Non-equilibrium stochastic mechanics: Emergence and renormalization

Wout Merbis*
Dutch Institute for Emergent Phenomena (DIEP)
Institute of Physics (IoP), University of Amsterdam

October 9, 2023

Abstract:

These notes serve to organize my thoughts on the lecture module prepared for the DITP course: *Advanced Topics in Theoretical Physics* with the theme "Emergence and the Renormalization Group". In the lectures we will explore quantum methods for stochastic mechanics, where we will focus on the theme of 'emergence and renormalization'. First, we will provide a highly selective and biased historical and conceptual background on *emergence*. What does this even mean? What are different kinds of emergence? After this, we will focus on bringing techniques familiar to theoretical physicists (such as second quantization, Fock space, path integrals and quantum field theory) back into the classical domain. We will apply these techniques to reaction-diffusion models of non-equilibrium stochastic systems, which may find applications in a diverse set of scientific fields, including physical chemistry, theoretical ecology, epidemiology, game theory and socio-economical models of complex systems. We conclude with discussing renormalization group methods in this setting, which links back to the topic of emergence, as we will see universality of critical dynamics appear near dynamical phase transitions, similar to that of equilibrium systems.

1 Introduction

A course fitting in the theme *'Emergence and the renormalization group'* could potentially span a large number of topics in many different fields of science. *Emergence* is, after all, not really a scientific method or field, but rather a perspective on the sciences summarize by the credo: "the whole is more than the sum of its parts". This could apply to the science of any complex (read: composite) system, where many constituents come together to interact in non-trivial ways, resulting in a separation of scales in the system (the whole vs. the parts). Quantum mechanical interactions of many elementary particles form complex molecules. These are subject to bio-molecular reaction networks, leading to the emergence of cellular life. Neuronal cells, in turn, form intricate networks in the brain, leading to information processing capabilities and the emergence of intelligence and consciousness. And

*w.merbis@uva.nl

these intelligent and conscious entities interact in the economical and political arena to form collaborative works, impossible to achieve by the isolated effort of individuals. The renormalization group, too, has seen many use cases and applications since its inception in elementary particle physics. By now, renormalization is an well-established part of the theoretical physicists toolbox, with applications ranging from the block spin approach to the Ising model, to recent applications of renormalization group methods to neuronal data.

It is therefore very hard to limit oneself to any one subject, since all of this is extremely fascinating and constitute active fields of research. The aim of these lectures, then, is to first give a conceptual, almost philosophical, overview of emergence in general [1,2] and then focus on a concrete example of what theoretical physicists can bring to the table. Namely, we can apply methodology (analytical or numerical) from physics to study systems which are traditionally within other domains of science (chemistry, biology, neuroscience or even sociology). The context we will do so in here is that of non-equilibrium stochastic systems, which have found various applications outside physics, such as in chemical reaction networks, population dynamical models, epidemiology, modeling polarization in political science, and many other settings. We will discuss a general mathematical framework to describe the Markovian dynamics of such systems using a mathematical language familiar from quantum mechanics and quantum field theory. Ultimately, this will allow us to formulate the renormalization group equations for non-equilibrium systems and study the critical behavior of dynamical models in a way similar to the renormalization of quantum field theories and the analysis of equilibrium critical systems.

2 Emergence: some philosophical context

At all different length scales in Nature, entirely new properties appear as the result of the collective behavior of the system at a smaller scale. This leads to the view that there is a hierarchy in the sciences, where at each level entirely new laws, concepts and generalizations are necessary [3]. The hierarchy posits that the elementary building blocks of systems studied by one science form the main object of study for a more 'fundamental' or 'pure' science. It is not uncommon among physicists to have a predisposition that this chain of inter-dependencies ends at elementary particle physics and other sciences are less fundamental. The domains of other (natural) sciences are reducible to a more fundamental level, and ultimately this process repeats until one arrives at the fundamental laws of Nature, as uncovered by physicists. This way of thinking is illustrated in Figure 1, which might be one of the simplest illustrations of *reductionism* possible.

From a philosophical perspective, it is important to distinguish (at least two) different kinds of reductionism, *ontological* (referring to 'being') and *epistemological* (referring to 'knowing'). An ontological reductionist would claim that all sciences are, *in principle* reducible to more fundamental components, i.e. a complex system *is* in essence the sum of their parts. Epistemological reductionism constitutes a slightly stronger statement: it posits that any higher-level behavior of a physical system can ultimately be *explained* by the states, properties and laws of its parts. So not only are systems reducible to its parts, additionally the connection between the whole and the parts are knowable and understandable in terms of maps which translates between them.

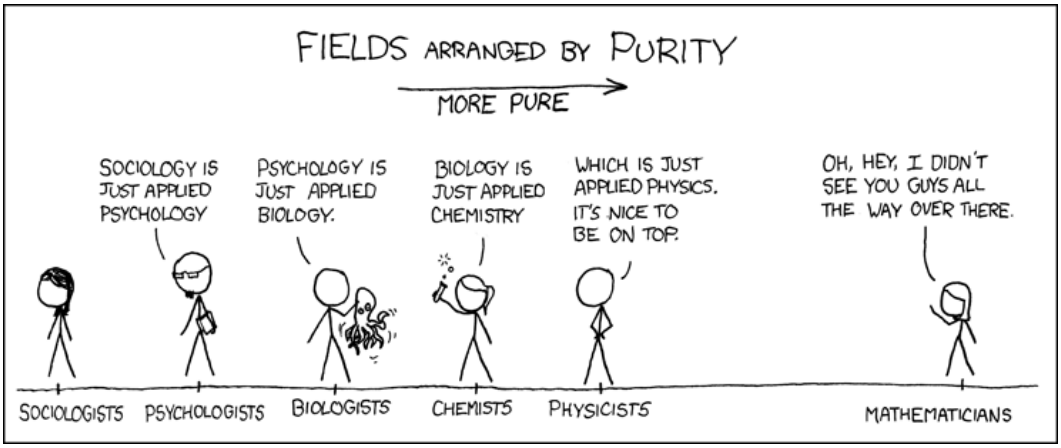


Figure 1: This xkcd comic illustrates how some physicists perceive the sciences.

At first sight, emergence can then be defined as a denial of ontological reductionism, as was done by some of the early British emergentists in the 1920's. For instance, C. D. Broad wrote in 1925:

> "the characteristic properties of the whole $R(A, B, C)$ [where R marks their structural arrangement] cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form $R(A, B, C)$. **C. D. Broad, 1925** [4]

This point of view, which denies ontological reductionism, can be called **ontological emergence** (or: strong emergence) [5]. Here, higher-level properties of a system are not reducible to the properties of the parts and are governed by new laws and new principles, independent of the laws governing the constituents. The system might very well be composed out of the part (as in: supervenient), but its behavior and characteristics are independent of it. This point of view, however, was quickly loosing terrain, partly due to advances in quantum mechanics, which started to unravel and explain chemical properties of substances in terms of the quantum description of atoms. Shortly after Broad, Paul Dirac wrote:

> "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws lead to equations much too complicated to be soluble" **P. Dirac, 1929** [6]

What is clear from this quote is that some form of reductionism is supposed, referring to the underlying physical laws (i.e. those of quantum mechanics), but there might be obstacles in our computational or explanatory ability to uncover the mapping between fundamental laws and higher-level properties of a system. This thus implies a failure of epistemological reductionism and leads to **epistemological emergence**, sometimes referred to as **weak emergence**. It implies that higher level phenomena (like biological cells) are not explainable in terms of elementary particles and forces due to epistemic limitations, such as the lack of computational power or analytical tractability.

Most physicists, including Einstein, Dirac and Schrödinger, held similar views on this topic, combining ontological reductionism with epistemological emergence [2]. Schrödinger held a lifelong interest in biology (and he is sometimes credited as being the first to propose that there exist genetic material as a code for constructing living materials). He wrote in his lectures on the physical basis of life:

> "What I wish to make clear ... is, in short, that from all we have learnt about the structure of living matter, we must be prepared to find it working in a manner that cannot be reduced to the ordinary laws of physics. And that not on the ground that there is any 'new force' or what not, directing the behaviour of the single atoms within a living organism, but because the construction is different from anything we have yet tested in the physical laboratory" **E. Schrödinger, 1944** [7]

This quote seems to lean towards ontological emergence by suggesting that living matter cannot be reduced to ordinary laws of physics. However, like Dirac, Schrödinger was ultimately convinced that the laws of physics could ontologically reduce biology, even if there were epistemological obstacles to explaining how exactly [2].

The above quote of Schrödinger seems to imply that the whole is not necessarily more, but *different* than the sum of the parts. This come back in Philip W. Anderson's important essay "More is Different" from 1972 [3], which reignited the emergence/reductionist debate[1]. In it, he starts by setting out the reductionist point of view on the sciences, after which he quickly points out the fallacy of epistemological reductionism:

> "The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe." **Anderson, 1972** [3]

As is, this does not tell us whether epistemological reductionism fails due ontological or epistemic reasons, i.e. can we not reconstruct the universe because it is too complex to do so, or because it cannot be done in principle? He does, at some point, seem to make a case that research into any of the higher-level domains is not less fundamental than for instance, fundamental particle physics: 'At each stage entirely new laws, concepts, and generalizations are necessary, requiring inspiration and creativity to just as great a degree as in the previous one' [3].

While the ontological nature of Andersons work seems to still lean towards reductionism, an interesting new twist is brought by his analysis of symmetry breaking. In the limit of $N \to \infty$, systems undergo a sharp and singular transition towards states where the microscopic symmetries and equations are violated. These microscopic symmetries leave behind only certain characteristics in the macroscopic state, for example, long wavelength vibrations such as sound waves. As such, higher-level entities will naturally have less symmetries than the microscopic description. This leads to the observation that the whole 'becomes not only more than but very different from the sum of its parts' [3]. One example he brings up for this is

[1] This debate culminated in Anderson's fierce opposition to the Superconducting Super Collider before the US congress, which contributed to the cancellation of the project, to much frustration of leading reductionists such as S. Weinberg.

the parity symmetry breaking of sugars produced by living matter. If the same sugars are synthesized using a chemical reaction, both parities are present in equal abundance. It takes living matter to break the symmetry and only produce sugars of a single parity.

In a much-discussed essay [8], R. Laughlin and D. Pines built on Andersons point of view and call into question the usefulness of a fundamental 'Theory of Everything'. In their opinion, this ultimate theory consists of the Schrödinger equation of conventional nonrelativistic quantum mechanics, which ultimately describes the everyday world of human beings (note the implicit ontological reductionism). However, they quickly note that no amount of computational power would be able to solve this equation even for a modest number of particles, which leads them to conclude that: "we have succeeded in reducing all of ordinary physical behavior to a simple, correct Theory of Everything only to discover that it has revealed exactly nothing about many things of great importance" [8].

While this point of view at first hand looks again like ontological reductionism combined with epistemological emergence, they later make the case that in some situations higher-level organizing principles (or *protectorates*) do exist, opening the door to ontological emergence at least a little bit.

> "The emergent physical phenomena regulated by higher organizing principles have a property, namely their insensitivity to microscopics, that is directly relevant to the broad question of what is knowable in the deepest sense of the term. The low-energy excitation spectrum of a conventional superconductor, for example, is completely generic and is characterized by a handful of parameters that may be determined experimentally but cannot, in general, be computed from first principles. **Laughlin & Pines, 2000** [8]

Other protectorates mentioned are (among others): superfluidity, band insulation, ferromagnetism, antiferromagnetism, and the quantum Hall states. At low energies, excited states of these systems are described by particles which satisfy similar properties of elementary particles (they carry conserved quantum numbers, obey Fermi or Bose statistics and scatter off one another according to simple rules). These excitations are not fundamental. They do not exist outside to context of the stable state of matter in which they live. However, the underlying (high-energy) microscopic theory may have no measurable consequences at low energies. In other words, the nature of the underlying theory cannot be deduced from the low-energy effective action, not until one raises the energy scale sufficiently to escape the protectorate. This connects directly to renormalization and universality. Under the renormalization group (RG), irrelevant operators of the microscopic theory have no effect whatsoever on the macroscopic theory. Furthermore, close to the RG fixed point, universality is observed, which states that the physical behavior of some systems close to criticality is the same, regardless of microscopic details.

2.1 Modern frameworks for understanding emergence

The distinction between strong and weak emergence, as focused on in the historical debate above, is according to some more recent authors neither useful nor complete. There is hence a call for more nuance in the different relations which may exist between high-level macroscopic phenomena and the description of the systems microscopics. For many physicists, science is

inherently reductionist and hence ontological emergence is a hard pill to swallow. After all, if there are phenomena which are not ultimately reducible to the elementary laws of Nature, than new laws are necessary to explain these phenomena. Furthermore, we are lacking a clear and uncontroversial example of a strongly emergent phenomena, although some authors see consciousness as the proof that ontological emergence does exist in Nature [9].

On the other hand, the definition for epistemological (or weak) emergence might be too loose to describe what one would intuitively consider to be emergent. For instance, no one would be able to explain all properties of a chair from the configurations of its atoms, but does that make the comfortableness of the chair an emergent property? Intuitively that would sound silly, as the chair is obviously intended to be comfortable by construction. There is no surprising element here. This tends to the idea that weak emergent properties might be observer and time dependent; it seems to depend on our inability to provide an explanation of *spontaneously occurring* and *surprising* higher-level phenomena in terms of lower-level system.

To move towards a more rigorous understanding of the possible relations which may exist between higher-level and lower-level explanations of a system, it is instructive to draw that following diagram:

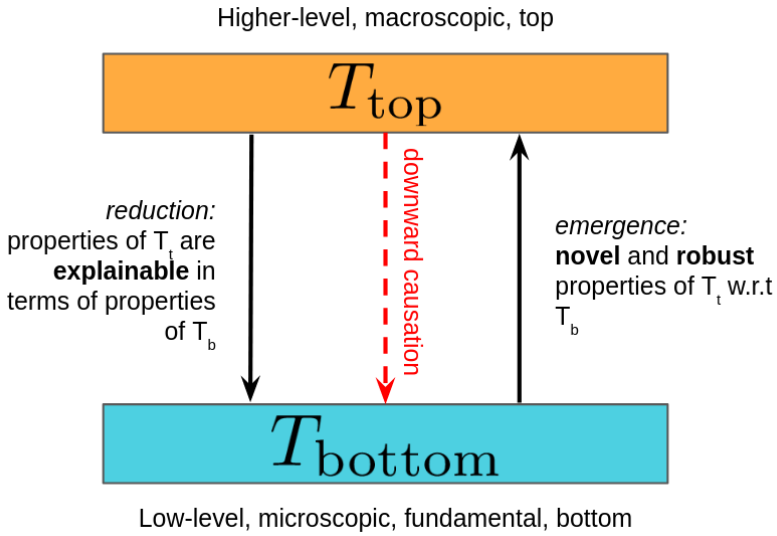


Figure 2: Schematic distinction between top and bottom theories, after [10]

Here, we distinguish two 'theories' about a phenomena in physical reality: a top theory T_{top} describing the high-level phenomena, and a bottom theory T_{bottom}, which concerns itself with the microscopic description of the system. Any mathematical theory always is accompanied with an interpretation which links the theory to a *domain* in the physical world. Emergence is the situation when properties, facts or phenomena of the top theory are **novel** and **robust** with respect to the bottom theory. The novelty here covers the spontaneously occurring and surprising element which we would intuitively ascribe to emergent properties. By additionally requiring robustness, we make sure that small changes in the bottom theory do not effect or influence the emergent properties of the top theory, just as we had seen in the discussion of renormalization and universality above.

The distinction between ontological and epistemological emergence in the picture above then becomes related to the distinction between the theory (which provides an explanation, and is hence epistemic) and the domain of the theory (which relates to physical reality, i.e. ontology). If the two theories describe the same domain, emergence is epistemological, whereas if the top theory and bottom theory deal with different domains, emergence is ontological.

A frequent (and for a physicist quite natural) assumption underlying the discussion so far is that the causal link of explanation goes from microscopic towards macroscopics. I.e. it is often supposed that the smaller system exerts causal influence over the higher-level description of the system, but this is merely an assumption. It might be logically possible that the causal arrow points (partly) the other way, and that the high-level system has some form of causal control over the microscopic configurations of the system. This possibility is called **downward causation** and implies that the low-level laws of T_{bottom} are incomplete as a guide to describe both the low-level and the high-level evolution of processes in the world [9].

Besides downward causation, emergence could also be characterized as a *causal decoupling* between the top theory and the bottom theory. Suppose that the bottom theory is characterized by a set of random variables $\{X_t^i\}$, with $i = 1, \ldots, n$, subject to some unknown time evolution. We can observe the system at a larger scale, defined by a course grained variable V_t, which is composed out of the set $\{X_t^i\}$, for instance as the average value of some observable. The following schematic picture for the systems time evolution applies:

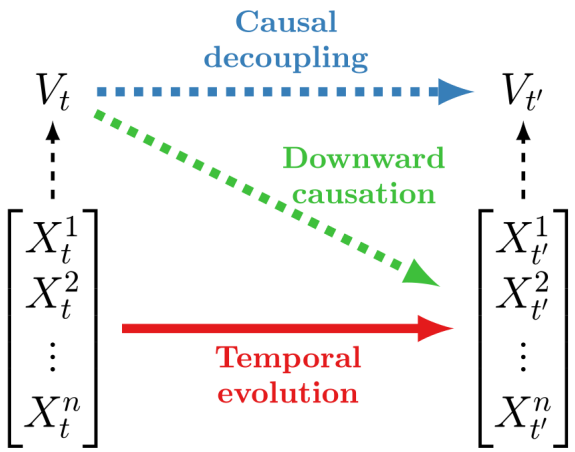


Figure 3: Schematic description of causal decoupling and downward causation in terms of a dynamical system for the top theory V_t and for the bottom theory $\{X_t^i\}$. Picture taken from [11].

Here, we again see the possibility of downward causation, but one could also ask what is a better predictor of the future macroscopic observable $V_{t'}$? It could be that knowing the state V_t is necessary and sufficient to make accurate predictions on $V_{t'}$, without referring to the systems components at all. In that case, when collective properties have irreducible causal power over other collective properties, we can speak of **causal decoupling** [11]. Accordingly, one can define *causal emergence* when a top theory feature V_t has irreducible causal power, i.e. when it exerts causal influence that is not mediated by any of the parts of the system.

The advantage of speaking in terms of causal influence in dynamical systems is that it is possible to make this concrete in terms of information theoretic quantities, coming from

multi-variate partial information decomposition [11]. Without going into too much detail, in information theory it is possible to distinguish between information which is shared between a set of random variables (redundancy) and information which is contained in the collection, but not in any of its parts (synergy). A measure of causal emergence is then formally defined as the synergy minus the redundancy in the system, and this can be used as an ignorant (as in unbiased) statistical measure of emergence, making no assumption on what the actual emergent properties are [11].

What should be clear from the horribly incomplete and incredibly biased discussion presented here, is that even after more than a hundred years of debate on the topic, *emergence* still constitutes an active topic of debate for scientists and for philosophers. Current efforts are being made to ground the concept into a more quantitative framework, which might help to guide the future discussion, but it is clear that the last word has not yet been said in this discussion. This is partly what makes this such an interesting topic to study, as opposing view are still held to this day and fiercely defended.

2.2 Discussion exercises

For the first exercise, as an ice-breaker for the group, we will have a discussion on various topics relating to emergence in the class.

Discussion topics:

1. Are emergence and reductionism mutually exclusive?
2. What are good examples of weak emergence without strong emergence?
3. What is a good example of downward causation (in physics or outside) and why?
4. Does strong emergence (or ontological emergence) exist? If so, what are examples of strong emergence? What would the existence of strongly emergent phenomena imply on the fundamental laws of Nature as uncovered by physicists?
5. Is it more fundamental to be a particle physicist than to be (for instance) a theoretical ecologist?
6. Is a theory of everything (if it even exists) actually useful in describing anything (i.e. higher-level phenomena)?

3 Stochastic mechanics: from quantum mechanics to chemistry and biology

As we saw in the previous section, emergence deals with connection of system at different scales and the relationships that can exist between the whole system and its parts. Often a top level description (whole) may be linked to a lower level description (parts, i.e. more elemental/fundamental). For the condensed matter physicist, this involves (quantum) statistical mechanics, which aims to describe large assemblies of (more or less) elementary constituents

interacting with one another. Similarly, chemical reaction networks often involve large numbers of molecules as reactants and reaction products. In population biology, the dynamics of ecosystems of interacting species ultimately depends on the assembly of all individuals, which interact with each other in a dynamical environment.

These situations show similarities in their mathematical description, which we aim to unravel in the following part of the lectures. Particularly, stochastic models of chemical reaction networks and population dynamics show a similarity with the description of quantum mechanical systems. The microscopic description is probabilistic, but stochastic in the sense of classical probability theory, not quantum mechanical in the sense of amplitudes. We can, however, use an analogy with quantum mechanics to describe situations where the basic entities (molecules or species) change in abundance over time, and use a Fock space representation to describe the probability distributions governing the systems dynamics [12, 13].

In the framework of emergence described in the last section, there are here also (at least) two levels of description for the systems dynamics. At the higher level (top theory), the dynamics of the population (or densities of chemical compounds) is described by the **rate equations**: a set of ordinary differential equations describing the change in densities of different species over time. This description emerges as a large N limit of the **master equation**; a microscopic theory which deals with the probabilistic change in abundances due to the interaction of individual components. The macroscopic rate equations are deterministic, continuous and often non-linear. The microscopic master equation gives a stochastic (as in: probabilistic) description of the system where entities vary discretely. This allows one not only to track average quantities, but also study fluctuations from the mean.

3.1 Chemical reaction networks

A chemical reaction network (CRN), loosely speaking, consists out of:

- A discrete and finite set of *species* S, which contains two (not mutually exclusive) subsets: the reactants $\mathcal{R}$ and the products $\mathcal{P}$
- A set of (real and positive) reaction rates $r \in \mathbb{R}^+$, which can be seen understood as maps from the reactants to the product.

There are several equivalent ways to represent a CRN graphically. The best way to show this is by example. For instance, a simple reaction network consists out of a single reaction for the set species $S = \{\mathrm{H}, \mathrm{OH}, \mathrm{H_2O}\}$:

$$\mathrm{H} + \mathrm{OH} \to \mathrm{H_2O}\,. \tag{1}$$

This reaction can be represented graphically as a *Perti net*, which is a special type of graph (a directed bipartite graph). The vertices are of two kinds: species or transitions. Species are connected with transitions by a directed edge if they are used up in the reaction (i.e. if they are the reactants). Likewise, transitions are connected to species if the reaction creates those species (i.e. if they are the reaction products). In case of the example above, the Petri net is drawn in Figure 4.

Another example from epidemiology is the susceptible-infected-susceptible model. This simple model of infectious decease spreading supposes the population is divided into people

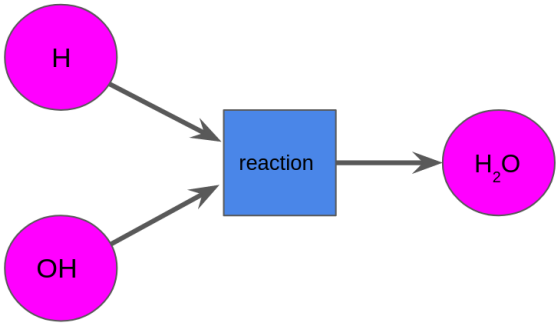


Figure 4: The Petri net for the chemical reaction given in (1).

infected by the decease (I) and the rest is susceptible to an infection (S). There are now two possible transitions, *infection* and *recovery*, which can be represented as chemical reactions:

$$\textit{infection}: \qquad \mathrm{S} + \mathrm{I} \to 2\mathrm{I}\,, \tag{2a}$$

$$\textit{recovery}: \qquad \mathrm{I} \to \mathrm{S}\,. \tag{2b}$$

In terms of a Petri net, this corresponds to the graph of Figure 5.

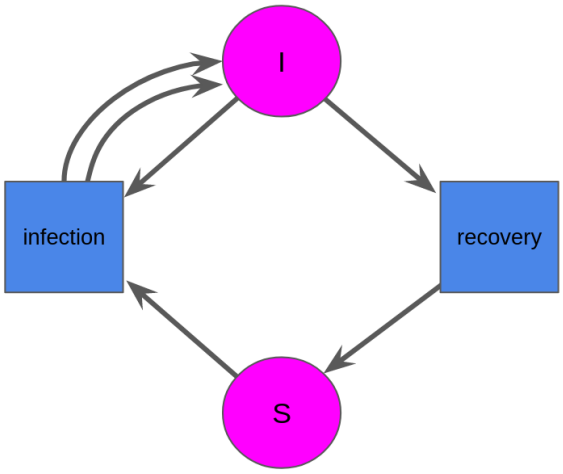


Figure 5: The Petri net for the chemical reaction network defining the SIS model (2).

Notice how here both an infected and a susceptible are needed for the infection reaction, and hence the output are two infected, illustrated by the two edges going from the infection vertex to the I vertex.

Chemists (or population biologists for that matter) often deal with a great number of particles. Hence, they are not always interested in knowing the exact number of particles of each species, but rather they want to know how the expected number of the reactants and products change over time. After all, when the expected numbers are very big, the random fluctuations will be very small.[2] This will result in the *rate equations* for the CRN, which can be derived just from looking at the Perti net in the following way:

[2]This is the law of large numbers

For each species $i \in S$, we'd like to know the rate of change $\frac{dx_i(t)}{dt}$ for the expected number of elements of the species x_i. Each reaction in the network connected to this species will contribute to the rate of change, positively if the species is the reaction product and negatively if it is the reactant. However, the argument is obviously the same for each reaction, so we can focus on a single reaction first. Let's suppose that the i-th species is involved in this reaction (with rate r) m_i times as reactant and n_i times as product. Then, the rate of change of x_i due to this reaction is:

$$\frac{dx_i(t)}{dt} = r(n_i - m_i)x_1^{m_1} \dots x_k^{m_k} \tag{3}$$

Here the factor $(n_i - m_i)$ tracks the amount of species i created (or destroyed) in the reaction. The product of all species $x_1^{m_1} \dots x_k^{m_k}$ appears because the reaction can only occur once all of its inputs are present, and then its rate is proportional to the product of the number of things it takes as an input. Finally, r is the reaction specific rate.

Hence, in our first example, there is only a single reaction, but there are three species so we can derive three equations:

$$\frac{dx_{\mathrm{H}}}{dt} = -r x_{\mathrm{H}} x_{OH}\,, \tag{4a}$$

$$\frac{dx_{\mathrm{OH}}}{dt} = -r x_{\mathrm{H}} x_{OH}\,, \tag{4b}$$

$$\frac{dx_{\mathrm{H_2O}}}{dt} = +r x_{\mathrm{H}} x_{OH}\,. \tag{4c}$$

The generalization to multiple transitions is quite straightforward, although it does becomes a bit notational heavy. Now each transition (label by τ) has its own rate $r(\tau)$, and consumes a number $m_i(\tau)$ and produces a number $n_i(\tau)$ of species i. To obtain the rate of change in the expected number of species i, one simply sums (3) over all transitions τ:

$$\frac{dx_i(t)}{dt} = \sum_{\tau} r(\tau)(n_i(\tau) - m_i(\tau))x_1^{m_1(\tau)} \dots x_k^{m_k(\tau)} \tag{5}$$

For example, for the reaction network defining the SIS model of epidemiology, if we let α be the infection rate and β the recovery rate, we immediately obtain from (5):

$$\frac{dx_{\mathrm{I}}}{dt} = \alpha x_{\mathrm{I}} x_{\mathrm{S}} - \beta x_{\mathrm{I}}\,, \tag{6a}$$

$$\frac{dx_{\mathrm{S}}}{dt} = -\alpha x_{\mathrm{I}} x_{\mathrm{S}} + \beta x_{\mathrm{I}}\,. \tag{6b}$$

To reduce the notational complexity, we will often forgo writing (τ) and let this be clear from context. Furthermore, x_i, m_i and n_i are vectors in $\mathbb{R}^d$, where $d = |S|$. So we will often use the shorthand notation:

$$\mathbf{x}^{\mathbf{m}} \equiv x_1^{m_1} \dots x_k^{m_k}, \tag{7}$$

In other words, raising a vector to a vector power will mean taking element-wise exponents.

The most generic CRN is now abstractly represented by a set of reactants $R_i(\tau) \in S$, with multiplicities $m_i(\tau)$ (or in chemistry: stoichiometric coefficients) and a set of products $P_i(\tau) \in S$ with multiplicities $n_i(\tau)$, which are connected by a set of scalar quantities (the

reaction rates) $r(\tau) : R_i(\tau) \times P_i(\tau) \to \mathbb{R}^+$ which dictate the rate with which each reaction occurs, i.e for each reaction labeled by τ we have:

$$\sum_i m_i(\tau) R_i(\tau) \xrightarrow{r(\tau)} \sum_i n_i(\tau) P_i(\tau)\,. \tag{8}$$

The rate equations for the expected number of things x_i for each species are then derived from the general formula:

$$\frac{dx_i(t)}{dt} = \sum_\tau r(\tau)(n_i(\tau) - m_i(\tau))\mathbf{x}^{\mathbf{m}(\tau)}\,. \tag{9}$$

3.1.1 Exercises

1. Draw the Perti nets for the following chemical reaction networks:

   (a)

   $$\mathrm{S} + \mathrm{I} \to 2\mathrm{I}\,, \tag{10a}$$
   $$\mathrm{I} \to \mathrm{R}\,. \tag{10b}$$

   (b)

   $$\mathrm{A} + \mathrm{B} \to 2\mathrm{A}\,, \tag{11a}$$
   $$\mathrm{A} + \mathrm{B} \to 2\mathrm{B}\,. \tag{11b}$$

   (c)

   $$\mathrm{R} \to 2\mathrm{R}\,, \tag{12a}$$
   $$\mathrm{R} + \mathrm{F} \to 2\mathrm{F}\,, \tag{12b}$$
   $$\mathrm{F} \to \emptyset \tag{12c}$$

2. For each picture above, think of a scenario (for instance in population dynamics, epidemiology, or sociology) which is modeled by that Perti net.

3. For each of the Perti nets in question 1, derive the rate equations.

4. Consider the rate equations (6) for the SIS model.

   (a) Find a conserved quantity from the set of equations (hint: conserved quantities do not change in time)

   (b) Use the conserved quantity to eliminate x_S in favor of x_I in equation (6a)

   (c) Can you write down a Petri net for a single species which is governed by the same equation? If not, why not? If yes, draw the Perti net and give an interpretation of the transitions.

   (d) Identify the steady state solutions x_0 (obeying $dx_\mathrm{I}/dt = 0$)

   (e) When are the steady state solution stable? In other words, for what values of the dimensionless ratio $\lambda = \alpha/\beta$ do linearized perturbations around the steady state solution ($x_\mathrm{I} = x_0 + \epsilon$) with $\epsilon \ll 1$, decay?

   (*) Integrate the differential equation and give an explanation of its behavior. (*Hint*: the obtained equation is famously called the logistic equation.)

3.2 A Fock space for the copy number representation

So far, we have only scratched the surface of chemical reaction kinetics. One could write entire books[3] on the kinetics of chemical reactions and its implications in chemistry. However, here we are interested in connecting the rate equations, which are a manifestation of what is called *mass action kinetics*, to a microscopic theory, where we describe the number of particles of a given species as fluctuating randomly (i.e. stochastically). To do so, we will employ a language which is quite similar to that of the Fock space representation of quantum mechanics, although there are essential and important differences with the quantum case, which we will highlight below. First, let's get to the physical description of the system.

We are now interested in the microscopics of the reactions, which implies that we do want to track the number of particles in the system. However, we are not sure exactly when the reactions take place, we just know the probability of two elements reacting. This means we wish to describe the system by the probability $P(\mathbf{n}, t)$ of having n_i elements of species i, i.e. the *copy number* for species i, at any given time t. Moreover, we want to know the dynamics of this probability; how does the likelihood of having n_i elements of species i change in time?

To arrive at an equation for $P(\mathbf{n}, t)$, we first construct a vector space for the probability distribution in the copy number representation, using creation and annihilation operators familiar from second quantization in bosonic systems [13,16,17]. This essentially exploits the fact that the number of elements of a certain species can only change by an integer. We start with defining the empty state $|0\rangle$, the creation operators $\hat{a}_i^\dagger$ and annihilation operators $\hat{a}_i$ for species i such that:

$$\hat{a}_i|0\rangle = 0\,, \qquad |\mathbf{n}\rangle = \prod_{i\in S}(\hat{a}_i^\dagger)^{n_i}|0\rangle \equiv (\hat{\mathbf{a}}^\dagger)^{\mathbf{n}}|0\rangle\,. \tag{13}$$

Here $\hat{\mathbf{a}}^\dagger$ is a vector of creation operators, one for each species and $\mathbf{n} \in \mathbb{Z}^d$ is a vector containing all of their copy numbers. We also use the notational shortcut defined in (7).

The state $|\mathbf{n}\rangle$ represents the system containing $\mathbf{n} = (n_1, n_2, \ldots)$ copies of the species under consideration. As each one of them is in principle indistinguishable, when we remove (annihilate) a particle of species i, there are n_i equivalent ways to do so and hence:

$$\hat{a}_i|\mathbf{n}\rangle = n_i|n_1, \ldots, n_i - 1, \ldots, n_k\rangle\,. \tag{14}$$

On the other hand, there is only a single way to add a particle of species i, and hence:

$$\hat{a}_i^\dagger|\mathbf{n}\rangle = |n_1, \ldots, n_i + 1, \ldots, n_k\rangle\,. \tag{15}$$

The above relations imply that the creation and annihilation operators satisfy bosonic commutation relations:

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}\,, \qquad [\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0\,. \tag{16}$$

In words, the ladder operators for different species commute, whereas like species ($i = j$) satisfy canonical commutation relation. Note that there is nothing quantum about non-commuting operators by itself. Here, the non-commutativity just means that there is one more way to first add a particle and then remove one, than there is to first remove a particle and then add one.

