M.Phys. C6 Theory Option/M.Math.Phys. Advanced Quantum Theory Lecture Notes for 2020-2021

Lectures given by Sid Parameswaran in Michaelmas 2020
Lecture Notes written by F.H.L. Essler
The Rudolf Peierls Centre for Theoretical Physics
Oxford University, Oxford OX1 3PU, UK

November 19, 2020

These notes cover Parts I-IV of the lectures (Functional Methods; Transfer Matrices; Many-Particle Quantum Systems; and Landau Theory), which are common to both M.Phys. C6 Theory Option and to the M.Math.Phys. Advanced Quantum Theory.
Please report errors and typos to sid.parameswaran@physics.ox.ac.uk.

©2015 F.H.L. Essler.

Contents

I Functional Methods in Quantum Mechanics **4**

1 Some Mathematical Background **4**
1.1 Functionals . . . 4
1.2 Functional differentiation . . . 5
1.3 Multidimensional Gaussian Integrals . . . 7
1.4 Homework Questions 1&2 . . . 8

2 Path Integrals in Quantum Mechanics **9**
2.1 The Propagator . . . 9
2.1.1 Propagator as a "Functional Integral" . . . 11
2.2 Quantum Mechanics à la Feynman . . . 12
2.3 Classical Limit and Stationary Phase Approximation . . . 12
2.4 The Propagator for Free Particles . . . 12
2.5 Homework Questions 3-5 . . . 14

3 Path Integrals in Quantum Statistical Mechanics **16**
3.1 Harmonic Oscillator at $T > 0$: a first encounter with Generating Functionals 17
3.1.1 Imaginary Time Green's Function of the Harmonic Oscillator . . . 21
3.2 Homework Question 6 . . . 23
3.3 Correlation Functions . . . 24
3.3.1 Wick's Theorem . . . 26
3.4 Probability distribution of position . . . 26
3.5 Perturbation Theory and Feynman Diagrams . . . 28
3.5.1 Partition Function of the anharmonic oscillator . . . 28
3.6 Homework Question 7 . . . 32

II Path Integrals and Transfer Matrices **35**

4 Relation of D dimensional quantum systems to $D+1$ dimensional classical ones **35**
4.1 Some Facts from Statistical Physics . . . 35
4.2 Quantum Mechanical Particle . . . 35

5 The Ising Model **38**
5.1 Statistical mechanics of the Ising model . . . 39
5.2 The One-Dimensional Ising Model . . . 39
5.2.1 Transfer matrix approach . . . 40
5.2.2 Averages of observables in the transfer matrix formalism . . . 41
5.2.3 The related zero-dimensional quantum model . . . 42
5.3 The Two-Dimensional Ising Model . . . 43
5.3.1 Transfer Matrix Method . . . 43
5.3.2 Spontaneous Symmetry Breaking . . . 44
5.3.3 The Related One-Dimensional Quantum Problem . . . 46
5.4 Homework Questions 8-10 . . . 48

III Many-Particle Quantum Mechanics **50**

6 Second Quantization **50**
6.1 Systems of Independent Particles . . . 50
6.1.1 Occupation Number Representation . . . 51
6.2 Fock Space . . . 52
6.2.1 Creation and Annihilation Operators . . . 52
6.2.2 Basis of the Fock Space . . . 53
6.3 Homework Questions 11-13 . . . 53
6.3.1 Change of Basis . . . 54
6.4 Second Quantized Form of Operators . . . 55
6.4.1 Occupation number operators . . . 55
6.4.2 Single-particle operators . . . 55
6.4.3 Two-particle operators . . . 58
6.5 Homework Question 14 . . . 60

7 Application I: The Ideal Fermi Gas **61**
7.1 Quantization in a large, finite volume . . . 61
7.1.1 Ground State . . . 62
7.1.2 Excitations . . . 63
7.1.3 Density Correlations . . . 64
7.2 Homework Questions 15-16 . . . 66

8 Application II: Weakly Interacting Bosons **68**
8.1 Ideal Bose Gas . . . 68
8.2 Bogoliubov Approximation . . . 68
8.3 Bogoliubov Transformation . . . 69
8.4 Ground State and Low-lying Excitations . . . 70
8.5 Ground state correlation functions . . . 71
8.6 Depletion of the Condensate . . . 72

9 Application III: Spinwaves in a Ferromagnet **72**
9.1 Heisenberg model and spin-rotational SU(2) symmetry . . . 73
9.2 Exact ground states . . . 73
9.3 Spontaneous Symmetry Breaking . . . 75
9.4 Holstein-Primakoff Transformation . . . 75
9.4.1 Heisenberg Antiferromagnet . . . 76
9.5 Homework Questions 17-18 . . . 77

10 Path Integral for interacting Bose systems **78**
10.1 Coherent States . . . 78
10.2 Partition Function . . . 79

IV Landau Theory of Phase Transitions **81**

11 Phase Transitions **81**

12 The Ising Model **81**
12.1 Peierls Argument . . . 82
12.2 Mean Field Theory . . . 84
12.3 Solution of the self-consistency equation for $h = 0$ . . . 85
12.4 Vicinity of the Phase Transition . . . 86

13 Critical Behaviour and Universality **86**
13.1 Universality . . . 87

14 Landau Theory **88**
14.1 Thermodynamic Equilibrium . . . 89
14.2 Beyond the Landau Free Energy . . . 90
14.3 Saddle Point Approximation . . . 92
14.4 Mean Field Exponents . . . 92
14.5 Homework Questions 19-21 . . . 95

15 Other Examples of Phase Transitions **96**
15.1 Isotropic-Nematic Transition in Liquid Crystals . . . 96
15.2 Superfluid Transition in Weakly Interacting Bosons . . . 98

16 Ising Model and Φ^4 Theory **98**

17 Transverse-Field Ising Model **100**
17.1 Symmetries . . . 100
17.2 Phase Diagram . . . 101
17.2.1 Disordered Phase . . . 101
17.2.2 Ordered Phase . . . 102
17.3 Duality . . . 103
17.4 Mean Field Theory . . . 106
17.5 Jordan-Wigner 'Fermionization' of the TFIM . . . 109
17.6 Solution of the Ising Model via Bogoliubov Transformation . . . 110
17.7 Continuum Limit and Critical Theory . . . 111
17.7.1 Path Integral and Critical Theory for the TFIM . . . 114
17.7.2 Correlation Functions, Structure Factors, and Long-Range Order . . . 118
17.7.3 Ising Symmetry in $CoNb_2O_6$ . . . 120

Some general remarks:

These notes aim to be self-contained. Homework questions are marked in red, and are placed at appropriate positions in the text, i.e. to work them out you will require only the preceeding material. Passages marked in blue give details on derivations we don't have time to go through in the lectures, or present material that goes beyond the core of the course. In some cases this material will be very useful for particular homework problems. All of the material covered in the course can be found in some form or other in a variety of books. These is no book that covers everything. Some useful references are

Sections of the text written in blue are not formally examinable.

- Path Integrals

  R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals*, Dover Books.

  J. Zinn-Justin, *Path Integrals in Quantum Mechanics*, Oxford.

- Many-particle QM

  R.P. Feynman, *Statistical Mechanics: A Set of Lectures*, Westview Press.

  A. Altland and B.D. Simons, *Condensed Matter Field Theory*, Cambridge.

- Landau Theory of Phase Transitions

  M. Kardar, *Statistical Physics of Fields*, Cambridge.

Part I
Functional Methods in Quantum Mechanics

1 Some Mathematical Background

Functional Methods form a central part of modern theoretical physics. In the following we introduce the notion of functionals and how to manipulate them.

1.1 Functionals

What is a *functional*? You all know that a real function can be viewed as a *map* from e.g. an interval $[a, b]$ to the real numbers

$$f : [a, b] \to \mathbb{R} \ , \quad x \to f(x). \tag{1}$$

A functional is similar to a function in that it maps all elements in a certain domain to real numbers, however, the nature of its domain is very different. Instead of acting on all points of an interval or some other subset of the real numbers, the domain of functionals consists of (suitably chosen) classes of functions. In other words, given some class $\{f\}$ of functions, a functional F is a map

$$F : \{f\} \to \mathbb{R} \ , \quad f \to F[f]. \tag{2}$$

We now consider two specific examples of functionals.

1. The distance between two points. A very simple functional F consists of the map which assigns to all paths between two fixed points the length of the path. To write this functional explicitly, let us consider a simple two-dimensional situation in the (x, y) plane and choose two points (x_1, y_1) and (x_2, y_2). We consider the set of paths that do not turn back, i.e. paths along which x increases monotonically as we go from (x_1, y_1) to (x_2, y_2). These can be described by the set of functions $\{f\}$ on the interval $[x_1, x_2]$ satisfying $f(x_1) = y_1$ and $f(x_2) = y_2$. The length of a path is then given by the well-known expression

$$F[f(x)] = \int_{x_1}^{x_2} dx' \, \sqrt{1 + f'(x')^2} \ . \tag{3}$$

2. *Action Functionals.* These are very important in Physics. Let us recall their definition in the context of classical mechanics. Start with n generalised coordinates $\mathbf{q}(t) = (q_1(t), \ldots, q_n(t))$ and a Lagrangian $L = L(\mathbf{q}, \dot{\mathbf{q}})$. Then, the action functional $S[\mathbf{q}]$ is defined by

$$S[\mathbf{q}] = \int_{t_1}^{t_2} dt\, L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \ . \tag{4}$$

It depends on classical paths $\mathbf{q}(t)$ between times t_1 and t_2 satisfying the boundary conditions $\mathbf{q}(t_1) = \mathbf{q}_1$ and $\mathbf{q}(t_2) = \mathbf{q}_2$.

1.2 Functional differentiation

In both the examples given above a very natural question to ask is what function *extremizes* the functional. In the first example this corresponds to wanting to know the path that minimizes the distance between two points. In the second example the extremum of the action functional gives the solutions to the classical equations of motion. This is known as *Hamilton's principle*. In order to figure out what function extremizes the functional it is very useful to generalize the notion of a derivative. For our purposes we define the *functional derivative* by

$$\boxed{\frac{\delta F[f(x)]}{\delta f(y)} = \lim_{\epsilon \to 0} \frac{F[f(x) + \epsilon \delta(x-y)] - F[f(x)]}{\epsilon} \ .} \tag{5}$$

Here, as usual, we should think of the δ-function as being defined as the limit of a test function, e.g.

$$\delta(x) = \lim_{a \to 0} \frac{1}{\sqrt{\pi} a} e^{-x^2/a^2} \ , \tag{6}$$

and take the limit $a \to 0$ only in the end (after commuting the limit with all other operations such as the $\lim_{\epsilon \to 0}$ in (5)). Importantly, the derivative defined in this way is a linear operation which satisfies the product and chain rules of ordinary differentiation and commutes with ordinary integrals and derivatives. Let us see how functional differentiation works for our two examples.

1. The distance between two points. In analogy with finding stationary points of functions we want to extremize (3) by setting its functional derivative equal to zero

$$0 = \frac{\delta F[f(x)]}{\delta f(y)} \ . \tag{7}$$

We first do the calculation by using the definition (5).

$$\frac{\delta F[f(x)]}{\delta f(y)} = \lim_{\epsilon \to 0} \int_{x_1}^{x_2} dx' \ \frac{\sqrt{1 + [f'(x') + \epsilon \delta'(x'-y)]^2} - \sqrt{1 + [f'(x')]^2}}{\epsilon}. \tag{8}$$

The Taylor expansion of the square root is $\sqrt{1+2\epsilon} = 1 + \epsilon + \ldots$, which gives

$$\sqrt{1 + [f'(x') + \epsilon \delta'(x'-y)]^2} \ = \ \sqrt{1 + [f'(x')]^2} + \frac{\epsilon f'(x') \delta'(x'-y)}{\sqrt{1 + [f'(x')]^2}} + \mathcal{O}(\epsilon^2) \ , \tag{9}$$

where $\delta'(x)$ is the derivative of the delta-function and $\mathcal{O}(\epsilon^2)$ denote terms proportional to ϵ^2. Substituting this back into (8) we have [1]

$$\frac{\delta F[f(x)]}{\delta f(y)} = \int_{x_1}^{x_2} dx' \ \frac{\delta'(x'-y) f'(x')}{\sqrt{1 + [f'(x')]^2}} = -\frac{d}{dy} \frac{f'(y)}{\sqrt{1 + [f'(y)]^2}}. \tag{11}$$

[1] In the last step we have used

$$\int_a^b dx' \delta'(x'-y) g(x') = -g'(y) \ , \tag{10}$$

which can be proved by "integration by parts".

