k \left(\widehat{\mathcal{I}}(G) - T^{\sigma_d(G)}[\widehat{\mathcal{I}}(G)] \right) \quad (2.118)$$

where as per above, $T^{\sigma_d(G)}$ is the Taylor operator, a polynomial of order $\sigma_d(G)$ evaluated at an arbitrary normalisation point. The monomials i.e. *counterterms* are

$$\mathcal{C}_n^p = -\frac{1}{n!} \int_k \partial_{\mathbf{p}}^n \mathcal{I}(G)|_{\mathbf{p}=0} \quad (2.119)$$

Combinatorics of renormalization A particularly mathematical description appears in [Connes and Marcolli, 2008] which follows in the tradition of the Connes-Kreimer theory;

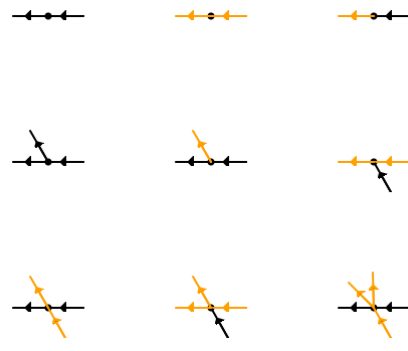
[T]he renormalization procedure corresponds exactly to the Birkhoff factorization of a loop $\gamma(z) \in \mathfrak{G}$ associated to the unrenormalized theory evaluated in complex dimension $d - z$, where d is the dimension of space-time and $z \neq 0$ is the complex

parameter used in dimensional regularization. The group $\mathfrak{G}$ is defined through its Hopf algebra of coordinates, which is the Hopf algebra of Feynman graphs of the theory. The Birkhoff factorization of the loop gives a canonical way of removing the singularity at $z = 0$ and obtaining the required finite result for the physical observables. This gives renormalization a clear and well defined conceptual meaning.

The BPHZ scheme and the forest formula give a recursive expression to renormalise graphs with nested sub-divergences. Connes-Kreimer renormalization enriches the mathematical description of this procedure in terms of a Hopf algebraic structure.

Generating Feynman graphs for reaction-diffusions

While the treatment of Feynman graphs in this chapter may have seemed rather abstract, it was done this way in support of defining an algorithmic method (Python code) for generating all Feynman diagrams corresponding to an arbitrary chemical reaction system as formulated in Chapter 1. Assuming transport for each species by linear diffusion with diffusion constant D_s , and given a reaction network in tensor form (e.g. matrix 1.37), we can generate all infinitesimal generators and their representations as tree-level diagrams. The tree level-diagrams are the primitive generators for all Feynman graphs $\mathfrak{G}_L$ to loop order L by way of the shuffle product, which acts on (coloured) incidence matrices to produce all bridgeless Feynman graphs of the form $\mathcal{G}(\mathbf{E}, \mathcal{C})$. Delta functions which exploit the *causal symmetry* and eliminate the imaginary loop frequencies, are determined from the edge basis matrix $\mathcal{E}$ and the parametric integral representation can be written down using these delta functions and the Symanzik polynomials. The resulting integral can be integrated parametrically or otherwise. The field theory discussed in the next chapter considers the *trace of a superprocess*. The primitive diagrams are;

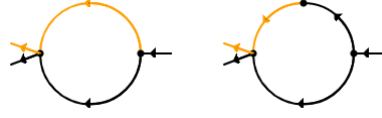

(2.120)

Applying the shuffle product seeded with n corollas and contracting internal edges produces m additional corollas such as this one

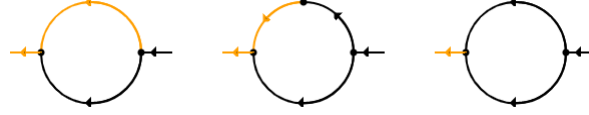

(2.121)

Using $n + m$ primitives in the iterative algorithm with suitable termination conditions, about 40 distinct diagrams are generated at first loop order without bridges (1PI). Ignoring graph residues, the procedure produces 7 *distinct* loops (integrals). Graphs are grouped by residues in the perturbation theory. For example the set of graphs below have the same

residue


(2.122)

as do


(2.123)

Adding additional primitive diagrams and additional species leads to an explosion in the number of diagrams in the set $\mathfrak{G}_L$ of all graphs generated to loop order L . Generally it will be quite tedious to enumerate diagrams and the algorithmic approach discussed in this chapter allows for the more tedious aspects to be automated. From the many diagrams that *can* be generated, there will be a relatively small number of distinct loops and depending on physical considerations, these may not all need to be integrated. To determine the scaling of the moments of the trace of the superprocess in Chapter 3 for example, only two graphs enter the picture and both of these map to the same integral (loop isomorphic at the normalization point).

Individual diagrams can be integrated parametrically or otherwise. Ultimately what we really want is not the possibly divergent amplitude of the graph but the *renormalization factor* Z_λ corresponding to all graphs that have the residue of the corolla Γ_λ . This factor will be used in the renormalization group analysis of the spatial moments. Z_λ is determined for example under a minimal subtraction scheme as a function of the sum of *counterterms* from all graphs with residue Γ_λ . While for many reaction-diffusion-type problems only a small number of loops and renormalization factors might need to be considered, using parametric integration/renormalization can be useful in algorithmic methods as demonstrated for example in the case of ϕ^4 at six loops [Kompaniets and Panzer, 2016].

3 Analysis of Scaling

While many-body systems in principle contain infinite degrees of freedom, in the study of critical phenomena one observes dynamical phases where microscopic degrees of freedom behave collectively, thus reducing the effective degrees of freedom. This collective behaviour is normally observed near a *continuous phase transition*.²⁹ The elimination of what turn out to be irrelevant degrees of freedom may map multiple systems into the same *universality class*. Of particular interest to us are non-equilibrium systems where the system starts at some point in phase space and then relaxes into stationary states in ways that are difficult to predict or model as determined by microscopic interactions. In Chapter 1, I defined such processes as compound Poisson processes living on a graph. Reaction-diffusion processes were introduced in the context of mean-field solutions while the perturbative field-theoretic methods were introduced in terms of the need to go beyond mean-field and in particular, quantify the impact of fluctuations on the macroscopic behaviour of the process. There are three potential outcomes; (i) the fluctuations are *irrelevant* so the mean-field precisely describes the macroscopic behaviour (ii) the opposite is true and the *relevant* fluctuations affect the macroscopic behaviour or (iii) *marginality* i.e. fluctuations cause typically logarithmic deviation from mean-field. Fluctuations may become relevant at a particular dimension d_c approached from below and such a d_c is called the *upper critical dimension*.³⁰ In dynamics, when a system approaches a critical temperature, the relaxation time τ diverges with correlation length ξ^z where z is the *dynamical critical exponent*.

In reaction-diffusion processes it is interesting to consider the interplay between reactions and diffusions. For example, diffusion may limit reactions as potential reactants are transported away from potential reaction sites. Any theory which fails to properly account for the role of diffusion is failing to properly *renormalize* the reaction rates so as to account for dissipation and reentrant properties of diffusing particles. If a bare equation of motion allows the rate couplings to be renormalized, we can, in principle, derive a renormalized equation of motion and analyse the long-term behaviour of the process using Renormalization Group (RG) techniques.

Example 3.0.1. Coagulation, again:^{*} Diffusion-controlled chemical reactions are adequately described by mean-field-type rate equations in higher dimensions, but in lower dimensions the fluctuations become relevant. A number of authors have studied coagulation reactions [Peliti, 1986, Droz and Sasvári, 1993, Lee, 1994, Deem and Park, 1998] using renormalization group (RG) methods. Suppose a single species system. We can first look at the time evolution of particles in the mean-field with reaction rate χ , $\partial_t n(t) = -\chi n^2(t)$ and then try to take into account the role of fluctuations. The rate equations suggest a power law decay in particles $n(t) = t^{-k}$ for some exponent k (corresponding the the number of coalescing particles in the reaction). Using methods described in Chapter 1, Peliti

^{*}In Chapter 1, see Examples 1.3.5, 1.3.6

[1986] gives an interacting Lagrangian density with interaction terms for coagulation;

$$\frac{1}{4}\lambda_0\tilde{\phi}^2\phi^2 + \frac{1}{2}\lambda_1\tilde{\phi}\phi^2 \quad (3.1)$$

One observes that neither monomial increases the number of fields $\tilde{\phi}^m\phi^n$ with $n \geq m$ such that combining any such terms in perturbative expansions is in some sense *contractive*. Terms absent from the initial Lagrangian but generated in the expansion are *irrelevant*. By dimensional analysis, both couplings λ_0, λ_1 have inverse length dimension $(2-d)$ and d appears as the upper critical dimension.³¹ Diagrams are summed in a geometric sum $\Gamma = [1 + \lambda\mathcal{I}(G)]^{-1}$ which implies an expression for renormalized rates $(\lambda_0^R, \lambda_1^R) = (\lambda_0, \lambda_1)\Gamma$. Lee [1994], through perturbative calculations finds actually at $d = d_c$, with $\kappa = 1/(k-1)$ and diffusion constant D

$$n(t) \sim \left[\frac{(k-2)!}{4\pi k^\kappa} \right]^\kappa \left[\frac{\ln t}{Dt} \right]^\kappa \quad (3.2)$$

Recall k , corresponds to $kA \rightarrow lA$ with $k = 2$ a special case for coagulation •

I provide a brief overview of the renormalization group machinery and apply it to study the scaling of (the moments of) distinct sites visited by a superprocess. This is work carried out in our group [Bordeu et al., 2019]. These problems are rooted in statistical physics and the study of critical phenomena. In Chapter 2 an algorithm was developed to carry out more efficient Monte Carlo simulation of the *Abelian sandpile model* that avoided the very slow equilibration epoch and some efforts to validate the algorithm are discussed in this Chapter.

3.1 Scaling symmetry and the renormalization group

The Wilson tradition³² of the renormalization group [Wilson, 1971] as applied to critical phenomena, allows for the calculation of critical exponents in a small parameter $\varepsilon = (d_c - d)$ where d_c is the critical spacetime dimension. The *renormalization group* (RG) functions are solved by determining the renormalization constants from Feynman integral calculations. RG analysis can determine if possible scaling regimes are given by the $\text{IR}^\dagger$ stable fixed points of the RG equations [Zinn-Justin, 1996, Vasil'ev, 2004, Kleinert and Schulte-Frohlinde, 2001]. A major contribution by Wilson was to provide an intuitive picture for attacking problems in physics. Wilson's work, which earned him a Nobel prize, was a major step forward from the mean-field theory that was developed by Landau.³³ In the Wilson picture, the scaling in a physical model needs to consider not just a physical scale but an observer's scale such as a lattice spacing. Wilson proposes an iterative procedure to integrate out irrelevant degrees of freedom at each scale to identify the essential physics.³⁴ Elements of this idea are sketched in Figure 3.1.

Remark 34. The renormalization group exploits a symmetry, namely, *scale invariance*. In Chapter 1, a *time-homogenous symmetry* was exploited in the semigroup of an evolutionary reaction diffusion process. In such cases the partial differential equation *might* eventually be solved by the method of characteristics, which exploits the symmetry of the one-parameter continuous group. In fact, all solutions to differential equations can be understood as

[†]Long-time, large-scale

the identification of inherent symmetries [Olver, 2000] and the case of the *renormalization group equation* is no different.

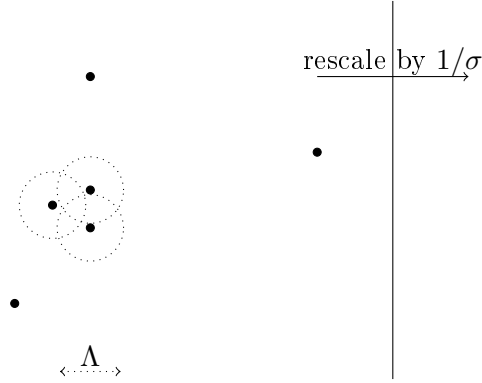


Figure 3.1: An annulus $[a, \sigma a]$ defines a scale range between a short-distance cutoff a e.g. lattice spacing and some scaling of it $\Lambda = \sigma \times a$. In order to remove short-range interaction effects one integrates out the degrees of freedom on the annulus and then rescales the system by a factor of inverse σ which creates an *effective* theory with a cutoff Λ . Short-distance behaviour describes a regime shorter than the correlation length but larger than the cut-off scale and this regime is governed by ultraviolet divergences. At the critical point, the short-distance behaviour extends to infinity and ultraviolet divergences are thus relevant to understand long-distance phenomena in many-body systems near the critical point

The renormalization group

We now look at some specific details of performing calculations. The goal is to modify the (infinite) theory in such a way that the theory will become finite after taking the limit $\Lambda \rightarrow \infty$ but without changing the physical properties or symmetries of the model. This process is called renormalization and the systematic scheme for carrying out calculations is described by the *renormalization group* (RG). RG transformations admit *fixed points*; all statistical models which, after these transformations, converge toward the same fixed point, belong to the same universality class. Our objective is to find *renormalized couplings*, which may be shown to have a stable fixed point of order ε . If we can demonstrate critical scaling, we can determine universal characteristics of the critical behaviour via the critical exponents and the normalized scaling functions [Vasil'ev, 2004]. We can encounter either a situation where the propagator needs to be renormalized, which is referred to as *wave-function renormalization* or in the absence of this, the couplings may be *multiplicatively renormalizable*.

The standard approach to studying critical behaviour using the renormalization group [Zinn-Justin, 1996] at an overview level proceeds as follows; the *bare* and *renormalized* vertex functions are related by a factor $Z(\lambda_R, \varepsilon)$ which has dependence on $\varepsilon = d - d_c$ and the renormalized rate coupling λ_R . The bare and renormalized vertex function relationship is of the form $\Gamma_B = Z\Gamma_R$, which corresponds to a so-called *multiplicative renormalization*.

The bare and renormalized couplings are related by

$$\lambda_B = \frac{\mu^\varepsilon \lambda_R}{Z(\lambda_R)} \quad (3.3)$$

where μ is a physical scale parameter. The renormalization factors or "Z-factors" will be determined using *dimensional regularisation* and *minimal subtraction* in the partitioning of divergent and convergent parts of Feynman graphs and will in this scheme take the form of a Laurent expansion in ε or rather λ_R/ε . They take the form

$$Z_i = 1 + \sum_{n=1}^{\infty} f(\lambda_R(\mu)) \varepsilon^{-n} \quad (3.4)$$

for some kinematic scale parameter μ . The two important renormalization group functions are the *beta* and *zeta* function. The zeta function gives the so-called *anomalous dimension* of the field as a variation in $\ln(Z)$. Up to constant pre-factors

$$\zeta(\lambda_R) = \tilde{\mathcal{D}}_\mu \ln(\lambda_R) \quad (3.5)$$

I introduce an operator $\mathcal{D}_\mu = \mu \partial_\mu$, in some circles referred to as the *Euler-Cauchy operator*, which may be applied at a point that fixes bare quantities $\tilde{\mathcal{D}}_\mu = \mu \partial_\mu|_B$ where the tilde distinguishes this evaluation point. We can write (3.5) as

$$\zeta(\lambda_R) = \beta(\lambda_R) \partial_{\lambda_R} \ln Z(\lambda_R), \quad \beta(\lambda_R) = \mathcal{D}_\mu \lambda_R \quad (3.6)$$

so as to introduce the second RG function, the beta function. These *dimensionless* RG functions, which do not depend on the scale parameter μ of the group, are of central importance. The *critical point* λ^* is the zero of the beta function and the *critical exponent* is the evaluation of the zeta function at this point. In the study of critical phenomena, the RG analysis requires a dimensional analysis of the Lagrangian which follows.

Dimensional analysis The Lagrangian density takes the form

$$\mathcal{L} = \int dt d^d x \Phi(\partial_t + \Delta) \tilde{\Phi} + \sum_{\Lambda} \mathcal{Q}_\Lambda[\Phi, \tilde{\Phi}] \quad (3.7)$$

where Φ and $\tilde{\Phi}$ represent fields for all of s species. The Lagrangian and each term in the Lagrangian is necessarily dimensionless and with diffusive scaling ($L^2 = T$)

$$[\mathcal{L}] = 1 = L^0, \quad [d^d x dt] = L^d T = L^{d+2} \quad (3.8)$$

Each term that has positive dimensional index i in L^i is relevant.

From the system of equations given by each term in the Lagrangian, it is necessary to determine the canonical (engineering) dimension of each of the couplings Vasil'ev [2004]. When we know the dimension of the fields, then the dimension of the couplings follows. From the free Lagrangian, given the dimension of the integration measure, with terms such as $\tilde{\phi}_a \partial_t \phi_a$ for each species,

$$[\tilde{\phi}_a \phi_a] = L^{-d}, \quad \forall a \quad (3.9)$$

There is insufficient information to completely determine the fields though, so we might impose that some couplings in the Lagrangian are dimensionless in order to resolve the dimensions of the fields. In [Bordeu et al., 2019] we make a particular choice, which is given in the following example.

Example 3.1.1. In the case of the superprocess studied below there are 2 species each with 2 fields. We need four equations to resolve four unknown field dimensions. Two of these equations come from the free Lagrangian as mentioned above. Choose to fix the dimension of two couplings $[\lambda_a] = [\lambda_b] = 1$ in the Lagrangian interaction for the (i) branching and (ii) transmutation terms (by a superprocess scaling ansatz [Bordeu et al., 2019]). Solving the system with this choice, the field dimensions for the two species are

$$\begin{bmatrix} [\tilde{\phi}_A] & [\phi_A] \\ [\tilde{\phi}_B] & [\phi_B] \end{bmatrix} = \begin{bmatrix} L^{-2} & L^{2-d} \\ L^{-4} & L^{4-d} \end{bmatrix} \quad (3.10)$$

By plugging these field dimensions into $\mathcal{L}$, the dimension of all couplings are determined. We find that at $d_c = 4$, all couplings beyond the two we fixed are *irrelevant* for $d \geq d_c$. One of the couplings is *marginal* at the critical dimension [Bordeu et al., 2019]. The general vertex is a monomial in these fields. Representing the general vertex in terms of powers of each field, we can write an expression for the dimension of the general vertex using the dimensions from (3.10)

$$[\Gamma \begin{bmatrix} \tilde{m} & m \\ \tilde{n} & n \end{bmatrix}] = L^{d(m+n-1)+2(\tilde{m}-m+2\tilde{n}-2n-1)} \quad (3.11)$$

Considering the exponent of L above, we have as a result the inequality for terms to be non relevant at $d_c = 4$

$$d(m+n-1)|_{d=4} + 2(\tilde{m}-m+2\tilde{n}-2n-1) \leq 0 \quad (3.12a)$$

$$m + \tilde{m} + 2\tilde{n} \leq 3 \quad (3.12b)$$

This dimensional analysis restricts diagrams that must be considered from all diagrams that we can generate.³⁵ The field theory needs to include all vertices with non-negative values for $m, n, \tilde{m}, \tilde{n}$ for which inequality (3.12) is satisfied. For example in the case where $\tilde{n} = 0$ and there are no incoming B particles, we must only consider diagrams that have at most $m + \tilde{m} \leq 3$ half-edges (of species A). The list of viable corollas will be given below●

3.2 Numerical studies

3.2.1 The Abelian sandpile

The *Abelian sandpile model* [Dhar, 1990] is a generalisation of the Bak-Tang-Wisenfeld (BTW) [Bak et al., 1987] model of self-organised criticality³⁶ on a (2D) lattice. The dynamics of the ASM were described in Chapter 2. These dynamics are inherently deterministic but stochasticity is introduced in Markov Chain Monte Carlo sampling of sandpile statistics by uniformly sampling random sites to *drive* the process. Driving the process means depositing a particle at a site. Given the update rule, this may cause particles to be redistributed to neighbouring particles and this can lead to a cascade of redistribution or toppling events. The key observables of the process relate to the toppling sequences, referred to as *avalanches*, induced by driving the system at some site.

The motivation for the current study stems from the significantly large *equilibration* time for the Abelian Sandpile; this is the time it takes to first reach a *recurrent configuration*, at which point meaningful statistics begin to emerge. The equilibration time does not scale favourably $\sim L^4$ drives. A spanning tree by contrast, may be generated within the *cover time* of the graph using Wilson's algorithm. To efficiently compute ASM statistics for large lattices, the ASM system is initialised and driven from recurrent configurations mapped from uniform spanning trees as discussed in Chapter 2.

Scaling in sandpile models Within self-organised criticality, the probability density for observable x is expected to display simple scaling (in system size L) of the form

$$\mathcal{P}_x = \mathcal{A}x^{-\alpha}\mathcal{G}\left(\frac{x}{\mathcal{B}L^z}\right) \quad (3.13a)$$

where observables of interest include the avalanche duration l , size s and *range* (alias trace or distinct sites) a . The universal critical exponents α, z are sought for each observable along with the universal scaling function $\mathcal{G}$ and factors $\mathcal{A}, \mathcal{B}$. When fitting, we add corrections to the functional form. An iterative routine attempts to fit with additional correction terms to see if the fitting (goodness of fit) can be improved. For example the following functions can be used in the fitting

$$\begin{aligned} f_2(L, \alpha, \beta) &= \alpha L^\beta \\ f_3(L, \alpha, \beta, \gamma) &= \alpha L^\beta \left(1 + \frac{\gamma}{L}\right) \\ f_4(L, \alpha, \beta, \gamma, \lambda) &= \alpha L^\beta \left(1 + \frac{\gamma}{L} + \frac{\lambda}{L^2}\right) \\ f_5(L, \alpha, \beta, \gamma, \lambda, \nu) &= \alpha L^\beta \left(1 + \frac{\gamma}{L} + \frac{\lambda}{L^2} + \frac{\nu}{L^3}\right) \end{aligned}$$

Provisional data for the observables, size s , duration l and trace a of the avalanche are given in the tables following; in particular, the moment scaling in system (lattice) size. The article in preparation [Amarteifio and Pruessner, 2019] provides further commentary on these data, which yield the exponents for scaling behaviour in the Abelian sandpile model through standard relations [Pruessner, 2012]. One begins by considering the *gap* between each moment (the value of μ in each row in the tables), which corresponds to the index parameter α in (3.13).

ASM scaling in avalanche area a_n

	μ	σ^2
a^1	1.556	13735.073
a^2	3.532	4801.659
a^3	5.522	2899.453
a^4	7.585	3884.738
a^5	9.586	3472.779

Table 3.1: *ASM avalanche area scaling:* For each moment, the mean and variance are shown for the scaling in the area (or trace) of avalanches.

ASM scaling in avalanche duration l^n

	μ	σ^2
l^1	0.842	103.237
l^2	2.503	4.501
l^3	4.286	5.734
l^4	6.165	5.227
l^5	8.11	5.018 8

Table 3.2: *ASM avalanche duration:* For each moment, the mean and variance are shown for the scaling in the duration of avalanches.

ASM scaling in avalanche size s^n

	μ	σ^2
s^1	1.999	5.529
s^2	4.793	8.012
s^3	7.664	4.773
s^4	10.377	6391.842

Table 3.3: *ASM avalanche size:* For each moment, the mean and variance are shown for the scaling in the size of avalanches. Size is defined as the number of sand grains involved in the event.

3.2.1.1 Algorithm validation results

In this section I will focus on the validation that was done to check the consistency of the ASM algorithm that was developed in Chapter 2.

A number of tests were devised to validate the recurrent configurations generated by the algorithm (as described in Chapter 2 and listed in Appendix A). The sample space of spanning trees/recurrent configurations grows extremely fast with increasing lattice side L .³⁷ Therefore I combine explicit checks for very small configurations with statistical tests for larger configurations. For small configurations (e.g. 3×3) one can sample say 10^6 configurations and ensure that of 192 (for $L = 3$) known configurations, each configuration is sampled uniformly. Wilson's algorithm samples *uniform* spanning trees and these should be mapped to *uniform* recurrent configurations. For larger configurations, I appeal to properties of sandpiles. Expected asymptomatic figures from Priezzhev [1994] are given in Table 3.4. These can be compared with sample recurrent configurations generated by the hybrid algorithm as shown in Table 3.5; the recurrent configurations generated by the hybrid algorithm should approach the expected site height distribution.

1	7,367,793
2	17,393,485
3	30,631,101
4	44,607,621

Table 3.4: *Expected ASM heights:* Expected (asymptotic) site height distribution on recurrent ASM configuration of size $L = 10^6$.

	Mean	st.dev
1	73809.611282	191.931382
2	174264.807627	371.657324
3	306488.059173	450.654951
4	445437.301013	320.047521

Table 3.5: *Checking site heights against expected values:* Mean and standard deviation for 10^6 samples of recurrent configurations on a lattice of size $L = 10^3$ matching theoretically expected asymptomatic figures.

Height symmetry sampling Recurrent configurations should be statistically symmetric; if the heights at corners are sampled in a configuration ensemble the top left should match the top right. A control test can be used by introducing a bias or a "drift", where directions on a lattice are chosen non-uniformly in the hybrid algorithm. By generating many recurrent configurations and counting the number of sites of size 4 for example, symmetry is observed when there is no bias. Without (known) bias the values are (statistically) equivalent.

	Top Left	Top Right
1	144228.9	144295.1
2	251031.4	251044.2
3	302365.3	302579.1
4	302374.4	302081.6

Table 3.6: *Symmetry test without bias:* An *insignificant* differences between height counts top left and top right supports *uniform* sampling. For example observe $|144228.9 - 144295.1|/144262.0 \approx 5 \times 10^{-4}$. Alternatively, the normalized deviation from the normalized mean (of the left and right values) should approach 0.5 and in this case it is 0.4998.

With an introduced bias they are quite different as shown in the table below.

	Top Left	Top Right
1	121515.3	182776.8
2	192301.7	119157.9
3	312229.6	406419.6
4	373953.4	291645.7

Table 3.7: *Symmetry test with bias:* Significant differences between height counts, top left and top right, indicate a bias in configuration sampling. In this case, the difference is of the order 5×10^{-1} , which is significant. Alternatively, the normalized deviation from the normalized mean (of the left and right values) should approach 0.5 but in this case it is 0.2988.

Remark 35. Various other tests were considered such as the distributions of the width of the first level of the spanning tree (i.e. the number of edges immediately connected to v_∞) as generated by Wilson's algorithm. If there is no bias, the number of these edges appears to approach an expected value.

3.2.2 The trace of a superprocess

A branching random walk is a special case of a *superprocess* [Etheridge, 2000]. The trace (range) i.e. distinct sites visited by the process is an interesting observable. The volume traced out by this process is referred to as the *branching Wiener sausage*. In order to count the distinct sites visited, a branching random walk with a single diffusing species A is augmented with a secondary *immobile* species B , which is deposited on unvisited sites. The field theory considers only the *critical* regime of the process. For a branching random walk, the critical point is independent of space. The critical point is inherited from the critical point of the underlying branching mechanism³⁸ - as discussed below this is a Galton-Watson-Bienaymé branching process. At the critical point, the mean offspring

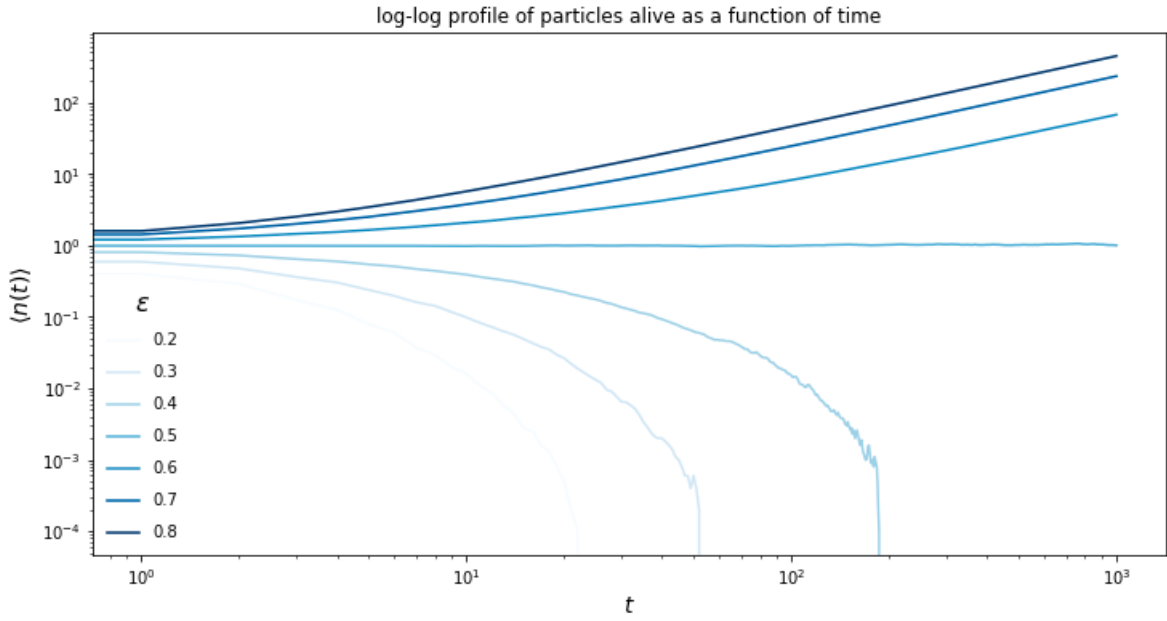


Figure 3.2: *Critical regimes for observables in branching processes:* Shown are profiles of the *logarithm of the ensemble average* of the number of particles observed as a function of *logarithmic time* for different parameter regimes in a branching process (or simple birth and death process) initialised with a single particle. The time-dependant solutions for the probability or moments of particle numbers appear with a factor $e^{(\beta-\varepsilon)t}$. For the critical regime with $\beta = \varepsilon$, any exponential factor drops out leaving a power law dependence on time. What this picture does not show is the divergence in the variance of the observable at the critical point. In general, dynamics of critical processes are strongly influenced by such fluctuations.

of the *dyadic* (0 or 2 offspring) branching process is $\mu = 1$. The sub/super/critical regimes are plotted in Figure 3.2. Properties of branching processes and simple birth and death processes can be found in standard reference texts on branching processes [Harris, 2002, Athreya et al., 2004] or textbooks on stochastic processes such as [Allen, 2010].

3.2.2.1 What is a superprocess?

The name "superprocess" was coined by Dynkin in the 1980s but describes processes studied decades earlier [Etheridge, 2000]. The following defines a general superprocess;

Definition 16. A (α, d, β) -superprocess is a stochastic process on $\mathbb{R} \times \mathbb{R}^d$ characterised by a branching mechanism parametrised by β and a (Levi) α -stable process describing particle transport.

Remark 36. A special case is the *Dawson-Watanabe superprocess* or *super-Brownian motion*. This is the case for which $\alpha = 2$ and the branching mechanisms is a dyadic critical Galton-Watson-Bienaymé branching process. The parameter $\alpha = 2$ in this case corresponds to Brownian motion on the (Levi) α -stable spectrum, as discussed in §C.12. In discrete space, this process is a *branching random walk*.

A superprocess is constructed as follows;

- Each particle moves on a graph, independently of all other particles as a Brownian motion i.e. transition directions are chosen uniformly from the adjacency matrix of the graph.
- Particle lifetimes follow Poisson statistics. They are exponentially distributed and there is a Poisson probability of observing particle deaths in some small time interval.
- At time of death, each particle produces offspring according to a probability generating function $\Phi(z) = \sum_n p_k z^n$. In the dyadic branching processes, the generating function allows for only the case of $p(X = 2) = \beta$, $p(X = 0) = \varepsilon$. In the critical case $\beta = \varepsilon$. For the processes as described, this critical point is independent of d .

This superprocess can also be formulated as a compound Poisson processes that includes hopping as one of the random events. In such case, there are three mutually exclusive events; hopping, branching and extinction. For the process to be critical the rate of branching and extinction must be equal and the hopping rate effectively scales time. Formulated either as a single compound Poisson process or as a compound poisson process with Brownian motion, the superprocess is a Markov process, being a collection of Markov processes.

Remark 37. This process is in some sense ideal because particles diffuse in space and particle numbers *diffuse* in phase space in that particle numbers in any δt can increase or decrease by 1. So the entire process is similar in a weak sense to a diffusion in phase space (though not in general isotropic).

Statistics of ordinary diffusion that make sense for $d \leq n$ will make sense for the super-diffusion in $n + 2$, an observation that is described as the *dimension gap* [Etheridge, 2000]. Ordinary diffusion has upper critical dimension $d = 2$. The upper critical dimension of a super-diffusion is therefore 4.

Using the perturbation theory and the renormalization group we have determined the scaling of the correlation functions for the trace of the *critical* superprocess on arbitrary graphs [Bordeu et al., 2019]. In Chapter 2, the general method for generating and computing graphs was given. From the dimensional analysis above, the restricted set of corollas and their one-loop graphs are given in the next section. the one-loop graph is considered to be sufficient for the study of critical properties as additional loops add accuracy but do not introduce different characteristics Vasil'ev [2004]. The RG recipe is applied to determine the scaling of the correlation functions.