[3] In fact, some people have done exactly this, see for instance [14, 15]

An important operator is the number operator $\hat{N}_i = \hat{a}_i^\dagger \hat{a}_i$ as it counts the number of particles of species i in the state $|\mathbf{n}\rangle$:

$$\hat{N}_i|\mathbf{n}\rangle = n_i|\mathbf{n}\rangle\,. \tag{17}$$

In other words, the state $|\mathbf{n}\rangle$ is an eigenvector for the number operator N_i with eigenvalue n_i.

We are now ready to define the main object of study, namely the probability vector $|\phi(t)\rangle$, which is defined as:

$$|\phi(t)\rangle = \sum_{\mathbf{n}} P(\mathbf{n},t)|\mathbf{n}\rangle\,. \tag{18}$$

Here $P(\mathbf{n},t)$ is the probability of having $\mathbf{n}$ copies of each species at time t.

To continue, we need a few more important ingredients. So far we have defined vectors, but to compute expectation values (or any other scalar quantity of interest) we also need dual vectors to take inner products. The full story is a bit more complicated[4], but for our purpose we can simply define the dual vacuum to be annihilated by the creation operators, and construct a Fock space with annihilation operators:

$$\langle 0|\hat{a}_i^\dagger = 0\,, \qquad \langle \mathbf{n}| = \prod_{i\in S}\langle 0|\hat{a}_i^{n_i} \equiv \langle 0|\hat{\mathbf{a}}^{\mathbf{n}}\,. \tag{19}$$

Additionally, we require $\langle 0|0\rangle = 1$. The states $\langle \mathbf{m}|$ and $|\mathbf{n}\rangle$ are orthogonal (exercise 3 below), but the normalization here differs a little bit from the one in quantum mechanics. Here one picks up a factor $\mathbf{n}! \equiv \prod_i n_i!$ in the inner product:

$$\langle \mathbf{m}|\mathbf{n}\rangle = \mathbf{n}!\delta_{\mathbf{m},\mathbf{n}}\,. \tag{20}$$

This implies that a resolution of the identity should be defined with the normalization factor of $1/\mathbf{n}!$:

$$\mathbb{1} = \sum_{\mathbf{n}} \frac{1}{\mathbf{n}!}|\mathbf{n}\rangle\langle\mathbf{n}|\,. \tag{21}$$

Now we have vectors, which may signify the probability of having a certain number of entities of a given species, and we have dual vectors, which we may use to take inner products and compute scalar quantities. We can now start to think what this means in terms of probability distributions and how to compute useful quantities.

First, we need a special dual state, which will signify the sum over all components of $|\phi(t)\rangle$. Due to the normalization of probability, we will have that $\sum_{\mathbf{n}} P(\mathbf{n},t) = 1$ for all t. In our quantum inspired language, this will be represented by contracting $|\phi(t)\rangle$ with a **flat state** $\langle \mathbf{1}|$:

$$\sum_{\mathbf{n}} P(\mathbf{n},t) = \langle \mathbf{1}|\phi(t)\rangle = 1\,. \tag{22}$$

What properties do we require from the flat state? Since it sums uniformly over all possible states, when the probability distribution is peaked on a single state (let's say $|\mathbf{n}\rangle$), then the flat state will need to contain the dual vector $\frac{1}{\mathbf{n}!}\langle\mathbf{n}|$ (note the normalization factor is due to (20)). Another property we could require of the flat state is that it should not matter how

[4]This is because contrary to quantum mechanics, the vector (18) is an element of the L^1 space of normalizable function, whose dual is L^∞

many states are present in the system, it should count them all. So if we increase the number of states by creating another particle of species i with $\hat{a}_i^\dagger$, it should not affect the outcome of summing over all state. To put this in mathematical terms, the flat state should be a left eigenvector for the creation operators:

$$\langle \mathbf{1}|\hat{a}_i^\dagger = \langle \mathbf{1}|\,, \qquad \forall i \in S\,. \tag{23}$$

These requirements fix the flat state to be:

$$\langle \mathbf{1}| = \sum_{\mathbf{n}} \frac{1}{\mathbf{n}!}\langle \mathbf{n}| = \langle 0|e^{\sum_i \hat{a}_i}\,. \tag{24}$$

Please verify for yourself that this definition satisfies the eigenvector equation (23).

With the flat state defined, we are able to compute expectation values of observables. Any observable $A(\mathbf{n})$ can first be cast in terms of creation and annihilation operators $\hat{A}(\hat{a}_i, \hat{a}_i^\dagger)$. The expectation value of A at time t is then computed as:

$$\langle A\rangle_t = \sum_{\mathbf{n}} A(\mathbf{n})P(\mathbf{n},t) = \langle \mathbf{1}|\hat{A}(\hat{a}_i, \hat{a}_i^\dagger)|\phi(t)\rangle\,. \tag{25}$$

For example, if we wish to compute the expected value of the number of elements of species i, then the observable of interest is the number operator $\hat{N}_i = \hat{a}_i^\dagger \hat{a}_i$ and the expectation value becomes:

$$\langle N_i\rangle = \langle \mathbf{1}|\hat{N}_i|\phi(t)\rangle = \langle \mathbf{1}|\hat{a}_i|\phi(t)\rangle\,. \tag{26}$$

The last equation here follows from the definition of $\hat{N}_i$ along with equation (23).

We have seen already a lot of mathematical machinery being defined, which we will shortly bring to use in deriving the dynamical equation for $|\phi(t)\rangle$, but we wish to define one last state which will be important in the remaining sections: the **coherent state**. it is defined analogously to coherent states in quantum mechanics, although (as will be a theme in these lectures) its interpretation is slightly different. A coherent state $|\mathbf{z}\rangle$ is defined as:

$$|\mathbf{z}\rangle = e^{\sum_{i\in S} z_i \hat{a}_i^\dagger}|0\rangle\,. \tag{27}$$

Here, the exponential of an operator is defined in the usual way $e^{\hat{a}_i^\dagger} = \sum_n \frac{1}{n!}(\hat{a}_i^\dagger)^n$. Hence, the coherent state is a function of a 'species vector' z_i with $i \in S$ and it hence defines (when properly normalized) a parameterized probability distribution over the Fock space. In fact, as you will see in exercise 6 below, it constitutes a very special distribution, namely the *Poisson distribution*, one for each species i. Recall that the Poisson distribution gives the probability of having a certain number of elements in a given interval of time, if they are generated at random with a fixed rate.

Similar to the construction above, we can also define the dual coherent states as:

$$\langle \mathbf{z}| = \langle 0|e^{\sum_{i\in S} z_i \hat{a}_i}\,. \tag{28}$$

From this definition, we immediately observe that the flat state (24) is a special case of a dual coherent state with all $z_i = 1$. We will encounter other uses for the more general dual coherent state later.

Now we have all the ingredients we need to formulate the stochastic master equation for reaction systems.

3.2.1 Exercises:

1. Show that equations (13), (14) and (15) lead to the commutation relations for the creation and annihilation operators (16).

2. Show that the state $|\mathbf{n}\rangle$ (defined in (13)) is an eigenvector for the number operator $\hat{N}_i = \hat{a}_i^\dagger \hat{a}_i$ for species i.

3. Show that the states $\langle \mathbf{n}|$ and $|\mathbf{m}\rangle$ are orthogonal, using their definitions and the commutation relations (16)

4. Show that $\sum_{\mathbf{n}} \frac{1}{\mathbf{n}!}\langle \mathbf{n}| = \langle 0|e^{\sum_i \hat{a}_i}$ by using the appropriate definitions and the usual definition of the exponential of an operator $e^{\hat{a}_i} = \sum_n \frac{1}{n!}\hat{a}_i^n$

5. Proof that the coherent state $|\mathbf{z}\rangle = e^{\sum_i z_i \hat{a}_i^\dagger}|0\rangle$ is an eigenvector for the annihilation operator $\hat{a}_i$ with eigenvalue z_i:
$$\hat{a}_i|\mathbf{z}\rangle = z_i|\mathbf{z}\rangle \tag{29}$$

6. Consider the case where we have only a single species A and a coherent state $|z\rangle = e^{z\hat{a}^\dagger}|0\rangle$, where $\hat{a}^\dagger$ is the creation operator for A.

   (a) Compute $c(z)$ such that the state $|\phi(z)\rangle = c(z)|z\rangle$ forms a normalized probability distribution (such that $\langle \mathbf{1}|\phi(z)\rangle = 1$).

   (b) Compute the probability $P(n)$ of having n elements of species A in the state $|\phi(z)\rangle$. In other words, compute the components $P(n)$ of $|\phi(t)\rangle = \sum_n P(n)|n\rangle$. What distribution does $\phi(z)$ signify?

   (c) What is the expected number of elements of species A in the distribution $|\phi(z)\rangle$? (hint: this is computed by $\langle \mathbf{1}|\hat{N}|\phi(z)\rangle$)

   (d) What is the variance in the expected number of elements of species A in the distribution $|\phi(z)\rangle$? In other words, what is $\langle \mathbf{1}|\hat{N}^2|\phi(z)\rangle - \langle \mathbf{1}|\hat{N}|\phi(z)\rangle^2$

3.3 The master equation

The master equation describes how the probability of having a certain number of elements of any given species changes in time. Generically, it will hence have the form:

$$\frac{d}{dt}|\phi(t)\rangle = \hat{H}(\{\hat{a}_i^\dagger, \hat{a}_i\})|\phi(t)\rangle \tag{30}$$

This equation shows great similarity with the (imaginary time) Schrödinger equation, except that here, the 'Hamiltonian' $\hat{H}$ is not necessarily Hermitian. In fact, calling $\hat{H}$ a Hamiltonian may be misleading, since it is in general not Hermitian (as Hamiltonians in quantum mechanics are), nor is it a function specifying the systems energy in equilibrium (as Hamiltonians in equilibirum statistical mechanics are). In the literature $\hat{H}$ is actually referred to by many different names, which may depend on the discipline the papers are directed to. We might call it: a quasi-Hamiltonian, the infinitesimal generator, the Markov generator, the transition rate matrix, or the time-evolution operator, among others. Here we will mostly stick to calling $\hat{H}$ *the generator*. In any case, its name is of secondary importance, what matters more are it's properties.

The generator $\hat{H}$ should satisfy a number of properties, such that Eq. (30) can describe the time evolution of a probability distribution. First of all, we would like the total amount of probability to be conserved in time. This means that after evolving the system for some time t, starting from an arbitrary normalized state $|\phi(0)\rangle$, the state should remain normalized. Since the formal solution of (30) is $|\phi(t)\rangle = \exp(\hat{H}t)|\phi(0)\rangle$, this implies that $\langle \mathbf{1}|e^{\hat{H}t}|\phi(0)\rangle = 1$, but $|\phi(0)\rangle$ is arbitrary, so:

$$\langle \mathbf{1}|e^{\hat{H}t} = \langle \mathbf{1}|\,, \quad \Rightarrow \quad \langle \mathbf{1}|\hat{H} = 0\,. \tag{31}$$

To put in words, the sum over all states of $\hat{H}$ should vanish, or: the flat state is a left eigenvector of any generator, with eigenvalue 0.

Furthermore, we cannot have that $\hat{H}$ will generate distributions which contain negative probabilities for some choice of copy numbers. This implies that any operation implemented by $\hat{H}$ which effectively creates or annihilates particles of a certain species should occur with positive sign.

To understand this better, consider the situation where $\hat{H}$ creates a single particle of species A with rate α, corresponding to the reaction $\emptyset \to A$. We would write a term in $\hat{H}$ proportional to $\alpha\hat{a}^\dagger$, with $\hat{a}^\dagger$ the creation operator for A. If we start with an initial state where there are no particles, then $|\phi(0)\rangle = |0\rangle$. After an infinitesimal time δt, we will have $|\phi(\delta t)\rangle = |0\rangle + \alpha\,\delta t\,|1\rangle + \mathcal{O}((\delta t)^2)$. The probability in this state to have one particle is $\propto \alpha\,\delta t$, which shows that $\alpha \geq 0$, otherwise this becomes negative.

So any generator $\hat{H}$ which conserves total probability and respects the non-negativity of the distributions should satisfy (31) and at the same time have non-negative terms for transitions which create or annihilate states. That means that any such transition should be balanced by a negative term which effectively leaves the number of elements of each species unchanged, otherwise the sum over all states in (31) would give something positive (not zero).

Lets make this a bit more concrete with some examples and then give the general recipe to turn any CRN into a generator $\hat{H}$. Consider again the simplest possible case of creating a single element of species A with rate α. We already knew that $\hat{H}$ in that case should contain a term $\alpha\hat{a}^\dagger$, however, by itself this will not satisfy (31). In fact,

$$\langle \mathbf{1}|\alpha a^\dagger = \alpha\langle \mathbf{1}|\,, \tag{32}$$

by using equation (23). So, the proper generator for this process is actually

$$\hat{H} = \alpha(\hat{a}^\dagger - 1)\,. \tag{33}$$

This generator is easily verified to satisfy $\langle \mathbf{1}|\hat{H} = 0$ as it should. Moreover, the formal solution $|\phi(t)\rangle = e^{\hat{H}t}|\phi(0)\rangle$, starting from an initial state where $|\phi(t=0)\rangle = |0\rangle$ is the coherent state:

$$|\phi(t)\rangle = e^{-\alpha t}e^{\alpha t\,\hat{a}^\dagger}|0\rangle = e^{-\alpha t}|\alpha t\rangle\,. \tag{34}$$

Hence, we see explicitly that the process where a single particle is created at random with a fixed rate is described by the coherent state, which we know from the last section (exercise 6) to be equivalent to the Poisson distribution.

As another example, consider the reaction $A + B \to B + B$ with reaction rate r. The reaction can be thought of as first annihilating one element of species A and one of species B,

followed by creating two elements of species B. So, the generator should have a term $r\hat{b}^\dagger\hat{b}^\dagger\hat{b}\hat{a}$, where $\hat{a}(\hat{b})$ are the annihilation operators for species $A(B)$ and likewise for the creation operators.

Once again, this term by itself will not satisfy (31), so we should balance the transition with a 'diagonal' term $r\hat{b}^\dagger\hat{a}^\dagger\hat{b}\hat{a}$ to obtain the generator:

$$\hat{H} = r(\hat{b}^\dagger\hat{b}^\dagger - \hat{b}^\dagger\hat{a}^\dagger)\hat{b}\hat{a}\,, \tag{35}$$

The first term reflects the positive change in probability of obtaining two B elements in the reaction, while the second term signifies the negative change in the probability of A, B pairs remaining in the solution.

In general, for each reaction τ in the reaction network specified by the transition:

$$\sum_i m_i(\tau)R_i(\tau) \xrightarrow{r(\tau)} \sum_i n_i(\tau)P_i(\tau)\,, \tag{36}$$

we need to add the following terms:

- The term which implements the transition with a positive sign: $r(\tau)(\hat{\mathbf{a}}^\dagger)^{\mathbf{n}(\tau)}\hat{\mathbf{a}}^{\mathbf{m}(\tau)}$
- The negative diagonal term which ensures total probability is conserved and the loss of probability of keeping the reactants is taken into account: $r(\tau)(\hat{\mathbf{a}}^\dagger)^{\mathbf{m}(\tau)}\hat{\mathbf{a}}^{\mathbf{m}(\tau)}$.

Hence, to turn any CRN into a infinitesimal generator $\hat{H}$, we can use the general formula:

$$\hat{H} = \sum_\tau r(\tau)\left[(\hat{\mathbf{a}}^\dagger)^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^\dagger)^{\mathbf{m}(\tau)}\right]\hat{\mathbf{a}}^{\mathbf{m}(\tau)} \tag{37}$$

Here (and above) we are once more using the notation that $\hat{\mathbf{a}}^{\mathbf{n}} = \prod_i \hat{a}_i^{n_i}$. Any generator $\hat{H}$ which satisfies the property of conserving non-negative probabilities under time-evolution is called **infinitesimal stochastic**. Hence, instead of Hermitian, in stochastic mechanics time evolution is generated by infinitesimal stochastic operators! However, it is perfectly possible for an operator to be both Hermitian and infinitesimal stochastic. In that special case, the operator will define both a stochastic system (through the master equation) *and* a quantum system (through the Schrödinger equation) and it is called a *Dirichlet operator*.

The formal solution of the master equation defines a *one-parameter family semigroup* of linear operators $\hat{U} = e^{\hat{H}t}$. This operator is a **stochastic** operator, in the sence that is satisfies $\langle\mathbf{1}|\hat{U}(t) = \langle\mathbf{1}|$. Note that we already encountered stochastic operators in (23); any creation operator $\hat{a}_i^\dagger$ is stochastic, since it does not influence the total amount of probability when acting on a probability vector from the left. This implies that whenever we are computing expectation values of some observables, due to the contraction with $\langle\mathbf{1}|$ from the left, we can set all $\hat{a}_i^\dagger \to 1$ after normal ordering[5].

So now we have the general recipe to turn any CRN into an infinitesimal generator. We will use this in the following sections to obtain the field theory formulation when we add spatial dependence to the problem and describe reaction-diffusion systems. But first, we can play around a bit more with the obtained mathematical formalism and describe the link to the macroscopic rate equations (including which assumptions are needed for this) and we digress a little bit to discuss the statistical mechanics of non-equilibrium stochastic systems through the lens of what is called *large deviation theory*.

[5]Here normal ordering places creation operators to the left of annihilation operators by using the commutation relations (16).

3.3.1 Exercises:

1. For all the reactions in (3.1.1), write down the generator $\hat{H}$

2. We have encountered already many similarities between stochastic and quantum systems, however the two are definitely not the same. To make the distinction crystal clear, what are the stochastic analogues of the following quantum mechanical properties and statements?

   - The wavefunction $|\psi(t)\rangle$ and its amplitudes ψ_n, such that $|\psi(t)\rangle = \sum_n \psi_n |n\rangle$ and $\langle\psi(t)|\psi(t)\rangle = 1$ for all t.
   - The expectation values for observables $\bar{O} = \langle\psi|\hat{O}|\psi\rangle$
   - A Hermitian operator satisfying $\hat{H}^\dagger = \hat{H}$
   - A one-parameter unitary group $\hat{U}(t)$, such that $\hat{U}^{-1} = \hat{U}^\dagger$. (This one's actually a bit tricky since stochastic systems are not generally invertible, what is the stochastic analogue of this statement?)

3. Consider the process:
$$\mathrm{A} \to \emptyset\,. \tag{38}$$
with rate α (called the linear death process)

   (a) Write down the generator for this process

   (b) Suppose that initially our system has k particles of species A, such that $|\phi(0)\rangle = |k\rangle = (\hat{a}^\dagger)^k|0\rangle$. What is $|\phi(t)\rangle$?
   Hint: there are two commutator identities which make this question a lot easier to solve. First, you may use the Campbell identity:
$$e^X Y e^{-X} = \sum_{n=0}^{\infty} \frac{[(X)^n, Y]}{n!}\,, \tag{39}$$
   where $[(X)^n, Y] = \underbrace{[X, \ldots [X, [X}_{n}, Y]] \ldots]$ is the nested commutator. Then, if Y is an integer power of some other operator $Y = A^k$, then we may write:
$$\sum_{n=0}^{\infty} \frac{[(X)^n, A^k]}{n!} = \left(\sum_{n=0}^{\infty} \frac{[(X)^n, A]}{n!} \right)^k\,. \tag{40}$$

   (c) Derive from $|\phi(t)\rangle$ the probability of having $n < k$ particles at time t.

   (d) What is the expected number of particles at time t?

   (e) What is the variance in the number of particles at time t?

4. Usually, the master equation is formulated in terms of $\frac{d}{dt}P(n,t)$ and derived heuristically from arguing how the probability of having n particles changes as a result of the reactions present in the system.

   For instance, for the death process of the previous question, one would say that $\frac{d}{dt}P(n,t)$ receives $(n+1)$ positive contributions from $P(n+1,t)$ with rate α, because if we are

in a state with $n+1$ particles (which happens with probability $P(n+1,t)$), then each of the $n+1$ particles could decay with rate α, leading to the state with n particles.

Likewise, $P(n,t)$ receives n negative contributions from $P(n,t)$ with rate α, because if we are in the state with n particles, each of these could decay and take us out of the n particle state. So we have heuristically argued that:

$$\frac{d}{dt}P(n,t) = \alpha(n+1)P(n+1,t) - \alpha n P(n,t)\,. \tag{41}$$

(a) Derive this equation from the second quantized master equation by normal ordering and equating like powers of $\hat{a}^\dagger$ on both sides.

(b) Repeat this for the (linear) birth-death process, where there are two reactions:

$$\mathrm{A} \to \emptyset\,, \tag{42a}$$

$$\mathrm{A} \to 2\mathrm{A}\,. \tag{42b}$$

3.4 Limit to rate equations

We have now seen two approaches to model systems which are described by a chemical reaction network. In section 3.1 we have discussed the *rate equations* which tell us how the total number of each species changes in time (or, if we divide by the total number of elements, it tells us how concentrations of things change in time). The rate equations are deterministic, which implies that given an initial composition of elements, the equations determine completely and uniquely the future evolution of the system.

In the last two sections, we have explored another approach to the same system, one which is stochastic. We have formulated a *master equation*, which dictates how the probability of having a certain number of elements of each species changes in time. This formulation is probabilistic, hence given an initial composition of elements, we can obtain distributions over the final outcome. Although it might appear as this gives less information on the state of the system, it actually gives us more information, because from the complete distribution we can obtain not only the expected number of things at a given time, but also the variance and any higher moment of the distribution. Additionally, we can answer questions like: "what is the probability of observing more than x elements of species Y?"

Within the framework of emergence presented in section 2.1, we can think of the rate equations as providing a top theory for chemical reaction networks. It tells us how macroscopic observables such as concentrations change in time, without referring to the individual constituents, i.e. the change in concentrations depends only on the concentrations itself. The bottom theory in this case is the stochastic master equation. This theory *does* refer to the microscopic constituents and gives a probabilistic description of the system. In this section, we wish to connect the two descriptions and see under what assumptions can we go from the micro theory to the macro theory.

The macroscopic variables are actually straightforward to define from the microscopic theory. For the total number of elements of a certain species, we can simply take the expected number of elements of this species and so, in accordance with the notation above, we will define:

$$x_i(t) = \langle N_i \rangle_t = \langle \mathbf{1} | \hat{N}_i | \phi(t) \rangle\,. \tag{43}$$

The next step is to work out the dynamical equations. Lets first look at what the microscopic theory would give using the master equation (30):

$$\frac{d}{dt}x_i(t) = \langle \mathbf{1}|\hat{N}_i\frac{d}{dt}|\phi(t)\rangle = \langle \mathbf{1}|\hat{N}_i\hat{H}|\phi(t)\rangle = \langle \mathbf{1}|[\hat{N}_i,\hat{H}]|\phi(t)\rangle\,. \tag{44}$$

In the last line here, we have used the property of the generator $\langle \mathbf{1}|\hat{H} = 0$ to express the answer in terms of the commutator. Hence, the time-evolution of the expected value of the number operator (in fact, of any operator) is given by the expectation value of the commutator with the generator $\hat{H}$, another parallel with quantum mechanics!

Let's see how this works out with some simple example first. Consider the linear birth-death process of the last exercise:

$$\mathrm{A} \to \emptyset\,, \tag{45a}$$

$$\mathrm{A} \to 2\mathrm{A}\,. \tag{45b}$$

From the rules of section 3.1, we can derive the rate equation for $x = x_\mathrm{A}$ to be:

$$\dot{x}(t) = -(\alpha - \beta)x(t)\,, \tag{46}$$

where α is the death rate and β birth rate. The generator for this process is obtained by following the prescription of section 3.3:

$$\hat{H} = \alpha(1-\hat{a}^\dagger)\hat{a} + \beta(\hat{a}^\dagger\hat{a}^\dagger - \hat{a}^\dagger)\hat{a}\,. \tag{47}$$

We can now derive the equation for $x(t)$ straight from computing the expected value of the commutator of $\hat{N}$ with $\hat{H}$:

$$\begin{aligned}\frac{d}{dt}x(t) &= \langle \mathbf{1}|[\hat{N},\hat{H}]|\phi(t)\rangle = -\alpha\langle \mathbf{1}|\hat{a}|\phi(t)\rangle + \beta\langle \mathbf{1}|\hat{a}|\phi(t)\rangle \\ &= -(\alpha-\beta)x(t)\,.\end{aligned} \tag{48}$$

Everything seems to work out immediately, which is often more than we can wish for. Indeed, the simplicity of starting with a reaction with only a single reactant hides a very important subtlety in the derivation of the rate equations.

A second example shows how this simple reasoning can fail. Consider now the SIS process introduced in (2) (with Petri net displayed in Figure 5). In this case, the generator is

$$\hat{H} = \alpha(\hat{a}_\mathrm{I}^\dagger\hat{a}_\mathrm{I}^\dagger - \hat{a}_\mathrm{S}^\dagger\hat{a}_\mathrm{I}^\dagger)\hat{a}_\mathrm{S}\hat{a}_\mathrm{I} + \beta(\hat{a}_\mathrm{S}^\dagger - \hat{a}_\mathrm{I}^\dagger)\hat{a}_\mathrm{I}\,. \tag{49}$$

Now, if we wish to obtain the rate equation for the total number of infected x_I, we should work out carefully the expectation value for the commutator of $\hat{N}_\mathrm{I}$ with the generator $\hat{H}$. We find that:

$$\frac{d}{dt}x_\mathrm{I}(t) = -\beta x_\mathrm{I}(t) + \alpha\langle \mathbf{1}|\hat{a}_\mathrm{S}\hat{a}_\mathrm{I}|\phi(t)\rangle\,. \tag{50}$$

Comparing with (6), we see that we get the correct answer, only if we say that

$$\langle \mathbf{1}|\hat{a}_\mathrm{S}\hat{a}_\mathrm{I}|\phi(t)\rangle \cong \langle \mathbf{1}|\hat{a}_\mathrm{S}|\phi(t)\rangle\langle \mathbf{1}|\hat{a}_\mathrm{I}|\phi(t)\rangle = x_\mathrm{S}(t)x_\mathrm{I}(t)\,. \tag{51}$$

This equation only holds if the number of S and I are not correlated, such that the expectation of their product equals the product of the expectation values. So, we see that the rate

equations of the macroscopic theory suppose that the (co)variances between the total number of elements of each species should vanish. This assumption is generally called a **mean field approximation**. It gives a simplification of the process, which is a good first approximation, but of course, it is generically not true exactly! In the case of the SIS model, we would actually expect the number of elements of S and I to be (negatively) correlated; after all, the I's can infect the S 'particles' and hence if there are more infected, it is less likely for a susceptible to stay susceptible.

In general, for the chemical reaction networks generated by (37), one can derive (exercise 2) that:

$$\frac{d}{dt}x_i(t) = \sum_\tau r(\tau)(n_i(\tau) - m_i(\tau))\langle \mathbf{1}|\hat{\mathbf{a}}^{\mathbf{m}(\tau)}|\phi(t)\rangle\,. \tag{52}$$

If we focus on a single species for illustrative purposes, this implies that the general rate equations are derivable from the master equation only when:

$$\langle \mathbf{1}|\hat{a}^m|\phi(t)\rangle - \langle \mathbf{1}|\hat{a}|\phi(t)\rangle^m = 0\,. \tag{53}$$

This is not likely the most realistic situation in cases where variances ($m = 2$ above) are large. But in many cases, the law of large numbers and the *central limit theorem* can offer a helping hand. The law of large numbers states (roughly) that for n *independent and identically distributed* (iid) random variables $\{X_n\}$, the sample mean $S_n = \frac{1}{n}\sum_n X_n$ will converge to the mean μ of the distribution from which the X_n are sampled. Or: $\lim_{n\to\infty} S_n = \mu$. The central limit theorem states (equally roughly) that in the same large n limit, the distribution of the sample mean S_n becomes Gaussian with standard deviation $\sigma/\sqrt{n}$, where σ is the standard deviation of the distribution for the random variables. Or:

$$S_n \to \mu + \xi/\sqrt{n}\,, \tag{54}$$

where ξ is normally distributed a $\mathcal{N}(0, \sigma^2)$. This in particular implies that the expectation value of any power of the sample mean S_n^k will tend to the power of the mean μ^k, with subleading corrections proportional to $\frac{1}{n}$. These corrections vanish in the large n limit.

In our case, we can apply this when we assume that the number of elements of each species is very large, or likewise, if we are considering the dynamics of concentrations of species in a large volume expansion of the CRN. In chemistry this is called the van Kampen expansion, after [18]. In these situations, and assuming the elements are iid, we can approximate $\langle \hat{a}^k\rangle$ as $\langle \hat{a}\rangle^k$ and in this limit we recover the rate equations (5) from the master equation description of the system.

Hence, we have seen that we can relate the top theory (rate equations) to the bottom theory (master equations) by a mean-field approximation, which becomes more accurate after a limiting procedure (large volume expansion). Through the limit we loose information on the microscopics of the system, as (co)variances and higher-moments will not come to matter anymore. As emergent phenomena, we do gain the property of having a closed and deterministic system, described by a small set of ordinary differential equations, which was not true for the microscopic system.

3.4.1 Exercises:

1. Within the framework of emergence, what are generic emergent features of the macroscopic rate equations not present in the microscopic stochastic models?

2. Work out the rate equation of $x_i(t)$ in a general chemical reaction network with generator $\hat{H}$ given in (37). When do you reproduce equation (5)?

3. Consider again the SIS model and the equation for $x_{\mathrm{I}}(t)$ given in (50). Suppose that we do not wish to assume that $\langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle \cong \langle \mathbf{1}|\hat{a}_{\mathrm{S}}|\phi(t)\rangle\langle \mathbf{1}|\hat{a}_{\mathrm{I}}|\phi(t)\rangle$. Then, to describe the systems evolution, we would also need an equation for the SI pairs: $x_{\mathrm{SI}} = \langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle$. Derive this equation from the master equation with generator (49)

4. This exercise we will explore *Noether's theorem* for stochastic processes, based partly on the work of [12, 19]. Noether's theorem relates symmetries to conservation laws. In quantum mechanics, Noether's theorem can loosely be phrased as:

$$[\hat{O}, \hat{H}] = 0 \Leftrightarrow \partial_t \langle \hat{O} \rangle = 0\,. \tag{55}$$

In words: any observable which commutes with the (quantum) Hamiltonian has an expectation value which does not change in time (i.e. is conserved). The implication works both ways, so any conserved quantity also corresponds to an observable which commutes with the Hamiltonian. (Noether's theorem is, of course, more general than this, but for here we will do with conservation in time.)

(a) Show that the relation $[\hat{O}, \hat{H}] = 0 \Rightarrow \partial_t \langle \hat{O} \rangle = 0$ also works in stochastic mechanics, so whenever $\hat{H}$ is an infinitesimal generator and $\langle \hat{O} \rangle$ is computed in the L_1 norm as: $\langle \hat{O} \rangle = \langle \mathbf{1}|\hat{O}|\phi(t)\rangle$.

Above we have shown that Noether's theorem works one way for stochastic systems. Unfortunately, the converse is not necessarily true. To show this, consider the following reaction network:

$$\mathrm{B} \to \mathrm{A} \tag{56a}$$

$$\mathrm{B} \to \mathrm{C} \tag{56b}$$

where (importantly) both reaction occur with the same rate α.

(b) Define an observable $\hat{O} = \hat{N}_B + 2\hat{N}_C$. Show that in the above system, we have $\partial_t \langle \hat{O} \rangle = 0$ even though $[\hat{O}, \hat{H}] \neq 0$.

(c) Is the second moment of the observable $\hat{O}$ (defined as $\langle \hat{O}^2 \rangle$) also conserved? I.e. what is $\partial_t \langle \hat{O}^2 \rangle$?

(d) Can you provide a physical explanation as to why Noethers theorem fails in this case?

The above situation shows that we need supplementary conditions to show that the conservation of a quantity leads to a observable which commutes with the Hamiltonian. In [19] it was shown that, for instance $\partial_t \langle \hat{O}^2 \rangle = 0$ is sufficient and so that the Noether's theorem for stochastic mechanics can be formulated as

$$[\hat{O}, \hat{H}] = 0 \Leftrightarrow \partial_t \langle \hat{O} \rangle = 0 = \partial_t \langle \hat{O}^2 \rangle\,. \tag{57}$$

(*) Proof Noether's theorem for stochastic mechanics. You may use that observables in stochastic mechanics are 'diagonal' in the sense that we can always expand them as $\hat{O} = \sum_n \frac{1}{n!} O_n |n\rangle\langle n|$.

3.5 Generating functions

Ultimately, the probability vector we have been constructing gives us a formal way of representing an infinite sequence of numbers, which in our case represents the probability of our system to have a certain copy number $\mathbf{n}$. In mathematics, and specifically in probability theory, another useful representation of infinite sequences is through a *generating function*. The generating function encodes an infinite sequence of numbers (a_n) by treating them as coefficients of a formal power series of an undetermined and arbitrary variable z. If the infinite sequence a_n has a finite sum, then the formal power series will converge and if we are lucky we may even obtain a closed form expression for the generating function in terms of analytic functions of the variable z. In any case, we will see that it is possible to derive a partial differential equation for the generating function for the probability $P(\mathbf{n}, t)$ for a generic chemical reaction network. Afterwards, we will define the **moment generating function** for an observable, where the series coefficients a_n are not the probabilities $P(\mathbf{n}, t)$, but relate to the different statistical *moments* of an observable $\langle \hat{O}^n \rangle$.