The solution to (7) is thus

$$f'(y) = \text{const}, \tag{12}$$

which describes a straight line. In practice we don't really go back to the definition of the functional derivative any more than we use the definition of an ordinary derivative to work it out, but proceed as follows.

- We first interchange the functional derivative and the integration

$$\frac{\delta F[f(x)]}{\delta f(y)} = \int_{x_1}^{x_2} dx' \; \frac{\delta}{\delta f(y)} \sqrt{1 + [f'(x')]^2}. \tag{13}$$

- Next we use the chain rule

$$\frac{\delta \sqrt{1 + f'(x')^2}}{\delta f(y)} = \frac{1}{2\sqrt{1 + f'(x')^2}} \frac{\delta(1 + f'(x')^2)}{\delta f(y)} = \frac{f'(x')}{\sqrt{1 + f'(x')^2}} \frac{\delta f'(x')}{\delta f(y)}. \tag{14}$$

- Finally we interchange the functional and the ordinary derivative

$$\frac{\delta f'(x')}{\delta f(y)} = \frac{d}{dx'} \frac{\delta f(x')}{\delta f(y)} = \frac{d}{dx'} \delta(x' - y) \; . \tag{15}$$

  The last identity follows from our definition (5).

Now we can put everything together and arrive at the same answer (11).

Exercise 1

2. Next we want to try out these ideas on our second example and extremize the classical action (4) in order to obtain the classical equations of motion. We first interchange functional derivative and integration and then use the chain rule to obtain

$$\begin{aligned} \frac{\delta S[\mathbf{q}]}{\delta q_i(t)} &= \frac{\delta}{\delta q_i(t)} \int_{t_1}^{t_2} d\tilde{t} \, L(\mathbf{q}(\tilde{t}), \dot{\mathbf{q}}(\tilde{t})) && (16) \\ &= \int_{t_1}^{t_2} d\tilde{t} \left[\frac{\partial L}{\partial q_j}(\mathbf{q}, \dot{\mathbf{q}}) \frac{\delta q_j(\tilde{t})}{\delta q_i(t)} + \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \frac{\delta \dot{q}_j(\tilde{t})}{\delta q_i(t)} \right] && (17) \\ & && (18) \end{aligned}$$

   We now use that $\frac{\delta \dot{q}_j(\tilde{t})}{\delta q_i(t)} = \frac{d}{d\tilde{t}} \frac{\delta q_j(\tilde{t})}{\delta q_i(t)}$ and integrate by parts with respect to $\tilde{t}$

$$\begin{aligned} \frac{\delta S[\mathbf{q}]}{\delta q_i(t)} &= \int_{t_1}^{t_2} d\tilde{t} \left[\frac{\partial L}{\partial q_j}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{d\tilde{t}} \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \right] \frac{\delta q_j(\tilde{t})}{\delta q_i(t)} && (19) \\ &= \int_{t_1}^{t_2} d\tilde{t} \left[\frac{\partial L}{\partial q_j}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{d\tilde{t}} \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \right] \delta_{ij} \delta(\tilde{t} - t) = \frac{\partial L}{\partial q_i}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) \; . && (20) \end{aligned}$$

   In the second last step we have used

$$\frac{\delta q_j(\tilde{t})}{\delta q_i(t)} = \delta_{ij} \delta(\tilde{t} - t) \; , \tag{21}$$

   which follows straightforwardly from our general definition (5). Thus we conclude that the extrema of the classical action are given by paths that fulfil the equations of motion

$$\boxed{\frac{\partial L}{\partial q_i}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) = 0.} \tag{22}$$

   Nice.

1.3 Multidimensional Gaussian Integrals

As a reminder, we start with a simple one-dimensional Gaussian integral over a single variable y. It is given by

$$\boxed{I(z) \equiv \int_{-\infty}^{\infty} dy\, \exp(-\frac{1}{2}zy^2) = \sqrt{\frac{2\pi}{z}}\,,} \tag{23}$$

where z is a complex number with $\mathrm{Re}(z) > 0$. The standard proof of this relation involves writing $I(z)^2$ as a two-dimensional integral over y_1 and y_2 and then introducing two-dimensional polar coordinates $r = \sqrt{y_1^2 + y_2^2}$ and φ. Explicitly,

$$I(z)^2 = \int_{-\infty}^{\infty} dy_1\, \exp(-\frac{1}{2}zy_1^2) \int_{-\infty}^{\infty} dy_2\, \exp(-\frac{1}{2}zy_2^2) = \int_{-\infty}^{\infty} dy_1 \int_{-\infty}^{\infty} dy_2\, \exp(-\frac{1}{2}z(y_1^2+y_2^2)) \tag{24}$$

$$= \int_0^{2\pi} d\varphi \int_0^{\infty} dr\, r \exp(-\frac{1}{2}zr^2) = \frac{2\pi}{z}\,. \tag{25}$$

Next we consider n-dimensional Gaussian integrals

$$W_0(\mathbf{A}) \equiv \int d^n\mathbf{y}\, \exp\left(-\frac{1}{2}\mathbf{y}^T\mathbf{A}\mathbf{y}\right)\,, \tag{26}$$

over variables $\mathbf{y} = (y_1, \ldots, y_n)$, where $\mathbf{A}$ is a symmetric, positive definite matrix (all its eigenvalues are positive). This integral can be reduced to a product of one-dimensional Gaussian integrals by diagonalising the matrix $\mathbf{A}$. Consider an orthogonal rotation $\mathbf{O}$ such that $\mathbf{A} = \mathbf{O}\mathbf{D}\mathbf{O}^T$ with a diagonal matrix $\mathbf{D} = \mathrm{diag}(a_1, \ldots, a_n)$. The eigenvalues a_i are strictly positive since we have assumed that $\mathbf{A}$ is positive definite. Introducing new coordinates $\tilde{\mathbf{y}} = \mathbf{O}^T\mathbf{y}$ we can write

$$\mathbf{y}^T\mathbf{A}\mathbf{y} = \tilde{\mathbf{y}}^T\mathbf{D}\tilde{\mathbf{y}} = \sum_{i=1}^{n} a_i\tilde{y}_i^2\,, \tag{27}$$

where the property $\mathbf{O}^T\mathbf{O} = \mathbf{1}$ of orthogonal matrices has been used. Note further that the Jacobian of the coordinate change $\mathbf{y} \to \tilde{\mathbf{y}}$ is one, since $|\det(\mathbf{O})| = 1$. Hence, using Eqs. (23) and (27) we find for the integral (26)

$$W_0(\mathbf{A}) = \prod_{i=1}^{n} \int d\tilde{y}_i \exp(-\frac{1}{2}a_i\tilde{y}_i^2) = (2\pi)^{n/2}(a_1a_2\ldots a_n)^{-1/2} = (2\pi)^{n/2}(\det\mathbf{A})^{-1/2}\,. \tag{28}$$

To summarise, we have found for the multidimensional Gaussian integral (26) that

$$\boxed{W_0(\mathbf{A}) = (2\pi)^{n/2}(\det\mathbf{A})^{-1/2}\,,} \tag{29}$$

a result which will be of some importance in the following. We note that if we multiply the matrix $\mathbf{A}$ by a complex number z with $\mathrm{Re}(z) > 0$ and then follow through exactly the same steps, we find

$$W_0(z\mathbf{A}) = \left(\frac{2\pi}{z}\right)^{n/2} (\det\mathbf{A})^{-1/2}\,. \tag{30}$$

One obvious generalisation of the integral (26) involves adding a term linear in $\mathbf{y}$ in the exponent, that is

$$W_0(\mathbf{A},\mathbf{J}) \equiv \int d^n\mathbf{y}\, \exp\left(-\frac{1}{2}\mathbf{y}^T\mathbf{A}\mathbf{y} + \mathbf{J}^T\mathbf{y}\right)\,. \tag{31}$$

Here $\mathbf{J} = (J_1, \ldots, J_n)$ is an n-dimensional vector. Changing variables $\mathbf{y} \to \tilde{\mathbf{y}}$, where

$$\mathbf{y} = \mathbf{A}^{-1}\mathbf{J} + \tilde{\mathbf{y}} \tag{32}$$

this integral can be written as

$$W_0(\mathbf{A}, \mathbf{J}) = \exp\left(\frac{1}{2}\mathbf{J}^T\mathbf{A}^{-1}\mathbf{J}\right) \int d^n\tilde{\mathbf{y}}\, \exp\left(-\frac{1}{2}\tilde{\mathbf{y}}^T\mathbf{A}\tilde{\mathbf{y}}\right) . \tag{33}$$

The remaining integral is Gaussian without a linear term, so can be easily carried out using the above results. Hence, one finds

$$\boxed{W_0(\mathbf{A}, \mathbf{J}) = (2\pi)^{n/2}(\det\mathbf{A})^{-1/2} \exp\left(\frac{1}{2}\mathbf{J}^T\mathbf{A}^{-1}\mathbf{J}\right) .} \tag{34}$$

1.4 Homework Questions 1&2

Question 1. Consider paths $\mathbf{X} = \mathbf{X}(\tau)$, where τ is a parameter, and the functional

$$l[\mathbf{X}] = \int_{\tau_0}^{\tau_1} d\tau\, n(\mathbf{X}) \sqrt{\frac{d\mathbf{X}}{d\tau} \cdot \frac{d\mathbf{X}}{d\tau}} ,$$

where $n = n(\mathbf{X})$ is a function. (The minima of this functional can be interpreted as light rays propagating in a medium with refractive index n.)
a) Derive the differential equation which has to be satisfied by minimal paths $\mathbf{X}$.
b) Consider a two-dimensional situation with paths $\mathbf{X}(\tau) = (X(\tau), Y(\tau))$ in the x, y plane and a function $n = n_0 + (n_1 - n_0)\,\theta(x)$. (The Heaviside function $\theta(x)$ is defined to be 0 for $x < 0$ and 1 for $x \geq 0$. Recall that $\theta'(x) = \delta(x)$.) Solve the differential equation in a) for this situation, using the coordinate x as parameter τ along the path.
c) Show that the solution in b) leads to the standard law for refraction at the boundary between two media with refractive indices n_0 and n_1.

Question 2. a) Evaluate the Gaussian integral

$$\int_{-\infty}^{\infty} dx\; e^{-\frac{1}{2}zx^2} \tag{35}$$

for a complex constant z. What is the requirement on z for the integral to exist?
b) The gamma function Γ is defined by

$$\Gamma(s+1) = \int_0^{\infty} dx\, x^s e^{-x} .$$

c) Show that $\Gamma(1) = 1$ and $\Gamma(s+1) = s\Gamma(s)$. (Hence $\Gamma(n+1) = n!$)
d) Take s to be real and positive. Evaluate $\Gamma(s+1)$ in the *steepest descent approximation*: write the integrand in the form $e^{f(x)}$ and argue that for large $s \gg 1$ the dominant contribution to the integral arises from the minima of $f(x)$. Expand the function to quadratic order around the minimum, argue that you may extend the integration boundaries to $\pm\infty$, and then carry out the resulting integral. Your result is known as *Stirling's approximation*: it tells you what $n!$ is when n becomes large.
e)* The following extension is for complex analysis afficionados, so simply omit it if you haven't taken the short option. Take s to be complex with positive real part. Deform the contour in a suitable way, so that you can again apply a steepest descent approximation. Ponder the name of the method. What is Stirling's approximation for complex s?

2 Path Integrals in Quantum Mechanics

So far you have encountered two ways of doing QM:

1. Following Schrödinger, we can solve the Schrödinger equation for the wave function → Fun with PDEs...

2. Following Heisenberg, we can work with operators, commutation relations, eigenstates → Fun with Linear Algebra...

Historically it took some time for people to realize that these are in fact equivalent. To quote the great men: *I knew of Heisenberg's theory, of course, but I felt discouraged, not to say repelled, by the methods of transcendental algebra, which appeared difficult to me, and by the lack of visualizability. (Schrödinger in 1926)*

The more I think about the physical portion of Schrödinger's theory, the more repulsive I find it. What Schrödinger writes about the visualizability of his theory is probably not quite right, in other words it's crap. (Heisenberg, writing to Pauli in 1926)

There is a third approach to QM, due to *Feynman*. He developed it when he was a graduate student, inspired by a mysterious remark in a paper by Dirac. Those were the days! Feynman's approach is particularly useful for QFTs and many-particle QM problems, as it makes certain calculations much easier. We will now introduce it by working backwards. The central object in Feynman's method is something called a *propagator*. We'll now work out what this is using the Heisenberg/Schrödinger formulation of QM you know and love. After we have done that, we formulate QM à la Feynman.

