Hybrid Photonics Computing

Lecturer: Natalia Berloff

Notes by: Piper Fowler-Wright

Lent Term 2020

These notes cover the 16 lecture course on *Hybrid Photonics Computing* by Natalia Berloff. Bear in mind they are my own interpretation of the material and were made during the inaugural running of the course so may not accurately reflect its content in later years.

In addition to the lectures, three example sheets were provided which are available on the Part III Examples page maintained by DATMP.

Contents

1 Overview and Motivation

1.1 Introduction

In this course we address the problem of minimising a given function f of n variables. This is a hard problem in a sense made precise below (Section 3). The conventional algorithm of gradient descent is efficient at finding local minima, but not *global* ones. While quantum computers provide a speed-up, for such computationally expensive problems the gain is insignificant. We look for alternative approaches which may utilise either classical or quantum devices. Our strategy, sketched in Figure 1.1, makes use of the fact that any optimisation problem of this form may be mapped to the problem of finding the global minimum of the spin Hamiltonian of the XY-model.

As illustrated, we will discuss both mathematical models (networks) that realise the minimisation of the Hamiltonian, and physical systems capable of performing the optimisation in reality. The crux of the problem is that this method should obtain a global minima, not just a local one, and do so relatively efficiently.

1.2 Hybrid Photonics Computing

1.2.1 Limitations of Conventional Computing

In previous decades, the need to solve optimisation problems of larger dimensionality in less time has been met by the doubling of the number of transistors on (and so computing power of) microchips roughly every two years – an observation known as *Moore's law.* This is in part made possible by *Dennard scaling,* which describes how despite this increase the density of power consumption remains approximately constant. However, these trends have begun to break down as CMOS, the technology used to fabricate microchips, is reaching its physical limits: oxide layers may now be only several atoms thick with quantum effects such as electron tunnelling becoming problematic, leading to increased power consumption. Moreover, conventional computers based on von Neumann architecture have the fundamental limitation that instructions and data are handled sequentially (rather than simultaneously) by the central processing unit (CPU).

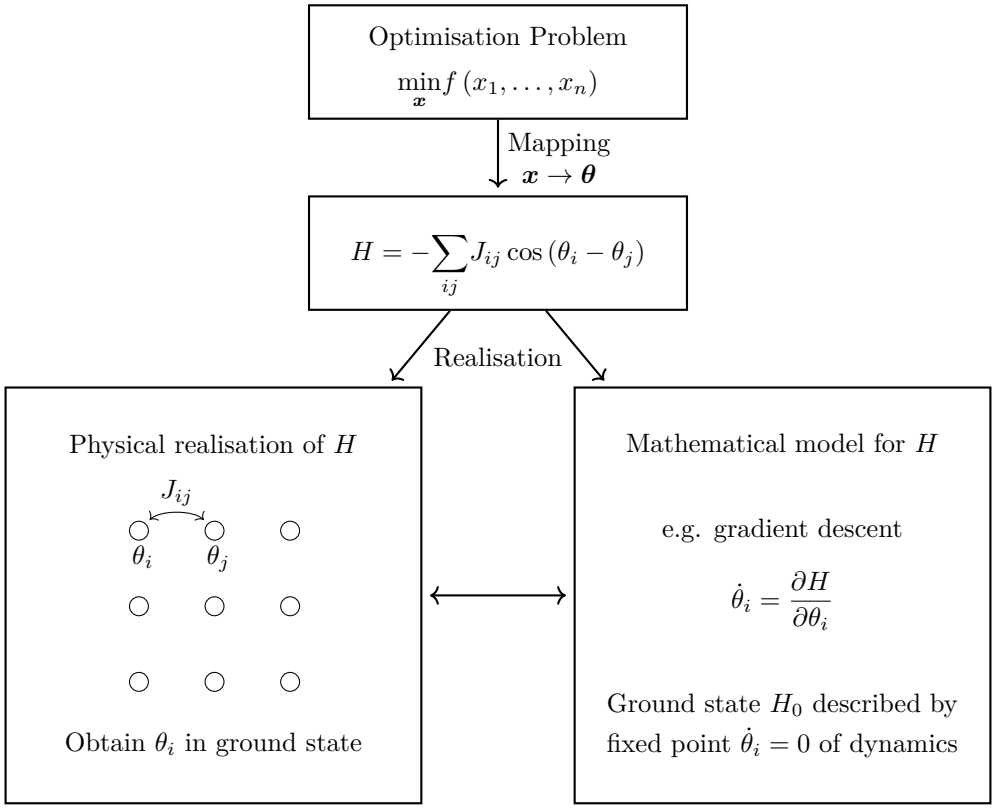


Figure 1.1: Schematic for solving optimisation problems. The problem is firstly mapped to a spin Hamiltonian H i.e. finding the ground state of H. There are two aspects to doing so: a mathematical model with dynamics describing evolution to the minimum of H and a physical system, here a network of coupled oscillators, capable of realising these dynamics. Performing the simulation, the phases θ_i should settle in the ground state of H, at which point they can be read off to obtain the solution to the original problem.

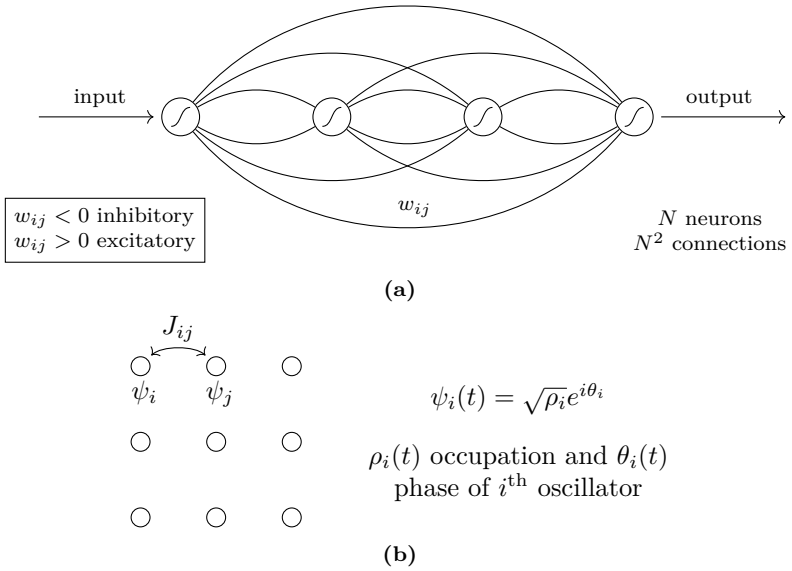


Figure 1.2: Non-conventional computing. (a) A neural network consists of layers of 'neurons' with programmable connections w_{ij}. This may be implemented in, for example, a photonic network where the neurons are made from materials with particular optical properties and electromagnetic waves are used to transfer information along the connections (we will study a type photonic network – a coherent Ising machine – in Section 11). (b) In a coherent network, oscillators characterised by a magnitude and phase are arranged in a graph and interact according to couplings J_{ij}.

1.2.2 Non-conventional and Neuromorphic Computing

Alternatives to von Neumann architecture include neural networks, capable of processing vast amounts of data in parallel (Figure 1.2a), and coherent networks, which consist of systems of coupled oscillators such lasers and Bose-Einstein condensates (Figure 1.2b). The goal is to design processors that can integrate and process massive amounts of information whilst using little energy.

A source of inspiration is the human brain, which boasts some 10^{15} synaptic connections (10^{11} neurons with 10^4 inputs per neuron) and achieves a computational efficiency[1] of $\sim 10^{18}$MACs^{-1} at only $\sim$ 20W. This implies an expenditure of less than 1 atto Joule (10^{-18}J) per MAC, some 8 orders of magnitude less than the $\sim$ 100pJ per MAC achieved by current supercomputers. The suggestion then, is that neuromorphic and other novel computing techniques have the potential to greatly surpass the performance of any conventional computer.

[1] MAC stands for multiply-accumulate, the common operation of multiplying two numbers together and adding it onto a running total.

2 Spin Hamiltonians and Non-linear Programming

2.1 The XY and Ising Models

The Hamiltonian appearing in Figure 1.1 is the two-dimensional XY model familiar from statistical physics. This describes classical spins (vectors of unit length) on a lattice with couplings J_{ij}:

$$H_{\mathrm{XY}} = -\sum_{ij} J_{ij}\boldsymbol{s}_i \cdot \boldsymbol{s}_j \equiv -\sum_{ij} J_{ij}\cos\left(\theta_i - \theta_j\right), \qquad \boldsymbol{s}_i = (\cos\theta_i, \sin\theta_i) \tag{2.1}$$

where each spin is parametrised by the angle $\theta_i \in [0, 2\pi)$ it makes with some reference axis and in the double sum i, j independently range over all sites of the lattice (see warning below). The system is normally taken to be translation-invariant, with periodic boundary conditions and J_{ij} the same when viewed from any site of the lattice. These couplings are in principle otherwise arbitrary, however the case $J_{ij} = 0$ unless i and j are nearest neighbours is most often considered.[2] Without loss of generality $J_{ii} = 0$ since $|\boldsymbol{s}_i|^2 = 1$ and adding a constant to the energy does not change the physics (minimisation problem).

When $J_{ij} > 0$, the spins want (i.e. to minimise the energy) to align in the same direction, whereas when $J_{ij} < 0$ they tend to align in opposite directions. In view of the situation when each spin corresponds to a magnetic moment, these two cases are referred to as ferromagnetic and antiferromagnetic couplings, respectively. However, the same language is used whether the spins represent superconducting qubits, electromagnetic oscillators or some other entity. Of course, finding the global minimum of (2.3) corresponds to finding the ground state (spin configuration) and its energy for the physical system described. Examples of the ground state for a few simple systems are given in Section 2.3.

Although the model is often written with a term

$$-\sum_i \boldsymbol{h}_i \cdot \boldsymbol{s}_i = -\sum_i h_i \cos\theta_i \tag{2.2}$$

describing the interaction of spins with an external field $\boldsymbol{h}_i$ that can vary from one site to the next, this term can always be incorporated in the spin-spin sum by introducing an additional spin $\boldsymbol{s}_k$ at fixed $\theta_k = 0$ that interacts with all other spins such that $J_{ik}\cos\left(\theta_i - \theta_k\right) = h_i \cos\theta_i$. Thus, we only need consider the form (2.1), possibly with long-range interactions.

The other spin model we will consider extensively is the two-dimensional *Ising model*. Here the spins are *discrete* variables, taking the values $s_i = \pm 1$:

$$H_I = -\sum_{ij} J_{ij} s_i s_j \tag{2.3}$$

which can be viewed as the XY Hamiltonian in the limit that θ_i is restricted to $\{0, \pi\}$. Again, the model is normally considered with nearest neighbour interactions on a physical lattice (typically, the square lattice) and an external field, but (2.3) is sufficiently general.

[2] We will work with Hamiltonians defined on arbitrary graphs for which the question of interaction length is moot (node arrangement does not necessarily correspond to a physical arrangement in space). Still, it is important to evaluate the graph's *connectivity* in mind that we want to realise these Hamiltonians in a real system.

2.1.1 Sum Notation: A Warning

In these notes

$$\sum_{ij} J_{ij} \tag{2.4}$$

indicates a double sum $\sum_i \sum_{j\neq i} J_{ij}$ in which both couplings J_{ij} and J_{ji} (which are equal) are included and self-terms are avoided. Other notations such as $\sum_{\langle i,j\rangle}$ and $\sum_{i\neq j}$ are found in the literature, with $J_{ii} = 0$ almost always implicit. The vanishing of J_{ij} when $i = j$ is also assumed then we write a single sum such as $\sum_j f_i J_{ij}$ in which i is a 'live' index.

Unfortunately, there is no standard convention and many take the same double-sum notation to imply $j < i$, meaning each pair of spins is included in the sum once only:

$$\sum_{ij} J_{ij} \leftrightarrow \sum_i \sum_{j<i} J_{ij} \tag{2.5}$$

Of course, the two conventions are related simply by a factor of $1/2$, and indeed if one sees the couplings written as $J_{ij}/2$ our convention is certainly being used. At any rate, one should always ask their lecturer whether they intend $j < i$ or not.

2.2 Quadratic Programming

A general non-linear programming problem is of the form

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{subject to} \quad g(\boldsymbol{x}) \leq 0 \text{ and } h(\boldsymbol{x}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^n \tag{2.6}$$

where both the *objective* function f and the constraints g, h may be non-linear. This optimisation problem has numerous applications not only in the mathematical and physical sciences but also in engineering, finance, logistics and industry.

We focus on *quadratic programming* problems (where f is quadratic and the constraints are linear) which describe more general problems close to their optimum. In particular, the two types of problem we initially encounter will be *quadratic continuous optimisation* (QCO) and *quadratic unconstrained binary optimisation* (QUBO). The former is most often formulated in terms of unit complex variables z_i as

$$\min_{|z_i|=1} -\boldsymbol{z}^\dagger Q \boldsymbol{z}, \quad \boldsymbol{z} = (z_1, z_2, \ldots, z_n),\ z_i \in \mathbb{C} \tag{2.7}$$

where $\dagger$ denotes the Hermitian conjugate. Writing $z_i = \cos\theta_i + i\sin\theta_i$, we realise this is the minimisation of some XY Hamiltonian

$$\min_{\theta_i} H_{\mathrm{XY}} = \min_{\theta_i} -\sum_{ij} J_{ij} \boldsymbol{s}_i \cdot \boldsymbol{s}_j \tag{2.8}$$

with $\boldsymbol{s}_i = (\cos\theta_i, \sin\theta_i)$. In QUBO, the variables are binary so unsurprisingly we are instead minimising an Ising Hamiltonian

$$\min_{\{s_i\}} H_I = \min_{\{s_i\}} -\sum_{ij} J_{ij} s_i s_j \tag{2.9}$$

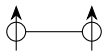
with arbitrary J_{ij} i.e. long-range interactions. We will also deal with examples of the binary type where constraints are imposed upon the variables.

2.3 Simple Spin Systems

In this section we consider the ground state of chains of spins in the XY and Ising models in the simple case of nearest neighbour interactions of constant strength $J_{ij} = \pm J$ with $J > 0$. We use a solid line to indicate a ferromagnetic coupling ($J_{ij} = +J$) and a dashed line an antiferromagnetic one ($J_{ij} = -J$).

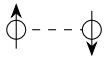
2.3.1 Two Spins

Firstly, with two spins for both models we have the ground state



$E = -2J$

in the ferromagnetic case and

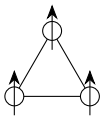


$E = -2J$

in the antiferromagnetic case. Technically, each of these states are degenerate: there is two-fold degeneracy for the Ising model, corresponding to flipping both spins (remember s_i takes only the values ± 1), whilst H_{XY} is left unchanged by any rotation or reflection of the spins ($O(2)$ symmetry). To remove this redundancy, we choose to always fix the first spin to be up.

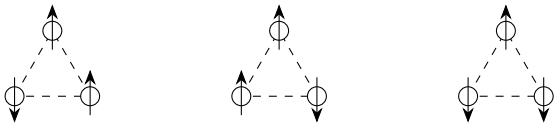
2.3.2 Triangle

The ferromagnetic solution is again the same for both models,



$E = -6J$

However the antiferromagnetic ground states now differ. For the Ising model, there is 3-fold degeneracy (after fixing $s_1 = +1$):



$E = -2J$

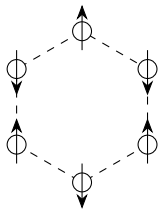
The system is said to be *frustrated,* because after placing two spins with antiferromagnetic coupling, the third cannot be placed so as to minimise its interactions with both of the other two. For the XY model there is only 2-fold degeneracy:

$E = -3J$

These states are obtained by, going around the triangle, adding a relative phase $+2\pi/3$ or $-2\pi/3$ to successive spins. The system is again frustrated to an extent, although all interactions are equal and the ground state energy lower than in the Ising model.

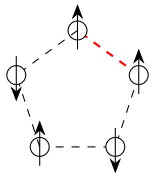
2.3.3 Antiferromagnetic Rings

Consider a ring of N spins with equal couplings $J_{ij} = -J$ when i, j are neighbours and $J > 0$ (the ferromagnetic case $J_{ij} = +J$ is trivial). When N is *even* both models have a unique ground state:



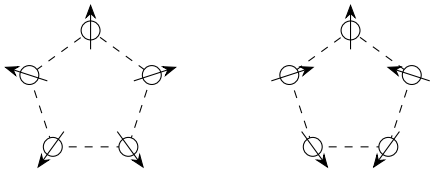
$$E = -2NJ$$

When N is odd frustration plays a role, resulting in degeneracy. There are N degenerate states for H_I, each with *one* ferromagnetic coupling and the rest antiferromagnetic:



$$E = -2(N-2)J$$

While for H_{XY} there is only 2-fold degeneracy, reflecting a choice of phase differences $2\pi/N$ or $-2\pi/N$ are you go around the N-gon:



$$E = 2N\cos(2\pi/N)J$$

3 Elements of Complexity Theory

In this section we classify the *difficulty* of the problems we are interested in solving. A scheme for doing so is provided by complexity classes, which sort problems according the asymptotics of the computational expense of finding or verifying a solution (with the best possible algorithm) as the size n of the problem's input increases. In other words, the behaviour $O(n)$ of the number of steps required to obtain or verify a solution, as $n \to \infty$. Here n refers to the size in memory i.e. number of binary digits of the input. For example, the single decimal input $x_1 = 12$ has binary representation 1100 (possibly up to a sign bit), so $n = 4$.

The primary complexity classes define collections of *decision* style questions. Unlike the optimisation questions we have considered so far, these have a binary output i.e. `yes` or `no`. As we shall

see, many optimisation problems (find min f) have a related decision style problem (is min $f > 0$?) and vice versa.

A second concept we will need to define the classes is that of a *turning machine.* This is in essence a finite state machine representing the operation of a classical[3] computer where data is read and processed sequentially: at each step the machine performs an action according the current input and a set of pre-defined (programmed) rules. The machine is assumed to have infinite memory; we are concerned with complexity in *time* (number of operations), not space.

3.1 Common Complexity Classes

The two most basic classes of decision problems are P and NP. P contains those decision problems that can be solved by a turning machine in polynomial time, i.e. $O(n^m)$ for some constant $m \geq 0$, and NP (non-deterministic polynomial) those for which a solution may be checked, but not necessarily obtained, in polynomial time. Simple examples of problems in P include testing whether two numbers are coprime (have a greatest common divisor of 1) or whether a string contains a particular substring. A well known example of an NP problem is integer factorisation, with a possible decision style formulation of "given integers N and k, is there an integer $f \leq k$ that divides N?"

Clearly P $\subseteq$ NP, so the famous problem *does $P = NP$?* is the question of whether, if we can verify a solution to a problem quickly, is it necessarily the case that a solution can be found quickly too? Common experience suggests otherwise: there are many problems $p \in$ NP for which no polynomial time algorithm is known. Integer factorisation provides on example, boolean satisfiability another (see below). Of course, such observations do not *prove* the non-existence of efficient algorithms, only that we are not aware of any. Indeed, proving P $\neq$ NP is not straightforward, despite the number of problems we *think* are in NP $\setminus$ P, for that very reason: it is difficult to determine how 'hard' a problem is beyond just looking at the best algorithms that have been devised to solve them.

Next we have NP-hard problems, which are 'at least as hard' of any problem in NP. More formally, a problem q is NP-hard if any $p \in$ NP can be reduced efficiently to q. Here *efficient reduction* typically refers to a polynomial time algorithm that transforms inputs of p to inputs of q such that the solution to the transformed problem solves the original problem, that is, an algorithm that transforms any instance of p to an instance of q with the *same* answer. This is known as a *Karp* reduction.

NP-hard problems are generally outside of NP, and may not even be decidable. This leads us to the final class we will discuss, NP-complete, which consists of problems that are both NP-hard *and* NP (Figure 3.1). Like any NP-hard problem, if you could construct an algorithm capable of solving an NP-complete problem in polynomial time, you would be able to solve any NP problem in polynomial time too. You could certainly do this if P = NP, but we believe this not to be the case. A common *hardness assumption* is that an NP-complete problem cannot be solved in polynomial time by *any* physical means (not limited to classical computers).[4] With this in mind, when designing a machine to solve NP-complete (or hard) problems using novel approaches, we are not looking to find a polynomial time solution (impossible), rather improve upon conventional methods either by

[3] Quantum computation has a separate set of classes. This is why integer factorisation is regarded as NP, despite Shor's ground-breaking algorithm for a polynomial-time solution on an ideal quantum computer.

[4] This is justified not only by the large number of NP-complete problems ($\gtrsim 30000$) with no known efficient method of solution known but also the omnipotence such a method would bring.