Formally, the generating function for an infinite series a_n is defined as:

$$G(a_n; z) = \sum_n a_n z^n \,. \tag{58}$$

This naturally generalizes to multiple variables. If we suppose there are k species, then:

$$G(a_{\mathbf{n}}; \mathbf{z}) = \sum_{\mathbf{n}} a_{\mathbf{n}} \mathbf{z}^{\mathbf{n}} = \sum_{\{n_1, n_2, \ldots, n_k\}} a_{\{n_1, n_2, \ldots, n_k\}} \prod_{i=1}^{k} z_i^{n_i} \tag{59}$$

The generating function is then a formal power series in k variables. We can now translate the probability vector $|\phi(t)\rangle$ into a *probability generating function*, where the components of the formal power series are the probabilities $P(\mathbf{n}, t)$. To do so, we can make use of the dual coherent state $\langle \mathbf{z}|$ defined in (28):

$$G(\mathbf{z}, t) = \sum_{\mathbf{n}} P(\mathbf{n}, t) \mathbf{z}^{\mathbf{n}} = \langle \mathbf{z} | \phi(t) \rangle \tag{60}$$

The probability generating function has a number of useful and interesting properties. Lets go through a few in the case of a single species to lighten up the notation. The generalization to multiple species is immediate. First, the normalization of $|\phi(t)\rangle$ is now translated to the property that the probability generating function evaluated at $z = 1$ equals 1:

$$G(z = 1, t) = \langle 1 | \phi(t) \rangle = 1 \tag{61}$$

Then, if we wish to recover the probability mass function $P(n, t)$ giving the probability of having n elements, we can do so by taking n derivatives with respect to z, followed by evaluating the result at $z = 0$

$$P(n, t) = \frac{1}{n!} \frac{d^n G(z, t)}{dz^n} \bigg|_{z=0} \tag{62}$$

The expectation value for the number of elements at time t is also extracted by taking derivatives, but now evaluated at $z = 1$:

$$\langle N \rangle_t = \frac{d}{dz} G(z, t) \bigg|_{z=1} \,. \tag{63}$$

If we wish to compute higher moments of the number operator, we can do so by the following relation:

$$\langle N^k \rangle_t = \left(z \frac{d}{dz} \right)^k G(z,t) \Big|_{z=1} . \tag{64}$$

This expression invites the identification of the number operator $N = \hat{a}^\dagger \hat{a}$ with $z\frac{d}{dz}$, such that creation operators correspond to multiplications with z and annihilation operators are differentiation with respect to z. Indeed, speaking in terms of probability generating functions is completely equivalent to making the following identifications in the formalism developed in the last sections:

$$\hat{a}_i^\dagger = z_i \,, \qquad \hat{a}_i = \frac{\partial}{\partial z_i} \,. \tag{65}$$

This representation of the creation and annihilation operators satisfies the canonical commutation relations (16) and allows us to write the general master equation in terms of a partial differential equation for the probability generating function:

$$\partial_t G(\mathbf{z},t) = \sum_\tau r(t) (\mathbf{z}^{\mathbf{n}(\tau)} - \mathbf{z}^{\mathbf{m}(\tau)}) \partial_{\mathbf{z}}^{\mathbf{m}(\tau)} G(\mathbf{z},t) \,. \tag{66}$$

Here, we are still using the shorthand notation for exponentials (7) and $\partial_{\mathbf{z}}^{\mathbf{m}} = \prod_{i \in S} \left(\frac{\partial}{\partial z_i} \right)^{m_i}$. Here are two simple examples which illustrate the computation of the probability generating function:

Example 1: The Poisson process $\emptyset \to A$. The equation for the probability generating function is:

$$\partial_t G(z,t) = \alpha (z-1) G(z,t) \,. \tag{67}$$

This is easily solved as $G(z,t) = e^{\alpha t (z-1)}$, which is the same result as (34) after replacing the creation operator with z.

Example 2 Besides the Poisson process, we may also obtain the solution to the pure death process $A \to \emptyset$ immediately from the differential equation for the probability generating function. The differential equation for $G(z,t)$ is now:

$$\partial_t G(z,t) = \alpha (1-z) \partial_z G(z,t) \,. \tag{68}$$

A trick to solve these kinds of equations is to find a characteristic curve c in terms of z and t, such that $G(z,t) = f(c(z,t))$. We can do this by supposing that z has some implicit time dependence, such that G is a function of $z(t)$ and t and the equation (68) splits into two equations:

$$\frac{d}{dt} G(z(t),t) = 0 \,, \qquad \frac{d}{dt} z(t) = \alpha (z-1) \,. \tag{69}$$

The latter equation is solved by $z(t) = 1 + c e^{\alpha t}$, which introduces a constant c. Solving for c gives:

$$c = (z-1) e^{-\alpha t} \,, \quad \text{such that:} \;\; G(z,t) = f(c(z,t)) \,. \tag{70}$$

One can verify readily that $f(c)$ obeys equation (68) when c is given as above. The final step is to consider the initial conditions. If at $t = 0$, there are k particles present, then we must

have $G(z,0) = z^k$ and this then fixes the functional form of $f(c)$ to be $(1 + c(z,0))^k$ such that:

$$G(z,t) = (1 + e^{-\alpha t}(z-1))^k \,, \tag{71}$$

as we have seen in exercise 3 of section 3.3.1, but now with $\hat{a}^\dagger$ replaced by z.

Next to the probability generating function, the formalism also allows one to compute **moment generating functions** for observables. The moment generating function for the observable $\hat{O}$ is defined as the generating function for the series $a_n = \langle \hat{O}^n \rangle / n!$. This is equivalent to the expectation value of $e^{s\hat{O}}$, if s is the variable of the power series. In the second quantized framework:

$$M_{\hat{O}}(s,t) = \langle e^{s\hat{O}} \rangle_t = \langle \mathbf{1} | e^{s\hat{O}} | \phi(t) \rangle \,. \tag{72}$$

For instance, if the observable is the number operator for species i, then we can express its moment generating function as:

$$M_{\hat{N}_i}(s_i,t) = \langle \mathbf{1} | e^{s_i \hat{N}_i} | \phi(t) \rangle \tag{73}$$

We can also construct the moment generating function for number operators of all species, which will then depend on a vector of variables $\mathbf{s}$:

$$M_{\hat{\mathbf{N}}}(\mathbf{s},t) = \langle \mathbf{1} | e^{\sum_i s_i \hat{N}_i} | \phi(t) \rangle \tag{74}$$

It is possible to show that this is equivalent to the following expression:

$$M_{\hat{\mathbf{N}}}(\mathbf{s},t) = \langle e^{\mathbf{s}} | \phi(t) \rangle \equiv \langle 0 | e^{\sum_i e^{s_i} \hat{a}_i} | \phi(t) \rangle \,. \tag{75}$$

Additionally, if we use the fact that $|\phi(t)\rangle = e^{\hat{H}t}|\phi(0)\rangle$ and (37) for the generic generator of a chemical reaction network, then:

$$M_{\hat{\mathbf{N}}}(\mathbf{s},t) = \langle \mathbf{1} | e^{\sum_i s_i \hat{N}_i} e^{\hat{H}t} | \phi(0) \rangle = \langle \mathbf{1} | e^{\tilde{H}(s)t} e^{\sum_i s_i \hat{N}_i} | \phi(0) \rangle \,, \tag{76}$$

with:

$$\tilde{H}(s) = \sum_\tau r(\tau) \left[e^{\sum_i s_i (n_i(\tau) - m_i(\tau))} (\hat{\mathbf{a}}^\dagger)^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^\dagger)^{\mathbf{m}(\tau)} \right] \hat{\mathbf{a}}^{\mathbf{m}(\tau)} \,. \tag{77}$$

So, the moment generating function is related to the exponential of a so-called *tilted* generator $\tilde{H}(s)$, which is easily obtained from the Markov generator $\hat{H}$ by replacing the creation operators and annihilation operators as:

$$\hat{a}_i^\dagger \to e^s \hat{a}_i^\dagger \,, \qquad \hat{a}_i \to e^{-s} \hat{a}_i \,. \tag{78}$$

Such replacements are sometimes called an exponential tilting, hence the name tilted generator.

3.5.1 Exercises

1. Proof the definition of the probability generating function (60) using the definition of $\langle \mathbf{z}|$ (28) and the definition of $|\phi(t)\rangle$.

2. Proof the identification of (65) for the probability generating function by showing that $\langle z|\hat{a}^\dagger|\phi(t)\rangle = z\langle z|\phi(t)\rangle$ and $\langle z|\hat{a}|\phi(t)\rangle = \frac{d}{dz}\langle z|\phi(t)\rangle$

3. Derive the differential equations for the generating functions corresponding to the processes of 3.1.1 question 1.

4. Proof that (74) and (75) are equivalent, using a corollary of the Campbell-Baker-Hausdorf formula:
$$e^X e^Y = e^Y \exp\left(\sum_{n=0}^{\infty} \frac{1}{n!}[X, Y^{(n)}]\right) , \tag{79}$$
where $[X, Y^{(n)}] = [\dots[[X, \underbrace{Y], Y], \dots Y}_{n}]$is the nested commutator

5.* Using a formula similar to (79), proof (76).

3.6 A little bit of Large Deviation Theory

We saw in the last section how the moment generating function for the number operator is given by the exponential of a tilted generator. A similar tilted generator can also be defined for another observable, called the *dynamical activity* K. This observable is not easily expressible in terms of creation and annihilation operators, because it is a so-called trajectory dependent observable. The dynamical activity counts the total number of reactions which have taken place at a given time t. Since the reactions happen as a random process, the dynamical activity K is also a random variable.

To obtain its moment generating function $M_K(s,t)$ (following [20, 21]), we have to adapt our formalism slightly. We should not only track the probabilities $|\phi(t)\rangle$, but we should define the vector $|\phi(t,k)\rangle$ which gives the probability of having a certain combination of copy numbers at time t AND having seen k reactions since $t = 0$. Then, the moment generating function for the dynamical activity becomes:

$$M_K(s,t) = \langle e^{sK}\rangle_t = \sum_{k=0}^{\infty} \langle \mathbf{1}|e^{sk}|\phi(t,k)\rangle = \langle \mathbf{1}|\tilde{\phi}(t,s)\rangle \tag{80}$$

Here we have defined the tilted probability vector $|\tilde{\phi}(t,s)\rangle = \sum_k e^{sk}|\phi(t,k)\rangle$ as the (discrete) Laplace transform of $|\phi(t,k)\rangle$. It is possible to show that for a general chemical reaction network, the tilted probability vector obeys the following (tilted) master equation:

$$\partial_t|\phi(t,s)\rangle = \tilde{H}(s)|\phi(t,s)\rangle\,, \tag{81}$$

with

$$\tilde{H}(s) = \sum_\tau r(\tau)\left[e^s(\hat{\mathbf{a}}^\dagger)^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^\dagger)^{\mathbf{m}(\tau)}\right]\hat{\mathbf{a}}^{\mathbf{m}(\tau)}\,. \tag{82}$$

In words: the tilted generator for the dynamical activity has all off-diagonal terms (which generate reactions) *exponentially tilted* by multiplying them with e^s. Now the formal solution to (81) is $|\phi(t,s)\rangle = e^{\tilde{H}(s)t}|\phi(0)\rangle$, assuming that initially there has been no activity, so $K = 0$ at $t = 0$. This allows one to write the moment generating function as the exponential of the tilted generator:

$$M_K(s,t) = \langle \mathbf{1}|e^{\tilde{H}(s)t}|\phi(0)\rangle \tag{83}$$

We can think of this quantity, (or any moment generating function for that matter) as the analogue to a partition function in equilibrium statistical mechanics. Only here the observable is not energy, with dual parameter inverse temperature (β), but it is the tilted generator, for which time serves as dual parameter. The interpretation is that this encapsulates information on the dynamical activity K (which has dual parameter s) at any time.

A modern approach to statistical mechanics is given in a subject called **large deviation theory**. We will not give an exhaustive review of this, but rather refer to the review of Hugo Touchette [22] for a historical and topical background on this fascinating subject, which connects statistical mechanics with the mathematics of probability theory and the analysis of rare events (i.e. the occurrence of events which great deviations from the mean value, hence the name).

In the context of this lecture, we are interested in how the dynamical activity will behave at late times, so in the limit of $t \to \infty$. The **large deviation principle** states that the probability of seeing events with activity k decays exponentially according to:

$$P(k,t) \propto \exp(-tI(k/t))\,. \tag{84}$$

Here $I(k)$ is the *rate function* (also called the Cramér function). The rate function has the property that its zeros define the expected values of k/t, as for these values the probability $P(k,t)$ does not decay in time. For all other values of k/t, the rate function determines the late time decay rate of the probability of seeing these events, and hence it contains information on how probabilities for observing events deviating from the mean behave.

The rate function is related to the moment generating function in the following way. It is the Legendre-Fenchel transform of an object called the *scaled cumulant generating function* (SCGF) $\Theta(s)$, or:

$$I(k) = \sup_s(k\,s - \Theta(s))\,, \tag{85}$$

where $\Theta(s)$ is defined as:

$$\Theta(s) = \lim_{t\to\infty} \frac{1}{t} \log M_K(s,t)\,. \tag{86}$$

This definition is equivalent to stating that the moment generating function $M_K(s,t)$ behaves as:

$$M_K(s,t) \propto e^{t\Theta(s)}\,, \tag{87}$$

when $t \to \infty$. As we have just seen, the moment generating function is also the exponential of the tilted generator (82), and hence the leading contribution to $M_K(s,t)$ will come from the largest eigenvalue of $\tilde{H}(s)$. This argument relates the SCGF to the leading eigenvalue of the tilted generator as a function of s. Hence to find $\Theta(s)$ one can do something similar to a ground state search in quantum mechanics, where one optimizes for the lowest eigenvalue of a Hamiltonian. Only in this case, the Hamiltonian is a tilted generator, which in general is not Hermitian, and neither is it infinitesimal stochastic, since the exponential tilting spoils this property. Hence, it is not always a very easy object to compute and might involve extensive numerical computations.

The analogy with equilibrium statistical mechanics we are laying bare here is that the MGF $M_K(s)$ is a type of (dynamical) partition function. Then the SCGF is the analog of free energy and the rate function (as log of a probability) is analogous to entropy. Taking

a late time limit might be thought of as increasing β, which then relates to the low energy regime, although this is only a suggestive analogy.

We are brushing a number of important details under the carpet here. The large deviation principle is not always guaranteed to exist, just as the leading eigenvectors of $\tilde{H}(s)$ are not always unique. In some cases, the eigenvalue decomposition of $\tilde{H}$ does not exist (as not all matrices are diagonalizable). Generally, chemical reaction networks (or Markov processes in general) with (one or multiple) absorbing states are going to cause problems, since the dynamics can get stuck in absorbing state, which makes the late time properties sensitive to the initial conditions. What we need is a system which is *ergodic*, such that every microscopic configuration can be reached from any other configuration by a series of transitions. Quite often, ergodicity is implicitly assumed for physical systems, but it is good practice to verify whether this assumption really holds. In the case of CRN, a system can always be made ergodic by including the reverse reactions, where the reverse reaction rates can be taken to be small.

4 Continuum limit: reaction-diffusion systems

In the last section we have described the dynamics of chemical reaction networks at two levels, the rate equations which give the deterministic dynamics of concentrations and the master equation, which gives the stochastic dynamics of the copy numbers. One underlying assumption in both of these modeling approaches, is the lack of spatial dependence in the problem. The description is hence valid under the assumption that spatial factors do not play a significant role, which implies that the reactants are homogeneously distributed (or well-mixed), and more importantly, remain so during the systems temporal evolution.

The purpose of this section is to relax the assumption of spatial homogeneity and introduce spatial dependence into the microscopic probabilistic model developed in the last section. This will lead to a classical, statistical field theory for *reaction-diffusion* systems. Now, instead of having a finite and constant probability of reacting with one-another, particles rely on diffusion to be brought into the vicinity of one another. These are also called *diffusion-limited* reactions. In chemistry, it means the reactants in the solution should not be stirred, but all elements are subject to *Brownian motion*. In population biology, we may think of members of a species performing a *random walk* trough an environment in order to find each other.

Below we will briefly review the relation between random walks and the diffusion equation and formulate this in terms of the second-quantized description of the last section. Then, following [17,23] we will turn the master equations for the reaction systems of the last section into a bona fide field theory for non-equilibrium reaction-diffusion systems. Finally, we derive the propagator of the free theory for reaciton-diffusion systems and note that this is nothing else than the Green's function for the diffusion equation. Much of this section is based on the review [24] and [25, 26] as well as chapter 2 of the book [27].

4.1 Random walks and the emergence of diffusion

There are several complementary ways to model random walks and their continuum limit to Brownian motion and the diffusion equation, which we will certainly not all discuss here.

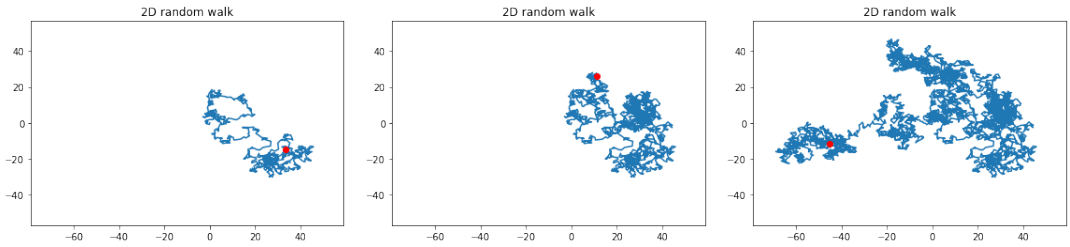


Figure 6: The trajectory of a random walk in two dimensions after 750 (left), 2000 (middle) and 5000 (right) hops.

For more details on this exemplary subject which lies at the heart of many processes in non-equilibrium statistical physics we refer to [27,28]. Here we will mostly be concerned with the minimal ingredients necessary for our field theory derivation of reaction-diffusion systems.

The microscopic diffusion process can be modeled as a random walk on a regular lattice by means of a master equation. We introduce a regular, d-dimensional lattice with lattice spacing h and suppose that a particle has equal probability of hopping from any site of the lattice to any neighboring site. Importantly, each walker's hopping probability is independent of the presence of any other walker. This means that we can treat the walkers as iid (independent and identically distributed), such that it suffices to study the evolution of a single walker. This single walker traces out a trajectory on the d-dimensional space, hopping from site to site and making twists and turns independent from what site they are at or what happened at the last hop. One example of such a trajectory in $d = 2$ is plotted in Figure 6.

Consider first the continuous time master equation for a single walker on a one-dimensional lattice. The probability $P(x,t)$ of the walker being in position x at time t is subject to the following continuous time master equation:

$$\frac{\partial}{\partial t}P(x,t) = \frac{D}{h^2}\left[P(x-h,t) + P(x+h,t) - 2P(x,t)\right] \tag{88}$$

where we have set (with some amount of hindsight) the hopping rate for the walker to $\frac{D}{h^2}$. The equation simply states that $P(x,t)$ can change in two different ways. The first two terms account for the walker coming into position x from a neighboring lattice site to the left or right, the negative term accounts for the walker being at position x and leaving either to the left or to the right, which accounts for the factor of 2,

While the above master equation may be solved (after Fourier transformation in space) in terms of modified Bessel functions of the first kind, what we are after in this section is the continuum limit, where we send the lattice spacing $h \to 0$. It is then instructive to Taylor expand $P(x \pm h, t)$ to second order, such that

$$P(x \pm h, t) = P(x,t) \pm h\partial_x P(x,t) + \frac{h^2}{2}\partial_x^2 P(x,t) + \dots \tag{89}$$

Plugging this into (88) and taking the limit $h \to 0$ gives the diffusion equation:

$$\frac{\partial}{\partial t}P(x,t) = D\partial_x^2 P(x,t)\,. \tag{90}$$

Generalization of the above argument to a d-dimensional lattice results in the d dimensional diffusion equation:

$$\partial_t P(\mathbf{x},t) = D\nabla^2 P(\mathbf{x},t)\,. \tag{91}$$

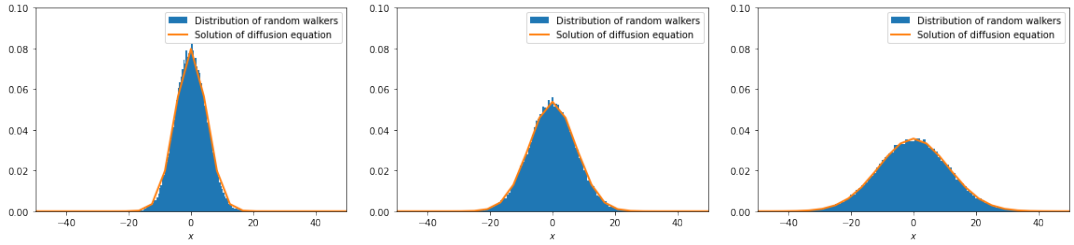


Figure 7: The distribution of 50000 random walkers starting at the origin, compared with the solution of the diffusion equation (91) at three different times.

Here $\nabla^2 = \vec{\nabla} \cdot \vec{\nabla}$ is the Laplacian. The solution to this equation is most easily found in terms of the Fourier transform

$$P(\mathbf{k}, t) = \int_{-\infty}^{\infty} d^d x P(\mathbf{x}, t) e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{92}$$

after which (91) becomes $\partial_t P(\mathbf{k}, t) = -Dk^2 P(\mathbf{k}, t)$ (with $k^2 = \mathbf{k} \cdot \mathbf{k}$), which simply gives:

$$P(\mathbf{k}, t) = e^{-Dk^2} P(k, 0)\,. \tag{93}$$

We take as initial condition a single walker at the origin, such that $P(\mathbf{x}, 0) = \delta^d(\mathbf{x})$ and $P(k, 0) = 1$. It is then not hard to perform the inverse Fourier transform and obtain:

$$P(\mathbf{x}, t) = \frac{1}{(4\pi D t)^{d/2}} e^{-\frac{x^2}{4Dt}}\,. \tag{94}$$

So we see that the solution to the diffusion equation takes the form of a (multivariate) normal distribution with vanishing mean and variance growing linearly in time as $2Dt$. Interestingly, here we see another instance of the central limit theorem. Namely, the continuous time and space distribution of a random walk is given by a Gaussian (normal) distribution, which holds because at each discretized lattice point, the next step of the walk is an independent and identically distributed random variable. Moreover, since all walkers are identical and independently distributed, the equation describing a macroscopic density of Brownian particles $\rho(\mathbf{x}, t)$ is identical to the diffusion equation (91) with P replaced by ρ. So we see here another instance of the emergence of deterministic macroscopic system (described by the diffusion equation) from a probabilistic microscopic model (the random walk, or Brownian motion). Indeed, by simulating a large number of random walkers starting at the same initial position, the empirical distribution of the walker positions at time t agrees with the Gaussian distribution (94) (see Figure 7).

In terms of the physical properties of the walker, (94) implies that the expected position of the walker does not change in time, but its mean square distance from the starting position grows linearly in time. So, at a time t, the walker can explore roughly a sphere around the origin of radius $\sqrt{Dt}$. The hopping rate is constant, however, so the number of sites visited grows linearly with time t. This implies that the density of sites visited within the exploration radius is proportional to $\rho \sim t/t^{d/2} = t^{1-d/2}$. From this heuristic argument, one is led to the conclusion that the number of spatial dimensions has an important effect on the property of the random walk, namely the question if the walker can explore all of space if given infinite time. If $d < 2$, ρ diverges as $t \to \infty$ and so each site is visited infinitely often. Similarly, in this regime the random walk is certain to return to its starting point, a

feature called *recurrence*. When $d > 2$, ρ decreases with time and hence there will be points which are never visited. In this case, there is non-zero probability for the random walk to never return to its starting position and the random walk is called *transient*. The marginal case where $d = 2$ turns out to be recurrent, although a more careful analysis is required to show that the density of visited sites will then diverge logarithmically in time. Indeed, more careful deliberations (see for instance [27]) lead to the following asymptotic behavior of the *survival probability* $\mathcal{S}(t)$, defined as the probability that the random walk has not returned to its starting position:

$$\mathcal{S}(t) \sim \begin{cases} A_d t^{d/2-1}\,, & d < 2\,, \\ A_2 (\ln t)^{-1}\,, & d = 2\,, \\ (1-\mathcal{R}) + C_d (1-\mathcal{R})^2 t^{1-d/2}\,, & d > 2\,. \end{cases} \tag{95}$$

Here A_d, A_2 are t-independent factors and $\mathcal{R}$ is the *eventual return* probability that the walker will ultimately reach its starting position.

In the context of the reaction-diffusion systems we are interested in, we only need to make a small modification to the master equations above. First, we wish to allow multiple particles at each lattice site, and ultimately, we should also allow particles of multiple species, such that they can interact if they find themselves at the same lattice site. Additionally, we should express the master equation in terms of creation and annihilation operators.

Because the addition of multiple species does not give any technical difficulties besides making the notation more cumbersome, lets consider just a single species A for now, but label the lattice sites by i, j etc. Now, the state of the system can be described by an array $\{n\} = \{n_1, n_2, \ldots\}$ which corresponds to having n_1 particles at the first lattice site, n_2 at the second site, etc. The probability $P(\{n\}; t)$ of the system being in the state labeled by occupation numbers $\{n\}$ at time t changes in time by the master equation:

$$\begin{aligned} \partial_t P(\{n\};t) = \frac{D}{h^2} \sum_{\langle i,j \rangle} \big[&(n_i+1) P(\{\ldots, n_i+1, n_j-1, \ldots\};t) - n_i P(\{n\};t) \\ &+ (n_j+1) P(\{\ldots, n_i-1, n_j+1, \ldots\};t) - n_j P(\{n\};t) \big] \end{aligned} \tag{96}$$

Here the summation is over nearest neighboring sites i and j. The first line on the r.h.s. accounts for particles hopping from i to j, while the second line accounts for particles hopping the other way. The probability vector $|\phi(t)\rangle$ will now be expanded into a Fock space for the lattice occupation numbers $\{n\}$, such that: $|\phi(t)\rangle = \sum_{\{n\}} P(\{n\};t) \prod_i (\hat{a}_i^\dagger)^{n_i} |0\rangle$. Hence, now we suppose that $\hat{a}_i, \hat{a}_i^\dagger$ are the annihilation and creation operators for species A, but at lattice site i. The random walk master equation (96) becomes:

$$\partial_t |\phi(t)\rangle = \frac{D}{h^2} \sum_{\langle i,j \rangle} \big[\hat{a}_j^\dagger \hat{a}_i - \hat{a}_i^\dagger \hat{a}_i + \hat{a}_i^\dagger \hat{a}_j - \hat{a}_j^\dagger \hat{a}_j \big] |\phi(t)\rangle \tag{97}$$

$$= -\frac{D}{h^2} \sum_{\langle i,j \rangle} (\hat{a}_i^\dagger - \hat{a}_j^\dagger)(\hat{a}_i - \hat{a}_j) |\phi(t)\rangle \equiv \hat{H}_0(D) |\phi(0)\rangle\,. \tag{98}$$

The generalization to multiple species is immediate, although notationally cumbersome. Now there is a set of occupation numbers for each species. Suppose that now we have a set of

species S, then each has an associated set of numbers $\{n\}_1, \{n\}_2$, etc. Lets denote that total set of $|S|$ occupation numbers $\{n\}_S$. Then $|\phi(t)\rangle$ is expanded in the basis:

$$|\{n\}_S\rangle = \prod_i \prod_{k\in S} (\hat{a}^\dagger_{i,k})^{n_{i,k}} |0\rangle \,. \tag{99}$$

where here $\hat{a}^\dagger_{i,k}$) is the creation operator for species k at lattice site i and $n_{i,k}$ is the number of particles of species k at lattice site i. Fortunately, the random walk process for each species is independent of any other species, and so the multi-species diffusion generator $\hat{H}_0(\{D\}_S)$ is just the sum of the single species diffusion generators, where each species k may have a different and independent diffusion constant D_k:

$$\hat{H}_0(\{D\}_S) = \sum_{k\in S} \hat{H}_0(D_k) \,. \tag{100}$$

You can see now that adding more species is not more complicated, but it is not particularly enlightening in terms of notation, so we proceed to perform the field theory limit for reactions of the type $kA \to \ell A$, where there is only a single species involved.

4.2 A field theory for reaction-diffusion systems

Just as a quantum field theory can be obtained from quantum mechanics, a statistical (classical) field theory can be obtained from the second-quantized representation of stochastic mechanics. Here we will outline the continuum limit leading to the field theory for reaction-diffusion systems with a single species. Hence, we are restricting ourselves to reaction systems of the type $kA \to \ell A$, however, the generalization to multiple reacting species is not difficult and follows along very similar lines. The prescription presented here relies on the coherent state representation (although a slightly different one compared to the one presented in section 3.2). Using the coherent states as resolution of the identity operator a path integral can be constructed, first discussed in this context for the birth-death processes ($A \to 2A$ and $A \to \emptyset$) by Peliti in [17]. Hence this framework is sometimes called the Peliti (or Doi-Peliti) path integral representation.

First, we add the diffusion generator $\hat{H}_0(D)$ of equation (98) to the reaction generator, and we suppose that only particles at the same lattice location can react. This implies that:

$$\begin{aligned} \hat{H} &= \hat{H}_0(D) + \hat{H}_I(\{\lambda\}) \,, \\ &= -\frac{D}{h^2} \sum_{\langle i,j\rangle} (\hat{a}^\dagger_i - \hat{a}^\dagger_j)(\hat{a}_i - \hat{a}_j) + \sum_{i,\tau} \lambda(\tau) \left[(\hat{a}^\dagger_i)^{\ell(\tau)} - (\hat{a}^\dagger_i)^{k(\tau)} \right] \hat{a}_i^{k(\tau)} \,. \end{aligned} \tag{101}$$

Here, τ labels the different reactions present in the system, each with their own reaction rate $\lambda(\tau)$ and stoichiometric coefficients $k(\tau)$ and $\ell(\tau)$. Somewhat suggestively, we have already named the diffusive part of the generator $\hat{H}_0$ and the reactive part $\hat{H}_I$, foreshadowing the fact that in the final action we can treat $\hat{H}_0$ non-perturbatively, as the 'free theory' for reaction-diffusion systems, but we should treat the final action perturbatively with respect to the non-linearities induced by the reactions in $\hat{H}_I$.

To proceed and build the Peliti path integral for observables such as $\langle A\rangle_t = \langle \mathbf{1}|\hat{A}|\phi(t)\rangle$, we introduce a new coherent state, now defined in terms of a complex valued variable η as:

$$|\eta\rangle = e^{-\frac{1}{2}|\eta|^2} e^{\eta \hat{a}^\dagger} |0\rangle \,, \quad \langle \eta| = \langle 0| e^{\eta^* \hat{a}} e^{-\frac{1}{2}|\eta|^2} \,. \tag{102}$$

Here the star denotes complex conjugation. The overlap between two of these states is simply computed as:

$$\langle \eta_1 | \eta_2 \rangle = e^{-\frac{1}{2}|\eta_1|^2 - \frac{1}{2}|\eta_2|^2 + \eta_1^* \eta_2} = e^{-\eta_1^*(\eta_1 - \eta_2)} e^{\frac{1}{2}|\eta_1|^2 - \frac{1}{2}|\eta_2|^2}\,, \tag{103}$$

where the last equality is a convenient rewriting which will be useful later. The reason for introducing these complex valued coherent states is so that we may define a useful resolution of the identity operator:

$$\mathbb{1} = \int \frac{d^2\eta}{\pi} |\eta\rangle\langle\eta|\,, \tag{104}$$

or, if we define one such coherent state for each lattice site i, then:

$$\mathbb{1} = \int \prod_i \left(\frac{d^2\eta_i}{\pi} \right) |\{\eta\}\rangle\langle\{\eta\}|\,, \tag{105}$$

where $d^2\eta_i = d\mathrm{Re}(\eta_i) d\mathrm{Im}(\eta_i)$ and $|\{\eta\}\rangle = |\eta_1\rangle \otimes |\eta_2\rangle \otimes \ldots$.

The path-integral for $\langle A \rangle_t$ is then obtained by first using the formal solution of the master equation $\langle A \rangle_t = \langle \mathbf{1} | \hat{A} e^{\hat{H}t} | \phi(0) \rangle$. Then, the time evolution induced by the generator $\hat{H}$ is divided into infinitely many slices of infinitesimal size:

$$e^{\hat{H}t} = \lim_{\Delta t \to 0} \left(e^{\hat{H}\Delta t} \right)^{t/\Delta t} = \lim_{\Delta t \to 0} \left(1 + \Delta t \hat{H} \right)^{t/\Delta t} \tag{106}$$

Then, we insert resolutions of the identity operator (105) in between each of the infinitesimally thin slices of $e^{\hat{H}\Delta t}$. Let τ label the slices, then:

$$\langle A \rangle_t = \mathcal{N}^{-1} \lim_{\Delta t \to 0} \int \left(\prod_{i,\tau} d^2\eta_{i,\tau} \right) \langle \mathbf{1} | \hat{A} | \{\eta\}_t \rangle \left[\prod_{\tau = \Delta t}^{t} \langle \{\eta\}_\tau | (1 + \Delta t \hat{H}) | \{\eta\}_{\tau - \Delta t} \rangle \right] \langle \{\eta\}_0 | \phi(0) \rangle\,. \tag{107}$$

Here $\mathcal{N}$ is a normalization factor, which can be determined in the end by requiring that $\hat{A} = 1$ gives unity. Let's discuss the terms appearing in this expression one by one.

First, the operator $\hat{A}$ can be made to only depend on annihilation operators $\hat{a}_i$. This was discussed below equation (37); any operator $\hat{A}(\{\hat{a}_i^\dagger\}, \{\hat{a}_i\})$ can be brought to a normal ordered form, where all creation operators are to the left of the annihilation operators. By the property $\langle \mathbf{1} | \hat{a}_i^\dagger = \langle \mathbf{1} |$, the creation operators can be set to one: $\{\hat{a}_i^\dagger\} \to 1$. Moreover, since the coherent state is an eigenvector of the annihilation operator $\hat{a}_i |\{\eta\}\rangle = \eta_i |\{\eta\}\rangle$, we have that $\langle \mathbf{1} | \hat{A} | \{\eta\}_t \rangle = \langle \mathbf{1} | \{\eta\}_t \rangle \hat{A}(1, \{\eta\}_t)$, or: we can simply replace all annihilation operators $\hat{a}_i$ in $\hat{A}$ with the complex variables $\eta_{i,t}$. The inner product $\langle \mathbf{1} | \{\eta\}_t \rangle$ gives $\prod_i \exp(-\frac{1}{2}|\eta_{i,t}|^2 + \eta_{i,t})$ by equation (103).