3.2.2.2 The relevant action

The chemical reaction network for a branching random walk is



From which the action can be written down by the methods discussed in Chapter 1

$$\mathcal{A}[z, \partial_z]_{BRW} = z(\partial_t + \Delta)\partial_z + (\varepsilon - \beta)z\partial_z - \beta z^2\partial_z \quad (3.16)$$

To study the trace of the superprocess, in addition, there is a conditional deposition of B particles, which generates the trace (range) of the superprocess. As A particles diffuse, if they encounter a site they deposit a B particle *but* only if there is no B particle present i.e. the site has not been visited before. In the literature there are a number of suggestions for managing such exclusion principles (site restrictions), which might be described as particles having a *fermionic* property. Täuber et al. [2005] suggests a method that would be compatible with the formulation as pure Poisson processes in the standard proclivity, which is to introduce a coagulation $B + B \rightarrow B$. van Wijland [2001] introduces an alternative method.³⁹ Another approach used in Nekovar and Pruessner [2016] is what I will refer to in this chapter as *the carrying capacity trick*. In the master equation, constant rates are replaced by the following map

$$\rho(n_b) = n_a(c - n_b)|_{c=1} \quad (3.17)$$

This map has the property that it makes the rate vanish when the number of B particles reaches capacity $c = 1$. Now in the master equation, consider the conditional reaction $A \xrightarrow{?} B$ implemented by the map $A \xrightarrow{\lambda\rho(n_B)} B$. From first principles, the master equation is

$$\partial_t P(n_a, n_b; t) = n_a(1 - (n_b - 1))P_t(n_a, n_b - 1) - n_a(1 - n_b)P_t(n_a, n_b) \quad (3.18)$$

Using the generating function method and going to fields $[\phi_a, \tilde{\phi}_b]$ the generator for the *ordinary* generating function is

$$\mathcal{Q} = -\phi_a^\dagger \phi_b + \phi_a^\dagger \phi_a \phi_b^\dagger \phi_b + \phi_a^\dagger \phi_a \phi_b^\dagger - \phi_a^\dagger \phi_a \phi_b^{2\dagger} \phi_b \quad (3.19)$$

To recover the factorial moment generator, make the substitution $\phi^\dagger = 1 + \tilde{\phi}$. For example first factoring *number operators* to expose a pair of $\tilde{\phi}_b$'s $= (\phi_b^\dagger - 1)$'s we can write this as

$$\mathcal{Q} = \phi_a^\dagger \phi_a (\phi_b^\dagger - 1) - \phi_a^\dagger \phi_a \phi_b^\dagger \phi_b (\phi_b^\dagger - 1) \quad (3.20)$$

and after some algebra the following 6 terms can be added to the action for BRW

$$-\mathcal{Q}_\Lambda = \Lambda \begin{bmatrix} +\tilde{\phi}_b \phi_a \\ -\tilde{\phi}_b \phi_b \phi_a \\ -\tilde{\phi}_b^2 \phi_b \phi_a \\ -\tilde{\phi}_b^2 \tilde{\phi}_a \phi_b \phi_a \\ -\tilde{\phi}_b \tilde{\phi}_a \phi_b \phi_a \\ +\tilde{\phi}_b \tilde{\phi}_a \phi_a \end{bmatrix} \quad (3.21)$$

where I have at this point managed the signs, consistent with the *semigroup* Equation (1.7). Terms in (3.21) describe events where A and B particles interact at some rate given by the rate matrix Λ . The complete action is represented by the primitive corollas in Figure 3.3. Respecting inequality (3.11), we restrict from the full set of corollas given in Figure 3.3. There are two vertices of interest having ≤ 3 external legs for the observables corresponding to the *production of B particles* i.e. the statistics of the trace. The interesting part of the interaction Lagrangian therefore contains the terms - observe signs coming from (3.21) -

$$\lambda_0 \times \text{---} \bullet \text{---} \bullet \text{---} \quad -\lambda_1 \times \text{---} \bullet \text{---} \bullet \text{---} \quad (3.22)$$

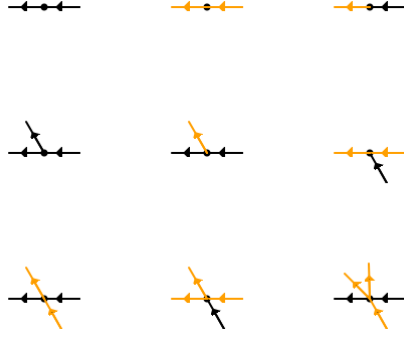


Figure 3.3: *Corollas (amputated vertices) of the Branching Wiener Sausage:* The propagators correspond to the free part of the Lagrangian and the interaction terms are used in a perturbative expansion.

Of the diagrams that are generated by the primitives, there are actually only *two* graphs, which both map to *one* relevant one-loop integral (by loop isomorphism at vanishing external momentum). Taking note of the corollas of which they are composed, the only two relevant graphs are given below with their residues;

$$\Re \left(\text{Diagram 1} \right) = \text{Corolla 1} \quad (3.23a)$$

$$\Re \left(\text{Diagram 2} \right) = \text{Corolla 2} \quad (3.23b)$$

Evaluated for $\hat{k} = \hat{\omega} = 0$ these graphs, which are loop isomorphic, evaluate to the same Feynman amplitude. This amplitude is used to renormalize both interactions. From Chapter 2, result (2.108) gave an example of an integral of interest at $d_c = 4$. Using this result there are two integrals for couplings λ_0 and λ_1 both of the form

$$\mathcal{I}_\lambda(\text{Loop}) = \lambda \cdot \frac{\mathcal{A}_d}{\mathfrak{s}(G)} \cdot \Gamma(\varepsilon/2) [1/\widetilde{\mathcal{M}}]^{\varepsilon/2} \quad (3.24)$$

for

$$\mathfrak{s}(\text{Loop}) = 2, \quad \mathcal{A}_d = (4\pi D)^{-d/2}|_{d=d_c=4} \quad (3.25)$$

where λ is the respective coupling rate for each of the corollas. As discussed in the next section, one seeks the *Z factors* for each coupling. From inspection of the graphs in (3.23), the renormalization of the vertex coupled to λ_0 depends on the renormalized vertex coupling λ_1 . Using a *minimal subtraction* scheme outlined in the next section, we find that the anomalous dimension yields $\gamma_\lambda = \varepsilon$ for $\lambda_R = \lambda_R^*$ and the dynamical critical exponent is $z = 2$ [Bordeu et al., 2019].

3.2.2.3 Minimal subtraction and RG functions

A renormalised coupling is introduced along with an arbitrary mass scale $\mu^{-\varepsilon}$

$$\lambda_R \leftarrow \lambda \mu^{-\varepsilon} \quad (3.26)$$

To renormalise, an ε -expansion and a replacement of the rate parameter with the renormalized one (3.26) yields

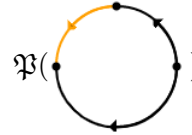
$$\mathcal{I}_\lambda = \mu^\varepsilon \lambda_R \cdot \frac{\mathcal{A}_d}{\mathfrak{s}(G)} \cdot \left[\frac{2}{\varepsilon} + \psi(1) + \mathcal{O}(\varepsilon) \right] \cdot \left[1 + \frac{\varepsilon}{2} \log \widetilde{\mathcal{M}}^{-1} + \mathcal{O}(\varepsilon^2) \right] \quad (3.27a)$$

$$= \mu^\varepsilon \lambda_R \cdot \frac{\mathcal{A}_d}{\mathfrak{s}(G)} \cdot \left[\frac{2}{\varepsilon} + \psi(1) + \log \widetilde{\mathcal{M}}^{-1} + \mathcal{O}(\varepsilon) \right] \quad (3.27b)$$

The expansion of the Gamma function uses (B.8).

Remark 38. It is worth emphasising that in this case, what $\widetilde{\mathcal{M}}$ actually is does not necessarily affect the value of the terms for $\varepsilon^n, n < 0$ which is what we are after in determining ε -poles in the expression.

The purpose of this exercise is to isolate poles in ε . In any renormalization scheme, some operator $\mathfrak{P}$ say, can be used to identify the divergent part of the Feynman integral. Under the *minimal subtraction scheme* (cf. Kleinert and Schulte-Frohlinde [2001], ch. 9)

$$\mathfrak{P}(\text{diagram}) = \mu^\varepsilon \lambda_R \cdot \frac{\mathcal{A}_d}{\mathfrak{s}(G)} \cdot \frac{2}{\varepsilon} \quad (3.28)$$


where I keep track of the symmetry factor in this expression as a convention. These terms have the form

$$\sum_{n=-k}^{-1} A_n \varepsilon^n \quad (3.29)$$

and the Z factors i.e. the renormalization constants (3.4) at the critical dimension $d_c = 4$ for $\lambda \in \{-\lambda_0, -\lambda_1\}$ for the branching and coagulation terms are

$$Z_\lambda = 1 + \frac{1}{(2\pi D)^2} \cdot \frac{\lambda \mu^{-\varepsilon}}{\varepsilon} \quad (3.30)$$

The Z factors are used in the determination of the RG function values and the renormalization group equations. Some well-known results can be found in [Zinn-Justin, 1996, Vasil'ev, 2004, Kleinert and Schulte-Frohlinde, 2001] but here I note that the *beta* function in first-order perturbation theory in these diagrams will have the form

$$\beta(\lambda) = \varepsilon \lambda_R + B \lambda_R^2 \quad (3.31)$$

with a non-trivial fixed point at

$$\lambda_R^* = \varepsilon / B \quad (3.32)$$

3.2.2.4 Scaling for superprocess on graphs

When discussing dimensional analysis above, I noted the scaling behaviour of the general vertex (a monomial in the field operators). We are interested in the scaling of (the statistics of) the trace of the superprocess. The observable in words is, *suppose we create A particles at sites - what is the probability of observing n correlated B particles*. Following the dimensional analysis, by scaling arguments [Bordeu et al., 2019], the scaling of n^{th} moment of the observable a (distinct sites visited) can be determined from the general vertex at the normalization point, parametrised with $\tilde{m} = 1$ for the creation of an A particle for the probing of n B particles.

$$\lim_{t \rightarrow \infty} \langle a^n \rangle(t) \propto L^2 L^{nd-4} \Gamma \begin{bmatrix} 0 & 1 \\ n & 0 \end{bmatrix} (1, D, \widehat{\{\lambda\}}; \{0, 0\}) \quad (3.33)$$

The scaling below ($\varepsilon > 0$) and above ($\varepsilon < 0$) the critical dimension $d_c = 4$ are given by

$$\lim_{t \rightarrow \infty} \langle a^n \rangle(t) \propto \begin{cases} L^{nd-2} & \varepsilon > 0 \\ L^{n4-2} & \varepsilon < 0 \end{cases} \quad (3.34)$$

Below the critical dimension, depending on the effective diffusion rate and its relation to *finite* lattice size, and assuming $nd - 2 > 0$,

$$\langle a^n \rangle(t, L) \propto \begin{cases} t^{(nd-2)/2} & Dt \ll L^2 \\ L^{nd-2} & Dt \gg L^2 \end{cases} \quad (3.35)$$

In [Bordeu et al., 2019] we consider the volume explored by a superprocess on general graphs. This is of both *theoretical* and *practical* interest, for example in understanding the spread of disease on social networks. Unlike diffusion on d -dimensional hypercubic lattices, with $d \in \mathbb{Z}$, graphs such as scale-free graphs or real world graphs such as social networks have non-integer dimension. To extend the scaling results to such graphs, one replaces d as used on regular lattices with the d_s , the graph's *spectral dimension* [Lindenberg et al., 1991, Burioni and Cassi, 2005], §C.5. This use of the spectral dimension is valid in the absence of an anomalous dimension.[‡] For a graph with non-integer dimension and L vertices, the moment-scaling with respect to the *trace of the superprocess* is found to be (with $L = \hat{L}^{1/d_s}$ giving the network size)

$$\langle a^n \rangle(t, \hat{L}) \propto \begin{cases} t^{(nd_s-2)/2} & \text{for } Dt \ll \hat{L}^{2/d_s} \\ \hat{L}^{(n-2/d_s)} & \text{for } Dt \gg \hat{L}^{2/d_s} \end{cases} \quad (3.36a)$$

In [Bordeu et al., 2019] we compared this analytical result with simulations for a selection of graphs and fractal lattices for which the spectral dimension is known (*preferential attachment* graph, random tree, Sierpinski carpet). Results are tabulated at the end of this chapter and a description of simulation setup follows immediately below.

[‡]We remind ourselves that the existence of an anomalous dimension follows from a renormalization of the propagator and the propagator was derived from the Laplacian operator which describes the diffusion of particles on graphs.

3.2.2.5 Numerics and graph generation algorithms

The superprocess consists of a branching, diffusion and extinction. A particle either diffuses at rate h , branches with rate β or becomes extinct with rate ϵ . I focus only on the critical case, which is at $\beta = \epsilon$. In the current implementation parameters are chosen such that $h + \beta + \epsilon = 1$ and h is chosen to be high e.g. 0.95. With this normalization we can treat rates as though they are probabilities. Each of these subprocesses is a Poisson process. As the system evolves, an exponential waiting time for an event is simulated in the standard way as $-\ln(1 - U)/N$ where $U \sim U(0, 1)$ and N is the size of the population. Time is stored as a `long double`. The process is *bosonic*. When branching occurs, offspring are added *locally* i.e. to the same site as the parent.

Regular Lattice For regular lattices we typically choose system sizes of $2^n - 1$ in order that the centre of the lattice is an integer coordinate where an initial particle is placed. An upper bound can be set that avoids excessive memory usage in higher dimensions. For example, simulations in two dimensions may safely exceed $L = 2^{10}$ while three and five dimension simulations may struggle depending on hardware. One might expect to use $L \leq 63$ for 5D simulations on contemporary hardware. The d -dimensional lattice is mapped to a one dimensional data structure using a mapping $f : \mathbb{R}^d \rightarrow \mathbb{R}$,

$$f(x_0, x_1, \dots, x_{d-1}) = \sum_{d=0}^{D-1} x_d L^d \quad (3.37)$$

With this map, a particle that hops 1 site in any direction in $\mathbb{R}^D$ jumps $\pm L^d$ in $\mathbb{R}^1$ for a given component axis.[§]

Remark 39. The inverse mapping $f^{-1}(i, d)$ exists for extracting component values from a given integer $i < L^D$.

Each site is mapped onto a *hierarchical* partitioning of L^D and plays a role in implementing boundary conditions. A bitmask is used to close boundaries on given axes. For example setting bits 0,1 closes the first component axis at both ends to create a toroidal axis (as opposed to closed/reflecting boundaries). It is illegal for a particle to hop between partitions. When a jump is computed along an axis such that $i' = i + L^d$, it is possible to check the condition $h(i, d) \stackrel{?}{=} h(i', d)$, where h gives the partitions for i and i' .

$$h(i) = \left\lfloor \frac{i}{L^d} \right\rfloor, \quad d = 1, \dots, D \quad (3.38)$$

Considering the mapping $\mathbb{R}^d \rightarrow \mathbb{R}$, particles can move in one direction in any plane. Care is taken to check that a particle moving in $\mathbb{R}$ does not jump between different partitions illegally.

Sierpinski Fractal Lattice The Sierpinski fractal $\mathcal{S}_{b,m}^{(d)}$ is implemented in d dimensions. In 2 dimensions it is called a carpet. It is parametrized by b and m and like fractals in general,

[§]Conventionally we assign even numbers in $2D$ to positive(forward) moves and odd numbers to negative (back) moves

it is generated recursively. In this case the rule is to partition a lattice or sub-lattice into p^2 partitions and erase $p^2 - m$ partitions. In these simulations, parameters b, m are restricted to certain values by convention and a lattice with sides that are powers of b is chosen. The Sierpinski carpet structure is illustrated in Figure 3.4.

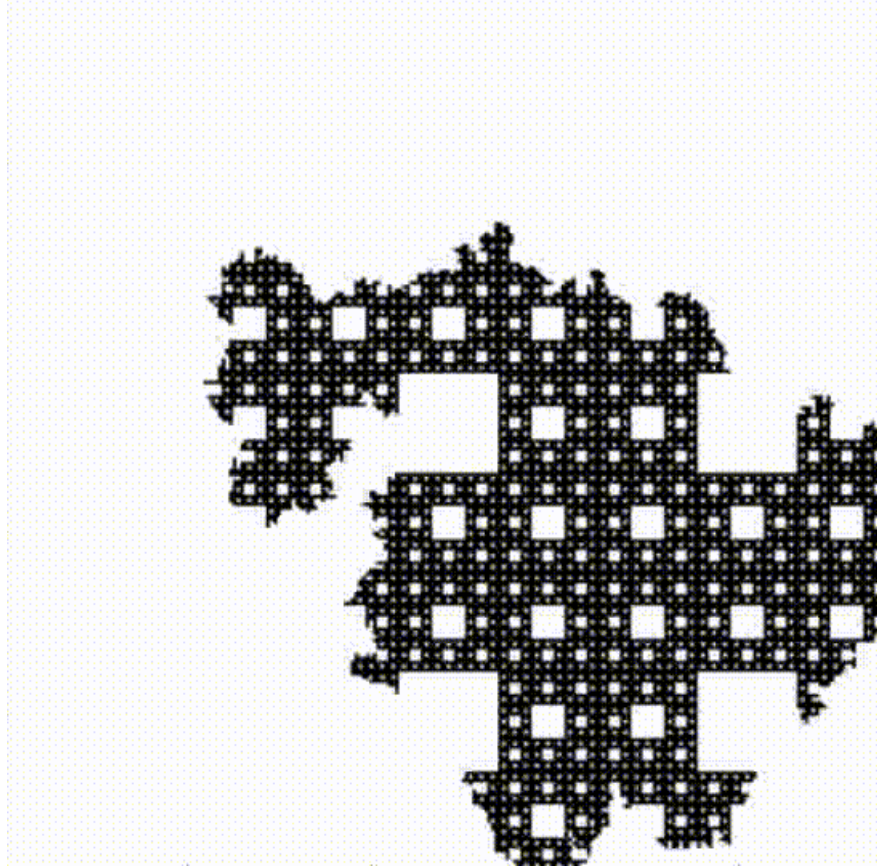


Figure 3.4: *superprocess on Sierpinski carpet:* A realization of the superprocess on a Sierpinski carpet generated with parameters $b = 3, m = 8$. Visited sites are marked black.

In the numerical implementation, the fractal lattice $\mathcal{S}_{3,8}$ is a simple modification of a regular lattice. The overall structure does not change but *holes* (inaccessible sites in the graph) are placed in the lattice. As such, a walker is reflected from regions of the plane that are not *accessible*. The existing boundary conditions for the extremities of the lattices are retained (open boundaries are used) but holes are *strictly inaccessible*. For any site i an accessibility rule is required

$$g(i) \rightarrow [0, 1] \quad (3.39)$$

For example when choosing $b = 3, m = 8$, the middle tile out of 9 tiles are removed. Given the recursive nature of the fractal, we can consider partitions at various zoom levels scaling in powers of 3. A rule of the form given by (3.39) is defined below. It determines if a particular lattice site falls in a hole of the fractal (in this example, the central tile) for *any* zoom level. Over all zoom levels, there is a k -hierarchy of lattices with $p = b^k$ tiles at each level. Now use the notation $p = b^k, p' = b^{k-1}$. We can check that the following

equality holds for *any* p

$$f_p(i) := \left\lfloor \frac{x_d(i) \bmod p}{p'} \right\rfloor = \lfloor b/2 \rfloor, \forall d \quad (3.40)$$

This defines if a site on $\mathcal{S}_{3,8}$ is inaccessible where the boolean result is understood to map to $[0, 1]$ and a site indexed i is mapped onto d components in D dimensions via $x_d(i)$. So, for a Sierpinski fractal on a hypercubic lattice one must check that this condition holds *for all* c but *for any* p . Initialising the process the walker is placed at a *random* direct offset from the centre on the open boundary of the middle square, which is itself, a hole.

Random Trees A random tree is a graph generated by a *critical* Galton-Watson-Bienaymé branching process X_t^N conditioned to have N vertices. Resulting structures are a graded subspace on the space of pruned, balanced binary trees i.e. for all balanced binary trees with N vertices, the critical process produces those trees of size N vertices for which (asymptotically) half of the vertices are *terminal* vertices (0 children) and half have 2 children.



Figure 3.5: *Some simple random trees:* Random trees are grown by a critical branching process. One way to think of the random tree is to start with a perfectly balanced binary tree with $2N + 1$ vertices. Visit each vertex and flip a fair coin; if the coin lands on heads, remove the sub-tree rooted at that vertex.

The process X_t^N , a sequence of branching and non-branching events, can conveniently be represented as a binary sequence. The algorithm to generate the binary tree can be understood as two conceptual stages. (i) In the first stage, we sample a realisation of the critical branching process (a binary string). This representation of the binary tree must, (ii) in the second stage, be mapped onto a suitable adjacency matrix structure that will efficiently embed the superprocess.[¶] For large graphs/large processes, an optimal adjacency structure generally consists of 1D direct-memory-access registers. There is one subtle part to generating such structures in an efficient single-pass algorithm that does not consume excessive memory. The two conceptual algorithm stages (implemented in a single pass in practice) are described below.

(i) The conditioned binary string is generated in the following way. The process X_t^N taking values in $\{0, 1\}$ is conditioned on $\sum X_t^N = B = (N - 1)/2$ branching events.^{||} The conditioned process will *visit* N vertices. The upper bound is checked during the process by making sure that we never branch $B + 1$ times. The lower bound (that the process does not die before generating enough vertices) can be checked by introducing a mapped process $\hat{X} : \hat{X}_i \in \{-1, 1\}$. Starting at 1, we ensure that the $E = 1 + \sum \hat{X}_t \geq 0$. In practice, the first check is applied when branching while the latter is applied when not branching. This simple algorithm samples *critical binary sequences* from X_t^N .

[¶]Some structures are slower for performing random walks than others!

^{||}The reader can check that is true for any binary tree

(ii) Having generated a valid critical binary sequence (tree representation), the sequence must be mapped onto an efficient adjacency matrix representation of the random tree. Three separate structures are used to represent the adjacency matrix. Some of these can be read off from the binary sequence, or during the mapping process using a single state variable but the general adjacency structure takes some care. To see the problem, consider for example a general binary tree $((...), (...))$ where the parse order (prefix, postfix or infix) does not alter the general argument. This corresponds to a root node with two subtrees of arbitrary size. The index for the left subtree in the adjacency matrix may immediately follow the root's partition in the adjacency matrix while the right subtree will have an arbitrary position that is not known until the left tree is fully resolved.** Some bespoke stack-like structure is typically required in an optimal solution and here a First In First Out (FIFO) queue turns out to be appropriate.

On every branch event, we can enqueue *two instances* (one for each resolved child) of the current node id to the front of the queue structure. These will be consumed by child vertices which dequeue from the queue. This is a FIFO structure. In the *breadth-first* generation of a binary tree, it so happens that a layer of parent ids are queued (from left to right) at layer d and systematically dequeued by children (from left to right) at layer $d + 1$. In a FIFO structure, child vertices that dequeue parent ids may safely branch and enqueue their own id ($\times 2$) for the next layer. Having dealt with this issue, the rest of the algorithm is straight-forward. During the generation of the tree, all visited vertices can populate their *degree* and *offset* entries, while also adding parent ids (half-edges) to their own partition in the adjacency matrix. They also update their parent's partition (the offset of which they can look up) with the *dual* half-edge. Ambiguously, the parent partition has two half-edge slots but these can be distinguished using the (child's) id parity (even left, odd right or vice versa).

Remark 40. Ergodicity test: An important test for non-regular adjacency matrices is to ensure a single walker eventually visits all vertices with probability 1 for all of some k trials.

A note on waiting times and random variables In all simulations the SIMD-orientated Fast Mersenne Twister random number generator [Saito and Matsumoto, 2008] is used to generate random numbers. Poisson waiting times are simulated for inter-event waiting times and data are flushed deterministically according to a write time

$$W(i) = t_0 \left(\frac{t_\infty}{t_0} \right)^{\frac{i}{n_b - 1}} \quad (3.41)$$

where t_0, t_∞ are parameters for minimum and maximum timescales set to $10^{-1}, 10^9$. The number of bins, n_b is set to 1000.

When drawing random numbers from the generator and projecting to a range n , one avoids introducing bias due to rounding via modulus errors in the following way. Given a range maximum of the generator g_∞ , resample from the generator in the rare case where

$$x \geq g_\infty - (g_\infty \bmod n) \quad (3.42)$$

**It is tempting and possible to use recursion for any tree problem, which renders the issue a little less perplexing. But in general this does not lead to optimal code since one loses control of the stack.

3.3 Results

The scaling in time follows

$$\langle a^n \rangle \sim t^{\alpha_n} \quad (3.43)$$

for regular graphs with integer dimension and non-regular graphs with $d \rightarrow d_s$ appears in the following two tables.

Scaling of visited sites by BRW in time on hypercubic lattice

exponent	d=1		d=2		d=3		d=5	
	numerics	theory	numerics	theory	numerics	theory	numerics	theory
α_1					0.47(2)	1/2	1.0(2)	1
α_2			0.98(3)	1	2.0(1)	2	2.9(3)	3
α_3	0.48(4)	1/2	2.0(1)	2	3.5(1)	7/2	4.8(4)	5
α_4	1.0(1)	1	2.9(1)	3	5.0(2)	5	6.7(7)	7
α_5	1.5(1)	3/2	3.9(1)	4	6.4(2)	13/2	9(1)	9

Table 3.8: *Scaling of visited sites in time by BRW on hypercubic lattice:* Scaling shown for lattices in dimension above and below the critical dimension $d_c = 4$. Comparisons are made between theory and numerical simulation.

Scaling of visited sites by BRW in time on non-regular graphs

exponent	Sierpinski		Random Tree		PA Graph	
	numerics	theory	numerics	theory	numerics	theory
α_1					1.0(1)	1
α_2	0.81(5)	0.86	0.35(7)	1/3	2.8(2)	3
α_3	1.71(7)	1.79	0.9(1)	1	4.8(4)	5
α_4	2.62(10)	2.72	1.6(2)	5/3	6.6(5)	7
α_5	3.54(14)	3.66	2.2(4)	7/3	8.5(9)	9

Table 3.9: *Scaling of visited sites in time by BRW on non-regular graphs:* Scaling is shown for the Sierpinski carpet, random tree and preferential attachment scale-free graph. The spectral dimensions for these graphs are, for Sierpinski carpet $d_s \approx 1.86$, for the random tree $d_s \approx 4/3$ and for the scale-free graph generated by the preferential attachment algorithm $d_s \approx 4$. Comparisons are made between theory and numerical simulation.

Scaling in system size follows, where b is the system size or number of particles,

$$\langle b^n \rangle \sim L^{\beta_n} \quad (3.44)$$

for regular graphs with integer dimension and non-regular graphs with $d \rightarrow d_s$ appears in the following two tables.

Scaling of visited sites by BRW as function of system size on hypercubic lattice

exponent	d=1		d=2		d=3		d=5	
	numerics	theory	numerics	theory	numerics	theory	numerics	theory
β_1					0.97(4)	1	1.9(2)	2
β_2			2.1(2)	2	3.9(1)	4	5.7(5)	6
β_3	0.96(4)	1	4.2(3)	4	6.8(2)	7	10(1)	10
β_4	1.93(7)	2	6.1(3)	6	9.8(3)	10	14(2)	14
β_5	2.93(8)	3	8.1(4)	8	12.7(4)	13	17(3)	18

Table 3.10: *Scaling of visited sites by a BRW as function of the system size on hypercubic lattices:* Scaling shown for lattices in dimension above and below the critical dimension $d_c = 4$. Comparisons are made between theory and numerical simulation.

Scaling of visited sites by BRW as function of system size on non-regular graphs

exponent	Sierpinski		Random Tree		PA Graph	
	numerics	theory	numerics	theory	numerics	theory
β_1					0.49(1)	1/2
β_2	1.9(1)	1.72	0.58(6)	1/2	1.49(1)	3/2
β_3	3.8(2)	3.59	1.6(1)	3/2	2.49(2)	5/2
β_4	5.7(3)	5.45	2.6(2)	5/2	3.49(2)	7/2
β_5	7.5(4)	7.31	3.7(2)	7/2	4.49(2)	9/2

Table 3.11: *Scaling of visited sites by a BRW as function of the system size on non-regular graphs:* Scaling is shown for the Sierpinski carpet, random tree and preferential attachment scale-free graph. The spectral dimensions for these graphs are, for Sierpinski carpet $d_s \approx 1.86$, for the random tree $d_s \approx 4/3$ and for the scale-free graph generated by the preferential attachment algorithm $d_s \approx 4$. Comparisons are made between theory and numerical simulation.

These results show strong agreement between theory and numerical simulation for regular lattices for both system size and scaling time. For non-regular graphs the agreement is

reasonable but not as good in all cases. For example the scaling of visited sites shows strong agreement between numerics and theory for the scale free graph (preferential attachment) but the agreement is not as strong for the Sierpinski fractal lattice. For performance reasons a relatively small fractal lattice was used, which may be a factor in this discrepancy.

In Chapter 1 it was stated that perturbation theory (Feynman diagram expansions) and the renormalization group are considered powerful tools in making predictions about interacting particle systems. In high energy physics, theory has been shown to match physical experiment to very high accuracy. In the theory of critical phenomena, the only viable validation method is numeric simulation. This thesis focused on some of the algebraic of carrying out this procedure in previous chapters. In this chapter, the final stages in the renormalization group analysis for the trace of a superprocess was outlined and techniques to simulate processes on different graphs were discussed. Having determined relevant diagrams, the renormalized vertex functions can be used to make scaling arguments about observables of interest. The predicted scaling behaviour has been verified against stimulations on different graphs.

4 Empirical studies with biological data

This chapter focuses on methods for extracting knowledge from volumetric time-lapse microscopy data. The material reflects my collaborative work in biology and biophysics [Duarte et al., 2018, Amarteifio et al., 2018].

I begin in the next section discussing image processing theory and practice. Theoretically, image processing is a rich topic involving signal processing, harmonic analysis, machine learning and mathematical morphology. In practice, image processing can be rather heuristic, an inevitability at an unpredictable data interface. While the boundaries between image processing and tracking methods are blurred in practice, for the purpose of discussion, suppose that the image processing stage determines concrete object shapes and locations, which are propagated on to the tracking phase. Having configured and applied the typically non-trivial image processing and object tracking pipeline, one finally extracts spatiotemporal data. At this stage various methods such as time series or clustering analysis can be used to describe the data. From the perspective of (statistical) physics one may also wish to model particles as Brownian particles and consider various diffusive regimes to characterise behaviour.

The first part of this chapter discusses image processing and tracking methods applied to a plant biology problem while the second part of the chapter uses anomalous diffusion and techniques from physics to analyse cancer cell motility given 3D track data.

4.1 Overview of image processing

I define an image, which will be assumed to be a microscopy image, as a five dimensional object, with 4 space-time dimensions and a colour channel;

$$\mathcal{F} := (t, z, y, x, c)$$

I will avoid directly discussing microscope technology, adding occasional comments to the end notes and focus directly on processing the raw image data. Occasionally we will take cross-sections of the image for example a 3D grey-scale image at a particular point in time. An image is stored on disk in a file format such as TIFF. An example of the dimension of a microscopy image is $|\mathcal{F}| = (64, 64, 800, 1024, 3)$. Typically the aspect-ratio is such that the resolution in the z -axis is much lower than in the x and y axes. Typically there are three colour channels such as red, green, blue. The time dimension for imaging methods such as *confocal microscopy*, sequenced over several hours every 15 minutes for example, may often be smaller than for *lightsheet microscopy*,⁴⁰ sequenced for up to a week in order to produce in the region of 500 images. A significant part of image processing involves the

convolution of images with *kernels* or *structural elements*, referred to as *filtering*;

$$\mathcal{F}' = \langle \kappa, \mathcal{F} \rangle$$

Image processing involves chaining together sequences of filtering and threshold operations

$$\begin{aligned}\mathcal{F}^{(1)} &\leftarrow \langle \kappa, \mathcal{F}^{(0)} \rangle_{\theta} \\ \mathcal{F}^{(2)} &\leftarrow \langle \kappa, \mathcal{F}^{(1)} \rangle_{\theta} \\ &\vdots \\ \mathcal{F}^{(n)} &\leftarrow \langle \kappa, \mathcal{F}^{(n-1)} \rangle_{\theta}\end{aligned}$$

The final image $\mathcal{F}^{(n)}$ may correspond to discretely *labelled* objects such as *blob centroids*, which can be passed to tracking algorithms.

Linear shift-invariant filters

Filters are the main workhorses in basic image processing. I derive some of these operators here. Of course, for the image processing practitioner, it is often completely unnecessary to consider such derivations. But for what follows, it is useful to observe (i) the role of convolution and (ii) how continuous test functions are discretized in image processing.

A time-invariant or shift-invariant linear filter $\mathcal{L}$ is equivalent to a convolution with an impulse response κ . for a continuous function f , its value at t is given by its response to a Dirac impulse; $f(t) = \int f(u)\delta(u-t)du$. In discrete terms, a signal can be seen as a superposition of responses to shifted delta functions. For a continuous, linear operator, we can identify the impulse response $\kappa(t)$

$$\mathcal{L}f(t) = \int f(u) \underbrace{\mathcal{L}\delta(u-t)}_{\kappa(t)} du \quad (4.2)$$

from which it follows by the convolution theorem that

$$\mathcal{L}f(t) = \int \kappa(u)f(t-u)du = \kappa * f(t) \quad (4.3)$$

Here κ is sometimes called the *point spread function* and its Fourier transform is the *transfer function*. The transfer function used in the signal processing context is more generally known as a multiplier e.g. Fourier multiplier. Image processing operations can equivalently be applied in spatial or frequency domain. Motivations for choosing one or the other are discussed briefly in the notes.⁴¹ In discrete image processing, a convolution operation is defined over a *neighbourhood*, for a kernel of size u . The convolution operation is depicted in Figure 4.1 together with examples of 1-dimensional kernels. In $2D$, a pixel at location i, j is updated according to,

$$\mathcal{R}_{ij} = \sum_{u,v} \kappa_{i-u,j-v} \mathcal{F}_{uv} \quad (4.4)$$

where $i-u, j-v$ are displacements in the neighbourhood of i, j . Here κ is the kernel and $\mathcal{F}$ is an image. $\mathcal{R}$ is the filter response. Kernels are relatively small with respect to the

image (e.g. $3 \times 3, 5 \times 5, 7 \times 7$). By convolution each pixel becomes a linear combination of its neighbouring pixels. Each pixel is "spread out" over the neighbourhood hence the term *point spread function*.