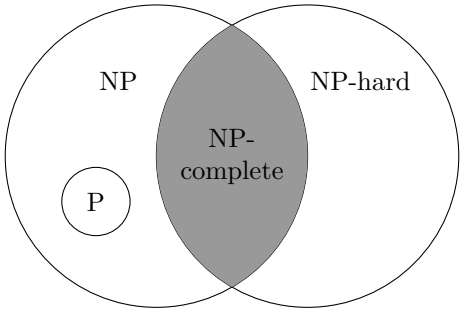


Figure 3.1: The intersection of NP and NP-hard defines NP-complete problems: decision problems that any $p \in$ NP can be efficiently reduced to, and with solutions that may be verified in polynomial time.

a) Finding solutions faster, or

b) Finding better solutions (for the optimisation style problems) with limited time or resources.

We will see many examples of NP-complete problems shortly. A classic one is the graph isomorphism problem: is G_1 isomorphic to G_2? Another is *3SAT*: given n boolean variables $\mathrm{x}_1, \ldots, \mathrm{x}_n$ and a set of logical *clauses* that relate at *most* 3 variables each, e.g. $\mathrm{x}_2 \vee \mathrm{x}_5 \vee \bar{\mathrm{x}}_6$ (bar denoting negation), is there a way to satisfy all clauses simultaneously? We look at the formulation of this problem in detail and an example in Section 4.7. Its NP-completeness (Cook-Levin Theorem) is particularly important for the fact that polynomial time reductions from 3SAT to many spin Hamiltonians such as H_{XY} exist. This demonstrates that these Hamiltonians offer NP-complete problems, which they should do as the basis of our envisioned problem solver (Figure 1.1).

A result established by G. De las Cuevas and T. Cubitt [1] that goes further is that the two-dimension Ising model on the square lattice with nearest neighbour interactions and an external field[5] is *universal* in the sense that its low energy sector can reproduce the complete physics (energy level structure and partition function) of any other classical spin model – the ultimate simulator. In the following section, we look at how some classic NP-hard and NP-complete problems can be *mapped* to the Ising model.

4 Ising Formulations of NP Problems

Following from the previous section, we consider NP-complete problems to be decision problems "does the ground state of H_I have energy 0 (or ≤ 0)?", and NP-hard problems to be optimisation problems "what is the ground state of H_I?". In this section we establish how some famous NP-complete and NP-hard problems can be mapped to the Ising model i.e. formulated in terms of these two questions.

In the following, $G = (E, V)$ denotes a graph (edge set E, vertex set V) of order $N = |V|$ which, unless otherwise stated, is simple, undirected and unweighted. A section title with **(D)** indicates a decision formulation of a problem (NP-complete) and **(O)** an optimisation one (NP-hard). Further problems (covering and packing) are addressed in Example Sheet 1.

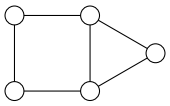
[5] As mentioned in Section 2.1, the external field subsumed with an additional spin, but at the expense of long-range interactions i.e. we would no longer have a square lattice with nearest neighbour interactions.

4.1 Partitioning Problems

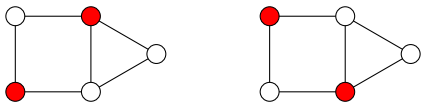
4.1.1 Max-Cut (O)

Partition the vertices of $G = (E, V)$ into two sets such that the number of edges between the two sets is maximised.

To illustrate the problem, consider the graph



There are evidently two solutions where the partitions have 5 connecting edges:



How to realise this result as the ground state of a spin model?

Assign to each vertex $u \in G$ a value $s_u = +1$ to indicate it is in one partition and $s_u = -1$ the other. Maximising the number of edges between the two partitions is then simply the problem of *minimising* the sum

$$\sum_{(uv)\in E} s_u s_v \tag{4.1}$$

which is just the minimisation of the Ising Hamiltonian $H_I = -\sum_{(uv)\in E} J_{uv} s_u s_v$ with antiferromagnetic couplings J_{uv} defined by the edges of the graph:

$$J_{uv} = \begin{cases} -1 & \text{if } (uv) \in E \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

Indeed, one can envision (Figure 4.1) a physical system with spins prepared on the vertices of the graph in some easy initial state and antiferromagnetic couplings are slowly[6] turned on, moving the system to a ground state describing the solution to the max-cut problem. As in the example shown here, this will in general be only one of many degenerate ground states, but for the purposes of optimisation we only need find one solution, not all possible solutions.

In addition to solving the NP-hard problem of maximising the number of connecting edges we can also answer the associated NP-complete decision problem,[7] "Does G have a cut of at least size K?", as this is simply the question of whether the energy of the ground state obtained above is at least as negative as $-K$.

[6] We will see the relevance of the initial state being 'easy' and the couplings turned on 'slowly' in Section 7. Note J_{uv} is just the adjacency matrix of the graph (Section 5.1).

[7] This is one of Karp's original NP-complete problems [2].

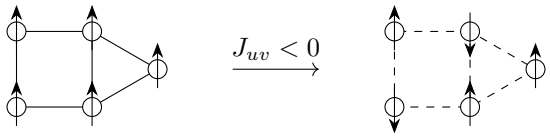


Figure 4.1: Spins are arranged in a graph with $s_u = +1$ initially and antiferromagnetic couplings established along the edges of the graph. The value of each spin in the resulting ground state tells us the vertices in each partition of a solution to the associated max-cut problem. Note in more general cases we have to careful that this procedure yields a global minimum of H_I, not just a local one.

4.1.2 Number Partitioning (D, O)

Given $S = \{n_1, \ldots, n_N\} : n_i > 0$, is there a partition into two disjoint subsets R, $S \setminus R$ such that the sum of the elements in each set is the same i.e.

$$\sum_{n_i \in R} n_i = \sum_{n_i \in S \setminus R} n_i \tag{4.3}$$

If not, find a partition that minimises the mismatch.

Let $s_i = 1$ if $n_i \in R$ and $s_i = -1$ otherwise and note that (4.3) holds if and only if the sum

$$\sum_{i}^{N} n_i s_i \tag{4.4}$$

vanishes. This is *not* a suitable Hamiltonian, since it is always lowered by removing numbers from R and indeed the ground state has all $s_i = -1$ which is clearly not a solution. To get around this, simply square the sum:

$$H_I = \left(\sum_{i}^{N} n_i s_i \right)^2 \tag{4.5}$$

Then H_I is bounded below by zero and *any* deviation from zero is penalise. Thus, by minimising H_I one not only answers the decision problem (is the ground state energy 0?) but also the optimisation problem in the case of no balanced solution: the ground state of H_I is a configuration of spins describing a partition of minimal mismatch.

4.1.3 Graph Partitioning (O)

Partition the vertices of a graph $G = (E, V)$ of even order N into equal subsets such that the number of edges between sets is minimised.

Again assign $s_u = 1$ to one set of a proposed partition and $s_u = -1$ to the other (Figure 4.2). Unlike max-cut problem we wish to minimise not maximise the number of connecting edges and so require ferromagnetic (positive) couplings between the spins along the edges of the graph. The sum

$$\sum_{(uv) \in E} \frac{1 - s_u s_v}{2} \tag{4.6}$$

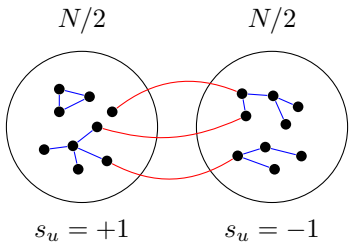


Figure 4.2: A possible partitioning of a graph into equal subsets. The total number of connections between the two partitions (red) is to be minimised.

does the trick.[8] In addition, we need the partitions to be equal. The simplest way to ensure this is with a term proportional to

$$\left(\sum_{u\in V} s_u\right)^2 \tag{4.7}$$

which vanishes in the case of equal partitions only. Together,

$$H_I = H_A + H_B = A\left(\sum_{u\in V} s_u\right)^2 + B\sum_{(uv)\in E}\frac{1 - s_u v s_v}{2} \tag{4.8}$$

where A and B are positive constants weighting each term. These are not arbitrary: the equality condition must hold absolutely to have a viable solution, so H_A should dominate H_B. In other words, it should never be favourable to violate the A constraint. To reason a loose lower bound on the ratio A/B, note that if Δ is the maximum degree of G then the greatest possible gain from violating the constraint by moving a vertex between sets is ΔB (the vertex had Δ edges to vertices in the other set), and the cost of doing so is *at least* 2^2A (one partition now has at least two more vertices than the other). Hence we should choose[9]

$$\frac{A}{B} \geq \frac{\Delta}{4} \tag{4.9}$$

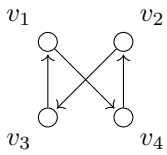
4.2 General Considerations

We have now seen several general features of mapping problems to Ising Hamiltonians:

- Positive terms in H_I are minimised, negative terms maximised
- Terms should be included to impose constraints, appropriately weighted
- Squaring terms is a useful tactic to ensure positive semi-definiteness

[8] $(1 - s_u s_v)/2$ was used instead of $-s_u s_v$ for convenience; this term vanishes if u, v (vertices with a common edge) are from the same partition and is 1 otherwise.

[9] A tighter bound can be obtained: due to the fact H_A is very penalising for many spins in the same partition, we only need to worry about moving one vertex from an equally partitioned ($H_A = 0$) situation, in which case we can replace $B\Delta$ by $B\min\{\Delta, N/2\}$ (the vertex cannot have more than $N/2$ edges to the other partition).



(a)

u \ i	1	2	3	4
1	1	0	0	0
2	0	0	1	0
3	0	0	0	1
4	0	1	0	0

(b)

Figure 4.3: (a) A directed graph with a Hamiltonian cycle and (b) values of the 4^2 binary variables defined in (4.11) describing the cycle.

So far, the number of spins required to encode the problem has corresponded directly to the number of items or vertices in the problem, N. As the next examples illustrate, this is not always the case, and moreover it is often necessary to introduce auxiliary spins to impose constraints (for example, in the Knapsack problem in the first example sheet).

The growth of the number of spins required to encode the problem with input size is obviously an important consideration, since in any proposed physical simulator more spins will necessitate larger hardware. Other factors to consider to this end are any separation of energy scales and the connectivity of the embedding graph (that is, the graph described by the couplings J_{uv}, which may or may not correspond to an actual graph in the problem).

Often it is convenient to phrase problems in terms of binary variables χ_u with possible values $0, 1$ instead of spin variables $s_u = \pm 1$. This makes no difference to the minimisation problem and one can always go between the two types of variables with

$$s_u = 2\chi_u - 1 \quad \leftrightarrow \quad \chi_u = \frac{s_u + 1}{2} \tag{4.10}$$

4.3 Hamiltonian Cycles

Recall that a Hamiltonian path on a graph $G = (E, V)$ is a sequence of edges between vertices such that every vertex is visited exactly once. In a Hamiltonian cycle, the path starts and ends at the same vertex (which is visited twice). We address the problem of Hamiltonian cycles on a *directed* graph, so $(uv) \in E$ means you can go from u to v, but not necessarily from v to u. An example is provided in Figure 4.3a.

4.3.1 Existence on a Directed Graph (D)

Does G have a Hamiltonian cycle?

As there are $N = |V|$ vertices each of which could potentially be visited at any of the N steps of a prospective cycle, we require N^2 variables.[10] Label the vertices in V with index $u = 1, \dots N$ and define

$$\chi_{u,i} = \begin{cases} 1 & u \text{ visited on } i^{\text{th}} \text{ step} \\ 0 & \text{otherwise} \end{cases} \tag{4.11}$$

[10] Well, there are $N + 1$ steps but the final is predetermined since we have a cycle. In fact, we can always take the first step to be vertex 1, fixing $\chi_{1,i} = \delta_{1,i}$ and reducing the number of variables required to $(N-1)^2$.

Our aim is to construct a Hamiltonian from the $\chi_{u,i}$ which vanishes only when these variables describe a true cycle. We need

- A term to ensure each vertex is visited exactly once (barring the first/last vertex)
- A term to ensure that one and only one vertex is visited at each step
- A term to ensure steps are only taken between vertices which actually share an edge

An appropriate Hamiltonian is then

$$H_{\text{cycle}} = \sum_{v=1}^{N} \underbrace{\left(1 - \sum_{i=1}^{N} \chi_{v,i}\right)^2}_{\substack{\text{visit vertex } v \\ \text{exactly once}}} + \sum_{i=1}^{N} \overbrace{\left(1 - \sum_{v=1}^{N} \chi_{v,i}\right)^2}^{\substack{\text{visit exactly one vertex} \\ \text{in the } i^{\text{th}} \text{ step}}} + \underbrace{\sum_{i=1}^{N} \sum_{(uv)\notin E} \chi_{u,i}\chi_{v,i+1}}_{\substack{\text{only travel along} \\ \text{edges actually in } E}} \tag{4.12}$$

where in the final sum $\chi_{u,N}\chi_{v,N+1}$ should be understood as $\chi_{u,N}\chi_{v,1}$, enforcing that the $(N+1)^{\text{th}}$ vertex visited is the same as the first i.e. we have a cycle (for the corresponding path problem, this sum would run to $i = N-1$). You should check each term acts as described so that $H_{\text{cycle}} \geq 0$ has a zero energy ground state only if a cycle exists.

4.3.2 Travelling Salesman (O)

> *Given weights w_{uv} for each edge $(uv) \in E$ of a complete graph, find the Hamiltonian cycle that minimises the sum of weights of edges in the cycle.*

As written, this problem is typically phrased for a complete graph, but we can work more generally with any graph containing Hamiltonian cycles.

The solution is simple extension of previous problem: H_{cycle} is used to impose that a prospective cycle is Hamiltonian and then

$$H_B = B \sum_{(uv)\in E} W_{uv} \sum_{i=1}^{N} \chi_{u,i}\chi_{v,i+1} \tag{4.13}$$

is added so that the ground state minimises the total weight. As with the partitioning problem, in

$$H_I = AH_{\text{cycle}} + H_B \tag{4.14}$$

we need $B > 0$ small enough such that the constraint of H_{cycle} is never violated. A simple bound is

$$A > B \max_{(uv)\in E} W_{uv} \tag{4.15}$$

since the cost of violating the cycle constraint once is at least A (travel along an edge not in E) and the maximum gain from doing so is given by the right hand side.

4.4 Colouring Problems

4.4.1 Graph Colouring (D)

Given an undirected graph G of order N and n colours, is it possible to colour each vertex so that no edge connects two vertices of the same colour?

We set up binary variables similar to the cycle problems to describe the nN possible colourings of the N vertices:

$$\chi_{u,i} = \begin{cases} 1 & u \text{ has colour } i \\ 0 & \text{otherwise} \end{cases} \tag{4.16}$$

The Hamiltonian to answer this decision problem is

$$H_I = \sum_{v=1}^{N} \Bigg(\underbrace{1 - \sum_{i=1}^{n} \chi_{v,i}}_{\substack{\text{vertex } v \text{ has} \\ \text{exactly one colour}}} \Bigg)^2 + \sum_{(uv)\in E} \underbrace{\sum_{i=1}^{n} \chi_{u,i}\chi_{v,i}}_{\substack{u \text{ and } v \text{ have} \\ \text{different colours}}} \tag{4.17}$$

where as indicated the first term ensures each vertex receives one colour and the second that neighbouring vertices have different colours.

4.5 Reduction of k-local Hamiltonians

The Ising and XY models are examples of 2-local Hamiltonians since each term couples two spins. More generally, we may have a model with interaction terms coupling up to k different spins, for example the 3-local

$$H = \sum_{ijk} J_{ijk} s_i s_j s_k, \quad s_i \in \{-1, 1\} \tag{4.18}$$

Any k-local Hamiltonian can be converted into a 2-local one with *polynomial overhead.*[11] We demonstrate this for the $3 \to 2$ case explicitly.

4.5.1 Reduction of 3-local Hamiltonians

Consider a general term wxy of a 3-local Hamiltonian with w, x, y, binary variables (recall we can easily go back and forth between these and the spins ± 1 traditionally used in statistical physics). For $x, y, z \in \{0, 1\}$ it is easy to show that (Appendix A)

$$xy = z \Leftrightarrow xy - 2xz - 2yz + 3z = 0 \tag{4.19}$$

and

$$xy \neq z \Leftrightarrow xy - 2xz - 2yz + 3z > 0 \tag{4.20}$$

[11] That this is possible jibes with the universality of H_I (Section 3.1) although we do not attempt reduction to nearest neighbour interactions on a square lattice, which is not straightforward.

This allows us to replace the product wxy with a series of two variable terms involving w, x, y and $z := xy$, as

$$wxy \mapsto wz + 2\left(xy - 2xz - 2yz + 3z\right) \tag{4.21}$$

where the weighting $2 > 1$ ensures that the second factor dominates the first i.e. asserts $z = xy$. So a 3-local term is transformed to a 2-local one at the cost of an additional variable (spin) and some trivial (polynomial) computations.

Next we discuss a notable problem for which a k-local Hamiltonian features.

4.6 Number Factoring

The widely used RSA cryptosystem is based on the difficulty of factorising a large integer N, constructed by multiplying two primes x and y (we say N is *semiprime*), when these primes are unknown. We consider the slightly more general problem of determining x, y given an n-bit integer $N = xy$ which is odd but not necessarily semiprime.

To construct a suitable Hamiltonian we simply inspect the steps in the forward operation of binary multiplication to obtain a condition for each digit of N. Begin by writing the binary representations of N (known) and x, y (unknowns):

$$N = N_n N_{n-1} \ldots N_1 1 \tag{4.22}$$

$$x = x_n \;\; x_{n-1} \ldots x_1 1 \tag{4.23}$$

$$y = y_n \;\; y_{n-1} \ldots y_1 1 \tag{4.24}$$

where the last digit is always 1 since each integer is odd. These final digits are such that

$$x_0 = 1, \;\; y_0 = 1, \;\; N_0 = x_0 y_0 = 1 \tag{4.25}$$

Call this the 0^{th} step. The next digit N_1 of N is determined by the last *two* digits of x and y:

$$x_1 y_0 + x_0 y_1 = N_1 + 2\ell_{11} \tag{4.26}$$

where the binary variable ℓ_{11} accounts for the fact that without this variable the equality only holds modulo 2: if $x_1 y_0 + x_0 y_1 = 2$ then $N_1 = 0$ and the 1 rolls over to the next column, hence the appearance of ℓ_{11} on the left hand side of the second step,

$$x_2 y_0 + x_1 y_1 + x_0 y_2 + \ell_{11} = N_2 + 2^1 \ell_{21} + 2^2 \ell_{22} \tag{4.27}$$

Here we needed *two* additional registers ℓ_{22} and ℓ_{21}, multiplied by appropriate powers of 2. In general for the k^{th} step k such variables appear on the right and $\lfloor k/2 \rfloor$ on the left ($\lfloor \cdot \rfloor$ the floor function):

$$x_k y_0 + x_{k-1} y_1 + \ldots + x_0 y_k + \ell_{k-1,1} + \ell_{k-2,2} + \ldots + \ell_{a,b} = N_k + 2\ell_{k1} + 2^2 \ell_{k2} + \ldots 2^k \ell_{kk} \tag{4.28}$$

where $\ell_{a,b}$ has indices

$$a = \left\lfloor \frac{k+1}{2} \right\rfloor \text{ and } b = \left\lfloor \frac{k}{2} \right\rfloor \tag{4.29}$$

This hierarchical procedure is continued until $k = n$ and the multiplication completes. Converting the equation at each stage to a constraint in H is simply a matter of moving everything to one side. We arrive at

$$\begin{aligned} H = &(x_0 y_0 - N_0)^2 + \dots \\ &\vdots \\ &+ \left(x_k y_0 + \dots + x_0 y_k + \ell_{k-1,1} + \dots + \ell_{a,b} - N_k - 2\ell_{k1} - \dots - 2^k \ell_{kk}\right)^2 \\ &\vdots \end{aligned} \tag{4.30}$$

Each term is squared in order to penalise any deviation from equality and thus convey the desired ground state. Since inside each parentheses terms are at most second order in the binary variables, the full Hamiltonian is 4-*local*. As discussed above, with additional work this could be reduced to a 2-local (Ising) Hamiltonian.

4.7 3SAT

> *Given k clauses C_1, C_2, …, C_k each involving at most 3 binary variables from $\mathrm{x}_1, \dots, \mathrm{x}_n$, determine whether there exists a tuple $(\mathrm{x}_1, \dots, \mathrm{x}_n)$ such that all clauses are simultaneously satisfied.*

In logic *clause* has a specific meaning: it is a disjunction[12] of literals such as $\mathrm{x}_1 \vee \mathrm{x}_2 \vee \bar{\mathrm{x}}_3$. Normally the clauses are presented in 'conjunctive normal form' $C_1 \wedge C_2 \wedge \dots \wedge C_k$ which evaluates `true` for a solution tuple.

To appreciate the boolean satisfiability problem we reverse engineer a boolean function to construct a simple instance of the problem.