The terms in the middle square brackets induce a similar replacement as with $\hat{A}$, but now for $\hat{H}$. Indeed, here we may write:

$$\begin{aligned} \langle \{\eta\}_\tau | (1 + \Delta t \hat{H}) | \{\eta\}_{\tau - \Delta t} \rangle &= \langle \{\eta\}_\tau | \{\eta\}_{\tau - \Delta t} \rangle (1 + \Delta t \hat{H}(\{\eta^*\}_\tau, \{\eta\}_{\tau - \Delta t})) \\ &= \langle \{\eta\}_\tau | \{\eta\}_{\tau - \Delta t} \rangle \exp(\Delta t \hat{H}(\{\eta^*\}_\tau, \{\eta\}_{\tau - \Delta t}))\,. \end{aligned} \tag{108}$$

where $\hat{H}(\{\eta^*\}_\tau, \{\eta\}_{\tau - \Delta t})$ is given by $\hat{H}$ with creation operators replaced as $\hat{a}_i^\dagger \to \eta_{i,\tau}^*$ and annihilation operators replaced by $\hat{a}_i \to \eta_{i,\tau - \Delta t}$. In the last line above we have restored the original exponential form of $\hat{H}\Delta t$. This is related to a subtle point, where we are actually

sneakily swapping a sum (in the definition of the exponential operator) with doing an integral (in this case integrating over the complex numbers $d^2\eta$). Such a procedure does not always neatly work out, so we should keep in the back of our heads the fact that we are actually *always* defining $e^{\hat{H}t}$ in terms of its power series. So the actual explicit computation of the final object implicitly implies a perturbative series, which can be cast into the form of a power series in λ.

The inner product in (108) is explicitly:

$$\langle\{\eta\}_\tau|\{\eta\}_{\tau-\Delta t}\rangle = \prod_i \langle\eta_{i,\tau}|\eta_{i,\tau-\Delta t}\rangle = e^{-\sum_i \eta^*_{i,\tau}(\eta_{i,\tau}-\eta_{i,\tau-\Delta t})} e^{\sum_i \frac{1}{2}|\eta_{i,\tau}|^2 - \frac{1}{2}|\eta_{i,\tau-\Delta t}|^2}\,. \tag{109}$$

Here the last line follows from (103). Now we see that multiplying over all the τ slices will have the effect of canceling out the terms in the second exponential, except the initial term $-\frac{1}{2}|\eta_{i,0}|^2$ and the final term $\frac{1}{2}|\eta_{i,t}|$. This final term, however, cancels the factor of $-\frac{1}{2}|\eta_{i,t}|^2$ which came from the inner product of $\langle\mathbf{1}|\{\eta\}_t\rangle$. The first exponential term on the right hand side in (109) can be expanded as a Taylor series in Δt to give the time derivative of $\eta_{i,\tau}$ as leading term: it becomes $e^{-\sum_i \eta^*_{i,\tau}\frac{d}{d\tau}\eta_{i,\tau}\Delta t + \mathcal{O}(\Delta t^2)}$.

Finally, the last term in (107) concerns with the initial conditions. If we start with a lattice filled uniformly at random, then the occupation at each lattice site is Poisson distributed. We already know that Poisson distributions are coherent states, where the parameter of the coherent state corresponds to the mean of the distribution. So for each lattice site to be filled with on average n_0 particles, the initial state $|\phi(0)\rangle = |n_0\rangle \sim e^{\sum_i n_0 \hat{a}_i^\dagger}|0\rangle$.[6] The inner product with $\langle\{\eta\}_0|$ again has the effect that all creation operators $\hat{a}_i^\dagger$ are replaced with annihilation operators and so, when taken together with the factors of $\exp(-\frac{1}{2}|\eta_{i,0}|^2)$ left over from (109):

$$\langle\{\eta\}_0|\phi(0)\rangle e^{-\sum_i \frac{1}{2}|\eta_{i,0}|^2} \sim e^{\sum_i \eta^*_{i,0}(n_0-\eta_{i,0})} \tag{110}$$

Let's put everything back together and then take the limit $\Delta t \to 0$. First, we denote the final time as t_f for clarity and use t as the time variable. Putting everything together:

$$\langle A\rangle_{t_f} = \mathcal{N}^{-1} \lim_{\Delta t\to 0} \int \left(\prod_{i,t} d^2\eta_{i,t}\right) A(\{\eta\}_{t_f}) e^{-S(\{\eta^*\},\{\eta\})}\,. \tag{111}$$

with

$$S = \sum_i \left(-\eta_{i,t_f} - \eta^*_{i,0}(n_0-\eta_{i,0}) + \sum_{t=\Delta t}^{t_f} \Delta t \left(\eta^*_{i,t}\frac{d}{dt}\eta_{i,t} - H(\{\eta^*_{i,t}\},\{\eta_{i,t-\Delta t}\})\right)\right) \tag{112}$$

We can now take the $\Delta t \to 0$ limit and replace the $\sum_t \Delta t$ sum by an integral $\int_0^{t_f} dt$. The η's then become continuous functions of t and the Δt time difference in $H(\{\eta^*_{i,t}\},\{\eta_{i,t-\Delta t}\})$ disappears, with the understanding that when it matters, we should think of the $\eta^*(t)$ fields as following the $\eta(t)$ field in time. The $\Delta t \to 0$ limit turns the measure $\prod_{i,t} d^2\eta_{i,t}$ into functional differentials $\mathcal{D}\eta_i\mathcal{D}\eta_i^*$ and we obtain:

$$\langle A\rangle_t = \mathcal{N}^{-1} \int \mathcal{D}\eta_i\mathcal{D}\eta_i^* A(\{\eta_i(t)\}) e^{-S(\{\eta^*\},\{\eta\})}\,. \tag{113}$$

[6] All normalization factors we can just hide in $\mathcal{N}$

where $\mathcal{N} = \int \mathcal{D}\eta_i \mathcal{D}\eta_i^* e^{-S(\{\eta^*\},\{\eta\})}$ and the *action* S reads as:

$$S = \sum_i \left(-\eta_i(t_f) - \eta_i^*(0)(n_0 - \eta_i(0)) + \int_0^{t_f} dt \left(\eta_i^*(t) \frac{d}{dt} \eta_i(t) - H(\{\eta_i^*(t)\}, \{\eta_i(t)\}) \right) \right) \tag{114}$$

We are almost done. What rests is the continuum space limit, and a discussion on the meaning of the initial and final terms in the action which are still remaining. The continuum limit is taken as

$$\sum_i \to \int h^{-d} d^d x \,, \qquad \eta_i(t) \to h^d \varphi(\mathbf{x}, t) \,, \qquad \eta_i^*(t) \to \tilde{\varphi}(\mathbf{x}, t) \,. \tag{115}$$

Here it should be clear from context the φ's are fields and not the probability vectors discussed earlier. The complex conjugate fields $\tilde{\varphi}$ and φ are to be treated as independent field variables.

By reinstating H as a diffusive part H_0 and a reaction part H_I as in (101) and performing the continuum limit in space, we get the final action:

$$S[\tilde{\varphi}, \varphi] = \int d^d x \left(-\varphi(t_f) - \tilde{\varphi}^*(0)(n_0 - \varphi(0)) + \int_0^{t_f} dt \left(\tilde{\varphi}(\partial_t - D\nabla^2)\varphi - H_I(\tilde{\varphi}, \varphi) \right) \right) \tag{116}$$

Here $H_I(\tilde{\varphi}, \varphi)$ represents the interaction generator $\hat{H}_I$ with $\hat{a}_i^\dagger \to \tilde{\varphi}(\mathbf{x}, t)$ and $\hat{a}_i \to \varphi(\mathbf{x}, t)$. Additionally, the interaction rates $\lambda(\tau)$ appearing in (101) will have to be rescaled as $\lambda(\tau) \to \lambda(\tau)/h^{d(k(\tau)-1)}$ to keep the interaction terms finite as $h \to 0$. The result is an interaction generator of the form

$$H_I(\tilde{\varphi}, \varphi) = \sum_\tau \lambda(\tau) \left[\tilde{\varphi}^{\ell(\tau)} - \tilde{\varphi}^{k(\tau)} \right] \varphi^{k(\tau)} \,. \tag{117}$$

We have now turned our single species reaction network into a field theory for a pair of scalar fields $\tilde{\varphi}, \varphi$. The generalization to multiple species is immediate; each species $k \in S$ gets a scalar field associated to it. The diffusion term is diagonal in the sense that there is one term for each species, each with its own diffusion constant D_k. The interaction terms in H_I may (and generally will) couple different species together. The upshot after this somewhat lengthy derivation, is that we can arrive at a very simple prescription for turning any CRN into a reaction-diffusion field theory. Just replace the creation operators $\hat{a}_i^\dagger$ by $\tilde{\varphi}_i$ fields and the annihilation operators $\hat{a}_i$ by φ_i fields and add a diffusion term for each species!

Many authors use a slightly different formulation of the reaction-diffusion action, one where the $\tilde{\varphi}$ fields are shifted by one: $\tilde{\varphi} \to 1 + \bar{\varphi}$. This field shift traces back to an observation first made in [16]. The flat state can be expressed as $\langle \mathbf{1}| = \langle 0| e^{\sum_i \hat{a}_i}$. It is possible to move the $e^{\sum_i \hat{a}_i}$ to the right in the expectation value, commuting it through any operator it comes across its path. Since $e^{\hat{a}_i} \hat{a}_i^\dagger = (\hat{a}_i^\dagger + 1) e^{\hat{a}_i}$, this has the net effect of shifting all creation operators by one: $\hat{a}_i^\dagger \to \hat{a}_i^\dagger + 1$. In field theory language, this has the stated effect of the field shift $\tilde{\varphi} \to 1 + \bar{\varphi}$, under which the time derivative part gets an additional final and initial time contribution:

$$\int_0^{t_f} dt \tilde{\varphi} \partial_t \varphi = \varphi(t_f) - \varphi(0) + \int_0^{t_f} dt \bar{\varphi} \partial_t \varphi \,. \tag{118}$$

The $\varphi(t_f)$ from this shift now cancels the $-\varphi(t_f)$ present in (116), but it generates another $\varphi(0)$ initial term. This term is then canceled by performing the shift in the initial conditions. The $\bar{\varphi}(0)(n_0 - \varphi(0))$ term remaining in the action (116) has the interpretation that $\bar{\varphi}(0)$ acts as a Lagrange multiplier field, enforcing the constraint $\varphi(0) = n_0$, and hence making sure the initial condition are satisfied.

4.3 Free propagator

With the above prescription, any CRN is mappable to a field theory, describing the reaction process in the diffusion-limited regime. This allows for the importation of field theory methods to the field of reaction-diffusion systems. Most notably, these methods are perturbation theory (computing Feynman diagrams) and the renormalization group flow. We cannot give a completely self-contained overview of these important methods, they are addressed in many quantum field theory textbooks (such as [29,30]), however we will try to opt for a more pedagogical (and hence less deep-diving) presentation here and in the next section.

The first point of discussion is the free (purely diffusive) theory, where there are no reactions present in the system (i.e. all reaction rates are set to zero). The free theory is important, because it provides an elementary building block of the perturbative expansion in terms of Feynman diagram, namely the *propagator*. The propagator corresponds to the pair correlations between the fields $\bar{\varphi}$ and φ and hence it represents mathematically the propagation of a $\bar{\varphi}$ field into a φ field. In terms of diagrams, this is represented as a line (with time running from right to left), where a $\bar{\varphi}$ field is created at the right and propagates into a φ field on the left.

Considering the bulk part of the free action (leaving the subtleties of the terms at $t = 0$ and $t = t_f$ for the end), we can obtain the propagator from the bilinear part of the action, as the Green's function of the diffusion equation:

$$(\partial_t - D\nabla^2_{\mathbf{x}})G(\mathbf{x}-\mathbf{x}', t-t') = \delta^d(\mathbf{x}-\mathbf{x}')\delta(t-t') \tag{119}$$

We can (just like in section 4.1) most easily obtain the propagator in momentum space after Fourier transforming the above equation as:

$$G(\mathbf{k},\omega) = \int d^d x \int dt e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} e^{i\omega(t-t')} G(\mathbf{x}-\mathbf{x}', t-t') \tag{120}$$

The equation (119) is then solved using the inverse Fourier transform and one can readily find that:

$$G(\mathbf{k},\omega) = \frac{1}{-i\omega + Dk^2}\,. \tag{121}$$

This is the momentum space propagator. It turns out to be convenient to invert the temporal Fourier transform and work in terms of $G(\mathbf{k}, t-t') = \int d\omega/(2\pi) G(\mathbf{k},\omega) e^{-i\omega(t-t')}$. In the complex frequency plane, there is a single pole at $\omega = -iDk^2$. Hence, the contour integration only picks up a contribution from this pole if the contour is closed in the lower-half ω plane. This should be done whenever $\mathrm{Re}(-i\omega(t-t')) < 0$ such that the exponent vanishes at infinity. In the lower half plane, $\mathrm{Im}(\omega) < 0$ so that the inverse frequency Fourier transform only gives a non-vanishing result whenever $t - t' > 0$:

$$G(\mathbf{k}, t-t') = e^{-Dk^2(t-t')}\Theta(t-t')\,. \tag{122}$$

Here $\Theta(t-t')$ is the Heaviside step function. Physically, the presence of the Heaviside step function represents causality: only earlier $\bar{\varphi}$ fields are connected to later φ fields. This can be shown by computing the correlation function:

$$\langle \varphi(\mathbf{x},t)\bar{\varphi}(\mathbf{x}',t')\rangle_0 = G(\mathbf{x}-\mathbf{x}', t-t') = \quad \varphi \text{ ——— } \bar{\varphi} \tag{123}$$

Here the subscript 0 on the left hand side indicates that the correlation function is computed in the free theory. The Feynman diagram on the right is the graphical representation of the propagator.

The causality requirement above has two important consequences: any term in the action where φ fields appear without an earlier $\bar{\varphi}$ field will vanish when averaged with the statistical weight e^{-S_0}. Hence, as observables are only expressed in terms of φ fields, the $\bar{\varphi}(0)\varphi(0)$ term in the initial conditions constraint can be safely dropped from the action, since pairing the φ fields in the observable with $\bar{\varphi}(0)$ would leave an unpaired $\varphi(0)$ field in the action. Secondly, it justifies an earlier hidden assumption where we secretly extended the time domain from $[0, t_f]$ to the entire real time axis (incl. negative t) when performing the Fourier transform (120).

With these last thoughts in mind, the final result for the path integral representation of any (normal ordered) observable $\hat{A}$ can be written as:

$$\langle A \rangle_t = \frac{1}{\mathcal{N}} \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi \; A(\varphi) e^{-S[\bar{\varphi},\varphi]} \tag{124}$$

with the action:

$$S[\bar{\varphi},\varphi] = \int d^d x \left\{ \int dt \left[\bar{\varphi}(\partial_t - D\nabla^2)\varphi - H_I(1+\bar{\varphi},\varphi) \right] - \bar{\varphi} n_0 \right\} . \tag{125}$$

The Euler-Lagrange equations for this action give the saddle-point approximation, which corresponds to the rate equations obtained in the mean-field approximation (exercise 4). In the next section, we will use this action to explore perturbative corrections to the saddle point equations and see where and when fluctuation become important. This will lead us to construct Feynman rules to compute contributions from (one-loop) Feynman diagrams and to study renormalization group methods for reaction-diffusion systems.

4.4 Exercises

1. Consider a random walk in one dimension, where the probability of moving to the left is not equal to the probability of moving to the right. Call the rate of moving to the left r_- and the rate of moving to the right r_+

   (a) What is the continuous-time master equation for $P(x,t)$ in this case?

   (b) Derive the continuum limit $h \to 0$ in this case. When do you get an extra term? What could this extra term signify physically?

2. Show that $\mathbb{1} = \int \frac{d^2\eta}{\pi} |\eta\rangle\langle\eta|$ using the identity:

   $$\delta_{n,m} = \frac{1}{\pi m!} \int d^2\eta \; e^{-|\eta|^2} \eta^{*m} \eta^n \, . \tag{126}$$

   Bonus points if you can proof this identity (Hint: decompose the complex number $\eta = re^{i\theta}$ and remember the integral representation of the Gamma function)

3. What are the field theories for the CRN's of question 3.1.1 question 1?

4. Derive the Euler-Lagrange equations from the (time-dependent part of the) general action (116) by computing $\delta L/\delta\varphi$ and $\delta L/\delta\tilde{\varphi}$ (where $S = \int d^d x dt L$) and setting the result to zero. Can you find a solution for $\tilde{\varphi}$ and use this to in the equation for φ? What do you find in the end?

5. Consider the field shift $\tilde{\varphi} = 1 + \bar{\varphi}$ in the action $H_I(\tilde{\varphi}, \varphi)$. For simplicity, consider just one transition with a fixed $\ell < k$.

   (a) What is the interaction term $H_I(\bar{\varphi}, \varphi)$ after the field shift?

   (b) Derive the Euler-Lagrange equations for φ and $\bar{\varphi}$ by varying the action with respect to $\delta\bar{\varphi}$ and $\delta\varphi$ and setting the resulting variation to zero. When using the trivial (constant) solution for $\bar{\varphi}$ what is the remaining equation for φ? Does this look familiar?

   (c) Are the Euler-Lagrange equations for $\varphi(\mathbf{x}, t)$ derived above the same as those derived from the action prior to the field shift?

6. Just as in quantum field theory, the correlation functions in the free theory can be obtained from functional differentiation with respect to a *generating functional* $Z(\bar{J}, J)$. Here we will derive the generating functional and use it to compute the propagator of the free theory.

   The generating functional in this case is defined as:

$$Z(\bar{J}, J) = \frac{1}{\mathcal{N}_0} \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi \exp\left[-S_0 + \int d^d x \int dt \left(\bar{J}(\mathbf{x}, t)\varphi(\mathbf{x}, t) + J(\mathbf{x}, t)\bar{\varphi}(\mathbf{x}, t)\right)\right] \tag{127}$$

   where $\mathcal{N} = \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi e^{-S_0}$ is a normalization factor. The purpose of this exercise is to show that:

$$Z(\bar{J}, J) = \exp\left\{\int d^d x d^d x' dt dt' \bar{J}(\mathbf{x}, t) G(\mathbf{x} - \mathbf{x}', t - t') J(\mathbf{x}', t')\right\} \tag{128}$$

   and use this to compute the free theory correlation functions $\langle\varphi(x, t)\bar{\varphi}(x', t')\rangle_0$ by taking functional derivatives with respect to J and $\bar{J}$.

   First, we define two Green's functions, G and $\bar{G}$ such that:

$$(\partial_t - D\nabla^2) G(\mathbf{x} - \mathbf{x}', t - t') = \delta^d(\mathbf{x} - \mathbf{x}')\delta(t - t') \tag{129}$$

$$(-\partial_t - D\nabla^2) \bar{G}(\mathbf{x} - \mathbf{x}', t - t') = \delta^d(\mathbf{x} - \mathbf{x}')\delta(t - t') \tag{130}$$

   and we recall the short hand notation for the convolution:

$$(f * g)(\mathbf{x}, t) = \int d\mathbf{x}' dt' f(\mathbf{x} - \mathbf{x}', t - t') g(\mathbf{x}', t') \,. \tag{131}$$

   (a) By using the field shifts $\varphi \to \varphi + G * J$ and $\bar{\varphi} \to \bar{\varphi} + \bar{G} * \bar{J}$, show that the generating functional (127) can be expressed as (128)

   (b) Using (128), proof equation (123) by functional differentiation with respect to J and $\bar{J}$ and afterward setting $J, \bar{J} = 0$

   (c) Using (128), compute the four point correlation function $\langle\varphi_1\varphi_2\bar{\varphi}_3\bar{\varphi}_4\rangle_0$ in the free theory (in terms of propagators G), where the subscripts simply indicate that each field is to be taken at different spacetime points $\varphi_1 \equiv \varphi(\mathbf{x}_1, t_1)$. Can you express the answer in terms of Feynman diagrams?

5 Renormalization group methods

The field theory for reaction-diffusion systems constructed in the last section constitutes a microscopic model for the chemical reactions taking place when particles rely on random walks to find each other. We have already seen in the exercises of the last section, that the macroscopic rate equation correspond to the saddle point solutions, which minimize the field theory actions. In this section, we will look at fluctuations and correlations around those saddle point solutions, as perturbative expansions. This will lead to the construction of Feynman diagrams for the reaction-diffusion systems, which leads into the analysis of scaling properties of the theory as we change the (inverse) length scale (or momentum space cutoff) in the model. In view of the length of the course, we cannot treat all cases in painstaking detail, so we will illustrate the general methodology at the hand of single-species annihilating reactions $kA \to \ell A$ with $\ell < k$ [31]. After this, we make some general comments on different kinds of systems which have been analyzed using renormalization group (RG) methods and the universality classes which emerge from those systems.

First, some general remarks are in order. Fluctuations and correlations are expected to become important close to a critical point in the theory. These critical points are characterized by diverging correlation lengths, which enables the (microscopic) fluctuations to influence the (macroscopic) thermodynamics properties of the system. The thermodynamic behavior of systems near the critical point is characterized by scaling functions and critical exponents, which exhibit universal properties. This means that these exponents are insensitive to (some) microscopic details of the system and characterize the behavior of many different systems in the vicinity of their critical point.

The renormalization group is a systematic procedure which can be applied to any field theory (quantum, classical, statistical or stochastic) to understand the impact of fluctuations and identify and compute universal properties in the system. Generically, a momentum-space cutoff κ is introduced to regulate short-distance (ultraviolet: UV) singularities. These singularities are an artifact of the continuum limit taken above; because we have taken the lattice spacing to go to zero, we essentially require particles to be infinitesimally close to each other to react. The introduction of a momentum scale κ makes the parameters of the theory (i.e. the reaction rates) κ dependent. One can then define renormalized effective parameters, which absorb the short-distance singularities. This allows one to study how the parameters depend on the momentum scale of the theory. Scale invariance is achieved when the effective (renormalized) parameters do not change with decreasing momentum scale and at this point the UV scaling properties translate to the long distance infrared (IR) regime.

An important concept in this discussion is the *upper critical dimension* of the system. We have already encountered an example of a critical dimension when discussing the random walk. In dimensions $d > 2$ the properties of walker changed from being able to explore all of space to not being able to. In the reaction-diffusion systems, there is also a critical dimension d_c. For $d > d_c$ the mean-field theory analysis is asymptotically valid *without approximation*. That means the perturbative expansion describing fluctuations in the system is benign when $t \to \infty$, or the late-time scaling properties of the system are not affected by loop corrections. For $d \leq d_c$, IR divergences appear which are more serious than the UV divergences stemming from the lattice spacing. These imply that at $d = d_c$ fluctuations have an important effect on the long-distance behavior of the system, and hence we should expect deviations from

mean-field theory results. The goal of RG is to extract the correct power laws associated to those IR singularities in, for instance, the particle density and other correlation functions. This can be done by relying on an underlying scale invariance in the system, as we will try to illustrate here for the example of the $kA \to \ell A$ annihilating reactions.

5.1 The perturbative expansion

As mentioned above, we will focus this subsection on the perturbative expansion and renormalization of the field theory for $kA \to \ell A$ diffusion limited reactions (see also [31]). This is to illustrate the more general procedure in light of an example applicable to more general reaction-diffusion systems and where things work out quite nicely. More details on generic field theory methods can be found in many textbooks, such as [29, 30].

The $kA \to \ell A$ field theory action, after the shift $\tilde{\varphi} = 1 + \bar{\varphi}$ is (see section 4.4 exercise 5):

$$S[\bar{\varphi}, \varphi] = \int d^d x \left\{ \int_0^{t_f} dt \ \left[\bar{\varphi}(\partial_t - D\nabla^2)\varphi + \sum_{i=1}^{k} \lambda_i \bar{\varphi}^i \varphi^k \right] - n_0 \bar{\varphi}(0) \right\} . \tag{132}$$

The coupling constants λ_i are related to the original reaction rate λ of the $kA \to \ell A$ transition as:

$$\lambda_i = \lambda \left[\binom{k}{i} - \binom{\ell}{i} \right] , \qquad i \leq \ell \tag{133a}$$

$$\lambda_i = \lambda \binom{k}{i} , \qquad i > \ell \, . \tag{133b}$$

In the computation of any observable, we treat the free (diffusion) part of the action non-perturbatively, but expand the reaction part. For instance, if our observable is the particle density, then $A = \varphi$ (recall the number operator is $\hat{a}_i^\dagger \hat{a}_i$ and we could set $\hat{a}_i^\dagger = 1$ in the expectation value!), then the path integral may be expanded in three parts as:

$$\begin{aligned} \langle \varphi(\mathbf{x}, t) \rangle &= \frac{1}{\mathcal{N}} \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi \ \varphi(\mathbf{x}, t) e^{-S_0} e^{-S_I(\{\lambda_i\})} e^{\int d^d y \ n_0 \bar{\varphi}(\mathbf{y}, 0)} \qquad (134) \\ &= \frac{1}{\mathcal{N}} \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi \ \varphi(\mathbf{x}, t) e^{-S_0} \left(1 - \int d^d y \int_0^t dt' \ \sum_{i=1}^{k} \lambda_i \bar{\varphi}(\mathbf{y}, t')^i \varphi(\mathbf{y}, t')^k + \dots \right) e^{\int d^d y \ n_0 \bar{\varphi}(\mathbf{y}, 0)} \end{aligned}$$

The first exponent is the free, diffusive action, which we will use as statistical weight for the computation of the expectation value. The second term represents the reaction interaction, which we treat perturbatively. The last term is the implementation of the initial conditions, which effectively generates $\bar{\varphi}$ fields at $t = 0$.

The effect of computing everything with S_0 as statistical weight means that we can use the propagator (123) to connect φ fields to earlier $\bar{\varphi}$ fields. We'll only get contributions from terms where pairs can be made (recall exercise 6 of section 4.4) with equal number of φ and $\bar{\varphi}$ fields, and the propagator also imposes causality: later $\bar{\varphi}$ fields contracted with earlier φ fields give vanishing contributions.

Each term in the expansion will now correspond to a specific *Feynman diagram*, which is a graphical representation of the contributions to the path integral. The rules for constructing

these diagrams and mapping them to a formula are generically called *Feynman rules*. Let's do the first few explicitly. A first contribution comes from no reactions:

$$\frac{1}{\mathcal{N}}\int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi\; e^{-S_0}\varphi(\mathbf{x},t)e^{\int d^dy\; n_0\bar{\varphi}(\mathbf{y},0)} = \frac{1}{\mathcal{N}}\int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi\; e^{-S_0}\varphi(\mathbf{x},t)\int d^dy\; n_0\bar{\varphi}(\mathbf{y},0)\,. \quad (135)$$

Here we have used that in the expansion of the initial conditions, only the linear term will contribute, since all other terms cannot be paired with φ fields. The result is simply given by the propagator

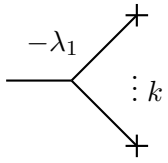
$$\int d^dy\; n_0 G(\mathbf{x}-\mathbf{y},t) = \quad \longrightarrow\!\!\times \quad (136)$$

The above integral simply evaluates to unity by Gaussian integration of the propagator. The cross represents the $\bar{\varphi}$ field inserted at $t=0$, which subsequently propagates to t. Physically it represents the contribution of the initial conditions (having a uniform field with density n_0) and no reactions taking place.

More interesting is when reaction takes place. The only terms $\sim \lambda_i\bar{\varphi}^i\varphi^k$ contributing to the density have $i=1$, as only one $\bar{\varphi}$ can be paired with the later φ field. Working this out explicitly, we get:

$$\frac{1}{\mathcal{N}}\int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi\; e^{-S_0}\varphi(\mathbf{x},t)\int d^dy\int_0^{t_f} dt\;(-\lambda_1)\bar{\varphi}(\mathbf{y},t)\varphi(\mathbf{y},t)^k e^{\int d^dy'\; n_0\bar{\varphi}(\mathbf{y}',0)} \quad (137)$$

$$= \int d^dy\int_0^t dt'\; G(\mathbf{x}-\mathbf{y},t-t')(-\lambda_1)\left(\int d^dy' G(\mathbf{y}-\mathbf{y}',t')\right)^k n_0^k\,. \quad (138)$$

$=$ 

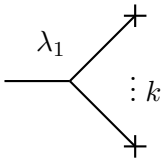
(139)

Here, to pair the k φ fields with the initial $\bar{\varphi}$ fields, we had to expand the exponential $e^{\int d^dy'\; n_0\bar{\varphi}(\mathbf{y}',0)}$ to order k. This gives a factor $\frac{1}{k!}$, which is conveniently cancelled by the $k!$ ways to pair the k φ fields with a barred counterpart at $t=0$.

It is often practically easier to compute the Feynman diagrams in the momentum-time representation, by Fourier transforming over the spatial variables. Then, the integral $\int d^dy$ will give a delta function over all incoming and outgoing momenta at a vertex. This represents physically momentum conservation at the vertex. The initial conditions are also simpler: $\int d^dy'\; n_0\bar{\varphi}(\mathbf{y}',0) = n_0\bar{\varphi}(\mathbf{k}=0,0)$, or the integration over space just gives the zero momentum mode. This makes sense as the initial conditions are uniform, so they should not depend on momentum. Moreover, propagators also conserve incoming and outgoing momentum:

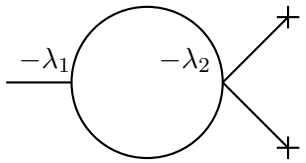
$$\langle\varphi(\mathbf{k},t)\bar{\varphi}(\mathbf{k}',t')\rangle = (2\pi)^d\delta(\mathbf{k}+\mathbf{k}')G(\mathbf{k},t-t') \quad (140)$$

with $G(\mathbf{k},t-t')$ given in (122). This makes the evaluation of the above diagram simply:



$= \int_0^t dt'\; G(0,t-t')(-\lambda_1)G(0,t')^k n_0^k = -\lambda_1 n_0^k t\,.$ (141)

More interesting still, are the diagrams with more vertices, which allows the possibility to create *loops* of particles with undetermined momentum. For instance, for $k = 2$, we get at second order in the couplings the first one-loop diagram:



(142)

Working this out carefully gives:

$$I_{\text{loop}} = \int_0^t dt'' \int_0^{t''} dt' G(0, t - t'')(-\lambda_1) \int \frac{d^d p}{(2\pi)^d} 2G(\mathbf{p}, t'' - t')G(-\mathbf{p}, t'' - t') \, (-\lambda_2)G(0, t')^2 n_0^2 \,, \tag{143}$$

The factor of 2 comes from the two different ways of attaching the propagators inside the loop, and hence constitutes a diagram dependent symmetry factor.

In general, we can define a set of *Feynman rules*, which help to translate any Feynman diagram to an integral. These are:

- For each propagator connecting an earlier $\bar{\varphi}(\mathbf{p}', t_1)$ field to a $\varphi(\mathbf{p}, t_2)$ field, multiply by the propagator $(2\pi)^d \delta^d(\mathbf{p} + \mathbf{p}')G(\mathbf{p}, t_2 - t_1)$.
- For each vertex with k incoming fields (coming from the right) and i outgoing fields (going to the left), multiply by $-\lambda_i$
- For each external end point (labeled by $\times$), multiply by n_0
- Impose momentum conservation at each vertex: multiply by $(2\pi)^d \delta(\sum_i \mathbf{p}_i)$ where i labels all momenta connected to that vertex
- Integrate over all momenta $\int \prod_i \frac{d^d p_i}{(2\pi)^d}$ and over all time points of the inserted vertices.
- Multiply by the symmetry factor (number of ways to attach propagators without changing the diagram)[7]

Following these Feynman rules, any diagram can be turned into an integral, which due to the properties of the propagators is readily computable in terms of Gaussian integrals. For instance, the first loop diagram (143) evaluates to:

$$I_{\text{loop}} = \frac{8\lambda_1 \lambda_2 n_0^2}{(8\pi D)^{d/2}} \frac{t^{2-d/2}}{(2-d)(4-d)} \,. \tag{144}$$

Compared to the *tree-level* (which means: zero loops) diagram computed in (141) with $k = 2$, we see that the one-loop contribution adds a term to the effective coupling of these contributions to the density which scales as $(\lambda_2/D^{d/2})t^{1-d/2}$. This contribution diverges as $t \to \infty$ when $d < d_c = 2$ (while for $d > 2$ it diverges as $t \to 0$). At $d = d_c$, the effective coupling goes

[7] This is a slightly different convention as is usual in standard QFT (see for instance [29]), where the convention is to divide the coupling term constants of order k by $1/k!$ and then one should *divide* by the symmetry factor instead of mulitiply.