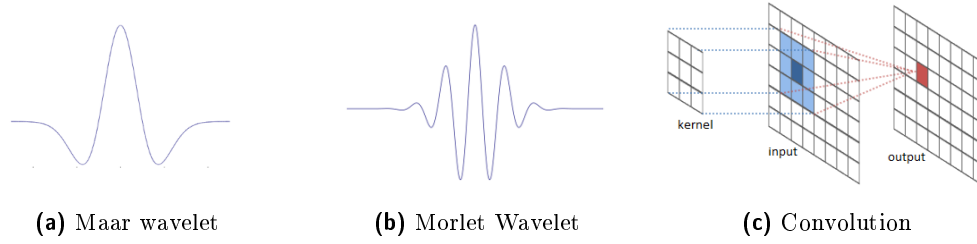


Figure 4.1: *Convolution* and some example kernels

Some discrete kernels

The **Laplacian** Δ measures the difference between $f(x)$ and the average of the value of f in the neighbourhood. By definition, $\Delta f = \sum \frac{\partial^2 f}{\partial x_i^2}$. In one dimension, from Taylor's theorem we can write the second derivative

$$\frac{d^2 f}{dx^2} = \lim_{h \rightarrow 0} \frac{1}{h^2} [f(x+h) + f(x-h) - 2f(x)] \quad (4.5)$$

This can be seen by considering the expansions for changes $(\pm h)$ on each axis

$$f(x \pm h) = f(x) \pm f'(x)h + f^{(2)}(x)\frac{h^2}{2} \pm f^{(3)}(x)\frac{h^3}{3!} + \dots \quad (4.6)$$

Adding these, alternating terms (odd derivatives) cancel. Rearranging, in the limit $h \rightarrow 0$ we have (4.5). The discrete approximation (extrapolating from 1D) on a lattice with coordination number q is

$$\Delta f = f(x+1) + f(x-1) - 2f(x) \rightarrow \sum_{\mathbf{e}}^q f(x+\mathbf{e}) - f(x) \quad (4.7)$$

For the (isometric) Laplacian kernel in 2D, the discrete representation follows,

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.8)$$

The **Gaussian** kernel approximates a continuous Gaussian function. For example, with a normalisation constant as the sum of cell values,

$$\frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 6 & 24 & 6 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 6 & 24 & 6 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (4.9)$$

The **Laplacian of Gaußian** or Maar wavelet (alias *Mexican Hat Wavelet*) is shown in Figure 2.1(a) for the 1D case. The continuous function

$$\Delta G = \frac{\partial^2}{\partial x^2} G + \frac{\partial^2}{\partial y^2} G = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} e^{-(x^2+y^2)/2\sigma^2} \quad (4.10)$$

can be approximated with a discrete kernel such as shown below. In practice, a σ parameter controls the aperture.

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & -16 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (4.11)$$

The Maar wavelet is used for detecting edges and blobs. For performance reasons an approximation may use the differences of two Gaußian filters with different σ values. This is a slightly less accurate filter but more efficient to compute.

Remark 41. In the blob detection routines described for data in this chapter, I use this *difference of Gaußians* approach for blob detection over the 3D volume, while ΔG with a large σ value is used to segment the root in 2D before projecting back to a 3D mask.

4.2 Lightsheet microscopy data

A single *Arabidopsis* primary root is grown and imaged on a custom-made light-sheet microscope setup [Sena et al., 2011, Baesso et al., 2018]. The root is hydroponically grown in a *perfusion chamber* maintained under constant light and constant temperature, with its liquid medium fully exchanged every 2 minutes. A full 3D scan in fluorescence of the root tip is composed of 60 optical sections $4\mu\text{m}$ apart, captured every 15 minutes for up to 7 days. To visualize mitotic events, an established transgenic line is used, expressing a fusion between the cyclin protein CYCB1;1 and the fluorescent protein GFP [Reddy et al., 2004]. In these plants, the fluorescent reporter CYCB1;1::GFP accumulates in a cell transitioning between G2 and M phases of the cell cycle, and is quickly degraded after entering mitosis. It is widely adopted as a reliable live marker for cell events.

Noise and histogram levels are used on normalised grey-scale images taking pixel values in $[0, 1]$. The noise shown in the top panel of Figure 4.2B is estimated using a wavelet-based estimator of the Gaußian noise standard deviation [Donoho and Johnstone, 1994, van der Walt et al., 2014]. The middle panel in Figure 4.2B describes the signal. This is a simple quasi-signal metric defined as a ratio between the 99th and 95th histogram percentile boundary values. If this value approaches 0 there is low signal/information. For values of this measure above a calibrated threshold, the data are more concentrated outside the 99th percentile of the data, corresponding to bright singularities in the image. A noise range is used to guide downstream strategies in the image processing pipeline. A normal noise band $[0.01, 0.045]$ is used. If the noise is very low e.g below 0.01, no denoising is required otherwise wavelet denoising will be applied. Excessive noise, exceeding 0.045, suggests light saturation and/or low signal-to-noise ratio typically due to a low number of biological

events. In such instances, additional thresholding is applied to the data before wavelet denoising. If the noise exceeds 0.1, the signal-to-noise ratio is so low that the frame is marked as degenerate. Figure 4.2A contrasts different noise levels in sample images and Figure 4.2B plots image properties over time for a given experiment.

Region of interest Bright blobs are scattered across the full image and may correspond to mitotic events within the root or debris floating beyond the root. For tracking accuracy and tracking speed, it is advisable to select a *region of interest* (ROI), drawing a box around the identified root. A 2D projection of the data is obtained by summing the 3D tensor along the z-axis. An adaptive threshold based on image histogram percentile ranges is used to construct a narrow-band filter for the 2D data. This range can vary erratically between frames as shown in the bottom time series in Figure 4.2B. A thin slice of the image data range (shaded band) is selected and an aggressive “Gaussian smoothing” (averaging neighbouring pixel values using a Gaussian kernel) with a large sigma value is used to find a mesh-like connected component corresponding to the region of activity in the root tip. The 2D mask of this largest component is extended to a 3D mask by projecting the 2D region back into the z-plane. The largest connected component in the thresholded image is identified as the root tip. These stages are illustrated in Figure 4.4. The image processing pipeline continues to process data only within the ROI. For cases where the region of activity is well-isolated (smaller region of interest), down-stream image processing is more efficient as the amount of volumetric data is reduced and debris beyond the root are filtered from the pipeline.

Object detection Segmentation plays a central role in many image processing pipelines and typically involves (i) thresholding and identifying background, (ii) using distance/gradient transformations with peak detection to identify markers and (iii) routines such as watershed [Beucher, 1994] to segment blob labels. For our data, the final segmentation will not be performed after identifying the peaks centroids. Given the variability in the data over a large frame sequence a simple “annealing thresholding” turns out to be effective. A threshold is increased iteratively to remove large connected components beyond a maximum perimeter length thus removing the *background* from the image. Being iterative, this is not as efficient as a linear filter but it is simple and robust to noise and variability in light-sheet data. In extreme cases where it is not possible to find a threshold level in this manner, the 99th percentile determines what is image *background* for removal. Having applied the threshold filter, centroids are identified (i) performing a difference of Gaussians to emphasise blob-like objects, (ii) applying a maximum filter and (iii) returning the coordinates of the local maxima (peaks) in the image. Part of this process is illustrated in Figure 4.6.

Remark 42. To allow for a fully automated routine that can cope with arbitrary datasets (in the scope of our light-sheet microscopy datasets) the emphasis in the centroid detection stage has been to avoid spurious centroid detections at the risk of under-sampling, while optimising for objects to be identified for at least two frames somewhere during the peak of their light intensity arc.

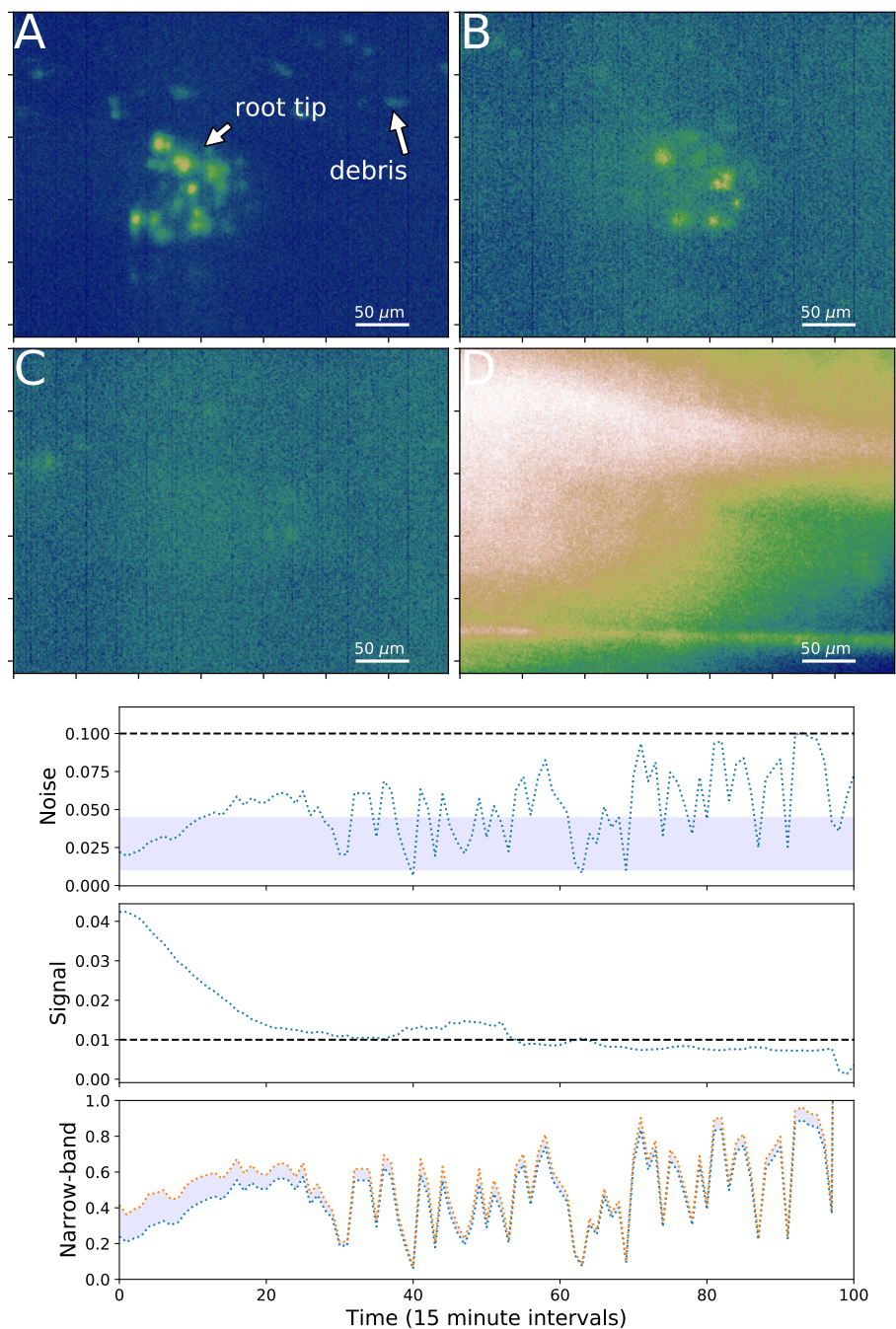


Figure 4.2: *Variation in frame properties:* The tiles top(a) show four sample frames. Frame A is a “good” image. Frame B is a noisy image. Frame C is a noisy image with a low number of objects and some noise. Frame D is over-saturated and will be marked as degenerate. On the bottom(b), the same story is told quantitatively over 100 consecutive sample frames from one imaging session. The *signal* and *noise* levels are plotted with threshold values. The *narrow-band* plot shows the region between the 95th and 99th percentile ranges of the image histogram. Notably, these values fluctuate dramatically.

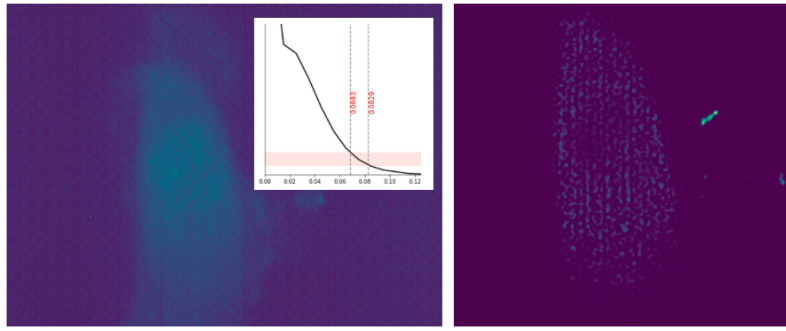


Figure 4.3: *Root Thresholding:* A rank-based adaptive threshold is used in conjunction with a Laplacian of a Gaussian to enhance large structures. Data intersecting the 95th and 99th percentile ranges are selected by band-pass filter.

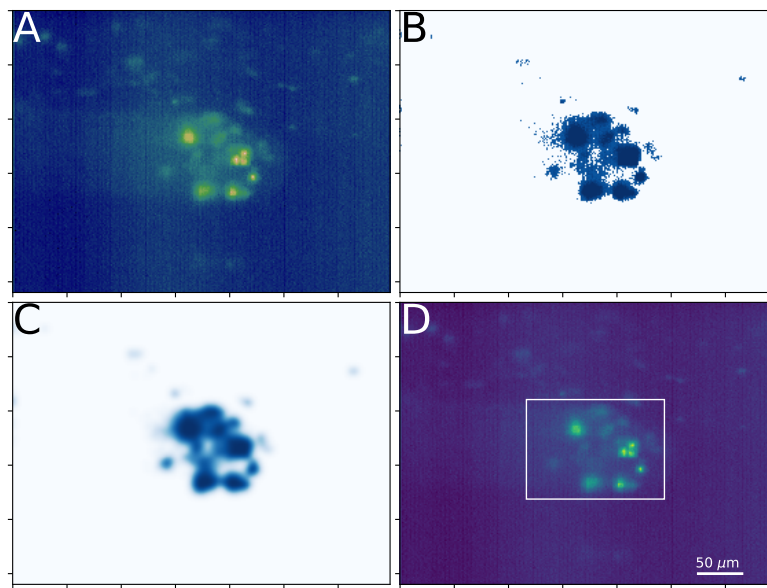


Figure 4.4: *Isolating the Region of Interest (ROI):* A narrow-band filter (B) and aggressive Gaussian smoothing (C) is used to find the largest connected component in the narrow band of the data corresponding to the region of activity in the root tip (D).

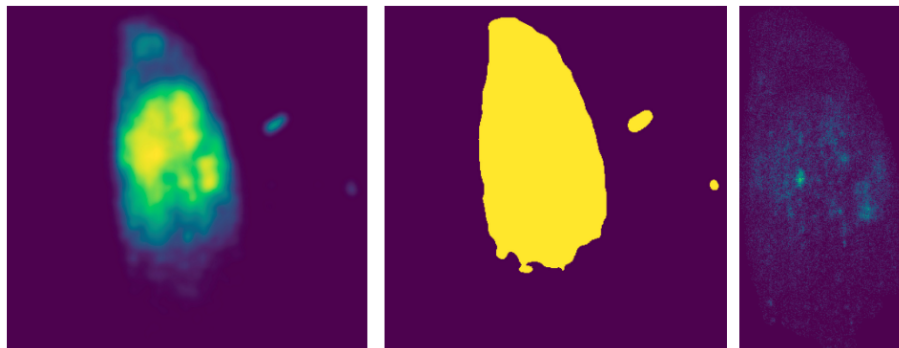


Figure 4.5: *Root masking:* A Gaussian filter is first used to ensure the integrity of the components (first image). A binary mask is constructed (middle image). The labelled giant component is used to construct a 3D mask or *Region of Interest*, which is used to isolate the root pixels in the original 3D data for further processing.

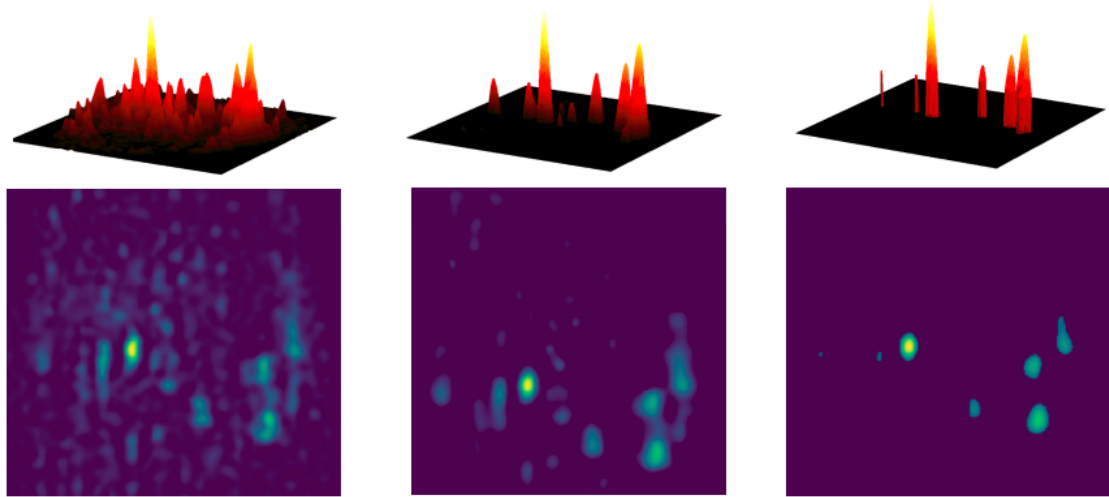


Figure 4.6: *Blob detection:* Illustration of blob detection using a *Difference of Gaussians*. The 3D projection emphasises how thresholding might then be used to keep peaks in the image.

4.3 Tracking mitotic events

Automated computer vision⁴² and tracking methods can match or exceed manual human-led tracking of objects in large image datasets. Due to the rate at which such data are being produced, these methods are of growing importance [von Wangenheim et al., 2017, Ulman et al., 2017]. The two types of data discussed in this chapter have been collected using fluorescence microscopy. On the one hand, the cancer cell data by *confocal microscopy* and on the other hand, the plant data by *light-sheet microscopy*. Fluorescence time-lapse microscopy typically exploits markers that emit light continuously for a period of time allowing for objects of interest to be tracked [Kanade et al., 2011, Wöll et al., 2013, Godinez and Rohr, 2015, Liu et al., 2017, von Wangenheim et al., 2017] notwithstanding occlusions or objects leaving and entering scenes. In this chapter I focus on tracking methods for (plant) data with *transient* fluorescent events where objects in the field of view fluoresce for a relatively short time to mark *events of interest*.

The two types of data discussed in this chapter have important differences from an image processing and tracking perspective. On the one hand we observe independently motile cancer cells, while on the other hand we observe plant cells in an almost-rigid growing tissue. Both data require tracking but the tracking requirements are quite different. In the cancer cell biology case, data are collected using fluorescent confocal microscopy where motile cancer cells are observed in an immobile specimen. This is the most common setup for single particle tracking experiments. A powerful class of stochastic tracking methods used in this case for example are *state space models*.⁴³ However in the plant biology case, effectively immobile cells are observed in a moving specimen; movement is due to actual root growth and experimental conditions i.e. the inability of the microscope to keep the specimen fixed in the field of view over extended imaging sessions. The tracking method discussed below considers this tracking problem.

The algorithm developed in this chapter is best understood as a hybrid concept combining stochastic *point-set registration* and *single particle tracking*. Single particle tracking

solutions may or may not require pre-registration. In the case that they do require pre-registration, point-set registration solutions may not be robust against lack of persistent points from frame to frame. In the case where single particle tracking methods are robust in the absence of pre-registration, they typically emphasise individual object motion models, which may not be optimal under certain data conditions *i.e.* transient events in rigid structures and low signal to noise ratio across extended images sequences.

Point-set registration [Besl and McKay, 1992, Fitzgibbon, 2003, Jian and Vemuri, 2011, Zhou et al., 2018] involves identifying *keypoint* correspondences in two frames. Keypoints may correspond to distinctive edges or blobs that exist in both frames. If points appear in only one frame they are treated as *outliers* in the registration problem. Given a keypoint correspondence (pairs of *inliers*), an affine transformation (*i.e.* a generalised transformation including translations, rotations, scalings) that takes one set of points to the other can be determined. The plant data contain few guaranteed inliers as fluorescence events have short life-times (several frames) and tissue structures may be difficult to identify.

Object *tracking* in time-lapse image data involves *detection* of objects in individual frames and solving the *data association* or *identity management* problem between frames. The goal of identity management is to link an object observed at one time point to what is found to be the same object observed at a later time point. Unique *identifiers* are propagated between married objects, *i.e.* objects determined to be the same object in each frame. The set of observations of a given object through time is referred to as its *lineage*. One early example of a tracking algorithm, *multiple hypothesis tracking* [Reid, 1979, Cox and Hingorani, 1996] takes an exhaustive, deterministic approach to consider possible lineage trees. Other examples may take a stochastic approach [Schulz et al., 2001], which can be more efficient and robust in the context of noisy data. Tracking approaches are greatly influenced by the nature of the data. For tracking real-life scenes, objects may be complex and may be distinguished on high-level image features, which can enhance tracking [Okuma et al., 2004]. In fluorescence time-lapse cell microscopy [Godinez and Rohr, 2015], objects may be virtually indistinguishable such that tracking relies more strongly on *motion models* alone. Myriad and diverse examples of tracking applications are discussed in the literature as seen for example in review articles such as [Chenouard et al., 2014, Ulman et al., 2017].

4.3.1 Tracking algorithm

The position, appearance and disappearance of transient objects are modelled as a random spatial process. The frame-to-frame displacement of objects are a sum of three random variables: (1) the global movement of the tissue within the field of view, (2) the global movement due the growth of the tissue and (3) small fluctuations of the object (cell) within the tissue. The goal of the algorithm is to *filter* the global movement. The strategy of the algorithm is to find point-set correspondences between the random process at time t and the random process at time $t - \tau$ where τ is a lag variable. Objects that appear in both frames are *inliers* and those that do not appear in both frames are *outliers*. Outliers can be either *debris* *i.e.* any object which is determined *not* to be an object of interest or an object of interest that has “exited” or “entered” the new frame. From the image processing stage, any object outside the region of interest will be an outlier. More generally, an outlier is *an object that cannot be explained by a global frame-frame transformation*.

There are two conceptual phases of the algorithm which may be mixed during optimisation; (i) proposal transformations are generated from the data using one or more strategies and (ii) the optimal transformation with respect to some objective is chosen as the global transformation. Inliers are those objects which can be mapped to each other by the global transformation. Identifiers are propagated between the inliers in each frame in the *linkage* phase. New identifiers are generated for outliers within the region of interest. These algorithm stages are discussed in more detail in the following subsections. Algorithm 3 (§A) provides an overview of the stages that will be discussed.

Proposing transformations

A RANSAC strategy iteratively extracts a random sample from all data and partitions inliers from outliers in the data sample [Fischler and Bolles, 1981]. Consider sets of points $U(t), V(t - \tau)$ at times offset by a lag τ . A subset of these points corresponds to true cells (i.e. excluding noise and debris) and a subset of these true cells will exist in both frames (due to short-lived events). There are N points at time t and M points at time $t - \tau$. A number of permutations $\mathcal{P}(N, k)$ are sampled from U and a number of permutations $\mathcal{P}(M, k)$ are sampled from V . Take $k \in \{1, 3\}$, thus sampling either single points or triangular *constellations* from each frame. Points take a *natural order* and the same points can only appear in one constellation. Translation sampling is illustrated in Figure 4.7 and Figure 4.8.

Let n be the number of k -constellations $c_i(t)$ sampled from $U(t)$ and m be the number of k -constellations $c_j(t')$ sampled from $V(t - \tau)$. In general, for two tensors $\mathbf{M}_1$ and $\mathbf{M}_2$ related by $\mathbf{M}_1 = \mathbf{A}\mathbf{M}_2$, the affine transformation $\mathbf{A}$ can be uniquely determined through a least squares method. The transformation γ_{ij} is the affine transformation, taking the point-set $c_i(t)$ to the point-set $c_j(t')$.

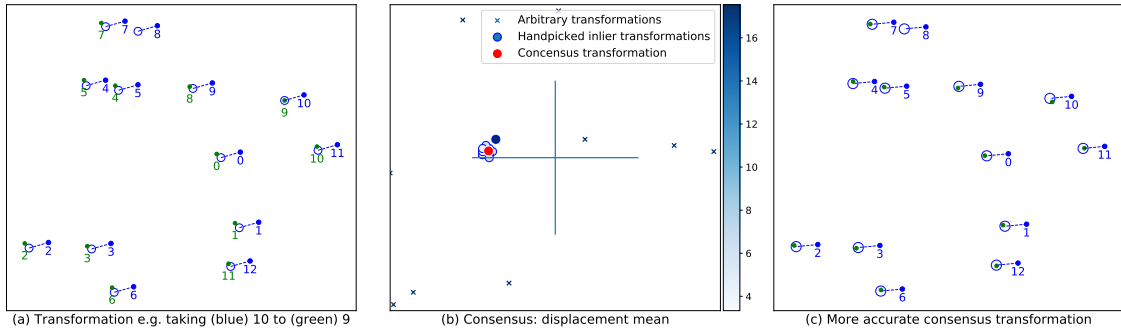


Figure 4.7: *Translation sampling with $k = 1$:* Panel (a) shows an arbitrary translation between a sample point in V and a sample point in U . Panel (b) shows the displacement of many such translations from the origin and includes a “consensus” translation i.e. the average of the top ranked translations. Panel (c) illustrates that in this case, the consensus translation is better than the one in panel (a) making better matches between circles and green dots. In this case, while not discernible from the image, the consensus translation turns out to be better than any of the other sampled translations.

Remark 43. Referring to Algorithm 3 (§A), on lines 5 and 6 constellations are sampled from the data in frames U and V and translations are computed. Sampling returns pairs

of points. For comparison with other stages, think of points as k -constellations with $k = 1$. As seen in Figure 4.7A, translations are generated for each pair of points. The translation taking a point in $v \in V$ to a point $u \in U$ is added to a list of candidate transformations. While translations and all general affine transformations could be generated in a single phase, it is more efficient to evaluate the linkages found from translations before sampling constellations for $k > 1$. Translations for $k = 1$ can be considered as a subset of transformations of $k = 3$ which should in general fit the data better. On lines 7 and 8 of Algorithm listing 3 (§A) translations are sampled and evaluated before full constellation sampling. The routines RANK and APPLY are discussed in Subsection 4.3.1. While stochastic sub-sampling could be used, for our data sizes it is efficient to simply consider all possible point pairings when $k = 1$.

General constellation sampling is preferably *seeded* with linkages found by translations so that likely outliers are excluded from consideration. It is worth emphasising that this not only reduces the candidate points to likely inlier points but *constellation congruences* can be identified i.e. the *same set of points* in one frame, can be paired to the *same set of points* in the other frame when proposing transformations. By first finding possible object linkages using translations only, constellation pairs $\mathcal{P}(N_L, 3), \mathcal{P}(M_L, 3)$ can then be sampled from the L inlier objects that appear to be in both frames. These can then be used to generate proposal transformations.

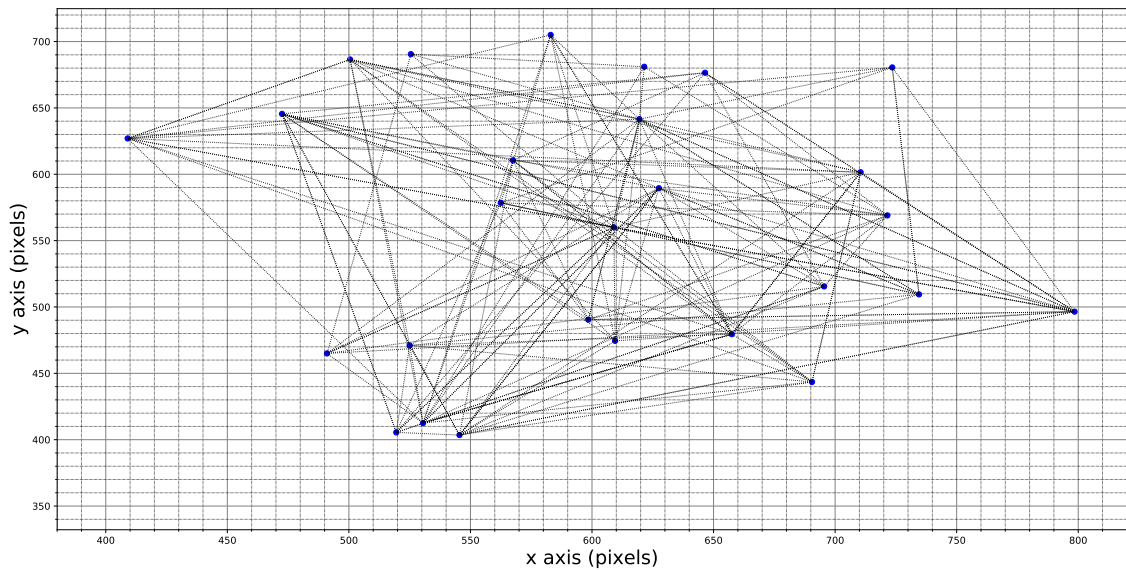


Figure 4.8: *Constellation sampling:* Random permutations $\mathcal{P}(M, k)$ with $k = 3$ produce triangular constellations in each frame. Affine transformations dimensions mapping point triplets in one frame to point triplets in another frame are added to proposal set Γ .

Ranking transformations

Transformation *ranking* is carried out using a modified least squares loss function. The *least squares loss* objective is generally described as the minimization of $\sum |\mathbf{y} - \hat{\mathbf{y}}|^2$, where $\mathbf{y}$ is a *proposal* vector and $\hat{\mathbf{y}}$ is the *target* vector. In the registration problem, the proposal

vectors correspond to the points in $U(t)$ under transformations from the set Γ . The target vectors are points in $V(t - \tau)$. The *least squares loss* objective is modified based on the following prior

1. We expect a global rigid transformation to explain the movement of object centroids up to small fluctuations due to individual object movement deemed negligible.
2. A distance between objects' centroids $> \varepsilon$ is assumed

The value of ε in the context of the sample data corresponds to a value slightly larger than the average blob's radius. When considering distances between proposal points and target points, only distances to the *first nearest neighbour within a radial distance ε* are considered. Each candidate transformation γ will be applied to *all* points $u_i(t) \in U(t), i \in \{1, 2, \dots, N\}$. Let $\tilde{u}_i = \gamma(u_i)$ then we use a cost function

$$C_\gamma = f \left(\sum_{i=1}^N \min(\varepsilon, |\tilde{u}_i - nn_V(\tilde{u}_i)|^2) \right) + d_\gamma \quad (4.12)$$

where $nn_V(\tilde{U}_i)$ is the nearest neighbour position in $V(t - \tau)$ to the projected point $\tilde{U}_i$. If the projected point has no nearest neighbour within a ball of radius ε , the *capped* distance $\varepsilon + 1$ is attributed.

Remark 44. Depending on the data, the cost function C_γ may be sufficient to rank transformations in certain scenarios. Raw transformation scores in $\mathbb{R}_+$, are approximated by scores in $\mathbb{Z}_+$ i.e. *binned* into discrete classes. These are then ranked using further criteria. Ranking sorts first on the binned score and then sub-sorts on say, translation displacement magnitude of transformations, in order to distinguish similarly performing transformations. Integer binning is a simple example of a linear binning function that is easy to implement and test. More adaptive approaches to binned ranking such as density-based clustering are potentially more robust but more difficult to test in non-trivial pipelines.

Assumptions about the data

Catchment region *Objects in the point cloud are expected to be separated by a minimum distance ε such that under global transformation single objects are matched.* The parameter ε can be chosen based on the data. For sparse point cloud data there will typically be only one match within the ball of radius ε . An ε value can be chosen such that it is possible to find multiple objects matched to one object. This assumes all matched points in the given catchment region correspond to the same object and that multiple points are the result of *e.g.* image processing anomalies.

Consensus translation *Transformations explain the global translation up to small Gaussian fluctuations in individual object movement due to true events or observation error.* Figure 4.7B gives an example of taking the “consensus” *i.e.* displacement mean of the best translations. This can often be better than any single translation given that each object undergoes a small random fluctuation such that translations between any point-pair cannot explain all object movements. If the position of an object is given as

$$\mathbf{p}_i(t) = \gamma(\mathbf{p}_i(\mathbf{t} - \tau) + \xi_i$$

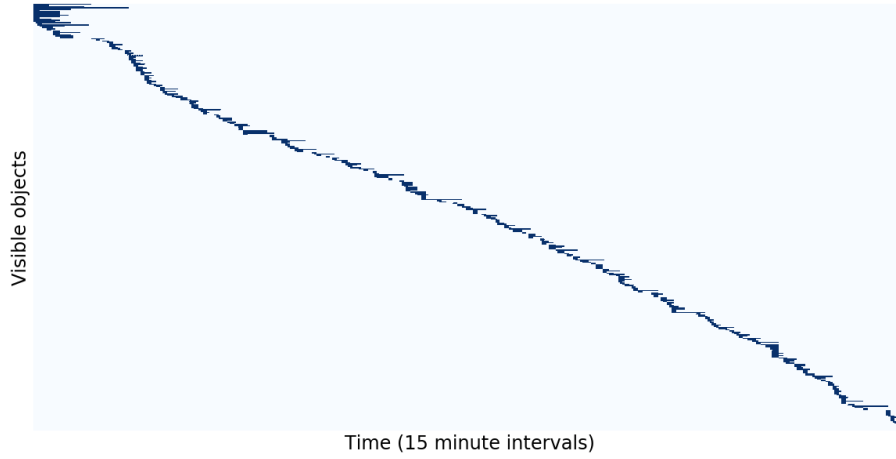


Figure 4.9: *Life matrix:* Rows (y-axis) correspond to detected individuals with time on the x-axis. The life matrix shows a generally consistent trend in the appearance and persistence of objects. At the beginning of the sequence, objects persist for an abnormally long time for biological/experimental reasons.

where $\xi_i \sim \mathcal{N}(0, \sigma_0)$ is the random fluctuation for one object, taking the “consensus” effectively filters out the *noise* to reveal the global movement.