4.7.1 Example and Mapping

Consider the boolean function of three boolean variables $\mathrm{x}_i \in \{0, 1\}$:

$$\varphi_{00}(\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3) = \begin{cases} 1 & \text{if } (\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3) = (1,0,0) \vee (0,1,0) \vee (0,0,1) \vee (1,1,0) \\ 0 & \text{otherwise} \end{cases} \tag{4.31}$$

Observe that φ_{00} can be written as a series of 3-clauses:

$$\varphi_{00}(\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3) = (\mathrm{x}_1 \vee \mathrm{x}_2 \vee \mathrm{x}_3) \wedge (\mathrm{x}_1 \vee \bar{\mathrm{x}}_2 \vee \bar{\mathrm{x}}_3) \wedge (\bar{\mathrm{x}}_1 \vee \mathrm{x}_2 \vee \bar{\mathrm{x}}_3) \wedge (\bar{\mathrm{x}}_1 \vee \bar{\mathrm{x}}_2 \vee \bar{\mathrm{x}}_3) \tag{4.32}$$

To see this, note that of the $2^3 = 8$ possible values of the tuple $(\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3)$ those for which φ_{00} evaluates `false` are

$$(0,0,0) \qquad (0,1,1) \qquad (1,0,1) \qquad (1,1,1) \tag{4.33}$$

i.e.

$$(\bar{\mathrm{x}}_1 \wedge \bar{\mathrm{x}}_2 \wedge \bar{\mathrm{x}}_3) \quad (\bar{\mathrm{x}}_1 \wedge \mathrm{x}_2 \wedge \mathrm{x}_3) \quad (\mathrm{x}_1 \wedge \bar{\mathrm{x}}_2 \wedge \mathrm{x}_3) \quad (\mathrm{x}_1 \wedge \mathrm{x}_2 \wedge \mathrm{x}_3) \tag{4.34}$$

[12] $A \vee B$ is a disjunction, $A \wedge B$ a conjunction.

So (as φ_{00} is `true` otherwise)

$$\varphi_{00} = \neg\left[(\bar{\mathrm{x}}_1 \wedge \bar{\mathrm{x}}_2 \wedge \bar{\mathrm{x}}_3) \vee (\bar{\mathrm{x}}_1 \wedge \mathrm{x}_2 \wedge \mathrm{x}_3) \vee (\mathrm{x}_1 \wedge \bar{\mathrm{x}}_2 \wedge \mathrm{x}_3) \vee (\mathrm{x}_1 \wedge \mathrm{x}_2 \wedge \mathrm{x}_3)\right] \tag{4.35}$$

Distributing the negation $\neg$ according to De Morgan's laws,

$$\varphi_{00} = \neg(\bar{\mathrm{x}}_1 \wedge \bar{\mathrm{x}}_2 \wedge \bar{\mathrm{x}}_3) \wedge \neg(\bar{\mathrm{x}}_1 \wedge \mathrm{x}_2 \wedge \mathrm{x}_3) \wedge \neg(\mathrm{x}_1 \wedge \bar{\mathrm{x}}_2 \wedge \mathrm{x}_3) \wedge \neg(\mathrm{x}_1 \wedge \mathrm{x}_2 \wedge \mathrm{x}_3) \tag{4.36}$$

$$= (\mathrm{x}_1 \vee \mathrm{x}_2 \vee \mathrm{x}_3) \wedge (\mathrm{x}_1 \vee \bar{\mathrm{x}}_2 \vee \bar{\mathrm{x}}_3) \wedge (\bar{\mathrm{x}}_1 \vee \mathrm{x}_2 \vee \bar{\mathrm{x}}_3) \wedge (\bar{\mathrm{x}}_1 \vee \bar{\mathrm{x}}_2 \vee \bar{\mathrm{x}}_3) \tag{4.37}$$

as claimed. A 3SAT problem is to determine whether there exists a triplet $(\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3)$ such that (4.37) is `true`. Of course, we already know the answer is positive: any of the tuples in (4.31) will do! But in general only the conjunctive normal form will be given. How might we encode this decision problem in an Ising Hamiltonian?

Returning to the notation of χ_i as a binary variable in a Hamiltonian, a naive way would be write down a non-negative term for each clause in (4.37) that vanishes only when the clause is satisfied:

$$H = (1-\chi_1)(1-\chi_2)(1-\chi_3) + (1-\chi_1)\chi_2\chi_3 + \chi_1(1-\chi_2)\chi_3 + \chi_1\chi_2\chi_3 \tag{4.38}$$

so that in total $H = 0$ is a possibility only if the satisfiability problem has a solution, in which case the ground state describes a solution set of literals. This is a 3-local Hamiltonian which we know can be reduced to a 2-local one, however the embedding graph is highly connected: for a general 3SAT problem each spin will be coupled to all others. To demonstrate the universality of the Ising Hamiltonian on the square lattice (reduction from the general 3SAT problem), Cuevas and Cubitt used a more sophisticated mapping – see their paper [1] for the details.

5 Network Dynamics

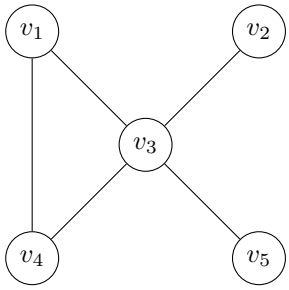
In this section we develop the basic mathematical tools for studying dynamics on *networks*. A network is an interconnected system of nodes each of which has a state, represented by some set of variables, that evolves deterministically according to a prescribed set of rules or equations. In other words, a graph with one more dynamic quantities at each vertex (node). The equations governing the dynamics at one node in general depend on the state of all other nodes in the network, or at least those that particular node is connected to. We closely follow the presentation of the topic by Mark Newman in Chapter 17 of his book *Networks* [3] which you may wish to consult for further details.

Looking ahead, in Section 6 we will introduce some specific networks capable of realising the minimisation of the Ising and XY Hamiltonians, before moving on to discuss physical systems with the potential to simulate them in later sections.

5.1 Graph Matrices

We begin by introducing two matrices used to represent finite graphs and properties of their spectra. These are the adjacency matrix $\boldsymbol{A}$ and the graph Laplacian $\boldsymbol{L}$. For a graph $G = (V, E)$ with $n = |V|$ vertices (enumerated as $v_1, v_2, \ldots, v_n$) and $m = |E|$ edges, both are $n \times n$ matrix with components

$$A_{ij} = \begin{cases} 1 & \text{if } (v_i v_j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$



$$A_{ij} = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \qquad L_{ij} = \begin{pmatrix} 2 & 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Figure 5.1: The adjacency matrix and graph Laplacian for a graph of 5 vertices.

and

$$L_{ij} = k_i \delta_{ij} - A_{ij} = \begin{cases} k_i & \text{if } i = j \\ -1 & \text{if } (v_i v_j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

respectively, where k_i is the degree of the vertex $v_i \in V$. Weights may be included in the description by assigning $A_{ij} = w_{ij}$ where w_{ij} is the weight of the edge between v_i and v_j, but we consider unweighted *undirected* graphs for which $\boldsymbol{A}$ and $\boldsymbol{L}$ are *symmetric* matrices with the above components. Examples of these matrices for a simple graph of 5 vertices are provided in Figure 5.1.

The adjacency matrix and graph Laplacian encode the *structure* of a network and as such feature prominently in any discussion of dynamics on that network. In particular, we will see that their spectra determine the *stability* of fixed points of the dynamics.

5.1.1 Spectra of Networks

As $\boldsymbol{A}$ and $\boldsymbol{L}$ are real symmetric matrices they have n (possibly degenerate) real eigenvalues. We denote the eigenvalues of $\boldsymbol{A}$ by κ_i, ordered as $\kappa_1 \geq \kappa_2 \geq \ldots \geq \kappa_n$ and those of $\boldsymbol{L}$ by λ_i, instead ordered as $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ (the inconsistency in the ordering is silly, but that's convention for you). We list some properties of these spectra that will be useful in the following. Firstly,

i) $\boldsymbol{A}$ necessarily has both positive and negative eigenvalues ($\operatorname{Tr} \boldsymbol{A} = 0$)

ii) $0 \leq \kappa_1 \leq k_{\max}$

iii) The eigenvalues of $\boldsymbol{L}$ are non-negative and the smallest eigenvalue $\lambda_1 = 0$ always

iv) $k_{\max} \leq \lambda_n \leq 2k_{\max}$

Here $k_{\max}$ is the maximum degree of the graph. We also note $\sum_j A_{ij} = k_i$, the degree of the i^{th} vertex, hence $\sum_{ij} A_{ij} = 2m = 2|E|$. Additional bounds on the first eigenvalue of the adjacency matrix can be obtained by evaluating the Rayleigh quotients[13]

$$\frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{\sum_i c_i \kappa_i}{\sum_i c_i^2} \leq \frac{\sum_i c_i^2 \kappa_1}{\sum_i c_i^2} = \kappa_1 \quad \text{and} \quad \frac{\boldsymbol{x}^T \boldsymbol{A}^2 \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \leq \kappa_1^2 \tag{5.3}$$

[13]For the second we used the fact that the largest eigenvalue of $\boldsymbol{A}^2$ is κ_1^2, not κ_n^2 (Perron-Frobenius Theorem).

κ_1	κ_2	κ_3	κ_4	κ_5
-1.81	-1	0	0.47	2.34

λ_1	λ_2	λ_3	λ_4	λ_5
0	1	1	3	5

(a)

$k_{\max} = 4$
$\langle k \rangle = 2$
$\sqrt{\langle k^2 \rangle} = 2.28$

(b)

Table 1: (a) Eigenvalues of the adjacency matrix and graph Laplacian for the graph shown in Figure 5.1 and (b) the maximum, mean and root mean squared degree for this graph. Non-integer values are given to two decimal places.

($\boldsymbol{x}$ an arbitrary vector expanded in the eigenbasis) with the choice $x_i = 1$ for each i:

$$\kappa_1 \geq \frac{\sum_{ij} A_{ij}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{2m}{n} = \langle k \rangle \quad \text{and} \quad \kappa_1^2 \geq \frac{\sum_i k_i^2}{n} = \langle k^2 \rangle \tag{5.4}$$

Since the variance $\langle k^2 \rangle - \langle k \rangle^2$ of the degree distribution cannot be negative, the latter gives a stronger condition i.e.

$$\kappa_1 \geq \sqrt{\langle k^2 \rangle} \geq \langle k \rangle \tag{5.5}$$

If v_j is a node of maximum degree, the slightly less obvious choice of

$$x_i = \begin{cases} \sqrt{k_{\max}} & \text{if } i = j \\ 1 & \text{if } A_{ij} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{5.6}$$

gives[14]

$$\sum_k A_{ik} x_k \geq \begin{cases} k_{\max} & \text{if } i = j \\ \sqrt{k_{\max}} & \text{if } A_{ij} = 1 \\ 0 & \text{otherwise} \end{cases} = \sqrt{k_{\max}} x_i \tag{5.7}$$

and so

$$\kappa_1 \geq \frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \geq \sqrt{k_{\max}} \tag{5.8}$$

This may offer a better bound than (5.5) or not depending on the graph under consideration.

As an example, the spectra of $\boldsymbol{A}$ and $\boldsymbol{L}$ for the graph in Figure 5.1 are given in Table 1; you can check that each of the above bounds are satisfied.

5.2 Dynamical Systems on Networks

The general set-up for this section will be a network with one of more dynamical variables x_i at each node such that these variables are coupled together along the edges of the network i.e. x_j features

[14] When $i = j$, $\sum_k A_{jk} x_k = k_{\max}$, since node j is of degree $k_{\max}$ and $x_i = 1$ when i is adjacent to j. For i adjacent to j, $\sum_k A_{ik} x_k = A_{ij} x_j + \sum_{k \neq j} A_{ik} x_k \geq A_{ij} x_j = 1 \cdot \sqrt{k_{\max}}$. In any other case, $\sum_k A_{jk} x_k$ is at least non-negative.

in the equation for $\dot{x}_i$ when j is adjacent to i (in what follows, a dot denotes a time derivative). This amounts to a high dimensional dynamical system which we recall[15] can be written as a system of coupled first order autonomous differential equations (remove higher derivatives with additional variables $y_i = \dot{x}_i$etc., and time with $\dot{y} = 1 \leftrightarrow y = t$).

In general, such a system (unless linear) does not have a closed-form solution and the most productive strategy is to identify fixed points or limit cycles of the dynamics and perform *linear stability analysis.* This provides useful information about the likely behaviour of the system at long times.

We begin with the simple case of a single variable per node,

$$\dot{x}_i = f_i(x_i) + \sum_j A_{ij} g_{ij}(x_i, x_j) \qquad 1 \leq i \leq n \tag{5.9}$$

The function f_i describes the intrinsic dynamics of the node (how it would evolve in isolation) and g_{ij} the contribution from connections. Typically, the nodes represent identical entities (spins, electrical oscillators etc.) meaning $f_i = f$ and $g_{ij} = g$ are the same for every node. Although A_{ij} is symmetrical (the graph is undirected), g is *not* necessarily symmetric in its arguments: $g(x_i, x_j) = (1 - x_i)x_j$ is a possibility, for example.

5.2.1 Linear Stability Analysis

Suppose we have been able to identify a fixed point $\{x_i^*\}$ of the dynamics such that

$$f(x_i^*) + \sum_j A_{ij} g(x_i^*, x_j^*) = 0 \quad \forall i \tag{5.10}$$

Linearise about this fixed point by writing $x_i = x_i^* + \epsilon_i$ with $\epsilon_i \ll 1$:

$$\frac{d\epsilon_i}{dt} = f\left(x_i^* + \epsilon_i\right) + \sum_j A_{ij} g\left(x_i^* + \epsilon_i, x_j^* + \epsilon_j\right) \tag{5.11}$$

$$\begin{aligned} &= f\left(x_i^*\right) + \sum_j A_{ij} g(x_i^*, x_j^*) \\ &+ \epsilon_i \left.\frac{df}{dx}\right|_{x=x_i^*} + \epsilon_i \sum_j A_{ij} \left.\frac{\partial g(u,v)}{\partial u}\right|_{(u,v)=(x_i^*, x_j^*)} + \sum_j A_{ij}\epsilon_j \left.\frac{\partial g(u,v)}{\partial v}\right|_{(u,v)=(x_i^*, x_j^*)} + \ldots \end{aligned} \tag{5.12}$$

Using (5.10) and defining

$$\alpha_i = \left.\frac{df}{dx}\right|_{x=x_i^*}, \quad \beta_{ij} = \left.\frac{\partial g(u,v)}{\partial u}\right|_{(u,v)=(x_i^*, x_j^*)}, \quad \gamma_{ij} = \left.\frac{\partial g(u,v)}{\partial v}\right|_{(u,v)=(x_i^*, x_j^*)} \tag{5.13}$$

we have, to leading order,

$$\frac{d\epsilon_i}{dt} = \left[\alpha_i + \sum_j \beta_{ij} A_{ij}\right]\epsilon_i + \sum_j \gamma_{ij} A_{ij} \epsilon_j \tag{5.14}$$

[15] The standard reference on nonlinear dynamics for the sciences is the textbook by S. Strogatz [4].

which can be written in matrix form as

$$\dot{\boldsymbol{\epsilon}} = \boldsymbol{M}\boldsymbol{\epsilon} \tag{5.15}$$

where $\boldsymbol{M}$ has elements

$$M_{ij} = \delta_{ij}\left[\alpha_i + \sum_r \beta_{ir} A_{ir}\right] + \gamma_{ij} A_{ij} \tag{5.16}$$

Next, write $\boldsymbol{\epsilon}$ as a linear combination of the right eigenvectors $\boldsymbol{e}_r$ of $\boldsymbol{M}$, which we assume is diagonalisable:

$$\boldsymbol{\epsilon}(t) = \sum_r c_r(t)\boldsymbol{e}_r \tag{5.17}$$

Then (5.15) yields

$$\sum_r \dot{c}_r \boldsymbol{e}_r = \sum_r \mu_r c_r(t)\boldsymbol{e}_r \tag{5.18}$$

where μ_r is the r^{th} eigenvalue. According to the linear independence of the eigenvectors,

$$\dot{c}_r = \mu_r c_r(t) \quad \forall r \tag{5.19}$$

which implies

$$c_r(t) = c_r(0)e^{\mu_r t} \tag{5.20}$$

for each r. We see that if the real parts of all the eigenvalues are positive each of the c_r and so $\boldsymbol{\epsilon}$ grow exponentially: a small perturbation from $\boldsymbol{x}^*$ grows with time making the fixed point unstable and 'repelling.' On the other hand, when all the eigenvalues have negative real parts, the c_r and so $\boldsymbol{\epsilon}$ decays: the fixed point is stable and attracts nearby motions. If some μ_r have positive real parts and others negative then there will be growth in some directions and decay in the other, a situation known as a saddle point. Since stability necessitates the decay of a perturbation in *any* direction from the fixed point, we reach the important conclusion that a *fixed point is stable if the real parts of all the eigenvalues μ_r are negative.*[16]

5.2.2 Symmetrical Fixed Points and Special Cases

So far our discussion has been quite general. We now analyse a few particular cases in more detail. The first assumption we make is that we have a *symmetric* fixed point, meaning $x_i^* = x^*$ is the same for each i. This satisfies (see (5.10))

$$f(x^*) + \underbrace{\sum_j A_{ij}}_{=k_i} g\left(x^*, x^*\right) = f\left(x^*\right) + k_i g\left(x^*, x^*\right) = 0 \tag{5.21}$$

[16] When all are non-positive but some vanish, it is necessary to work to higher order in $\boldsymbol{\epsilon}$ to determine the stability.

Since $f(x^*)$ is a constant, for this to be satisfied for all i we must have $f(x^*) = 0$ and either $k_i = k$ for all i (all nodes have the same degree) or $g(x^*, x^*) = 0$. We assume the latter – the case of a k-regular graph is addressed in Example Sheet 1.

Having a symmetric fixed point simplifies the stability analysis greatly since the quantities α_i, β_{ij} and γ_{ij} defined in (5.13) become index independent and the leading order expansion (5.14) reads

$$\frac{d\epsilon_i}{dt} = (\alpha + \beta k_i)\,\epsilon_i + \gamma \sum_j A_{ij}\epsilon_j \tag{5.22}$$

There are two special cases we consider:

1. $g(x_i, x_j) = g(x_j)$ depends on its second argument alone. Then $\beta = 0$ and

$$\frac{d\epsilon_i}{dt} = \alpha\epsilon_i + \gamma \sum_j A_{ij}\epsilon_j \tag{5.23}$$

or

$$\dot{\boldsymbol{\epsilon}} = (\alpha \boldsymbol{I} + \gamma \boldsymbol{A})\,\boldsymbol{\epsilon} \tag{5.24}$$

where $\boldsymbol{A}$ is the adjacency matrix and $\boldsymbol{I}$ the identity. Clearly an eigenvector of $\boldsymbol{A}$ with eigenvalue κ_r is an eigenvector of $(\alpha\boldsymbol{I} + \gamma\boldsymbol{A})$ with eigenvalue $\alpha + \gamma\kappa_r$, so for stability, i.e. all the eigenvalues (which are real) to be negative, we require

$$\alpha + \gamma\kappa_r < 0 \tag{5.25}$$

for all r. Since the $\boldsymbol{A}$ necessarily has both positive and negative eigenvalues (Section 5.1), we must have $\alpha < 0$. Then

$$\gamma > 0 \to \kappa_r < -\frac{\alpha}{\gamma} \tag{5.26}$$

$$\gamma < 0 \to \kappa_r > -\frac{\alpha}{\gamma} \tag{5.27}$$

The right-hand side of the first inequality is positive and the second negative, so these hold for *all* r provided they hold for the most positive κ_1 and most negative κ_n eigenvalues of the adjacency matrix, respectively:

$$\kappa_1 < -\frac{\alpha}{\gamma} \quad (\gamma > 0) \quad \text{and} \quad \kappa_n > -\frac{\alpha}{\gamma} \quad (\gamma < 0) \tag{5.28}$$

These two conditions can be combined in

$$\frac{1}{\kappa_n} < -\frac{\gamma}{\alpha} < \frac{1}{\kappa_1} \tag{5.29}$$

where $\alpha = f'(x^*)$ and $\gamma = g'(x^*)$. This is a single sufficient condition for dynamics about the fixed point to be stable and is known as a *master stability condition.* It depends on both the nature of the dynamics prescribed by the functions f and g as well as the structure of the network described by the eigenvalues of $\boldsymbol{A}$.

We can of course make use of what we know about about the eigenvalues of $\boldsymbol{A}$. For example, κ_1 is bounded below by both $\sqrt{k_{\max}}$ and $\langle k \rangle$, so in the case $\gamma > 0$ if one increases the mean degree or maximum degree of the graph the system will eventually become unstable; there is a limit to the connectively of a graph supporting stable motion.