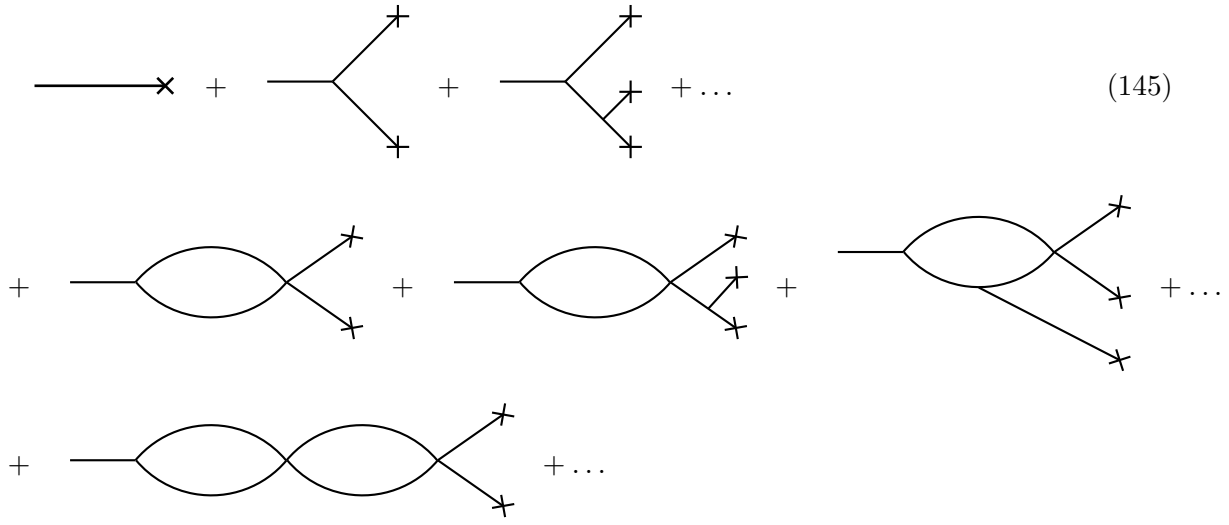
as $(\lambda_2/D)\ln(Dt)$, which diverges both at small and large t's; or we say it diverges both in the IR and the UV. The UV (short times, short distances) singularities can be cured by introducing a small distance cut-off, which is somehow a reasonable thing to do, as we originally had a lattice spacing, or particles usually have some finite interaction radius, setting a smallest distance scale in the theory. The IR divergences are more serious. This implies that in the late time regime, at or below the critical dimension of $d_c = 2$, the loop contributions start to become increasingly important and hence the naive perturbation theory in the λ values don't immediately make sense.

Conversely, for $d > d_c = 2$, the perturbative corrections from loop diagrams become less important in the asymptotic limit as $t \to \infty$. Hence, in this regime, perturbation theory is applicable and the overall scaling behavior of parameters in the theory are not affected by loop corrections. This is the reason for the earlier statement that mean-field theory is asymptotically exact above the critical dimension.

5.1.1 Tree level diagrams and the rate equations

Having specified the Feynman rules, we can now organize any perturbative expansion in terms of Feynman diagrams, starting with the fundamental building blocks: the propagators and the vertices. The latter we can read off immediately from the field theory action, a term $\sim \bar{\varphi}^i \varphi^k$ gives a vertex with k legs incoming from the right and i legs outgoing from the left.

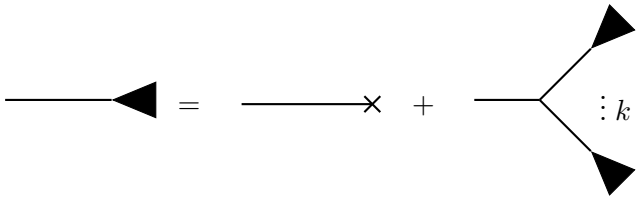
In the $kA \to \ell A$ reaction (with $k \geq 2$ and $\ell < k$) this immediately implies that the propagator, with only one incoming and one outgoing leg, cannot receive any loop corrections: there are no vertices with one incoming leg and multiple outgoing legs! So, the free theory propagator is exact to all orders in perturbation theory. The general expansion for the density, (with one outgoing φ field) can then generally be organized into an expansion with increasing number of loops. For instance, for $k = 2$:



(145)

The first line represents all the tree diagram, the second line those with one loop, the third line with two loops, etc. The contributions of the tree diagrams can be resumed by realizing the following. If we represent the sum over all tree diagrams as a triangle vertex (and call its

contribution $a_{\text{tree}}(t)$), then the following relation holds:



(146)

or, expressed in formula:

$$a_{\text{tree}}(t) = n_0 + \lambda_1 \int_0^t dt' \, (a_{\text{tree}}(t'))^k \tag{147}$$

Taking the time-derivative on both sides and using that $\lambda_1 = (k-\ell)\lambda$ in terms of the original reaction rate (see equation (133)), we obtain:

$$\frac{d}{dt} a_{\text{tree}}(t) = \lambda(k-\ell) a_{\text{tree}}(t)^k \,, \tag{148}$$

with initial conditions $a_{\text{tree}} = n_0$. This is just the rate equation for the density of A particles! Hence we have shown that the tree level diagrams give the mean-field result. The loop diagrams are thus responsible for corrections to the mean-field result, which become increasingly important when increasing the length and time scales (or for decreasing momentum/frequency scales) at or below the critical dimension.

5.1.2 Vertex functions

We have already seen an example where loop diagrams give effective corrections to a vertex in the section above. The idea of the renormalization group is to identify how physical observables, such as densities, fluctuations or other correlation functions, scale with time, momentum or length scale transformations. To this end, we should first identify the UV singularities, stemming from the short distance (large wave number) contributions to the loop integrals. We can then absorb these UV singularities into a renormalized coupling which will depend explicitly on a *normalization momentum scale* κ. Associated to this, there is a typical length scale κ^{-1} and a typical time scale $t_0 = 1/(D\kappa^2)$.

The first step is to identify primitive UV divergences as $\kappa \to \infty$; or which components of the diagrams are responsible for the divergences. To figure this out, we can perform a power-counting argument on the vertices. The action has to be dimensionless and as stated above:

$$[x] = \kappa^{-1} \,, \qquad [t] = \kappa^{-2} \,. \tag{149}$$

In the continuum limit of section 4, we have used the scalings (115), which imply that:

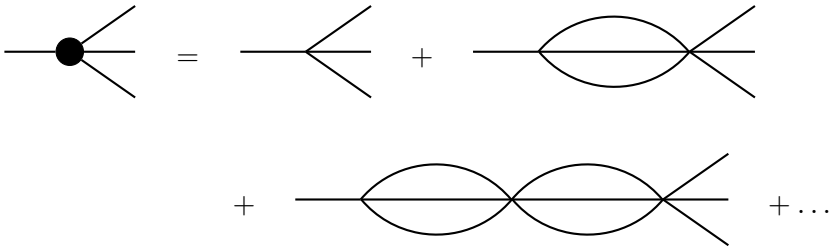
$$[\bar{\varphi}] = \kappa^0 \,, \qquad [\varphi] = \kappa^d \,. \tag{150}$$

All this taken together imply that the diffusive part of the action $\int d^dx \int dt \bar{\varphi}(\partial_t - D\nabla^2)\varphi$ is dimensionless, as it should be. In order for the interaction terms to be dimensionless, the vertices $\lambda_i \bar{\varphi}^i \varphi^k$ impose a scaling of the coupling constant as:

$$[\lambda_i] = \kappa^{2-(k-1)d} \tag{151}$$

Now, a loop correction to the vertex $\lambda_i \bar{\varphi}^i \varphi^k$ would generate a contribution proportional to $\lambda_i \lambda_k$. In order for this contribution to scale as the tree level contribution λ_i, the momentum integral over the loop must be of dimension $[\lambda_k]^{-1} = \kappa^{(k-1)d-2}$. Hence, if $(k-1)d - 2 > 0$, or if $d \geq d_c = 2/(k-1)$ the loops will give a contribution which diverges at $\kappa \to \infty$ and the vertex should be renormalized. Conversely, if $d < d_c$, IR singularities appear as $\kappa \to 0$. At the critical dimension $d = d_c$ the loop diagrams carry logarithmic UV and IR divergences.

Using the primitive divergencies, we can renormalize the coupling constants by absorbing the UV singularities. For the reaction $kA \to \ell A$ the procedure is actually quite simple, as the whole perturbative series can be resummed. As noted above, the propagator does not receive any loop corrections, hence this leave just the vertex functions $\Gamma^{(m,k)}$, representing all diagrams with k incoming and m outgoing legs. Without considering the propagators for the eternal lines, the vertex function represents the sum over all possible loop diagrams with the specified external lines. Graphically, for $m = 1$ and $k = 3$, the only contributions are:



(152)

Using the Feynman rules from the last section (only for the vertex part, not the external propagators), the explicit expression for this expansion is:

$$\Gamma^{(m,k)}(t_2 - t_1) = \lambda_m \delta(t_2 - t_1) - \lambda_m \lambda_k I(t_2 - t_1) + \lambda_m \lambda_k^2 \int_{t_1}^{t_2} dt I(t_2 - t') I(t' - t_1) - \dots \quad (153)$$

Where the integral is over the k propagators in the loop:

$$I(t) = k! \int \prod_{i=1}^{k} \left(\frac{d^d p_i}{(2\pi)^d} \right) (2\pi)^d \delta^d \left(\sum_{i=1}^{k} \mathbf{p}_i \right) e^{-\sum_{i=1}^{k} D p_i^2 t} \quad (154)$$

The above integral is most easiest evaluated by expressing the delta function in integral form $(2\pi)^d \delta^d \left(\sum_{i=1}^{k} \mathbf{p}_i \right) = \int d^d x \exp(i \sum_i \mathbf{x} \cdot \mathbf{p}_i)$ and then performing the k Gaussian integrals over the momenta, with a linear contribution coming from the delta function. The result is another Gaussian integral over $\mathbf{x}$ after which:

$$I(t) = B_k (Dt)^{-d/d_c} . \qquad \text{with: } B_k = k! k^{-d/2} (4\pi)^{-d/d_c} . \quad (155)$$

Here $d_c = 2/(k-1)$ as stated above. The higher-loop contributions are expressed in terms of convolutions of the single loop integral. This is most conveniently addressed using the Laplace transform:

$$\tilde{\Gamma}^{(m,k)}(s) = \int_0^\infty \Gamma^{(m,k)}(t) e^{-st} dt . \quad (156)$$

Laplace transforms have the useful property that convolutions of functions of t turn into products of the Laplace transformed functions. The infinite series (153) then turns into the geometric series for the Laplace transform of $I(t)$:

$$\tilde{\Gamma}^{(m,k)}(s) = \frac{\lambda_m}{1 + \lambda_k \tilde{I}(s)} , \qquad \tilde{I}(s) = B_k \Gamma(\epsilon/d_c) D^{-d/d_c} s^{-\epsilon/d_c} . \quad (157)$$

Here $\epsilon = d_c - d$ and $\Gamma(x) = \int_0^\infty dt\, t^{x-1}e^{-t}dt$ is the Gamma function.

Since in the original theory, there is only one coupling constant λ, the rate of the reaction process, we have that $\lambda_m \sim \lambda$ and $\lambda_k = \lambda$. We can now express everything in terms of a dimensionless *bare* coupling constant $g_0 = (\lambda/D)\kappa^{-2\epsilon/d_c}$ (recall (151)), where κ is the arbitrary momentum scale. The UV divergencies, which are now manifested by the pole of the Gamma function at $\epsilon \to 0$, are absorbed into a *renormalized* coupling constant g_R defined as:

$$g_R = \tilde{\Gamma}^{(k,k)}(s = D\kappa^2)\kappa^{-2\epsilon/d_c}/D = Z_g g_0\,, \qquad Z_g^{-1} = 1 + g_0 B_k \Gamma(\epsilon/d_c)\,. \tag{158}$$

Here the Laplace argument is set at the typical inverse time scale $s = 1/t_0 = D\kappa^2$, which defines the normalization point [31]. This renormalized coupling tells us how the effective interaction strength (in our case, reaction rates) change as a function of the momentum scales at which we probe the theory. It is a function of the normalization scale κ and its associated *β-function* shows how it changes with momentum:

$$\beta_g(g_R) = \kappa\frac{d}{d\kappa}g_R = 2g_R\left[-\frac{\epsilon}{d_c} + B_k\Gamma\left(1+\frac{\epsilon}{d_c}\right)g_R\right] \tag{159}$$

This β function, computed to all orders in perturbation theory, is finite for $\epsilon \to 0$ when expressed in terms of renormalized quantities. When it vanishes, the theory is manifestly scale invariant, which happens when $g_R = 0$ (the trivial fixed point, corresponding to pure diffusion with no reactions), or at the non-trivial value, when $g_R = g_R^*$ with:

$$g_R^* = [B_k\Gamma(\epsilon/d_c)]^{-1} \tag{160}$$

This is of order ϵ. Using these relations, any perturbative expansion in the bare coupling g_0 can be exchanged for one in the renormalized coupling, as:

$$g_0 = \frac{g_R}{1 - g_R/g_R^*} = g_R + \frac{g_R^2}{g_R^*} + \dots \tag{161}$$

Effectively, this equation can then be used to turn the perturbative expansion for observables into an expansion in ϵ, which is the small parameter of the theory. This is done through the computation of *Callan-Symanzik* equations, which determine how the density, density correlations or any other correlation function scales with the momentum scale κ. The details for this procedure are found in [24, 31], we unfortunately don't have the time to treat this here. However,the computation goes according to the following rough lines. First, one determines the Callan-Symanzik equation of the observable of interest, by requiring that the bare observable is independent on the renormalization scale κ and so $\kappa\frac{d}{d\kappa}O(t, n_0, D, \lambda) = 0$. This defines a partial differential equation for the renormalized observable $O(t, n_0, D, \kappa, g_R)$, whose scaling behaviour can be determined via the method of characteristics. This gives characteristic equations for the renormalized coupling g_R and densities n_0, which determine how the effective couplings change with time. This *running* of the coupling constant can in this case be solved exactly, showing that at late times, the theory *flows* to the non-trivial fixed point g_R^* and hence becomes scale-invariant when $t \to \infty$.

To solve the Callan-Symanzik (CS) equations, one still needs a known value for the observable for some value of the running parameters, which one can compute perturbatively in g_0. The CS equations and (161) then allow one to write the perturbative expansion for g_0

in terms of an expansion in terms of ϵ. Moreover, as the renormalized running coupling flows to the fixed point g_R^*, universal behavior emerges in the asymptotic regime, where for instance the density scales as $n \sim A_k t^{-d/2}$ for $d < d_c$ and as $A_k(\ln t/t)^{1/(k-1)}$ when $d = d_c$ [31]. These scaling exponent don't depend on the value of ℓ, nor on initial conditions and hence are universal.

5.2 Universality

We close this section, as well as these lecture notes, with some remarks on universality in the diffusion-reaction systems in more generality. We can distinguish at least two different kinds of universality, specific for two different kinds of reaction-diffusion systems. First, we may have relaxational models, where the process eventually reaches the absorbing state where there are no more reactants. These models include the single-species $kA \to \ell A$ with $\ell < k$ systems discussed above. Universal features in these models are the decay rates of densities or correlations, which show scaling exponents in t. Secondly, there are universal properties of systems with *absorbing state* phase transitions, where there is a competition between an annihilation and/or a coagulation process and the branching process. Examples include the SIS model, or single species birth-death process (perhaps supplemented with $A + A \to A$ reactions). Here, universality in scaling exponents may be found near the absorbing state transition, such as the exponent of the density or the correlation length.

Above we had already shown that the relaxational models have a non-trivial fixed point which determines the scaling exponent and amplitude at late times. The scaling exponents of the density are universal, in the sense that they do not depend on the microscopic details of the process such as the underlying lattice, initial densities or even the specific reaction types (i.e whether it is $A + A \to A$ or $A + A \to 0$). Furthermore, density exponents have been shown to be the same even for multi-species generalizations, such as $A + B \to 0$ reactions.

The universality of relaxational models driven by diffusion (below $d < d_c$) has an intuitive explanation. As we have seen in Section 4.1, the large-distance properties of the random walk are also universal, depending only on macroscopic properties as the diffusion constant. Relaxational reactions which are diffusion-limited ultimately rely on random walks for particles to find each other. For $d \leq d_c = 2$, any pair of walkers will find be able to find each other. Therefore, at late times, the effective reaction rate is governed by the probability of diffusive particles to propagate form large to vanishing separations. As this process doesn't depend on the species, the asymptotic density power-law decays also doesn't depend on the species. The number of particles which are required to meet does influence the critical dimension (as we have seen above, $d_c = 2/(k-1)$). This also has an intuitive explanation in terms of diffusion; for 3 particles to meet each other in continuous space with certainty, they must be constrained to move in one dimension. Above the critical dimension, two particles will actually not meet in the continuous limit and so the effective reaction rates will depend on the short-distance regulators which are chosen, such as the lattice or a finite reaction range.

Universal behavior also occurs near continuous dynamical phase transitions separating an active state (with non-zero density as $t \to \infty$) from an absorbing (inactive) state (see [32] for a comprehensive review). Generically, such phase transition are within the universality class of *directed percolation* (DP), although there are some exceptions. In directed bond percolation, random bonds are placed with probability p in a regular lattice, which can be

traversed in only one direction. When increasing p, a percolating cluster appears and the size and correlation lengths of these cluster satisfies scales with power-laws near the percolation threshold. The power law exponents are universal; they are found to be the same in many reaction-diffusion systems with an absorbing state transition. The critical dimension for the DP universality class is $d_c = 4$. Field theory methods and the renormalization group has been applied to compute the critical exponents in an expansion in $\epsilon = d_c - d$ in [33]. For many different reaction networks with absorbing states, the *effective* field theory is the same (it is the *Reggeon field theory*), and this explains the universal behavior. Under the RG group, irrelevant higher-order coupling constants will not contribute to the long-distance behavior of the theory and the effective action becomes equivalent for many different microscopic models.

5.3 Exercises

The final exercises are leading up to a one-loop calculation for the SIS model, which is not easy to do, so let's take it step by step. First, we have to know how to deal with the field theory when we have single-species reactants (i.e. $kA \to \ell A$ with $k = 1$). In this case, the propagator gets extra contributions from terms bilinear in the fields. In physics jargon these are mass terms, and they alter the propagator of the free, reactionless theory.

1. First consider a single species A subject to a death process: $A \to \emptyset$. We have already solved the homogeneously mixed (space independent) part in earlier exercises and since each reaction is independent of the presence of other A particles, also the field theory should be simple.

   (a) What is the field theory action for this reaction? Write down the action with reaction rate λ in terms of $\bar{\varphi}$-fields

   (b) What is the propagator for this action? Hint: Fourier transform the fields. The propagator in momentum space is the inverse of the bilinear term

   (c) Compute the expected density of φ as $\langle \varphi(\mathbf{k} = 0, t) \rangle$ in this theory. Do you get any loop corrections?

2. Consider now two species A and B, where A particles can decay into B particles with rate λ: $A \to B$.

   (a) What is the field theory action in this case? Write down the action in terms of $\bar{\varphi}$-fields

   (b) Find a linear combination of A and B fields such that all terms in the action are diagonal, i.e. such that the integrand of the action is written in the form $\bar{\varphi}_1 G_1^{-1} \varphi_1 + \bar{\varphi}_2 G_2^{-1} \varphi_2$. What are the propagators for the two fields?

   (c) Compute the density of A and B particles from the path integral. Do you get loop corrections?

3. Now we consider a more interesting example, the SIS model, which was given by the reactions $I \to S$ and $I + S \to 2I$.

   (a) Write down the field theory action in terms of $\bar{\varphi}$ fields.

(b) Similar to the $A \to B$ reaction above, the bilinear terms in the action are not diagonal. Perform the transformation which diagonalizes the bilinear fields and compute the propagators.

(c) Perform the same field redefinition in the interaction terms. Which terms are present in the action? What vertices do they correspond to?

(d) Do you expect loop corrections to the propagator? In other words: is it possible to construct loop diagram with only one incoming and one outgoing leg?

(e) What are the first loop diagrams you can write down? Can you compute their contributions? When does their contribution diverge?

6 Concluding remarks

In these lectures, we have come a long way and discussed emergence in general and some emergent features in the context of stochastic models of chemical reaction networks (CRN). We have now developed theoretical and mathematical methods for describing CRN at three different levels, first the rate equations, then the master equation for homogeneously mixed solutions and then the field theory description for diffusion-limited reactions. We can think of these three levels of descriptions as being linked to each other, where the mathematical theory of the former descriptive level *emerges* as a limit or mean-field approximation of the latter levels.

Along the way, we have also encountered other emergent properties, such as the appearance of universality in reaction-diffusion systems. One way to understand this universality was by employing the renormalization group methods to study the asymptotic properties of the systems. By following the renormalization group flow (i.e. by studying the effective couplings of the theory as one increases the length scales), scaling properties of observable quantities in the theory can be uncovered which do not depend on microscopic details of the system when close to an RG fixed point. Hence, we can say that the appearance of universal scaling exponents is an emergent feature under the RG flow; it is novel and robust with respect to the microscopic details of the theory.

There is, of course, a lot more which can be said and done along the lines of this lecture series. Field theory and RG methods are starting to now make their way into chemistry, population dynamics, epidemiological modeling and even neuronal dynamics and machine learning. There is a large literature on the application of field theory methods to diffusion-limited systems, but the general import of quantum techniques into the field of complex systems and the continuation and generalization of this research line is so far only limited by the imagination of the researchers.

References

[1] S. Gibb, R. F. Hendry, and T. Lancaster, *The Routledge handbook of emergence*. Routledge, 2019.

[2] R. C. Bishop, *The physics of emergence*. Morgan & Claypool Publishers, 2019.

[3] P. W. Anderson, "More is different: Broken symmetry and the nature of the hierarchical structure of science.," *Science* **177** (1972), no. 4047, 393–396.

[4] C. D. Broad, *The Mind and its Place in Nature*. Routledge, 1925.

[5] T. O'Connor, "Emergent Properties," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, ed. Metaphysics Research Lab, Stanford University, Winter 2021 ed., 2021.

[6] P. A. M. Dirac, "Quantum mechanics of many-electron systems," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **123** (1929), no. 792, 714–733.

[7] E. Schrodinger, *What is life? The physical aspect of the living cell*. Cambridge University Press, 1944.

[8] R. B. Laughlin and D. Pines, "The theory of everything," *Proceedings of the national academy of sciences* **97** (2000), no. 1, 28–31.

[9] D. J. Chalmers, *Strong and Weak Emergence*. Oxford University Press, 2006.

[10] S. de Haro, "Towards a theory of emergence for the physical sciences," *European Journal for Philosophy of Science* **9** (2019), no. 38,.

[11] F. E. Rosas, P. A. Mediano, H. J. Jensen, A. K. Seth, A. B. Barrett, R. L. Carhart-Harris, and D. Bor, "Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data," *PLoS computational biology* **16** (2020), no. 12, e1008289.

[12] J. Baez and J. D. Biamonte, *Quantum Techniques in Stochastic Mechanics*. WORLD SCIENTIFIC, May, 2017.

[13] M. Doi, "Second quantization representation for classical many-particle system," *Journal of Physics A: Mathematical and General* **9** (1976), no. 9, 1465.

[14] K. Narayanan and B. Lakshmikutty, *Stoichiometry and process calculations*. PHI Learning Pvt. Ltd., 2016.

[15] J. Ancheyta, *Chemical reaction kinetics: concepts, methods and case studies*. John Wiley & Sons, 2017.

[16] P. Grassberger and M. Scheunert, "Fock-space methods for identical classical objects," *Fortschritte der Physik* **28** (1980), no. 10, 547–578.

[17] L. Peliti, "Path integral approach to birth-death processes on a lattice," *Journal de Physique* **46** (1985), no. 9, 1469–1483.

[18] N. G. Van Kampen, "The expansion of the master equation," *Advances in chemical physics* **34** (1976) 245–309.

[19] J. C. Baez and B. Fong, "A noether theorem for markov processes," *Journal of Mathematical Physics* **54** (2013), no. 1,.

[20] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, "Dynamical first-order phase transition in kinetically constrained models of glasses," *Physical review letters* **98** (2007), no. 19, 195702.

[21] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, "First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories," *Journal of Physics A: Mathematical and Theoretical* **42** (2009), no. 7, 075007.

[22] H. Touchette, "The large deviation approach to statistical mechanics," *Physics Reports* **478** (2009), no. 1-3, 1–69.

[23] N. Goldenfeld, "Kinetics of a model for nucleation-controlled polymer crystal growth," *Journal of Physics A: Mathematical and General* **17** (1984), no. 14, 2807.

[24] U. C. Täuber, M. Howard, and B. P. Vollmayr-Lee, "Applications of field-theoretic renormalization group methods to reaction–diffusion problems," *Journal of Physics A: Mathematical and General* **38** (2005), no. 17, R79.

[25] J. Cardy, "Renormalisation group approach to reaction-diffusion problems," *arXiv preprint cond-mat/9607163* (1996).

[26] J. L. Cardy and U. C. Täuber, "Field theory of branching and annihilating random walks," *Journal of statistical physics* **90** (1998) 1–56.

[27] P. Krapivsky, S. Redner, and E. Ben Naim, *A Kinetic View of Statistical Physics*. Cambridge University Press, 2010.

[28] B. D. Hughes, *Random walks and random environments: random walks*, vol. 1. Oxford University Press, 1995.

[29] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*. Westview press, 1995.

[30] J. Zinn-Justin, *Quantum field theory and critical phenomena*, vol. 171. Oxford university press, 2021.

[31] B. P. Lee, "Renormalization group calculation for the reaction ka to oe," *Journal of Physics A: Mathematical and General* **27** (1994), no. 8, 2633.

[32] H. Hinrichsen, "Non-equilibrium critical phenomena and phase transitions into absorbing states," *Advances in physics* **49** (2000), no. 7, 815–958.

[33] H.-K. Janssen and U. C. Täuber, "The field theory approach to percolation processes," *Annals of Physics* **315** (2005), no. 1, 147–192.

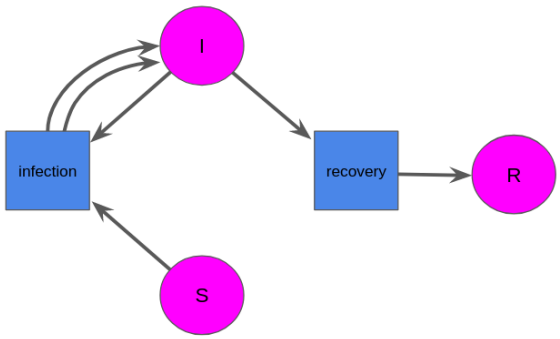
A Exercise solutions

3.1.1 Solutions

1. *Draw the Perti nets for the following chemical reaction networks:*

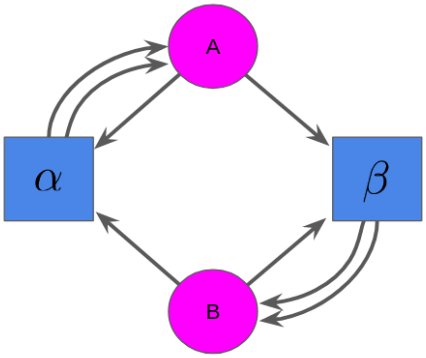
(a)

$$\mathrm{S} + \mathrm{I} \to 2\mathrm{I}\,, \tag{162a}$$
$$\mathrm{I} \to \mathrm{R}\,. \tag{162b}$$



(b)

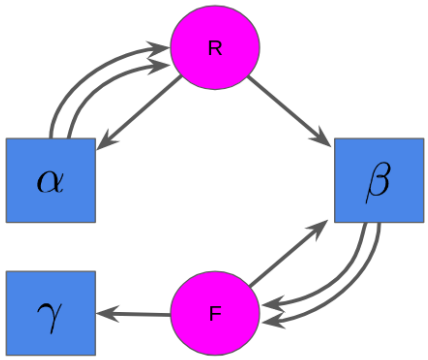
$$\mathrm{A} + \mathrm{B} \to 2\mathrm{A}\,, \tag{163a}$$
$$\mathrm{A} + \mathrm{B} \to 2\mathrm{B}\,. \tag{163b}$$



(c)

$$\mathrm{R} \to 2\mathrm{R}\,, \tag{164a}$$
$$\mathrm{R} + \mathrm{F} \to 2\mathrm{F}\,, \tag{164b}$$
$$\mathrm{F} \to \emptyset \tag{164c}$$

2. *For each picture above, think of a scenario (for instance in population dynamics, epidemiology, or sociology) which is modeled by that Perti net.*



(a) This is the SIR model of epidemiology. It models the spread of an infectious disease which gives immunity to individuals who recover. This is modeled by the transmission reaction of the SIS model and the recovery process introduces a new species R which models the recovered individuals. these can no longer be infected.

(b) This is an opinion model, sometimes called the voter model. It models the situation where individuals can hold opinions A and B. Through interaction between two individuals with differing opinions, either one (A) or the other (B) may change their opinion to match that of the opponent. In the end, the dynamical process will end when everyone has the same opinion.

(c) This is a simplified predator-prey model, where R stand for the preys (rabbits, for instance) and F stands for predators (foxes). t/he first reaction is a birth process, in which the rabbits may multiply. The second reaction is the predation interaction, where a fox eats a rabbit and as a consequence may reproduce (or in terms of ecologists: some biomass will transition from being rabbit to being a fox.). The last term is the death process for foxes. It is implicitly assumed that rabbits have access to infinite food supplies, so they may reproduce without limits and they only die due to predation.

3. *For each of the Perti nets in question 1, derive the rate equations.*

This is in principle a straight-forward application of equation (5). It gives:

(a)

$$\frac{dx_{\mathrm{I}}}{dt} = \alpha x_{\mathrm{I}} x_{\mathrm{S}} - \beta x_{\mathrm{I}}\,, \tag{165a}$$

$$\frac{dx_{\mathrm{S}}}{dt} = -\alpha x_{\mathrm{I}} x_{\mathrm{S}}\,, \tag{165b}$$

$$\frac{dx_{\mathrm{R}}}{dt} = +\beta x_{\mathrm{I}}\,. \tag{165c}$$

(b)

$$\frac{dx_{\mathrm{A}}}{dt} = (\alpha - \beta) x_{\mathrm{A}} x_{\mathrm{B}}\,, \tag{166a}$$

$$\frac{dx_{\mathrm{B}}}{dt} = (\beta - \alpha) x_{\mathrm{A}} x_{\mathrm{B}}\,. \tag{166b}$$

(c)

$$\frac{dx_{\mathrm{R}}}{dt} = \alpha x_{\mathrm{R}} - \beta x_{\mathrm{R}} x_{\mathrm{F}}\,, \tag{167a}$$

$$\frac{dx_{\mathrm{F}}}{dt} = \beta x_{\mathrm{R}} x_{\mathrm{F}} - \gamma x_{\mathrm{F}}\,. \tag{167b}$$

4. *Consider the rate equations* (6) *for the SIS model.*

   (a) *Find a conserved quantity from the set of equations (hint: conserved quantities do not change in time)*

   The total number of individuals $N = x_{\mathrm{I}} + x_{\mathrm{S}}$ is conserved, as can be verified easily by adding the two equations in (6).

   (b) *Use the conserved quantity to eliminate* x_{S} *in favor of* x_{I} *in equation* (6a)

   By writing $x_{\mathrm{S}} = N - x_{\mathrm{I}}$ one can obtain the following equation, which is the logistic differential equation, due to Verhulst:

$$\frac{dx_{\mathrm{I}}}{dt} = (N\alpha - \beta)x_{\mathrm{I}} - \alpha x_{\mathrm{I}}^2\,. \tag{168}$$

   (c) *Can you write down a Petri net for a single species which is governed by the same equation? If not, why not? If yes, draw the Perti net and give an interpretation of the transitions.*

   It's the Perti net for the reaction network:

$$\mathrm{I} \to 2\mathrm{I}\,, \tag{169a}$$

$$2\mathrm{I} \to \mathrm{I}\,. \tag{169b}$$

   Where the first reaction happens with rate $N\alpha - \beta$ and the second reaction happens with rate α. Note, though, that this is only valid when $N\alpha - \beta > 0$! In that regime, the first reaction can be thought of as reproduction (or a birth process) and the second transmission as inter specific competition, i.e. two I's compete for resources and only one survives.

   (d) *Identify the steady state solutions* x_0 *(obeying* $dx_{\mathrm{I}}/dt = 0$*)*

   There are two steady states, corresponding to the solutions of $(N\alpha - \beta)x_{\mathrm{I}} - \alpha x_{\mathrm{I}}^2 = 0$, which are:

$$x_0 = 0\,, \qquad x_0 = N - \frac{\beta}{\alpha} = N - \frac{1}{\lambda}\,. \tag{170}$$

   (e) *When are the steady state solution stable? In other words, for what values of the dimensionless ratio* $\lambda = \alpha/\beta$ *do linearized perturbations around the steady state solution (*$x_{\mathrm{I}} = x_0 + \epsilon$*) with* $\epsilon \ll 1$*, decay?*

   Since there are two steady state solutions we should analyze both of them by expanding around the solution as:

$$x_{\mathrm{I}} \approx x_0 + \epsilon(t) \tag{171}$$

   First, look at the solution at $x_{\mathrm{I}} = 0$. Inserting in (168) and ignoring terms quadratic in $\epsilon(t)$, we find easily that:

$$\dot{\epsilon}(t) = (N\alpha - \beta)\epsilon(t) \tag{172}$$

And hence, whenever $(N\alpha - \beta) < 0$ (or whenever $\alpha/\beta = \lambda < N$) any perturbation around $x_0 = 0$ will decay and hence in the regime $\lambda < N$ the solution without any infected $x_\mathrm{I} = 0$ is stable.

Now, let's analyze the second fixed point, $x_0 = N - \frac{\beta}{\alpha}$ and write:

$$x_\mathrm{I} \approx N - \frac{1}{\lambda} + \epsilon(t) \tag{173}$$

Inserting this in (168) gives:

$$\dot{\epsilon}(t) = -(N\alpha - \beta)\epsilon(t) \tag{174}$$

Now the situation is reversed: whenever $\lambda > N$ perturbations around this fixed point decay and hence in this regime the second fixed point is stable, where there is a non-zero fraction of infected in the population.