Minimum allowed frame detections *Transient events are expected to persist for a minimum number of frames greater than or equal to 1.* For the root data the minimum number of frames is 2. Objects that have been identified for less than 2 frames are excluded from the tracking result. Such object detections are termed *single-frame detections*.

Gaussian region of interest likelihood *Objects of interest in the xy -plane are expected to be found in the center of the field of view.* Objects may be detected near the original frame boundaries depending on the size of the ROI. Objects can be treated as inliers/outliers based on standard outlier detection methods under the assumption of normality.

Remark 45. Further details about the software implementation and open source code can be found in [Amarteifio et al., 2018]

4.3.2 Results

The *life matrix* shown in Figure 4.9 is a binary matrix taking a value 1 if an object is detected and 0 otherwise, where each row contains the lifetime of a single detected object. The data in the life matrix are conditioned (filtered) on individuals surviving for at least 2 frames. The tracker requires the frame-frame movement of objects of interest to follow the global transformation. Debris movement is not correlated with that of objects of interest and should not be identified over consecutive frames. These objects appear in the tracking output as single-frame detections and are discarded. The number of objects observed as a function of time and the average duration of events as a function of time are alternative representations of the data shown in the life matrix. These time series are

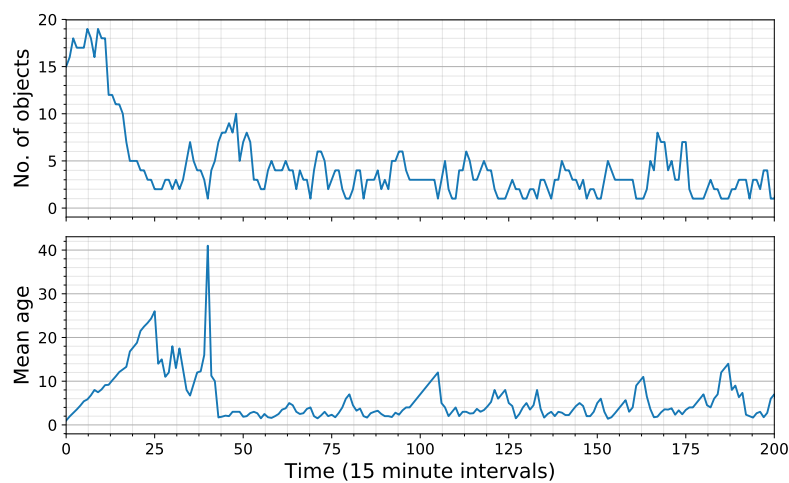


Figure 4.10: *Time series:* The population time series reflect the information shown in the life matrix. The high spike in the mean age is a result of one object surviving for a long time while others disappear. After this point more typical fluctuations in object activity are observed.

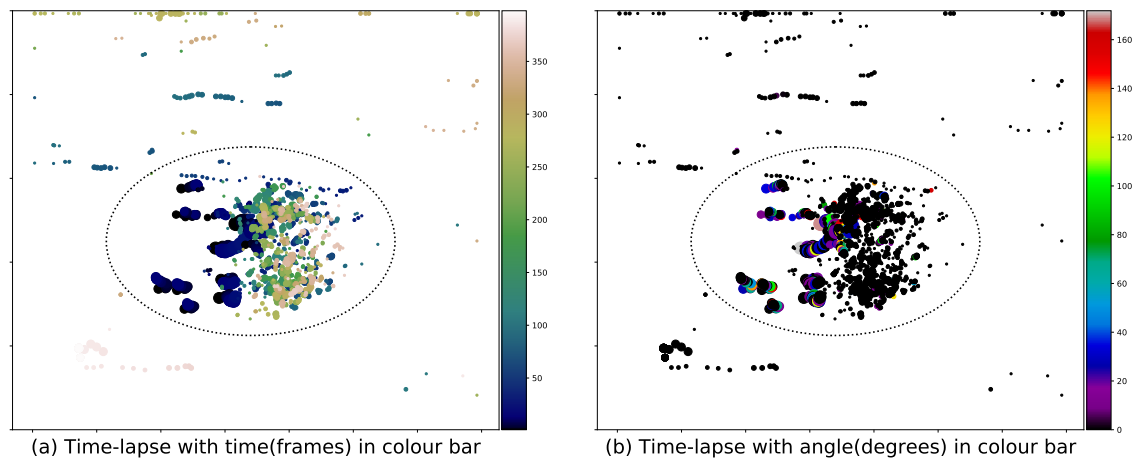


Figure 4.11: *Time-lapse plots:* Two different colourings are used to distinguish time(a) or differences between the object displacement vector and the global transformation vector(b). The ellipse represents the (Gaussian) statistical region of interest over many frames.

shown in Figure 4.10. Figure 4.11 shows time-lapse representations of the data produced by the tracker.

4.3.3 Testing against surrogate data

In order to test the tracking strategy against ground truth, surrogate data are used. Surrogate data must be representative of the biological data. A number of parameters are introduced to determine how data are generated and the tracker is tested for different points in the parameter space. A parameter n determines how many objects will be generated in a frame. Objects are distributed uniformly at random locations around the origin and added to the set of points U . A *random* Euclidean transformation (roto-translation) $\hat{\gamma}$ is generated using a parameter r to control the magnitude of rotation and a parameter s to control the magnitude of the translation. A set V is produced by applying $\hat{\gamma}$ to U . A parameter u determines the strength of (small) random fluctuations to individual object locations in V . A parameter p determines how many “noise” points are added to both U and V to represent debris or particles that exist in one but not the other frame. In the biological context this corresponds to cells that are no longer fluorescing or just beginning to fluoresce at time t . These parameters are summarised in the table below.

Param.	Description
ε	Minimum object separation
n	Number of points to generate in the point-set U
p	Number of “noise” points to add to the point-set U and the point-set V
r	Generate random (global) rotations in the range $[-\pi/r, \pi/r]$
s	Generate random (global) translations shifting axes in range $[0, s]$
u	Random uniform shift of points in V on each axis in range $[0, u]$

Table 4.1: *Surrogate data parameters:* Surrogate data are generated respecting a minimum object separation ε . A random Euclidean transformation determined by parameters (r, s, u) is applied to a set of points U to produce a new set of points V . To simulate debris or particle births and deaths, p points are added at random locations to both U and V . Added points also respect the minimum object separation ε .

The tracking objective is as follows: given U and V determine $\hat{\gamma}$. The tracking algorithm’s ability to do this will depend on the amount of noise points and the ferocity of the Euclidean transformation. Results for some parts of the parameter space are shown in Table 4.2. To test the tracker’s ability to find $\hat{\gamma}$ we establish a classification task. Under the correct global transformation, correct identifiers (with respect to ground truth) should be assigned to inliers and the identifier value -1 is assigned to outliers. The *loss*, which we aim to minimise, is averaged over samples and corresponds to the number of objects in V that do not have the correct identifier.

r	s = 2	s = 6	s = 10	s = 14	s = 18
12	0.2	1.00	1.54	1.70	2.22
16	0.0	0.12	0.56	0.72	1.26
20	0.0	0.00	0.36	0.74	1.30
24	0.0	0.00	0.16	0.74	1.20
28	0.0	0.00	0.10	0.74	1.10
32	0.0	0.00	0.00	0.64	1.00
36	0.0	0.00	0.00	0.62	0.92

Table 4.2: *Surrogate data test results (loss):* With $n = 10$, $p = 3$, $u = 1.0$, results for average of 10 samples for each point in the (r, s) parameter space. These results illustrate that for small rotations (approx 5 degrees) or for small translations the tracker is robust and degrades gracefully for larger rotations (approx 15 degrees).

4.3.4 Comparing with manual tracking

The automated tracking was compared with human-led manual tracking.* Algorithmic detection was verified by marking the position of all cells, without prior knowledge of the results of the algorithmic detection, in a subset of 50 frames by an experienced researcher using the built-in FIJI plugin *Cell Counter*. The human researcher marked $n = 439$ cells, compared to $m = 379$ cells detected algorithmically. The rate of false positives, i.e. cells detected algorithmically and not by the researcher was 12.4%. The rate of false negatives, i.e. cells detected by the researcher but not algorithmically was 24.4%. The average number of false positives and false negatives combined in a frame was 3.0, with an average number of total cells detected per frame was 11.0.

In the automated cell counting we prefer to under-count than over-count mitotic events. The emphasis in the biological investigation is to understand where and when true events occur within the root structure. The 12.4% false positive error is considered reasonable and would be reduced further if using some of the data filtering methods discussed above.

Evaluating transformations chosen by tracker

Constellation sampling with $k = 3$ is the default mode for finding a general affine transformation between constellations. This assumes there are at least $k = 3$ objects in each frame and to find a congruent constellation in each frame typically a higher number *e.g.* $2k$ is required. In addition to affine transformations between constellations, translations between single points are always added to transformation proposals. In some cases these may be better than any general affine transformation sampled by the tracker. In situations

*These tests were performed by Todd Fallesen

where there are too few points in either frame to find viable general affine transformations via constellation sampling, the tracker “falls back” to using translations between points as the sampling strategy. Additional transformations can be added to the proposal set using other strategies such as inclusion of an optimal transformation found in the previous time step. Tracking is most accurate when there are many objects in point clouds from which to generate transformations. Figure 4.12 shows an example of a sequence with *few* objects. If there are two objects in one frame and one object in the other, translations taking one object to either of the other objects will score similarly. There is no way to clearly distinguish them, keeping in mind that registration and tracking is being performed simultaneously. In the first-to-second frame transition in Figure 4.12, the actual object 211 is incorrectly labelled as *new* object 216 and the new object is incorrectly labelled 211. The identity transformation, which was not selected here, scores similarly but not necessarily better than the exhibited rival transformation as both transformations match one object. Binning as discussed in section 4.3.1 helps to discriminate between similar transformations based on displacement. The transformations will be in the same score “bin” but the identity will be preferred by the tracker because it has the smaller displacement.

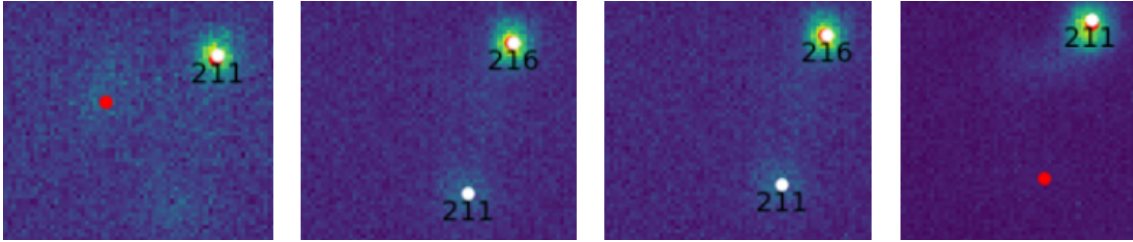


Figure 4.12: *Dealing with low information when ranking transformations:* The image sequence (left to right) demonstrates how identifiers can be incorrectly assigned when there are too few objects with white dots (at t) and red dots (at $t - 1$). The *identity transformation* would make the correct assignment in this example.

Lag analysis can identify possible tracker calibration issues. The tracker considers transformations between consecutive frames, $t - 1 \leftarrow t$. Suppose for the purpose of analysis we introduce a lag τ and consider transformations between frames $t - \tau \leftarrow t$. As a confidence indicator, compare the propagation of object identifiers between frames separated by different lags. Conditioned on objects being more than τ frames old, compare identifiers propagated over frames for $\tau \in \{1, 2, 3\}$. Then count the number of objects where identifiers are in agreement between different lags. For example it is reasonable to expect the following equivalence,

$$t - 3 \xleftarrow{\lambda_1} t - 2 \xleftarrow{\lambda_1} t - 1 \xleftarrow{\lambda_1} t \equiv t - 3 \xleftarrow{\lambda_3} t \quad (4.13)$$

where with an abuse of notation λ_τ corresponds to propagation of object identifiers using optimal transformations at lag τ . If the lag analysis shows *significant* differences, it may be due to accumulation of errors, in turn due to poorly chosen parameters. Lag analysis can identify specific instances of violating frames and objects for troubleshooting. The lag analysis result for a sequence of sample frames is shown in Figure 4.13 showing generally good agreement between different lags for suitably chosen parameters. While large lag-lag differences suggest calibration issues, small discrepancies are expected. For example if there is a sufficient accumulation of individual object fluctuations between lags, which place an object outside of a given catchment region(ε value), an object may be treated as an outlier and given a fresh identifier for larger lags.

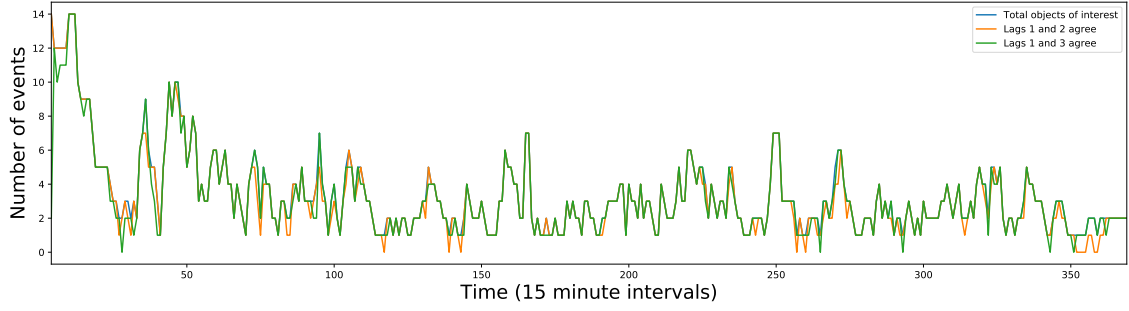


Figure 4.13: *Dis/agreement between lags:* Lag-1 identifiers for objects of interest (excluding single-frame detections) are compared with identifiers assigned by $\lambda_\tau : \tau \in \{2, 3\}$ transformations for objects that are older than τ . 734 objects appeared in the sample sequence. $42 \approx 6\%$ showed lag1-2 disagreements and excluding these, $29 \approx 4\%$ showed lag1-3 disagreements in specific frames.

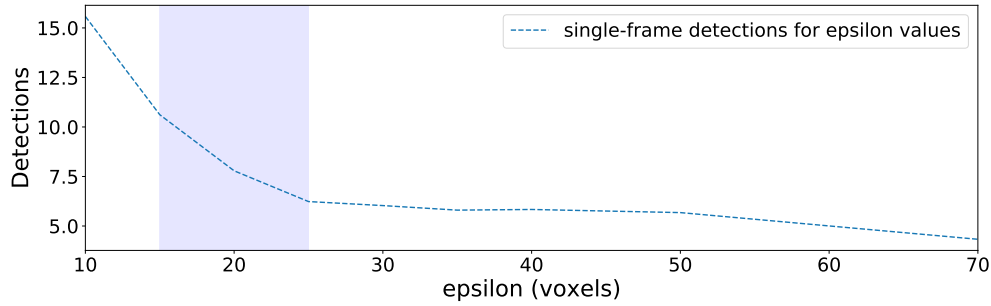


Figure 4.14: *Choosing an ε value:* The region 20 ± 5 , which is shaded in the plot, highlights the region where the reduction in single-frame detections begins to decrease before levelling out. This region can also be interpreted as the range of observed object radii. There is a trade-off in choosing an ε value. A value of $\varepsilon \ll 20$ results in excessive single frame detections as the tracker cannot accommodate small fluctuations in cell movement because the catchment region is too small, while values for $\varepsilon > 25$ merge lineages erroneously because the catchment region is too large.

Anisotropic matching tolerance considers the trade-off when choosing the ε value. When the value is very small, the tracker is very strict, marrying objects that sit within a small ball of radius ε as illustrated in Figure 4.14. If the input data are accurate and if a bound on fluctuations away from the global transformation is known and small enough, then a small ε value is sensible. If however there are unpredictable fluctuations for example due to inconsistent centroid/peak detections in noisy data, then a large epsilon value is required to tolerate these fluctuations and to avoid generating new identifiers erroneously. For some datasets it may be difficult to find a value for ε that balances conflicting objectives. One way to mitigate this is to use a larger ε value and then in a post-processing step replace the isotropic ball of radius ε with an anisotropic catchment region. This penalises objects that are moving in the opposite direction to the global transformation while being more tolerant of objects moving in a similar direction to the global transformation. This is illustrated in Figure 4.15 where it is clear object 44 is moving in a different direction to the others. This deviation in movement may or may not be explained as a small fluctuation in the object's position.

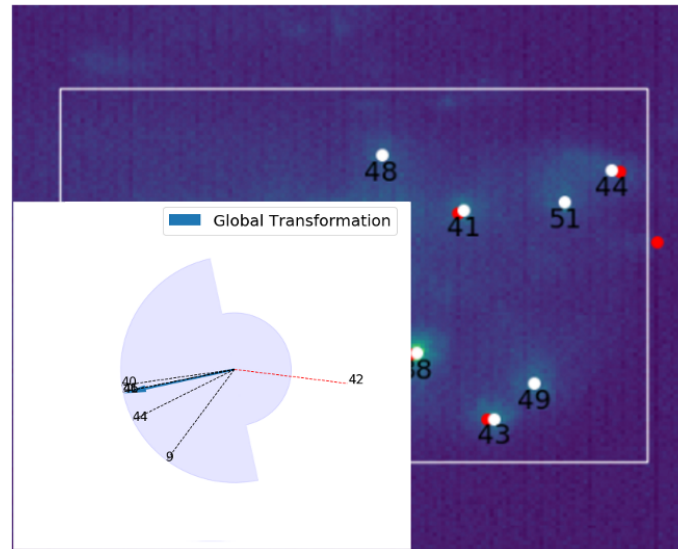


Figure 4.15: *Anisotropic matching tolerance:* Differences in angles between the global transformation and the displacements between married objects are compared. White dots and red dots correspond to objects at time t and $t - 1$ respectively. Most red dots are to the left of white dots which corresponds to a global transformation vector of about 170° from the positive x-axis vector. Object 44 has a very different angle to the rest and in the image, the red dot appears to the right. Large displacements in conjunction with large angle differences for a number of consecutive frames suggests debris.

Finalising the dataset

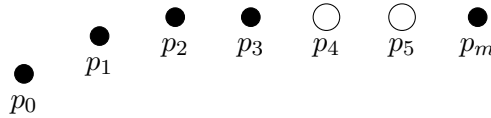
Data filtering can be applied in a post-processing stage. Figure 4.11 illustrates post-processing of the full datasets. The ellipse shows the region containing the most likely mitotic events occurring in the root tip. The time colouring Figure 4.11A illustrates when and where events occur and indicates (roughly) if plotted points correspond to the same object. In Figure 4.11B, large angle differences between object displacement vectors and a global transformation vector appear in lighter colours (*e.g.* red) and are indicative of possible inlier misclassification. Objects outside of (or far from) the ellipse and objects that show large angle differences are candidates for removal. In the case of angle differences, actual displacements and ε values can be taken into account when filtering objects as described in the preceding paragraph and illustrated in Figure 4.15. Fluctuations of the ROI area in each frame is an indicating of how well the image processing step can localise the root. When the signal to noise ratio is good, the root region can be accurately detected and there should be significant fluorescent activity corresponding to biological events within a (relatively) small ROI, which varies smoothly between frames. On the other hand if there is too much noise or few (biological) events, it may be difficult to localise the root from frame to frame resulting in a relatively large ROI or one that varies non-smoothly. In the later case, debris may lie within the excessively large ROI and be picked up by the tracker. In Figure 4.11, this debris appears for example in the top part of the images, particularly towards the later stages of the experiment, due to drift in microscope calibration or other factors. The root grows in the positive x direction (left to right) and debris are typically static or moving with gravity in the opposite direction as seen by some trail-like patterns roughly parallel to the x-axis.

4.4 Characterising cancer cell motility

In this section I describe methods used for the statistical analysis of 3D+time tracked cancer cells. There are two cell types, T-cell acute lymphoblastic leukaemia (T-ALL) and acute myeloid leukaemia (AML) cells. Datasets corresponding to these cell types are each divided into two sub-types; either they have not previously been exposed to chemotherapy or they have. This yields four experimental *conditions*; *T-ALL no chemo*, *T-ALL post chemo*, *AML no chemo*, *AML post chemo*. Each of these conditions are monitored in experiments over different positions in different mice. For a given mouse *position*, an experiment is carried out in two *stages*. The position is monitored for 1.5 hours in a *PRE* stage. A treatment is applied. After 15 minutes pause, the position is monitored for 3 hours in a *POST* stage. During monitoring, a selection of cells are identified and tracked to produce a set of four-dimensional tracks (t, x, y, z) for each [position], for each [stage], for each [condition]. Spatial coordinates are in microns and time step indices correspond to 3-minute intervals. We will study the *displacements* i.e. jumps and the *distinct sites visited* by cells in a volume. Track displacements are quantified in terms of *mean squared displacement* and further characterisation of the displacement densities uses a fractal measure. The *distinct sites visited* is analogous to measuring the fractal dimension of a curve. The *distinct sites visited* metric is introduced to describe spatial (region) correlations and how tracks explore space. Diffusion and anomalous diffusion is discussed as a way analysis cell movements in single particle tracking. In order to measure differences between displacement distributions, a family of stable distributions is used to fit track data. This is the third measure, in addition to the mean squared displacement and distinct sites visited, used to discriminate between cell movements under different conditions.

4.4.1 Displacements

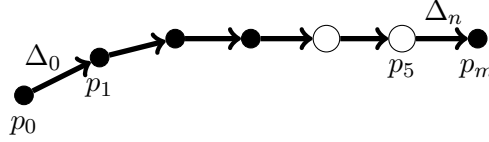
Experiments produce tracked cells in three dimensions. Raw tracks consist of data points labelled with a *global* clock time and three spatial coordinates $p(t_c, x, y, z)$. Due to noisy data or various unavoidable inaccuracies in tracking, there may be gaps in time producing *incomplete* tracks. As such, there is a need to (time) interpolate all tracks assuming constant speed for *missing* data points.[†] This produces *complete* tracks with m data points. Throughout this chapter, it can be assumed that all tracks are complete. In the example below, points p_4 and p_5 are interpolated points.



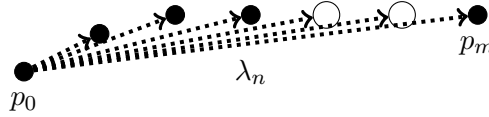
After time interpolation there is one data point per time point for the interval over which the track is observed.

[†]Of 24,365 data points in the track data, 2,059 are interpolated points.

The *local displacements* are the vectors connecting the data points p_t to data points p_{t+1} . For a track consisting of m data points there are of course $n = m - 1$ local displacements, the last of which is Δ_n . These are vector lengths i.e. the euclidean distance between each pair of points.



The *lagged displacement* is the displacement between each point, $p_1, \dots, p_m$ and the first point, p_0 . The last such point is λ_n .



One simple interpretation of mean squared displacement up to time t is the average of the squared-lagged displacements

$$msd_t := \frac{1}{t} \sum_{j=1}^t \lambda_j^2 \quad (4.14)$$

Remark 46. The speed can be characterised in a similar way in terms of the cumulative *local* displacements $\frac{1}{t} \sum_{j=1}^t \Delta_j^2$.

The *time-averaged mean squared displacement* ($tmsd$), which is averaged over all lags τ and times t . The angled brackets $\langle \cdot \rangle$ denote the ensemble mean.[‡]

$$msd(t, \tau) = \langle |\mathbf{r}(t + \tau) - \mathbf{r}(t)|^2 \rangle_{m,n} \quad (4.15)$$

Remark 47. Consider the $m \times m$ distance matrix D_{ji} . This is the matrix of Euclidean distances from each point i in a track of length m to another point j on the track. This matrix is symmetric and the main diagonal contains 0s (i.e. when $i = j$ this distance to self is 0). As such we can restrict our attention to the *upper triangular matrix* $D_{j \geq i}$. Subsets of displacements sit along the diagonals of this matrix. The main diagonal forms the set of 0-lag displacements. The next diagonal forms the set of 1-lag displacements (and so on). The last diagonal i.e the single matrix cell in the corner, is the *track displacement*. We can shift each row i by i columns to the left to produce the lower triangular *lag matrix*.

Remark 48. When considering distributions of “random” displacements we should *not* in general treat them as being independent or identically distributed (*iid*) nor should we treat the *process* that generates them as being stationary.

[‡]Explicitly, the summation is $\frac{1}{t-\tau} \sum_{i=1}^{t-\tau}$

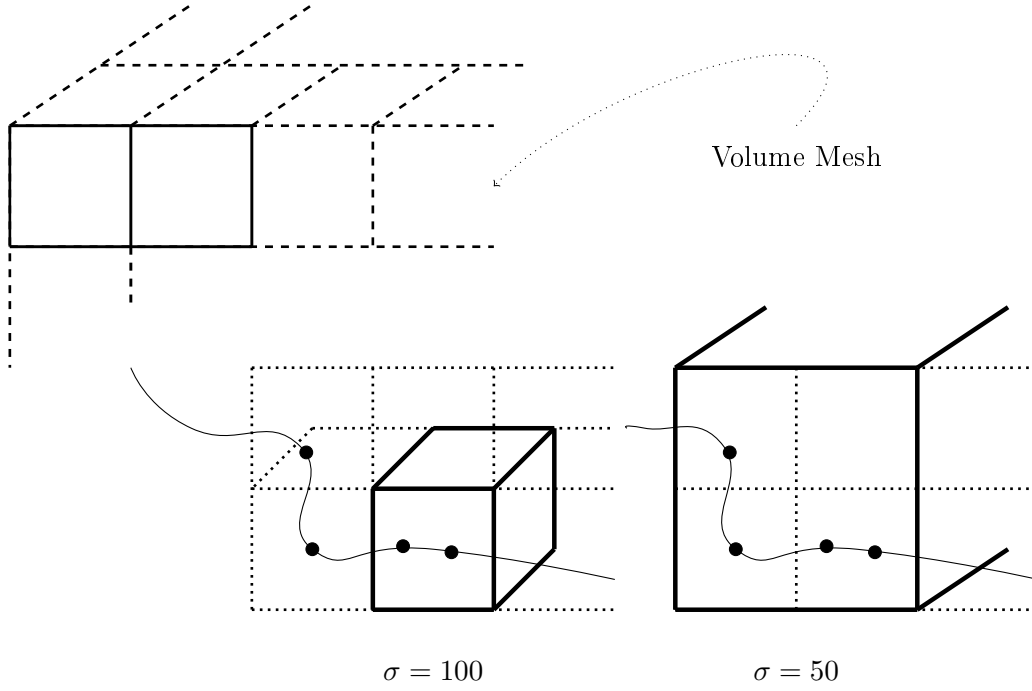


Figure 4.16: *Measuring distinct sites visited at different scales:* The illustration shows sites for $\sigma = 100$ and $\sigma = 50$. A volumetric mesh is constructed for each scale σ . The same track (tracked data points marked with dots) is shown passing through sites at two different scales. Four data points for a sample track are shown falling within sites for each scale. At scale $\sigma = 100$, 4 data points *fall in* 3 sites. At scale $\sigma = 50$, the same 4 data points fall in 1 site.

4.4.2 Distinct sites visited

The distinct sites visited measure depends on an inverse-scale parameter σ used to partition the mouse position volume into small *volume elements* or *sites*. This calculation is carried out similarly to the method of *box counting* when computing the fractal dimension [Falconer, 2004].

Remark 49. For microscopy data, it is important to take into account the *aspect ratio* in x, y, z i.e. the differences in resolution. Typically the z resolution is lower than the xy resolution. For example we might have $x : y : z = 1 : 1 : 0.2$. With this in mind, consider partitioning the volume with inverse scale parameter $\sigma = 10$. We slice the x and y axes into 10 segments and the z axis into 2. At a more *fine-grained* scaled $\sigma = 100$, we partition the x and y axes into 100 segments and the z axis into 20.

For each track, consider which points p_j visit a particular site and count the distinct sites visited for each track and divide by the length of the track to produce a distinct site visited metric *at a given scale*. This is illustrated in Figure 4.16.

$$dsv_\sigma := \frac{\#\{s_0^{(\sigma)}, s_1^{(\sigma)}, \dots, s_k^{(\sigma)}\}}{n} \quad (4.16)$$

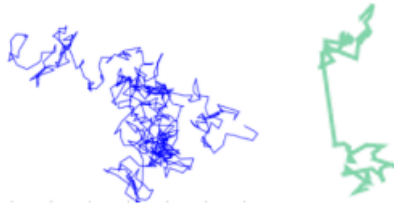


Figure 4.17: *Comparing linear and anomalous diffusion:* a Brownian motion on the left and a “typical” cell track on the right. Distribution of *jumps*, *angles*, *speed* etc. can be measured to understand how the particle explores space. The track on the right could be referred to as *Lévy-like* in that, qualitatively, it consists of more occurrences of large jumps than might be expected from Brownian motion.

4.4.3 Anomalous diffusion

In single particle tracking (SPT) analysis it is common to treat cells, molecules, and other objects as diffusing *particles* allowing for the adoption of methods from physics to study motility. The vanilla physical theory might treat cells as simple Brownian particles and consider the mean squared displacement, in this case the expectation of the variance, as a way to quantify how a particle explores space. Exploration of space will be our primary focus as it may offer perspective on an apparent cell *strategy*. In the absence of knowing precisely how cells make decisions, by treating observed trajectories as stochastic processes, we would like to determine if a certain process is more *efficient* than another. This study considers diffusions that are not (only) Brownian i.e. *anomalous diffusion*. This widening of scope better accounts for the variability we see in biological behaviour. There are ramifications for calculation and interpretation of measurements such as *msd*.

Given that jump distributions are not-only-Brownian, as part of the analysis I consider one method to fitting data to distributions that uses the Lévy alpha stable spectrum, which admits both Brownian motion and certain power law jump distributions. *Lévy flights* and *Lévy walks* (the name *Lévy flight* was proposed by Benoît Mandelbrot and the name *Lévy walk* by Shlesinger et al. [1982]) appear frequently within ecology and biology to model animal or cell movement despite having infinite second moments.

For Brownian motion, upon measuring the mean squared displacement, one expects the Einstein-Smoluchowski relation $\langle |\mathbf{r}(\tau)|^2 \rangle = 2D\tau^\alpha$ to hold, where t is lagged time and D is the diffusion constant/coefficient of the moving particle. For normal diffusion, the exponent $\alpha = 1$. *Anomalous diffusion* (cf. [Eliazar and Klafter, 2011, Méndez et al., 2013, Zaburdaev et al., 2015]) applies in the case where the exponent is *not* 1. In cases where $\alpha > 1$ we have *superdiffusion* and in cases where $\alpha < 1$ we have *subdiffusion*.

Classification of the diffusion *regime* is of interest in the biological context; for example superdiffusion might suggest *active transport* while subdiffusion may be indicative of one or all of crowding, corralling, preferential interactions, network-like or porous environments [Burov et al., 2011]. In ecology/biology, diffusive regimes have been linked to strategy; for example, there are studies that show the efficiency of subdiffusive transport [Guigas and Weiss, 2008] and indeed the much discussed and much debated Lévy foraging hypothesis

(cf. [Viswanathan et al., 1999, Pyke, 2015] for origins and recent discussion) emphasises the advantages of superdiffusion to search. One study by Harris et al. [2012] considers the migratory behaviour of T-cells. In particular they show that $CD8^+T$ T-cell conforms to so-called Lévy strategies observed in as other organisms.

The suitability of utilising (time averaged) mean squared displacement in single particle tracking analysis is widely discussed in the literature [Destainville and Salomé, 2006, Michalet, 2010, Burov et al., 2011, Kepten et al., 2013, 2015, Meroz and Sokolov, 2015, Gottwald and Melbourne, 2016]. One particular concern in biology for example is the heterogeneity among particles leading to a scattering of msd values [Kepten et al., 2013]. Another issues is the breakdown of ergodicity assumptions in subdiffusive tracks [Burov et al., 2011, Meroz and Sokolov, 2015], which challenges the idea of using msd as an estimator. Some authors such as Elliott et al. [2011] discuss complementary mechanisms that can be used for more robust analysis. Complications may depend on whether one is concerned with single tracks, for which data may be small, or ensemble tracks for which data may be more abundant (but heterogeneity becomes an issue). In certain cases it has been noted that on small time scales ($10^{-1}s$) only subdiffusion is observed while over larger time periods ($10^{-1}s$) mixtures of diffusion types (including normal diffusion) are observed (cf. [Selhuber-Unkel et al., 2009] and references within). In the case of continuous time random walk, a widely used (anomalous) random walk model, although the ensemble msd is subdiffusive the temporal msd is linear in time. He et al. [2008] re-evaluates the role of the diffusion coefficient when analysing (anomalous) diffusion in single particle tracking experiments describing it as a random variable. They observe that, unlike in the case of normal diffusion, it is *not* representative of the ensemble average due to ergodicity breaking leading them to propose a generalisation of the Einstein relation [He et al., 2008, Burov et al., 2011]. These observations suggest the need to be careful when drawing any conclusions as to the diffusive behaviour of a system by considering only the temporal msd [Meroz and Sokolov, 2015].