2. The coupling depends on its arguments in an antisymmetric way as $g(x_i, x_j) = g(x_i) - g(x_j)$. In this case $\gamma = -\beta$ and

$$\frac{d\epsilon_i}{dt} = (\alpha + \beta k_i)\,\epsilon_i - \beta \sum_j A_{ij}\epsilon_j \qquad \text{(from (5.14))} \tag{5.30}$$

$$= \alpha\epsilon_i + \beta \sum_j (k_i\delta_{ij} - A_{ij})\,\epsilon_j \tag{5.31}$$

$$\rightarrow \dot{\boldsymbol{\epsilon}} = (\alpha \boldsymbol{I} + \beta \boldsymbol{L})\,\boldsymbol{\epsilon} \tag{5.32}$$

and so stability is now controlled by the eigenvalues λ_r of the graph Laplacian:

$$\alpha + \beta\lambda_r < 0 \tag{5.33}$$

Since the smallest eigenvalue of the graph Laplacian is always zero (Section 5.1) we must again have $\alpha < 0$ and the remaining eigenvalues satisfying

$$\beta\lambda_r < -\alpha \rightarrow \frac{1}{\lambda_r} > -\frac{\beta}{\alpha} \tag{5.34}$$

which is true if the inequality holds for the largest eigenvalue λ_n, hence the master stability condition

$$\frac{1}{\lambda_n} > -\frac{\beta}{\alpha} = -\left.\frac{g'(x)}{f'(x)}\right|_{x=x^*} \tag{5.35}$$

5.2.3 Multiple Variables Per Node

The basic ideas are the same when there are m variables per node, only the notation becomes more cumbersome. A vector $\boldsymbol{x}^i$ with components x^i_μ is assigned to each node obeying

$$\dot{\boldsymbol{x}}^i = \boldsymbol{f}(\boldsymbol{x}^i) + \sum_j A_{ij}\boldsymbol{g}(\boldsymbol{x}^i, \boldsymbol{x}^j) \tag{5.36}$$

where $\boldsymbol{f}$ and $\boldsymbol{g}$ are now vector valued functions. Again assuming a symmetric fixed point $\boldsymbol{x}^i = \boldsymbol{x}^*$ and linearising with $\boldsymbol{x}^i = \boldsymbol{x}^* + \boldsymbol{\epsilon}^i$, we have, for each component ϵ^i_μ of $\boldsymbol{\epsilon}^i$

$$\frac{d\epsilon^i_\mu}{dt} = \left[\epsilon^i_1 \frac{\partial f_\mu(\boldsymbol{x})}{\partial x_1} + \epsilon^i_2 \frac{\partial f_\mu(\boldsymbol{x})}{\partial x_2} + \ldots\right]_{\boldsymbol{x}=\boldsymbol{x}^*}$$
$$+ \sum_j A_{ij}\left[\epsilon^i_1 \frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial u_1} + \epsilon^i_2 \frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial u_2} + \ldots + \epsilon^j_1 \frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial v_1} + \epsilon^j_2 \frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial v_2} + \ldots\right]_{\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{x}^*} \tag{5.37}$$

$$= \sum_\nu \left[\left(\left.\frac{\partial f_\mu(\boldsymbol{x})}{\partial x_\nu}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} + k_i \left.\frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial u_\nu}\right|_{\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{x}^*}\right)\epsilon^i_\nu + \sum_j A_{ij}\epsilon^j_\nu \left.\frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial v_\nu}\right|_{\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{x}^*}\right] \tag{5.38}$$

$$= \sum_{j\nu} \left[\delta_{ij}\left(\alpha_{\mu\nu} + k_i\beta_{\mu\nu}\right) + A_{ij}\gamma_{\mu\nu}\right]\epsilon^j_v \qquad 1 \le j \le n,\ 1 \le \nu \le m \tag{5.39}$$

where we noted that j runs over the index labelling the nodes and ν the index labelling the variables at each node, and defined the coefficients

$$\alpha_{\mu\nu} = \left.\frac{\partial f_\mu(\boldsymbol{x})}{\partial x_\nu}\right|_{\boldsymbol{x}=\boldsymbol{x}^*}, \quad \beta_{\mu\nu} = \left.\frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial u_\nu}\right|_{\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{x}^*}, \quad \gamma_{\mu\nu} = \left.\frac{\partial g_\mu(\boldsymbol{u},\boldsymbol{v})}{\partial v_\nu}\right|_{\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{x}^*} \tag{5.40}$$

We see that stability will be governed by the spectrum of some $nm \times nm$ matrix (writing (5.39) in matrix form is a matter of concatenating the $\boldsymbol{\epsilon}^i$ vectors to form a 'super' vector). We jump straight to the two special cases considered above:

1. $\boldsymbol{g}(\boldsymbol{x}^i, \boldsymbol{x}^j) = \boldsymbol{g}(\boldsymbol{x}^j) \to \beta_{\mu\nu} = 0 \ \forall \mu, \nu$ and

$$\frac{d\epsilon^i_\mu}{dt} = \sum_{j\nu} \left[\delta_{ij}\alpha_{\mu\nu} + A_{ij}\gamma_{\mu\nu}\right] \epsilon^j_\nu \tag{5.41}$$

At each $\mu = 1, \ldots, m$ we can expand the vector[17] $\boldsymbol{\epsilon}_\mu = (\epsilon^1_\mu, \epsilon^2_\mu, \ldots \epsilon^n_\mu)$ in terms of the eigenvectors $\boldsymbol{e}_r$ of the adjacency matrix:

$$\boldsymbol{\epsilon}_\mu = \sum_r c^r_\mu(t)\boldsymbol{e}_r \quad \text{i.e.} \quad \epsilon^i_\mu = \sum_r c^r_\mu(t) e^i_r \tag{5.42}$$

This expresses the perturbation in each variable x^i_μ as a linear combination of the eigenvectors similar to before, only with a separate set of coefficients c^r_μ for each type of variable. Substituting we find

$$\sum_r \frac{dc^r_\mu}{dt} e^i_r = \sum_{rj\nu} \left[\delta_{ij}\alpha_{\mu\nu} + A_{ij}\gamma_{\mu\nu}\right] c^r_\nu(t) e^j_r \tag{5.43}$$

$$= \sum_{r\nu} \left[\alpha_{\mu\nu} + \kappa_r\gamma_{\mu\nu}\right] c^r_\nu(t) e^i_r \tag{5.44}$$

Again note $r \in \{1, \ldots, n\}$ while $\nu \in \{1, \ldots, m\}$. At each r we have

$$\dot{\boldsymbol{c}}^r = \left[\boldsymbol{\alpha} + \kappa_r\boldsymbol{\gamma}\right] \boldsymbol{c}^r \tag{5.45}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are $mm \times mm$ matrices[18] with elements $\alpha_{\mu\nu}$ and $\gamma_{\mu\nu}$, and $\boldsymbol{c}^r = (c^r_1, c^r_2, \ldots, c^r_m)$ (this is a vector in the sense $\boldsymbol{x}^i$ is, not $\boldsymbol{\epsilon}_\mu$). We we can view this vector equation, which captures dynamics in the vicinity of the fixed point, as a decoupled set of n separate systems, one for each eigenvector κ_r of the adjacency matrix. In order for the fixed point of the entire system to be stable, each of these individual systems must be stable, meaning their eigenvalues need to have negative real parts.

With this in mind define the *master stability function* $\sigma(\kappa)$ to be the most positive real part of an eigenvalue of $\boldsymbol{\alpha} + \kappa\boldsymbol{\gamma}$, where κ is a real variable. This function is straightforward to calculate

[17] This is a vector with components across the upper index i which enumerates the nodes, rather than a vector across the lower index μ at fixed i (such as $\boldsymbol{\epsilon}^i$) which enumerates the variables at any particular node.

[18] These are just the Jacobian matrices of $\boldsymbol{f}$ and $\boldsymbol{g}$, so it is unsurprising to see them in the equation determining the stability of the fixed point.

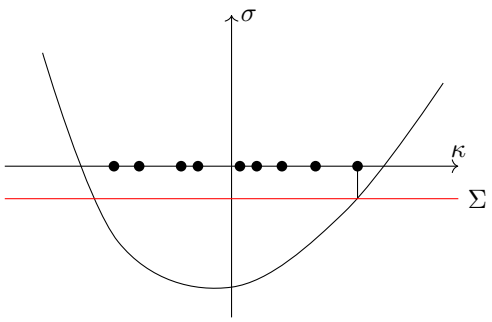


Figure 5.2: Sketch of a simple monotonic master stability function. At each eigenvalue κ_r of the adjacency matrix (black dot) σ is negative, meaning the associated fixed point is stable. Equivalently, note that the 'super master' stability function, indicated here with a horizontal (red) line, has a negative value.

numerically: solving an eigenvalue problem has polynomial time complexity and at any rate m is unlikely to be large. Overall stability then requires $\sigma(\kappa_r) < 0$ for each r or

$$\Sigma = \max_r \sigma\left(\kappa_r\right) < 0 \tag{5.46}$$

Being droll, we might refer to Σ as a 'super master' stability function. A simple example is shown in Figure 5.2.

2. $\boldsymbol{g}(\boldsymbol{x}^i, \boldsymbol{x}^j) = \boldsymbol{g}(\boldsymbol{x}^i) - \boldsymbol{g}(\boldsymbol{x}^j) \Rightarrow \gamma_{\mu\nu} = -\beta_{\mu\nu}$. The analysis is the same, except starting from

$$\frac{d\epsilon^i_\mu}{dt} = \sum_{j\nu} \left[\delta_{ij}\alpha_{\mu\nu} + L_{ij}\gamma_{\mu\nu}\right] \epsilon^j_\nu \tag{5.47}$$

and resulting in

$$\dot{\boldsymbol{c}}^r = \left[\boldsymbol{\alpha} + \lambda_r \boldsymbol{\beta}\right] \boldsymbol{c}^r \tag{5.48}$$

so that the appropriate master stability function $\sigma(\lambda)$ is the most positive real part of an eigenvalue of $\boldsymbol{\alpha} + \lambda\boldsymbol{\beta}$, and if this function is negative for all eigenvalues λ_r of the graph Laplacian the fixed point is stable.

5.2.4 Limit Cycles and Synchronisation

An important possibility of we have not yet discussed is that of a limit cycle, a periodic motion of the variables of the system which, like a fixed point, can be stable or unstable, attracting or repelling. Naturally in a real system only stable limit cycles are observed, since for an unstable cycle external perturbations, no matter how small, will always be destabilising and result in the eventual breakdown of the periodic motion.

We are particularly interested in the phenomena of *synchronisation,* when a system of coupled oscillators undergo periodic motion *in union* with each other – a special type of limit cycle where it is not just the overall system that cycles through the same set of values, but also the variables at each node, doing so in time with each other.

In our study of the dynamical system

$$\frac{d\boldsymbol{x}^i}{dt} = \boldsymbol{f}(\boldsymbol{x}^i) + \sum_j A_{ij}\boldsymbol{g}(\boldsymbol{x}^i, \boldsymbol{x}^j) \tag{5.49}$$

on a network, the coupling between nodes is essential for the possibility of synchronisation. To see this is the case, suppose $A_{ij} = 0$ and that the equation

$$\frac{d\boldsymbol{x}^i}{dt} = \boldsymbol{f}(\boldsymbol{x}^i) \tag{5.50}$$

has a periodic solution $\boldsymbol{x}^i = \boldsymbol{s}(t)$ such that $\boldsymbol{s}(t+\tau) = \boldsymbol{s}(t)\ \forall t$ (the equivalent of a symmetric fixed point for limit cycles). Then $\boldsymbol{s}(t+\phi)$ is also a solution, for any value of ϕ. Consequently we can have solutions of the form

$$\boldsymbol{x}^i = \boldsymbol{s}(t+\phi_i) \tag{5.51}$$

where ϕ_i has a different value for each i i.e. all nodes follow the same motion with the same period but not, in general, in time with each other. This is intuitive: without interactions the oscillator at one node has no no way of knowing what its neighbours are doing!

Now consider adding interactions, for simplicity in the form $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(y)$ we have considered several times already. Then $\boldsymbol{x}^i = \boldsymbol{s}(t)$ for all i is a solution of (5.49) (but $\boldsymbol{x}^i = \boldsymbol{s}(t+\phi_i)$ isn't), since

$$\boldsymbol{g}(\boldsymbol{x}^i, \boldsymbol{x}^j) = \boldsymbol{s}(t) - \boldsymbol{s}(t) = 0 \tag{5.52}$$

Thus we have a synchronised state where all nodes are following the same periodic motion and at the same point in that motion at any given time. This is just one possible way in which synchronisation can arise; we will see another in the form of the Kuramoto model in the next section.

The stability of a limit cycle can be analysed in a similar way to that of a fixed point. For our simple synchronised state we write $\boldsymbol{x}^i(t) = \boldsymbol{s}(t) + \boldsymbol{\epsilon}^i$ and linearise (5.49) to obtain

$$\frac{d\epsilon^i_\mu}{dt} = \sum_{j\nu} \left[\delta_{ij}\alpha_{\mu\nu}(t) + L_{ij}\beta_{\mu\nu}(t)\right] \epsilon^j_\nu \tag{5.53}$$

as in (5.47), except now

$$\alpha_{\mu\nu}(t) = \left.\frac{\partial f_\mu}{\partial x_\nu}\right|_{\boldsymbol{x}=\boldsymbol{s}(t)}, \quad \beta_{\mu\nu}(t) = \left.\frac{\partial g_\mu}{\partial x_\nu}\right|_{\boldsymbol{x}=\boldsymbol{s}(t)} \tag{5.54}$$

are evaluated at $\boldsymbol{x} = \boldsymbol{s}(t)$ rather than $\boldsymbol{x}^*$, and so are *time-dependent.* Moreover, they are periodic with period τ. Repeating the steps in the analysis of above (expanding $\boldsymbol{\epsilon}_\mu = (\epsilon^1_\mu, \ldots, \epsilon^n_\mu)$ in the eigenbasis of $\boldsymbol{L}$), one obtains the matrix equation

$$\frac{d\boldsymbol{c}^r}{dt} = \left[\boldsymbol{\alpha}(t) + \lambda_r\boldsymbol{\beta}(t)\right] \boldsymbol{c}^r(t) \tag{5.55}$$

Due to the time dependence of the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the solutions for the coefficients $c^r_\mu(t)$ are no longer (in an appropriate coordinate system) simply exponential functions, making a complete

analysis challenging. However, we can quite easily derive a *sufficient* condition for stability by defining the stability function $\sigma_t(\lambda)$ as the most positive real part of an eigenvalue of $\alpha + \lambda\beta$ *at time* t. This varies with period τ and at any given time the stability calculation "$\sigma_t(\lambda_r) < 0\ \forall r$?" tells us whether the perturbations are decaying or growing at that instant. If $\sigma_t(\lambda_r)$ is negative for all r at all times then the cycle is certainly stable, hence the sign of the master stability functions

$$\sigma(\lambda) = \max_{t\in[0,\tau]} \sigma_t(\lambda), \quad \Sigma = \max_r \sigma(\lambda_r) \tag{5.56}$$

can be used to confirm the stability of a cycle: if $\Sigma < 0$ then the motion is stable, with perturbations decaying in all directions at every point along the cycle. To reiterate, this is a sufficient condition, not a necessary one: overall stability does not require that the system be instantaneously stable at every moment. Indeed, it is possible for perturbations around the synchronised state to grow at certain times, provided they decay (more rapidly) at others so that in total the motion is stable. While it may be possible to obtain a stronger condition in certain cases, often one must resort to numerical simulation in order to determine stability in a complex system.

6 Networks for the Minimisation of Spin Hamiltonians

Having studied the general features of dynamical systems on networks, we now examine three specific networks that will be relevant to our proposed simulator i.e. the minimisation of the Ising and XY Hamiltonians (repeated below for convenience). Referring back to Figure 1.1, this section concerns the 'mathematical model' box and subsequent sections the physical realisation.

$$H_{\mathrm{XY}} = -\sum_{ij} J_{ij}\cos\left(\theta_i - \theta_j\right) \qquad \boldsymbol{s}_i = \left(\cos\theta_i, \sin\theta_i\right),\ \theta_i \in [0, 2\pi) \tag{6.1}$$

$$H_I = -\sum_{ij} J_{ij} s_i s_j \qquad s_i \in \{-1, 1\} \tag{6.2}$$

6.1 Hopfield

We consider a Hopfield network to be a network in which continuous inputs u_i are related to an output via a sigmoid function such as $v_i = g(u_i) = \tanh u_i$ (Figure 6.1a). The dynamics are normally prescribed for the input, which evolves as

$$\frac{du_i}{dt} = \sum_j J_{ij} v_j - u_i + h_i \tag{6.3}$$

where h_i is an external field and J_{ij} is symmetric. Observe that $\dot{u}_i$ may be written as a gradient:

$$\dot{u} = -\frac{\partial E}{\partial v_i} \quad \text{or} \quad \dot{\boldsymbol{u}} = -\nabla_{\boldsymbol{v}} E \tag{6.4}$$

where

$$E = -\frac{1}{2}\sum_{ij} J_{ij} v_i v_j + \sum_i \int_0^{v_i} g^{-1}(v)dv + \sum_i h_i v_i \tag{6.5}$$

We identify E as a energy (or, in the context of dynamical systems, a *Lyapunov function*), since it is bounded below, monotonically decreasing and such that the system, via the equations of motions (6.3), will evolve to one of its (possibly local) minima. To see this, note

$$\frac{d}{dt}\int^{v_i(t)} g^{-1}(v)dv = \frac{dv_i}{dt}g^{-1}(v_i) \equiv \frac{dv_i}{dt}u_i \tag{6.6}$$

hence

$$\frac{dE}{dt} = -\sum_i \frac{dv_i}{dt}\left[\sum_j J_{ij}v_j - u_i + h_i\right] \tag{6.7}$$

$$= -\sum_i \frac{dv_i}{dt}\frac{du_i}{dt} \tag{6.8}$$

$$= -\sum_i \left(g^{-1}\right)'(v_i)\left(\frac{dv_i}{dt}\right)^2 \tag{6.9}$$

$$\leq 0 \tag{6.10}$$

Therefore $\dot{u}_i = 0 \to \dot{E} = 0$ and E is decreasing so the system evolves to attain a minimum of E.

You would have noticed that E is close to, but not quite, an Ising Hamiltonian. To remove the anonymous integral terms, we rescale the argument of g using a positive parameter λ (Figure 6.1b) as

$$v_i = g(\lambda u_i),\ \lambda > 0 \Rightarrow u_i = \frac{1}{\lambda}g^{-1}(v_i) \tag{6.11}$$

Taking $\lambda \to \infty$,

$$\frac{du_i}{dt} = \sum_j J_{ij}v_j - \frac{u_i}{\lambda} + h_i \to \sum_j J_{ij}v_j + h_i \tag{6.12}$$

and

$$E = -\frac{1}{2}\sum_{ij} J_{ij}v_iv_j + \frac{1}{\lambda}\sum_i \int_0^{v_i} g^{-1}(v)dv + \sum_i h_iv_i \tag{6.13}$$

$$\to -\frac{1}{2}\sum_{ij} J_{ij}v_iv_j + \sum_i h_iv_i \tag{6.14}$$

and an Ising Hamiltonian emerges.

Hopfield networks provide a model for understanding human memory but for our purposes are crucial to the operation of the coherent Ising machine discussed in Section 11. Note that although the system will converge to a minimum of H_I, we have not established a reason for it to converge to a global one, and indeed one of the drawbacks of the Ising machine we discuss is that local minima are likely to be found.

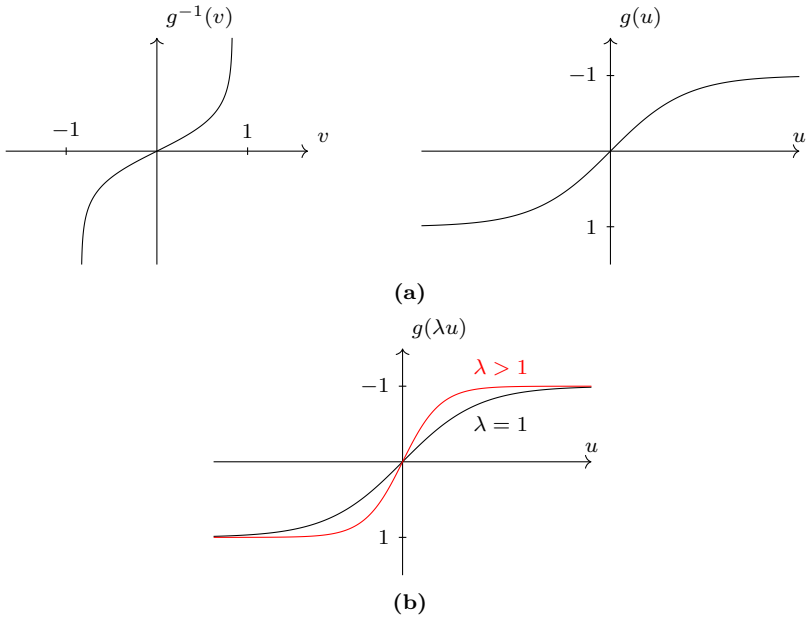


Figure 6.1: (a) Input-output relations for the Hopfield network and (b) scaling the input with $\lambda > 1$ results in a steepening of the output profile.