In the usual SIS model, one has that the transmission rate α as defined here scales as $\frac{1}{N}$, such that the number of infected grows proportional to the *fraction* of susceptibles in the population. If we define $\tilde{\alpha} = \alpha/N$ and $\tilde{\lambda} = \tilde{\alpha}/\beta$ then the above analysis show that

- When $\tilde{\lambda} < 1$, the dynamics flows to a stable fixed point where all the population is healthy, which is called the absorbing phase
- When $\tilde{\lambda} > 1$, the dynamics flows to a stable non-trivial fixed point with a finite fraction of infected in the populace, which is called the endemic phase

The fate at the critical point $\tilde{\lambda} = 1$ is not clear from a linear analysis, but when taking quadratic term in ϵ into account one sees that $\dot{\epsilon} < 0$ and so also the critical SIS model will eventually flow towards the healthy absorbing state. This is a first simple example of a dynamical phase transition, which will come bakc over the course of the lectures (including stochastic fluctuations and the critical behavior).

(B) *Integrate the differential equation and give an explanation of its behavior. (Hint: the obtained equation is famously called the logistic equation.)*

The logistic equation is one of the rare treats where a non-linear differential equation has an exact closed-form solution. To find its solution we can rearrange (168) a little bit and formally integrate:

$$(N\alpha - \beta)\int_0^t dt' = \int_{x_0}^{x(t)} \frac{dx'}{x'\left[1 - \frac{x'}{(N-\frac{1}{\lambda})}\right]}\,. \tag{175}$$

Next, we have to massage the fraction on the right hand side a bit and split it in two:

$$\int_{x_0}^{x(t)} dx' \frac{1}{x'\left[1 - \frac{x'}{(N-\frac{1}{\lambda})}\right]} = \int_{x_0}^{x(t)} dx' \frac{1 - \frac{x'}{(N-\frac{1}{\lambda})} + \frac{x'}{(N-\frac{1}{\lambda})}}{x'\left[1 - \frac{x'}{(N-\frac{1}{\lambda})}\right]} = \tag{176}$$

$$\int_{x_0}^{x(t)} dx' \frac{1 - \frac{x'}{(N-\frac{1}{\lambda})}}{x'\left[1 - \frac{x'}{(N-\frac{1}{\lambda})}\right]} + \int_{x_0}^{x(t)} dx' \frac{\frac{x'}{(N-\frac{1}{\lambda})}}{x'\left[1 - \frac{x'}{(N-\frac{1}{\lambda})}\right]}$$

$$\int_{x_0}^{x(t)} dx' \frac{1}{x'} + \int_{x_0}^{x(t)} dx' \frac{1}{\left[(N-\frac{1}{\lambda}) - x'\right]}$$

Now we have two integrals which we can perform. They give us logarithms! So:

$$(N\alpha - \beta)t = \log(x')|_{x_0}^{x(t)} - \log((N - \frac{1}{\lambda}) - x')|_{x_0}^{x(t)} \tag{177}$$

$$= \log\left[\frac{x(t)}{(N - \frac{1}{\lambda}) - x(t)}\right] - \log\left[\frac{x_0}{(N - \frac{1}{\lambda}) - x_0}\right] \tag{178}$$

Now all that is left to do is to take the exponent on both sides and solve for $x(t)$.

$$e^{(N\alpha-\beta)t} = \frac{x(t)}{(N - \frac{1}{\lambda}) - x(t)}\left(\frac{(N - \frac{1}{\lambda})}{x_0} - 1\right) \tag{179}$$

$$\Rightarrow e^{(N\alpha-\beta)t}\left((N - \frac{1}{\lambda}) - x(t)\right) = x(t)\left(\frac{(N - \frac{1}{\lambda})}{x_0} - 1\right) \tag{180}$$

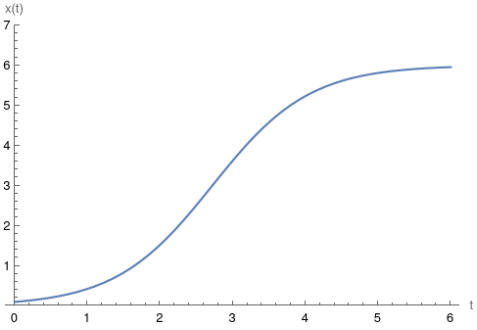
$$\Rightarrow e^{(N\alpha-\beta)t}(N - \frac{1}{\lambda}) = x(t)\left(\frac{(N - \frac{1}{\lambda})}{x_0} - (1 - e^{-(N\alpha-\beta)t})\right) \tag{181}$$

$$\Rightarrow x(t) = \frac{e^{(N\alpha-\beta)t}(N - \frac{1}{\lambda})}{\left(\frac{(N-\frac{1}{\lambda})}{x_0} - (1 - e^{(N\alpha-\beta)t})\right)} \tag{182}$$

Or, after rearranging a little bit more, we find the logistic function as an answer:

$$x(t) = \frac{(N - \frac{1}{\lambda})}{1 + \frac{(N-\frac{1}{\lambda})-x_0}{x_0}e^{-(N\alpha-\beta)t}} \tag{183}$$

In the endemic phase, when $N\alpha - \beta > 0$, the logistic function above has a characteristic S-shaped curve, where it starts with exponential growth e^{rt} with rate $r = (N\alpha - \beta)$ and continuous to grow until it reaches its *carrying capacity* $K = (N - \frac{1}{\lambda})$ in our case. A typical graph of the function for values in the endemic region would look like this: When $N\alpha - \beta < 0$, we see instead immediate

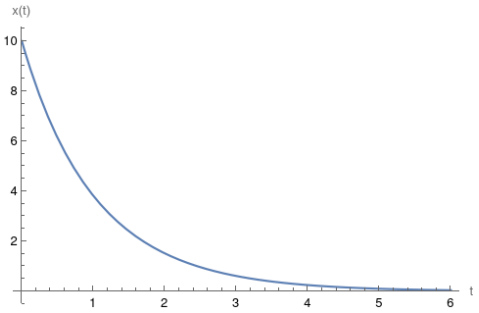


exponential decay to the trivial fixed point $x = 0$

3.2.1 Solutions

1. *Show that equations* (13), (14) *and* (15) *lead to the commutation relations for the creation and annihilation operators* (16).

   The commutation relations only make sense in the operator sense, where it is supposed that the operators act on arbitrary states from the left, as in the definition of the Fock



space basis states $|\mathbf{n}\rangle$. So we should compute the action of the commutator on an arbitrary basis vector and find that it gives an eigenvalue equation.

$$\begin{aligned}[\hat{a}_i, \hat{a}_i^\dagger]|\mathbf{n}\rangle &= \hat{a}_i\hat{a}_i^\dagger|\mathbf{n}\rangle - \hat{a}_i^\dagger\hat{a}_i|\mathbf{n}\rangle \\ &= \hat{a}_i|n_1, \ldots, n_i + 1, \ldots, n_k\rangle - n_i\hat{a}_i^\dagger|n_1, \ldots, n_i - 1, \ldots, n_k\rangle \\ &= (n_i + 1)|\mathbf{n}\rangle - n_i|\mathbf{n}\rangle = |\mathbf{n}\rangle\end{aligned} \tag{184}$$

This shows that $[\hat{a}_i, \hat{a}_i^\dagger] = 1$. Likewise, it is easy to demonstrate that $[\hat{a}_i, \hat{a}_j^\dagger] = 0$ whenever $i \neq j$, which completes the proof that $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$.

Similarly, $[\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0$ is straightforward to show by applying the operators to an arbitrary state $|\mathbf{n}\rangle$.

2. *Show that the state* $|\mathbf{n}\rangle$ *(defined in* (13)*) is an eigenvector for the number operator* $\hat{N}_i = \hat{a}_i^\dagger\hat{a}_i$ *for species* i*.*

To show this, we apply $\hat{N}_i = \hat{a}_i^\dagger\hat{a}_i$ on the basis state $|\mathbf{n}\rangle$ and use the commutation relations (16) to move the annihilation operator $\hat{a}_i$ through to the right. First write out the definitions

$$\hat{N}_i|\mathbf{n}\rangle = \hat{a}_i^\dagger\hat{a}_i\prod_{j\in S}(\hat{a}_j^\dagger)^{n_j}|0\rangle \tag{185}$$

If $i \neq j$, the operators commute, so nothing has to be done. When $j = i$ we can write $\hat{a}_i\hat{a}_i^\dagger = \hat{a}_i^\dagger\hat{a}_i + 1$ and so for those terms we do get a nontrivial term:

$$\prod_{j\in S}\hat{a}_i^\dagger\hat{a}_i\hat{a}_j^\dagger(a_j^\dagger)^{n_j-1}|0\rangle = \prod_{j\in S}\hat{a}_i^\dagger(\hat{a}_j^\dagger\hat{a}_i + \delta_{ij})(a_j^\dagger)^{n_j-1}|0\rangle = \prod_{j\in S}\hat{a}_i^\dagger(\hat{a}_j^\dagger\hat{a}_i(a_j^\dagger)^{n_j-1}|0\rangle + |\mathbf{n}\rangle \tag{186}$$

We repeat this process for the first term on the left hand side until the annihilation operator is moved all the way to the left and annihilates with the vacuum. At each iteration, we get one $|\mathbf{n}\rangle$ whenever $j = i$ and so in the end we are just counting the number of elements of species i in $|\mathbf{n}\rangle$:

$$\hat{N}_i|\mathbf{n}\rangle = n_i|\mathbf{n}\rangle \tag{187}$$

3. *Show that the states* $\langle\mathbf{n}|$ *and* $|\mathbf{m}\rangle$ *are orthogonal, using their definitions and the commutation relations* (16)

This exercise is a straightforward application of the definitions. Start with writing out the definition of the basis states:

$$\langle\mathbf{m}|\mathbf{n}\rangle = \prod_{i\in S}\prod_{j\in S}\langle 0|\hat{a}_i^{m_i}(\hat{a}_j^\dagger)^{n_j}|0\rangle \tag{188}$$

Then, we wish to move all the annihilation operators $\hat{a}_i$ to the left of the creation operators $\hat{a}_j^\dagger$. If $i \neq j$, the operators commute, so nothing has to be done. When $j = i$ we can write $\hat{a}_i\hat{a}_i^\dagger = \hat{a}_i^\dagger\hat{a}_i + 1 = \hat{N}_i + 1$ and so we can write use the fact that $\hat{N}_i$ is diagonal in the basis states (see exercise above):

$$\prod_{i\in S}\prod_{j\in S}\langle 0|\hat{a}_i^{m_i}(\hat{a}_j^\dagger)^{n_j}|0\rangle = \prod_{i\in S}\prod_{j\in S}\langle 0|\hat{a}_i^{m_i-1}(\hat{N}_i+1)(\hat{a}_j^\dagger)^{n_j-1}|0\rangle \tag{189}$$

$$= \prod_{i\in S}\prod_{j\in S} n_i\langle 0|\hat{a}_i^{m_i-1}(\hat{a}_j^\dagger)^{n_j-1}|0\rangle \tag{190}$$

By repeating this k times, we continue to pull out eigenvalues of $\hat{N}_i + 1$, which decrease by one in every step. This process continues until either $m_i - k = 0$ or $n_i - k = 0$. Then three things can happen. Whenever $m_i < n_i$, we are left with creation operators $\hat{a}_i^\dagger$, which will annihilate the vacuum $\langle 0|$. Whenever $m_i > n_i$, we are left with annihilation operators $\hat{a}_i$, which annihilate the vacuum $|0\rangle$. So *only* when $m_i = n_i$ and we are both out of creation and annihilation operators do we get a non-zero result, namely $\langle 0|0\rangle = 1$.

All of this leads to:

$$\langle \mathbf{m}|\mathbf{n}\rangle = \prod_{i\in S}\prod_{j\in S}\delta_{m_i,n_j} n_i! = \mathbf{n}!\delta_{\mathbf{m},\mathbf{n}}\,. \tag{191}$$

4. *Show that* $\sum_{\mathbf{n}} \frac{1}{\mathbf{n}!}\langle \mathbf{n}| = \langle 0|e^{\sum_i \hat{a}_i}$ *by using the appropriate definitions and the usual definition of the exponential of an operator* $e^{\hat{a}_i} = \sum_n \frac{1}{n!}\hat{a}_i^n$

$$\sum_{\mathbf{n}} \frac{1}{\mathbf{n}!}\langle \mathbf{n}| = \prod_i \sum_{n_i} \frac{1}{n_i!}\langle 0|\hat{a}_i^{n_i} = \prod_i \langle 0|e^{\hat{a}_i} = \langle 0|e^{\sum_i \hat{a}_i} \tag{192}$$

5. *Proof that the coherent state* $|\mathbf{z}\rangle = e^{\sum_i z_i\hat{a}_i^\dagger}|0\rangle$ *is an eigenvector for the annihilation operator* $\hat{a}_i$ *with eigenvalue* z_i*:*

$$\hat{a}_i|\mathbf{z}\rangle = z_i|\mathbf{z}\rangle \tag{193}$$

For this exercise we first work out the commutator of $[\hat{a}_i, (\hat{a}_i^\dagger)^{n_i}]$, which is:

$$[\hat{a}_i, (\hat{a}_i^\dagger)^n] = \hat{a}_i(\hat{a}_i^\dagger)^n - (\hat{a}_i^\dagger)^n\hat{a}_i \tag{194}$$

$$= \hat{a}_i(\hat{a}_i^\dagger)^n - \hat{a}_i^\dagger\hat{a}_i(\hat{a}_i^\dagger)^{n-1} + \hat{a}_i^\dagger\hat{a}_i(\hat{a}_i^\dagger)^{n-1} - (\hat{a}_i^\dagger)^n\hat{a}_i$$

$$= [\hat{a}_i, \hat{a}_i^\dagger](\hat{a}_i^\dagger)^{n-1} + \hat{a}_i^\dagger\hat{a}_i(\hat{a}_i^\dagger)^{n-1} - (\hat{a}_i^\dagger)^n\hat{a}_i$$

$$= (\hat{a}_i^\dagger)^{n-1} + \hat{a}_i^\dagger\hat{a}_i(\hat{a}_i^\dagger)^{n-1} - (\hat{a}_i^\dagger)^n\hat{a}_i$$

$$= (\hat{a}_i^\dagger)^{n-1} + \hat{a}_i^\dagger\hat{a}_i(\hat{a}_i^\dagger)^{n-1} - (\hat{a}_i^\dagger)^2\hat{a}_i(\hat{a}_i^\dagger)^{n-2} + (\hat{a}_i^\dagger)^2\hat{a}_i(\hat{a}_i^\dagger)^{n-2} - (\hat{a}_i^\dagger)^n\hat{a}_i$$

$$= 2(\hat{a}_i^\dagger)^{n-1} + (\hat{a}_i^\dagger)^2\hat{a}_i(\hat{a}_i^\dagger)^{n-2} - (\hat{a}_i^\dagger)^n\hat{a}_i$$

$$\vdots \tag{195}$$

$$= n(\hat{a}_i^\dagger)^{n-1} \tag{196}$$

Using this result, we can write:

$$\hat{a}_i|\mathbf{z}\rangle = \prod_j \sum_{n_j} \frac{1}{n_j!}\hat{a}_i(z_j\hat{a}_j^\dagger)^{n_j}|0\rangle = \prod_j \sum_{n_j} \frac{1}{n_j!} z_j^{n_j}[\hat{a}_i, (\hat{a}_j^\dagger)^{n_j}]|0\rangle \tag{197}$$

$$= \sum_{n_i} \frac{n_i z_i}{n_i!}(z_i\hat{a}_i^\dagger)^{n_i-1}\left(\prod_{j\neq i}\sum_{n_j}\frac{(z_j\hat{a}_j^\dagger)^{n_j}}{n_j!}\right)|0\rangle = z_i|\mathbf{z}\rangle\,. \tag{198}$$

Likewise, one can show that the dual coherent state $\langle\mathbf{z}|$ satisfies:

$$\langle\mathbf{z}|\hat{a}_i^\dagger = z_i\langle\mathbf{z}|\,. \tag{199}$$

6. *Consider the case where we have only a single species A and a coherent state $|z\rangle = e^{z\hat{a}^\dagger}|0\rangle$, where $\hat{a}^\dagger$ is the creation operator for A.*

   (a) *Compute $c(z)$ such that the state $|\phi(z)\rangle = c(z)|z\rangle$ forms a normalized probability distribution (such that $\langle\mathbf{1}|\phi(z)\rangle = 1$).*

   Let's first compute $\langle\mathbf{1}|z\rangle$ and then use that $c(z) = 1/\langle\mathbf{1}|z\rangle$ gives $\langle\mathbf{1}|\phi(z)\rangle = 1$:

   $$\langle\mathbf{1}|z\rangle = \sum_n \langle 0|\frac{1}{n!}\hat{a}^n|z\rangle = \sum_n \frac{z^n}{n!}\langle 0|z\rangle = e^z \tag{200}$$

   The second equation follows from $\hat{a}|z\rangle = z|z\rangle$, as shown in the exercise above. the last equation uses that $\langle 0|z\rangle = 1$. Finally, $c(z) = e^{-z}$ so that

   $$|\phi(z)\rangle = e^{-z}|z\rangle\,, \tag{201}$$

   forms a normalized probability distribution.

   (b) *Compute the probability $P(n)$ of having n elements of species A in the state $|\phi(z)\rangle$. In other words, compute the components $P(n)$ of $|\phi(t)\rangle = \sum_n P(n)|n\rangle$. What distribution does $\phi(z)$ signify?*

   We may write

   $$|\phi(z)\rangle = e^{-z}|z\rangle = \sum_n \frac{e^{-z}}{n!}z^n(\hat{a}^\dagger)^n|0\rangle = \sum_n \frac{e^{-z}}{n!}z^n|n\rangle\,. \tag{202}$$

   Therefore:

   $$P(n) = \frac{e^{-z}}{n!}z^n\,, \tag{203}$$

   which you can recognize as the Poisson distribution.

   (c) *What is the expected number of elements of species A in the distribution $|\phi(z)\rangle$? (hint: this is computed by $\langle\mathbf{1}|\hat{N}|\phi(z)\rangle$)*

   We can compute the expectation value for the number operator as $\langle\hat{N}\rangle = \langle\mathbf{1}|\hat{N}|\phi(z)\rangle$. Now we will use the tricks we have proven above:

   $$\langle\mathbf{1}|\hat{N}|\phi(z)\rangle = \langle\mathbf{1}|\hat{a}^\dagger\hat{a}|\phi(z)\rangle = \langle\mathbf{1}|\hat{a}|\phi(z)\rangle\,. \tag{204}$$

   The last equation follows from $\langle\mathbf{1}|\hat{a}^\dagger = \langle\mathbf{1}|$. Then using $\hat{a}|z\rangle = z|z\rangle$ we find immediately:

   $$\langle\mathbf{1}|\hat{N}|\phi(z)\rangle = z\langle\mathbf{1}|\phi(z)\rangle = z \tag{205}$$

   Or in words, the expectation value of the Poisson distribution with parameter z is z, which is a well-known fact.

(d) *What is the variance in the expected number of elements of species A in the distribution $|\phi(z)\rangle$? In other words, what is $\langle \mathbf{1}|\hat{N}^2|\phi(z)\rangle - \langle \mathbf{1}|\hat{N}|\phi(z)\rangle^2$*

This now requires to compute the expectation value of the square of the number operator. We can do this by first normal ordering:

$$\hat{N}^2 = \hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a} = \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} + \hat{a}^\dagger \hat{a} \tag{206}$$

Then, using again that $\langle \mathbf{1}|\hat{a}^\dagger = \langle \mathbf{1}|$ and $\hat{a}|z\rangle = z|z\rangle$ we have immediately that

$$\begin{aligned} \mathrm{Var}[\hat{N}] &\equiv \langle \mathbf{1}|\hat{N}^2|\phi(z)\rangle - \langle \mathbf{1}|\hat{N}|\phi(z)\rangle^2 && (207)\\ &= \langle \mathbf{1}|\hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a}|\phi(z)\rangle + \langle \mathbf{1}|\hat{a}^\dagger \hat{a}|\phi(z)\rangle - \langle \mathbf{1}|\hat{a}^\dagger \hat{a}|\phi(z)\rangle^2 && (208)\\ &= z^2 + z - z^2 = z && (209) \end{aligned}$$

Or, the variance of the Poisson distribution with parameter z is also z, which is also a well-known fact.

3.3.1 Solutions

1. *For all the reactions in* (3.1.1)*, write down the generator $\hat{H}$*

   (a) $\hat{H} = \alpha(\hat{a}_\mathrm{I}^\dagger \hat{a}_\mathrm{I}^\dagger - \hat{a}_\mathrm{S}^\dagger \hat{a}_\mathrm{I}^\dagger)\hat{a}_\mathrm{S}\hat{a}_\mathrm{I} + \beta(\hat{a}_\mathrm{R}^\dagger - \hat{a}_\mathrm{I}^\dagger)\hat{a}_\mathrm{I}$

   (b) $\hat{H} = \alpha(\hat{a}_\mathrm{A}^\dagger \hat{a}_\mathrm{A}^\dagger - \hat{a}_\mathrm{A}^\dagger \hat{a}_\mathrm{B}^\dagger)\hat{a}_\mathrm{A}\hat{a}_\mathrm{B} + \beta(\hat{a}_\mathrm{B}^\dagger \hat{a}_\mathrm{B}^\dagger - \hat{a}_\mathrm{A}^\dagger \hat{a}_\mathrm{B}^\dagger)\hat{a}_\mathrm{A}\hat{a}_\mathrm{B}$

   (c) $\hat{H} = \alpha(\hat{a}_\mathrm{R}^\dagger \hat{a}_\mathrm{R}^\dagger - \hat{a}_\mathrm{R}^\dagger)\hat{a}_\mathrm{R} + \beta(\hat{a}_\mathrm{F}^\dagger \hat{a}_\mathrm{F}^\dagger - \hat{a}_\mathrm{R}^\dagger \hat{a}_\mathrm{F}^\dagger)\hat{a}_\mathrm{R}\hat{a}_\mathrm{F} + \gamma(1 - \hat{a}_\mathrm{F}^\dagger)\hat{a}_\mathrm{F}$

2. *We have encountered already many similarities between stochastic and quantum systems, however the two are definitely not the same. To make the distinction crystal clear, what are the stochastic analogues of the following quantum mechanical properties and statements?*

   - *The wavefunction $|\psi(t)\rangle$ and its amplitudes ψ_n, such that $|\psi(t)\rangle = \sum_n \psi_n |n\rangle$ and $\langle\psi(t)|\psi(t)\rangle = 1$ for all t.*

     Stochastic analogue is the probability vector in the Fock space representation $|\phi(t)\rangle = \sum_{\mathbf{n}} P(\mathbf{n}, t)|\mathbf{n}\rangle$. Now, however, we have $\langle \mathbf{1}|\phi(t)\rangle = 1$ at all t, such that this vector is normalized in the L^1 norm (not the L^2 norm, as the wavefunction.)

   - *The expectation values for observables $\bar{O} = \langle\psi|\hat{O}|\psi\rangle$*

     Expectation values of operators are now $\bar{O} = \langle \mathbf{1}|\hat{O}|\phi(t)\rangle$.

     Note that this implies some funny stuff. If I have a normal ordered operator $\hat{O}(\hat{a}_i^\dagger, \hat{a}_i)$, where normal ordering implies that all creation operators $\hat{a}_i^\dagger$ are to the left of annihilation operators $\hat{a}_i$, then the expectation value of this operator is the same as the expectation value of the operator where all creation operators are set to 1: $\hat{O}(\hat{a}_i^\dagger = 1, \hat{a}_i)$. This follows immediately from $\langle \mathbf{1}|\hat{a}_i^\dagger = \langle \mathbf{1}|$.

     Likewise, the expectation value for the operator $\hat{O}$ equals to the expectation value of $(\hat{a}_i^\dagger)^k \hat{O}$ for any k.

   - *A Hermitian operator satisfying $\hat{H}^\dagger = \hat{H}$*

     Hermitian operators in quantum mechanics define a time-evolution for the system. In the stochastic case, time evolution is generated by infinitesimal stochastic operators $\hat{H}$ which satisfy $\langle \mathbf{1}|\hat{H} = 0$ and $\langle \mathbf{m}|\hat{H}|\mathbf{n}\rangle \geq 0$ for $\mathbf{m} \neq \mathbf{n}$.

- *A one-parameter unitary group $\hat{U}(t)$, such that $\hat{U}^{-1} = \hat{U}^\dagger$. (This one's actually a bit tricky since stochastic systems are not generally invertible, what is the stochastic analogue of this statement?)*

  The one-parameter unitary group can generally be written as $\hat{U}(t) = e^{-i\hat{H}t}$ (Stone's theorem) such that $\hat{H}$ is Hermitian (self-adjoint). This ensures the conservation of the L^2 norm under time-evolution in QM: $\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|e^{it\hat{H}^\dagger}e^{-it\hat{H}}|\psi(0)\rangle = \langle\psi(0)|\psi(0)\rangle = 1$.

  In stochastic mechanics, we don't have unitary time-evolution, but neither do we need this to preserve the vectors norm, because now the probability vector is normalized in the L^1 norm: $\langle\mathbf{1}|\phi(t)\rangle = 1$. To conserve this, we need the time-evolution operator $\hat{S}(t) = e^{t\hat{H}}$ to be a stochastic operator: $\langle\mathbf{1}|\hat{S}(t) = \langle\mathbf{1}|$, such that $\langle\mathbf{1}|\phi(t)\rangle = \langle\mathbf{1}|\hat{S}(t)|\phi(0)\rangle = \langle\mathbf{1}|\phi(0)\rangle$.

  Note that now, we don't need any statement on the inverse time-evolution operator. Generically, stochastic operators such as $\hat{S}(t) = e^{t\hat{H}}$ for $t > 0$ define a one-parameter semigroup, called the Markov semigroup.

3. *Consider the process:*

$$\mathrm{A} \to \emptyset\,. \tag{210}$$

with rate α

(a) *Write down the generator for this process*

$$\hat{H} = \alpha(1 - \hat{a}^\dagger)\hat{a}\,. \tag{211}$$

(b) *Suppose that initially our system has k particles of species A, such that $|\phi(0)\rangle = |k\rangle = (\hat{a}^\dagger)^k|0\rangle$. What is $|\phi(t)\rangle$?*

Formally, the system is solved as:

$$|\phi(t)\rangle = e^{t\hat{H}}|k\rangle = e^{\alpha t(1-\hat{a}^\dagger)\hat{a}}(\hat{a}^\dagger)^k|0\rangle\,. \tag{212}$$

We have to work out some commutators, which we'll do explicitly step by step. First note that we can use the Campbell identity:

$$e^X Y e^{-X} = \sum_{n=0}^{\infty} \frac{[(X)^n, Y]}{n!}\,, \tag{213}$$

where $[(X)^n, Y] = [X, \ldots [X, [X, Y]] \ldots]$ is the nested commutator. We suppose here that $X = t\hat{H}$ and $Y = (\hat{a}^\dagger)^k$, and use (213) with e^X multiplied from the right on both sides:

$$|\phi(t)\rangle = \sum_{n=0}^{\infty} \frac{[(t\hat{H})^n, (\hat{a}^\dagger)^k]}{n!} e^{t\hat{H}}|0\rangle = \sum_{n=0}^{\infty} \frac{[(t\hat{H})^n, (\hat{a}^\dagger)^k]}{n!}|0\rangle\,. \tag{214}$$

In the last equality we used that $e^{t\hat{H}}|0\rangle = |0\rangle$ as $\hat{H}$ annihilates the vacuum state. Now we use the other commutator identity given in the hint:

$$\sum_{n=0}^{\infty} \frac{[(X)^n, A^k]}{n!} = \left(\sum_{n=0}^{\infty} \frac{[(X)^n, A]}{n!}\right)^k\,, \tag{215}$$

with $A = \hat{a}^\dagger$ such that:

$$|\phi(t)\rangle = \left(\sum_{n=0}^{\infty} \frac{[(t\hat{H})^n, \hat{a}^\dagger]}{n!}\right)^k |0\rangle \tag{216}$$

This simplifies things a bit. All we have to do now is work out $[(t\hat{H})^n, \hat{a}^\dagger]$. The zeroth order term is just $\hat{a}^\dagger$. The first order term works out to:

$$[(t\hat{H})^1, \hat{a}^\dagger] = \alpha t[(1-\hat{a}^\dagger)\hat{a}, \hat{a}^\dagger] = \alpha t(1-\hat{a}^\dagger)\,. \tag{217}$$

Then the second order term follows as:

$$[(t\hat{H})^2, \hat{a}^\dagger] = (\alpha t)^2[(1-\hat{a}^\dagger)\hat{a}, (1-\hat{a}^\dagger)] = -(\alpha t)^2(1-\hat{a}^\dagger)\,. \tag{218}$$

We're starting to see the pattern now:

$$[(t\hat{H})^n, \hat{a}^\dagger] = -(-\alpha t)^n(1-\hat{a}^\dagger)\,. \tag{219}$$

This we plug into the sum (216) such that:

$$|\phi(t)\rangle = \left(\hat{a}^\dagger + \sum_{n=1}^{\infty} \frac{-(-\alpha t)^n}{n!}(1-\hat{a}^\dagger)\right)^k |0\rangle \tag{220}$$

$$= \left(1 - e^{-\alpha t}(1-\hat{a}^\dagger)\right)^k |0\rangle\,. \tag{221}$$

This is the final answer we are after at this point

(c) *Derive from $|\phi(t)\rangle$ the probability of having $n < k$ particles at time t.*

To get this, we must expand the answer above in powers of the creation operator $\hat{a}^\dagger$. This is straight-forward using the binomial series:

$$|\phi(t)\rangle = \left(1 - e^{-\alpha t}(1-\hat{a}^\dagger)\right)^k |0\rangle \tag{222}$$

$$= \left((1-e^{-\alpha t}) + e^{-\alpha t}\hat{a}^\dagger)\right)^k |0\rangle \tag{223}$$

$$= \sum_{n=0}^{k} \binom{k}{n} (1-e^{-\alpha t})^{k-n}(e^{-\alpha t}\hat{a}^\dagger)^n|0\rangle \tag{224}$$

$$= \sum_{n=0}^{k} \binom{k}{n} (1-e^{-\alpha t})^{k-n}e^{-n\alpha t}|n\rangle \tag{225}$$

Recall that $|\phi(t)\rangle = \sum_n P(n,t)|n\rangle$, such that the probability of having n particles at time t is:

$$P(n,t) = \binom{k}{n}(1-e^{-\alpha t})^{k-n}e^{-n\alpha t}\,. \tag{226}$$

(d) *What is the expected number of particles at time t?*

One could of course compute this immediately from $P(n,t)$ now that we found this by evaluating:

$$\sum_n nP(n,t) = ke^{-\alpha t}\,, \tag{227}$$

but then we're not making use of our quantum language. Let's try to do this computation using some tricks instead. Recall that $\langle \mathbf{1}|\hat{a}^\dagger = \langle \mathbf{1}|$, so that every time that we hit a creation operator from the left with a flat state, we might as well replace it by a 1. Using this, we will write:

$$\begin{aligned}
\langle \hat{N} \rangle = \langle \mathbf{1}|\hat{a}|\phi(t)\rangle &= \langle \mathbf{1}|\hat{a}(1-e^{-\alpha t}(1-\hat{a}^\dagger))^k|0\rangle && (228)\\
&= \langle \mathbf{1}|(1-e^{-\alpha t}(1-\hat{a}^\dagger))\hat{a}(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle \\
&\quad + \langle \mathbf{1}|[\hat{a},(1-e^{-\alpha t}(1-\hat{a}^\dagger))](1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle && (229)\\
&= \langle \mathbf{1}|\hat{a}(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle + e^{-\alpha t}\langle \mathbf{1}|(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle && (230)\\
&= \langle \mathbf{1}|\hat{a}(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle + e^{-\alpha t} && (231)\\
&= \ldots \text{repeat another } k-1 \text{ times} && (232)\\
&= ke^{-kt} && (233)
\end{aligned}$$

Here we have used several times that $\langle \mathbf{1}|(1-e^{-\alpha t}(1-\hat{a}^\dagger)) = \langle \mathbf{1}|$.

(e) *What is the variance in the number of particles at time t?*

In a similar vein to the above question, we compute now $\mathrm{Var[N]} = \langle \mathbf{1}|\hat{\mathrm{N}}^2|\phi(\mathrm{t})\rangle - \langle \mathbf{1}|\hat{\mathrm{N}}|\phi(\mathrm{t})\rangle^2$. First recall that $\hat{N}^2 = \hat{a}^{\dagger 2}\hat{a}^2 + \hat{a}^\dagger\hat{a}$ and then the interesting new computation w.r.t above is:

$$\begin{aligned}
\langle \mathbf{1}|\hat{a}^2|\phi(t)\rangle &= \langle \mathbf{1}|\hat{a}^2(1-e^{-\alpha t}(1-\hat{a}^\dagger))^k|0\rangle && (234)\\
&= \langle \mathbf{1}|\hat{a}^2(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle \\
&\quad + \langle \mathbf{1}|[\hat{a}^2,(1-e^{-\alpha t}(1-\hat{a}^\dagger))](1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle && (235)\\
&= \langle \mathbf{1}|\hat{a}^2(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle \\
&\quad + 2e^{-\alpha t}\langle \mathbf{1}|\hat{a}(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle && (236)\\
&= \langle \mathbf{1}|\hat{a}^2(1-e^{-\alpha t}(1-\hat{a}^\dagger))^{k-1}|0\rangle + 2(k-1)e^{-2\alpha t} && (237)\\
&= \ldots \text{repeat another } k-1 \text{ times} && (238)\\
&= \sum_{n=1}^{k} 2(k-n)e^{-2kt} = (k^2-k)e^{-2\alpha t}\,. && (239)
\end{aligned}$$

Then we can put everything together by writing:

$$\begin{aligned}
\mathrm{Var[N]} &= \langle \mathbf{1}|\hat{N}^2|\phi(t)\rangle - \langle \mathbf{1}|\hat{N}|\phi(t)\rangle^2 && (240)\\
&= \langle \mathbf{1}|\hat{a}^2|\phi(t)\rangle + \langle \mathbf{1}|\hat{a}|\phi(t)\rangle - \langle \mathbf{1}|\hat{a}|\phi(t)\rangle^2 && (241)\\
&= (k^2-k)e^{-2\alpha t} + ke^{-\alpha t} - k^2e^{-2\alpha t} && (242)\\
&= ke^{-2\alpha t}(e^{\alpha t}-1) && (243)
\end{aligned}$$

You may verify by yourself that this gives the same answer as the sum:

$$\sum_n n^2 P(n,t) - \left(\sum_n nP(n,t)\right)^2\,, \qquad (244)$$

with $P(n,t)$ given in (226) above

4. *Usually, the master equation is formulated in terms of $\frac{d}{dt}P(n,t)$ and derived heuristically from arguing how the probability of having n particles changes as a result of the reactions present in the system.*

   For instance, for the death process of the previous question, one would say that $\frac{d}{dt}P(n,t)$ receives $(n+1)$ positive contributions from $P(n+1,t)$ with rate α, because if we are in a state with $n+1$ particles (which happens with probability $P(n+1,t)$), then each of the $n+1$ particles could decay with rate α, leading to the state with n particles.