A modelling "wrinkle" with Lévy walks/flights is that as jump distributions they have infinite second moments, which require an artificial truncation in modelling scenarios. This gives rise for example to the *Truncated Lévy walk* [Mantegna and Stanley, 1994] or *tempered stable processes* [Rosiński, 2007]. In their review of Lévy walks, Zaburdaev et al. [2015] root this issue in its historical context, in terms of the random walk. They remind us that this trait of infinite propagation speeds is shared with the important diffusion equation (going back to Fourier), which could not easily be reconciled with the random walk (as conceived by Pearson). They discuss various reconciliation efforts that took place throughout the twentieth century culminating with the *continuous time random walk* [Montroll and Weiss, 1965]. Having a particle wait for a random time before moving ensures the ingredients required to model subdiffusion.

Remark 50. On truncations and ballistic cones: In what may appear to be a mere technicality, I briefly describe *truncation* as it pertains to making models *physical* (i.e. realistic). The truncated Lévy flight was introduced to maintain certain of the Lévy flight's desirable properties (e.g. modelling fat tails) while avoiding the "infinity problem". However, Zaburdaev et al. [2015] observe that the truncation is *arbitrary* and cannot always be justified in the given application context. They offer a resolution by way of a space-time coupling $x = v\tau$ (finite velocity). A ballistic cone is constructed by adding a time cost for the time it takes a walker to traverse some distance. Interestingly, this also provides perspective on the different anomalous diffusive regimes; in the case that the mean displacement length is

finite but the wait time is anomalously long, we find ourselves in the subdiffusive regime. On the other hand, in the case where the average waiting time is finite but the jumps are anomalously long, we find ourselves in the superdiffusive regime. In the subdiffusive case the width of the trace grows like $t^{\gamma/2}$ and in the superdiffusive case like $t^{\gamma/2\beta}$.

With an interest in evaluating the null hypothesis that tracked leukaemia cells are *not* (only) Brownian particles, one approach might be to consider comparing the data to, on the one hand, the distribution of Brownian increments (simulated from a Gaussian distribution) and on the other hand some other *fat-tailed* distributions such as Lévy distributions. Fat-tailed distributions also known as *heavy-tailed*, *leptokurtic* or *multi-scaled* distributions, have tails which decay much more slowly than the Normal distribution.

Ying et al. [2009] tested single particle track data against general *Weibull*, *chi*, the *long-short*, and the *generalized extreme value* (GEV) distributions. Their choice of GEV was to account for the observation of fat tails i.e. *extreme values*. In this study I consider the use of the *Lévy α -stable* family (Appendix C) of distributions due to the fact that by tuning a single relevant parameter (of four parameters in total) one can use the parameter α to directly describe the *rate of decay* of the tail (Nolan [2017]). This is one way to measure the deviation from "normality" and the likelihood of observing the sorts of extreme values that a Lévy walk (for example) may produce.

Remark 51. There are technical difficulties when relating such distributions to physical systems. Refer for example to [Stanislavsky et al., 2008, Gajda and Magdziarz, 2010] for a discussion on subdiffusion and α -stability. However, the stable spectrum (which subsumes certain relevant distributions used in the literature to model motility) provides one way to quantify differences between distributions, in terms of tail decay, under various experimental conditions and the question is whether or not this statistical test *discriminates* between cell types before and after treatment. A practical motivation for adopting this family of distributions in the current study, is the availability of numerical methods for simulating and fitting regions of parameter space even if closed forms are not available.

4.4.4 Results

Three tests were applied to the data; (i) comparing data points to distributions on the Lévy α -stable spectrum (ii) measuring the distinct sites visited in a volume and (iii) measuring the mean-squared displacement. In summary, we find that the treatment has a distinctive impact on T-ALL post chemotherapy, while AML shows a less significant differences in most cases.

Comparing distributions: Local displacement distributions have been fitted against a family of *Lévy α -stable distributions* using the method of Maximum Likelihood Estimation to extract most likely parameters. For our purposes, the so-called stability index, *alpha*, is of primary interest as this can be used to classify the jump distribution in terms of *tail decay*. In the fitting routine the software library [Miotto, 2018] was used. The method of Maximum Likelihood Estimation is used to find the most likely parameters given the model. We remind ourselves that MLE minimises $\frac{\partial \log \mathcal{L}(\theta)}{\partial \theta}$ with $\log \mathcal{L}(\theta) = \sum_i \log P(x_i | \theta_1, \dots, \theta_n)$, at the zeroes of all parameters. To compare distributions, the *Kullback-Liebler divergences*

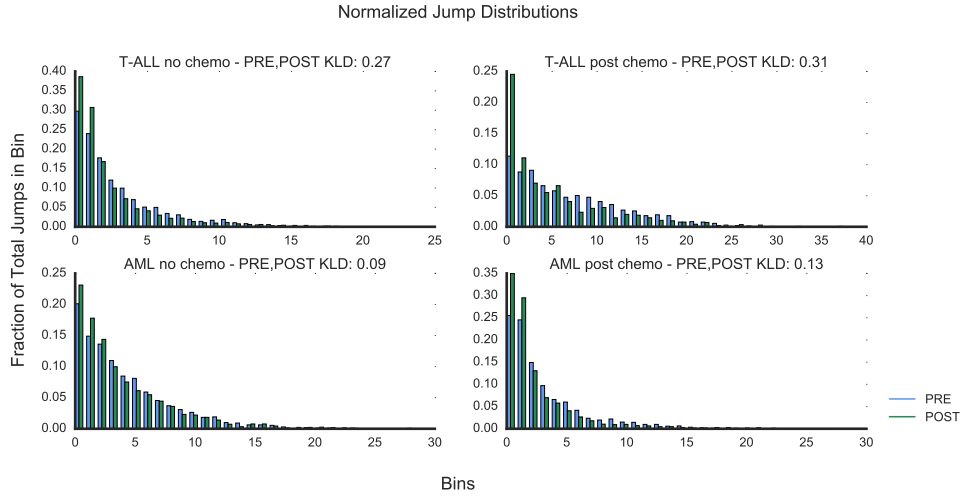


Figure 4.18: The displacement densities (normalised histograms) for each of the conditions: These compare before (blue/lighter) and after (green/darker) treatment. We are interested primarily in detecting differences as a result of treatment. The Kullback-Liebler Divergence (KLD) is a measure of this difference (§4.4.4 for details). The impact of treatment is greatest for the T-ALL types as the KLD values are higher. For T-ALL, the KLD values are close to 0.3, for AML they are close to 0.1. The differences seem to correspond to results of other tests but are not striking *per se*.

(KLD) is a $[0,1]$ -scaled measure of the significance of the differences between displacement distributions. When computing KLD and other entropic measures for continuous distributions, density-based methods can optionally be used to circumvent *ad hoc* binning of continuous data [Kraskov et al., 2004]. Density-based algorithms use a k -nearest-neighbour approach, which requires the parameter k to be specified. To employ this approach here, for the track data, I sampled 1000 displacements from each condition and applied the algorithm with $k = 7$. Experiments yield varying data sizes (number of data points), and sizes could effect the algorithm. Having analysed sensitivity to different samples, KLD was robust over permutations of the sample space. The sample size of 1000 was chosen based on the size of the smallest experiential data set.

The distinct sites visited metric partitions the position volume on a number of scales. The number of unique volume elements or “sites” at different scales that were visited by individual tracks are counted. For a given track of length n , by dividing the distinct count of sites it visits by its length, a number between 0 and 1 is produced. Unity corresponds to a track visiting a new site (at a particular scale) with each observed jump. The fractal “scaling” of a given track(set) can be analysed by plotting the number of distinct sites visited as a function of the scale parameter. This method is similar to *box counting* when measuring the fractal dimension of a curve in space or the *filling ratio* discussed by Méndez et al. [2013].

The distinct sites visited are analysed in Figures 4.19, 4.20, 4.21. Box plots compare distinct site visited values for ensembles and for a chosen reference scale i.e at the fine-grained scale ($\sigma = 100$). Two-way ANOVA p -values are reported as a measure of statistical significance within conditions and between stages.

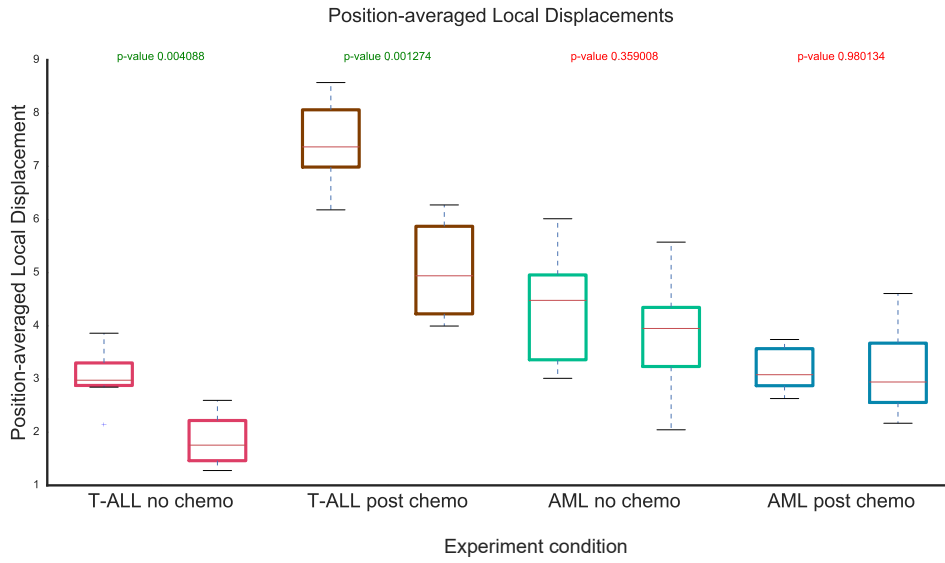


Figure 4.19: *Box plots of position-averaged displacements:* This shows a coarse-grained perspective of the displacements. These were computed by first taking an average within each mouse position and then showing the mean and standard deviation over the set of positions *per* condition. Consistent with the changes within displacement densities, we see most notable differences in T-ALL. p-values from ANOVA were (T-ALL no chemo, 0.004), (T-ALL post chemo, 0.001), (T-ALL post chemo, 0.004), (AML no chemo, 0.359), (AML post chemo, 0.980), confirming significance (95% confidence) for T-ALL but not for AML.

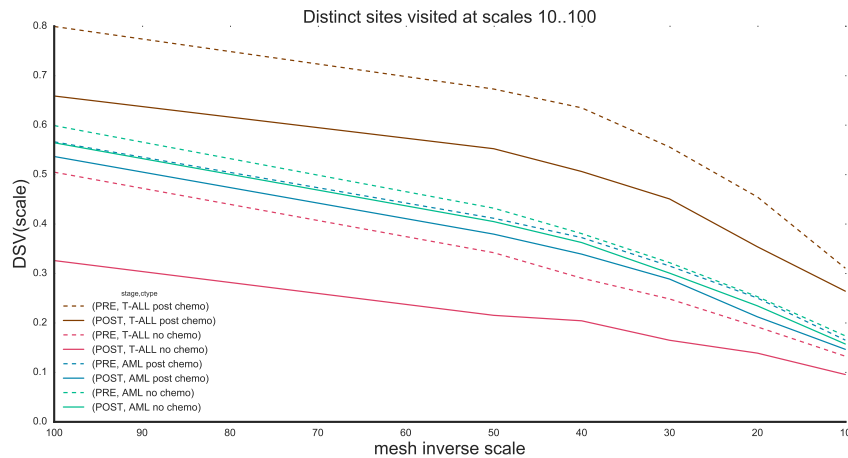


Figure 4.20: *Distinct sites visited as a function of the inverse scale parameter σ :* From left to right, fine-grained to coarse-grained, we see the expected decline in slope at coarse scales. Both the comparative rates of decline and how the distinct site visited *levels* change before and after treatment is consistent with other results, where we see the most dramatic changes in T-ALL after treatment.

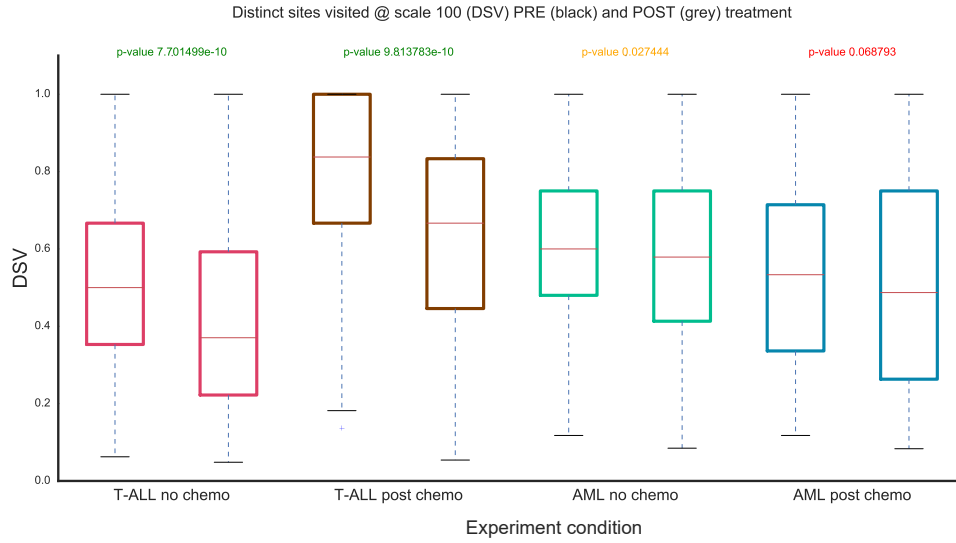


Figure 4.21: The box plot above compares distinct sites visited at a fine-gained scale ($\sigma = 100$). The differences correspond to the gaps between the (start of) lines in the scaling behaviour shown in the previous figure.

The ensemble time-averaged mean squared displacements as a function of lag are compared across experiment types. Subdiffusive behaviour is observed in most cases with the exception of T-ALL *post chemo*, which showed a transition to superdiffusive.

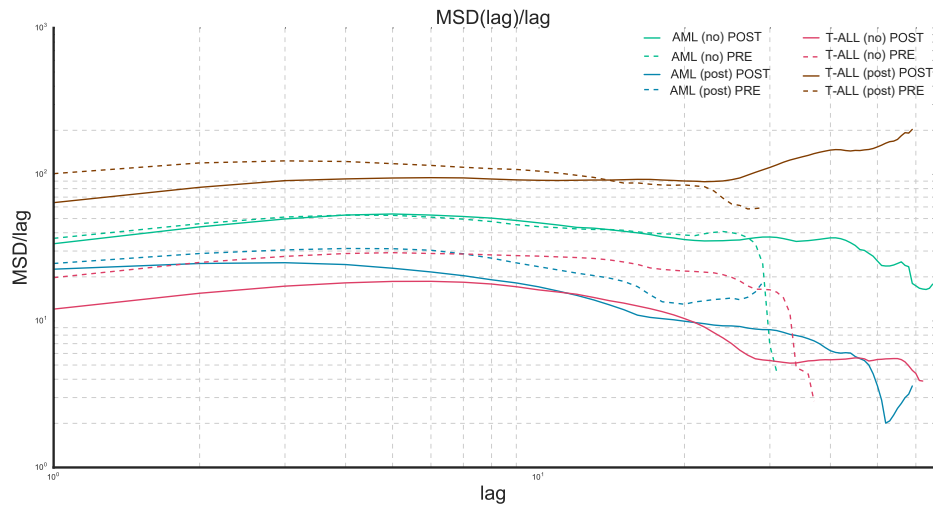


Figure 4.22: The log-log plot of time averaged mean squared displacement as a function of the lag (x axis) divided by the lag for each condition. These trends show power law fits of the data. Each condition e.g. (T-ALL no chemo) takes a fixed colour, while PRE/POST pairs can be distinguished by line type; dashed for PRE and solid for POST. We can see the T-ALL *post chemo* POST stands out, with a slope that does not decline. T-ALL *post chemo* POST appears *superdiffusive* as opposed to *subdiffusive* (all other cases).

4.5 Summary

The two studies in this chapter were presented in the thesis *introduction* as problems of the *first* and *second kind* in bioinformatics.

Considering a problem of the first kind, raw microscopy data must be processed and particles must be identified and tracked in order to extract knowledge. This is a difficult problem given the variability in 4D microscopy data and often, as in this chapter, bespoke image processing steps and tracking algorithms are required. In the case discussed, noisy and un-registered light-sheet microscopy data from experiments that used transient fluorescent markers presented a unique challenge. An algorithm, which was tested against surrogate data and compared with human-led tracking, was developed to extract spatiotemporal data corresponding to mitotic events in *arabidopsis* roots. Having extracted such data, a problem of the second kind emerges, namely, how to make statements about spatiotemporal patterns and behaviour. For the *arabidopsis* root case study, this problem of the second kind and indeed any statements about the biology was out of scope and will be addressed in a future article.

The second part of this chapter did address a problem of the second kind. Given cancer cell tracks, various metrics were constructed to quantify differences between different cell lines and different conditions (before and after chemotherapy). For example, the mean-squared displacement was used to show that T-ALL cells are more susceptible than AML cells to the treatment and become super-diffusive post-treatment, thus being expelled from the bone marrow. This is an example of a characterisation of dynamics that goes beyond analysing speed and displacement and offers additional insight to the biologist.

As a final remark it should be noted that Chapter 4 was very different in almost every respect to the preceding chapters. However all chapters are united in their attempt to understand spatiotemporal processes.

5 Discussion

Chapter 3 was in some sense a motivation and the fulcrum of the thesis. It provided an example of applying the field theory machinery to study the trace of a superprocess. The remainder of the thesis operated at two extremes away from this specific case study. On the one hand, Chapter 1 and Chapter 2 spent time developing abstractions of the field theoretic method, which in principle can assist in its application to arbitrary biological processes, or at least those that can be modelled as critical chemical reaction networks. At the other extreme, Chapter 4 directly analysed real, biological, spatiotemporal data.

In Chapter 1, my objective was to emphasise the Markov property and Markov semigroup, which leads to a formulaic way to write down the action and the path integral. For a wide class of chemical reaction networks for which particles are transported by linear diffusion, this leads in principle to a method to write down the action immediately. From the action, methods in Chapter 2 allow for L -loop graphs to be generated systemically. Using graph polynomials in parametric representations or otherwise, many of these graphs can be integrated easily⁴⁴, which leads to the determination of the *renormalization group functions* used in Chapter 3. These in turn allow for the scaling behaviour of the vertex functions to be determined and tested against numerical simulation of the reaction-diffusion processes.

While the renormalization group combined with the Doi-Peliti formalism has proven itself in making statements about the scaling of primitive processes such as *coagulations* or simple compound processes such as *branching and annihilation random walks* or the *Gribov process*, the general flexibility and efficiency in applications to arbitrary processes is perhaps questionable. It would be interesting to consider more studies that apply these methods to general chemical reaction networks or biological processes in order to quantify the impact of the geometry of space on various *critical* "chemical" processes more generally. In this thesis I have tried to emphasise algebraic and algorithmic aspects that might aid in the efficiency with which these methods can be applied.

In future I would seek a clearer overall picture of the renormalization group applied to spatial Markov processes. It is encouraging that one can get quite far in a renormalised perturbation theory by thinking only about continuous Lie groups; firstly exploiting the conservation of energy in evolutionary processes and then exploiting self-similarity and scale invariance in the renormalization group equations. Sandwiched between these two one-parameter groups is the determination of canonical dimensions of fields. At the time of writing I do not have a clear (non-heuristic) picture of this, which certainly inhibits a fully algorithmic treatment of the application to arbitrary processes. This little detail holds some sway over whether or not a renormalizable field theory can be found for an arbitrary process.

Summary of contributions

I review the main contributions of the work. In brief, the following are novel; a number of algorithms are developed to track cells in noisy microscopy data, to evaluate analytical models of spatial processes on different graphs and to enable efficient Monte Carlo sampling of the Abelian sandpile model. Theoretical contributions include a number of graph theoretical perspectives of existing methods in statistical physics, a derivation and use of parametric Feynman integrals and algebraic and algorithmic developments of the coherent state path integral formalism and its application to spatial Markov processes.

5.1 Algebraic aspects of the Doi-Peliti/pQFT formalism

By *algebraic* I refer to a sufficiently abstract description of the formalism, which lends itself to algorithmic treatment. As for the case of the parametric integral representation, I was initially motivated by the possibility of carrying out the field-theoretic procedure algorithmically for many graphs and at higher loop orders. This algorithmic approach was covered in several stages; in Chapter 1, focus on Poisson statistics leads to a general form of infinitesimal generator for the semigroup. A chemical reaction network is a compound Poisson process, a Markov process. But the coherent state path integral formalism is also based on a *Poisson ansatz*, which says that the configuration (in a reaction-diffusion process or a Euclidean field theory) evolves such that particles interact in a bath of particles described by a spatial (Poisson) point process i.e. "complete spatial randomness". This idea was discussed by Glauber [1963] and picked up by Doi [1976a]. The combinatorial description of the infinitesimal generator follows from the Poisson ansatz in both the ordinary and factorial moment generator cases. Depending on who you ask, the quantisation of chemical processes takes place on either a Bargmann-Fock space in the Doi [1976a] picture or on a formal power series basis by way of the *Umbral calculus* (cf. Blasiak [2005], Roman [2005]). In any case, the infinitesimal generators and the Lagrangian by extension can be generated algebraically from the reaction rates, which bootstraps the generation of a group of graphs by way of a suitable graph product operation. A graph (shuffle) product can be defined between two Feynman graphs $\mathcal{G}(\mathbf{E}, \mathcal{C})$ (or sets thereof) and a suitable termination condition can be used when iteratively generating all relevant graphs to some given loop order. Quantities such as symmetry factors are easily computed from the incidence matrix of the graph. I have prototyped parts of the basic algorithm in `Python`.

5.2 Parametric Feynman rules for out-of-equilibrium graphs

My initial motivation for exploring a parametric representation for Feynman graphs in non-equilibrium field theories was to have more powerful methods to efficiently compute graphs at higher loop order and for graphs produced for complex chemical reaction networks. The derivation in the non-equilibrium case and application to spatial Markov processes appears for the first time in this thesis. It turned out however, I could find little motivation in the literature either to (i) carry out calculations beyond one-loop or (ii) consider a multitude of diagrams, beyond the ones already considered for a few well-known processes in the

field. Nonetheless, the parametric representation has an elegant form due to the fact that the frequency variable integrations, the fingerprint of causality in non-equilibrium field theories, can be represented as delta functions in the alpha parameters. The parametric form is particularly concise and well suited to symbolic methods.

$$\mathcal{I}(G) = \Omega_d \cdot \Gamma(-\sigma_d(G')) \int_{[\alpha]} \frac{\prod_e \alpha_e^{n_e-1}}{\Gamma(n_e)} \frac{1}{(\Psi')^{d/2}} \left(\frac{\Psi'}{\Phi'} \right)^{-\sigma_d(G')} \cdot \prod_L \delta(\Sigma_l) \quad (5.1)$$

While parametric integrals (and parametric renormalization) in the context of the coherent state path integral formalism may be interesting to study for their own sake, a suitable physical problem is perhaps required to motivate it. This may be tied in turn to the generality and effectiveness of the formalism for *arbitrary* Markov processes as called into question above.

5.3 Efficient sandpile sampling

I presented results in Chapter 3 on the scaling of observables in the Abelian sandpile [Dhar, 1990] for 2D lattices $L \times L$ up to $L = 32 \times 10^3$. The developed algorithm exploits two primary results in the literature. The first, a bijection due to Majumdar and Dhar [1992] demonstrated that there exists an algorithm that maps any recurrent configuration to a spanning tree and *vice versa*. The second is Wilson’s fast algorithm for generating spanning trees [Wilson, 1996]. An algorithm to simply generate a spanning tree is not enough for efficient Monte Carlo sampling. Rather, one needs to *uniformly* sample spanning trees of a graph. Wilson’s algorithm does this sampling faster than the *cover time* of the graph. Treating the spanning tree as an arborescence rooted at v_∞ , a C algorithm has been developed to perform a breadth first traverse of the vertices of the tree to compute the height variable z_i in the recurrent configuration. While it takes $\sim L^4$ *drives* to equilibrate into a recurrent configuration via ordinary ASM dynamics initialised on an empty lattice, the developed algorithm generates recurrent configurations on the order of the cover time, orders of magnitude faster. The motivation for doing this, beyond the need for speed, is to achieve good statistics. Generally, in ASM simulations, increasingly large lattices show little convergence and scaling seems to suffer from significant finite size corrections and corrections to scaling. The precise nature of these corrections is not known, yet is normally assumed to be of power-law or logarithmic form. Thus larger system sizes must be used in numerical studies, which require more CPU and RAM. The efficient Monte Carlo sampling algorithm developed in this work aims to address these concerns and presents up-to-date results on scaling in large systems. The spanning tree bijection leads to a theoretically fast approach and a careful implementation minimises the memory footprint.

5.4 Scaling of superprocess trace on graphs

Chapter 3 outlined results from our field-theoretic study of the volume explored by a superprocess on arbitrary graphs [Bordeu et al., 2019]. By exploiting results about the *spectral dimension* of general graphs, we have been able to extend the results of a renormalization group analysis of the critical scaling behaviour in $\varepsilon = d - 4$ dimensions to derive the scaling

behaviour of the time-dependant trace of the superprocess on arbitrary graphs. This is of practical interest, for example in studying the spread of disease on social networks.

For a graph with non-integer dimension and L vertices, the moment-scaling with respect to the *trace of the superprocess* is found to be (with $L = \hat{L}^{1/d_s}$ giving the network size)

$$\langle a^n \rangle(t, \hat{L}) \propto \begin{cases} t^{(nd_s-2)/2} & \text{for } Dt \ll \hat{L}^{2/d_s} \\ \hat{L}^{(n-2/d_s)} & \text{for } Dt \gg \hat{L}^{2/d_s} \end{cases} \quad (5.2a)$$

This result has been validated using numerical simulation on graphs, using graph generating algorithms discussed in Chapter 3.

5.5 Fuzzy tracking of root cells

In Chapter 4 I presented a new tracking method in the context of a sample dataset where mitotic events were tracked in *Arabidopsis* roots. Challenges arise due to the transient nature of fluorescent events and variation in image properties in light-sheet microscopy experiments lasting up to one week. For example, the data exhibit a lack of guaranteed, permanent features in image sequences. These challenges influenced the design of the tracking algorithm and image pre-processing stages.

The lack of fiducial markers requires the construction of fuzzy markers through sub-sampling of point clouds. A suitable cost function evaluates affine transformations between fuzzy fiducial markers. This approach is distinct from standard registration or single particle tracking methods, which put greater emphasis on the existence of a constant signal.

The algorithm fills a gap in tracking methods and should be particularly useful in the context of light-sheet microscopy data. In a recent review, Berthet and Maizel [2016] observe the under-utilisation in plant biology of light-sheet fluorescence microscopy despite its potential, stating in particular the desirable properties of fast imaging and low phototoxicity over extended imaging periods. They outline salient challenges in processing light-sheet microscopy images, some of which have been dealt with in our fuzzy tracking method.

5.6 Quantifying cancer cell motility

In Chapter 4, methods commonly used in statistical physics were used as part of a study of acute myeloid leukaemia (AML) cells. In contrast with T-cell acute lymphoblastic leukaemia (T-ALL), chemoresistant AML cells become less motile. Unlike T-ALL, the *in vivo* exploratory behaviour of chemoresistant AML cells is unaffected by exposure to chemotherapy. The results expand our understanding of AML cells in bone marrow micro-environment interactions, highlighting unique traits of leukaemia of different lineages.

A number of methods were used to compare spatiotemporal dynamics for different cell lineages before and after treatment (fitting to distributions, fractal measurements, mean squared displacement). Focusing on the mean squared displacement, I summarise

the interesting result that motility can be characterised on a diffusivity spectrum i.e. sub/super/diffusive. These regimes characterise spatiotemporal dynamics in a manner that admits useful physical and biological interpretation; linear diffusion describes pure random motion. If particles are behaving sub-diffusively this suggests that they are spending more time crowding locally rather than moving as freely as a Brownian particle. If particles are super-diffusive, they move more ballistically than Brownian particles. In the case of leukaemia cells for example, accumulation of cells is not favoured. One prefers to disperse cancer cells from the bone marrow.

In closing

While at times the work has covered methodologies that appear disturbingly disparate, it is perhaps an honest reflection of the contemporary applied sciences. We are increasingly obliged, as individuals or in collaborative efforts, to tackle real-world problems from the two fronts of theoretical models and real data.

In this age of increasingly large and varied data sets and rapidly developing experimental techniques, we are still learning how to assimilate varieties of data and different methodologies. In the context of analysing spatiotemporal biological processes by both theoretical and empirical means, it would be interesting in future to "close the loop"; starting with observations of spatiotemporal processes and reaction networks in biology (modelled as compound Poisson chemical reaction networks), compare real data with a (field theoretic) analytical quantification of how diffusion (on graphs) influence interactions, which in turn give rise to global patterns. For example in encountering new data for an arbitrary biological reaction network for which microscopic interactions are mapped out, one hopes to "quickly" make predictions about the impact (on particle numbers over time) of diffusive reactant transport in space or on graphs. It appears there is still much work to do on all fronts in order to achieve this. Yet in principle, being able to characterise the impact of the geometry of space, on arbitrary chemical reaction networks, justifies the use of sophisticated field-theoretic tools when developing analytical insight beyond the mean-field.

Outlook

This section is necessarily speculative. With regards to my view of the future, there are avenues both in terms of applications and abstractions, which might be pursued. Clearly, with the volume of data that is now being produced in biology, there are opportunities to apply the various data analysis methodologies discussed in this work more widely. This applies in the case of microscopy imagery, giving rise to data analysis problems of the first and second kind. However I do not offer any further commentary on this avenue.

There are two aspects of the work, which I would look to abstract and develop further and I have alluded to these throughout the work; on the one hand, the quantization of generalised (spatial) chemical reaction networks on graphs and on the other hand, algebraic and algorithmic approaches to the renormalization of such processes.

On the latter, it would be interesting to consider parametric integrals and the combinatorics of graphs that can be generated by reaction-diffusion processes having M species and reactions $kA \rightarrow lA : k + l \leq 3$. Addressing this as an abstract procedure might lead to a maturation of the formalism for wider application in biology. There are many tricks in the literature for dealing with Feynman integrals and I would like to enumerate solutions for the diagrams encountered from perturbative expansions of any multi-species chemical reaction network action. Given their formulation as Markov processes (consider the form of the propagators and how they are combined), these diagrams will be restricted to a specific form. It is possible to algorithmically generate all parametric integrals, which then need to be integrated. While many are easily integrating, integrating *any* such integral at two loops for example requires further research. It might be sufficient for some of the analysis to write down the form of the integral and simply Taylor expand the *quasimass* as was done in Chapter 3, which simplifies things dramatically.

On the former, the coherent state path integral seems to represent a particular *instance* of a more a general framework, which does not seem to me to be well understood. It is difficult to make concise and objective statements because this depends on the individual's taste and particular training. Consider an arbitrary physical problem e.g. reaction diffusion processes on lattices or random walks in a potential. In probability (stochastic processes) or statistical physics, these problems can typically be regarded as a problem of enumerating/counting states and weighted transitions between states. Calculus and functional analysis have been the primary tool here and in some cases, formal power series (of *binomial type*) have been used as a basis for (to find solutions to) such problems. Emphasising the combinational nature of the problem, one seeks generating functions that yield the statistics of interest. In the special case of the coherent state path integral, the factorial moment generating function was the appropriate choice. However, it appeared as though it was convenient (or necessary perhaps) to map this generating function to a canonical basis, namely the ordinary generating function. For each generating function (basis) there is a dual Heisenberg-Weyl algebra that "diagonalizes" the basis. In mapping between bases, one can consider the modification to this dual algebra. In the special case of the coherent state path integral formalism, from (1.28) we observed that a shift to the operator $z \rightarrow z + 1$ allowed us to express the *factorial moment generating function* as a (canonical) *ordinary generate function*, from which expansions in Feynman diagrams followed. This is not how it is explained in the literature, where the *doi-shift*, which produces the same outcome, is often mysteriously introduced as a convenience. Appropriate tools and language must be chosen to carry out the generalisation of the coherent state path integral formalism and I would look to carry out research using methods from combinatorics for the following reason; (i) As combinatorics (generating functions, *symbolic method* [Flajolet and Sedgewick, 2009], *umbra* [Roman, 2005] and *species* Bergeron et al. [1998]) focus directly on the enumeration of discrete states it seems a very direct and appropriate language to describe process quantisation on lattices. (ii) Some of these powerful methods in combinatorics have only been developed since the 1970s (after the techniques used in coherent state path integral were developed) and are not well known in physics. (iii) The language and considerations of quantum field theory comes with much baggage that is perhaps superfluous for classical processes.