2.1 The Propagator

Our starting point is the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle. \tag{36}$$

We recall that the wave function is given by

$$\psi(\vec{x},t) = \langle\vec{x}|\psi(t)\rangle. \tag{37}$$

Eqn (36) can be integrated to give

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle \tag{38}$$

The *time-evolution operator* in QM is thus (assuming that H is time-independent)

$$U(t;t_0) = e^{-\frac{i}{\hbar}H(t-t_0)}. \tag{39}$$

A central object in Feynman's approach is the *propagator*

$$\boxed{\langle\vec{x}'|U(t;t_0)|\vec{x}\rangle,} \tag{40}$$

where $|\vec{x}\rangle$ are the simultaneous eigenstates of the position operators $\hat{x}$, $\hat{y}$ and $\hat{z}$. The propagator is the probability amplitude for finding our QM particle at position $\vec{x}'$ at time t, if it started at position $\vec{x}$ at time t_0. To keep notations simple, we now consider a particle moving in one dimension with time-independent Hamiltonian

$$H = \hat{T} + \hat{V} = \frac{\hat{p}^2}{2m} + V(\hat{x}). \tag{41}$$

We want to calculate the propagator

$$\langle x_N|U(t;0)|x_0\rangle. \tag{42}$$

It is useful to introduce small time steps

$$t_n = n\epsilon \ , \quad n = 0, \ldots, N, \tag{43}$$

where $\epsilon = t/N$. Then we have by construction

$$U(t;0) = \left(e^{-\frac{i}{\hbar}H\epsilon}\right)^N . \tag{44}$$

The propagator is

$$\begin{aligned} \langle x_N|U(t;0)|x_0\rangle &= \langle x_N|e^{-\frac{i}{\hbar}H\epsilon}\cdots e^{-\frac{i}{\hbar}H\epsilon}|x_0\rangle \\ &= \int dx_{N-1}\cdots\int dx_1\ \langle x_N|e^{-\frac{i}{\hbar}H\epsilon}|x_{N-1}\rangle\langle x_{N-1}|e^{-\frac{i}{\hbar}H\epsilon}|x_{N-2}\rangle\ldots\langle x_1|e^{-\frac{i}{\hbar}H\epsilon}|x_0\rangle, \end{aligned} \tag{45}$$

where we have inserted $N-1$ resolutions of the identity in terms of position eigenstates

$$\mathbf{1} = \int dx\ |x\rangle\langle x| \ . \tag{46}$$

This expression now has a very nice and intuitive interpretation, see Fig. 1: The propagator, i.e. the

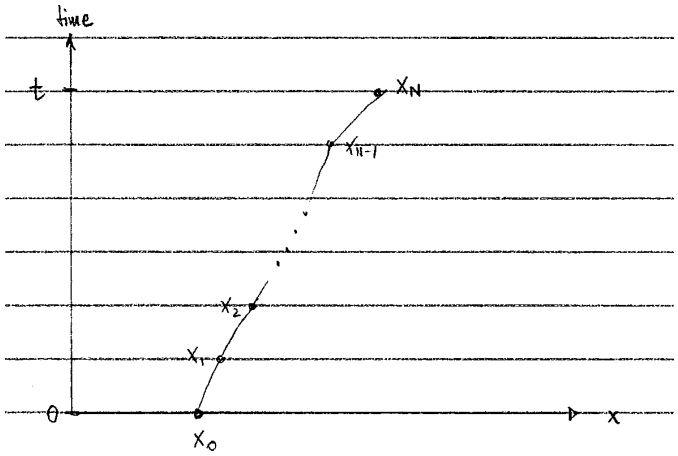


Figure 1: Propagator as sum over paths.

probabilty amplitude for finding the particle at position x_N and time t given that it was at position x_0 at time 0 is given by the sum over all "paths" going from x_0 to x_N (as $x_1,\ldots,\ x_{N-1}$ are integrated over).

In the next step we determine the "infinitesimal propagator"

$$\langle x_{n+1}|e^{-\frac{i}{\hbar}H\epsilon}|x_n\rangle. \tag{47}$$

Importantly we have $[\hat{T},\hat{V}] \neq 0$ and concomitantly

$$e^{\alpha(\hat{T}+\hat{V})} \neq e^{\alpha\hat{T}}e^{\alpha\hat{V}}. \tag{48}$$

However, using that ϵ is infinitesimal, we have

$$\begin{aligned} e^{-\frac{i}{\hbar}\epsilon(\hat{T}+\hat{V})} &= 1 - \frac{i}{\hbar}\epsilon(\hat{T}+\hat{V}) + \mathcal{O}(\epsilon^2) \ , \\ e^{-\frac{i}{\hbar}\epsilon\hat{T}}e^{-\frac{i}{\hbar}\epsilon\hat{V}} &= 1 - \frac{i}{\hbar}\epsilon(\hat{T}+\hat{V}) + \mathcal{O}(\epsilon^2). \end{aligned} \tag{49}$$

So up to terms of order ϵ^2 we have

$$\langle x_{n+1}|e^{-\frac{i}{\hbar}H\epsilon}|x_n\rangle \simeq \langle x_{n+1}|e^{-\frac{i}{\hbar}\hat{T}\epsilon}e^{-\frac{i}{\hbar}\hat{V}\epsilon}|x_n\rangle = \langle x_{n+1}|e^{-\frac{i}{\hbar}\hat{T}\epsilon}|x_n\rangle e^{-\frac{i}{\hbar}V(x_n)\epsilon}, \tag{50}$$

where we have used that $\hat{V}|x\rangle = V(x)|x\rangle$. As $\hat{T} = \hat{p}^2/2m$ it is useful to insert a complete set of momentum eigenstates [2] to calculate

$$\begin{aligned}\langle x_{n+1}|e^{-\frac{i}{\hbar}\hat{T}\epsilon}|x_n\rangle &= \int \frac{dp}{2\pi\hbar}\langle x_{n+1}|e^{-\frac{i\hat{p}^2\epsilon}{2m\hbar}}|p\rangle\langle p|x_n\rangle = \int \frac{dp}{2\pi\hbar}e^{-\frac{ip^2\epsilon}{2m\hbar}-i\frac{p}{\hbar}(x_n-x_{n+1})} \\ &= \sqrt{\frac{m}{2\pi i\hbar\epsilon}}e^{\frac{im}{2\hbar\epsilon}(x_n-x_{n+1})^2}. \end{aligned} \tag{51}$$

In the second step we have used that $\hat{p}|p\rangle = p|p\rangle$ and that

$$\langle x|p\rangle = e^{\frac{ipx}{\hbar}}. \tag{52}$$

The integral over p is performed by changing variables to $p' = p + \frac{m}{\epsilon}(x_n - x_{n+1})$ (and giving ϵ a very small imaginary part in order to make the integral convergent). Substituting (51) and (50) back into our expression (45) for the propagator gives

$$\boxed{\langle x_N|U(t;0)|x_0\rangle = \lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar\epsilon}\right]^{\frac{N}{2}}\int dx_1\ldots dx_{N-1}\exp\left(\frac{i\epsilon}{\hbar}\sum_{n=0}^{N-1}\frac{m}{2}\left(\frac{x_{n+1}-x_n}{\epsilon}\right)^2 - V(x_n)\right)} \tag{53}$$

Note that in this expression there are no operators left.

2.1.1 Propagator as a "Functional Integral"

The way to think about (53) is as a *sum over trajectories*:

- $x_0, \ldots, x_N$ constitute a discretization of a *path* $x(t')$, where we set $x_n \equiv x(t_n)$.
- We then have
$$\frac{x_{n+1}-x_n}{\epsilon} = \frac{x(t_{n+1})-x(t_n)}{t_{n+1}-t_n} \simeq \dot{x}(t_n), \tag{54}$$
and
$$\epsilon\sum_{n=0}^{N-1}\frac{m}{2}\left(\frac{x_{n+1}-x_n}{\epsilon}\right)^2 - V(x_n) \simeq \int_0^t dt'\left[\frac{m}{2}\dot{x}^2(t') - V(x)\right] \equiv \int_0^t dt'\mathcal{L}[\dot{x},x], \tag{55}$$
where $\mathcal{L}$ is the *Lagrangian* of the system. In classical mechanics the time-integral of the Lagrangian is known as the *action*
$$S = \int_0^t dt'\mathcal{L}. \tag{56}$$
- The integral over $x_1, \ldots x_{N-1}$ becomes a *functional integral*, also known as a *path integral*, over all paths $x(t')$ that start at x_0 at time $t' = 0$ and end at x_N at time $t' = t$.
- The prefactor in (53) gives rise to an overall (infinite) normalization and we will denote it by $\mathcal{N}$.

These considerations lead us to express the propagator as the following *formal expression*

$$\boxed{\langle x_N|U(t;0)|x_0\rangle = \mathcal{N}\int\mathcal{D}x(t')\; e^{\frac{i}{\hbar}S[x(t')]}.} \tag{57}$$

What is in fact meant by (57) is the limit of the discretized expression (53). The ultimate utility of (57) is that it provides a compact notation, that on the one hand will allow us to manipulate functional integrals, and on the other hand provides a nice, intuitive interpretation.

[2] We use a normalization $\langle p|k\rangle = 2\pi\hbar\delta(p-k)$, so that $\mathbf{1} = \int\frac{dp}{2\pi\hbar}|p\rangle\langle p|$.

2.2 Quantum Mechanics à la Feynman

Feynman's formulation of Quantum Mechanics is based on the single postulate that the probability amplitude for propagation from a position x_0 to a position x_N is obtained by summing over all possible paths connecting x_0 and x_N, where each path is *weighted* by a phase factor $\exp\left(\frac{i}{\hbar}S\right)$, where S is the *classical action* of the path. This provides a new way of thinking about QM!

2.3 Classical Limit and Stationary Phase Approximation

An important feature of (57) is that it gives us a nice way of thinking about the classical limit "$\hbar \to 0$" (more precisely in the limit when the dimensions, masses, times etc are so large that the action is huge compared to $\hbar$). To see what happens in this limit let us first consider the simpler case of an ordinary integral

$$g(a) = \int_{-\infty}^{\infty} dt\ h_1(t) e^{iah_2(t)}, \tag{58}$$

when we take the real parameter a to infinity. In this case the integrand will oscillate wildly as a function of t because the phase of $\exp\left(iah_2(t)\right)$ will vary rapidly. The dominant contribution will arise from the points where the phase changes slowly, which are the stationary points

$$h_2'(t) = 0. \tag{59}$$

The integral can then be approximated by expanding around the stationary points. Assuming that there is a single stationary point at t_0

$$g(a \gg 1) \approx \int_{-\infty}^{\infty} dt \left[h_1(t_0) + (t-t_0)h_1'(t_0) + \ldots\right] e^{iah_2(t_0) + i\frac{ah_2''(t_0)}{2}(t-t_0)^2}, \tag{60}$$

Changing integration variables to $t' = t - t_0$ (and giving a a small imaginary part to make the integral converge at infinity) as obtain a Gaussian integral that we can take using (23)

$$g(a \gg 1) \approx \sqrt{\frac{2\pi i}{ah_2''(t_0)}} h_1(t_0) e^{iah_2(t_0)}\ . \tag{61}$$

Subleading contributions can be evaluated by taking higher order contributions in the Taylor expansions into account. If we have several stationary points we sum over their contributions. The method we have just discussed is known as *stationary phase approximation*.

The generalization to path integrals is now clear: in the limit $\hbar \to 0$ the path integral is dominated by the vicinity of the *stationary points* of the action S

$$\frac{\delta S}{\delta x(t')} = 0. \tag{62}$$

The condition (62) precisely defines the *classical trajectories* $x(t')$!