   Likewise, $P(n,t)$ receives n negative contributions from $P(n,t)$ with rate α, because if we are in the state with n particles, each of these could decay and take us out of the n particle state. So we have heuristically argued that:

$$\frac{d}{dt}P(n,t) = \alpha(n+1)P(n+1,t) - \alpha n P(n,t)\,. \tag{245}$$

   (a) *Derive this equation from the second quantized master equation by normal ordering and equating like powers of $\hat{a}^\dagger$ on both sides.*

   The second quantized master equation is:

$$\frac{d}{dt}|\phi(t)\rangle = H|\phi(t)\rangle\,, \tag{246}$$

   with

$$H = \alpha(1-\hat{a}^\dagger)\hat{a}\,. \tag{247}$$

   We also know that the probability vector expanded in the Fock space is

$$|\phi(t)\rangle = \sum_{n=0}^{\infty} P(n,t)|n\rangle = \sum_{n=0}^{\infty} P(n,t)(\hat{a}^\dagger)^n|0\rangle\,. \tag{248}$$

   So we may write:

$$\frac{d}{dt}|\phi(t)\rangle = \sum_{n=0}^{\infty}\frac{d}{dt}P(n,t)|n\rangle = \sum_{n=0}^{\infty} P(n,t)\alpha(1-\hat{a}^\dagger)\hat{a}(\hat{a}^\dagger)^n|0\rangle \tag{249}$$

$$= \sum_{n=0}^{\infty} P(n,t)\alpha\left(1-\hat{a}^\dagger\right)[\hat{a},(\hat{a}^\dagger)^n]|0\rangle \tag{250}$$

$$= \sum_{n=0}^{\infty} P(n,t)\alpha\left(1-\hat{a}^\dagger\right)n(\hat{a}^\dagger)^{n-1}|0\rangle \tag{251}$$

$$= \sum_{n=0}^{\infty} P(n,t)\alpha\left(n|n-1\rangle - n|n\rangle\right) \tag{252}$$

$$= \sum_{n=0}^{\infty} \alpha(n+1)P(n+1,t)|n\rangle - \alpha n P(n,t)|n\rangle\,. \tag{253}$$

   Here we have used $[\hat{a},(\hat{a}^\dagger)^n] = n(\hat{a}^\dagger)^{n-1}$ and in the last equality we have shifted the sum over the first term as $n \to n+1$ in order to get the like powers of $\hat{a}^\dagger$ explicit. Now we can simply equate the coefficients of the basis states $|n\rangle$ on both sides and arrive at (245).

(b) *Repeat this for the (linear) birth-death process, where there are two reactions:*

$$\mathrm{A} \to \emptyset\,, \tag{254a}$$

$$\mathrm{A} \to 2\mathrm{A}\,. \tag{254b}$$

Repeating the steps above in a similar fashion, but now for the generator

$$\hat{H} = \alpha(1-\hat{a}^\dagger)\hat{a} + \beta(\hat{a}^\dagger\hat{a}^\dagger - \hat{a}^\dagger)\hat{a}\,, \tag{255}$$

gives the master equation of the linear birth-death process:

$$\frac{d}{dt}P(n,t) = \alpha(n+1)P(n+1,t) + \beta(n-1)P(n-1,t) - (\alpha+\beta)nP(n,t)\,. \tag{256}$$

3.4.1 Solutions

1. *Within the framework of emergence, what are generic emergent features of the macroscopic rate equations not present in the microscopic stochastic models?*

   Ordinary differential equations which are continuous, deterministic and often non-linear. The microscopic description is probabilistic and has a discrete state space, so elements change discretely in time, (one more or less each time step). Like a jump process.

2. *Work out the rate equation of $x_i(t)$ in a general chemical reaction network with generator $\hat{H}$ given in* (37)*. When do you reproduce equation* (5)*?*

   The definition of x_i in terms of microscopic constants is the expectation value of the number operator $x_i(t) = \langle \mathbf{1}|\hat{N}_i|\phi(t)\rangle$. Let's work out it's time derivative

$$\frac{d}{dt}x_i(t) = \frac{d}{dt}\langle \mathbf{1}|\hat{N}_i|\phi(t)\rangle = \langle \mathbf{1}|[\hat{a}_i,\hat{H}]|\phi(t)\rangle \tag{257}$$

$$= \sum_\tau r(\tau)\langle \mathbf{1}|\left[\hat{a}_i, \left((\hat{\mathbf{a}}^\dagger)^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^\dagger)^{\mathbf{m}(\tau)}\right)\hat{\mathbf{a}}^{\mathbf{m}(\tau)}\right]|\phi(t)\rangle \tag{258}$$

$$= \sum_\tau r(\tau)\langle \mathbf{1}|\left[\hat{a}_i, \left((\hat{\mathbf{a}}^\dagger)^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^\dagger)^{\mathbf{m}(\tau)}\right)\right]\hat{\mathbf{a}}^{\mathbf{m}(\tau)}|\phi(t)\rangle \tag{259}$$

$$= \sum_\tau r(\tau)\left(n_i(\tau) - m_i(\tau)\right)\langle \mathbf{1}|\hat{\mathbf{a}}^{\mathbf{m}(\tau)}|\phi(t)\rangle \tag{260}$$

   This gives the rate equation (5) whenever we suppose that

$$\langle \mathbf{1}|\hat{\mathbf{a}}^{\mathbf{m}(\tau)}|\phi(t)\rangle = \prod_i \langle \mathbf{1}|\hat{a}_i|\phi(t)\rangle^{n_i} = \prod_i x_i^{n_i}\,. \tag{261}$$

   This is the mean-field approximation, which holds generically in the limit of large volumes, when the total number of elements goes to infinity $N \to \infty$.

3. *Consider again the SIS model and the equation for $x_{\mathrm{I}}(t)$ given in* (50)*. Suppose that we do not wish to assume that $\langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle \cong \langle \mathbf{1}|\hat{a}_{\mathrm{S}}|\phi(t)\rangle\langle \mathbf{1}|\hat{a}_{\mathrm{I}}|\phi(t)\rangle$. Then, to describe the systems evolution, we would also need an equation for the SI pairs: $x_{\mathrm{SI}} = \langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle$. Derive this equation from the master equation with generator* (49)

To work out the equation for SI pairs $x_{\mathrm{SI}} = \langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle$, we work out the commutator of $\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}$ with $\hat{H} = \alpha(\hat{a}_{\mathrm{I}}^\dagger\hat{a}_{\mathrm{I}}^\dagger - \hat{a}_{\mathrm{S}}^\dagger\hat{a}_{\mathrm{I}}^\dagger)\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}} + \beta(\hat{a}_{\mathrm{S}}^\dagger - \hat{a}_{\mathrm{I}}^\dagger)\hat{a}_{\mathrm{I}}$.

$$\frac{dx_{\mathrm{SI}}}{dt} = \langle \mathbf{1}| \left[\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}, \alpha(\hat{a}_{\mathrm{I}}^\dagger\hat{a}_{\mathrm{I}}^\dagger - \hat{a}_{\mathrm{S}}^\dagger\hat{a}_{\mathrm{I}}^\dagger)\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}} + \beta(\hat{a}_{\mathrm{S}}^\dagger - \hat{a}_{\mathrm{I}}^\dagger)\hat{a}_{\mathrm{I}}\right] |\phi(t)\rangle \tag{262}$$

$$\begin{aligned} &= \alpha\langle \mathbf{1}|[\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}, \hat{a}_{\mathrm{I}}^\dagger\hat{a}_{\mathrm{I}}^\dagger]\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle - \alpha\langle \mathbf{1}|[\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}, \hat{a}_{\mathrm{S}}^\dagger\hat{a}_{\mathrm{I}}^\dagger]\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle \\ &\quad + \beta\langle \mathbf{1}|[\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}, \hat{a}_{\mathrm{S}}^\dagger]\hat{a}_{\mathrm{I}}|\phi(t)\rangle - \beta\langle \mathbf{1}|[\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}, \hat{a}_{\mathrm{I}}^\dagger]\hat{a}_{\mathrm{I}}|\phi(t)\rangle \end{aligned} \tag{263}$$

$$\begin{aligned} &= 2\alpha\langle \mathbf{1}|\hat{a}_{\mathrm{I}}^\dagger\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle - \alpha\langle \mathbf{1}|(\hat{a}_{\mathrm{S}}^\dagger\hat{a}_{\mathrm{S}} + \hat{a}_{\mathrm{I}}^\dagger\hat{a}_{\mathrm{I}} + 1)\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle \\ &\quad + \beta\langle \mathbf{1}|\hat{a}_{\mathrm{I}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle - \beta\langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle \end{aligned} \tag{264}$$

$$= \alpha\langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle - \alpha\langle \mathbf{1}|\hat{a}_{\mathrm{I}}\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle + \beta\langle \mathbf{1}|\hat{a}_{\mathrm{I}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle - (\alpha+\beta)\langle \mathbf{1}|\hat{a}_{\mathrm{S}}\hat{a}_{\mathrm{I}}|\phi(t)\rangle$$

$$= \alpha x_{\mathrm{SSI}}(t) - \alpha x_{\mathrm{ISI}}(t) + \beta x_{\mathrm{II}}(t) - (\alpha+\beta)x_{\mathrm{SI}}(t) \tag{265}$$

So we see that the expectation value for SI-pairs will depend on II-pairs, but also on triples SSI and ISI. Of course, this interdependence will continue, as the expectation value of triples will depend on quartic terms, etc. Several mean-field type of approximations have been proposed to close the set of equations at different levels, i.e. one can approximate the triples as products of pairs and single values. In the end, each higher level will give a slightly better approximation compared to the previous one.

4. *This exercise we will explore Noether's theorem for stochastic processes, based partly on the work of [12, 19]. Noether's theorem relates symmetries to conservation laws. In quantum mechanics, Noether's theorem can loosely be phrased as:*

$$[\hat{O}, \hat{H}] = 0 \Leftrightarrow \partial_t\langle\hat{O}\rangle = 0\,. \tag{266}$$

In words: any observable which commutes with the (quantum) Hamiltonian has an expectation value which does not change in time (i.e. is conserved). The implication works both ways, so any conserved quantity also corresponds to an observable which commutes with the Hamiltonian. (Noether's theorem is, of course, more general than this, but for here we will do with conservation in time.)

(a) *Show that the relation $[\hat{O}, \hat{H}] = 0 \Rightarrow \partial_t\langle\hat{O}\rangle = 0$ also works in stochastic mechanics, so whenever $\hat{H}$ is an infinitesimal generator and $\langle\hat{O}\rangle$ is computed in the L_1 norm as: $\langle\hat{O}\rangle = \langle\mathbf{1}|\hat{O}|\phi(t)\rangle$.*

This is quite straightforward. Basically it follows directly from the time evolution as $\partial_t\langle\hat{O}\rangle = \langle\mathbf{1}|[\hat{O}, \hat{H}]|\phi(t)\rangle$. In fact, $[\hat{O}, \hat{H}] = 0$ implies $\partial_t\langle\hat{O}^n\rangle = 0$ for any n.

Above we have shown that Noether's theorem works one way for stochastic systems. Unfortunately, the converse is not necessarily true. To show this, consider the following reaction network:

$$\mathrm{B} \to \mathrm{A} \tag{267a}$$

$$\mathrm{B} \to \mathrm{C} \tag{267b}$$

where (importantly) both reaction occur with the same rate α.

(b) *Define an observable* $\hat{O} = \hat{N}_B + 2\hat{N}_C$. *Show that in the above system, we have* $\partial_t\langle\hat{O}\rangle = 0$ *even though* $[\hat{O}, \hat{H}] \neq 0$. *Can you also provide a physical explanation as to why this fails in this case?*

Let's just work out the commutator. First, we need the generator, which is given as the following

$$\hat{H} = \alpha(\hat{a}_A^\dagger + \hat{a}_C^\dagger - 2\hat{a}_B^\dagger)\hat{a}_B\,. \tag{268}$$

Then, the commutator is:

$$\begin{aligned}
[\hat{N}_B + 2\hat{N}_C, \hat{H}] &= \alpha[\hat{N}_B + 2\hat{N}_C, (\hat{a}_A^\dagger + \hat{a}_C^\dagger - 2\hat{a}_B^\dagger)\hat{a}_B] && (269)\\
&= \alpha\left([\hat{N}_B, (\hat{a}_A^\dagger + \hat{a}_C^\dagger)\hat{a}_B] + 2[\hat{N}_C, \hat{a}_C^\dagger\hat{a}_B]\right) && (270)\\
&= \alpha\left((\hat{a}_A^\dagger + \hat{a}_C^\dagger)[\hat{N}_B, \hat{a}_B] + 2[\hat{N}_C, \hat{a}_C^\dagger]\hat{a}_B\right) && (271)\\
&= -\alpha(\hat{a}_A^\dagger - \hat{a}_C^\dagger)\hat{a}_B && (272)
\end{aligned}$$

So, the commutator is not zero, but when we compute the expected value of the commutator, we may replace creation operators by 1 and see that the expectation value vanishes, so $\partial_t\langle\hat{O}\rangle = 0$. However, if we work out the variance in $\hat{O}$, we see actually that $\partial_t\langle\hat{O}^2\rangle \neq 0$.

(c) *Is the second moment of the observable* $\hat{O}$ *(defined as* $\langle\hat{O}^2\rangle$*) also conserved? I.e. what is* $\partial_t\langle\hat{O}^2\rangle$*?*

The time derivative of the second moment of $\hat{O}$ is computed as follows:

$$\begin{aligned}
\partial_t\hat{O}^2 &= \langle\mathbf{1}|[\hat{O}^2, \hat{H}]|\phi(t)\rangle = \langle\mathbf{1}|[\hat{O}, \hat{H}]\hat{O} + \hat{O}[\hat{O}, \hat{H}]|\phi(t)\rangle && (273)\\
&= -\alpha\langle\mathbf{1}|(\hat{a}_A^\dagger - \hat{a}_C^\dagger)\hat{a}_B(\hat{N}_B + 2\hat{N}_C) + (\hat{a}_B + 2\hat{a}_C)(\hat{a}_A^\dagger - \hat{a}_C^\dagger)\hat{a}_B|\phi(t)\rangle && (274)\\
&= -\alpha\langle\mathbf{1}|2\hat{a}_C(\hat{a}_A^\dagger - \hat{a}_C^\dagger)\hat{a}_B|\phi(t)\rangle && (275)\\
&= -2\alpha\langle\mathbf{1}|\hat{a}_B|\phi(t)\rangle = -2\alpha\langle N_B(t)\rangle\,. && (276)
\end{aligned}$$

Here, in going from the second to the third line, we are using that $\langle\mathbf{1}|(\hat{a}_A^\dagger - \hat{a}_C^\dagger) = 0$, so that the only surviving contribution comes from normal ordering the C operators. The final result is clearly not zero if $\langle N_B(t)\rangle \neq 0$, so the variance will grow proportionally to the number of B elements in the system.

(d) *Can you provide a physical explanation as to why Noethers theorem fails in this case?*

According to the reaction network, B decays into A or C with equal probability, so if we start with any B, after a sufficient amount of time has passed, we have 50% chance of ending up with an A and 50% chance of having C. The observable $\hat{O}$ counts B's and it add 2 times C, so its expectation will not change under the reaction, it either counts 100% B or two times 50% C. However, the variance does chance, as we are not sure whether B has decayed to A or B, and so Noether theorems fails for this case.

The above situation shows that we need supplementary conditions to show that the conservation of a quantity leads to a observable which commutes with the Hamiltonian.

In [19] it was shown that, for instance $\partial_t\langle\hat{O}^2\rangle = 0$ is sufficient and so that the Noether's theorem for stochastic mechanics can be formulated as

$$[\hat{O},\hat{H}] = 0 \Leftrightarrow \partial_t\langle\hat{O}\rangle = 0 = \partial_t\langle\hat{O}^2\rangle\,. \tag{277}$$

(*) *Proof Noether's theorem for stochastic mechanics. You may use that observables in stochastic mechanics are 'diagonal' in the sense that we can always expand them as $\hat{O} = \sum_n \frac{1}{n!} O_n |n\rangle\langle n|$.*

Above we already showed that $[\hat{O},\hat{H}] = 0$ implies $\partial_t\langle\hat{O}^n\rangle = 0$ for any n. All that remains to show is the converse, i.e. that $\partial_t\langle\hat{O}\rangle = 0 = \partial_t\langle\hat{O}^2\rangle$ implies $[\hat{O},\hat{H}] = 0$.

To do so, let's write out the components of $[\hat{O},\hat{H}]$ explicitly and insert a resolution of the identity:

$$\begin{aligned}
\langle\mathbf{m}|[\hat{O},\hat{H}]|\mathbf{n}\rangle &= \langle\mathbf{m}|\hat{O}\hat{H} - \hat{H}\hat{O}|\mathbf{n}\rangle && (278)\\
&= \sum_{\mathbf{k}} \frac{1}{\mathbf{k}!}\langle\mathbf{m}|\hat{O}|\mathbf{k}\rangle\langle\mathbf{k}|\hat{H}|\mathbf{n}\rangle - \langle\mathbf{m}|\hat{H}|\mathbf{k}\rangle\langle\mathbf{k}|\hat{O}|\mathbf{n}\rangle && (279)\\
&= (O_{\mathbf{m}} - O_{\mathbf{n}})\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle && (280)
\end{aligned}$$

Here, in the last line, we used the diagonality of $\hat{O} = \sum_{\mathbf{n}} \frac{1}{\mathbf{n}!} O_{\mathbf{n}}|\mathbf{n}\rangle\langle\mathbf{n}|$. For above to vanish, we see that either $\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle = 0$, i.e. there is no transition between the states $\mathbf{m}$ and $\mathbf{n}$, or the value of the operator $\hat{O}$ in the state $\mathbf{m}$ and $\mathbf{n}$ should be the same, such that $(O_{\mathbf{m}} - O_{\mathbf{n}})$ vanishes.

Now consider the following sum:

$$\sum_{\mathbf{m}} \frac{1}{\mathbf{m}!}(O_{\mathbf{m}} - O_{\mathbf{n}})^2\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle\,. \tag{281}$$

This infinite sum over $\mathbf{n}$ actually only consists out of positive terms, because if $\mathbf{m} = \mathbf{n}$, then the factor $(O_{\mathbf{m}} - O_{\mathbf{n}})^2$ vanishes, whereas if $\mathbf{m} \neq \mathbf{n}$, both $(O_{\mathbf{m}} - O_{\mathbf{n}})^2$ and $\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle$ are positive. So if the above sum is zero, then each term in the sum is zero, which is sufficient to show that $(O_{\mathbf{m}} - O_{\mathbf{n}})\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle = 0$ for all $\mathbf{m},\mathbf{n}$.

Next we expand the square in the sum above and use the identity $\langle\mathbf{1}| = \sum_{\mathbf{m}} \frac{1}{\mathbf{m}!}\langle\mathbf{m}|$:

$$\begin{aligned}
\sum_{\mathbf{m}} \frac{1}{\mathbf{m}!}(O_{\mathbf{m}} - O_{\mathbf{n}})^2\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle &= \sum_{\mathbf{m}} \frac{1}{\mathbf{m}!}(O_{\mathbf{m}}^2\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle - 2O_{\mathbf{n}}O_{\mathbf{m}}\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle + O_{\mathbf{n}}^2\langle\mathbf{m}|\hat{H}|\mathbf{n}\rangle)\\
&= \langle\mathbf{1}|\hat{O}^2\hat{H}|\mathbf{n}\rangle - 2O_{\mathbf{n}}\langle\mathbf{1}|\hat{O}\hat{H}|\mathbf{n}\rangle + O_{\mathbf{n}}^2\langle\mathbf{1}|\hat{H}|\mathbf{n}\rangle \quad (282)\\
&= 0\,.
\end{aligned}$$

Here the last equality follows by the assumption that $\partial_t\langle\mathbf{1}|\hat{O}^2|\mathbf{n}\rangle = 0 = \partial_t\langle\mathbf{1}|\hat{O}|\mathbf{n}\rangle$, together with the property of the infinitesimal generator $\langle\mathbf{1}|\hat{H} = 0$. This concludes the proof.

3.5.1 Solutions

1. *Proof the definition of the probability generating function* (60) *using the definition of* $\langle\mathbf{z}|$ (28) *and the definition of* $|\phi(t)\rangle$.

This is quite straightforward, if one uses the fact that the dual coherent state is a (left) eigenvector of the creation operator: $\langle \mathbf{z}|\hat{a}_i^\dagger = z_i\langle \mathbf{z}|$. It then follows immediately that:

$$\langle \mathbf{z}|\phi(t)\rangle = \sum_{\mathbf{n}} P(\mathbf{n},t)\langle \mathbf{z}|\prod_i (\hat{a}_i^\dagger)^{n_i}|0\rangle = \sum_{\mathbf{n}} P(\mathbf{n},t)\prod_i z_i^{n_i} = G(\mathbf{z},t)\,. \tag{283}$$

2. *Proof the identification of* (65) *for the probability generating function by showing that* $\langle z|\hat{a}^\dagger|\phi(t)\rangle = z\langle z|\phi(t)\rangle$ *and* $\langle z|\hat{a}|\phi(t)\rangle = \frac{d}{dz}\langle z|\phi(t)\rangle$

   The first equation $\langle z|\hat{a}^\dagger|\phi(t)\rangle = z\langle z|\phi(t)\rangle$ follows directly from $\langle \mathbf{z}|\hat{a}_i^\dagger = z_i\langle \mathbf{z}|$, which we have used above. What remains to be shown is $\langle z|\hat{a}|\phi(t)\rangle = \frac{d}{dz}\langle z|\phi(t)\rangle$. we can do this by expanding the definition of $\langle \mathbf{z}|$ and applying one annihilation operator from the right:

$$\langle z|\hat{a} = \sum_{n=0}^{\infty}\langle 0|\frac{1}{n!}(z\hat{a})^n\hat{a} = \sum_{n=0}^{\infty}\langle 0|\frac{1}{n!}z^n\hat{a}^{n+1} \tag{284}$$

$$= \sum_{n=0}^{\infty}\langle 0|\frac{d}{dz}\left(\frac{1}{(n+1)!}z^{n+1}\right)\hat{a}^{n+1} = \frac{d}{dz}\langle 0|\left(\sum_{n=0}^{\infty}\frac{1}{(n+1)!}(z\hat{a})^{n+1}\right) \tag{285}$$

$$= \frac{d}{dz}\left(\sum_{n=0}^{\infty}\langle 0|\frac{1}{n!}(z\hat{a})^n\right) - \frac{d}{dz}\langle 0| = \frac{d}{dz}\langle z|\,. \tag{286}$$

   This is true regardless of $|\phi(t)\rangle$, which does not depend on z itself, so also $\langle z|\hat{a}|\phi(t)\rangle = \frac{d}{dz}\langle z|\phi(t)\rangle$. The generalization to multiple species is straightforward.

3. *Derive the differential equations for the generating functions corresponding to the processes of 3.1.1 question 1.*

   (a) SIR process:

$$\partial_t G(z_S,z_I,z_R;t) = \alpha(z_I^2 - z_I z_S)\partial_{z_I}\partial_{z_S}G(\mathbf{z};t) + \beta(z_R - z_I)\partial_{z_I}G(\mathbf{z};t)\,. \tag{287}$$

   (b) Voter model:

$$\partial_t G(z_A,z_B;t) = (\alpha z_A^2 + \beta z_B^2 - (\alpha+\beta)z_A z_B)\partial_{z_A}\partial_{z_B}G(\mathbf{z};t)\,. \tag{288}$$

   (c) Predator-Prey model:

$$\partial_t G(z_R,z_F;t) = \alpha(z_R^2 - z_R)\partial_{z_R}G(\mathbf{z};t) + \beta(z_F^2 - z_R z_F)\partial_{z_R}\partial_{z_F}G(\mathbf{z};t) \tag{289}$$

$$+ \gamma(1-z_F)\partial_{z_F}G(\mathbf{z};t)\,. \tag{290}$$

4. *Proof that* (74) *and* (75) *are equivalent, using a corollary of the Campbell-Baker-Hausdorf formula:*

$$e^X e^Y = e^Y \exp\left(\sum_{n=0}^{\infty}\frac{1}{n!}[X,Y^{(n)}]\right)\,, \tag{291}$$

   where $[X,Y^{(n)}] = [\ldots[[X,\underbrace{Y],Y],\ldots Y}_{n}]$ *is the nested commutator*

We are asked to show that:

$$\langle \mathbf{1}|e^{\sum_i s_i \hat{N}_i}|\phi(t)\rangle = \langle 0|e^{\sum_i e^{s_i}\hat{a}_i}|\phi(t)\rangle \tag{292}$$

To do so, we write out $\langle \mathbf{1}| = \langle 0|e^{\sum_i \hat{a}_i}$ and use the given identity to work out the commutators in the exponent

$$\langle \mathbf{1}|e^{\sum_i s_i \hat{N}_i}|\phi(t)\rangle = \langle 0|e^{\sum_j \hat{a}_j} e^{\sum_i s_i \hat{N}_i}|\phi(t)\rangle \tag{293}$$

$$= \langle 0|e^{\sum_i s_i \hat{N}_i} \exp\left(\sum_{n=0}^{\infty} \frac{1}{n!}\left[\sum_j \hat{a}_j, \left(\sum_i s_i \hat{N}_i\right)^{(n)}\right]\right)|\phi(t)\rangle \tag{294}$$

Let's work out the commutators in the exponential:

$$\left[\sum_j \hat{a}_j, \left(\sum_i s_i \hat{N}_i\right)^{(0)}\right] = \sum_i s_i \hat{N}_i \tag{295}$$

$$\left[\sum_j \hat{a}_j, \left(\sum_i s_i \hat{N}_i\right)^{(1)}\right] = \sum_{i,j} s_i \delta_{ij} \hat{a}_j = \sum_i s_i \hat{a}_i \tag{296}$$

$$\left[\sum_j \hat{a}_j, \left(\sum_i s_i \hat{N}_i\right)^{(2)}\right] = \left[\left[\sum_j \hat{a}_j, \left(\sum_i s_i \hat{N}_i\right)^{(1)}\right], \sum_k s_k \hat{N}_k\right] = \sum_i s_i^2 \hat{a}_i \tag{297}$$

Now we see that the pattern repeats and

$$\left[\sum_j \hat{a}_j, \left(\sum_i s_i \hat{N}_i\right)^{(n)}\right] = \sum_i s_i^n \hat{a}_i \tag{298}$$

This leads to the following equalities above:

$$\langle \mathbf{1}|e^{\sum_i s_i \hat{N}_i}|\phi(t)\rangle = \langle 0|e^{\sum_i s_i \hat{N}_i} \exp\left(\sum_i \sum_{n=0}^{\infty} \frac{1}{n!} s_i^n \hat{a}_i\right)|\phi(t)\rangle \tag{299}$$

$$= \langle 0|e^{\sum_i e^{s_i}\hat{a}_i}|\phi(t)\rangle\,. \tag{300}$$

The last equality follows from $\langle 0|e^{\sum_i s_i \hat{N}_i} = \langle 0|$ by definition of the action of the creation operator on the dual vacuum state $\langle 0|\hat{a}_i^\dagger = 0$.

5.* *Using a formula similar to* (291)*, proof* (76)*.*

The identity we wish to use now is the BCH formula the other way around:

$$e^X e^Y = \exp\left(\sum_{n=0}^{\infty} \frac{1}{n!}[X^{(n)}, Y]\right) e^X\,, \tag{301}$$

Then, we may write:

$$\langle \mathbf{1}|e^{\sum_i s_i \hat{N}_i}|\phi(t)\rangle = \langle \mathbf{1}|e^{\sum_i s_i \hat{N}_i} e^{\hat{H}t}|\phi(0)\rangle \tag{302}$$

$$= \langle \mathbf{1}| \exp\left(\sum_{n=0}^{\infty} \frac{[(\sum_i s_i \hat{N}_i)^{(n)}, \hat{H}t]}{n!}\right) e^{\sum_i s_i \hat{N}_i}|\phi(0)\rangle \tag{303}$$

The general Hamiltonian was given in (37) and is repeated here:

$$\hat{H} = \sum_{\tau} r(\tau) \left[(\hat{\mathbf{a}}^{\dagger})^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^{\dagger})^{\mathbf{m}(\tau)} \right] \hat{\mathbf{a}}^{\mathbf{m}(\tau)} . \tag{304}$$

Let's work out the commutators with the number operators:

$$[(\sum_i s_i \hat{N}_i)^{(0)}, \hat{H}t] = \hat{H}t \tag{305}$$

$$[(\sum_i s_i \hat{N}_i)^{(1)}, \hat{H}t] = \sum_{\tau} \sum_i r(\tau) t [\hat{N}_i, \{ (\hat{\mathbf{a}}^{\dagger})^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^{\dagger})^{\mathbf{m}(\tau)} \} \hat{\mathbf{a}}^{\mathbf{m}(\tau)}] \tag{306}$$

$$= \sum_{\tau} \sum_i r(\tau) t s_i (n_i(\tau) - m_i(\tau)) (\hat{\mathbf{a}}^{\dagger})^{\mathbf{n}(\tau)} \hat{\mathbf{a}}^{\mathbf{m}(\tau)} \tag{307}$$

$$[(\sum_i s_i \hat{N}_i)^{(2)}, \hat{H}t] = \sum_{\tau} \sum_{i,j} r(\tau) t s_i (n_i(\tau) - m_i(\tau)) s_j (n_j(\tau) - m_j(\tau)) (\hat{\mathbf{a}}^{\dagger})^{\mathbf{n}(\tau)} \hat{\mathbf{a}}^{\mathbf{m}(\tau)}$$

$$= \sum_{\tau} r(\tau) t \left(\sum_i s_i (n_i(\tau) - m_i(\tau)) \right)^2 (\hat{\mathbf{a}}^{\dagger})^{\mathbf{n}(\tau)} \hat{\mathbf{a}}^{\mathbf{m}(\tau)} \tag{308}$$

$$[(\sum_i s_i \hat{N}_i)^{(n)}, \hat{H}t] = \sum_{\tau} r(\tau) t \left(\sum_i s_i (n_i(\tau) - m_i(\tau)) \right)^n (\hat{\mathbf{a}}^{\dagger})^{\mathbf{n}(\tau)} \hat{\mathbf{a}}^{\mathbf{m}(\tau)} \tag{309}$$

Using this in equation (303) leads to the sought after relationship

$$M_{\hat{\mathbf{N}}}(\mathbf{s}, t) = \langle \mathbf{1} | e^{\sum_i s_i \hat{N}_i} e^{\hat{H}t} | \phi(0) \rangle = \langle \mathbf{1} | e^{\tilde{H}(s)t} e^{\sum_i s_i \hat{N}_i} | \phi(0) \rangle , \tag{310}$$

with:

$$\tilde{H}(s) = \sum_{\tau} r(\tau) \left[e^{\sum_i s_i (n_i(\tau) - m_i(\tau))} (\hat{\mathbf{a}}^{\dagger})^{\mathbf{n}(\tau)} - (\hat{\mathbf{a}}^{\dagger})^{\mathbf{m}(\tau)} \right] \hat{\mathbf{a}}^{\mathbf{m}(\tau)} . \tag{311}$$

4.4 Solutions

1. *Consider a random walk in one dimension, where the probability of moving to the left is not equal to the probability of moving to the right. Call the rate of moving to the left r_- and the rate of moving to the right r_+*

   (a) *What is the continuous-time master equation for $P(x,t)$ in this case?*
   The random walk is now described by a master equation:

   $$\frac{\partial}{\partial t} P(x,t) = [r_+ P(x-h,t) + r_- P(x+h,t) - (r_+ + r_-) P(x,t)] \tag{312}$$

   (b) *Derive the continuum limit $h \to 0$ in this case. When do you get an extra term? What could this extra term signify physically?*
   The continuum limit now gives:

   $$\frac{\partial}{\partial t} P(x,t) = h(r_- - r_+) \partial_x P(x,t) + \frac{1}{2} h^2 (r_+ + r_-) \partial_x^2 P(x,t) + h^3\text{-terms} \tag{313}$$

   The limit $h \to 0$ is finite for the first two terms if $(r_- - r_+) \propto \frac{1}{h}$, but $(r_+ + r_-) \propto 1/h^2$, or if $r_\pm = D/h^2 \pm \gamma/2h$, in that case the limit gives:

   $$\frac{\partial}{\partial t} P(x,t) = -\gamma \partial_x P(x,t) + D \partial_x^2 P(x,t) \tag{314}$$

which is the *convection-diffusion equation*. So the additional term describes convection and γ is the bias velocity.