6.2 Kuramoto Model

The Kuramoto model describes a large number of coupled oscillators which tend to *synchronise* (Section 5.2.4). Each oscillator is characterised by a phase θ_i and a intrinsic frequency ω_i with governing equations

$$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_j \sin(\theta_j - \theta_i) \tag{6.15}$$

where N is the total number of oscillators and K controls the strength of the sinusoidal couplings.

The collective dynamics of the entire population and conveniently described by a single complex order parameter

$$r(t)e^{i\phi(t)} := \frac{1}{N} \sum_j e^{i\theta_j(t)} \tag{6.16}$$

$\phi(t)$ defines the average phase and r the extent of phase coherence, with $r \to 0$ and $r \to 1$ corresponding to the limits of incoherent and coherent oscillation, respectively. Multiplying by $e^{-i\theta_i}$, from the imaginary part

$$r \sin(\phi - \theta_i) = \frac{1}{N} \sum_j \sin(\theta_j - \theta_i) \tag{6.17}$$

and so the equations of motion can be written

$$\dot{\theta}_i = \omega_i + Kr \sin(\phi - \theta_i) \tag{6.18}$$

in which each oscillator interacts with all others through the collective quantities r and ϕ only. A positive feedback loop is evident in which, as r increases i.e. the population becomes more coherent so does the strength of the coupling, tending to increase the coherence further.

Further details of the model are discussed in the second example sheet.[19] Here we note that if a steady state solution for the order parameter is sought where $r(t)$ is a constant and $\phi(t)$ rotates uniformly at frequency Ω, then moving to a coordinate system rotating at Ω we have $\phi \equiv 0$ and from the equation

$$\dot{\theta}_i = \omega_i - kr \sin\theta_i \tag{6.19}$$

we see that those oscillators such that $\omega_i < Kr$ have the possibility of becoming phase locked $\dot{\theta}_i = 0 \rightarrow \sin\theta_i = Kr/\omega_i$, which corresponds to rotation with a constant phase difference i.e. coherent oscillation in the original frame. Meanwhile those oscillators with $\omega_i > Kr$ have changing phase and continue to drift around unit circle in the rotating frame. Thus, a state of *partial synchronisation.* As K is increased more and more oscillators are recruited into the coherent pack, until eventually $r \approx 1$ in the steady state. In general there are (constant) phase differences between the oscillators but in the case $\omega_i = \omega \ \forall i$ these are zero and for large enough coupling the system can exhibit a perfectly synchronised state in the sense described in the previous section.

Finally, in the original form of the equation of motions (6.15) we recognise the XY Hamiltonian:

$$\dot{\theta}_i = \omega_i - \sum_j \sin(\theta_i - \theta_j) \tag{6.20}$$

$$= \omega_i + \frac{1}{2}\frac{\partial}{\partial\theta}\sum_{kj} J_{kj}\cos(\theta_k - \theta_j) \tag{6.21}$$

$$= \omega_i - \frac{1}{2}\frac{\partial H_{\mathrm{XY}}}{\partial\theta_i} \tag{6.22}$$

In other words, the dynamics follow the gradient descent of the XY Hamiltonian such that the long term steady state behaviour described above corresponds to a minimum of H_{XY}.

6.3 Stuart-Landau Oscillators

We have now seen how both the Ising model and the XY Hamiltonian can be realised in the dynamics of a network. The third model we look at, that of Stuart-Landau (SL), has the potential to minimise both and captures the essential dynamics of the physical non-equilibrium systems we consider for this purpose. The model describes the state of N oscillators $\phi_i = \sqrt{\rho_i}e^{i\theta_i}$ with both phase θ_i and occupation ρ_i via the coupled equations

$$\dot{\psi}_i = \underbrace{\left(\gamma_i \overbrace{-|\psi_i|^2}^{\text{non-linear dissipation}}\right)}_{\text{gain}}\psi_i + \underbrace{\sum_j J_{ij}\psi_j}_{\text{coupling}} \tag{6.23}$$

As indicated, the dynamics intrinsic to a single oscillator has both a positive and negative term, the former corresponding to a gain and the latter a loss, specifically non-linear dissipation, controlling

[19] Extensive analysis of the model and Kuramoto's original work may be found in Strogatz's millennium review [5].

the growth of ψ_i. Simple interactions between oscillators are described by the matrix J_{ij} which we assume to be real.

To analyse this system, multiply both sides of (6.23) by $\sqrt{\rho_i}e^{-i\theta_i}$ and separate real and imaginary parts:

$$\sqrt{\rho_i}e^{-i\theta_i}\left(\frac{1}{2\sqrt{\rho_i}}\dot{\rho}_i e^{i\theta_i} + i\sqrt{\rho_i}\dot{\theta}_i e^{i\theta_i}\right) = (\gamma_i - \rho_i)\,\rho_i + \sum_j J_{ij}\sqrt{\rho_i\rho_j}e^{i(\theta_j-\theta_i)} \tag{6.24}$$

$$\Rightarrow \frac{1}{2}\dot{\rho}_i = (\gamma_i - \rho_i)\,\rho_i + \sum_j J_{ij}\sqrt{\rho_i\rho_j}\cos(\theta_i - \theta_j) \tag{6.25}$$

and

$$\dot{\theta}_i = -\sum_j J_{ij}\sqrt{\frac{\rho_j}{\rho_i}}\sin(\theta_i - \theta_j) \tag{6.26}$$

The phase equations look similar to the couplings in the Kuramoto model (6.15) (which we know relate to H_{XY}) except for the ratio of occupations $\sqrt{\rho_j/\rho_i}$. To address this, the gains γ_i should be adjusted in order to drive all oscillators to the *same* occupancy ρ_{th} (the subscript abbreviates 'threshold'). A simple scheme would have the gains vary according to

$$\dot{\gamma}_i = (\rho_{\text{th}} - \rho_i)\,\gamma_i \tag{6.27}$$

such that $\rho_i = \rho_{\text{th}}$ provides a fixed point at each i (Figure 6.2). Then (6.26) describes Kuramoto oscillators whilst the value of ρ_{th}, known *a priori* (i.e. chosen), may be related to the gains and couplings via (6.25) at threshold:

$$\rho_i = \rho_{\text{th}} \Rightarrow (\gamma_i - \rho_{\text{th}})\,\rho_{\text{th}} + \rho_{\text{th}}\sum_j J_{ij}\cos(\theta_i - \theta_j) = 0 \tag{6.28}$$

$$\Rightarrow \rho_{\text{th}} = \gamma_i + \sum_j J_{ij}\cos(\theta_i - \theta_j) \tag{6.29}$$

and so the *total occupancy* $N\rho_{\text{th}}$, which typically relates directly to a density or intensity (e.g. of light emitted) in a real system, is

$$\underbrace{N\rho_{\text{th}}}_{\text{fixed}} \equiv \sum_i \rho_i = \underbrace{\sum_{i=1}^{N}\gamma_i}_{\text{controlled}} + \underbrace{\sum_{ij} J_{ij}\cos(\theta_i - \theta_j)}_{-H_{\text{XY}}} \tag{6.30}$$

Thus, if we can adjust the γ_i to *minimise* the total gain at fixed occupancy, $-H_{\text{XY}}$ is maximised hence H_{XY} *minimised.* One has the picture of increasing the gain from below the energy profile of H_{XY} (Figure 6.3); close to or at threshold the evolution of each phase follows the steepest descent to a minimum of H_{XY}. Establishing a procedure to do this – in particular to reach a *global* minimum – is not trivial due to the dynamic character of the gains.[20]

[20] There are many clever proposals. A small parameter ϵ is normally inserted as $\dot{\gamma}_i = \epsilon(\ldots)$ so that changes are slow and the system has time to explore many minima. Noise may also be added to the SL equations (6.23) to prevent the system from becoming stuck at a non-global minimum (by rendering such minima unstable).

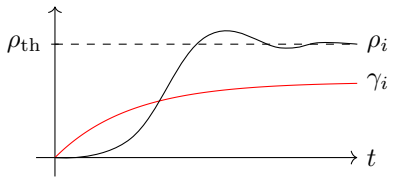


Figure 6.2: Density of a single oscillator with $\psi_i(0) = 0$ as the i^{th} gain is varied. γ_i is adjusted according to (6.27) so that the final value of ρ_i is ρ_{th}.

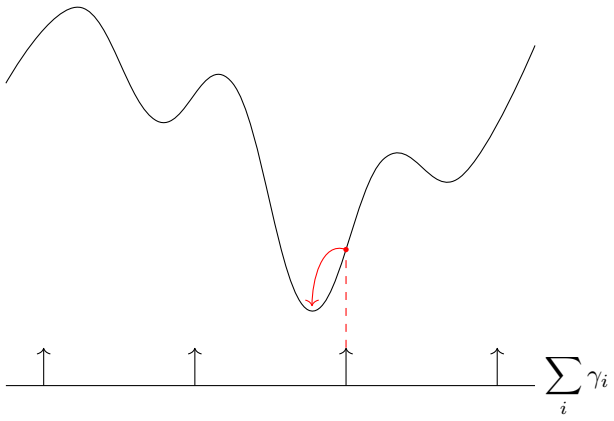


Figure 6.3: Minimising the gain to minimise H_{XY}. Arrows indicate possible choices for increasing the gain. Only some (one shown) will result in subsequent evolution to the ground state.

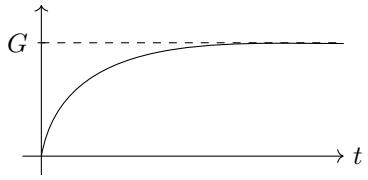


Figure 6.4: Amplitude $h(t)$ of the second resonant term in (6.23).

6.3.1 Second Resonance

Remembering that the Ising Hamiltonian can be obtained from the XY model in the limit that θ_i becomes a discrete variable (Section 2.1), it is simple to augment the SL-model to realise H_I instead of H_{XY}: add a conjugate term $h(t)\psi_i^*$, known as a *second resonance* term, to the right-hand side of (6.23). As the name implies, this results in a term of double frequency in the equations for ρ_i and θ_i:

$$\frac{1}{2}\dot{\rho}_i = (\gamma_i - \rho_i)\,\rho_i + \sum_j J_{ij}\sqrt{\rho_i\rho_j}\cos\left(\theta_i - \theta_j\right) + h(t)\cos(2\theta_i) \tag{6.31}$$

and

$$\dot{\theta}_i = -\sum_j J_{ij}\sqrt{\frac{\rho_j}{\rho_i}}\sin(\theta_i - \theta_j) - h(t)\sin(2\theta_i) \tag{6.32}$$

The function $h(t)$ is controlled to converge to some constant value G (Figure 6.4). At the fixed point we then have the total long-time density

$$N_{\rho_{\text{th}}} = \sum_i \gamma_i + \sum_{ij}\left[J_{ij}\cos\left(\theta_i - \theta_j\right) + G\cos(2\theta_j)\right] \tag{6.33}$$

The effect of the second resonance term then is, for sufficiently large G, to restrict each θ_i to the values 0 and π, since deviations from these values are now penalised. Consequently, we recover an Ising Hamiltonian as $\cos(\theta_i - \theta_j) \to \cos\theta_i\cos\theta_j = s_i s_j$. Successfully minimising the gain now minimises H_I.

One could instead envision adding a third or fourth etc. resonant term ($\psi_i^{*(n-1)}$ for n^{th} resonance), resulting in the so-called Potts models where θ_i is restricted to q values. Realising higher resonances in a physical system (i.e. pumping at $q\theta_i$) is technologically challenging however.

That concludes are look at the SL model, which is applicable to a myriad of real systems. Indeed, the majority of systems we investigate in subsequent sections will have SL oscillators at their core.

7 Physical Simulators: Equilibrium Systems

While later on we will focus on non-equilibrium systems, in particular for the purpose of constructing a polariton graph simulator, it is worth addressing the role *equilibrium* systems may have in simulations via the protocol of *Quantum Adiabatic Optimisation* (QAO), not least because this has long been successfully implemented in a real device.

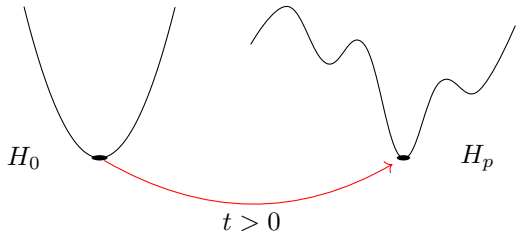


Figure 7.1: In AQO the Hamiltonian perturbed adiabatically so that a system initially in the ground state of $H_0 = H(0)$ arrives in the ground state of $H_p = H(T)$ after a time T.

7.1 Quantum Adiabatic Optimisation

Suppose there is a quantum Hamiltonian H_p whose ground state encodes the solution of some optimisation problem and a second Hamiltonian H_0 whose ground state is both known and easy to prepare. For example, H_p could be a quantum version one of Ising Hamiltonians set out in Section 4 and H_0 the same Hamiltonian but with identical short-range ferromagnetic couplings only.

The idea of QAO is to start with the system in the ground state of H_0 and change the Hamiltonian from H_0 to H_p over some time T (Figure 7.1). The *quantum adiabatic theorem* tells us that, provided this is done slowly enough i.e. adiabatically,[21] the system will remain in its instantaneous eigenstate throughout the process and so arrive at the ground state of H_p at $t = T$.

Of course the crucial question is *how slowly.* It is typically found [6] that that the process needs to be exponentially slow in the problem size n:

$$T = O\left(e^{\alpha n^\beta}\right) \quad \text{as} \quad n \to \infty \tag{7.1}$$

where $\alpha, \beta > 0$ depend on the structure of the problem and the physical system in use. Naturally, the hope is that these parameters are smaller than those for classical algorithms used to solve the problem.

Exactly how to perform the adiabatic change e.g. tune the couplings J_{ij} is an active area of research. A simple scheme might have the Hamiltonian changing linearly:

$$H(t) = \left(1 - \frac{t}{T}\right) H_0 + \frac{t}{T} H_p \tag{7.2}$$

but this is not the only approach (Figure 7.2).

The 'D-Wave Systems' company have successfully implemented a form of AQO in superconducting integrating circuits, although the jury is still out as to whether this form of quantum computing offers any speed-up over classical algorithms. Hence we are prompted to explore other possibilities.

[21] An adiabatic process is one (a series of sufficiently small perturbations) performed sufficiently slowly *such that* the system remains in an instantaneous eigenstate at all times. Note that this is different from the use of the term in classical statistical mechanics: the quantum definition is closer to the thermodynamic concept of *quasistatic change* (infinitesimally close to equilibrium at every instant) and has *no* direct relation to heat exchange.

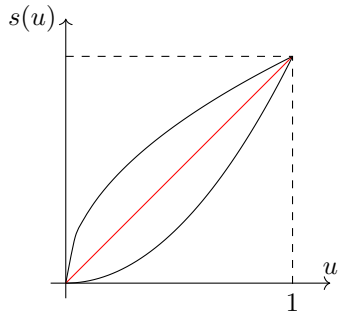


Figure 7.2: The Hamiltonian is generally changed according to

$$H(t) = (1 - s(t/T))\, H_0 + s(t/T) H_p$$

where s is increasing on $[0, 1]$. Three different possibilities for this function are shown with the linear case (7.2) in red.

8 Condensation in Physical Systems

In this section we investigate the phenomena of Bose-Einstein condensation. It is worth stepping back to consider the nature of the systems considered in this course (Figure 8.1). We will start with the equilibrium – atomic condensates – before making our way to exciton-polariton condensates, which are out-of-equiiblrium systems with driving and dissipation, although in some sense not as far from equilibrium as other lossy systems such as lasers.

We take a heuristic approach, seeking to understand the basis principles and develop an operational theory, not belabour the detailed many-body physics. We begin with an overview of the length scales characterising different views of a gas.

8.1 Matter Fields

Refer to Figure 8.2. On one end of the spectrum we have the classical, mean field description of a gas in terms of macroscopic (coarse-grained) variables such as $\rho(\boldsymbol{x}, t)$, $v(\boldsymbol{x}, t)$ and thermodynamic variables of state (pressure, entropy etc.). On the other we have the quantum description where the system is fully specified by a many-body wavefunction $\psi(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N, t)$. The systems we study will be quantum in nature, but at the same time well described by a macroscopic field of a single spatial variable, due to the phenomenon of Bose-Einstein condensation.

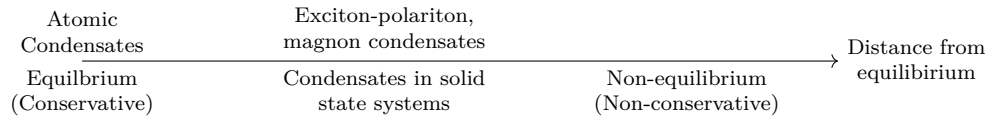


Figure 8.1: Range of systems considered in this course and several examples (top). Equilibrium systems are conservative with respect to energy and particle number, non-conservative systems are not. In between the two extremes are a range of near- and out-of-equilibrium systems.

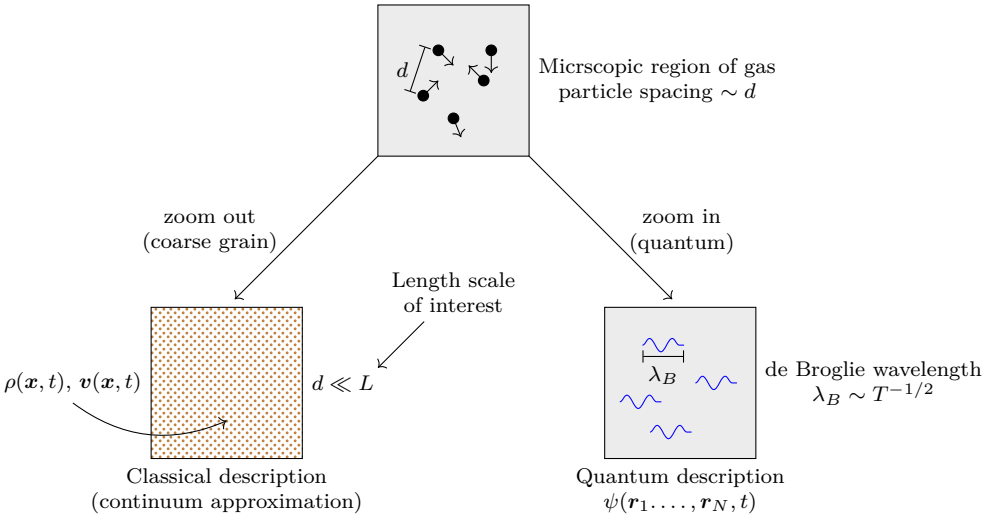


Figure 8.2: Length scales in descriptions of a gas.

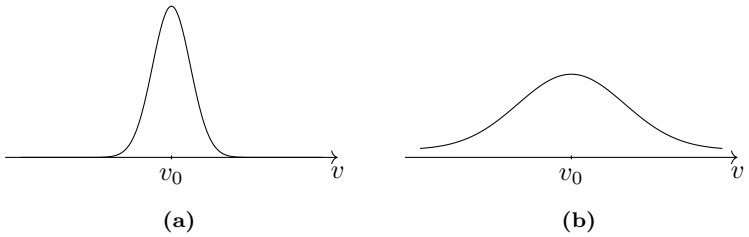


Figure 8.3: (a) The experimental signature of Bose-Einstein condensation is a sharply peaked velocity or momentum distribution as many particles occupy the same quantum state. (b) This is in contrast to the broad statistical distribution (Maxwell-Boltzmann) of a thermal gas.

8.2 Bose-Einstein Condensation

Two important parameters in a system of bosons (particles allowed to share the same quantum state) are the inter-particle spacing d and thermal de Broglie wavelength λ_B, which is roughly the average de Broglie wavelength of the particles at a given temperature.[22] It is observed that, below a critical temperature T_c at which $\lambda_B \sim d$, the ground state becomes *macroscopically occupied* meaning many $N \gg 1$ particles occupy the lowest energy quantum state (Figure 8.3). One can imagine that, as $\lambda_B \sim d$, the wavefunctions of individual particles overlap in space to produce a single, coherent matter wave (collective mode). This state of matter, the Bose-Einstein condensate, can consequently be described by a single field or wavefunction $\psi(\boldsymbol{r}, t)$.

[22] For more on particle statistics and wave-particle duality, consult any introductory quantum mechanics textbook.