2.4 The Propagator for Free Particles

We now wish to calculate the functional integral (57) for a free particle, i.e.

$$V(x) = 0. \tag{63}$$

Going back to the explicit expression (53) we have

$$\langle x_N|U(t;0)|x_0\rangle \ = \ \lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar\epsilon}\right]^{\frac{N}{2}} \int dx_1 \ldots dx_{N-1} \exp\left(\frac{i\epsilon}{\hbar}\sum_{n=0}^{N-1}\frac{m}{2}\left(\frac{x_{n+1}-x_n}{\epsilon}\right)^2\right)\ . \tag{64}$$

It is useful to change integration variables to

$$y_j = x_j - x_N \ , \quad j = 1, \ldots, N-1, \tag{65}$$

which leads to an expression

$$\langle x_N|U(t;0)|x_0\rangle \ = \ \lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar\epsilon}\right]^{\frac{N}{2}}\int d\mathbf{y}\exp\left(-\frac{1}{2}\mathbf{y}^T\mathbf{A}\mathbf{y}+\mathbf{J}^T\cdot\mathbf{y}\right)e^{\frac{im}{2\hbar\epsilon}(x_0-x_N)^2}. \tag{66}$$

Here

$$\mathbf{J}^T = \big(\frac{im}{\hbar\epsilon}(x_N - x_0), 0, \ldots, 0\big), \tag{67}$$

and $\mathbf{A}$ is a $(N-1)\times(N-1)$ matrix with elements

$$A_{jk} = \frac{-im}{\epsilon\hbar}\left[2\delta_{j,k} - \delta_{j,k+1} - \delta_{j,k-1}\right]. \tag{68}$$

For a given N (66) is a multidimensional Gaussian integral and can be carried out using (34)

$$\langle x_N|U(t;0)|x_0\rangle \ = \ \lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar\epsilon}\right]^{\frac{N}{2}}(2\pi)^{\frac{N-1}{2}}\left[\det(\mathbf{A})\right]^{-\frac{1}{2}}\exp\left(\frac{1}{2}\mathbf{J}^T\mathbf{A}^{-1}\mathbf{J}\right)e^{\frac{im}{2\hbar\epsilon}(x_0-x_N)^2}. \tag{69}$$

The matrix $\mathbf{A}$ is related to the one dimensional lattice Laplacian, see below. Given the eigenvalues and eigenvectors worked out below we can calculate the determinant and inverse of $\mathbf{A}$ (homework problem). Substituting the results into (69) gives

$$\boxed{\langle x_N|U(t;0)|x_0\rangle = \sqrt{\tfrac{m}{2\pi i\hbar t}}e^{\frac{im}{2\hbar t}(x_0-x_N)^2}.} \tag{70}$$

For a free particle we can evaluate the propagator directly in a much simpler way.

$$\begin{aligned}\langle x_N|U(t;0)|x_0\rangle \ &= \ \int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\langle x_N|e^{-i\frac{\hat{p}^2t}{2m\hbar}}|p\rangle\langle p|x_0\rangle = \int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}e^{-i\frac{p^2t}{2m\hbar}-i\frac{p(x_0-x_N)}{\hbar}} \\ &= \ \sqrt{\frac{m}{2\pi i\hbar t}}e^{\frac{im}{2\hbar t}(x_0-x_N)^2}.\end{aligned} \tag{71}$$

Exercise 2

The matrix A is related to the one dimensional *Lattice Laplacian*. Consider functions of a variable $z_0 \le z \le z_N$ with "hard-wall boundary conditions"

$$f(z_0) = f(z_N) = 0. \tag{72}$$

The Laplace operator D acts on these functions as

$$Df \equiv \frac{d^2f(z)}{dz^2}. \tag{73}$$

Discretizing the variable z by introducing $N-1$ points

$$z_n = z_0 + na_0 \ , \quad n = 1, \ldots, N-1 \tag{74}$$

where $a_0 = (z_N - z_0)/N$ is a "lattice spacing", maps the function $f(z)$ to a $N-1$ dimensional vector

$$f(z) \to \mathbf{f} = (f(z_1), \ldots, f(z_{N-1})). \tag{75}$$

Recalling that

$$\frac{d^2f}{dz^2}(z) = \lim_{a_0\to 0} \frac{f(z+a_0)+f(z-a_0)-2f(z)}{a_0^2} \ , \tag{76}$$

we conclude that the Lapacian is discretized as follows

$$Df \to a_0^{-2}\mathbf{\Delta f} \ , \tag{77}$$

where

$$\Delta_{jk} = \delta_{j,k+1} + \delta_{j,k-1} - 2\delta_{j,k}. \tag{78}$$

Our matrix $\mathbf{A}$ is equal to $\frac{im}{\epsilon\hbar}\mathbf{\Delta}$. The eigenvalue equation

$$\mathbf{\Delta a}_n = \lambda_n \mathbf{a}_n, \quad n = 1, \ldots, N-1 \tag{79}$$

gives rise to a recurrence relation for the components $\mathbf{a}_{n,j}$ of $\mathbf{a}_n$

$$a_{n,j+1} + a_{n,j-1} - (2+\lambda_n)a_{n,j} = 0. \tag{80}$$

The boundary conditions $a_{n,N} = a_{n,0} = 0$ suggest the ansatz

$$a_{n,j} = C_n \sin\left(\frac{\pi n j}{N}\right). \tag{81}$$

Substituting this in to (80) gives

$$\lambda_n = 2\cos\left(\frac{\pi n}{N}\right) - 2 \ , \quad n = 1, \ldots, N-1. \tag{82}$$

The normalized eigenvectors of $\mathbf{\Delta}$ are

$$\mathbf{a}_n = \frac{1}{\sqrt{\sum_{j=1}^{N-1}\sin^2\left(\frac{\pi n j}{N}\right)}} \begin{pmatrix} \sin\left(\frac{\pi n}{N}\right) \\ \sin\left(\frac{2\pi n}{N}\right) \\ \vdots \\ \sin\left(\frac{\pi(N-1)n}{N}\right). \end{pmatrix} = \sqrt{\frac{2}{N}} \begin{pmatrix} \sin\left(\frac{\pi n}{N}\right) \\ \sin\left(\frac{2\pi n}{N}\right) \\ \vdots \\ \sin\left(\frac{\pi(N-1)n}{N}\right). \end{pmatrix} \tag{83}$$

2.5 Homework Questions 3-5

Question 3. Consider a free quantum mechanical particle moving in one dimension. The Hamiltonian is

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} \ . \tag{84}$$

We have shown in the lecture that the propagator can be represented in the form

$$\langle x_N|e^{-\frac{i}{\hbar}tH}|x_0\rangle \ = \ \lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar\epsilon}\right]^{\frac{N}{2}} \int dx_1 \ldots dx_{N-1} \exp\left(\frac{i\epsilon}{\hbar}\sum_{n=0}^{N-1}\frac{m}{2}\left(\frac{x_{n+1}-x_n}{\epsilon}\right)^2\right) \ . \tag{85}$$

a) Change variables from x_j to $y_j = x_j - x_N$ to bring it to the form

$$\langle x_N|e^{-\frac{i}{\hbar}tH}|x_0\rangle \ = \ \lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar\epsilon}\right]^{\frac{N}{2}} \int d\mathbf{y} \exp\left(-\frac{1}{2}\mathbf{y}^T\mathbf{A}\mathbf{y} + \mathbf{J}^T\cdot\mathbf{y}\right) e^{\frac{im}{2\hbar\epsilon}(x_0-x_N)^2}. \tag{86}$$

Give expressions for $\mathbf{J}$ and $\mathbf{A}$.
b) Carry out the integrals over y_j to get an expression for the propagator in terms of $\mathbf{A}$ and $\mathbf{J}$.

c) Work out the eigenvalues λ_n and eigenvectors $\mathbf{a}_n$ of the matrix $\mathbf{A}$. You may find helpful hints in the lecture notes.
d) What is $\det(\mathbf{A})$? A useful identity you may use is

$$\prod_{j=1}^{N-1} 2\sin(\pi j/2N) = \sqrt{N}. \tag{87}$$

Now work out $\mathbf{J}^T\mathbf{A}^{-1}\mathbf{J}$ by working in the eigenbasis of $\mathbf{A}^{-1}$ (Hint: write this as $\mathbf{J}^T\mathbf{A}^{-1}\mathbf{J} = \mathbf{J}^T\mathbf{O}^T\mathbf{O}\mathbf{A}^{-1}\mathbf{O}^T\mathbf{O}\mathbf{J}$, where $\mathbf{O}^T\mathbf{O} = 1$ and $\mathbf{O}\mathbf{A}^{-1}\mathbf{O}^T$ is a diagonal matrix you have already calculated above.). A useful identity you may use is

$$\sum_{j=1}^{N-1} \cos^2(\pi j/2N) = \frac{N-1}{2}. \tag{88}$$

e) Use the result you have obtained to write an explicit expression for the propagator.

Question 4. Denote the propagator by

$$K(t,x;t'x') = \langle x|e^{-\frac{i}{\hbar}H(t-t')}|x'\rangle. \tag{89}$$

Show that the wave function $\psi(t,x) = \langle x|\Psi(t)\rangle$, where $|\Psi(t)\rangle$ is a solution to the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle, \tag{90}$$

fulfils the integral equation

$$\psi(t,x) = \int_{-\infty}^{\infty} dx'\ K(t,x;t'x')\ \psi(t',x'). \tag{91}$$

Question 5. Diffraction through a slit. A free particle starting at $x = 0$ when $t = 0$ is determined to pass between $x_0 - b$ and $x_0 + b$ at time T. We wish to calculate the probabilty of finding the particle at position x at time $t = T + \tau$.
a) Argue on the basis of Qu 5. that the (un-normalized) wave function can be written in the form

$$\psi(T+\tau,x) = \int_{-b}^{b} dy\ K(T+\tau,x;T,x_0+y)\ K(T,x_0+y;0,0)\ , \tag{92}$$

where

$$K(t,x;t'x') = \langle x|e^{-\frac{i}{\hbar}H(t-t')}|x'\rangle. \tag{93}$$

b) Using that the propagation for $0 \le t < T$ and $T \le t < T+\tau$ is that of a free particle, obtain an explicit integral representation for the wave function.
c) Show that the wave function can be expressed in terms of the *Fresnel integrals*

$$C(x) = \int_0^x dy\ \cos(\pi y^2/2)\ , \quad S(x) = \int_0^x dy\ \sin(\pi y^2/2)\ . \tag{94}$$

Hint: make a substitution $z = \alpha y + \beta$ with suitably chosen α and β.

Derive an expression for the ratio $P(T+\tau,x)/P(T+\tau,x_0)$, where $P(T+\tau,x)dx$ is the probability of finding the particle in the interval $[x, x+dx]$ at time $T+\tau$.
d)* If you can get hold of *Mathematica* (the default assumption is that you will not), plot the result as a function of the dimensionless parameter $x/[b(1+\tau/T)]$ for $x_0 = 0$ and different values of the ratio

$$\gamma = \frac{mb^2(1+\tau/T)}{\hbar\tau}\ . \tag{95}$$

Discuss your findings.

3 Path Integrals in Quantum Statistical Mechanics

Path integrals can also be used to describe quantum systems at finite temperatures. To see how this works we now consider a quantum mechanical particle coupled to a heat bath at a temperature T. An important quantity in Statistical Mechanics is the *partition function*

$$Z(\beta) = \mathrm{Tr}\left[e^{-\beta H}\right], \tag{96}$$

where H is the Hamiltonian of the system, Tr denotes the trace over the Hilbert space of quantum mechanical states, and

$$\beta = \frac{1}{k_B T}. \tag{97}$$

Ensemble averages of the quantum mechanical observable $\mathcal{O}$ are given by

$$\langle \mathcal{O} \rangle_\beta = \frac{1}{Z(\beta)} \mathrm{Tr}\left[e^{-\beta H}\mathcal{O}\right]. \tag{98}$$

Taking the trace over a basis of eigenstates of H with $H|n\rangle = E_n|n\rangle$ gives

$$\begin{aligned} \langle \mathcal{O} \rangle_\beta &= \frac{1}{Z(\beta)} \sum_n e^{-\beta E_n} \langle n|\mathcal{O}|n\rangle \ , \\ Z(\beta) &= \sum_n e^{-\beta E_n}. \end{aligned} \tag{99}$$

Assuming that the ground state of H is non-degenerate we have

$$\lim_{T\to 0} \langle \mathcal{O} \rangle_\beta = \langle 0|\mathcal{O}|0\rangle \ , \tag{100}$$

where $|0\rangle$ is the ground state of the system. Let us consider a QM particle with Hamiltonian

$$H = \frac{\hat{p}^2}{2m} + V(\hat{x}), \tag{101}$$

coupled to a heat bath at temperature T. The partition function can be written in a basis of position eigenstates

$$Z(\beta) = \int dx \langle x|e^{-\beta H}|x\rangle = \int dx \int dx' \ \langle x|x'\rangle \ \langle x'|e^{-\beta H}|x\rangle. \tag{102}$$