Through a generalisation of the coherent state path integral formalism, one can in principle use the diagrammatic perturbation theory for a wide class of problems. Here I ask, *if* for any Markovian reaction network we could easily (e.g. algorithmically) generate the

perturbation series and the corresponding Feynman loops (at least at one loop) so as to carry out the renormalization group program, could we make predictions about species concentrations for relevant biological processes? My suggestion for next steps is simply this; select and simulate a number of chemical reaction networks as studied by followers of the Feinberg [1972] school, say. Those of interest should be Markovian (compound Poisson). Generate the Lagrangian and perturbation series (diagrams) and the one-loop parametric integrals. While this would normally be a tedious procedure by hand, following the techniques discussed in this thesis, much of it can be automated. From here the objective is to make sense of the perturbation series and renormalized amplitudes in the context of what is known about the chemical reaction network. For example what can we say about the renormalization of complex balanced networks in a volume?

Becoming even more speculative, in the distant future, it is not hard to imagine very high resolution microscopy imaging that can be used to track perhaps tens of thousands of "particles" of different species *in vivo*. In this way we might observe entire biological reaction networks *in vivo*. At this point having access to analytical models that describe the dynamics and changing concentration of particle species, as they diffuse in geometric space, might be particularly interesting.

Bibliography

- David J Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM Journal on Discrete Mathematics*, 1990.
- Shlomo Alexander and Raymond Orbach. Density of states on fractals: «fractons». *Journal de Physique Lettres*, 43(17):625–631, 1982.
- Linda JS Allen. *An introduction to stochastic processes with applications to biology*. CRC Press, 2010.
- Saoirse Amarteifio and Gunnar Pruessner. Exploiting dhar’s spanning tree bijection and wilson’s algorithm to compute abelian sandpile statistics efficiently. *in preparation*, 2019.
- Saoirse Amarteifio, Todd Fallesen, Giovanni Sena, and Gunnar Pruessner. Tracking transient fluorescent events in structured point clouds. *bioRxiv*, page 419572, 2018.
- Krishna B Athreya, Peter E Ney, and PE Ney. *Branching processes*. Courier Corporation, 2004.
- Louis Bachelier. *Théorie de la spéculation*. Gauthier-Villars, 1900.
- Paolo Baesso, Ricardo S Randall, and Giovanni Sena. Light sheet fluorescence microscopy optimized for long-term imaging of arabidopsis root development. In *Root Development*, pages 145–163. Springer, 2018.
- John C Baez and Jacob Biamonte. Quantum techniques for stochastic mechanics. *arXiv preprint*, 2012.
- Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of the $1/f$ noise. *Physical Review Letters*, 1987.
- Andrea Baronchelli, Michele Catanzaro, and Romualdo Pastor-Satorras. Bosonic reaction-diffusion processes on scale-free networks. *Physical Review E*, 78(1):016111, 2008.
- Nicolas Behr, Gérard HE Duchamp, and Karol A Penson. Combinatorics of chemical reaction systems. *arXiv preprint*, 2017.
- François Bergeron, F Bergeron, Gilbert Labelle, Pierre Leroux, et al. *Combinatorial species and tree-like structures*, volume 67. Cambridge University Press, 1998.
- Beatrice Berthet and Alexis Maizel. Light sheet microscopy and live imaging of plants. *Journal of Microscopy*, 263(2):158–164, 2016.

- Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor Fusion IV: Control Paradigms and Data Structures*, volume 1611, pages 586–607. International Society for Optics and Photonics, 1992.
- Serge Beucher. Watershed, hierarchical segmentation and waterfall algorithm. In *Mathematical morphology and its applications to image processing*, pages 69–76. Springer, 1994.
- Irénée-Jules Bienaymé. De la loi de multiplication et de la durée des familles. *Soc. Philomat. Paris Extraits, Sér.*, 5:37–39, 1845.
- Pawel Blasiak. *Combinatorics of boson normal ordering and some applications*. PhD thesis, Université Pierre et Marie Curie, 2005.
- Pawel Blasiak and Philippe Flajolet. Combinatorial models of creation-annihilation. *Séminaire Lotharingien de Combinatoire*, 65(B65c):1–78, 2011.
- Christian Bogner. *Mathematical aspects of Feynman integrals*. PhD thesis, Mainz, 2009.
- Christian Bogner and Stefan Weinzierl. Feynman graph polynomials. *International Journal of Modern Physics A*, 25(13):2585–2618, 2010.
- NN Bogoliubov and OS Parasiuk. On the multiplication of the causal function in the quantum theory of fields. *Acta Mathematica*, 97:227–266, 1957.
- Ignacio Bordeu, Saoirse Amarteifio, Rosalba Garcia-Millan, Benjamin Walter, Nanxin Wei, and Gunnar Pruessner. Volume explored by a branching random walk on general graphs. *submitted*, 2019.
- Michael Borinsky. *Graphs in perturbation theory: Algebraic structure and asymptotics*. PhD thesis, Humboldt-Universität zu Berlin, 2018.
- Andrei Broder. Generating random spanning trees. In *Foundations of Computer Science, 1989., 30th Annual Symposium on*, pages 442–447. IEEE, 1989.
- Francis Brown. The massless higher-loop two-point function. *Communications in Mathematical Physics*, 287(3):925–958, 2009.
- Francis Brown. Periods and feynman amplitudes. *arXiv preprint*, 2015.
- Francis Brown and Dirk Kreimer. Angles, scales and parametric renormalization. *Letters in Mathematical Physics*, 103(9):933–1007, 2013.
- Robert Brown. A brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.

- Raffaella Burioni and Davide Cassi. Random walks on graphs: ideas, techniques and results. *Journal of Physics A: Mathematical and General*, 38(8):R45, 2005.
- Krzysztof Burnecki and Aleksander Weron. Fractional lévy stable motion can model sub-diffusive dynamics. *Physical Review E*, 82(2):021130, 2010.
- Stas Burov, Jae-Hyung Jeon, Ralf Metzler, and Eli Barkai. Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking. *Physical Chemistry Chemical Physics*, 13(5):1800–1812, 2011.
- John Cardy. Reaction-diffusion processes. *Lectures held at Warwick University, unpublished*, 2006.
- John L Cardy and RL Sugar. Directed percolation and reggeon field theory. *Journal of Physics A: Mathematical and General*, 13(12):L423, 1980.
- Pierre Cartier. Mathemagics. In *Noise, oscillators and algebraic randomness*, pages 6–67. Springer, 2000.
- Michele Catanzaro, Marián Boguñá, and Romualdo Pastor-Satorras. Diffusion-annihilation processes in complex networks. *Physical Review E*, 71(5):056104, 2005.
- Seth Chaiken. A combinatorial proof of the all minors matrix tree theorem. *SIAM Journal on Algebraic Discrete Methods*, 3(3):319–329, 1982.
- Alexei V Chechkin, Ralf Metzler, Joseph Klafter, Vsevolod Yu Gonchar, et al. Introduction to the theory of lévy flights. *Anomalous transport: Foundations and applications*, pages 129–162, 2008.
- Nicolas Chenouard, Ihor Smal, Fabrice De Chaumont, Martin Maška, Ivo F Sbalzarini, Yuanhao Gong, Janick Cardinale, Craig Carthel, Stefano Coraluppi, Mark Winter, et al. Objective comparison of particle tracking methods. *Nature Methods*, 2014.
- Alain Connes and Dirk Kreimer. Hopf algebras, renormalization and noncommutative geometry. In *Quantum field theory: perspective and prospective*, pages 59–109. Springer, 1999.
- Alain Connes and Dirk Kreimer. Renormalization in quantum field theory and the riemann–hilbert problem i: The hopf algebra structure of graphs and the main theorem. *Communications in Mathematical Physics*, 210(1):249–273, 2000.
- Alain Connes and Dirk Kreimer. Renormalization in quantum field theory and the riemann–hilbert problem ii: The β -function, diffeomorphisms and the renormalization group. *Communications in Mathematical Physics*, 216(1):215–241, 2001.
- Alain Connes and Matilde Marcolli. *Noncommutative geometry, quantum fields and motives*, volume 55. American Mathematical Soc., 2008.

- Robert Cori and Yvan Le Borgne. The sand-pile model and tutte polynomials. *Advances in Applied Mathematics*, 2003.
- Lngemar J Cox and Sunita L Hingorani. An efficient implementation of reid’s multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1996.
- Donald A Dawson. Introductory lectures on stochastic population systems. *arXiv preprint*, 2017.
- C De Dominicis, E Brézin, and J Zinn-Justin. Field-theoretic techniques and critical dynamics. i. ginzburg-landau stochastic models without energy conservation. *Physical Review B*, 12(11):4945, 1975.
- Michael W Deem and Jeong-Man Park. Effect of static disorder and reactant segregation on the $a + b \rightarrow 0$ reaction. *Physical Review E*, 57(3):2681, 1998.
- Max Delbrück. Statistical fluctuations in autocatalytic reactions. *The Journal of Chemical Physics*, 8(1):120–124, 1940.
- Nicolas Destainville and Laurence Salomé. Quantification and correction of systematic errors due to detector time-averaging in single-molecule tracking experiments. *Biophysical Journal*, 90(2):L17–L19, 2006.
- Deepak Dhar. Self-organized critical state of sandpile automaton models. *Physical Review Letters*, 1990.
- Reinhard Diestel. *Graph theory (4th ed.)*. Springer-Verlag Heidelberg, New York, 2010.
- Masao Doi. Second quantization representation for classical many-particle system. *Journal of Physics A: Mathematical and General*, 9(9):1465, 1976a.
- Masao Doi. Stochastic theory of diffusion-controlled reaction. *Journal of Physics A: Mathematical and General*, 9(9):1479, 1976b.
- David L Donoho and Jain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.
- MD Donsker and SRS Varadhan. Asymptotics for the wiener sausage. *Communications on Pure and Applied Mathematics*, 28(4):525–565, 1975.
- Max Dresden. Renormalization in historical perspective: the first stage. In *Renormalization*, pages 29–55. Springer, 1993.
- Michel Droz and A McKane. Equivalence between poisson representation and fock space formalism for birth-death processes. *Journal of Physics A: Mathematical and General*, 1994.

- Michel Droz and László Sasvári. Renormalization-group approach to simple reaction-diffusion phenomena. *Physical Review E*, 48(4):R2343, 1993.
- Delfim Duarte, Saoirse Amarteifio, Heather Ang, Isabella Y Kong, Nicola Ruivo, Gunnar Pruessner, Edwin D Hawkins, and Cristina Lo Celso. Defining the in vivo characteristics of acute myeloid leukemia cells behavior by intravital imaging. *Immunology and Cell Biology*, 2018.
- EB Dynkin. Markov processes as a tool in field theory. *Journal of Functional Analysis*, 50(2):167–187, 1983.
- Albert Einstein. Investigations on the theory of the brownian movement. *Annalen der Physik*, 1905.
- David Elderfield. Exact macroscopic dynamics in non-equilibrium chemical systems. *Journal of Physics A: Mathematical and General*, 18(11):2049, 1985.
- Ido Eliazar and Joseph Klafter. Anomalous is ubiquitous. *Annals of Physics*, 326(9):2517–2531, 2011.
- Lindsay CC Elliott, Moussa Barhoum, Joel M Harris, and Paul W Bohn. Trajectory analysis of single molecules exhibiting non-brownian motion. *Physical Chemistry Chemical Physics*, 13(10):4326–4334, 2011.
- Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194. Springer Science & Business Media, 1999.
- Alison Etheridge. *An introduction to superprocesses*. American Mathematical Soc., 2000.
- Kenneth Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.
- Martin Feinberg. Complex balancing in general kinetic systems. *Archive for Rational Mechanics and Analysis*, 49(3):187–194, 1972.
- Richard P Feynman, Robert B Leighton, and Matthew Sands. *Lectures on Physics, vol. III*. Addison-Wesley Reading, MA, 1965.
- Adolph Fick. On liquid diffusion. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(63):30–39, 1855.
- Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- Andrew W Fitzgibbon. Robust registration of 2d and 3d point sets. *Image and Vision Computing*, 21(13-14):1145–1153, 2003.

- Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009.
- Martin Friesen. *Semigroup methods for spatial birth-and-death processes*. PhD thesis, Universität Bielefeld, 2016.
- Janusz Gajda and Marcin Magdziarz. Fractional fokker-planck equation with tempered α -stable waiting times: Langevin picture and computer simulation. *Physical Review E*, 82(1):011117, 2010.
- Lazaros K Gallos and Panos Argyrakis. Absence of kinetic effects in reaction-diffusion processes in scale-free networks. *Physical Review Letters*, 92(13):138301, 2004.
- CW Gardiner and S Chaturvedi. The poisson representation. i. a new technique for chemical master equations. *Journal of Statistical Physics*, 1977.
- Roy J Glauber. Coherent and incoherent states of the radiation field. *Physical Review*, 131(6):2766, 1963.
- William J Godinez and Karl Rohr. Tracking multiple particles in fluorescence time-lapse microscopy images via probabilistic data association. *Medical Imaging, IEEE Transactions on*, 2015.
- Marcel Golz. Private Communication, 2018.
- Georg A Gottwald and Ian Melbourne. On the detection of superdiffusive behaviour in time series. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(12):123205, 2016.
- Peter Grassberger. On phase transitions in schlögl’s second model. *Zeitschrift für Physik B Condensed Matter*, 47(4):365–374, 1982.
- Peter Grassberger and M Scheunert. Fock-space methods for identical classical objects. *Fortschritte der Physik*, 28(10):547–578, 1980.
- Peter Grassberger and Kurt Sundermeyer. Reggeon field theory and markov processes. *Physics Letters B*, 77(2):220–222, 1978.
- George Green. Researches on the vibration of pendulums in fluid media. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, 13(1):54–62, 1835.
- VN Gribov. A reggeon diagram technique. *Sov. Phys. JETP*, 26(414-422):27, 1968.
- Gernot Guigas and Matthias Weiss. Sampling the cell with anomalous diffusion - the discovery of slowness. *Biophysical Journal*, 94(1):90–94, 2008.
- Peter N Gutter. *Stochastic modeling of scientific data*. CRC Press, 1995.

- John Harris, Jeffry L Hirst, and Michael Mossinghoff. *Combinatorics and Graph Theory*. Springer Science & Business Media, New York, 2009.
- Tajie H Harris, Edward J Banigan, David A Christian, Christoph Konradt, Elia D Tait Wojno, Kazumi Norose, Emma H Wilson, Beena John, Wolfgang Weninger, Andrew D Luster, et al. Generalized lévy walks and the role of chemokines in migration of effector cd8+ t cells. *Nature*, 486(7404):545, 2012.
- Theodore E Harris. Contact interactions on a lattice. *The Annals of Probability*, 1974.
- Theodore E Harris. *The theory of branching processes*. Courier Corporation, 2002.
- Thomas Hawkins. *Emergence of the theory of Lie groups: An essay in the history of mathematics 1869–1926*. Springer Science & Business Media, 2012.
- Y He, S Burov, Ralf Metzler, and E Barkai. Random time-scale invariant diffusion and transport coefficients. *Physical Review Letters*, 101(5):058101, 2008.
- Gudrun Heinrich. Sector decomposition. *International Journal of Modern Physics A*, 23(10):1457–1486, 2008.
- Klaus Hepp. Proof of the bogoliubov-parasiuk theorem on renormalization. *Communications in Mathematical Physics*, 2(1):301–326, 1966.
- David Hochberg and M-P Zorzano. Path integral evaluation of the one-loop effective potential in field theory of diffusion-limited reactions. *Physica A: Statistical Mechanics and its Applications*, 378(2):238–254, 2007.
- Gerard’t Hooft. Dimensional regularization and the renormalization group. *Nuclear Physics: B*, 61:455–468, 1973.
- Gerard’t Hooft and Martinus Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44(1):189–213, 1972.
- Fritz Horn and Roy Jackson. General mass action kinetics. *Archive for Rational Mechanics and Analysis*, 47(2):81–116, 1972.
- Claude Itzykson and Jean-Michel Drouffe. *Statistical field theory: Volume 1, From Brownian Motion to Renormalization and Lattice Gauge Theory*, volume 2. Cambridge University Press, 1991.
- Claude Itzykson and Jean-Bernard Zuber. *Quantum field theory*. McGraw-Hillm New York, 1980.
- David M Jackson, Achim Kempf, and Alejandro H Morales. Combinatorial aspects of the legendre and fourier transforms in perturbative quantum field theory. *arXiv preprint*, 2018.
- Hans-Karl Janssen. On the nonequilibrium phase transition in reaction-diffusion systems

- with an absorbing stationary state. *Zeitschrift für Physik B Condensed Matter*, 42(2):151–154, 1981.
- Antal A Járai. Thermodynamic limit of the abelian sandpile model on $\mathbb{Z}^d$. *Markov Process. Related Fields*, 2005.
- Iwan Jensen and Ronald Dickman. Series analysis of the generalized contact process. *Physica A: Statistical Mechanics and its Applications*, 203(2):175–188, 1994.
- Bing Jian and Baba C Vemuri. Robust point set registration using gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–1645, 2011.
- Leo P Kadanoff. Scaling laws for ising models near T_c . *Physics Physique Fizika*, 2(6):263, 1966.
- Alex Kamenev. Keldysh and doi-peliti techniques for out-of-equilibrium systems. In *Strongly Correlated Fermions and Bosons in Low-Dimensional Disordered Systems*, pages 313–340. Springer, 2002.
- Alex Kamenev. *Field theory of non-equilibrium systems*. Cambridge University Press, 2011.
- Takeo Kanade, Zhaozheng Yin, Ryoma Bise, Seungil Huh, Sungeun Eom, Michael F Sandbothe, and Mei Chen. Cell image analysis: Algorithms, system and applications. In *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pages 374–381. IEEE, 2011.
- K Kang. K. kang and s. redner, phys. rev. lett. 52, 955 (1984). *Physical Review Letters*, 52:955, 1984.
- Leonid Veniaminovich Keldysh. Diagram technique for nonequilibrium processes. *Journal of Experimental and Theoretical Physics*, 47:1018, 1964.
- Jonathan A Kelner and Aleksander Madry. Faster generation of random spanning trees. In *Foundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on*, 2009.
- David G Kendall. On the generalized “birth-and-death” process. *The annals of mathematical statistics*, pages 1–15, 1948.
- Eldad Kepten, Irena Bronshtein, and Yuval Garini. Improved estimation of anomalous diffusion exponents in single-particle tracking experiments. *Physical Review E*, 87(5):052713, 2013.
- Eldad Kepten, Aleksander Weron, Grzegorz Sikora, Krzysztof Burnecki, and Yuval Garini. Guidelines for the fitting of anomalous diffusion mean square displacement graphs from single particle tracking experiments. *PLOS ONE*, 10(2):e0117722, 2015.

- Gustav Kirchhoff. Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird. *Annals of Physics*, 148(12):497–508, 1847.
- Hagen Kleinert. *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*. World scientific, 2009.
- Hagen Kleinert and Verena Schulte-Frohlinde. *Critical Properties of ϕ^4 theories*. World Scientific, 2001.
- Mikhail Kompaniets and Erik Panzer. Renormalization group functions of ϕ^4 theory in the ms-scheme to six loops. *arXiv preprint*, 2016.
- Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical review E*, 69(6):066138, 2004.
- Supriya Krishnamurthy and Eric Smith. Solving moment hierarchies for chemical reaction networks. *Journal of Physics A: Mathematical and Theoretical*, 50(42):425002, 2017.
- Lev Davidovich Landau. On the theory of phase transitions. *Journal of Experimental and Theoretical Physics*, 7:19–32, 1937.
- Gregory F Lawler and Vlada Limic. *Random walk: a modern introduction*. Cambridge University Press, 2010.
- Benjamin P Lee. Renormalization group calculation for the reaction ka to oe. *Journal of Physics A: Mathematical and General*, 27(8):2633, 1994.
- Thomas Milton Liggett. *Interacting particle systems*. Springer Science & Business Media, 2012.
- Katja Lindenberg, Wen-Shyan Sheu, and Raoul Kopelman. Scaling properties of diffusion-limited reactions on fractal and euclidean geometries. *Journal of Statistical Physics*, 65(5-6):1269–1283, 1991.
- An-An Liu, Yao Lu, Mei Chen, and Yu-Ting Su. Mitosis detection in phase contrast microscopy image sequences of stem cell populations: A critical review. *IEEE Transactions on Big Data*, 3(4):443–457, 2017.
- Jesper Lützen. *The prehistory of the theory of distributions*. Springer Science & Business Media, 2012.
- Satya N Majumdar and Deepak Dhar. Equivalence between the abelian sandpile model and the q-0 limit of the potts model. *Physica A: Statistical Mechanics and its Applications*, 185(1):129–145, 1992.
- Rosario N Mantegna and H Eugene Stanley. Stochastic process with ultraslow convergence to a gaussian: the truncated lévy flight. *Physical Review Letters*, 73(22):2946, 1994.

- Paul Cecil Martin, ED Siggia, and HA Rose. Statistical dynamics of classical systems. *Physical Review A*, 8(1):423, 1973.
- Naoki Masuda, Mason A Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Physics Reports*, 716:1–58, 2017.
- Daniel C Mattis and M Lawrence Glasser. The uses of quantum field theory in diffusion-limited reactions. *Reviews of Modern Physics*, 70(3):979, 1998.
- Michel Le Bellac. *Quantum and statistical field theory*. Clarendon Press Oxford, 1991.
- AG McKendrick. Studies on the theory of continuous probabilities, with special reference to its bearing on natural phenomena of a progressive nature. *Proceedings of the London Mathematical Society*, 2(1):401–416, 1914.
- AG McKendrick. Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, 44:98–130, 1925.
- Vicenç Méndez, Daniel Campos, and Frederic Bartumeus. *Stochastic foundations in movement ecology: anomalous diffusion, front propagation and random searches*. Springer Science & Business Media, 2013.
- Yasmine Meroz and Igor M Sokolov. A toolbox for determining subdiffusive mechanisms. *Physics Reports*, 573:1–29, 2015.
- Xavier Michalet. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Physical Review E*, 82(4):041914, 2010.
- José Miotto. pylevy. <https://github.com/josemiotto/pylevy>, 2018.
- Elliott W Montroll and George H Weiss. Random walks on lattices. ii. *Journal of Mathematical Physics*, 6(2):167–181, 1965.
- Noboru Nakanishi. *Graph theory and Feynman integrals*. Gordon and Breach, New York, 1971.
- John W Negele and Henri Orland. *Quantum many-particle systems*. Addison-Wesley, 1988.
- Stefan Nekovar and Gunnar Pruessner. A field-theoretic approach to the wiener sausage. *Journal of Statistical Physics*, 163(3):604–641, 2016.
- J. P. Nolan. *Stable Distributions - Models for Heavy Tailed Data*. Birkhauser, Boston, 2017. In progress, Chapter 1 online at fs2.american.edu/jpnolan/www/stable/stable.html.
- A Nordsieck, WE Lamb Jr, and GE Uhlenbeck. On the theory of cosmic-ray showers i the furry model and the fluctuation problem. *Physica*, 7(4):344–360, 1940.

- SP Obukhov. The problem of directed percolation. *Physica A: Statistical Mechanics and its Applications*, 101(1):145–155, 1980.
- Jun Ohkubo. One-parameter extension of the doi-peliti formalism and its relation with orthogonal polynomials. *Physical Review E*, 86(4):042102, 2012.
- Kenji Okuma, Ali Taleghani, Nando De Freitas, James J Little, and David G Lowe. A boosted particle filter: Multitarget detection and tracking. In *Computer Vision-ECCV 2004*. Springer, 2004.
- Peter J Olver. *Applications of Lie groups to differential equations*. Springer Science & Business Media, 2000.
- Otso Ovaskainen, Dmitri Finkelshtein, Oleksandr Kutoviy, Stephen Cornell, Benjamin Bolker, and Yuri Kondratiev. A general mathematical framework for the analysis of spatiotemporal point processes. *Theoretical Ecology*, 7(1):101–113, 2014.
- Erik Panzer. *Feynman integrals and hyperlogarithms*. PhD thesis, Humboldt Universität, 2015.
- L Peliti. Path integral approach to birth-death processes on a lattice. *Journal de Physique*, 46(9):1469–1483, 1985.
- L Peliti. Renormalisation of fluctuation effects in the $a + a \rightarrow a$ reaction. *Journal of Physics A: Mathematical and General*, 19(6):L365, 1986.
- David Perkinson, Qiaoyu Yang, and Kuai Yu. G-parking functions and tree inversions. *arXiv preprint*, 2013.
- Michael J Plank and Richard Law. Spatial point processes and moment dynamics in the life sciences: a parsimonious derivation and some extensions. *Bulletin of Mathematical Biology*, 77(4):586–613, 2015.
- Vyatcheslav B Priezzhev. Structure of two-dimensional sandpile. i. height probabilities. *Journal of Statistical Physics*, 1994.
- Gunnar Pruessner. *Self-organised criticality: theory, models and characterisation*. Cambridge University Press, 2012.
- Graham H Pyke. Understanding movements of organisms: it’s time to abandon the lévy foraging hypothesis. *Methods in Ecology and Evolution*, 6(1):1–16, 2015.
- A Quetelet. Sur quelques propriétés curieuses que présentent les résultats d’une série d’observations, faites dans la vue de déterminer une constante, lorsque les chances de rencontrer des écarts en plus et en moins sont égales et indépendantes les unes des autres. *Bulletins de l’Académie Royale Belge*, 19:303–317, 1852.
- G Venugopala Reddy, Marcus G Heisler, David W Ehrhardt, and Elliot M Meyerowitz.

- Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *arabidopsis thaliana*. *Development*, 131(17):4225–4237, 2004.
- Sidney Redner. *A guide to first-passage processes*. Cambridge University Press, 2001.
- Donald B Reid. An algorithm for tracking multiple targets. *Automatic Control, IEEE Transactions on*, 1979.
- Steven Roman. *The umbral calculus*. Springer, New York, 2005.
- Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- Jan Rosiński. Tempering stable processes. *Stochastic Processes and their Applications*, 117(6):677–707, 2007.
- Gian-Carlo Rota and Jianhong Shen. On the combinatorics of cumulants. *Journal of Combinatorial Theory*, 2000.
- Gian-Carlo Rota and Timothy C Wallstrom. Stochastic integrals: a combinatorial approach. *The Annals of Probability*, pages 1257–1283, 1997.
- Ernest Rutherford, Hans Geiger, and H Bateman. Lxxvi. the probability variations in the distribution of α particles. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 20(118):698–707, 1910.
- Mutsuo Saito and Makoto Matsumoto. Simd-oriented fast mersenne twister: a 128-bit pseudorandom number generator. In *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pages 607–622. Springer, 2008.
- Friedrich Schlögl. Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für physik*, 253(2):147–161, 1972.
- Dirk Schulz, Wolfram Burgard, Dieter Fox, and Armin B Cremers. Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, 2001.
- Julian Schwinger. Brownian motion of a quantum oscillator. *Journal of Mathematical Physics*, 2(3):407–432, 1961.
- Christine Selhuber-Unkel, Pernille Yde, Kirstine Berg-Sørensen, and Lene B Oddershede. Variety in intracellular diffusion during the cell cycle. *Physical Biology*, 6(2):025015, 2009.
- Giovanni Sena, Zak Frentz, Kenneth D Birnbaum, and Stanislas Leibler. Quantitation of cellular dynamics in growing arabidopsis roots with light sheet microscopy. *PLOS ONE*, 6(6):e21303, 2011.

- Michael F Shlesinger, Joseph Klafter, and YM Wong. Random walks with infinite spatial and temporal moments. *Journal of Statistical Physics*, 27(3):499–512, 1982.
- Sloane. The on-line encyclopedia of integer sequences (sloane). <http://oeis.org/A007341>, 1999. [Online; accessed 2015-04-04].
- Valdimir A Smirnov. *Feynman integral calculus*. Springer, Berlin, 2006.
- Vladimir A Smirnov. *Evaluating feynman integrals*. Springer, Berlin, 2005.
- Eric Smith and Supriya Krishnamurthy. Flows, scaling, and the control of moment hierarchies for stochastic chemical reaction networks. *Physical Review E*, 96(6):062102, 2017.
- Eric Smith and Supriya Krishnamurthy. Path-reversal, doi-peliti generating functionals, and dualities between dynamics and inference for stochastic processes. *arXiv preprint*, 2018.
- Marian Smoluchowski. Drei vortrage uber diffusion. brownsche bewegung und koagulation von kolloidteilchen. *Z. Phys.*, 17:557–585, 1916.
- ER Speer and MJ Westwater. Generic feynman amplitudes. In *Annales de l’IHP Physique théorique*, volume 14, pages 1–55. Gauthier-Villars, 1971.
- Aleksander Stanislavsky, Karina Weron, and Aleksander Weron. Diffusion and relaxation controlled by tempered α -stable processes. *Physical Review E*, 78(5):051106, 2008.
- Gianluca Stefanucci and Robert Van Leeuwen. *Nonequilibrium many-body theory of quantum systems: a modern introduction*. Cambridge University Press, 2013.
- Paul M Stevenson. Dimensional analysis in field theory. *Annals of Physics*, 132(2):383–403, 1981.
- Nico Stollenwerk and Vincent Jansen. *Population biology and criticality*. Imperial College Press, London, 2011.
- Daniel W Stroock. *An introduction to Markov processes*. Springer, New York, 2013.
- Kurt Symanzik. Euclidean quantum field theory. In *Local Qunatum Field theory*. Academic Press, New York, 1969.
- Oleg V Tarasov. Connection between feynman integrals having different values of the space-time dimension. *Physical Review D*, 54(10):6479, 1996.
- Uwe C Täuber. *Critical dynamics: a field theory approach to equilibrium and non-equilibrium scaling behavior*. Cambridge University Press, 2014.
- Uwe C Täuber, Martin Howard, and Benjamin P Vollmayr-Lee. Applications of field-

- theoretic renormalization group methods to reaction–diffusion problems. *Journal of Physics A: Mathematical and General*, 38(17):R79, 2005.
- Hale F Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, 1959.
- Vladimir V Uchaikin. Self-similar anomalous diffusion and levy-stable laws. *Physics-Uspekhi*, 46(8):821, 2003.
- Vladimir V Uchaikin and Vladimir M Zolotarev. *Chance and stability: stable distributions and their applications*. Walter de Gruyter, 2011.
- Vladimír Ulman, Martin Maška, Klas EG Magnusson, Olaf Ronneberger, Carsten Haubold, Nathalie Harder, Pavel Matula, Petr Matula, David Svoboda, Miroslav Radojevic, et al. An objective comparison of cell-tracking algorithms. *Nature Methods*, 14(12):1141, 2017.
- Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. scikit-image: image processing in Python. *PeerJ*, 2:e453, June 2014. ISSN 2167-8359. doi: 10.7717/peerj.453.
- Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*. Elsevier, 1992.
- Frédéric van Wijland. Field theory for reaction-diffusion processes with hard-core particles. *Physical Review E*, 63(2):022101, 2001.
- Aleksandr Nikolaevich Vasil’ev. *The field theoretic renormalization group in critical behavior theory and stochastic dynamics*. CRC press, 2004.
- Gandimohan M Viswanathan, Sergey V Buldyrev, Shlomo Havlin, MGE Da Luz, et al. Optimizing the success of random searches. *Nature*, 401(6756):911, 1999.
- Daniel von Wangenheim, Robert Hauschild, Matyáš Fendrych, Vanessa Barone, Eva Benkova, and Jiří Friml. Live tracking of moving samples in confocal microscopy for vertically grown roots. *eLife*, 6, 2017.
- Henry William Watson and Francis Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.
- Markus F Weber and Erwin Frey. Master equations and the theory of stochastic path integrals. *Reports on Progress in Physics*, 80(4):046601, 2017.
- Rud Wegscheider. Über simultane gleichgewichte und die beziehungen zwischen thermodynamik und reactionskinetik homogener systeme. *Monatshefte für Chemie und verwandte Teile anderer Wissenschaften*, 32(8):849–906, 1911.

- Steven Weinberg. High-energy behavior in quantum field theory. *Physical Review*, 118(3):838, 1960.
- Kay Jörg Wiese. Coherent-state path integral versus coarse-grained effective stochastic equation of motion: From reaction diffusion to stochastic sandpiles. *Physical Review E*, 93(4):042117, 2016.
- David Bruce Wilson. Generating random spanning trees more quickly than the cover time. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, 1996.
- Kenneth G Wilson. Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture. *Physical Review B*, 4(9):3174, 1971.
- Dominik Wöll, Christoph Kölbl, Beate Stempfle, and Andreas Karrenbauer. A novel method for automatic single molecule tracking of blinking molecules at low intensities. *Physical Chemistry Chemical Physics*, 15(17):6196–6205, 2013.
- Wenxia Ying, Gabriel Huerta, Stanly Steinberg, and Martha Zúñiga. Time series analysis of particle tracking data for molecular motion on the cell membrane. *Bulletin of Mathematical Biology*, 71(8):1967, 2009.
- CK Yun, B Kahng, and D Kim. Annihilation of two-species reaction–diffusion processes on fractal scale-free networks. *New Journal of Physics*, 11(6):063025, 2009.
- V Zaburdaev, S Denisov, and J Klafter. Lévy walks. *Reviews of Modern Physics*, 87(2):483, 2015.
- Yichao Zhang, Zhongzhi Zhang, Jihong Guan, and Shuigeng Zhou. Diffusion–annihilation processes in weighted scale-free networks with an identical degree sequence. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(10):P10001, 2011.
- Zhiyong Zhou, Jianfei Tu, Chen Geng, Jisu Hu, Baotong Tong, Jiansong Ji, and Yakang Dai. Accurate and robust non-rigid point set registration using students mixture model with prior probability modeling. *Scientific Reports*, 8(1):8742, 2018.
- W Zimmermann. Convergence of bogolyubov’s method of renormalization in momentum space 1969 comm. *Mathematical Physics*, 15:208, 1969.
- Wolfhart Zimmermann. The power counting theorem for minkowski metric. *Communications in Mathematical Physics*, 11(1):1–8, 1968.
- Jean Zinn-Justin. *Quantum field theory and critical phenomena*. Clarendon Press, 1996.
- Jean Zinn-Justin. *Phase transitions and renormalization group*. Oxford University Press, 2007.