8.3 Modelling Atomic Condensates

8.3.1 The Gross-Pitaevskii Equation

Atomic condensates at ultracold temperatures ($T \ll T_c$) can be accurately described by a complex order parameter or condensate wavefunction $\psi(\boldsymbol{r}, t)$ which is the solution to the non-linear Schrödinger equation (NLSE)

$$i\hbar\frac{\partial\psi(\boldsymbol{r},t)}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r},t) + U_0|\psi(\boldsymbol{r},t)|^2\right)\psi(\boldsymbol{r},t) \tag{8.1}$$

where $V(\boldsymbol{r}, t)$ is the external, trapping potential and $U_0 > 0$ ($U_0 < 0$) a repulsive (attractive) contact potential. The latter models pairwise interactions between particles in the condensate as

$$U(\boldsymbol{r}_1 - \boldsymbol{r}_2) = U_0\delta(\boldsymbol{r}_1 - \boldsymbol{r}_2) \tag{8.2}$$

This is a reasonable approximation for the actual interactions provided the condensate gas is *dilute,* meaning $d, \lambda_B \gg a_s$ where the *s-wave scattering length* a_s offers an effective length scale of interactions.

(8.1) known as the Gross-Pitaevskii Equation (GPE) equation and is found to model a large variety of systems not limited to condensates (e.g. plasmas and water waves). We reason it at a phenomenological, mean-field level, as arising from the time-dependent variational principle applied to the energy functional

$$E = \int\left\{\frac{\hbar^2}{2m}|\nabla\psi|^2 + V|\psi|^2 + \frac{U_0}{2}|\psi|^4\right\}d\boldsymbol{r} \quad \rightarrow \quad i\hbar\frac{\partial\psi}{\partial t} \equiv \frac{\delta E[\psi,\psi^*]}{\delta\psi^*} \tag{8.3}$$

The terms of E describe kinetic, potential and interaction energies respectively. In the case $V(\boldsymbol{r}, t) = V(\boldsymbol{r})$ is time-independent, E is conserved. The modulus squared $|\psi|^2$ has the interpretation of the particle density $n(\boldsymbol{r}, t)$ so that

$$\int|\psi|^2 d\boldsymbol{r} = N \tag{8.4}$$

describes the total number of particles in the condensate and mN its mass, with N taken to be fixed and large. Moreover, the condensate is ascribed a momentum

$$\boldsymbol{p} = \frac{i\hbar}{2}\int\{\psi\nabla\psi^* - \psi^*\nabla\psi\}\,d\boldsymbol{r}, \tag{8.5}$$

and this is also conserved (refer to the second example sheet).

8.3.2 Hydrodynamic Description

Frequently it is useful to write ψ in polar form $\psi(\boldsymbol{r}, t) = \sqrt{n(\boldsymbol{r}, t)}e^{iS(\boldsymbol{r},t)}$ (in this context referred to as the *Madelung transformation*), where in addition to the density we now have the phase $S(\boldsymbol{r}, t)$ with which is associated a flow velocity

$$\boldsymbol{v} = \frac{\hbar}{m}\nabla S \tag{8.6}$$

in line with the definition of current in single-particle quantum mechanics.

There is a nice analogy for BEC with inviscid, irrotational fluid flow, as seen from substituting the Madelung transformation into the GPE (8.1) and separating real and imaginary parts (Appendix B):

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\boldsymbol{v}) = 0 \tag{8.7}$$

$$m\frac{\partial \boldsymbol{v}}{\partial t} = -\nabla\left(\frac{1}{2}mv^2 + V + U_0 n - \frac{\hbar^2}{2m}\frac{\nabla^2\sqrt{n}}{\sqrt{n}}\right) \tag{8.8}$$

which are just the Euler equations with an anomalous 'quantum pressure' term. A quantum addition is also manifest in the energy of the flow

$$\int\left\{\underbrace{\frac{\hbar^2}{2m}\left(\nabla\sqrt{n}\right)^2}_{\text{Quantum KE}} + \underbrace{\frac{1}{2}mnv^2}_{\substack{\text{Classical}\\ \text{KE}}} + \underbrace{Vn}_{\text{PE}} + \underbrace{\frac{U_0}{2}n^2}_{\substack{\text{Interaction}\\ \text{energy}}}\right\}d\boldsymbol{r} \tag{8.9}$$

8.3.3 Stationary Solutions

Assuming a time-independent potential, looking for stationary solutions $\psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})e^{-i\mu t/\hbar}$ to (8.1) we can derive a *time-independent* GPE

$$\mu\psi = -\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r}) + V(\boldsymbol{r})\psi(\boldsymbol{r}) + U_0|\psi|^2\psi(\boldsymbol{r}) \tag{8.10}$$

The constant μ is known as the chemical potential and has the interpretation of energy per particle added to the condensate, as seen by multiplying (8.10) by $\psi^*(\boldsymbol{r})$ and integrating over $\boldsymbol{r}$:

$$\mu\int|\psi|^2 d\boldsymbol{r} = \int\left\{\frac{\hbar^2}{2m}|\nabla\psi|^2 + V|\psi|^2 + U_0|\psi|^4\right\}d\boldsymbol{r} \tag{8.11}$$

$$\Rightarrow \mu = \frac{1}{N}\left(E_K + E_P + 2E_{\text{int}}\right) \tag{8.12}$$

In fact, continuing from our reasoning for the GPE we can see the time-independent GPE as arising from the minimisation of E subject to a constant number of particles:

$$\frac{\delta}{\delta\psi^*}\left[E - \mu N\right] = 0 \quad \rightarrow \quad \frac{\delta E}{\delta\psi^*} = \mu\frac{\delta}{\delta\psi^*}\int\psi\psi^* d\boldsymbol{r} = \mu\psi \tag{8.13}$$

where μ has the role of a Lagrange multiplier, consistent with the definition of the chemical potential in field theories.

In summary, the GPE is used to model a coherent state of a large number of identical bosons, suitable in the dilute ($a_s \ll d$) case at ultracold temperatures ($T \ll T_c \rightarrow N \gg 1$). Before addressing how such conditions may be achieved in experiment, we look at solutions to the time-independent GPE for several simple systems.

8.3.4 The Thomas-Fermi Approximation, Healing Length and Harmonic Traps

1. Uniform condensates.

Firstly setting $V \equiv 0$, (8.10) admits a constant solution $\psi = \psi_0$ where

$$\mu\psi_0 = U_0\psi_0^3 \Rightarrow \psi_0 = \sqrt{\frac{\mu}{U_0}} \tag{8.14}$$

More generally, the phase of ψ is an arbitrary constant and

$$|\psi| = \psi_0 = \begin{cases} \sqrt{\frac{\mu - V}{U_0}} & \text{if } \mu - V > 0 \\ 0 & \text{otherwise} \end{cases} \tag{8.15}$$

Despite its simplicity a constant solution is the first port of call in the presence of strong repulsive interactions or far from an external potential capable of producing changes in ψ since in these situations a constant is likely the minimiser of the energy (8.3) (the kinetic energy term is always positive, so ψ will not vary unless there is good reason for it to do so viz. an external potential that favours certain spatial configurations).

2. Condensate in a one dimensional well.

In a setup likely familiar from single-particle quantum mechanics,[23] we take

$$V = \begin{cases} 0 & 0 < x < L \\ \infty & \text{otherwise} \end{cases} \tag{8.16}$$

where the well length $L \gg \xi$ where ξ is a length scale over which the condensate can vary which we will define in a moment.

Far away from the edges then we expect a uniform solution $\psi = \psi_0$ and a density $n_0 = \mu/U_0$ as above. Within $\sim \xi$ of either edge the gradient term cannot be ignored however as ψ must decrease to 0 at each edge (where $V = \infty$). The standard way to proceed with the remaining time-independent GPE ($V = 0$, $\mu = n_0U_0$)

$$n_0U_0\psi = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U_0n_0\psi \tag{8.17}$$

is to introduce a dimensionless condensate function f which changes over distances of order unity via $\psi(x) = \psi_0 f(x/\xi) = \psi_0 f(u)$:

$$-\frac{\hbar^2}{2mU_0n_0\xi^2}f''(u) + \left(f^2 - 1\right)f = 0 \qquad (u = x/\xi) \tag{8.18}$$

Allowing us to identify the length scale in the problem

$$\xi^2 = \frac{\hbar^2}{2mU_0n_0} \tag{8.19}$$

This is known as the coherence or *healing length* since it characterises the distance over which ψ can change in response to a perturbation. We are left with the dimensionless form

$$-f'' + \left(f^2 - 1\right)f = 0 \tag{8.20}$$

[23] In the limit of negligible interactions $U_0 \sim 0$ the GPE reduces to the usual Schrödinger equation and so we get the familiar result $\psi \sim \sin\left(\frac{n\pi x}{L}\right)$ for a quantum particle in a box.

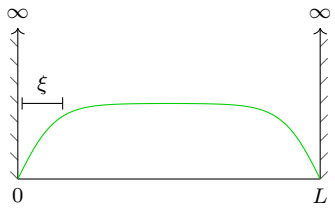


Figure 8.4: Condensate in a one dimensional well. Appreciable changes of ψ occur within $\sim \xi$ of either edge only.

You can check $f(u) \sim \tanh\left(u/\sqrt{2}\right)$ satisfies this equation and that the particular solution with boundary conditions $f(0) = f(L/\xi) = 0$ and $f \to 1$ near the centre of the well is

$$f(x/\xi) = \begin{cases} \tanh\left(\frac{x}{\xi\sqrt{2}}\right) & 0 < x < \frac{L}{2} \\ -\tanh\left(\frac{x-L}{\xi\sqrt{2}}\right) & \frac{L}{2} < x < L \end{cases} \tag{8.21}$$

In other words, $\psi = \psi_0 \tanh\left(s/\sqrt{2}\xi_0\right)$ where s is the distance from either edge, with $\psi \sim \psi_0$ for $s \gg \xi$ as expected (Figure 8.4).

3. Condensate in a harmonic potential.

We address the case of a harmonic potential which is applicable to the confinement of atomic condensates by magnetic traps (Section 8.4). In three dimensions,

$$V(\boldsymbol{r}) = \frac{m}{2}\omega_r^2 r^2 \qquad r^2 = x^2 + y^2 + z^2 \tag{8.22}$$

Firstly, in the absence of interactions we have the Schröndinger equation for the quantum harmonic oscillator (except with μ in place of E)

$$\mu\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + \frac{1}{2}m\omega_r^2 r^2\psi \tag{8.23}$$

which we know (or are told) has ground state

$$\psi(r) = \frac{\sqrt{N}}{\pi^{3/4}\ell_r^{3/2}} e^{-r^2/2\ell_r^2} \tag{8.24}$$

where we note the Gaussian is normalised to N, not 1, and

$$\ell_r = \sqrt{\frac{\hbar}{m\omega_r}} \tag{8.25}$$

is a characteristic length scale of the oscillator, obtained via similar analysis to that used to deduce ξ above (set KE $\sim$ PE). The chemical potential is easily read off from

$$-\frac{\hbar^2}{2m}\nabla^2\psi + \frac{1}{2}m\omega_r^2 r^2\psi = \frac{3\hbar^2}{2m\ell_r^2}\psi = \underbrace{\frac{3}{2}\hbar\omega_r}_{\mu}\psi \tag{8.26}$$

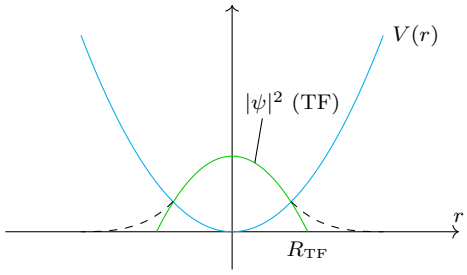


Figure 8.5: The TF approximation for a strongly interacting condensate in a harmonic trap. The actual condensate profile (determined numerically) is indicated with a dashed line.

corresponding to an energy

$$E = N\mu = \frac{3}{2}N\hbar\omega_r \tag{8.27}$$

of the condensate in the trap.

The other regime in which we can get analytic results is that of strong repulsive interactions, which for this problem means that the 'interaction parameter' Na_s/ℓ_r satisfies

$$\frac{Na_s}{\ell_r} \gg 1 \tag{8.28}$$

Neglecting the gradient term we get a solution of density

$$|\psi|^2 = \begin{cases} \frac{2\mu^2 - m\omega_r^2 r^2}{2U_0} & \text{if } 2\mu < m\omega_r^2 r^2 \\ 0 & \text{otherwise} \end{cases} \tag{8.29}$$

which, since the potential is changing, is not just a constant (Figure 8.5). Neglecting the kinetic energy in this way is known as the *Thomas-Fermi (TF) approximation,* and

$$R_{\text{TF}} = \sqrt{\frac{2\mu}{m\omega_r^2}} \tag{8.30}$$

the TF radius.

8.4 Preparing Atomic Condensates

There are two main steps in preparing an atomic condensate. Firstly, laser or *doppler* cooling is used to cool the gas down to temperatures in the μK range (Figure 8.6a). During this time the gas is confined by a magnetic trap.

Laser cooling alone cannot produce the nK temperatures required for condensation in an atomic gas, however.[24] Instead, the cooling lasers are turned off and the trapping steadily potential *reduced:* as the depth of the potential decreases the most energetic particles are able to escape the trap,

[24] This is due to the 'recoil-limit'. M. Fox provides an excellent introduction to cooling, trapping and BEC in his textbook on Quantum Optics [7].

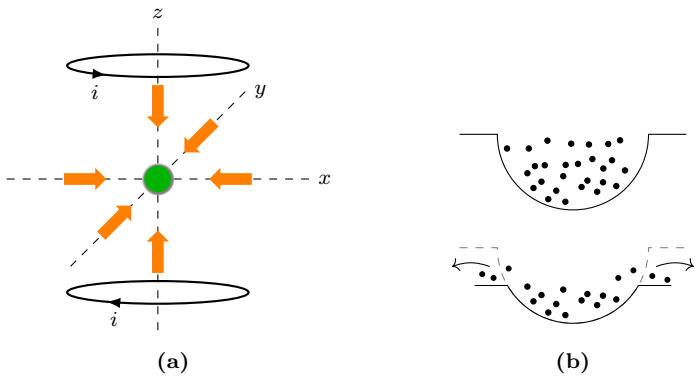


Figure 8.6: (a) Doppler cooling. In a magneto-optic trap three pairs of counter propagating laser beams annul atom velocities in the x,y and z directions whilst a magnetic field generated by current carrying coils confines the atoms to a small region at the intersection of the beams. (b) Evaporative Cooling. The laser beams are switched off and the height of the magnetic trap reduced, allowing the most energetic atoms to escape and reducing the temperature further. After Fox [7].

leaving behind the slow moving, low energy ones which thus form a gas of reduced temperature. This is known as *evaporative cooling* (Figure 8.6b).

A typical atomic BEC experiment may be performed at temperatures $\sim$ nK with condensates of anywhere from 10^3 to 10^9 atoms confined to a region of size $\sim 10\mu$m. Isotopes of alkali metals, such as ^{87}Rb or ^{23}Na, are commonly used.

9 Optical Lattices

The standing wave pattern generated by the inference of counter propagating laser beams may be used to trap neutral atoms in a structure know as an optical lattice (Figure 9.1). These structures have potential uses in quantum information processing and, our favourite subject, quantum simulation.

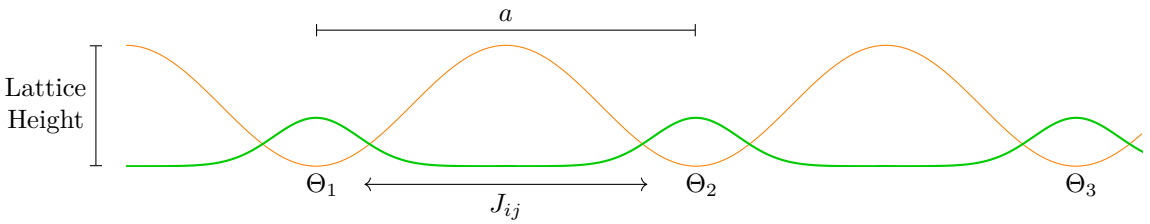


Figure 9.1: Optical Lattice. In a tight-binding approximation, a condensate is localised at each site of the lattice with interactions between neighbours.

9.1 Tight Binding

In an optical lattice the spatially dependent intensity described by the inference pattern of laser beams induces a spatially dependent potential energy for ultracold neutral atoms via the AC-Stack effect. This potential is ideal (lossless) and in simple cases has a sinusoidal form, for example

$$V = \sum_{\ell=x,y,z} V_{0,\ell} \sin^2\left(\frac{\pi \ell}{a}\right) \tag{9.1}$$

with atoms confined to the minima of this function. The two key parameters the lattice are its periodicity, controlled by the lattice constant a (half the laser wavelength), and its height, dictated by the magnitude of V.

When the depth of each well is not too large, atoms readily move between minima via quantum tunnelling and interact, resulting in coherence across the entire lattice and a superfluid-like state. If the depth of each well is increased, the atoms become increasingly localised at the minima and the system transitions into a Mott insulator state.

In the regime of large lattice height we can use a *tight-binding approximation,* writing the total N atom condensate $\psi(\boldsymbol{r}, t)$ as a sum over single atom condensates at each site of the lattice $\boldsymbol{R}_i$, with time-dependent coefficients:

$$\psi(\boldsymbol{r}, t) = \sqrt{N} \sum_{\boldsymbol{R}_i} \varphi_i(t) \Theta(\boldsymbol{r} - \boldsymbol{R}_i) \tag{9.2}$$

where

$$\int |\Theta|^2 d\boldsymbol{r} = 1 \tag{9.3}$$

i.e. Θ is a solution to the GPE with $N = 1$ (which we assume is known), and

$$\varphi_i(t) = \sqrt{n_i(t)} e^{i\theta_i(t)} \qquad n_i = \frac{N_i}{N} \tag{9.4}$$

with N_i the number of particles at the i^{th} site.

The tight-binding approximation at *lowest* order consists of this ansatz together with the prescription that interactions between sites due to the kinetic ($\sim \nabla^2$) and potential ($\sim V$) terms of the GPE arise for *nearest* neighbours only, whilst the overlap between wavefunctions at different sites is (otherwise) negligible.[25] The details are worked out in Example Sheet 2; from the GPE one obtains

$$i\hbar \frac{\partial \varphi_i}{\partial t} = -K\left(\varphi_{i-1} + \varphi_{i+1}\right) + \left(\epsilon_i + U|\varphi_i|^2\right) \varphi_i \tag{9.5}$$

where, denoting $\Theta_i(\boldsymbol{r}) \equiv \Theta(\boldsymbol{r} - \boldsymbol{R}_i)$,

$$K = -\int \left[\frac{\hbar^2}{2m} \nabla\Theta_i \cdot \nabla\Theta_{i+1} + \Theta_i V \Theta_{i+1}\right] d\boldsymbol{r} \tag{9.6}$$

[25] Some intuition I find helpful: in Figure 9.1 it is clear that Θ_1 and Θ_2 have no overlap (when one is non-zero the other is zero), but the integral of $\Theta_1\Theta_2$ *weighted by the potential* V may be non-zero, since V overlaps with both Θ_1 and Θ_2 (gradient overlap is similarly plausible, although the diagram is not drawn accurately enough to reflect this).

is the nearest-neighbour tunnelling rate,

$$\epsilon_i = \int \left[\frac{\hbar^2}{2m} \left(\nabla \Theta_i \right)^2 + \Theta_i V \Theta_i \right] d\boldsymbol{r} \tag{9.7}$$

are on-site energies and

$$U = U_0 \int \Theta^4 d\boldsymbol{r} \tag{9.8}$$

This is the lowest order discrete non-linear Schrödinger equation (DNSLE).

9.2 Realising the XY Hamiltonian

Including long-range interactions, the DNSLE (9.5) generalises to

$$i\hbar \frac{\partial \varphi_i}{\partial t} = -\sum_j J_{ij} \varphi_j + \left(\epsilon_i + U |\varphi_i|^2 \right) \varphi_i \tag{9.9}$$

which looks very similar to our equations for a system of SL oscillators (6.23), albeit with imaginary time dynamics. It should come as no surprise then that, if an *equal* number of atoms are loaded at each lattice site, the ground state of this equation realises the minimum of the XY Hamiltonian:

$$i\hbar \frac{\partial \varphi_i}{\partial t} = \frac{\delta E}{\delta \varphi_i^*} = \frac{\delta}{\delta \varphi_i^*} \int H d\boldsymbol{r} = \frac{\partial H}{\partial \varphi_i^*} \tag{9.10}$$

where

$$H = -\frac{1}{2} \sum_{ij} J_{ij} \varphi_i \varphi_j^* + \sum_i \left[\epsilon_i |\varphi_i|^2 + \frac{U}{2} |\varphi_i|^4 \right] \tag{9.11}$$

$$= \underbrace{-\frac{1}{2} \sum_{ij} J_{ij} \rho \cos\left(\theta_i - \theta_j \right)}_{H_{\mathrm{XY}}} + \underbrace{\sum_i \left(\epsilon_i + \frac{U}{2} \rho \right) \rho}_{\text{some constant}} \tag{9.12}$$

with ρ the constant density at each site.

A quantum annealing protocol not too dissimilar from that discussed in Section 7.1 (perhaps involving diabatic change to an extent) may be used to tune the precise couplings J_{ij} from a simpler Hamiltonian e.g. with antiferromagnetic interactions only.