Here

$$\langle x'|e^{-\beta H}|x\rangle \tag{103}$$

is very similar to the propagator

$$\langle x'|e^{-\frac{i(t-t_0)}{\hbar}H}|x\rangle. \tag{104}$$

Formally (103) can be viewed as the propagator in *imaginary time* $\tau = it$, where we consider propagation from $\tau = 0$ to $\tau = \beta\hbar$. Using this interpretation we can follow through precisely the same steps as before and obtain

$$\langle x_N|e^{-\beta H}|x_0\rangle = \lim_{N\to\infty} \left[\frac{m}{2\pi\hbar\epsilon}\right]^{\frac{N}{2}} \int dx_1 \dots dx_{N-1} \exp\left(-\frac{\epsilon}{\hbar}\sum_{n=0}^{N-1} \frac{m}{2}\left(\frac{x_{n+1}-x_n}{\epsilon}\right)^2 + V(x_n)\right), \tag{105}$$

where now

$$\epsilon = \frac{\hbar\beta}{N}. \tag{106}$$

We again can interpret this in terms of a sum over paths $x(\tau)$ with

$$x(\tau_n) = x_n \ , \quad \tau_n = n\epsilon. \tag{107}$$

Going over to a continuum description we arrive at an imaginary-time functional integral

$$\boxed{\langle x_N|e^{-\beta H}|x_0\rangle = \mathcal{N}\int \mathcal{D}x(\tau)\ e^{-\frac{1}{\hbar}S_E[x(\tau)]},} \tag{108}$$

where S_E is called *Euclidean action*

$$S_E[x(\tau)] = \int_0^{\hbar\beta} d\tau \left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 + V(x)\right], \tag{109}$$

and the path integral is over all paths that start at x_0 and end at x_N. Substituting (108) into the expression for the partition function we find that

$$\boxed{Z(\beta) = \mathcal{N}\int \mathcal{D}x(\tau)\ e^{-\frac{1}{\hbar}S_E[x(\tau)]},} \tag{110}$$

where we integrate over all *periodic paths*

$$x(\hbar\beta) = x(0). \tag{111}$$

The restriction to periodic paths arises because $Z(\beta)$ is a trace. Please note that the notation $\int \mathcal{D}x(\tau)$ means very different things in (108) and (110), as a result of involving very different classes of paths that are "integrated over". In the first case the path integral is over all paths from x_0 to x_N, in the latter over all periodic paths starting an an arbitrary position. In practice it is always clear from the context what paths are involved, and hence this ambiguous notation should not cause any confusion.

3.1 Harmonic Oscillator at $T > 0$: a first encounter with Generating Functionals

We now consider the simplest case of the potential $V(x)$, the harmonic oscillator

$$H = \frac{\hat{p}^2}{2m} + \frac{\kappa}{2}\hat{x}^2. \tag{112}$$

The physical quantities we want to work out are the averages of powers of the position operator

$$\langle \hat{x}^n\rangle_\beta = \frac{\int dx\ \langle x|e^{-\beta H}\hat{x}^n|x\rangle}{\int dx\langle x|e^{-\beta H}|x\rangle} = \frac{\int dx\ x^n\ \langle x|e^{-\beta H}|x\rangle}{\int dx\langle x|e^{-\beta H}|x\rangle}. \tag{113}$$

If we know all these moments, we can work out the probability distribution for a position measurement giving a particular result x. At zero temperature this is just given by the absolute value squared of the ground state wave function. The coupling to the heat bath will generate "excitations" of the harmonic oscillator and thus affect this probability distribution. We have

$$\langle x|e^{-\beta H}|x\rangle = \mathcal{N}\int \mathcal{D}x(\tau)\ e^{-\frac{1}{\hbar}\int_0^{\hbar\beta} d\tau\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 + \frac{\kappa}{2}x^2\right]}, \tag{114}$$

where the path integral is over all paths with $x(0) = x(\beta\hbar)$. Integrating by parts we can write the action as

$$-\frac{1}{\hbar}S_E[x(\tau)] = -\frac{1}{\hbar}\int_0^{\hbar\beta} d\tau \left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 + \frac{\kappa}{2}x^2\right] = -\frac{1}{2}\int_0^{\hbar\beta} d\tau\ x(\tau)\hat{D}x(\tau) - \frac{m}{2\hbar}x(\tau)\dot{x}(\tau)\Big|_0^{\hbar\beta}, \tag{115}$$

where

$$\hat{D} = -\frac{m}{\hbar}\frac{d^2}{d\tau^2} + \frac{\kappa}{\hbar}. \tag{116}$$

The contributions from the integration boundaries in (115) don't play a role in the logic underlying the following steps leading up to (124) and work out in precisely the same way as the "bulk" contributions. In order to show that we're not dropping anything important we'll keep track of them anyway. We now define the *generating functional*

$$W[J] \equiv \mathcal{N}\int \mathcal{D}x(\tau)\; e^{-\frac{1}{\hbar}S_E[x(\tau)]+\int_0^{\hbar\beta} d\tau\; J(\tau)x(\tau)}. \tag{117}$$

Here the functions $J(\tau)$ are called *sources*. The point of the definition (117) is that we can obtain $\langle \hat{x}^n\rangle_\beta$ by taking functional derivatives

$$\langle \hat{x}^n\rangle_\beta = \frac{1}{W[0]}\frac{\delta}{\delta J(0)}\cdots\frac{\delta}{\delta J(0)}\bigg|_{J=0} W[J]. \tag{118}$$

We now could go ahead and calculate the generating functional by going back to the definition of the the path integral in terms of a multidimensional Gaussian integral. In practice we manipulate the path integral itself as follows. Apart from the contribution from the integration boundaries the structure of (115) is completely analogous to the one we encountered for Gaussian integrals, cf (31). The change of variables (32) suggests that we should shift our "integration variables" by a term involving the inverse of the integral operator $\hat{D}$. The latter corresponds to the Green's function defined by

$$\hat{D}_\tau G(\tau-\tau') = \delta(\tau-\tau')\;, \quad G(0) = G(\beta\hbar). \tag{119}$$

We then change variables in the path integral in order to "complete the square"

$$y(\tau) = x(\tau) - \int d\tau'\; G(\tau-\tau')J(\tau'). \tag{120}$$

Under this change of variables we have

$$\begin{aligned}\int d\tau\; x(\tau)\hat{D}_\tau x(\tau) - 2\int d\tau\; x(\tau)J(\tau) \;&=\; \int d\tau\; y(\tau)\hat{D}_\tau y(\tau) - \int d\tau d\tau'\; J(\tau)G(\tau-\tau')J(\tau') \\ &\quad +\text{boundary terms.}\end{aligned} \tag{121}$$

Exercise 3

Verify that

$$\begin{aligned}\int d\tau\; y(\tau)\hat{D}_\tau y(\tau) \;&=\; \int d\tau\; x(\tau)\hat{D}_\tau x(\tau) + \int d\tau d\tau' d\tau''\; G(\tau-\tau')J(\tau')\hat{D}_\tau G(\tau-\tau'')J(\tau'') \\ &\quad - \int d\tau d\tau' \Big[x(\tau)\hat{D}_\tau G(\tau-\tau')J(\tau') + G(\tau-\tau')J(\tau')\hat{D}_\tau x(\tau)\Big] \\ &= \int d\tau\; x(\tau)\hat{D}_\tau x(\tau) + \int d\tau d\tau'\; G(\tau-\tau')J(\tau')J(\tau) - 2\int d\tau\; x(\tau)J(\tau) \\ &\quad + \frac{m}{\hbar}x(\tau)\dot{x}(\tau)\Big|_0^{\hbar\beta} - \frac{m}{\hbar}y(\tau)\dot{y}(\tau)\Big|_0^{\hbar\beta},\end{aligned} \tag{122}$$

In the last step you need to use (119) and integrate by parts twice to simplify the last term in the second line.

Putting everything together we arrive at

$$-\frac{1}{\hbar}S_E[x(\tau)] + \int_0^{\hbar\beta} d\tau\; J(\tau)x(\tau) = -\frac{1}{\hbar}S_E[y(\tau)] + \frac{1}{2}\int_0^{\hbar\beta} d\tau d\tau' J(\tau)G(\tau-\tau')J(\tau'). \tag{123}$$

On the other hand, the Jacobian of the change of variables (120) is 1 as we are shifting all paths by the same constant (you can show this directly by going back to the definition of the path integral in terms of multiple Gaussian integrals). Hence we have $\mathcal{D}y(\tau) = \mathcal{D}x(\tau)$ and our generating functional becomes

$$\boxed{W[J] = W[0]\; e^{\frac{1}{2}\int d\tau d\tau'\; J(\tau)G(\tau-\tau')J(\tau')}.} \tag{124}$$

Now we are ready to calculate (118). The average position is zero

$$\langle \hat{x} \rangle_\beta = \frac{1}{W[0]} \frac{\delta}{\delta J(0)} \bigg|_{J=0} W[J] = \frac{1}{2W[0]} \int d\tau d\tau' \left[\delta(\tau) G(\tau - \tau') J(\tau') + J(\tau) G(\tau - \tau') \delta(\tau') \right] W[J] \bigg|_{J=0} = 0. \tag{125}$$

Here we have used that

$$\frac{\delta J(\tau)}{\delta J(\tau')} = \delta(\tau - \tau'). \tag{126}$$

The expression (125) vanishes, because we have a "left over" J and obtain zero when setting all sources to zero in the end of the calculation. By the same mechanism we have

$$\langle \hat{x}^{2n+1} \rangle_\beta = 0. \tag{127}$$

Next we turn to

$$\begin{aligned} \langle \hat{x}^2 \rangle_\beta &= \frac{1}{W[0]} \frac{\delta}{\delta J(0)} \frac{\delta}{\delta J(0)} \bigg|_{J=0} W[J] \\ &= \frac{1}{W[0]} \frac{\delta}{\delta J(0)} \bigg|_{J=0} \frac{1}{2} \int d\tau d\tau' \left[\delta(\tau) G(\tau - \tau') J(\tau') + J(\tau) G(\tau - \tau') \delta(\tau') \right] W[J] = G(0). \end{aligned} \tag{128}$$

So the mean square deviation of the oscillator's position is equal to the Green's function evaluated at zero.

1: Aside: Propagators vs. Green's Function vs. Correlation functions

A frequent point of confusion is between propagators, correlation functions, and Green's functions. Often, QFT textbooks use these terms interchangeably, and in particular refer to Green's functions of classical equations of motion (such as the ones defined above) simply as 'propagators'. This is true in some sense" given the value of a field at some time, the Green's function tells you how to determine the field at some other time, by 'propagating' it forward in time. (In our QM examples, x is a *field*, since QM can be thought of as a 0+1D QFT; one part of the confusion in translating between QM and QFT is that the position x in QM is a *field*, whereas in a QFT x labels a field.) Viewed in the language of path integrals in imaginary time, this is the same as a correlation function. Given certain analytic properties of how to treat poles, the imaginary-time and real-time expressions can be linked by analytic continuation (a trick known as 'Wick rotation').
However, what we have called the propagator

$$K(x,t;x',t') = \langle x|\hat{U}(t,t')|x'\rangle = \langle x|e^{-\frac{i}{\hbar}H(t-t')}|x'\rangle \tag{129}$$

is a little bit different: it is really the representation of the unitary time evolution operator (as encoded by the Sc hrödinger equation) in the position basis. Crucially, it is *not* a Green's function of the Schrödinger equation. To see this, consider the example of the free particle propagator (70)

$$K(x,t;x',t') = \sqrt{\frac{m}{2\pi i\hbar|t-t'|}} e^{\frac{im}{2\hbar|t-t'|}(x-x')^2}. \tag{130}$$

One can readily check that

$$\left(i\hbar\frac{\partial}{\partial t} - H_x\right) K(x,t;x',t') = \left(i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\right) K(x,t;x',t') = 0 \tag{131}$$

where we use H_x to remind ourselves that we are working in the position basis. In other words, K is a *homogeneous* solution to the differential equation. A Green's function $G_S(x,t;x',t')$ for the Schrödinger equation should instead satisfy

$$\left(i\hbar\frac{\partial}{\partial t} - H_x\right) G_S(x,t;x't') = \delta(t-t')\delta(x-x') \tag{132}$$

– i.e., it is a solution with a δ-function source term. Consider now the object $\frac{1}{i\hbar}\Theta(t-t')K(x,t;x';t')$. Evidently, we have

$$\left(i\hbar\frac{\partial}{\partial t} - H_x\right) \Theta(t-t')K(x,t;x';t') = \delta(t-t')K(x,t;x';t') \tag{133}$$

where we have used that K is a homogenous solution of $i\hbar\partial_t - H_x$ and the fact that the derivative of a step function is a δ-function. Examining the limiting behaviour of K for $|t-t'| \to 0$ (the only part of it that is relevant to the RHS of the expresssion above, due to the $\delta(t-t')$ prefactor) we see that it reduces to a spatial δ-function: in other words, $G_S(x,t;x',t') \equiv \Theta(t-t')K(x,t;x';t')$ is a Green's function for the Schrödinger operator $i\hbar\partial_t - H_x$.