Notes

¹**The Gribov process:** This process goes by different names. It was introduced by Gribov to model hadron interactions at very high energies i.e. *Reggeon field theory* [Gribov, 1968] but is also widely used to model certain reaction-diffusion processes on a lattice. It contains essential features of growth phenomena and exhibits a nonequilibrium continuous phase transition phenomenon in the transition between survival and extinction. The correspondence between Reggeon field theory and Markov lattice processes has been noted in the literature [Grassberger and Sundermeyer, 1978]. This process relates to *contact processes* [Harris, 1974, Jensen and Dickman, 1994], *directed percolation* [Obukhov, 1980, Cardy and Sugar, 1980], Schlögl's first reaction [Schlögl, 1972], cellular automata and many applications in the natural sciences. The universality class of local growth processes with absorbing states is referred to as the *directed percolation* class [Grassberger, 1982, Janssen, 1981]. The Gribov process was one of a number of processes used in an interesting approach that advertised the construction of effective field theories in reaction-diffusion systems in a manner that circumvents the constructions of Feynman diagrams [Hochberg and Zorzano, 2007].

²In quantum field theory **canonical quantization** considers *observables* of a quantum system as self-adjoint operators acting on a Hilbert space and the elements of the space are the *states* of the system. The probability that an observable $\mathcal{O}$ maps to a state n is given by the expectation $\langle n | \mathcal{O} | n \rangle \in \mathcal{R}$. Field operators are defined in tandem with the space in which they act. In this thesis, instead of a Hilbert space, I consider a formal power series $\mathcal{F}$ that gives statistics of states of a system and a corresponding differential algebra $[D, M]$ that acts on it.

³**Detailed balance:** Applying to reversibility in equilibrium systems, does not apply in general to systems out of equilibrium.

⁴**Contrasting language in the literature relating to Markov processes and their second quantization:** What I refer to here as the *infinitesimal generator* is referred to by Doi [1976a] as the *Liouvillian*. Also to be found in the literature is the (*quasi*-)*Hamiltonian*. The term *semigroup* is not to my knowledge used in the *coherent state path integral formalism* literature. The semigroup equation is often referred to as a "Schrödinger-type" equation. As Schrödinger's equation is driven by a *unitary* operator it is of course an example of a semigroup in quantum physics. The semigroup is frequently exploited in the study of Markov processes (cf. [Engel and Nagel, 1999]). It also appears in the terminology of *stochastic mechanics* [Baez and Biamonte, 2012].

⁵**Non markovian processes** can be described through Markov processes by enlarging the state space making the Markov theory widely useful.

⁶**A semigroup** is a group in the sense of a continuous one-parameter Lie group. In the *group analysis of differential equations*, an idea originally developed by Sophus Lie and given a more modern treatment by the Ovsiannikov school in Russian and also Birkhoff in the United States, one solves differential equations in a systematic way that exploits symmetries of the problem. This method subsumes the various *ad hoc* methods that have been developed for solving ordinary or partial differential equations over the centuries while also giving information about inherent Noether and Lie-Bäcklund symmetries. Lie's insight was that the one parameter of the group corresponds to the constant of integration. In the case of evolutionary Markov processes, the one parameter is time and in the renormalization group, a scale parameter is used. Corresponding to the time parameter, the Noether symmetry i.e. conserved quantity is energy. As we are interested in positive time $t \geq 0$ the term *semigroup* is used. The ideas of semigroups were developed by Hille and Yosida and a good introduction can be found in Engel and Nagel [1999]. Lie's method of solving differential equations is described in some detail by Olver [2000] and an interesting history of Lie's contribution to these ideas can be found in [Hawkins, 2012].

⁷This quantization picture as inspired by quantum mechanics/field theory in turn goes back to Dirac.

⁸The name **master equation** is often used in the natural sciences to describe the differential form of the Chapman-Kolmogorov equation. Early appearances of this language can be found in Nordsieck et al. [1940]. Van Kampen [1992] comments on the use of the word 'master' and also refers to this as the *M-equation*.

⁹**Euclidean/Statistical field theory:** I use two terms interchangeably. The ideas of statistical theory are collected for example in [Itzykson and Drouffe, 1991].

¹⁰**Second quantization** I quote the description given by Negele and Orland [1988]: *The quantum mechanics of a single particle is usually formulated in terms of the position operator $\hat{x}$ and the momentum operator $\hat{p}$. All other operators of physical interest may be expressed in terms of these operators, and a natural representation for quantum mechanics, the coordinate representation, is defined in terms of eigenfunctions of the position operator, [...], an analogous formalism is developed for systems composed of many identical particles. For these systems it is useful to define operators which create or annihilate a particle in specified states. Operators of physical interest may be expressed in terms of these creation and annihilation operators, in which case they are said to be expressed in "second quantized" form. The eigenstates of the annihilation operators are coherent states. A natural representation for the quantum mechanics of many-particle systems, the holomorphic representation, is defined in terms of these coherent states*• I use the language of Markov semigroups acting on formal power series, which are the generating functions corresponding to configurations such as those described by coherent states. The transition from the "coordinate representation" of single particles to the "coherent state representation" of many identical particles discussed by Negele and Orland [1988] is implemented by a change of generating function i.e. a transition to a *factorial moment* generating function.

¹¹**On the discovery and rediscovery of Fock space methods for classical systems:** Peliti [1985] in his exposition of Doi's approach [Doi, 1976a,b] mentioned a motivation to present the method due to the fact it had been rediscovered by others who were unaware of its existence. Indeed Grassberger and Scheunert [1980] who Peliti cites, develop a very similar approach mentioning only at the end of their article that they only later became aware of the earlier work of Doi [1976a].

¹²The word **quasiprobability** is often used as in this case by Gardiner and Chaturvedi [1977]. I am not sure this word needs to be used as perhaps the notion of a *generalised function* or a *distribution* in the sense of Laurent Schwartz should be used. An interesting history on the theory of distributions can be found in [Lützen, 2012].

¹³**Anomalous dimension** is "the difference" between the predicted scaling by mean-field (Landau) and effect of fluctuations.

¹⁴Historical notes and references for the **development of the path integral** can be found for example in [Kleinert, 2009] (ch. 2 notes).

¹⁵Compared to relativistic quantum field theory, **Euclidean field theory** is qualified by a *Wick rotation* to imaginary time, while the path measure in field theory is otherwise complex. Actually this is comparable to a shift from the Schrödinger equation to a diffusion equation which recovers Wiener's path measure.

¹⁶**The statistics of first passage probabilities and times** with respect to random walks and diffusion are discussed in textbooks on random processes and I do not review them in the main text. Burioni and Cassi [2005] provide a useful account when extending the idea to random walks an general graphs.

¹⁷**Fields:** A field is a vector valued function. It is reasonable to think of it as a distribution in the Laurent Schwartz sense. An observable is a real functional over the field. Relativistic field theories use Minkowski space-time coordinates. A wick rotation makes the transformation $t \rightarrow it$ to transform to Euclidean space-time coordinates and one considers these as statistical fields i.e. a distribution (or wave function) describing fluctuations around the expectation value.

¹⁸For example the non-linear *Klein-Gordon equation* can be approached in this manner. To see this observe that the Euler-Lagrange equation is a certain difference between operators vanishing (stationary solutions). We can equate the nonlinear(interacting) and bilinear(free) parts.

¹⁹Zinn-Justin [2007] emphasises **the important role of experiment in constructing the QFT:** *let us point out that for more than 15 years theoretical progress had been stopped by the problem of divergences in QFT. However, once experiment started producing decisive information, in two years a complete and consistent framework for perturbative calculations was set up.*

²⁰ The action S in our context corresponds to the generator of the semigroup of a compound Poisson process with a spatiotemporal propagation operator. This semigroup is approximated via the Lie-Kato-Trotter formula. In the first mapping, $\mathbb{P}$, the semigroup which was mapped algebraically in the mean-field sense is mapped onto a path integral. Considering a birth and death process, the mean field solution which is well known has also been considered in the spatial context and is known to give wave solutions. This is pre-field theory. In the field theory we wish to model both space and the impact of fluctuations when going beyond the mean-field. In principle there are infinite degrees of freedom in the system.

²¹**Benefits of parametric representations in out-of-equilibrium field theories.** It is not surprising that some of these methods have not been considered in the critical dynamics applications, which are typically less computationally demanding than applications in high-energy physics. However I am interested in combinatoric interpretations of the coherent state path integral formalism. The combinatorial ideas began already in Chapter 1 where I defined the mapping to a path integral in terms of Heisenberg-Weyl semigroups acting on formal power series. In Chapter 2 the combinatorics of graphs and graph polynomials continue this story.

²² Although this is not the approach taken in this thesis, a simple algorithm to construct a valid flow for arbitrary composite graphs could exploit the 2-trees of the graph: Choose any 2-tree of graph G for which there is non-zero momentum flow across the cutset. Each such pair contains at least one external vertex of G . Choose tree T_1 to be a *source tree* with open boundaries on the cutset and rooted at external vertex v_a and the tree T_2 to be a *sink tree* with closed boundary on the cutset and rooted at external vertex v_b . The sink tree is a proper *arborescence* with all momenta flowing towards v_a . The source tree has all momenta flowing away from v_b . This algorithm avoids creating cyclic flows and for every circuit it is possible to choose a vertex pair that are source and sink duals on that circuit, where momentum always flows from source to sink. The condition for non-zero momentum on the cutset has been encountered in the discussion on graph polynomials (and implicitly in the proof of the all minors matrix theorem) - In this case it ensures that all trees contain at least one external vertex. To properly use this in theories that have creation and annihilation fields for example it would be necessary to redefine v_∞ and consider a v_{source} and a v_{sink} which serves to preserve the edge(field) type at all levels of the graph. In some cases the external field data may not be required in the problem in which case the 2-tree algorithm described will work for the definition of v_∞ already given in Chapter 2.

²³ Many modern textbooks focus on momentum representations in outlining **Feynman rules** although the parametric representations appear increasingly in modern research in attempts to find mathematical structure or for *symbolic computation* (can be integrated on a computer) [Brown, 2009, 2015].

²⁴for example this has been used in defining Hopf algebras in field theories such as QED

²⁵**Isometry and edge pairing combinatorics:** The distinction between the definitions of the $\mathfrak{R}_f$ and $\mathfrak{R}$ is important. Recall isometry is defined at the level of the adjacency structure with respect to vertices. When enumerating possible (coloured) half-edge pairings between two graphs, the vertex that the edge is incident to matters. However half-edges of the same colour on the same vertex are identical when pairing edges as any choice yields the same adjacency structure.

²⁶**The Leibniz rule and the binomial theorem:** Cartier [2000] presents interesting observations following the correspondence between Leibniz's formula and the binomial theorem in the context of symbolic/operator methods.

²⁷**The idea of renormalization in physics:** Green [1835] observes that due to interaction with a fluid medium, an object still obeys physical laws e.g. Newton's laws such as $F = ma$. However there is effective correction to the mass so that rather than using a so called "bare" mass m_b in the expression, one considers $m = m_0 + \delta m$. This is an early observation for the need to "renormalize" model parameters.

²⁸**Regularization methods:** A number of methods to regularize Feynman integrals are discussed in textbooks on quantum field theory, for example [Kleinert and Schulte-Frohlinde, 2001] (§8), but dimensional regularization is favoured by many authors. It has numerous advantages, chiefly; the preservation of symmetries in the problems, easy identification of divergences and it plays well with the minimal subtraction scheme. It also regularizes IR (as well as UV) divergences in massless theories. Kleinert and

Schulte-Frohlinde [2001] (§8.2.2) also provide an overview of *analytic regularization* which follows from the Schwinger representation of the propagator.

²⁹**Phase transitions** describe transitions between structural states of matter due to some external conditions or control parameter such as changing temperature or pressure. A simple example is a transition between liquid and gaseous states. They are classed as either first order or second order (continuous), the latter being the ones we care about. These second order phase transitions are characterized by divergent *susceptibility* (how it responds structurally to an applied external field), infinite *correlation length*, and a power law decay of correlations near the critical point.

³⁰**Intuition for upper critical dimension:** In the case of reaction diffusion processes, criticality is understood in terms of changing particle numbers due to spontaneous events or reactions. Reactions assume particles can meet in geometric space. If the dimensions of space are particularly high, particles can diffuse around and never actually cross paths. The dimension approached from below where this first occurs is called the *upper critical dimension*. This relates to the discussion in Chapter 1 about interaction probabilities for particles in super/Brownian motion.

³¹Actually Lee [1994] writes this more generally as the coupling dimension being $2 - (k - 1)d$ for general reaction $kA \rightarrow lA$ and $d_c = 2/(k - 1)$ which coincides with Kang and Redner [Kang, 1984].

³²**History of Wilson's renormalization group:** A brief history of the ideas leading up to Wilson's breakthrough can be found for example in chapter 1 of Vasil'ev [2004].

³³**Note on the Landau mean-field theory:** Landau [1937] developed a unified theory of critical phenomena to explain *universality* as observed in complex multi-body systems. This is the observation that the system behaviour is determined not by the microscopic details of interaction but by general properties such as dimension of space, symmetries, and the nature of the interaction be it long or short range. However Landau's theory does not completely capture the critical properties of complex systems - for example it fails to properly predict critical exponents of the 2D Ising model. The major breakthrough building on the early work of Landau was the observation by Wilson in the 1970s that, even for strong interactions, some specific critical traits were obtainable via renormalization group methods.

³⁴**Renormalization group ambiguity:** In a sense, there is an ambiguity if the renormalization group, which provides not a single solution but a family of solutions in some scale parameter. A nice discussion of this ambiguity and a "low tech" introduction to the renormalization group is given by Stevenson [1981].

³⁵In Bordeu et al. [2019] we take the route of making a topological argument which adds additional constraints. This is equivalent to algorithmically generating all *viable* diagrams using the methods in Chapter 2 and restricting the viable set using the inequality (3.12) from dimensional analysis.

³⁶**Self-organised criticality** describes dynamical systems that evolve to critical points without the need to tune a control parameter as in other critical systems. In the case of the Abelian sandpile model, depositing sandgrains, the system moves to a critical point until it becomes unstable and leads to avalanches - this cycle repeats.

³⁷**On number of spanning trees for regular lattices:** In particular note the number of spanning trees in an $L \times L$ grid for $L = 1..5$ i.e. 1, 4, 192, 100352, 557568000 - See online Sloane [1999]

³⁸**On the d dependence of the critical point in parameteric space:** That the critical point of the (spatial) branching random walk is determined by the critical point of the underlying branching process (non-spatial) should be treated as a special case in reaction-diffusion processes. Criticality can be thought of in terms of the production of "mass"; loosely speaking, if the number of particles neither grows nor shrinks exponentially but is produced at a stable rate that is "just right" then the process is critical. In this superprocess (BRW) there are no binary reactions and the mass growth and decay depend only on individual particles. Hence it does not matter if all particles are on a single site or dispersed on the graph. In reaction-diffusion processes that contain binary reactions such as the *Gribov process*, the critical point may depend on the dimensionality of space and perhaps the geometry of the graph.

³⁹**van Wijland's hard core particles:** modification of the free part of the action to add tempering delta distributions that enforce site restrictions.

⁴⁰ A description, advantages and disadvantages of **light-sheet fluorescence microscopy** are discussed in [Berthet and Maizel, 2016]. High-speed optical sectioning and adaptability in experimental set-up and arrangement with respect to the specimen to provide a more physiologically accurate environment and increase the viability of the plant. They describe the disadvantages, primarily light scattering, decreasing lateral resolution and contrast leading to the requirement of careful calibration per specimen per setup. Because the sample is illuminated from the side, any light-absorbing/scattering-structure in the illumination paths can generate stripes and/or degrade image quality generating shadows and strips in the image. Fluorescence emission from such shadowed regions is consequently weaker, which has to be taken in account during image analysis, such as in the case of quantification based on fluorescence intensity. Maintaining physiological conditions, while also preventing the sample from moving out of the field of view during the entire time period of the recording, is an experimental challenge.

⁴¹ The Fourier transformed form of the *point spread function* is called the *transfer function*. This is a useful object in the sense that the convolution in the spatial domain becomes the more efficient multiplication in the frequency domain. For each wave number, it gives the (complex) factor by which a periodic structure is multiplied (amplitude and phase) using the filter operation. Image filters are normally designed in the spatial domain because unlike in 1D signal processing, context is lost after the transformation. For very high dimensions, the *Discrete Fast Fourier Transform* can sometimes be faster than convolution. It may be desirable even in low dimensions to work in the frequency domain - for example threshold operations might be more intuitive in some cases.

⁴² **Computer vision** as a field has been invigorated by developments in deep (convolutional) neural networks, which have found their way into biological applications for example the U-Net architecture [Ronneberger et al., 2015]. I do not review these methods here.

⁴³ Two prominent examples of state-space models are those based on *Kalman filters* and those based on *particle filters*. The latter are more flexible being less prescriptive about the structure of distributional priors while the Kalman filter is optimal under suitable conditions.

⁴⁴ Furthermore, it may be that given the choice of normalisation point in these theories i.e. zero external momenta and taking masses to zero at the critical point, the many diagrams that are generated by an S-species system turn out to be projected onto graph isometrics from the perspective of the field theory. These then require that one enumerate different colourings of L loops graphs which have a bounded number of vertices/propagators that need to be considered.

A Algorithms

Algorithm 1 BFT Tree to Recurrent Configuration

```

/*init: tier is  $v_\infty$  of Dhar graph size  $N^*$ /
1: procedure TREETRAVERSE(tier)
2:   Depths  $\leftarrow list[N]$ 
3:   LatticeData  $\leftarrow list[N]$ 
4:   prevTier  $\leftarrow tier$ 
5:   depth  $\leftarrow 0$ 
6:   while ( $w \leftarrow Length(tier) > 0$ ) do
7:     for  $i \in \{1, \dots, w\}$  do
8:       Depth[ $i$ ]  $\leftarrow depth$ 
9:       LatticeData[ $i$ ]  $\leftarrow COMPUTEZ(i, depth)$ 
10:    depth  $\leftarrow depth + 1$ 
11:    prevTier  $\leftarrow tier$ 
12:    tier  $\leftarrow ChildrenOf(tier)$ 
13:
14: procedure COMPUTEZ(i,depth)
15:   candidateParents  $\leftarrow 0$ 
16:   e  $\leftarrow 0$   $\triangleright$  e is an edge
17:   q  $\leftarrow CoordinationNumberOf(i)$ 
18:   parentIndex  $\leftarrow 0$ 
19:   burntInPast  $\leftarrow 0$ 
20:   while e  $< q$  do
21:     ngh  $\leftarrow GetNgh(i, \mathbf{e})$ 
22:     if Depth[ngh]  $== (depth - 1)$  then
23:       if TreeAt(s)  $== \mathbf{e}$  then
24:         parentIndex  $\leftarrow candidateParents$ 
25:         candidateParents  $\leftarrow candidateParents + 1$ 
26:       if LatticeDataAt(ngh)  $\neq 0$  then
27:         burntInPast  $\leftarrow burntInPast + 1$ 
28:       e  $\leftarrow \mathbf{e} + 1$ 
29:
30:   Burnt  $\leftarrow (q - burntInPast - candidateParents)$ 
31:   z  $\leftarrow Burnt + parentIndex$ 
32: return z

```

Algorithm 2 Wilson's Algorithm

```

1: procedure RANDOMTREEWITHROOT
2:   for  $i \leftarrow 1$  to  $n$  do
3:      $InTree[i] \leftarrow false$ 
4:    $Next[r] \leftarrow nil$ 
5:    $InTree[r] \leftarrow true$ 
6:   for  $i \leftarrow 1$  to  $n$  do
7:      $u \leftarrow i$ 
8:     while not  $InTree[u]$  do
9:        $Next[u] \leftarrow RandomSuccessor(u)$ 
10:       $u \leftarrow Next[u]$ 
11:     $u \leftarrow i$ 
12:    while not  $InTree[u]$  do
13:       $InTree[u] \leftarrow true$ 
14:       $u \leftarrow Next[u]$ 
  return  $Next$ 

```

Algorithm 3 Transient point cloud tracking algorithm

High-level description of algorithm using 1-indexing

*Standard geometry or data structure functions are *lower camel case**

Other FUNCTIONS are explained in the main text

*Key tensor *sets* in **bold font** shorthand are understood from function calls*

```

1:  $results := \{\}$ 
2: for  $i \leftarrow lag + 1, frameCount$  do
3:    $U := frames(i)$ 
4:    $V := frames(i - lag)$ 
5:    $uv := cartesianPointProduct(U, V)$ 
6:    $tr1 := translationsFor(uv)$  ▷ For points, k=1
7:    $ranked := RANK(tr1, U, V, epsilon)$ 
8:    $linkages := APPLY(ranked, U, V, epsilon)$ 
9:    $kuv := CONSTELLATIONS(U, V, linkages, k)$  ▷ constellation pairs
10:   $trk := transformsFor(kuv)$ 
11:   $other := PRIORS(linkages, U, V)$ 
12:   $transforms := concatenate(tr1, trk, other)$ 

13:   $ranked := RANK(transforms, U, V, epsilon)$ 
14:   $result := APPLY(ranked, U, V, epsilon)$ 
15:   $results := concatenate(results, result)$ 

16:  $results \leftarrow MINFRAMEDETECTIONFILTERING(result, minFrames)$ 
17:  $results \leftarrow ROIBELIEFFILTERING(result)$ 
18:  $results \leftarrow ANISOTROPICCATCHMENTFILTERING(result)$ 

```

B Useful Results

B.1 The Gaussian integral

The Gaussian integral takes the form

$$\int_{-\infty}^{\infty} dx e^{-x^2} = \sqrt{\pi} \quad (\text{B.1})$$

or in n dimensions

$$\int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \dots \int_{-\infty}^{\infty} dx_n e^{-Ax^2} = \frac{(\pi)^n}{(\prod_i^n a_i)^{1/2}} \quad (\text{B.2})$$

for components in n dimensions. For a matrix representation of A one writes $\sqrt{\det A}$ in the denominator of the result of the multivariate Gaussian. The multivariate Gaussian in d spacetime dimensions appears in this scaled form as

$$\int \dots \int \exp\left(-\frac{1}{2} \mathbf{q}^T \mathbf{A} \mathbf{q}\right) dq_1 \dots dq_d = \frac{(2\pi)^{d/2}}{\sqrt{\det \mathbf{A}}} \quad (\text{B.3})$$

B.2 Gamma Functions

The Gamma function is in fact the Mellin transform of the negative exponential function

$$\Gamma(z) = \int_0^{\infty} dt t^{z-1} e^{-t} \quad (\text{B.4})$$

and arises frequently in the evaluation of Feynman diagrams. The Mellin-Barnes representation exploits formula for propagator forms. Take for example the formula

$$\frac{1}{(a+b)^\lambda} = \frac{1}{\Gamma(\lambda)} \frac{1}{2\pi i} \int_{-i\infty}^{+i\infty} dz \Gamma(\lambda+z) \Gamma(-z) \frac{b^z}{a^{\lambda+z}} \quad (\text{B.5})$$

where a, b take the place of terms such as $m^2, -k^2$ appearing in propagators. The Mellin-Barnes integral is related to various hyper-geometric series and plays a useful role in asymptotic expansions.

Some useful values for the Gamma function are $\Gamma(0) = \infty$, $\Gamma(1) = 1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(n) = (n-1)!$ and the recurrence

$$z\Gamma(z) - \Gamma(z+1) = 0 \quad (\text{B.6})$$

Making use of this recurrence, the volume of the d -dimensional sphere of unit radius can usefully be represented as

$$V_d = \pi^{d/2} / \Gamma(\frac{d}{2} + 1) \quad (\text{B.7a})$$

$$= \frac{2}{d \Gamma(\frac{d}{2}) \pi^{-\frac{d}{2}}} \quad (\text{B.7b})$$

The expansion of the Gamma function is

$$\Gamma(-n + \varepsilon) = \frac{(-1)^n}{n!} \left\{ \varepsilon^{-1} + \frac{\varepsilon}{2} \left[\frac{\pi^2}{3} + \psi^2(n+1) - \psi'(n+1) + \mathcal{O}(\varepsilon^2) \right] \right\} \quad (\text{B.8})$$

where $\psi = \Gamma'(z)/\Gamma(z)$ is the Euler Digamma function.

The Euler-Beta function is

$$B(a, b) = \int_0^1 dz \frac{z^{a-1}}{(z+1)^{a+b}} = \int_0^1 dy y^{a-1} (1-y)^{b-1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad (\text{B.9})$$

B.3 Combinatorial

The binomial theorem:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (\text{B.10})$$

Permutations are written

$$n_{(k)} = \frac{n!}{(n-k)!} \quad (\text{B.11})$$

Stirling numbers of the first kind are written

$$s(k, l) = \left[\begin{matrix} k \\ l \end{matrix} \right] \quad (\text{B.12})$$

and count permutations indexed by the number of their l disjoint cycles. They can be defined algebraically in terms of the permutations by writing

$$x_{(k)} = \sum_{l=0}^k \left[\begin{matrix} k \\ l \end{matrix} \right] x^l \quad (\text{B.13})$$

The Stirling numbers of the second kind are written

$$S(k, l) = \left\{ \begin{matrix} k \\ l \end{matrix} \right\} \quad (\text{B.14})$$

These describe the number of ways to partition a set of k objects into l non-empty subsets. They can be defined

$$x^k = \sum_{l=0}^k \left\{ \begin{matrix} k \\ l \end{matrix} \right\} x_{(l)} \quad (\text{B.15})$$

Normal ordered operation relationship (*proof in appendix of Wiese [2016]*):

$$e^{\lambda \hat{n}} \equiv e^{\lambda z \partial_z} = \mathcal{N}(e^{(e^\lambda - 1)z \partial_z}) \quad (\text{B.16})$$

C Derivations, proofs and further examples

C.1 Homogeneous Poisson process

A Poisson process $\mathcal{P}(\lambda)$ is a memoryless random(point) process. Event waiting times are exponentially distributed $W_t \sim M(\lambda)$.

Introducing a shorthand, I will denote an interval $[t, t+\delta] = [t, t'] = \bar{t}$ and by a severe abuse of notation I will count the number of events in an interval with $X(\bar{t})$ and the number of events *by* a certain time as $X(t)$. A Poisson process is defined by the postulates below;

1. The probability that exactly one event occurs in an interval $P(X(\bar{t}) = 1) = \lambda\delta t + o(\delta t)$
2. $P(X(\bar{t}) > 1) = o(t)$
3. Occurrences of events at time t are independent of occurrences of events at time t' .

For example,* to determine the probability of zero events *by* a certain time, which I denote $p_0(t')$, we can decompose all possible paths in the following way. Consider the complement of the probability of there being *one or more* events in the interval combined with the probability of there being exactly zero events by the start of the interval;

$$p_0(t') = \underbrace{(1 - [P(X(\bar{t}) = 1) \cup P(X(\bar{t}) > 1)])}_{p_0(\bar{t})} \times \underbrace{P(X(t) = 0)}_{p_0(t)} \quad (\text{C.1a})$$

$$= p_0(t)[1 - \lambda\delta t + o(\delta t)] \quad (\text{C.1b})$$

By the Poisson process postulates, the probability of there being exactly one event in an interval is $\lambda\delta t + o(\delta t)$ and the probability of there being more than one is $o(\delta t)$. Now the definition of the derivative tells us that $p_0(t)$ is $e^{-\lambda t}$, that is we subtract p_0 from both sides and the result follows;

$$\dot{p}_0(t) = \lim_{\delta t \rightarrow 0} \frac{p_0(t') - p_0(t)}{\delta t} = -\lambda p_0(t) + p_0(t) \cdot o(\delta t)/\delta t \quad (\text{C.2})$$

Similarly, for $p_1(t')$ we consider two mutually exclusive events; either one event happened by the start of the interval (i.e. time t) and then nothing happened in the interval $\bar{t}$ OR nothing happened by the start of the interval and then one event occurred in the interval

*There is a certain tediousness to going through these basic steps. However it is instructive to take a step back and consider some of the microscopic details with more complex scenarios in mind.

$$p_1(t') = P(X(t) = 1 \cap X(\bar{t}) = 0) + P(X(t) = 0 \cap X(\bar{t}) = 1) \quad (\text{C.3a})$$

$$= p_1(t) \times [1 - \lambda\delta t + o(\delta t)] + p_0(t) \times (\lambda\delta t + o(\delta t)) \quad (\text{C.3b})$$

This can be solved by using the previous result for $p_0(t)$. We have

$$\dot{p}_1(t) = \lambda(p_0(t) - p_1(t)) \quad (\text{C.4})$$

Dividing through by $p_0(t) = e^{-\lambda t}$ and exploiting the inverse of the chain rule,

$$e^{\lambda t} \dot{p}_0(t) + \lambda e^{\lambda t} p_1(t) = \lambda \quad (\text{C.5a})$$

$$\frac{d(e^{\lambda t} p_1(t))}{dt} = \lambda \quad (\text{C.5b})$$

$$p_1(t) = \frac{1}{1!} (\lambda t)^1 e^{-\lambda t} \quad (\text{C.5c})$$

I have added the redundant power and factorial of 1 foreshadowing the fact that one could repeat this process *ad nauseam* to arrive at the definition of the Poisson distribution.

$$\mathcal{P}_n(\lambda) = P(X(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad (\text{C.6})$$

Recalling (C.4), the forward Kolmogorov equation for a Poisson process can be written as

$$\dot{P}(n, t) = -\lambda P(n, t) + \lambda P(n-1, t), \quad n \geq 1 \quad (\text{C.7})$$

On substitution of the generating function representation of the probability,

$$\sum_{n=0} \dot{P}(n, t) z^n = -\lambda \sum_{n=0} P(n, t) z^n + \lambda \sum_{n=1} P(n-1, t) z^n \quad (\text{C.8})$$

Generating function method We would like to write this in terms of operators acting on the definition of the generating function and we are almost there. One way to see this is to align the probability state e.g. $P(n-1)$ with the power in the test variable z_i . In this case "factoring out" z from the second term on the right hand side gives an expression in terms of the definition of the generating function

$$\dot{\Pi} = -\lambda \Pi + z \lambda \Pi = \lambda(z-1) \Pi \quad (\text{C.9})$$

For this infinitesimal generator $\mathcal{Q} = \lambda(z-1)$, the (semigroup) solution recovers the Poisson distribution

$$P(z, t) = P(z, 0) e^{-\lambda(z-1)t} \quad (\text{C.10})$$

C.2 Lattice Laplacian operator

$\nabla^2 = \Delta$ measures the difference between $f(x)$ and the average of the value of f in the neighbourhood. Generally this is defined as $\Delta f = \sum \frac{\partial^2 f}{\partial x_i^2}$. In one dimension, from Taylor's theorem we have the definition of the second derivative

$$\frac{d^2 f}{dx^2} = \lim_{h \rightarrow 0} \frac{1}{h^2} [f(x+h) + f(x-h) - 2f(x)] \quad (\text{C.11})$$

This can be seen by considering expansions for changes $(\pm h)$ in each direction

$$f(x \pm h) = f(x) \pm f'(x)h + f^{(2)}(x)\frac{h^2}{2} \pm f^{(3)}(x)\frac{h^3}{3!} + \dots \quad (\text{C.12})$$

Now adding, alternating terms (odd derivatives) cancel. Rearranging, in the limit $h \rightarrow 0$ we have (C.11). The discrete approximation on a 1D lattice is

$$\Delta f = f(x+1) + f(x-1) - 2f(x) \rightarrow \sum_{\mathbf{e}}^q f(x+\mathbf{e}) - f(x) \quad (\text{C.13})$$

C.3 A note on random walks

The basic definition of the random walk (cf. [Lawler and Limic, 2010]), as a random process in discrete space and time say, is defined on $\mathbb{Z}^d$ with unit vector jumps $\mathbf{e}_k : k \in 1, \dots, d$. The walk of length n is a sum of n random events starting at initial position x_0

$$S_n = x_0 + X_1 + \dots + X_n \quad (\text{C.14})$$

where the transition probability of the discrete Markov chain is

$$\mathbb{P}\{S_{n+1} = x | S_n = x'\} = \frac{1}{2d}, \quad : x - x' \in \{\pm \mathbf{e}_1, \dots, \pm \mathbf{e}_d\} \quad (\text{C.15})$$

A number of interesting properties of random walks are worthy of mention. Basic properties of 1D random walks and solutions for first passage problems such as the *gambler's ruin* can be found in any introductory text on random processes. In d dimensions, the first passage times and probabilities lead to a wealth of interesting problems/solutions as discussed for example in [Redner, 2001]. A good analysis of random walks and intersection probabilities can be found in [Lawler and Limic, 2010]. The important *central limit theorem* (law of large numbers and convergence to normality) appears in considering the behaviour of sums of stable *normal* random variables, such as those describing random walk displacements $\delta(t)$. Incidentally, later in this appendix the *alpha stable distributions* are reviewed - these generalise limit theorems for distributions that are summable in this way but are not normal. The Fokker-Planck equation can be derived from the limit of sums of trajectories of many random walkers (or Brownian particles). I do not reproduce these core results, which are so readily accessible in the excellent texts of Redner [2001] or Lawler and Limic [2010].

Remark 52. A useful generalisation of the random walk model is the *continuous time random walk* model of Montroll and Weiss [1965], which considers arbitrary jump distributions and waiting times between jumps. This occurs in modelling *anomalous diffusion* as considered in Chapter 4.