10 Polariton Condensates

Polaritons are quasi-particles that result from the strong coupling between *photons* and electronic excitations (electron-hole bound states) in a semiconductor known as *excitons.* Informally, a photon can be absorbed to create an exciton, which subsequently relaxes to emit a photon, which can go on to create another exciton, and so on. In the strong coupling regime this reversible process occurs so rapidly that the normal modes of the systems are not excitons or photons, but superpositions of these i.e. polaritons.

10.1 Condensation

At low enough densities, polaritons behave as bosons and may exhibit Bose-Einstein condensation[26] and associated phenomena such as long range phase coherence, superfluidity and pattern formation. That the effective mass of polaritons is typically far smaller than atomic mass (they are 'part matter part light') means that condensation can occur at relatively high temperatures (K rather than μK) – in some cases, even at room temperature. The condensates formed are dilute, weakly interacting Bose gases with a contact interaction $U\delta(\boldsymbol{r})$ of the kind we have been considering (describing polariton-polariton interactions).

In order to actually observe condensation, it is necessary to confine photons in a microcavity with excitons (bound in quantum wells) at the anti-nodes of the light modes. This facilitates strong couplings and allows the polaritons to survive long enough to cool and condense. Key in the description of microcavity polaritons is their dispersion, which has two branches as shown in Figure 10.1. Note that, due to their confinement, photons in the cavity have a non-zero effective mass. For a high quality cavity, order of magnitude estimates are

$$m_{\text{pho.}} \sim 10^{-4} m_e, \quad m_{\text{pol.}} \sim 2m_{\text{pho.}}, \quad U_0 \sim 10^{-3}\text{meV}(\mu\text{m})^2 \tag{10.1}$$

with polariton lifetimes of $5 - 10$ps. Polaritons also have two possible polarisation states (see Example Sheet 3), but often this degree of freedom can be ignored in a real experiment.

10.2 Injection

Due to imperfect confinement of the photon component (the cavity mirrors are not perfect reflectors), polaritons have a finite lifetime and have to be continuously repopulated. In this way polariton condensates are non-equilibrium steady states emitting coherent light, lying somewhere between an equilibrium BEC and a laser. Although losses from the cavity (photon emission) are not ideal, they are also the means by which measurements of the condensate can be obtained, namely via the emission angle of the outgoing photons which corresponds to the conserved in-plane momentum.

There are several ways in which the system can be pumped i.e. polaritons injected into the condensate. The key distinction is whether the pumping is *coherent or incoherent.* In the former case, the phase of the pump laser is imparted to the condensate, whereas it is not in the case of incoherent pumping, typically due to a multitude of small scattering events between injection and the final low momentum state of the polariton (Figure 10.2). We will be most interested in incoherent pumping at high energies.

[26] A technicality here is that true BEC does not occur in two dimensional systems (as we consider), at least not if they are infinite. However for finite systems in $d = 2$ with a suitable trap the principal features of condensation are recovered. The review by J. Keeling and Natalia [8] provides an accessible introduction to the phenomenology.

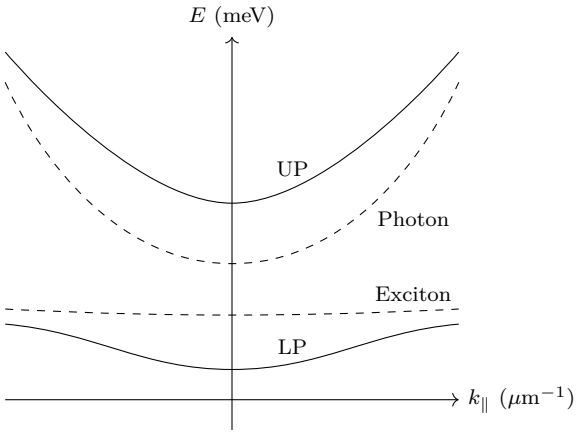


Figure 10.1: Sketch of the upper (UP) and lower (LP) branches of the dispersion relation for polaritons in a microcavity. This is obtained by solving the Schröndinger equations for coupled and photon wavefunctions:

$$E_k^{\text{LP, UP}} = \frac{1}{2}\left[\left(\hbar\omega_0 + \epsilon + \frac{\hbar^2 k^2}{2m}\right) \mp \left(\left(\hbar\omega_0 - \epsilon + \frac{\hbar^2 k^2}{2m}\right)^2 + g^2\right)^{1/2}\right]$$

where k is the in-plane polariton wavevector, $\hbar\omega_0$ the energy of the bottom of the photon band, ϵ the bare exciton energy and g the photon-exciton conversion rate (at resonance, $\epsilon = \hbar\omega_0$) [8].
The photon and exciton dispersions are indicated with a dashed line. The latter is essentially flat due to the large mass of the electron compared to $m_{\text{pho.}}$, $m_{\text{pol.}}$ (recall large effective mass $\leftrightarrow$ small curvature).

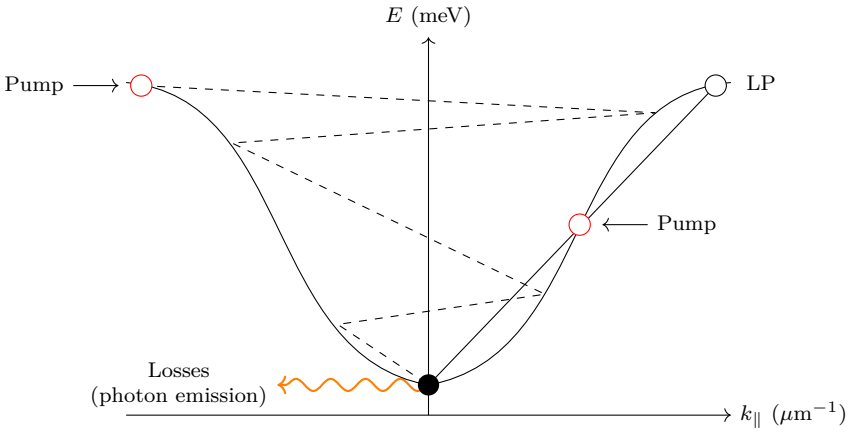


Figure 10.2: Injection into the LP branch may be coherent or incoherent. For example, polaritons may be injected coherently at a 'magic-angle' or momentum (right) from which they can scatter into one zero and one high momentum state in a single event. Alternatively, polaritons may be injected incoherently at higher energies (left), subsequently relaxing towards a low momentum state via many scattering events (phonon emission).

10.3 Modelling Polariton Condensates

10.3.1 The Driven Dissipative GPE

We have discussed how polariton condensates are non-equilibrium steady states. Building on our model of atomic condensates, we add imaginary terms describing pumping and dissipation to the GPE:

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V + U_0|\psi|^2\right)\psi + i\left[P_{\text{inc.}}(\boldsymbol{r}) - \gamma_c - \sigma|\psi|^2\right]\psi \tag{10.2}$$

As before the first three terms of the right-side are associated with kinetic energy, an external potential and a contact interaction. The polariton experiments we are interested in are performed using clean (ordered) materials without trapping potentials in which case V is negligible.

Amongst the new terms, $P_{\text{inc.}}(\boldsymbol{r})$ is an incoherent pump field which populates the condensate without fixing its phase; $P_{\text{inc.}}\psi$ is replaced by $iP_{\text{coh.}}e^{-\omega_p t}$ in the case of coherent pumping at frequency ω_p. The coefficients γ_c and σ annotate linear and non-linear losses, respectively. The former is associated with photon decay and the latter with inelastic scattering with the pump and the depletion of the exciton reservoir (Section 10.3.4). Fundamentally, non-linear losses must be present to an extent otherwise the system would be unstable to runaway condensate growth as soon as $P_{\text{inc.}}$ exceeded κ.

10.3.2 Steady State Outflow

Unlike equilibrium condensates, polariton condensates have steady *outflows* of particles maintained by the pumping. This is important as the primary mechanism by which condensates can interact over relatively long distances as is essential for our purposes. Assuming a basic Gaussian pumping profile $P_{\text{inc.}} = p_0 e^{-\alpha r^2}$ in two dimensions, we can determine a characteristic outflow velocity for a single condensate. Note that to produce multiple condensates in a graph, a device known as a spatial light modulator (SLM) is used to pump at precise locations in the 2D plane.

We look for stationary solutions $\psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})e^{-i\mu t}$ to the driven dissipative GPE:

$$\mu\hbar\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 + V + U_0|\psi|^2\right)\psi + i\left[P_{\text{inc.}}(\boldsymbol{r}) - \gamma_c - \sigma|\psi|^2\right]\psi \tag{10.3}$$

As was done for the ordinary GPE (Section 8.3), we use the Mandlung transformation $\psi = \sqrt{n}e^{iS}$ obtaining, from the imaginary part,[27]

$$\nabla\cdot(n\boldsymbol{v}) = -\left(P_{\text{inc.}} - \gamma_c - \sigma n\right)n \tag{10.4}$$

Away from the pumping spot we can neglect $P_{\text{inc.}}$ and non-linear losses, leaving

$$\nabla\cdot(n\boldsymbol{v}) = -\gamma_c n \tag{10.5}$$

On account of the pumping profile, the problem is circularly symmetric so that $\boldsymbol{v} = u(r)\hat{\boldsymbol{r}}$ where $\hat{\boldsymbol{r}}$ is the radial unit vector and

$$\frac{1}{r}\frac{d(rnu)}{dr} = -\gamma_c n \tag{10.6}$$

[27] Compared to the analysis in Appendix B we have $\mu\sqrt{n}e^{iS}$ on the left-hand side instead of $i\hbar\dot{\psi}$ (which upon cancelling e^{iS} is real) and the three new terms $\propto in$ on the right (having multiplied by $\sqrt{n}$). Hence the absence of $\dot{n}$ and the addition of $(P_{\text{inc.}} - \gamma_c - \sigma n)\,n$ in (10.5) compared to (8.7).

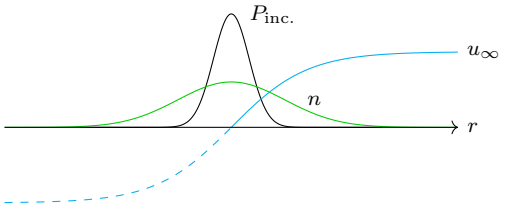


Figure 10.3: Pumping, density and velocity profiles for a single condensate with Gaussian pumping.

On the other hand, resolving the real part gives (cf. (B.9) with $V = 0$),

$$\hbar\mu = -\frac{\hbar^2}{2m}\left(\frac{\nabla^2\sqrt{n}}{\sqrt{n}} - (\nabla S)^2\right) + U_0 n \tag{10.7}$$

$$= \frac{\hbar^2}{2m}u(r)^2 + U_0 n \tag{10.8}$$

where we neglected the kinetic term in a TF approximation. As $r \to \infty$ we expect the density n to decay, meaning we must have $u \to u_\infty$ such that

$$u_\infty^2 = \frac{2m\mu}{\hbar} \tag{10.9}$$

and so the characteristic outflow velocity is fixed by the chemical potential. The associated wavevector $k_c = mu_\infty/\hbar$ is the radius of the circle described by the outflow in Fourier space. In view of these asymptotics we can approximate $u \sim u_\infty \tanh(\ell r)$ where ℓ is the distance over which the velocity profile changes.

Returning to (10.6), for large r we have

$$\frac{\partial(rn)}{\partial r} \approx -\frac{\gamma_c n r}{u_\infty} \tag{10.10}$$

$$\Rightarrow \frac{1}{n}\frac{dn}{dr} \approx -\frac{1}{r} - \frac{\gamma_c}{u_\infty} \tag{10.11}$$

and so the decay of the condensate density follows

$$n(r) \sim \frac{1}{r}\exp\left(-\frac{\gamma_c}{u_\infty}r\right) \tag{10.12}$$

A complete analytical expression for $n(r)$ is derived in [9] using asymptotic matching. The predicted forms of $n(r)$ and $u(r)$ are sketched alongside the pumping profile in Figure 10.3.

10.3.3 Condensate Coupling: Polariton Multiplets

In a graph of two or more polariton vertices (i.e. pumping centres) the principle governing the coupling established between the condensates is that

> *Polariton condensation occurs at the phase-configuration that carries the highest occupation* $M = \int |\psi|^2 d\boldsymbol{r}$.

This is due to the phenomena of *stimulated relaxation* in bosonic systems: the probability for a particle to relax in a particular state grows with the population of that state.[28]

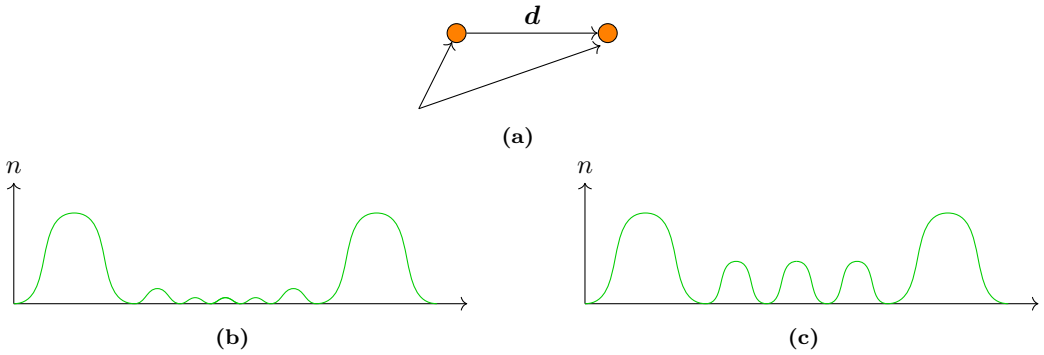


Figure 10.4: A polariton dyad. (a) Two pumping spots separated by $\boldsymbol{d}$. The origin of our coordinate system is placed between the two centres. We expect either (b) destructive or (c) constructive interference.

We apply this principle explicitly to the case of two coupled condensates (Figure 10.4). In particular, we are interested in what determines whether there is *destructive* (π phase difference, antiferromagnetic coupling) or *constructive* (0 phase difference, ferromagnetic coupling) between the condensates.

We approximate the wavefunction (stationary solution) for the entire system as

$$\psi = \psi_0\left(\boldsymbol{r} + \frac{d}{2}\right) + e^{i\Delta\theta}\psi_0\left(\boldsymbol{r} - \frac{d}{2}\right) \tag{10.13}$$

where ψ_0 is the wavefunction for a single condensate and $\Delta\theta$ the phase difference. Using Plancherel's theorem, we can express the total occupation in terms of the Fourier transform $\hat{\psi}_0(\boldsymbol{k})$ of a single condensate:

$$\begin{aligned} M &= \int \frac{d^2k}{(2\pi)^2}|\psi(\boldsymbol{k})|^2 && (10.14)\\ &= \int \frac{d^2k}{(2\pi)^2}|\psi_0(\boldsymbol{k})|^2\left|e^{i\boldsymbol{k}\cdot\boldsymbol{d}/2} + e^{i\Delta\theta}e^{-i\boldsymbol{k}\cdot\boldsymbol{d}/2}\right|^2 && (10.15)\\ &= \int \frac{d^2k}{(2\pi)^2}|\psi_0(\boldsymbol{k})|^2\left[2 + \cos(\boldsymbol{k}\cdot\boldsymbol{d} - \Delta\theta)\right] && (10.16)\\ &= 2\int \frac{d^2k}{(2\pi)^2}|\psi_0(\boldsymbol{k})|^2 + \cos\Delta\theta\frac{1}{2\pi^2}\int |\hat{\psi}_0(\boldsymbol{k})|^2\cos(\boldsymbol{k}\cdot\boldsymbol{d})d^2\boldsymbol{k} && (10.17) \end{aligned}$$

where the integral over $\sin(\boldsymbol{k}\cdot\boldsymbol{d})$ vanished since this is odd under $\boldsymbol{k}\to-\boldsymbol{k}$ whilst $\hat{\psi}_0(\boldsymbol{k})$ is even. In the first term we recognise the occupation of a single condensate in isolation:

$$\int \frac{d^2k}{(2\pi)^2}|\hat{\psi}_0(\boldsymbol{k})|^2 = \int |\psi_0(\boldsymbol{r})|^2 d\boldsymbol{r} = M_0 \tag{10.18}$$

[28]The effect of Bose stimulation (transition rates to a state of N bosons are proportional to $N+1$ [10]) is most familiar in the operation of an optical laser (stimulated emission).

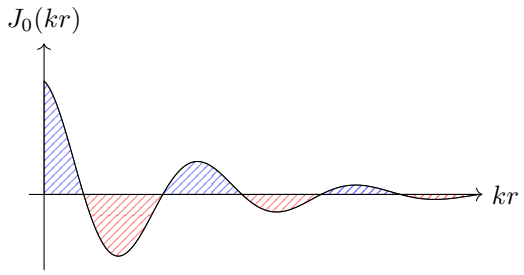


Figure 10.5: The sign of $J_0(kr)$ at $kr = k_c d$ determines whether there is a ferromagnetic (blue) or antiferromagnetic (red) coupling.

and in the second the Fourier transform of $|\hat{\psi}_0(\boldsymbol{k})|^2$:

$$\frac{1}{2\pi^2}\int |\hat{\psi}_0(\boldsymbol{k})|^2 \cos(\boldsymbol{k}\cdot\boldsymbol{d})d^2k = \frac{1}{2\pi^2}\mathcal{F}\left(|\hat{\psi}_0(\boldsymbol{k})|^2\right) \tag{10.19}$$

where again there is no component $\propto \sin(\boldsymbol{k}\cdot\boldsymbol{d})$ since $|\hat{\psi}_0(\boldsymbol{k})|^2$ is even. Moreover, for the pumping profile we have considered this function is circularly symmetric, meaning its Fourier transform can be written as to 2π times its Hankel transform of order zero:

$$\mathcal{F}\left(|\hat{\psi}_0(\boldsymbol{k})|^2\right) = 2\pi\int_0^\infty |\hat{\psi}_0(k)|^2 J_0(kd)kdk \tag{10.20}$$

where J_0 is the zeroth-order Bessel function of the first kind. As ψ_0 is sharply peaked in Fourier space at the characteristic wavevector k_c, we can approximate

$$|\hat{\psi}_0(k)|^2 \approx |\hat{\psi}_0(k_c)|^2\delta(k-k_c) \to \mathcal{F}\left(|\hat{\psi}_0(k)|^2\right) \approx 2\pi|\hat{\psi}_0(k_c)|^2 J_0(k_c d)k_c \tag{10.21}$$

All together, we have

$$M = 2M_0 + J\cos\Delta\theta \tag{10.22}$$

where M_0 is a constant and

$$J = \frac{1}{\pi}\int_0^\infty |\hat{\psi}_0(k)|^2 J_0(kd)kdk \approx \frac{k_c}{\pi}|\hat{\psi}_0(k_c)|^2 J_0(k_c d) \tag{10.23}$$

Hence, the preferred phase is determined by where $k_c d$ sits in relation to the zeros of the Bessel function (Figure 10.5): if $J_0(k_c d) > 0$ then $\Delta\theta = 0$ maximises M and we have a ferromagnetic coupling $J > 0$, whereas if $J_0(k_c d) < 0$ it is $\Delta\theta = \pi$ that is favoured and $J < 0$.

The upshot is that the coupling between two condensates can be controlled via pumping and in particular the distance between pumped spots. For Q condensates, the generalisation of (10.22) is

$$M = \int \left|\sum_i \psi_0(|\boldsymbol{r}-\boldsymbol{r}_i|)\right|^2 d\boldsymbol{r} = QM_0 + \sum_{ij} J_{ij}\cos(\theta_i - \theta_j) \tag{10.24}$$

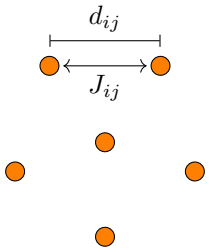


Figure 10.6: A primitive polariton graph simulator where the coupling between nodes is controlled by pumping spot characteristics and separation. We will see below that geometry alone is not sufficient to produce arbitrary couplings and more sophisticated pumping techniques may be required in order to avoid polariton-exciton interactions.

with $J_{ij} \propto J_0(k_c d_{ij})$ ($d_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$ the distance between the i^{th} and j^{th} condensates). Then, that the system maximises M implies it minimises

$$-\sum_{ij} J_{ij} \cos(\theta_i - \theta_j) = H_{\text{XY}} \tag{10.25}$$

We have now seen all the basic ingredients of a *polariton graph simulator* (Figure 10.6):

- Map an optimisation problem to coupling strengths of an XY Hamiltonian
- Arrange polariton vertices in a graph with distances d_{ij} in accordance with these couplings
- Let the polaritons condense
- Read out the phases θ_i which describe $\min_{\theta_i} H_{\text{XY}}$ and so a solution to the original problem

10.3.4 Coupled Polariton-Excition Reservoir System

Polariton condensation is in fact a two step process, involving

1. The population of a reservoir of 'hot' excitons, and
2. Scattering from the reservoir to the condensate

A more accurate model, which reflect these steps, is of a condensate evolution equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + U_0|\psi|^2\psi + \hbar g_R n_R \psi + \frac{i\hbar}{2}\left(R_R n_R - \gamma_c\right)\psi \tag{10.26}$$

coupled to the rate equation for the density of the exciton reservoir $n_R(\boldsymbol{r}, t)$:

$$\frac{\partial n_R}{\partial t} = -\left(\gamma_R + R_R|\psi|^2\right) n_R + P(\boldsymbol{r}) \tag{10.27}$$

where, in addition to those terms we have seen before, γ_R annotates linear dissipation of n_R, R_R captures scattering into the condensate and g_R is the strength of repulsive polariton-exciton interactions.