However, notice that this is not especially useful to us: a Green's function solves a linear differential equation with driving terms, but this is not what we want to do with regard to the Schrödinger equation. (In contrast, we frequently *do* want to do this with respect to classical equations of motion.) Rather than a problem with a driving term, we want to solve an *initial value* problem: *given* $\psi(x',0)$, what is $\psi(x,t)$? The answer is

$$\psi(x,t) = \int_\infty^\infty dx' K(x,t;x',0)\psi(x',0). \tag{134}$$

To see this, we can use the property that $(i\hbar\partial_t - H_x)K(x;t,x',t') = 0$, which means $\psi(x,t)$ satisfies the time-dependent Schrödinger equation, and the limiting property $\lim_{t\to t'} K(x,t;x',t') = \delta(x-x')$ to confirm that $\lim_{t\to 0}\psi(x,t) = \psi(x,0)$ so that $\psi(x,t)$ satisfies the correct initial value condition.

We could have also *defined* what we have called the propagator directly using the language of differential equations: K is simply the solution to the Schrödinger equation $(i\hbar\partial_t - H_x)K(x;t,x',t') = 0$ (a homogenous PDE) with the time-Dirichlet boundary condition$K(x,t;x't') = \delta(x-x')$. In the PDE literature, such a solution is usually termed the *kernel*. (134) makes it clear what the kernel propagates: a quantum state (written as a wavefunction) forward in time. In other words, what we have called the propagator is the kernel of the Schrödinger PDE, not its Green's function.

Often in field theory, one works with various types of Green's functions (things that solve an inhomogenous PDE with δ-function sources) but they are naturally written in terms of step functions and kernels. The kernels themselves are usually related to correlators. Therefore, the distinction is often ignored, but this leads to a perennial source of confusion. The key is to understand what equation you are solving, and what you are using the quantity in question for – the rest is just names, and to paraphrase the father of a famous physicist, there is a difference between knowing the name of something and knowing something!

3.1.1 Imaginary Time Green's Function of the Harmonic Oscillator

To determine $G(\tau)$ we need to solve the differential equation (119). As $G(0) = G(\beta\hbar)$ we are dealing with a periodic function and therefore may employ a Fourier expansion

$$G(\tau) = \frac{1}{\sqrt{\beta\hbar}} \sum_{n=-\infty}^{\infty} g_n e^{i\omega_n \tau}, \tag{135}$$

where the *Matsubara frequencies* ω_n are

$$\omega_n = \frac{2\pi n}{\beta\hbar}. \tag{136}$$

Substituting this into the differential equation gives

$$\hat{D}G(\tau) = \frac{1}{\sqrt{\beta\hbar}} \sum_{n=-\infty}^{\infty} g_n e^{i\omega_n \tau} \left[\frac{m\omega_n^2}{\hbar} + \frac{\kappa}{\hbar}\right] = \delta(\tau). \tag{137}$$

Taking the integral $\int_0^{\hbar\beta} d\tau e^{-i\omega_k \tau}$ on both sides fixes the Fourier coefficients and we obtain

$$G(\tau) = \frac{1}{\beta\kappa} \sum_{n=-\infty}^{\infty} \frac{\omega^2}{\omega^2 + \omega_n^2} e^{i\omega_n \tau}\ , \tag{138}$$

where $\omega = \sqrt{\kappa/m}$. Using contour integration techniques this can be rewritten as

$$G(\tau) = \frac{\hbar\omega}{2\kappa} \left[\frac{e^{\omega|\tau|}}{e^{\hbar\beta\omega} - 1} + \frac{e^{-\omega|\tau|}}{1 - e^{-\hbar\beta\omega}}\right]. \tag{139}$$

Setting $\tau = 0$ gives

$$G(0) = \frac{\hbar\omega}{2\kappa \tanh\left(\beta\hbar\omega/2\right)} = \frac{\hbar\omega}{\kappa} \left[\frac{1}{e^{\beta\hbar\omega} - 1} + \frac{1}{2}\right]. \tag{140}$$

We can relate this result to things we already know: using equipartition

$$\langle H\rangle_\beta = \langle T\rangle_\beta + \langle V\rangle_\beta = 2\langle V\rangle_\beta = \kappa\langle \hat{x}^2\rangle_\beta = \kappa G(0), \tag{141}$$

we find that the average energy of the oscillator at temperature T is

$$\langle H\rangle_\beta = \hbar\omega \left[\frac{1}{e^{\beta\hbar\omega} - 1} + \frac{1}{2}\right]. \tag{142}$$

Recalling that the Hamiltonian of the harmonic oscillator can be expressed as

$$H = \hbar\omega\big(\hat{n} + \frac{1}{2}\big), \tag{143}$$

where $\hat{n} = a^\dagger a$ is the number operator, we recover the Bose-Einstein distribution

$$\langle \hat{n}\rangle_\beta = \frac{1}{e^{\beta\hbar\omega} - 1}. \tag{144}$$

2: Aside: Contour integration for imaginary time Green's function.

When working with finite-temperature Green's functions, it is convenient to work in (Matsubara) frequency space. This means that to go back to imaginary time, one frequently needs to compute

Matsubara summations of the form

$$S = \frac{1}{\beta} \sum_{n-\infty}^{\infty} g(i\omega_n) \tag{145}$$

where $\omega_n = \frac{2\pi n}{\beta}$ and g is some (for now unspecified) function. The key to this is to use contour integration. Recall that the basic result behind the utility of contour integration is that a contour integral is completely determined by the values of its integrand at the poles, i.e.

$$\oint_{\mathcal{C}} f(z)dz = \frac{1}{2\pi i} \sum_{\text{poles } z_n \text{ of } f \in \mathcal{C}} \text{Res}[f(z)]|_{z=z_n}\,, \tag{146}$$

where we use the standard convention that contour integrals are positive when the contour is traversed in an anticlockwise sense. The strategy is as follows: first, we find a function $h(z)$ that has the following properties:

1. the only singularities of $h(z)$ are simple poles along the imaginary axis at the Matsubara frequencies, i.e. at $z_n = i\frac{2\pi n}{\beta\hbar}$
2. all residues of $h(z)$ are identical: $\text{Res}[h(z_n)] = \zeta$;
3. the function $f(z) = g(z)h(z)$ has the property that $|g(z)h(z)| \to 0$ faster than $1/|z|$ as $|z| \to \infty$.

We will also assume that any singularities of $g(z)$ lie away from the imaginary axis. (Note that g more or less must have singularities:a function that is bounded and analytic (as most physical correlators will be) in the entire complex plane is constant, i.e. every 'interesting' function will have singularities!) To write our sum S as a contour integral, we integrate the product $f(z)$ over a contour $\mathcal{C}_a$ that runs parallel to the imaginary axis from $-i\infty$ to $i\infty$ and 'encircles' it (Fig 2a; we can also view this as hugging the imaginary axis and making little detours around the poles – the two are the same because of the properties of analytic functions.) We have

$$\mathcal{I} = \oint_{\mathcal{C}_a} f(z) = 2\pi i \sum_n \text{Res}[g(z)h(z)]_{z=i\omega_n} = 2\pi i \sum_n g(i\omega_n)\zeta = 2\pi i\zeta\beta S \tag{147}$$

where we have successively used (i) the residue theorem, (ii) our assumption that g has no poles on the imaginary axis and the choice of poles of h; and (iii) the fact that all the residues of h are identical and equal to ζ.
Now, we evaluate $\mathcal{I}$ in a different way. Recall that we can deform the contour in the complex plane, as long as we don't cross any singularities; the value of the contour integral is unchanged. Using this freedom, we deform $\mathcal{C}_a$ into $\mathcal{C}_b$. We push $\mathcal{C}_b$ out as far as it will go, in particular pulling it far from any singularities of h; by 'pinching' around singularities of g (whether poles or branch cuts), we can deform $\mathcal{C}_b$ all the way until it consists just of the 'pinches' and an integral around a large circle of radius $R \to \infty$ (Fig.2b). This gives us another expression for $\mathcal{I}$

$$\mathcal{I} = \mathcal{I}_R + \sum_{\mathcal{C}_s \text{ encircling singularities of g}} \oint_{\mathcal{C}_s} f(z)dz. \tag{148}$$

We now use the property (3) related to our choice of h above: this allows us to take $\mathcal{I}_R = 0$. Thus, comparing (147) and (148), we have

$$S = \frac{1}{2\pi i\beta\zeta} \sum_{\mathcal{C}_s \text{ encircling singularities of } g} \oint_{\mathcal{C}_s} g(z)h(z)dz \tag{149}$$

Crucially, the RHS only depends on the analytic structure of g, and we can evaluate the necessary integrals using standard complex analysis. Usually, there will only be a finite number of siingularities,

meaning that it is much easier to do this than perform a complicated infinite sum. In the special case where all the singularities of g are simple poles, we can use the residue theorem on the RHS, to find

$$S = -\frac{1}{\beta\zeta} \sum_{\text{poles } z_k \text{ of } g} h(z_k)\text{Res}[g(z_k)], \tag{150}$$

where the $-$ comes from the fact that the poles of g are encircled in the clockwise sense. Branch cuts can be handled with slightly more effort.
Evidently, the above approach relies on our ability to choose a simple form for h. There are two especially convenient choices:

$$h_+(z) = \frac{1}{e^{\beta\hbar z} - 1} \quad \text{or} \quad h_-(z) = \frac{1}{1 - e^{-\beta\hbar z}}. \tag{151}$$

Notice that $h_\pm(z)$ are respectively exponentially vanishing for $\text{Re}(z) \to \pm\infty$, and have residue $\zeta = \frac{1}{\beta}$ at each of their poles at $z = i\omega_n$. Our choice of which of $h_\pm$ to use is determined by the choice of g in (145). For the example of the Green's function, $f(z)$ is only convergent as demanded by property (3) above for one choice of $h = h_\pm$, depending on the sign of τ, but there is always such a choice. You are encouraged to try this out in order to reproduce the Green's function above.

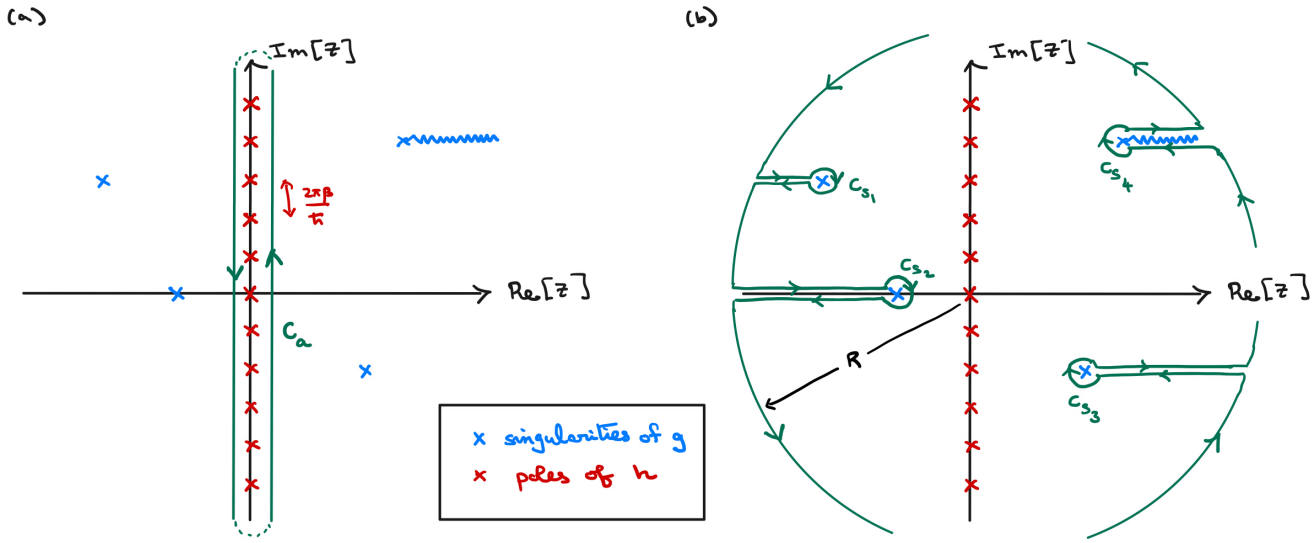


Figure 2: Contours used to evaluate Matsubara sums.