C.4 Overview of diffusion

Fick's law tells us that the flux density $\mathbf{J}$ is proportional to the concentration gradient, with flow proceeding from high to low concentration

$$\mathbf{J} = -D\nabla c \quad (\text{C.16})$$

The continuity equation tells us that flux is conserved

$$\frac{\partial c}{\partial t} + \nabla \mathbf{J} = 0 \quad (\text{C.17})$$

Combining these we arrive at the diffusion or *heat* equation.

$$\frac{\partial c}{\partial t} = -\nabla \cdot \mathbf{J} = -\nabla \cdot (-D \nabla c) \quad (\text{C.18a})$$

$$= D \nabla^2 c \quad (\text{C.18b})$$

A trivial one-dimensional solution to the diffusion equation is a Gaussian distribution of the concentration $c(x, t)$. For example a narrow variance where all particles are concentrated around one peak approaches a *delta function* or there may be a much flatter Gaussian where particles have spread out from the origin. For N particles and diffusion constant D ,

$$c(x, t) = \frac{N}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \quad (\text{C.19})$$

Let us take the Fourier transform of the concentration

$$c(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \tilde{c}(k, t) e^{ikx} \quad (\text{C.20})$$

and substituting into the 1-dimensional form of the diffusion equation

$$\int_{-\infty}^{\infty} e^{ikx} \left[\frac{\partial \tilde{c}}{\partial t}(k, t) + D \underbrace{\frac{\partial^2}{\partial x^2}}_{k^2} \tilde{c}(k, t) \right] = 0 \quad (\text{C.21})$$

up to constant prefactor. With spacetime translation invariance one finds

$$\tilde{c}(k, t) = e^{-Dk^2 t} \tilde{c}(k, 0) \quad (\text{C.22})$$

which is the inverse Fourier transform of a Gaussian. Making use of the convolution theorem i.e $\tilde{h}(k) = \sqrt{2\pi} \tilde{f}(k) \tilde{g}(k)$,

$$\tilde{c}(k, t) = \sqrt{2\pi} \tilde{\mathcal{P}}(k, t) \tilde{c}(k, 0) \quad (\text{C.23})$$

where one observes the emergence of the propagator (Green's function) such that

$$\mathcal{P}(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{ikx} e^{-Dk^2 t} \quad (\text{C.24a})$$

$$= \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \quad (\text{C.24b})$$

Remark 53. Some observations are in order:

1. The (inverse) Fourier transform of a Gaussian is also a Gaussian
2. The propagator can be interpreted as the probability density function for the probability that a particle reaches position x at time t , given it started at position x_0 at $t = 0$
3. When there are (possibly moving) domain boundaries, the propagator can depend not just on a delta between positions but can depend separately on x and x_0 or the start and end time.

Suppressing the obvious spatial dependence, Einstein considers a drift current $j = c\mu F$. The term μ can be thought of as the strength of the Stokes friction. With sufficient force $F = \nabla\phi$ in a potential, the gradient will grow to cancel the drift so that

$$j \equiv \mu c \nabla\phi - D\partial_x c = 0 \quad (\text{C.25})$$

This has a simple exponential solution for the concentration. Observing that the steady state should agree with the thermal equilibrium Boltzmann distribution

$$c = A e^{-\frac{\phi}{k_B T}} \quad (\text{C.26})$$

for potential ϕ , this leads to a diffusion constant

$$D = -\mu c \cdot \left(\frac{dc}{d\phi} \right)^{-1} = \mu k_B T \quad (\text{C.27})$$

By the *equipartition theorem* namely $\langle v^2 \rangle = k_B T$ one considers the covariance (auto-correlation or two point correlation) as a delta function

$$\langle x(t)x(s) \rangle = 2D\delta(t-s) \quad (\text{C.28})$$

If Einstein's current (C.25) is substituted into the continuity equation (C.17) then we arrive at a special case of the Fokker-Plank equation

$$\partial_t c = -v\partial_x c + D\partial_{xx} c \quad (\text{C.29})$$

also called the *convection-diffusion equation*. One useful algebraic representation of this is to take the combined Fourier-Laplace transform to arrive at

$$c(k, s) = \frac{1}{s + ivk + dk^2} \quad (\text{C.30})$$

Remark 54. One interesting observation which does not appear in the standard literature on diffusion is that made by Blasiak [2005] (note 5, pp. 13) with respect to the operator calculus. The operator calculus is implicitly used in this thesis in the *generating function method* described in Chapter 1. An exponential of the Laplace operator and a characteristic function basis leads by the methods discussed in [Blasiak, 2005] to the heat equation.

C.5 Hausdorff and spectral dimension

Consider a graph G and a vertex $v \in V(G)$. Denote $p_t(G, v)$ the return probability of a random walk on G back to v .

C.6 Chemical reaction networks out of equilibrium

The systems described in this thesis are formulated as chemical reaction networks. They are studied through the lens of (non-equilibrium) critical phenomena. Terms such as detailed

and complex balance, absorbing states, and phase transitions are briefly reviewed in this section in the context of chemical reaction networks.

Each Poisson event is formulated as an *elementary reaction*



This reaction is considered to be reversible and paired with a reverse reaction to produce a *reaction pathway*



In Chapter 1 we introduced the Gribov process, which additionally includes an elementary reaction where particles decay at a rate ε ,



Here there is no reverse spontaneous creation of particles to complement the decay and we can imagine a situation where the process can enter an (empty) *absorbing state* where there are no particles and nothing further can happen.

It is well known that Markov processes (chains) can enter *stationary states*, which are the states the system relaxes into in the long time limit. Simple processes such as the reaction pathway C.32 above may have reaction rates β, χ such that the reactions balance each other which is called *detailed balance* (i.e. forward and backward rates of a certain process must be equal at equilibrium), a particular case of stationarity.

With this background, we can define non-equilibrium systems as those systems whose elementary reactions do not satisfy detailed balance. If a system enters a regime that does not satisfy detailed balance, it may return to an equilibrium state that does satisfy detailed balance (such systems are called *out of equilibrium systems*) or it may not return (such systems are called *far from equilibrium systems*).

Consider the formation of ant trails in foraging ants. Trails require deposition of pheromone, which is re-reinforced by trail following ants. As long as there is food, a trail will form between food and nest. At first foraging ants will explore the area and no significant trail will be formed. But if sufficient ants discover the same food source and return to the nest then a trail will emerge roughly along the shortest path between food and nest. This trail formation can be understood as a phase transition into an ordered state. Energy must be pumped into the system to preserve the trail.

In more complex reaction networks we can extend the notion of reversibility between state pairs to find reversible *paths* between any two *complex states* (for example the complex $A + A$). This is called *weakly reversible*. Complex balance is the condition that the summed rate at which a complex is produced is equal to the summed rate at which it is destroyed [Horn and Jackson, 1972].

A nice overview of chemical reaction networks and stochastic dynamics can be found in [Baez and Biamonte, 2012].

The Hausdorff dimension In the box counting sense, given $|B_R(G, v)|$, the size of a ball of radius R centred on v , the Hausdorff dimension is given by

$$d_h = \lim_{R \rightarrow \infty} \frac{\ln |B_R(G, v)|}{\ln R} \quad (\text{C.34})$$

Spectral dimension The term *spectral dimension* has been introduced by Alexander and Orbach [1982]. The spectral dimension can be defined in terms of the return probability of the random walker on reasonably large graphs as

$$d_s = -2 \lim_{t \rightarrow \infty} \frac{\ln p_t(G, v)}{\ln t} \quad (\text{C.35})$$

C.7 All matrix tree theorem (proof)

Theorem C.1. (*All minors matrix-tree theorem*) For square matrices $\mathbf{E}_{[IJ]}$,

$$\det(\mathbf{E}_{[IJ]}) = \begin{cases} \pm 1 & \text{for the } |\chi| \text{-tree} = E \setminus I \text{ with } |\chi| \text{ connected components,} \\ & \text{each tree having one vertex in } \chi. \\ 0 & \text{otherwise} \end{cases} \quad (\text{C.36})$$

Proof. Distinguish cases where a subset of the graph denoted by the minor $\mathbf{E}_{[X,Y]}$ contains circuits (or certain paths) or does not contain circuits (or certain paths). Observe that the incidence matrix has ± 1 for all half-edges that pass through a vertex. In cases where there are circuits, the contributions of such terms in the determinant must cancel. Similarly, for a path $\pi(x, x')$, $x, x' \in X$, only the end points, which do not appear in $\mathbf{E}_{[X,Y]}$ can contribute as other half edges passing vertices on the path must cancel in the determinant. For the case where there are no cycles and no paths with *both* endpoints in the restricted set X , but instead paths for which $x \in X$ and $x' \notin X$ then there is one non-zero entry corresponding to the row $e \notin Y$ connecting $x' \in X$ to $x \notin X$ such that

$$\det \mathbf{E}_{[X,Y]} = \pm \det \mathbf{E}_{[X \cup \{x\}, Y \cup \{e\}]} \quad (\text{C.37})$$

where $\mathbf{E} \setminus (Y \cup \{e\})$ is a k -tree with one vertex of $(Y \cup \{x\})$ in each component.

This argument can be applied iteratively until $\mathbf{E}$ becomes scalar and the determinant has rank 1 • □

C.8 Schwinger representation

When deriving the parametric representation for the equilibrium case we began

$$\mathcal{P} = \alpha_1 k_1^2 + \alpha_2 (\hat{k} - k_1)^2 + \alpha_1 m_1 + \alpha_2 m_2 \quad (\text{C.38a})$$

$$= \underbrace{\alpha_1 + \alpha_2}_{\Psi} \left(\underbrace{k_1 - \hat{k} \frac{\alpha_2}{\alpha_1 + \alpha_2}}_k \right)^2 + \underbrace{\hat{k}^2 \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} + \alpha_1 m_1^2 + \alpha_2 m_2^2}_{\Phi/\Psi} \quad (\text{C.38b})$$

We needed to complete the square driven by the need to eliminate the $k\hat{k}$ term in the expansion

$$a_1 k_1^2 + a_2 (\hat{k}^2 - 2\hat{k}k_1 + k_1^2) \quad (\text{C.39})$$

Rearranging

$$(\alpha_1 + \alpha_2) \left[k_1 - \frac{2\alpha_2 \hat{k} k_1}{(\alpha_1 + \alpha_2)} + \underbrace{\frac{\alpha_2 \hat{k}^2}{(\alpha_1 + \alpha_2)}}_{T_1} \right] \quad (\text{C.40})$$

a suitable factor to complete the squares inside the square brackets and eliminate the crossterm is

$$\left(k_1 - \underbrace{\frac{\alpha_2 \hat{k}}{(\alpha_1 + \alpha_2)}}_{T_2} \right)^2 - \underbrace{\frac{\alpha_2^2 \hat{k}^2}{(\alpha_1 + \alpha_2)}}_{T_3} \quad (\text{C.41})$$

where $T_2^2 - T_3$ recovers T_1 .

I move now to the the non-equilibrium case. These involve adding the coefficient $D_e \geq 0$ to the k variables and adding in loop and external frequency variables for each edge.

Below, identifying the quasimass $\mathcal{M}$,

$$\alpha_1 D_1 k_1^2 + \alpha_2 D_2 (\hat{k} - k_1)^2 + \underbrace{\alpha_1 (m_1 - i\omega_1 - i\hat{\omega}_1) + \alpha_2 (m_2 + i\omega_2 + i\hat{\omega}_2)}_{\mathcal{M}} \quad (\text{C.42})$$

One interpretation here is then that we simply add the frequency variables to the mass coefficients in φ and we scale the alpha parameters in Ψ by possibly vanishing D .

Each term in momentum k appears with D_e on each edge α_e . Thus we have

$$\mathcal{P} = \underbrace{D_1 \alpha_1 + D_2 \alpha_2}_{\Psi'} \left(\underbrace{k_1 - \hat{k} \frac{D_2 \alpha_2}{D_1 \alpha_1 + D_2 \alpha_2}}_k \right)^2 + \underbrace{\hat{k}^2 \frac{D_1 \alpha_1 D_2 \alpha_2}{D_1 \alpha_1 + D_2 \alpha_2} + \mathcal{M}}_{\Phi'/\Psi'} \quad (\text{C.43a})$$

$$= \underbrace{(D_1 \alpha_1 + D_2 \alpha_2)}_{\Psi'} k^2 + \underbrace{\hat{k}^2 \frac{D_1 \alpha_1 D_2 \alpha_2}{D_1 \alpha_1 + D_2 \alpha_2} + \mathcal{M}}_{\Phi'/\Psi'} \quad (\text{C.43b})$$

In certain cases if we can argue for 0 external k momentum, we sometimes use

$$\mathcal{P} = \underbrace{(D_1\alpha_1 + D_2\alpha_2)}_{\Psi'} k^2 + \underbrace{\mathcal{M}}_{\Phi'/\Psi'} \quad (\text{C.44})$$

Writing the full integral

$$\int_k \int_\omega \int_\alpha e^{-\left((D_1\alpha_1 + D_2\alpha_2)k^2 + \alpha_1(m_1 - i\omega_1 - i\hat{\omega}_1) + \alpha_2(m_2 + i\omega_2 + i\hat{\omega}_2)\right)} \quad (\text{C.45})$$

Performing the Gaussian integral one finds

$$\frac{\pi^2}{(2\pi)^d} \int_\omega \int_\alpha \frac{1}{(D_1\alpha_1 + D_2\alpha_2)^2} e^{-[\alpha_1(m_1 - i\omega_1 + i\hat{\omega}) + \alpha_2(m_2 + i\omega_2 + i\hat{\omega})]} \quad (\text{C.46})$$

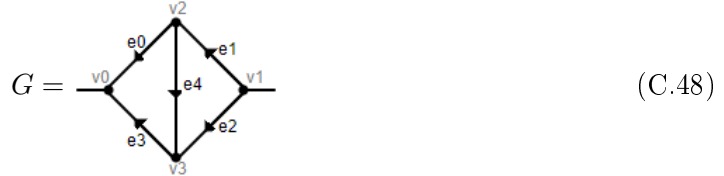
The omega integration produces a delta function in the alpha parameters $\delta(\alpha_1 - \alpha_2)$ and the choice of hyperplane means we can set one of the alpha parameters equal to 1 resulting in.

$$= \frac{\pi^2}{(2\pi)^d} \int_\alpha \frac{1}{(D_1\alpha_1 + D_2\alpha_2)^2} e^{-[\alpha_1(m_1 - i\omega_1 + i\hat{\omega}) + \alpha_2(m_2 + i\omega_2 + i\hat{\omega})]} \cdot \delta(\alpha_1 - \alpha_2) \quad (\text{C.47a})$$

$$= \frac{\pi^2}{(2\pi)^d} \int_\alpha \frac{1}{(\alpha_2[D_1 + D_2])^2} e^{-[\alpha_2(m_1 + m_2 + 2i\omega)]} \quad (\text{C.47b})$$

C.9 Some graph matrices

Consider the graph



For internal edges, I restrict the incident matrix to

$$\mathbf{E}' = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix} \quad (\text{C.49})$$

The diagonal matrix of alpha symbols is simply

$$\mathbf{D} = \begin{bmatrix} \alpha_0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_1 & 0 & 0 & 0 \\ 0 & 0 & \alpha_2 & 0 & 0 \\ 0 & 0 & 0 & \alpha_3 & 0 \\ 0 & 0 & 0 & 0 & \alpha_4 \end{bmatrix} \quad (\text{C.50})$$

The Laplacian matrix is constructed by the minor $\mathbf{E} = \mathbf{E}'_{[0;]}$ (deleting vertex/column 0) and is

$$\mathbf{L} = \mathbf{E}\mathbf{D}\mathbf{E}^T = \begin{bmatrix} \alpha_1 + \alpha_2 & -\alpha_1 & -\alpha_2 \\ -\alpha_1 & \alpha_0 + \alpha_1 + \alpha_4 & -\alpha_4 \\ -\alpha_2 & -\alpha_4 & \alpha_2 + \alpha_3 + \alpha_4 \end{bmatrix} \quad (\text{C.51})$$

The dual matrix which is useful for constructing edge complements when mapping between Kirchhoff and Symanzik type polynomials is

$$\tilde{\mathbf{L}} = \mathbf{E}\mathbf{D}^{-1}\mathbf{E}^T = \begin{bmatrix} \frac{1}{\alpha_2} + \frac{1}{\alpha_1} & -\frac{1}{\alpha_1} & -\frac{1}{\alpha_2} \\ -\frac{1}{\alpha_1} & \frac{1}{\alpha_4} + \frac{1}{\alpha_1} + \frac{1}{\alpha_0} & -\frac{1}{\alpha_4} \\ -\frac{1}{\alpha_2} & -\frac{1}{\alpha_4} & \frac{1}{\alpha_4} + \frac{1}{\alpha_3} + \frac{1}{\alpha_2} \end{bmatrix} \quad (\text{C.52})$$

Note the first Symanzik polynomial is given by

$$\Psi = \det \mathbf{D} \det \tilde{\mathbf{L}} = \alpha_0 \alpha_1 + \alpha_0 \alpha_2 + \alpha_0 \alpha_4 + \alpha_1 \alpha_3 + \alpha_1 \alpha_4 + \alpha_2 \alpha_3 + \alpha_2 \alpha_4 + \alpha_3 \alpha_4$$

Remark 55. Recall this polynomial is homogenous in the alpha parameters of degree L , the first Betti number, and describes the cutset for each of the spanning 1-trees in the graph.

C.10 On the Umbral calculus

The *umbral calculus*⁴⁵ is discussed briefly as an alternative *semantics* for expressing canonical quantization based on operators that obey the Heisenberg-Weyl commutation relation. In the context of field theory, this calculus can be used to discretize continuum processes and to construct representations of Lie algebras on lattices. Operating at the level of formal power series it does not consider convergence properties, as may be the case for defining processes on function spaces such as Hilbert spaces, yet can be mapped to such formulations. These series are often studied under the heading of *Appell* or *Sheffer sequences*. Of particular interest are those sequences of so-called *binomial type* that satisfy a semigroup-type property

$$p_n(s+t) = \sum_{k=0}^n \binom{n}{k} p_k(s) p_{n-k}(t), \quad p_0(s) = 1, \quad p_n(0) = 0 \quad \forall n \geq 1 \quad (\text{C.53})$$

Remark 56. Sequences of binomial type include those sequences that appear so often in physics such as *Hermite* and *Laguerre polynomials*. In the context of this thesis it is worth noting a study by Ohkubo [2012] that considers modifications to the formalism of Doi [1976a] to consider different bases in the "generating function method". He considers some of these series of binomial type (although not using the language of the umbral calculus).

Remark 57. On the relevance of the umbral calculus and sequences of binomial type to the study of random/stochastic processes see also [Rota and Wallstrom, 1997, Rota and Shen, 2000]

Following Roman [2005], the umbral algebra $\mathcal{L}$ is an algebra of linear operators acting on the algebra $\mathcal{F}$ of formal power series in a test variable or acting on real or complex polynomials $\mathcal{P}$ corresponding to a field $\mathbb{F}$ of characteristic zero e.g. on $\mathbb{R}$ or $\mathbb{C}$. From here we will make statements that apply to these various spaces using $\mathcal{F}$. Let $\mathcal{P}^*$ be the vector space of all linear functionals L on $\mathcal{P}$. For a linear functional L , we write

$$f_L(t) = \sum \frac{\langle L|x^k \rangle}{k!} t^k \quad (\text{C.54})$$

and we have that the map $L \rightarrow f_L(t)$ is an isomorphism from $\mathcal{P}^* \rightarrow \mathcal{F}$. Following Roman, we will think of $\mathcal{F}$ as denoting both spaces. The shift operator T and coordinate operator X act on $\mathcal{F}$. $\mathcal{F}[\partial]$ can be seen also as an operator where we write say;

$$\mathcal{F}[\partial] = \sum a_k \frac{\partial^k}{k!} \quad (\text{C.55})$$

C.11 Factorial moment generating functions

An alternative proof showing correspondence between factorial moment generating functions and ordinary ones appears in [Behr et al., 2017]. This is illuminating in that the proof goes via the Stirling numbers. Using the summation convention from Chapter 1, the generator of the exponential generating function is

$$\mathcal{Q}_{em} = (e^{z(l-k)} - 1) \left[\begin{matrix} i \\ l \end{matrix} \right] \partial_z^l \quad (\text{C.56})$$

Using identities for Stirling numbers and the generator of the exponential moment generating function, we can determine the form of the factorial moment generating function generator $\mathcal{Q}_{fm}$ and semigroup (time evolution).

We have met the Stirling numbers. They have the following orthogonality relationship;

$$\sum_{n=0}^k \left[\begin{matrix} k \\ n \end{matrix} \right] \left\{ \begin{matrix} n \\ l \end{matrix} \right\} = \delta_{k,l} \quad (\text{C.57})$$

For any sequences a_n, b_n related by

$$b_n = \left\{ \begin{matrix} n \\ k \end{matrix} \right\} a_k, \quad a_k = \left[\begin{matrix} k \\ n \end{matrix} \right] b_n \quad (\text{C.58})$$

as is this case for the number operators and factorial moments, their exponential generating functions $\mathcal{A}(x), \mathcal{B}(x)$ relationship is as follows

$$\mathcal{B}(x) = \mathcal{A}(e^x - 1), \quad \mathcal{A}(x) = \mathcal{B}(\ln(x + 1)) \quad (\text{C.59})$$

which implies that $\mathcal{M}(t; z) = \mathcal{F}(t; e^\lambda - 1)$ and their time evolution is related similarly having appropriately transformed generators. The mapping $z \rightarrow e^\lambda - 1$ is somewhat reminiscent of the normal ordering relationship (B.16).

Now recalling $\mathcal{Q}_{em}$, we consider a change of variable $\zeta := e^z - 1 \implies z = \ln(\zeta + 1)$. This implies in turn that

$$\partial_z^n = \left\{ \begin{matrix} n \\ \underline{k} \end{matrix} \right\} (\zeta + 1)^k \partial_\zeta^k \quad (\text{C.60})$$

This follows from judicious use of change of variables;

$$\zeta(z) = e^z - 1 \implies \partial_z = (\partial_z \zeta) \partial_\zeta = \underbrace{(\zeta + 1)}_{e^z} \partial_\zeta \quad (\text{C.61})$$

Terms appear in a product n times. Now $u = \zeta + 1 \implies \zeta = u - 1$ and $\partial_u = \partial_\zeta$. So $\partial_z^n = \hat{n}^n$ can be mapped by the Stirling formula using $u \rightarrow \zeta + 1$ while keeping the ∂_ζ invariant.

Now making the substitution (recall the definition of $\mathcal{Q}_{em}$ and the z value at which it is evaluated below),

$$\mathcal{Q}_{ef} = \mathcal{Q}_{em}|_{z \rightarrow \zeta := e^z - 1} = (\zeta + 1)^{l-k} - 1 \left[\begin{matrix} k \\ j \end{matrix} \right] \partial_z^j \quad (\text{C.62a})$$

$$= (\zeta + 1)^{l-k} - 1 \left[\begin{matrix} k \\ j \end{matrix} \right] \left\{ \begin{matrix} j \\ l \end{matrix} \right\} (\zeta + 1)^l \partial_\zeta^l \quad (\text{C.62b})$$

$$= (\zeta + 1)^{l-k} - 1 (\zeta + 1)^k \partial_\zeta^k \quad (\text{C.62c})$$

$$= ((\zeta + 1)^l - (\zeta + 1)^k) \partial_\zeta^k \quad (\text{C.62d})$$

In the third line the orthogonality relationship for the Stirling numbers is used. The last line simplifies to an expression to give the generator for the factorial moment generating function.

C.12 The Lévy α -stable spectrum

Stable distributions (see for example Uchaikin and Zolotarev [2011], Uchaikin [2003], Chechkin et al. [2008], Nolan [2017]) are of interest due to the manner in which they generalise beyond linear diffusion (Brownian motion) and how they can be used to model anomalously diffusive behaviour (Burnecki and Weron [2010]). *Stability* refers to the *closure* property, which is to say, they form a *basin of attraction* under sums of random variables. This leads to a generalised central limit theorem akin to the Gaussian one. The probability density decays asymptotically in power-law form and has a diverging variance that can be helpful to characterise naturally occurring, often fractal, phenomena. The generalised central limit theorem guarantees convergence to a Lévy stable distribution despite this diverging variance. In general the Lévy probability density function may be expressed in terms of it's (four parameter) characteristic function $p_{\alpha,\beta}(k; \delta, \gamma)$. In terms of the Fourier transform we can write it as

$$\exp \left[i\delta k - \gamma^\alpha |k|^\alpha \left(1 - i\beta \frac{k}{|k|} \phi(k, \alpha) \right) \right] \quad (\text{C.63})$$

This is defined for the closed interval, $\alpha \in [0, 2]$. In the case $\alpha = 1$ we have $\phi(k, \alpha) = -2 \ln |k| \pi$ otherwise we have $\phi(k, \alpha) = \tan(\pi \alpha / 2)$. The parameters have the following description; δ and γ are the shift and scale parameters. $\beta \in [-1, 1]$ is the skewness parameter and α is the *stability* index. The distribution is symmetric when $\beta = 0$ and when $\beta \pm 1$, is described as *extremal*. In the literature δ, γ are sometimes written as μ, σ which unfortunately can mislead suggesting they relate to the corresponding Gaussian parameters which they do not (Nolan [2017]). Below are three special cases for which forms in terms of elementary functions can be given;

Gaussian, for $\alpha = 2$

$$p_2(z) = \frac{1}{\sqrt{4\pi}} \exp\left(-\frac{z^2}{4}\right) \quad (\text{C.64})$$

Cauchy, for $\alpha = 1, \beta = 0$

$$p_{1,0}(x) = \frac{1}{\pi(1+x^2)} \quad (\text{C.65})$$

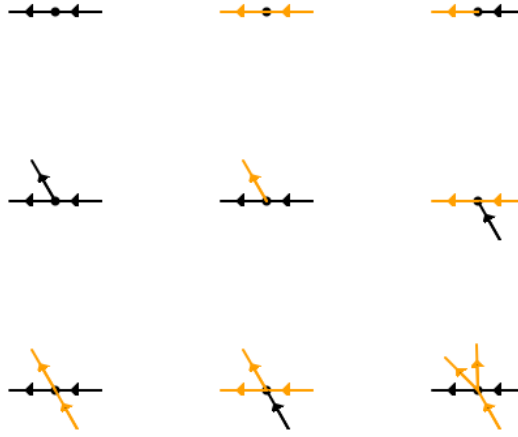
Lévy-Smirnov, for $\alpha = 1/2, \beta = 1$

$$p_{1/2,1}(x) = \frac{1}{\sqrt{2\pi}} x^{-3/2} \exp\left(-\frac{1}{2x}\right), x \geq 0 \quad (\text{C.66})$$

D Diagrams for BWS

Some diagrams for the analysis of the Branching Wiener Sausage are shown in this appendix to illustrate the basic combinatorics involved in enumerating diagrams.

The primitive graphs below come from the action derived from the infinitesimal generators of the semigroup and they are called *corollas* or *amputated vertices*. The colours correspond to the two species of the theory - black is the A particle and yellow is the B immobile particle.



The shuffle product over permutations of primitives generates $\mathcal{G}_l$ to loop order l . In this study $l = 1$. The first two, the propagators from the free part of the action, are not used. From permutations of the remaining 7 primitives it is possible to generate a great many diagrams and various restrictions are put in place, firstly for the maximum loop number and then the number of external legs that are permissible. It is worth pointing out that at tree level, new vertex diagrams are generated. We can contract over internal edges *of the same species* to produce new corollas. For example

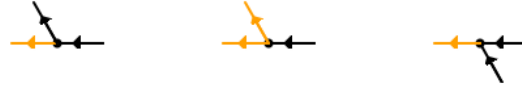
$$\text{Diagram 1} \otimes \text{Diagram 2} = \text{Diagram 3} \quad (\text{D.1})$$

To be clear

$$\text{Diagram 4} \otimes \text{Diagram 5} \neq \text{Diagram 6} \quad (\text{D.2})$$

During the contraction, we identify the vertices separated by an internal edge of the same species. As a counter example, the transmutation vertex will not be contracted because while it is 2-valent, the incoming and outgoing edges are different species and not contracted.

New 3-valent primitives are


(D.3)

Having generated all l -loop graphs, some book keeping is required to organise them by having either the same residue or the same loop topology. Below I show examples of some of the distinct loops that can be generated, although the same loops (ignoring residues) can be generated by other combinations.

Loop construction

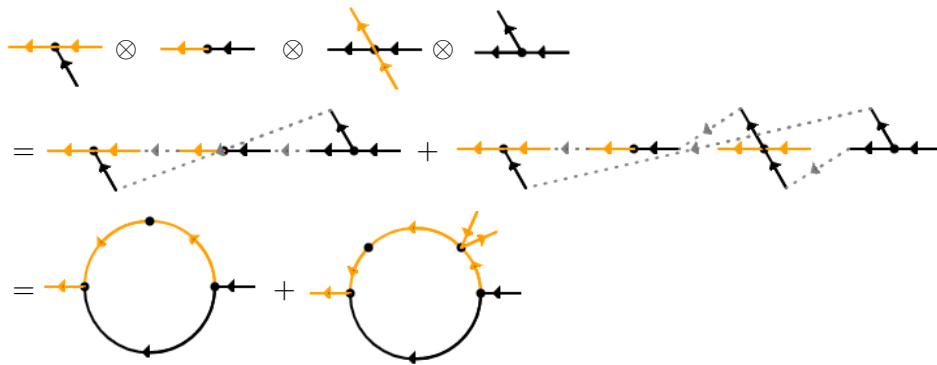
The shuffle product of two graphs results in a set of graphs. Decorated half-edges on the graph residues $\mathfrak{R}(G)$ are merged in all possible ways.



To apply the shuffle product over permutations of graph sets algorithmically, we can *fold left* on the shuffle product to apply the operator over all combinations. Fold-left is functional programming jargon for first applying a binary operation on the right-most pair of items to form a new item, and then apply the binary operation on the next pair, and so on. The underbraces show this binary pairing being applied from right to left. Each binary operation is as described immediately above.

$$\begin{array}{c}
 F_1 \otimes F_2 \otimes \underbrace{F_3 \otimes F_4}_{F'_3} \\
 \underbrace{F'_2 = F_2 \otimes F'_3}_{F'_1 = F_1 \otimes F'_2}
 \end{array}$$

In the particular case where only 1-loop graphs are required, the following example applies the product over 4 primitives

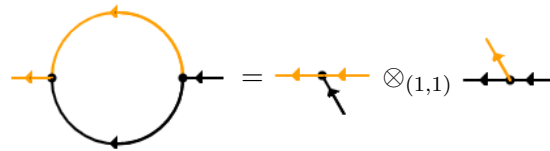


The intermediate representation emphasises which vertices are combined and over which half-edges, while the last representation emphasises the loop structure.

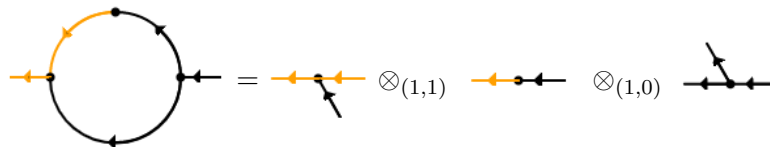
Distinct loops

The distinct loops demonstrated here may be constructed in different ways because we may, in some cases, ignore the residues of the graphs, which may differ depending on what graphs are combined. Each of these loops corresponds to a distinct integral

1

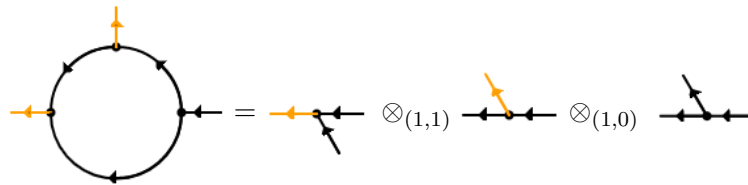


2



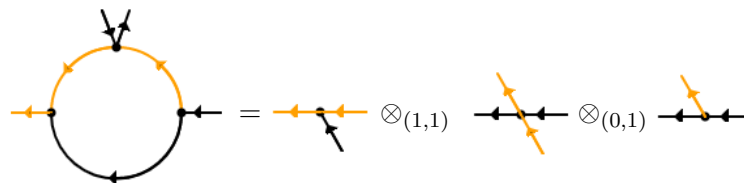
This diagram provides an example of a 2-valent vertex that is not contracted because the species transmute across the vertex.

3

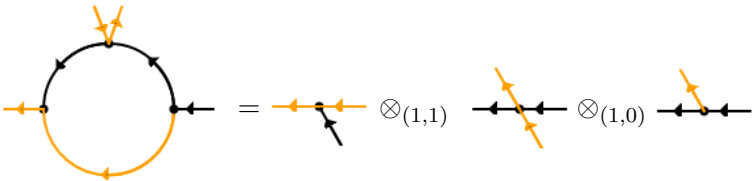


This example demonstrates the use of a new amputated vertex (the first of three) generated from the primitive generators.

4

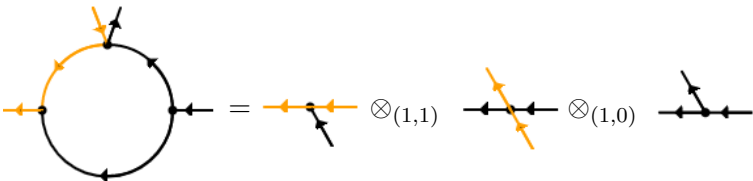


5



Notice that this uses the same primitives as I4 except the product far right uses a different pairing.

6



7