We now look at the how this system can be analysed with a tight-binding approximation in several regimes, and comment on the implications for our polariton graph simulator.

10.3.5 Tight-binding in a Polariton Network

We choose units such that $\hbar = 1$ and de-dimensionalise the above system such that

$$i\dot{\psi} = -\nabla^2\psi + |\psi|^2\psi + gn_R\psi + \frac{i}{2}\left[n_R - 1\right]\psi \tag{10.28}$$

$$\dot{n}_R = -\left(b_0 + b_1|\psi|^2\right)n_R + P \tag{10.29}$$

This involves introducing a length scale $\boldsymbol{r} \to \ell_0\boldsymbol{r}$ and rescaling ψ, t and other coefficients (we do not concern ourselves with the details but note that $g_R \to g$ and the new coefficients b_0, b_1 will be set by the properties of the physical system). Next we make the ansatz

$$\psi(\boldsymbol{r}, t) = \sum_i a_i(t)\varphi_i(\boldsymbol{r}_i) \tag{10.30}$$

$$P(\boldsymbol{r}, t) = \sum_i f_i(t)p_i(\boldsymbol{r}) \tag{10.31}$$

$$n_R(\boldsymbol{t}, t) = \sum_i \kappa_i(t)n_i(\boldsymbol{r}_i) \tag{10.32}$$

where $\varphi_i(\boldsymbol{r}) \equiv \varphi(|\boldsymbol{r} - \boldsymbol{r}_i|)$ is the wavefunction of an isolated condensate at $\boldsymbol{r} = \boldsymbol{r}_i$ such that $\int |\varphi_i|^2 d\boldsymbol{r} = 1$ and p_i, n_i are pumping and densities at $\boldsymbol{r}_i$ compatible with this normalisation. The idea with the above substitution is that we want to reduce the spatially dependent pair of equations for the entire system to a set of simpler, spatially independent and identical coupled pairs of equations, one for each node (oscillator) of the network.

Multiplying (10.28) by φ_i^*, (10.29) by $|\varphi_i|^2$ and integrating over the entire system subject to

$$i \neq j \Rightarrow \int \varphi_i\varphi_j^* d\boldsymbol{r} \ll \int n_i\varphi_j\varphi_i^* d\boldsymbol{r} = \ell_{ij} \quad \text{and} \quad \int \nabla^2\varphi_j\varphi_i^* d\boldsymbol{r} \ll \int \nabla^2\varphi_i\varphi_i^* d\boldsymbol{r} = d \tag{10.33}$$

it can be shown that, at each i

$$\dot{\psi} = -i|\psi_i|^2\psi + (1 - ig)\left(R_i\psi_i + \sum_{ij} J_{ij}\psi_j\right) - \psi_i \tag{10.34}$$

and

$$\dot{R}_i = b_0\left(\gamma_i - R_i - \xi R_i|\psi_i|^2\right) \tag{10.35}$$

where we introduced

$$\psi_i = a_i(t)e^{-idt} \tag{10.36}$$

to remove a term involving d, and

$$\gamma_i = \frac{f_i}{b_0}\int p|\varphi|^2 d\boldsymbol{r}, \quad \xi = \frac{b_1}{\ell b_0}\int n|\varphi|^4 d\boldsymbol{r}, \quad R_i = \ell\kappa_i, \quad \ell = \ell_{ii} \tag{10.37}$$

The couplings J_{ij} are determined to be

$$J_{ij} = \begin{cases} 0 & \text{when } i = j \\ (R_i\ell_{ij} + R_j\ell_{ji}^*)/\ell & \text{otherwise} \end{cases} \tag{10.38}$$

It will be convenient to denote

$$\mathcal{J}_{ij}e^{i\nu_{ij}} = \frac{(R_i\ell_{ij} + R_j\ell^*_{ji})}{\ell} \tag{10.39}$$

Aside from the leading non-linear term, (10.34)-(10.35) are the Lang-Kobayashi equations seen in laser theory.

A limit of relevance is that of fast reservoir relaxation, meaning $\gamma_R \gg \gamma_c$ which corresponds to $b_0 \gg 1$. In this case we can use the stationary state solution for the reservoirs,

$$R_i = \frac{\gamma_i}{1 + \xi|\psi_i|^2} \approx \gamma_i - \xi\gamma_i|\psi_i|^2 \tag{10.40}$$

for small densities (close to threshold). Substituting into the dynamical equation for the oscillator at i,

$$\dot{\psi}_i = i\left(g\xi\gamma_i - 1\right)|\psi_i|^2\psi_i - \xi\gamma_i|\psi_i|^2\psi_i + (1 - ig)\left[\gamma_i\psi_i + \sum_{ij} J_{ij}\psi_i\right] - \psi_i \tag{10.41}$$

We have essentially a system of SL oscillators $\psi_i = \sqrt{\rho_i}e^{i\theta_i}$ since, upon separating real and imaginary parts,

$$\frac{1}{2}\dot{\rho}_i = \left(\gamma_i - 1 - \xi\gamma_i\rho_i\right)\rho_i + \sum_{ij}\tilde{J}_{ij}\sqrt{\rho_i\rho_j}\cos(\theta_{ij} - \nu_{ij} + \alpha) \tag{10.42}$$

and

$$\dot{\theta}_i = \left(g\xi\gamma_i - 1\right)\rho_i - g\gamma_i - \sum_{ij}\tilde{J}_{ij}\sqrt{\frac{\rho_j}{\rho_i}}\sin(\theta_{ij} - \nu_{ij} + \alpha) \tag{10.43}$$

with the notation

$$\theta_{ij} = \theta_i - \theta_j, \quad \tan\alpha = g, \quad \tilde{J}_{ij} = \frac{\mathcal{J}_{ij}}{\cos\alpha} \tag{10.44}$$

Making density adjustments $\dot{\gamma}_i = \epsilon(\rho_{\text{th}} - \rho_i)$ as discussed in Section 6.3 we have, in the phase equations

$$\dot{\theta}_i = \left(g\xi\gamma_i - 1\right)\rho_{\text{th}} - g\gamma_i - \sum_{ij}\tilde{J}_{ij}\sin(\theta_{ij} - \nu_{ij} + \alpha) \tag{10.45}$$

where we made the assumption $\nu_{ij} \approx 0$ (or $\ell_{ij} \approx \ell^*_{ji}$). These describe a system of Sakaguchi-Kuramoto oscillators. The familiar Kuramoto network is recovered in the limit $\alpha = 0 \leftrightarrow g \approx 0$ i.e. negligible repulsions between the condensates and their reservoirs.

As we saw in Section 6.2, it is only the Kuramoto model and so this final regime[29] $g \approx 0$, that is the true minimiser of the XY Hamiltonian i.e. gives the gradient descent of H_{XY}. Also as discussed there, a term $h(t)\psi^*_i$ can be added (second resonant pumping) to discretise the phases $\theta_i \to 0, \pi$.

[29] $g \approx 0$ can be realised in experiment using a clever pumping technique: pump at the four corners of a square so that polaritons condense at the centre of that square, away from the reservoirs.

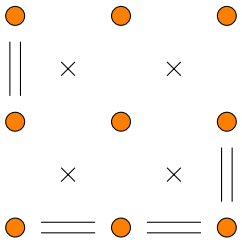


Figure 10.7: Dissipative gates (=) and barriers (×) are setup on a square lattice to facilitate or impede the dissipation between condensates and so control the couplings.

It is important to note that the couplings depend on the pumping $P(\boldsymbol{r}, t)$. This poses a challenge because each γ_i (thus the pumping) was adjusted to equalise the densities, meaning we cannot fix (or know) the couplings from the outset. Thus, similar to the control of the γ_i, it is necessary to subsequently adjust the J_{ij} in order to obtain the couplings J^0_{ij} describing the actual problem.

A second practical consideration here is that, for systems of more than three condensates plane geometry alone is not enough to realise arbitrary couplings. Instead, the lattice is normally taken to be regular and the dissipative properties of paths between condensates altered to setup the couplings (Figure 10.7). Alternatively, a SLM may be used to direct light between condensates to establish remote couplings. This has the advantage of producing real couplings ($\nu_{ij} = 0$) independent of pumping and is a promising research direction.

11 A Coherent Ising Machine

In this final section we look at promising technology for a coherent Ising machine, a device to minimise the Ising Hamiltonian H_I.

11.1 Basic Components

The fundamental elements of the machine are shown in Figure 11.1. An optical parametric oscillator (OPO) is pumped above threshold in a pulsed manner, creating a sequence of regularly spaced pulses. This is achieved by choosing a loop length such that the pulses are amplified as they

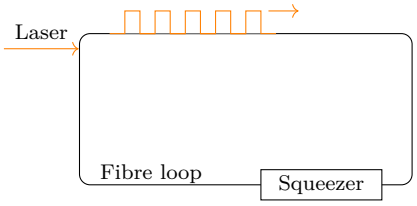


Figure 11.1: The primary element of the coherent Ising machine is an optical parametric oscillator (OPO) consisting of a fibre loop and a device known as a degenerate parametric amplifier (squeezer). The interacting elements are pulses sent along the fibre.

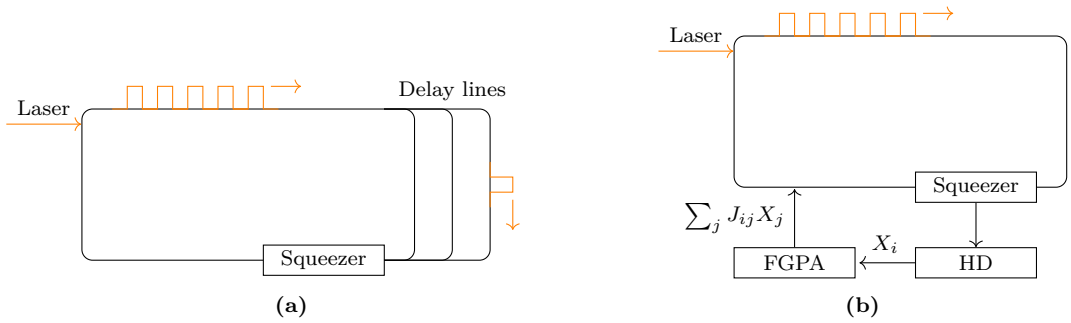


Figure 11.2: (a) Time-delays may be introduced by sending select pulses down additional loops of fibre. (b) A field-programmable gate array (FPGA) takes quadrature measurements from a Homodyne detector (HD) and imposes displacements $\sum_{ij} J_{ij}X_j$ on the signal.

circulate, with macroscopic occupation occurring for the maximum gain phase configuration (this is analogous to the bosonic stimulation with saw in Section 10.3.3). The parametric amplifier is phase sensitive, projecting or 'squeezing' the pulse phases onto $\{0, \pi\}$.

11.2 Interactions and Measurement

Interactions may be introduced by adding *time-delays* to some of the pulses. These are easy to implement with additional loops of fibre (Figure 11.2a), but the number of possible couplings is limited by the number of delay lines. Instead we consider the use of a field-programmable gate array (FPGA) to implement the couplings via displacements (Figure 11.2b) which allows for all of the pulses to be coupled, although has the disadvantage that the pulses have to be formed and measured before the interactions can be programmed.

In order to make these measurements, a portion of the optical energy entering the amplifier is deflected to a *homodyne* detector which returns the position quadrature[30] of each pulse X_i.

11.3 Modelling – A Hopfield Network

Each pulse has an amplitude $a_i = \frac{1}{\sqrt{2}}(X_i + iP_i)$ where X_i and P_i are the phase space quadratures. The change in amplitude per round trip is described by

$$\Delta a_i = \underbrace{\omega a_i^*}_{\substack{\text{parametric}\\ \text{gain}}} - \underbrace{\left(\gamma + s|a_i|^2\right)}_{\text{losses}} a_i + \underbrace{\xi \sum_j J_{ij}X_j}_{\text{displacements}} + \underbrace{\frac{f_i}{\sqrt{2}}}_{\text{noise}} \tag{11.1}$$

The corresponding changes in the quadratures are

$$\Delta X_i = \omega X_i - \gamma X_i - s\left(X_i^2 + P_i^2\right) X_i + \xi \sum_j J_{ij}X_j + \mathrm{Re}\, f_i \tag{11.2}$$

$$\Delta P_i = -\omega P_i - \gamma P_i - s\left(X_i^2 + P_i^2\right) P_i + \mathrm{Im}\, f_i \tag{11.3}$$

[30] For those unfamiliar with quantum optics, this is the cosine ($\propto \cos\varphi_i$) component of the electric field. The textbook by M. Fox [7] provides an excellent introduction to optical phase space and homodyne detection.

Provided the non-linearity (second harmonic generation) does not contribute much, if a sigmoid $g(u)$ is applied to the displacements then

$$\Delta X_i = \xi \sum_j J_{ij} g\left(X_j\right) + (\omega - \gamma) X_i + h_i \qquad (h_i = \operatorname{Re} f_i) \tag{11.4}$$

We recognise a Hopfield network with discrete time steps, which we recall from Section 6.1 minimises H_I. As mentioned then, a first issue is that local minima are likely to be found, although this can be mitigated with, for example, suitable noise. A second is that with unequal amplitudes it is difficult to ensure that the correct Ising Hamiltonian is minimised, since the couplings are modified according to displacements unknown at the start of the machine's operation.

A Product of Binary Variables

(Return to Section 4.5.1)

Let $x, y, z \in \{0, 1\}$ and A, B be the statements

$$\text{A:} \quad xy = z \quad \text{and} \quad \text{B:} \quad \underbrace{xy - 2xz - 2yz + 3z}_{=f(x,y,z)} = 0 \tag{A.1}$$

Suppose A is true. We have

$$\begin{aligned} f(x, y, z) &= f(x, y, xy) && \text{(A.2)} \\ &= xy - 2x^2y - 2y^2x + 3xy && \text{(A.3)} \\ &= \underbrace{2xy\left(2 - (x^2 + y^2)\right)}_{\tilde{f}(x,y)} && \text{(A.4)} \end{aligned}$$

Then clearly $\tilde{f}(0,0) = \tilde{f}(1,0) = \tilde{f}(0,1) = \tilde{f}(1,1) = 0$ and so B is true. From the contrapositive (not B implies not A)

$$f(x, y, z) > 0 \Rightarrow xy \neq z \tag{A.5}$$

Next suppose B is true. We have

$$f(x, y, z) = 0 \Rightarrow z\,(2(x + y) - 3) = xy \tag{A.6}$$

As this is symmetric under $x \leftrightarrow y$, there are three cases to consider:

$$\begin{aligned} &\bar{x} \wedge \bar{y} \Rightarrow -3z = 0 \Rightarrow z = 0 = xy && \text{(A.7)} \\ &\bar{x} \wedge y \vee x \wedge \bar{y} \Rightarrow -z = 0 \Rightarrow z = 0 = xy && \text{(A.8)} \\ &x \wedge y \Rightarrow z = 1 = xy && \text{(A.9)} \end{aligned}$$

So B implies A. The contrapositive is

$$xy \neq z \Rightarrow f(x, y, z) \neq 0 \Leftrightarrow f(x, y, z) > 0 \tag{A.10}$$

since f is non-negative. Hence (4.19) and (4.20) have been established. You can verify both using the truth table below.

x	0	0	0	0	1	1	1	1
y	0	0	1	1	0	0	1	1
z	0	1	0	1	0	1	0	1
$f(x,y,z)$	0	3	0	1	0	1	1	0

B The Mandelung Transformation

(Return to Section 8.3.2)

Writing $\psi(\boldsymbol{r},t)=\sqrt{n(\boldsymbol{r},t)}e^{iS(\boldsymbol{r},t)}$ we calculate

$$\dot{\psi}=\frac{1}{2}\frac{\dot{n}}{\sqrt{n}}e^{iS}+i\sqrt{n}\dot{S}e^{iS} \tag{B.1}$$

$$\nabla\psi=\frac{1}{2}\frac{\nabla n}{\sqrt{n}}e^{iS}+i\sqrt{n}(\nabla S)e^{iS} \tag{B.2}$$

$$\nabla^2\psi=\left(\frac{1}{2}\frac{\nabla^2 n}{\sqrt{n}}-\frac{1}{4}\frac{(\nabla n)^2}{n^{3/2}}+i\frac{\nabla n\cdot\nabla S}{\sqrt{n}}+i\sqrt{n}(\nabla^2 S)-\sqrt{n}(\nabla S)^2\right)e^{iS} \tag{B.3}$$

and then from the GPE

$$i\hbar\frac{\partial\psi(\boldsymbol{r},t)}{\partial t}=\left(-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r},t)+U_0|\psi(\boldsymbol{r},t)|^2\right)\psi(\boldsymbol{r},t) \tag{B.4}$$

$$\begin{aligned}\Rightarrow i\hbar\left[\frac{1}{2}\frac{\dot{n}}{\sqrt{n}}+i\sqrt{n}\dot{S}\right]=&-\frac{\hbar^2}{2m}\left[\frac{1}{2}\frac{\nabla^2 n}{\sqrt{n}}-\frac{1}{4}\frac{(\nabla n)^2}{n^{3/2}}+i\frac{\nabla n\cdot\nabla S}{\sqrt{n}}+i\sqrt{n}\nabla^2 S-\sqrt{n}(\nabla S)^2\right]\\&+(V+U_0 n)\sqrt{n}\end{aligned} \tag{B.5}$$

The imaginary part gives

$$\frac{1}{2}\frac{\dot{n}}{\sqrt{n}}=-\frac{\hbar}{2m}\left(\frac{\nabla n\cdot\nabla S}{\sqrt{n}}+\sqrt{n}\nabla^2 S\right) \tag{B.6}$$

$$\Rightarrow\dot{n}=-\nabla n\cdot\left(\frac{\hbar}{m}\nabla S\right)-n\nabla\cdot\left(\frac{\hbar}{m}\nabla S\right)\equiv-\nabla\cdot(n\boldsymbol{v}) \tag{B.7}$$

That is

$$\frac{\partial n}{\partial t}+\nabla\cdot(n\boldsymbol{v})=0 \tag{B.8}$$

As for the real part,

$$-\hbar\dot{S}=\overbrace{\frac{\hbar^2}{2m}(\nabla S)^2}^{\frac{1}{2}mv^2}+V+U_0 n-\frac{\hbar^2}{2m}\overbrace{\left(\frac{\nabla^2 n}{2n}-\frac{(\nabla n)^2}{4n^{5/2}}\right)}^{(\nabla^2\sqrt{n})/\sqrt{n}} \tag{B.9}$$

$$\Rightarrow m\frac{\partial\boldsymbol{v}}{\partial t}\equiv\hbar\nabla\dot{S}=-\nabla\left(\frac{1}{2}mv^2+V+U_0 n-\frac{\hbar^2}{2m}\frac{\nabla^2\sqrt{n}}{\sqrt{n}}\right) \tag{B.10}$$

Finally in the expression for energy (8.3) we note that the kinetic energy density splits up into a classical and non-classical part:

$$\frac{\hbar^2}{2m}|\nabla\psi|^2=\frac{\hbar^2}{2m}\frac{(\nabla n)^2}{4n}+\frac{\hbar^2}{2m}n(\nabla S)^2\equiv\frac{\hbar^2}{2m}\left(\nabla\sqrt{n}\right)^2+\frac{1}{2}mnv^2 \tag{B.11}$$

References

[1] G. De las Cuevas and T. S. Cubitt, Science **351**, 1180 (2016).

[2] R. M. Karp, in *Complexity of computer computations: proceedings of a symposium on the complexity of computer computations*, edited by R. E. Miller, J. W. Thatcher, and J. D. Bohlinger (Springer, 1972), pp. 85–103.

[3] M. Newman, *Networks*, 2nd ed. (Oxford University Press, 2018).

[4] S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*, Studies in Nonlinearity (CRC Press, 2000).

[5] S. H. Strogatz, Physica D: Nonlinear Phenomena **143**, 1 (2000).

[6] A. Lucas, Frontiers in Physics **2**, 10.3389/fphy.2014.00005 (2014).

[7] M. Fox, *Quantum optics an introduction* (Oxford University Press, 2006).

[8] J. Keeling and N. G. Berloff, Contemporary Physics **52**, 131 (2011).

[9] N. G. Berloff et al., Nature Materials **16**, 1120 (2017).

[10] H. Miesner et al., Science **279**, 1005 (1998).