3.2 Homework Question 6

Question 6. In this question the objective is to evaluate the Feynman path integral in one of the relatively few cases, besides those treated in lectures, for which exact results can be obtained. The system we consider consists of a particle of mass m moving on a circle of circumference L. The quantum Hamiltonian is

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}$$

and wavefunctions obey $\psi(x + L) = \psi(x)$. We want to determine the imaginary time propagator

$$\langle x_1 | \exp(-\beta H) | x_2 \rangle \,.$$

a) What are the eigenstates and eigenvalues of H? As we are dealing with a free particle, we can determine the propagator as in the lectures in a simple way by inserting resolutions of the identity in terms of the eigenstates of H. Show that this leads to the following result

$$\langle x_1|\exp(-\beta H)|x_2\rangle = \sum_{n=-\infty}^{\infty} \frac{1}{L}\exp\left(-\frac{\beta(2\pi n)^2\hbar^2}{2mL^2} + 2\pi i n\frac{[x_1-x_2]}{L}\right). \tag{152}$$

b) Next, approach this using a path integral in which paths $x(\tau)$ for $0 \le \tau \le \beta\hbar$ satisfy the boundary conditions $x(0) = x_1$ and $x(\beta\hbar) = x_2$. The special feature of a particle moving on a circle is that such paths may wind any integer number l times around the circle. To build in this feature, write

$$x(\tau) = x_1 + \frac{\tau}{\beta\hbar}[(x_2 - x_1) + lL] + s(\tau),$$

where the contribution $s(\tau)$ obeys the simpler boundary conditions $s(0) = s(\beta\hbar) = 0$ and does *not* wrap around the circle. Show that the Euclidean action for the system on such a path is

$$S[x(\tau)] = S_l + S[s(\tau)] \qquad \text{where} \qquad S_l = \frac{m}{2\beta\hbar}[(x_2-x_1)+lL]^2 \qquad \text{and} \qquad S[s(\tau)] = \int_0^{\beta\hbar} \mathrm{d}\tau \frac{m}{2}\left(\frac{\mathrm{d}s}{\mathrm{d}\tau}\right)^2.$$

c) using the results of b) show that

$$\langle x_1|\exp(-\beta H)|x_2\rangle \;= \mathcal{Z}_0 \sum_{l=-\infty}^{\infty} \exp\left(-\frac{m}{2\beta\hbar^2}[(x_1-x_2)+lL]^2\right), \tag{153}$$

where $\mathcal{Z}_0$ is the diagonal matrix element $\langle x|\mathrm{e}^{-\beta H}|x\rangle$ for a *free* particle (i.e. without periodic boundary conditions) moving in one dimension.

d) Argue on the basis of the result you obtained in Qu 3. for the propagator of a free particle that

$$\mathcal{Z}_0 = \left(\frac{m}{2\pi\beta\hbar^2}\right)^{1/2}. \tag{154}$$

e) Show that the expressions in Eq. (152) and Eq. (153) are indeed equal. To do so, you should use the *Poisson summation formula*

$$\sum_{l=-\infty}^{\infty} \delta(y-l) = \sum_{n=-\infty}^{\infty} \mathrm{e}^{-2\pi i n y}$$

(think about how to justify this). Introduce the left hand side of this expression into Eq. (153) by using the relation, valid for any smooth function $f(y)$,

$$\sum_{l=-\infty}^{\infty} f(l) = \int_{-\infty}^{\infty} \mathrm{d}y \sum_{l=-\infty}^{\infty} \delta(y-l) f(y),$$

substitute the right hand side of the summation formula, carry out the (Gaussian) integral on y, and hence establish the required equality.

3.3 Correlation Functions

It is clear from the above that we can calculate more general quantities from the generating functional $W[J]$, namely

$$\frac{1}{W[0]}\prod_{j=1}^{n}\frac{\delta}{\delta J(\tau_j)}\bigg|_{J=0} W[J] = \frac{\mathcal{N}}{W[0]}\int \mathcal{D}x(\tau)\prod_{j=1}^{n} x(\tau_j)\; e^{-\frac{1}{\hbar}S_E[x(\tau)]} \tag{155}$$

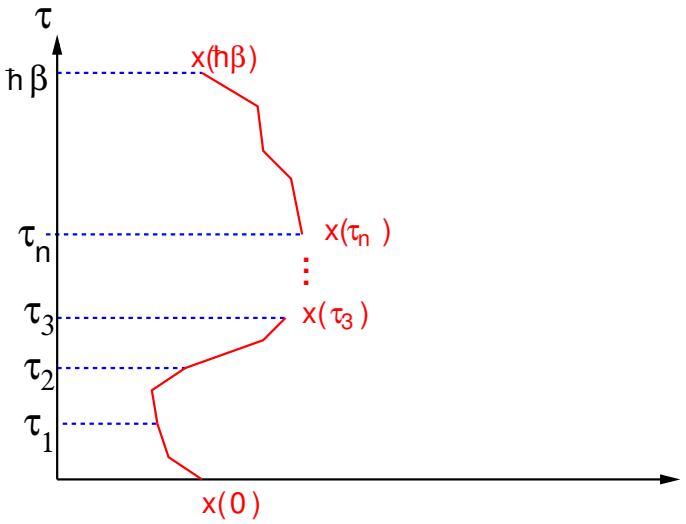


Figure 3: Path integral corresponding to (155).

What is their significance? Graphically, the path integral in (155) is represented in Fig. 3. It consists of several parts. The first part corresponds to propagation from $x(0)$ to $x(\tau_1)$ and the associated propagator is

$$\langle x(\tau_1)|e^{-H\tau_1/\hbar}|x(0)\rangle. \tag{156}$$

The second part corresponds to propagation from $x(\tau_1)$ to $x(\tau_2)$, and we have a multiplicative factor of $x(\tau_1)$ as well. This is equivalent to a factor

$$\langle x(\tau_2)|e^{-H(\tau_2-\tau_1)/\hbar}\hat{x}|x(\tau_1)\rangle. \tag{157}$$

Repeating this analysis for the other pieces of the path we obtain

$$\left[\prod_{j=1}^{n}\langle x(\tau_{j+1})|e^{-H(\tau_{j+1}-\tau_j)/\hbar}\hat{x}|x(\tau_j)\rangle\right]\langle x(\tau_1)|e^{-H\tau_1/\hbar}|x(0)\rangle\ , \tag{158}$$

where $\tau_{n+1}=\hbar\beta$. Finally, in order to represent the full path integral (155) we need to integrate over the intermediate positions $x(\tau_j)$ and impose periodicity of the path. Using that $\mathbf{1}=\int dx|x\rangle\langle x|$ and that $W[0]=Z(\beta)$ we arrive at

$$\begin{aligned}\frac{1}{Z(\beta)}\int dx(0)\langle x(0)|e^{-H(\beta-\tau_n)/\hbar}\hat{x}e^{-H(\tau_n-\tau_{n-1})/\hbar}\hat{x}\ldots\hat{x}e^{-H(\tau_2-\tau_1)/\hbar}\hat{x}e^{-H\tau_1/\hbar}|x(0)\rangle \\ =\frac{1}{Z(\beta)}\mathrm{Tr}\left[e^{-\beta H}\bar{x}(\tau_n)\bar{x}(\tau_{n-1})\ldots\bar{x}(\tau_1)\right],\end{aligned} \tag{159}$$

where we have defined operators

$$\bar{x}(\tau_j)=e^{H\tau_j/\hbar}\hat{x}e^{-H\tau_j/\hbar}. \tag{160}$$

There is one slight subtlety: in the above we have used implicitly that $\tau_1<\tau_2<\ldots<\tau_n$. On the other hand, our starting point (155) is by construction symmetric in the τ_j. The way to fix this is to introduce a *time-ordering* operation T_τ, which automatically arranges operators in the "right" order. For example

$$T_\tau\bar{x}(\tau_1)\bar{x}(\tau_2)=\theta(\tau_1-\tau_2)\bar{x}(\tau_1)\bar{x}(\tau_2)+\theta(\tau_2-\tau_1)\bar{x}(\tau_2)\bar{x}(\tau_1), \tag{161}$$

where $\theta(x)$ is the Heaviside theta function. Then we have

$$\boxed{\frac{1}{W[0]}\prod_{j=1}^{n}\frac{\delta}{\delta J(\tau_j)}\bigg|_{J=0}W[J]=\frac{1}{Z(\beta)}\mathrm{Tr}\left[e^{-\beta H}T_\tau\bar{x}(\tau_1)\bar{x}(\tau_2)\ldots\bar{x}(\tau_n)\right].} \tag{162}$$

Finally, if we *analytically continue* from imaginary time to real time $\tau_j \to it_j$, the operators $\bar{x}(\tau)$ turn into *Heisenberg-picture* operators

$$\hat{x}(t) \equiv e^{\frac{it}{\hbar}H}\hat{x}e^{-\frac{it}{\hbar}H}. \tag{163}$$

The quantities that we get from (162) after analytic continuation are called *n-point correlation functions*

$$\boxed{\langle T\hat{x}(t_1)\hat{x}(t_2)\dots\hat{x}(t_n)\rangle_\beta \equiv \frac{1}{Z(\beta)}\mathrm{Tr}\left[e^{-\beta H}T\hat{x}(t_1)\hat{x}(t_2)\dots\hat{x}(t_n)\right].} \tag{164}$$

Here T is a time-ordering operator that arranges the $\hat{x}(t_j)$'s in chronologically increasing order from right to left. Such correlation functions are the central objects in both quantum field theory and many-particle quantum physics.

3.3.1 Wick's Theorem

Recalling that

$$W[J] = W[0]\ e^{\frac{1}{2}\int d\tau d\tau'\ J(\tau)G(\tau-\tau')J(\tau')}, \tag{165}$$

then taking the functional derivatives, and finally setting all sources to zero we find that

$$\frac{1}{W[0]}\prod_{j=1}^{n}\frac{\delta}{\delta J(\tau_j)}\bigg|_{J=0} W[J] = \sum_{P(1,\dots,n)} G(\tau_{P_1}-\tau_{P_2})\dots G(\tau_{P_{n-1}}-\tau_{P_n})\ . \tag{166}$$

Here the sum is over all possible pairings of $\{1,2,\dots,n\}$ and $G(\tau)$ is the Green's function (138). In particular we have

$$\langle T_\tau \bar{x}(\tau_1)\bar{x}(\tau_2)\rangle_\beta = G(\tau_1-\tau_2). \tag{167}$$

The fact that for "Gaussian theories" [3] like the harmonic oscillator n-point correlation functions can be expressed as simple products over 2-point functions is known as *Wick's theorem.*

3.4 Probability distribution of position

Using Wick's theorem it is now straightforward to calculate all moments $\langle\hat{x}^{2n}\rangle_\beta$ for the harmonic oscillator. It is instructive to calculate the corresponding probability distribution directly. Let us first work out the relevant expectation value to consider. Let $|\psi\rangle$ be an arbitrary state and consider

$$\langle\psi|\delta(\hat{x}-x_0)|\psi\rangle = \int dx\langle\psi|\delta(\hat{x}-x_0)|x\rangle\langle x|\psi\rangle = \int dx\ \delta(x-x_0)\langle\psi|x\rangle\langle x|\psi\rangle = |\psi(x_0)|^2. \tag{168}$$

So the expectation value of the delta-function indeed gives the correct result for the probability distribution of a position measurement, namely the absolute value squared of the wave function. We then have

$$\begin{aligned}\langle\delta(\hat{x}-x_0)\rangle_\beta &= \int_{-\infty}^{\infty}\frac{dk}{2\pi}\langle e^{ik(\hat{x}-x_0)}\rangle_\beta \\ &= \int_{-\infty}^{\infty}\frac{dk}{2\pi}e^{-ikx_0}\frac{\mathcal{N}}{W[0]}\int\mathcal{D}x(\tau)\ e^{-\frac{1}{2}\int_0^{\hbar\beta}d\tau\ x(\tau)\hat{D}x(\tau)+\int_0^{\hbar\beta}d\tau\ x(\tau)ik\delta(\tau)}.\end{aligned} \tag{169}$$

This is a special case of our generating functional, where the source is given by $J(\tau) = ik\delta(\tau)$. We therefore can use (124) to obtain

$$\begin{aligned}\langle\delta(\hat{x}-x_0)\rangle_\beta &= \int_{-\infty}^{\infty}\frac{dk}{2\pi}e^{-ikx_0}\frac{W[ik\delta(\tau)]}{W[0]} = \int_{-\infty}^{\infty}\frac{dk}{2\pi}e^{-ikx_0}e^{\frac{1}{2}\int_0^{\hbar\beta}d\tau\int_0^{\hbar\beta}d\tau'\left(ik\delta(\tau)\right)G(\tau-\tau')\left(ik\delta(\tau')\right)} \\ &= \frac{1}{\sqrt{2\pi G(0)}}e^{-x_0^2/2G(0)}.\end{aligned} \tag{170}$$

[3] These are theories in which the Lagrangian is quadratic in the generalized co-ordinates.