2. *Show that* $\mathbb{1} = \int \frac{d^2\eta}{\pi} |\eta\rangle\langle\eta|$ *using the identity:*

$$\delta_{n,m} = \frac{1}{\pi m!} \int d^2\eta e^{-|\eta|^2} \eta^{*m} \eta^n \,. \tag{315}$$

Bonus points if you can proof this identity (Hint: decompose the complex number $\eta = re^{i\theta}$ *and remember the integral representation of the Gamma function)*

Let's work it out step by step using the definitions:

$$\mathbb{1} = \int \frac{d^2\eta}{\pi} |\eta\rangle\langle\eta| \tag{316}$$

$$= \int \frac{d^2\eta}{\pi} \sum_{m,n} e^{-|\eta|^2} \frac{\eta^n}{n!} |n\rangle\langle m| \frac{\eta^{*m}}{m!} \rightarrow \text{ use identity} \tag{317}$$

$$= \sum_{m,n} \delta_{n,m} \frac{1}{m!} |n\rangle\langle m| = \sum_m \frac{1}{m!} |m\rangle\langle m| = \mathbb{1} \qquad \text{(by equation (21))} \tag{318}$$

3. *What are the field theories for the CRN's of question 3.1.1 question 1?*

From an earlier question, we had already found the generators. Now we simply have to replace annihilation operators as $\hat{a}_i \rightarrow \varphi_i$ and creation operators $\hat{a}_i^\dagger \rightarrow \tilde{\varphi}_i$, add the diffusion terms and integrate over all space and time to obtain the action. Afterwards, one can make the shift $\tilde{\varphi} \rightarrow \bar{\varphi} + 1$.

(a) SIR field theory:

$$S[\tilde{\varphi}, \varphi] = \int d^d x \int_0^{t_f} dt \Big\{ \sum_{X=\{\mathrm{S,I,R}\}} \tilde{\varphi}_X (\partial_t - D_X \nabla^2) \varphi_X - \alpha(\tilde{\varphi}_\mathrm{I}^2 - \tilde{\varphi}_\mathrm{I}\tilde{\varphi}_\mathrm{S})\varphi_\mathrm{I}\varphi_\mathrm{S} - \beta(\tilde{\varphi}_\mathrm{R} - \tilde{\varphi}_\mathrm{I})\varphi_\mathrm{I} \Big\} \tag{319}$$

Or, in terms of the barred fields:

$$S[\bar{\varphi}, \varphi] = \int d^d x \int_0^{t_f} dt \Big\{ \sum_{X=\{\mathrm{S,I,R}\}} \bar{\varphi}_X (\partial_t - D_X \nabla^2) \varphi_X - \alpha(\bar{\varphi}_\mathrm{I} - \bar{\varphi}_\mathrm{S})\varphi_\mathrm{I}\varphi_\mathrm{S} - \alpha(\bar{\varphi}_\mathrm{I}^2 - \bar{\varphi}_\mathrm{I}\bar{\varphi}_\mathrm{S})\varphi_\mathrm{I}\varphi_\mathrm{S} \tag{320}$$

$$- \beta(\bar{\varphi}_\mathrm{R} - \bar{\varphi}_\mathrm{I})\varphi_\mathrm{I} \Big\} \tag{321}$$

(b) Voter model field theory:

$$S[\tilde{\varphi}, \varphi] = \int d^d x \int_0^{t_f} dt \Big\{ \sum_{X=\{\mathrm{A,B}\}} \tilde{\varphi}_X (\partial_t - D_X \nabla^2) \varphi_X - (\alpha\tilde{\varphi}_\mathrm{A}^2 + \beta\tilde{\varphi}_\mathrm{B}^2 - (\alpha+\beta)\tilde{\varphi}_\mathrm{A}\tilde{\varphi}_\mathrm{B})\varphi_\mathrm{A}\varphi_\mathrm{B} \Big\} \tag{322}$$

In terms of barred fields:

$$S[\bar{\varphi},\varphi] = \int d^dx \int_0^{t_f} dt \Big\{ \sum_{X=\{\mathrm{A,B}\}} \bar{\varphi}_X(\partial_t - D_X\nabla^2)\varphi_X$$
$$- ((\alpha-\beta)\bar{\varphi}_\mathrm{A} + (\beta-\alpha)\bar{\varphi}_\mathrm{B})\varphi_\mathrm{A}\varphi_\mathrm{B} \tag{323}$$
$$- (\alpha\bar{\varphi}_\mathrm{A}^2 + \beta\bar{\varphi}_\mathrm{B}^2 - (\alpha+\beta)\bar{\varphi}_\mathrm{A}\bar{\varphi}_\mathrm{B})\varphi_\mathrm{A}\varphi_\mathrm{B} \Big\} \tag{324}$$

(c) Predator-prey field theory:

$$S[\tilde{\varphi},\varphi] = \int d^dx \int_0^{t_f} dt \Big\{ \sum_{X=\{\mathrm{R,F}\}} \tilde{\varphi}_X(\partial_t - D_X\nabla^2)\varphi_X$$
$$- \alpha(\tilde{\varphi}_\mathrm{R}^2 - \tilde{\varphi}_\mathrm{R})\varphi_\mathrm{R} - \beta(\tilde{\varphi}_\mathrm{F}^2 - \tilde{\varphi}_\mathrm{R}\tilde{\varphi}_\mathrm{F})\varphi_\mathrm{R}\varphi_\mathrm{F} - \gamma(1-\tilde{\varphi}_\mathrm{F})\varphi_\mathrm{F} \Big\} \tag{325}$$

In terms of barred fields:

$$S[\bar{\varphi},\varphi] = \int d^dx \int_0^{t_f} dt \Big\{ \sum_{X=\{\mathrm{R,F}\}} \bar{\varphi}_X(\partial_t - D_X\nabla^2)\varphi_X$$
$$- \alpha(\bar{\varphi}_\mathrm{R}^2 + \bar{\varphi}_\mathrm{R})\varphi_\mathrm{R} - \beta(\bar{\varphi}_\mathrm{F} - \bar{\varphi}_\mathrm{R} + \bar{\varphi}_\mathrm{F}^2 - \bar{\varphi}_\mathrm{R}\bar{\varphi}_\mathrm{F})\varphi_\mathrm{R}\varphi_\mathrm{F} + \gamma\bar{\varphi}_\mathrm{F}\varphi_\mathrm{F} \Big\} \tag{326}$$

4. *Derive the Euler-Lagrange equations from the (time-dependent part of the) general action* (116) *by computing $\delta L/\delta\varphi$ and $\delta L/\delta\tilde{\varphi}$ (where $S = \int d^dxdtL$) and setting the result to zero. Can you find a solution for $\tilde{\varphi}$ and use this to in the equation for φ? What do you find in the end?*

The field equations (or equations of motion) for φ and $\tilde{\phi}$ are derived from functionally differentiating the action w.r.t the fields. By varying w.r.t. φ, we get the equation for the diffusion part from $\tilde{\varphi}$:

$$-(\partial_t + D\nabla^2)\tilde{\varphi} = \frac{\delta H_I}{\delta\varphi} \tag{327}$$

Now we vary the interaction term of the action

to obtain:

$$-(\partial_t + D\nabla^2)\tilde{\varphi} = \sum_\tau \lambda(\tau)k(\tau)\left[\tilde{\varphi}^{\ell(\tau)} - \tilde{\varphi}^{k(\tau)}\right]\varphi^{k(\tau)-1} \tag{328}$$

The solution is trivial: $\tilde{\varphi} = 1$ solves the system and corresponds to using a coherent state representation for the bra's with eigenvalue of 1, or: flat states.

Repeating the variation of the action, now varying w.r.t. $\tilde{\varphi}$, we obtain:

$$(\partial_t - D\nabla^2)\varphi = \sum_\tau \lambda(\tau)\left[\ell(\tau)\tilde{\varphi}^{\ell(\tau)-1} - k(\tau)\tilde{\varphi}^{k(\tau)-1}\right]\varphi^{k(\tau)}\,. \tag{329}$$

Now, plugging in the trivial solution $\tilde{\varphi} = 1$, we obtain the field equations for φ:

$$(\partial_t - D\nabla^2)\varphi = \sum_\tau \lambda(\tau)\left[\ell(\tau) - k(\tau)\right]\varphi^{k(\tau)}\,. \tag{330}$$

This equation should be familiar as the rate equation (now including a diffusion term) for the CRN! Compare for instance with (5) for a single species. Hence, we see the field equations, which give the analogue to the classical saddle point solutions are exactly the mean-field approximations which give the rate equations.

5. *Consider the field shift $\tilde{\varphi} = 1 + \bar{\varphi}$ in the action $H_I(\tilde{\varphi}, \varphi)$. For simplicity, consider just one transition with a fixed $\ell < k$.*

   (a) *What is the interaction term $H_I(\bar{\varphi}, \varphi)$ after the field shift?* The original interaction term was:

   $$H_I(\tilde{\varphi}, \varphi) = \lambda \left[\tilde{\varphi}^\ell - \tilde{\varphi}^k\right] \varphi^k \,. \tag{331}$$

   Now, we perform the field shift $\tilde{\varphi} = \bar{\varphi} + 1$:

   $$H_I(\bar{\varphi}, \varphi) = \lambda \left[(\bar{\varphi} + 1)^\ell - (\bar{\varphi} + 1)^k\right] \varphi^k \,. \tag{332}$$

   This works out to be:

   $$H_I(\bar{\varphi}, \varphi) = \sum_i \lambda_i \bar{\varphi}^i \varphi^k \tag{333}$$

   with the interaction parameters λ_i related to the original transmission rate λ as:

   $$\lambda_i = \lambda \left[\binom{\ell}{i} - \binom{k}{i}\right], \qquad i \leq \ell \tag{334}$$

   $$\lambda_i = -\lambda \binom{k}{i} \qquad i > \ell \tag{335}$$

   Most notably, $\lambda_0 = 0$ and so in the barred representation, there are no vertices where no barred fields are created as outgoing state.

   (b) *Derive the Euler-Lagrange equations for φ and $\bar{\varphi}$ by varying the action with respect to $\delta\bar{\varphi}$ and $\delta\varphi$ and setting the resulting variation to zero. When using the trivial (constant) solution for $\bar{\varphi}$ what is the remaining equation for φ? Does this look familiar?*

   This is essentially a repetition of the previous question, but now with barred fields, the equations are:

   $$-(\partial_t + D\nabla^2)\bar{\varphi} = \sum_i \lambda_i k \bar{\varphi}^i \varphi^{k-1} \tag{336}$$

   $$(\partial_t - D\nabla^2)\varphi = \sum_i \lambda_i i \bar{\varphi}^{i-1} \varphi^k \tag{337}$$

   The trivial solution for $\bar{\varphi}$ now is $\bar{\varphi} = 0$ and the final equation only receives a contribution from the sum over i when $i = 1$. Using this and the expression (334) for λ_i leads to the final equation for φ:

   $$(\partial_t - D\nabla^2)\varphi = \lambda(\ell - k)\varphi^k \tag{338}$$

   (c) *Are the Euler-Lagrange equations for $\varphi(\mathbf{x}, t)$ derived above the same as those derived from the action prior to the field shift?*

   Of course they are the same.

6. *Just as in quantum field theory, the correlation functions in the free theory can be obtained from functional differentiation with respect to a generating functional $Z(\bar{J}, J)$. Here we will derive the generating functional and use it to compute the propagator of the free theory.*

The generating functional in this case is defined as:

$$Z(\bar{J}, J) = \frac{1}{\mathcal{N}} \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi \exp\left[-S_0 + \int d^d x \int dt \left(\bar{J}(\mathbf{x}, t)\varphi(\mathbf{x}, t) + J(\mathbf{x}, t)\bar{\varphi}(\mathbf{x}, t)\right)\right] \tag{339}$$

where $\mathcal{N} = \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi e^{-S_0}$ *is a normalization factor. The purpose of this exercise is to show that:*

$$Z(\bar{J}, J) = \exp\left\{\int d^d x d^d x' dt dt' \bar{J}(\mathbf{x}, t) G(\mathbf{x} - \mathbf{x}', t - t') J(\mathbf{x}', t')\right\} \tag{340}$$

and use this to compute the free theory correlation functions $\langle\varphi(x, t)\bar{\varphi}(x', t')\rangle_0$ *by taking functional derivatives with respect to* J *and* $\bar{J}$*.*

First, we define two Green's functions, G *and* $\bar{G}$ *such that:*

$$(\partial_t - D\nabla^2) G(\mathbf{x} - \mathbf{x}', t - t') = \delta^d(\mathbf{x} - \mathbf{x}')\delta(t - t') \tag{341}$$

$$(-\partial_t - D\nabla^2) \bar{G}(\mathbf{x} - \mathbf{x}', t - t') = \delta^d(\mathbf{x} - \mathbf{x}')\delta(t - t') \tag{342}$$

and we recall the short hand notation for the convolution:

$$(f * g)(\mathbf{x}, t) = \int d\mathbf{x}' dt' f(\mathbf{x} - \mathbf{x}', t - t') g(\mathbf{x}', t')\,. \tag{343}$$

(a) *By using the field shifts* $\varphi \to \varphi + G * J$ *and* $\bar{\varphi} \to \bar{\varphi} + \bar{G} * \bar{J}$*, show that the generating functional* (339) *can be expressed as* (340)

Let's express the exponent of the path integral (339) after the field shift, schematically:

$$\int \Big\{ - (\bar{\varphi} + \bar{G} * \bar{J})(\partial_t - D\nabla^2)(\varphi + G * J) + \bar{J}(\varphi + G * J) + J(\bar{\varphi} + \bar{G} * \bar{J}) \Big\} \tag{344}$$

We can expand this expression and then use some simplifications using integration by parts. For instance, we have that:

$$\int \bar{G} * \bar{J}(\partial_t - D\nabla^2)\varphi = \int \varphi(-\partial_t - D\nabla^2)\bar{G} * \bar{J} = \int \varphi \bar{J} \tag{345}$$

Or, if you don't believe it, we can work this one out explicitly with all the integration variables (denoting now $x = \{\mathbf{x}, t\}$ to make the formulas at least a bit shorter)

$$\int dx \int dx'\, \bar{G}(x - x')\bar{J}(x')(\partial_t - D\nabla^2)\varphi(x) \tag{346}$$

$$= \int dx \int dx'\, (-\partial_t - D\nabla^2)[\bar{G}(x - x')\bar{J}(x')]\varphi(x) \tag{347}$$

$$= \int dx \int dx'\, \delta(x - x')\bar{J}(x')\varphi(x) = \int dx\, \varphi(x)\bar{J}(x) \tag{348}$$

Using this relation and similar ones, we can work out and simplify the action after the shift:

$$\begin{aligned}
&\int \Big\{ -(\bar\varphi + \bar G * \bar J)(\partial_t - D\nabla^2)(\varphi + G * J) + \bar J(\varphi + G * J) + J(\bar\varphi + \bar G * \bar J)\Big\} \\
&= \int \Big\{ -\bar\varphi(\partial_t - D\nabla^2)\varphi - \bar G * \bar J(\partial_t - D\nabla^2)\varphi - \bar\varphi(\partial_t - D\nabla^2)G * J \qquad (349)\\
&\qquad\qquad - \bar G * \bar J(\partial_t - D\nabla^2)G * J + \bar J(\varphi + G * J) + J(\bar\varphi + \bar G * \bar J)\Big\} \\
&= \int \Big\{ -\bar\varphi(\partial_t - D\nabla^2)\varphi - \bar J\varphi - \bar\varphi J - (\bar G * \bar J)J + \bar J\varphi + \bar J G * J + J\bar\varphi + J\bar G * \bar J\Big\} \\
&= \int \Big\{ -\bar\varphi(\partial_t - D\nabla^2)\varphi + \bar J G * J\Big\} \qquad (350)
\end{aligned}$$

Using this inside the path integral and noticing that the second term does not depend on $\bar\varphi$ and φ any more, we may write

$$Z(\bar J, J) = \frac{1}{\mathcal{N}} \int \mathcal{D}\bar\varphi \mathcal{D}\varphi e^{-S_0[\bar\varphi,\varphi]} e^{\int dx \int dx' \bar J(x) G(x-x') J(x')} \qquad (351)$$

The integration over e^{-S_0} is conveniently canceled by the normalization factor and we end up with the result stated in (340).

(b) *Using* (340)*, proof equation* (123) *by functional differentiation with respect to* J *and* $\bar J$ *and afterward setting* $J, \bar J = 0$

The correlation function is now easy to compute. By using the definition (339), it is obvious that:

$$\begin{aligned}
\langle\varphi(x)\bar\varphi(x')\rangle_0 &= \frac{1}{\mathcal{N}} \int \mathcal{D}\bar\varphi\mathcal{D}\varphi\ \varphi(x)\bar\varphi(x') e^{-S_0} \qquad (352)\\
&= \frac{\delta}{\delta\bar J(x)} \frac{\delta}{\delta J(x')} Z[\bar J, J]\Big|_{J,\bar J=0} \qquad (353)
\end{aligned}$$

Using the end results (340), we can compute the functional derivatives:

$$\begin{aligned}
\langle\varphi(x)\bar\varphi(x')\rangle_0 &= \frac{\delta}{\delta\bar J(x)} \frac{\delta}{\delta J(x')} e^{\int dy dy' \bar J(y) G(y-y') J(y')}\Big|_{J,\bar J=0} \qquad (354)\\
&= \frac{\delta}{\delta\bar J(x)} \int dy\, \bar J(y) G(y-x') e^{\int dy dy' \bar J(y) G(y-y') J(y')}\Big|_{J,\bar J=0} \qquad (355)\\
&= \left(G(x-x') + \int dy\, \bar J(y) G(y-x') \frac{\delta}{\delta\bar J(x)} \right) e^{\int dy dy' \bar J(y) G(y-y') J(y')}\Big|_{J,\bar J=0} \\
&= G(x-x') \qquad (356)
\end{aligned}$$

(c) *Using* (340)*, compute the four point correlation function* $\langle\varphi_1\varphi_2\bar\varphi_3\bar\varphi_4\rangle_0$ *in the free theory (in terms of propagators* G*), where the subscripts simply indicate that each field is to be taken at different spacetime points* $\varphi_1 \equiv \varphi(\mathbf{x}_1, t_1)$*. Can you express the answer in terms of Feynman diagrams?*

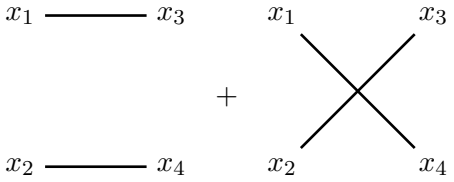
We're going to express the functional dependence of the J's in this answer also by lower case indices, meaning $J_1 = J(x_1) = J(\mathbf{x}_1, t_1)$ etc...

The sought after correlator is computed by performing four functional derivatives with respect to the source fields:

$$\langle\varphi_1\varphi_2\bar{\varphi}_3\bar{\varphi}_4\rangle_0 = \frac{\delta}{\delta\bar{J}_1}\frac{\delta}{\delta\bar{J}_2}\frac{\delta}{\delta J_3}\frac{\delta}{\delta J_4}Z[\bar{J},J]\Big|_{\bar{J},J=0} \tag{357}$$

$$\begin{aligned}
&= \frac{\delta}{\delta\bar{J}_1}\frac{\delta}{\delta\bar{J}_2}\frac{\delta}{\delta J_3}\int dy\bar{J}(y)G(y-x_4)Z[\bar{J},J]\Big|_{\bar{J},J=0} \\
&= \frac{\delta}{\delta\bar{J}_1}\frac{\delta}{\delta\bar{J}_2}\int dy\bar{J}(y)G(y-x_4)\int dy'\bar{J}(y')G(y'-x_3)Z[\bar{J},J]\Big|_{\bar{J},J=0} \\
&= \frac{\delta}{\delta\bar{J}_1}\Big(G(x_2-x_4)\int dy'\bar{J}(y')G(y'-x_3) \\
&\quad + \int dy\bar{J}(y)G(y-x_4)G(x_2-x_3)+\dots\Big)Z[\bar{J},J]\Big|_{\bar{J},J=0} \\
&= \left(G(x_2-x_4)G(x_1-x_3)+G(x_1-x_4)G(x_2-x_3)+\dots\right)Z[\bar{J},J]\big|_{\bar{J},J=0} \\
&= G(x_2-x_4)G(x_1-x_3)+G(x_1-x_4)G(x_2-x_3)
\end{aligned} \tag{358}$$

Here the dots above denote terms which will vanish when setting $\bar{J}, J = 0$. The result can be expressed in terms of the free propagator Feynman diagrams as:

$\langle\varphi_1\varphi_2\bar{\varphi}_3\bar{\varphi}_4\rangle_0 =$ 

(359)

Note that when comparing with free scalar field theory, there is one diagram missing, the one which connects x_1 with x_2 and x_3 with x_4. Hence, the rules for connecting free theory diagrams (i.e. Wicks theorem) is slightly different here: only barred fields are paired with unbarred fields! Due to this feature, any free standing φ in the action which cannot be paired up with an earlier $\bar{\varphi}$ field can safely be dropped from the action.

5.3 Solutions

The final exercises are leading up to a one-loop calculation for the SIS model, which is not easy to do, so let's take it step by step. First, we have to know how to deal with the field theory when we have single-species reactants (i.e. $kA \to \ell A$ with $k = 1$). In this case, the propagator gets extra contributions from terms bilinear in the fields. In physics jargon these are mass terms, and they alter the propagator of the free, reactionless theory.

1. *First consider a single species A subject to a death process: $A \to \emptyset$. We have already solved the homogeneously mixed (space independent) part in earlier exercises and since each reaction is independent of the presence of other A particles, also the field theory should be simple.*

   (a) *What is the field theory action for this reaction? Write down the action with reaction rate λ in terms of $\bar{\varphi}$-fields*

The answer is simply

$$S[\bar{\varphi},\varphi] = \int d^dx \left\{ \int_0^{t_f} dt \ \left[\bar{\varphi}(\partial_t - D\nabla^2)\varphi + \lambda\bar{\varphi}\varphi \right] - n_0\bar{\varphi}(0) \right\} \tag{360}$$

(b) *What is the propagator for this action? Hint: Fourier transform the fields. The propagator in momentum space is the inverse of the bilinear term*

Using the Fourier transformed form of the action, we have:

$$\int \frac{d^dp}{(2\pi)^d} \int \frac{d\omega}{2\pi} \, \bar{\varphi}(-\mathbf{p},-\omega)(-i\omega + D^2p^2 + \lambda)\varphi(\mathbf{p},\omega) \, . \tag{361}$$

This gives the propagator:

$$G(\mathbf{p},\omega) = \frac{1}{-i\omega + Dp^2 + \lambda} \, , \tag{362}$$

Now, inverting the frequency Fourier transform gives an additional exponential decay term:

$$G(\mathbf{p},t) = e^{-Dp^2t}e^{-\lambda t}\Theta(t) \, . \tag{363}$$

Which leads to the real space propagator:

$$G(\mathbf{x},t) = \frac{1}{(4\pi Dt)^{d/2}} e^{-x^2/(4Dt)} e^{-\lambda t}\Theta(t) \, . \tag{364}$$

(c) *Compute the expected density of φ as $\langle \varphi(\mathbf{k}=0,t) \rangle$ in this theory. Do you get any loop corrections?*

The answer is the path integral:

$$\langle \varphi(\mathbf{k}=0,t) \rangle = \frac{1}{\mathcal{N}} \int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi \ \varphi(0,t) e^{-S} e^{n_0\bar{\varphi}(0,0)} \tag{365}$$

$$= G(0,t)n_0 = e^{-\lambda t}n_0 \tag{366}$$

Which is the same result as computed from mean-field theory. There are no loop corrections, the free theory is exact.

2. *Consider now two species A and B, where A particles can decay into B particles with rate λ: $A \to B$.*

(a) *What is the field theory action in this case? Write down the action in terms of $\bar{\varphi}$-fields*

The field theory action now depends on two types of fields, φ_A and φ_B and their barred counterparts. Assuming they have equal diffusion constant:

$$S[\bar{\varphi},\varphi] = \int d^dx \Big\{ \int_0^{t_f} dt \ \left[\bar{\varphi}_A(\partial_t - D\nabla^2)\varphi_A + \bar{\varphi}_B(\partial_t - D\nabla^2)\varphi_B - \lambda(\bar{\varphi}_B - \bar{\varphi}_A)\varphi_A \right]$$
$$- n_{A,0}\bar{\varphi}_A(0) - n_{B,0}\bar{\varphi}_B(0) \Big\} \tag{367}$$

$$= \int d^dx \Big\{ \int_0^{t_f} dt \ \left[\bar{\varphi}_A(\partial_t - D\nabla^2 + \lambda)\varphi_A + \bar{\varphi}_B(\partial_t - D\nabla^2)\varphi_B - \lambda\bar{\varphi}_B\varphi_A \right]$$
$$- n_{A,0}\bar{\varphi}_A(0) - n_{B,0}\bar{\varphi}_B(0) \Big\} \tag{368}$$

(b) *Find a linear combination of A and B fields such that all terms in the action are diagonal, i.e. such that the integrand of the action is written in the form $\bar{\varphi}_1 G_1^{-1}\varphi_1 + \bar{\varphi}_2 G_2^{-1}\varphi_2$. What are the propagators for the two fields?*
The action above can be written as:

$$S[\bar{\varphi},\varphi] = \int d^dx \Big\{ \int_0^{t_f} dt \ \left[(\bar{\varphi}_A - \bar{\varphi}_B)(\partial_t - D\nabla^2 + \lambda)\varphi_A + \bar{\varphi}_B(\partial_t - D\nabla^2)(\varphi_A + \varphi_B)\right] - n_{A,0}(\bar{\varphi}_A(0) - \bar{\varphi}_B(0)) - (n_{B,0} + n_{A,0})\bar{\varphi}_B(0) \Big\} \tag{369}$$

So, if we redefine fields as:

$$\bar{\varphi}_1 = \bar{\varphi}_A - \bar{\varphi}_B\,, \qquad \varphi_1 = \varphi_A \tag{370}$$

$$\bar{\varphi}_2 = \bar{\varphi}_B\,, \qquad \varphi_2 = \varphi_A + \varphi_B \tag{371}$$

Together with $n_{A,0} = n_{1,0}$ and $n_{B,0} + n_{A,0} = n_{2,0}$, we see the following diagonal action:

$$S[\bar{\varphi},\varphi] = \int d^dx \Big\{ \int_0^{t_f} dt \ \left[\bar{\varphi}_1(\partial_t - D\nabla^2 + \lambda)\varphi_1 + \bar{\varphi}_2(\partial_t - D\nabla^2)\varphi_2\right] - n_{1,0}\bar{\varphi}_1(0) - n_{2,0}\bar{\varphi}_2(0) \Big\} \tag{372}$$

And the propagators are:

$$G_1(\mathbf{k},t) = e^{-Dk^2t - \lambda t}\Theta(t)\,, \tag{373}$$

$$G_2(\mathbf{k},t) = e^{-Dk^2t}\Theta(t)\,, \tag{374}$$

(c) *Compute the density of A from the path integral. Do you get loop corrections?*
Again, there are no loop corrections, and the density of A fields directly corresponds to the expectation value of φ_1. So:

$$\langle\varphi_1(\mathbf{k}=0,t)\rangle = \frac{1}{\mathcal{N}}\int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi\ \varphi_1(0,t)e^{-S}e^{n_{1,0}\bar{\varphi}_1(0,0)} \tag{375}$$

$$= G_1(0,t)n_{1,0} = e^{-\lambda t}n_{A,0} \tag{376}$$

Or the A particles decay exponentially, which is to be expected. For the density of B particles, we have to invert the relations (370) to compute $\langle\varphi_B\rangle = \langle\varphi_2 - \varphi_1\rangle$. It gives:

$$\langle\varphi_2 - \varphi_1\rangle = \frac{1}{\mathcal{N}}\int \mathcal{D}\bar{\varphi}\mathcal{D}\varphi\ (\varphi_2 - \varphi_1)e^{-S}e^{n_{1,0}\bar{\varphi}_1(0,0) + n_{2,0}\bar{\varphi}_2(0,0)} \tag{377}$$

$$= G_2(0,t)n_{2,0} - G_1(0,t)n_{1,0} = n_{B,0} + (1 - e^{-\lambda t})n_{A,0}\,. \tag{378}$$

So, the B particle density is what is was, plus the contribution form A particles decaying into B's exponentially. Again, we could have gotten all this from mean-field theory as there are no loop corrections, it's just to warm up for the next one and to show that the formalism works!

3. *Now we consider a more interesting example, the SIS model, which was given by the reactions $I \to S$ and $I + S \to 2I$.*

   (a) *Write down the field theory action in terms of $\bar{\varphi}$ fields.*

   First we use the tilde fields to get the interaction terms. The first reaction contributes $-\lambda_r(\tilde{\varphi}_S - \tilde{\varphi}_I)\varphi_I$ to the action, the second reaction contributes: $-\lambda_t(\tilde{\varphi}_I^2 - \tilde{\varphi}_I\tilde{\varphi}_S)\varphi_S\varphi_I$. In terms of the barred fields this becomes:

$$S = \int d^dx \Big\{ \int_0^{t_f} dt \, \Big[\bar{\varphi}_S(\partial_t - D\nabla^2)\varphi_S + \bar{\varphi}_I(\partial_t - D\nabla^2)\varphi_I - \lambda_r(\bar{\varphi}_S - \bar{\varphi}_I)\varphi_I - \lambda_t(1 + \bar{\varphi}_I)(\bar{\varphi}_I - \bar{\varphi}_S)\varphi_S\varphi_I \Big] - n_{I,0}\bar{\varphi}_I(0) - n_{S,0}\bar{\varphi}_S(0) \Big\}$$

   (b) *Similar to the $A \to B$ reaction above, the bilinear terms in the action are not diagonal. Perform the transformation which diagonalizes the bilinear fields and compute the propagators.*

   The same transformation as (370) diagonalizes the above action, but now with $A = I$ and $B = S$, or:

$$\bar{\varphi}_1 = \bar{\varphi}_I - \bar{\varphi}_S \,, \qquad \varphi_1 = \varphi_I \tag{379}$$

$$\bar{\varphi}_2 = \bar{\varphi}_S \,, \qquad \varphi_2 = \varphi_S + \varphi_I \tag{380}$$

   The bilinear part of the action becomes:

$$S_0 = \int d^dx \int_0^{t_f} dt \, \left[\bar{\varphi}_1(\partial_t - D\nabla^2 + \lambda_r)\varphi_1 + \bar{\varphi}_2(\partial_t - D\nabla^2)\varphi_2 \right] . \tag{381}$$

   Such that the propagators are:

$$G_1(\mathbf{k}, t) = e^{-Dk^2t - \lambda_r t}\Theta(t) \,, \tag{382}$$

$$G_2(\mathbf{k}, t) = e^{-Dk^2t}\Theta(t) \,, \tag{383}$$

   (c) *Perform the same field redefinition in the interaction terms. Which terms are present in the action? What vertices do they correspond to?*

   The interaction part of the action in the new fields can be worked out by first inverting the relations (379)

$$\bar{\varphi}_I = \bar{\varphi}_1 + \bar{\varphi}_2 \,, \qquad \varphi_I = \varphi_1 \tag{384}$$

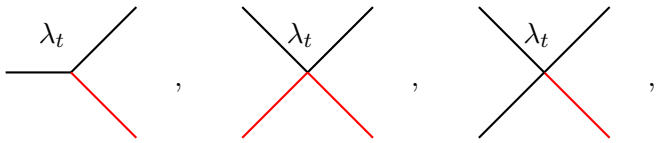
$$\bar{\varphi}_S = \bar{\varphi}_2 \,, \qquad \varphi_S = \varphi_2 - \varphi_1 \tag{385}$$

   Then, the interaction terms become:

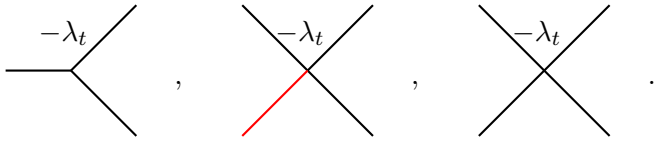
$$S_I = -\lambda_t \int d^dx \int_0^{t_f} dt \; \bar{\varphi}_1(1 + \bar{\varphi}_1 + \bar{\varphi}_2)(\varphi_2 - \varphi_1)\varphi_1 \,, \tag{386}$$

$$= -\lambda_t \int d^dx \int_0^{t_f} dt \; \left\{ (\bar{\varphi}_1 + \bar{\varphi}_1^2 + \bar{\varphi}_1\bar{\varphi}_2)\varphi_2\varphi_1 - (\bar{\varphi}_1 + \bar{\varphi}_1^2 + \bar{\varphi}_1\bar{\varphi}_2)\varphi_1^2 \right\} . \tag{387}$$

This implies that there are two types (or flavors) of fields coupled by six possible vertices. Denoting φ_1 propagators as black lines and φ_2 propagators as red lines, we have the following vertices to play around with:



(388)



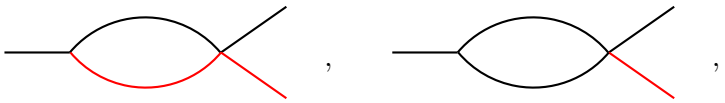
(389)

(d) *Do you expect loop corrections to the propagator? In other words: is it possible to construct loop diagram with only one incoming and one outgoing leg?*

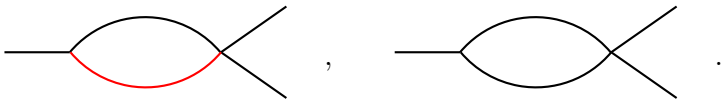
Using the vertices above it is not possible to make a diagram which connects a single ingoing leg to a single outgoing leg. All three point vertices require two incoming legs, so there are no loop corrections to the propagator.

(e) *What are the first loop diagrams you can write down? Can you compute their contributions? When does their contribution diverge?*

There are four one-loop corrections to the three-point vertices of the theory. They are:



(390)



(391)

Their contributions are computed following the Feynman rules and then performing some integrals over the times for the vertices. If all goes well, it will give IR divergences (for $t \to \infty$) when $d \leq 4$.