To go from the first to the second line we have taken the integrals over τ and τ' (which are straightforward because of the two delta functions) and finally carried out the k-integral using the one dimensional version of (34). We see that our probability distribution is a simple Gaussian with a variance that depends on temperature through $G(0)$. Note that at zero temperature (170) reduces, as it must, to $|\psi_0(x_0)|^2$, where $\psi_0(x)$ is the ground state wave function of the harmonic oscillator.

Exercise 4

If you have never worked out the harmonic oscillator at finite temperature before using "Heisenberg QM", here is how to do it. Defining creation and annihilation operators by

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right) \ , \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i\hat{p}}{m\omega}\right), \tag{171}$$

brings the Hamiltonian to the form

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right). \tag{172}$$

The creation/annihilation operators act on energy eigenstates $|n\rangle$ as

$$a|n\rangle = \sqrt{n}|n-1\rangle \ , \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \tag{173}$$

The energy eigenvalues are

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right). \tag{174}$$

The partition function is

$$Z(\beta) = \sum_{n=0}^{\infty} e^{-\beta E_n} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \tag{175}$$

Similarly you can calculate

$$\frac{1}{Z(\beta)}\mathrm{Tr}\left(e^{-\beta H}a^\dagger a\right) = \sum_{n=1}^{\infty} n e^{-\beta E_n} = \frac{1}{e^{\beta\hbar\omega} - 1}. \tag{176}$$

Time dependent operators are defined following (160)

$$\bar{a}(\tau) = e^{H\tau/\hbar} a e^{-H\tau/\hbar} \ , \quad \bar{a}^\dagger(\tau) = e^{H\tau/\hbar} a^\dagger e^{-H\tau/\hbar}. \tag{177}$$

Note that

$$\bar{a}^\dagger(\tau) \neq (\bar{a}(\tau))^\dagger. \tag{178}$$

Taking derivatives this gives the following equations of motion

$$\frac{d}{d\tau}\bar{a}(\tau) = \frac{1}{\hbar} e^{H\tau/\hbar}[H, a]e^{-H\tau/\hbar} = -\omega\bar{a}(\tau). \tag{179}$$

Integrating this as well as the analogous equation for $\bar{a}^\dagger(\tau)$ gives

$$\bar{a}(\tau) = e^{-\omega\tau}a \ , \quad \bar{a}^\dagger(\tau) = e^{\omega\tau}a^\dagger. \tag{180}$$

This allows us to compute imaginary-time dependent thermal averages

$$\begin{aligned} \frac{1}{Z(\beta)}\mathrm{Tr}\left(e^{-\beta H}\bar{a}(\tau)\bar{a}^\dagger(0)\right) &= \frac{e^{-\omega\tau}}{1 - e^{-\beta\hbar\omega}} \ , \\ \frac{1}{Z(\beta)}\mathrm{Tr}\left(e^{-\beta H}\bar{a}^\dagger(\tau)\bar{a}(0)\right) &= \frac{e^{\omega\tau}}{e^{\beta\hbar\omega} - 1}. \end{aligned} \tag{181}$$

Finally, recalling the relation between $\hat{x}$ and a and $a^\dagger$ we find

$$\frac{1}{Z(\beta)}\mathrm{Tr}\left(e^{-\beta H}\bar{x}(\tau)\bar{x}(0)\right) = \frac{\hbar\omega}{2\kappa}\left[\frac{e^{\omega\tau}}{e^{\beta\hbar\omega} - 1} + \frac{e^{-\omega\tau}}{1 - e^{-\beta\hbar\omega}}\right]. \tag{182}$$

3.5 Perturbation Theory and Feynman Diagrams

Let us now consider the anharmonic oscillator

$$H = \frac{\hat{p}^2}{2m} + \frac{\kappa}{2}\hat{x}^2 + \frac{\lambda}{4!}\hat{x}^4. \tag{183}$$

As you know from QM2, this Hamiltonian is no longer exactly solvable. What we want to do instead is perturbation theory for small $\lambda > 0$. As the Hamiltonian is of the form $H = \frac{\hat{p}^2}{2m} + V(\hat{x})$ our previous construction of the path integral applies. Our generating functional becomes

$$W_\lambda[J] = \mathcal{N}\int \mathcal{D}x(\tau)\; e^{-\frac{1}{\hbar}S_E[x(\tau)] + \int_0^{\hbar\beta} d\tau\; J(\tau)x(\tau) - \frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau\; x^4(\tau)}. \tag{184}$$

The partition function is

$$Z_\lambda(\beta) = W_\lambda[0]\ . \tag{185}$$

The idea is to expand (184) perturbatively in powers of λ

$$\begin{aligned}
W_\lambda[J] &= \mathcal{N}\int \mathcal{D}x(\tau)\; \left[1 - \frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau'\; x^4(\tau') + \ldots\right] e^{-\frac{1}{\hbar}S_E[x(\tau)] + \int_0^{\hbar\beta} d\tau\; J(\tau)x(\tau)} \\
&= \mathcal{N}\int \mathcal{D}x(\tau)\; \left[1 - \frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau'\; \left[\frac{\delta}{\delta J(\tau')}\right]^4 + \ldots\right] e^{-\frac{1}{\hbar}S_E[x(\tau)] + \int_0^{\hbar\beta} d\tau\; J(\tau)x(\tau)} \\
&= \mathcal{N}\int \mathcal{D}x(\tau)\; e^{-\frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau'\; \left[\frac{\delta}{\delta J(\tau')}\right]^4} e^{-\frac{1}{\hbar}S_E[x(\tau)] + \int_0^{\hbar\beta} d\tau\; J(\tau)x(\tau)} \\
&= e^{-\frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau'\; \left[\frac{\delta}{\delta J(\tau')}\right]^4} W_0[J].
\end{aligned} \tag{186}$$

We already know $W_0[J]$

$$W_0[J] = W_0[0]\; e^{\frac{1}{2}\int d\tau d\tau'\; J(\tau)G(\tau-\tau')J(\tau')}, \tag{187}$$

which will enable us to work out a perturbative expansion very efficiently.

3.5.1 Partition Function of the anharmonic oscillator

By virtue of (185) the perturbation expansion for $Z_\lambda(\beta)$ is

$$\begin{aligned}
Z_\lambda(\beta) &= e^{-\frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau'\; \left[\frac{\delta}{\delta J(\tau')}\right]^4} W_0[J]\Big|_{J=0} = Z_0(\beta) - \frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau'\; \left[\frac{\delta}{\delta J(\tau')}\right]^4\Big|_{J=0} W_0[J] \\
&\qquad + \frac{1}{2}\left[\frac{\lambda}{4!\hbar}\right]^2 \int_0^{\hbar\beta} d\tau' d\tau''\; \left[\frac{\delta}{\delta J(\tau')}\right]^4 \left[\frac{\delta}{\delta J(\tau'')}\right]^4\Big|_{J=0} W_0[J] + \ldots \\
&= Z_0(\beta)\left[1 + \lambda\gamma_1(\beta) + \lambda^2\gamma_2(\beta) + \ldots\right].
\end{aligned} \tag{188}$$

1. First order perturbation theory:

Carrying out the functional derivatives gives

$$\lambda\gamma_1(\beta) = -\frac{\lambda}{8\hbar}\int_0^{\hbar\beta} d\tau'\; [G(\tau-\tau)]^2 = -\frac{\lambda\beta}{8}[G(0)]^2\,. \tag{189}$$

This contribution can be represented graphically by a *Feynman diagram*. In order to do so we introduce the following elements:

1. The two-point function $G(\tau - \tau')$ is represented by a line running from τ to τ'.

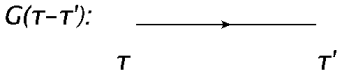


Figure 4: Graphical representation of the Green's function.

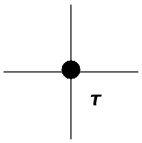


Figure 5: Graphical representation of the interaction vertex.

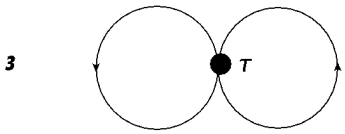


Figure 6: Feynman diagram for the 1st order perturbative contribution to the partition function.

2. The *interaction vertex* $-\frac{\lambda}{4!\hbar}\int_0^{\hbar\beta} d\tau$ is represented by by the 4-point vertex shown in Fig. 5.

Combining these two elements, we can express the integral $\lambda\gamma_1(\beta)$ by the diagram shown in Fig. 6.

Here the factor of 3 is a *combinatorial factor* associated with the diagram. We will discuss this more below. To see why it is quite natural to introduce Feynman diagrams in the way we have done let us go back to taking functional derivatives. Let us introduce a short-hand notation

$$W[J] = W[0]e^{\frac{1}{2}J_1G_{12}J_2} = W[0]\left[1 + \frac{1}{2}J_1G_{12}J_2 + \frac{1}{2^3}J_1G_{12}J_2J_3G_{34}J_4 + \ldots\right]. \tag{190}$$

The last term we have written is the one that gives rise to our first order diagram, so we have a factor $\frac{1}{2^3}$ to begin with. The first order contribution is worked out in Fig. 7. There are 4 ways of connecting the first

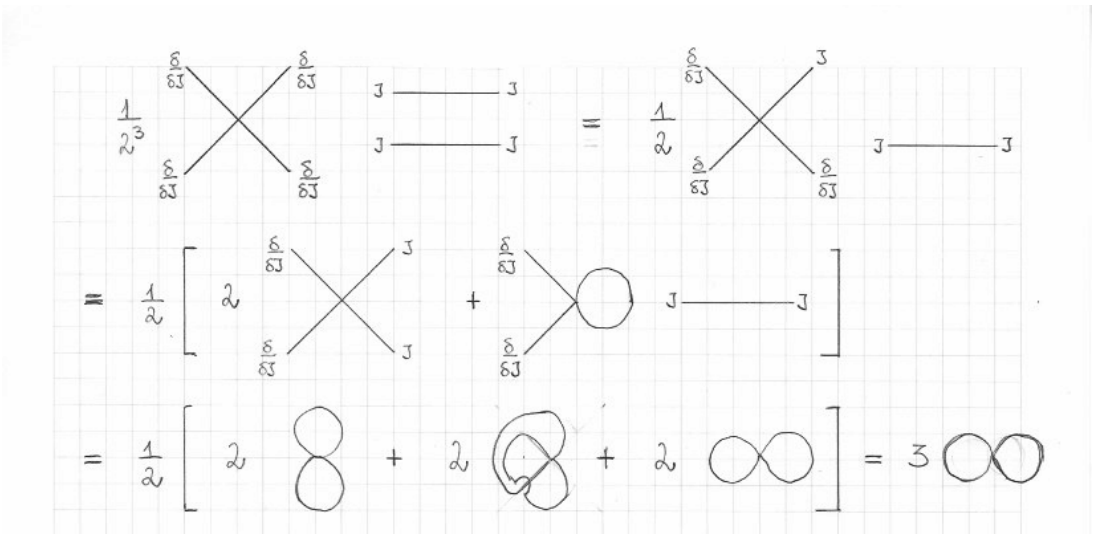


Figure 7: Combinatorics for the first order diagram

line to the vertex. In the next step, we can connect the second line in the vertex either to the remaining

external line (giving a factor of 2 as we could connect to either end), or to the line in the vertex that was connected in the first step. In the last step we complete the diagram.

Let us now turn to the combinatorial factor associated with the Feynman diagram. This is simply equal to the ways of connecting the lines in our vertex in such a way as to reproduce the Feynman diagram we want. As shown in Fig. 8 there are three ways.

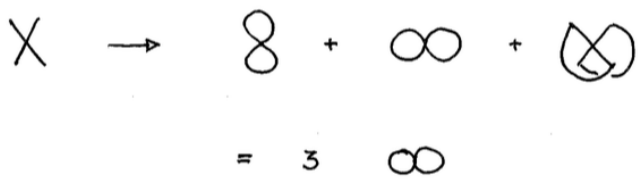


Figure 8: Combinatorial factor for the first order vacuum diagram.

2. Second order perturbation theory.

To second order we obtain a number of different contributions upon taking the functional derivatives. The full second order contribution is

$$\begin{aligned}\lambda^2\gamma_2(\beta) &= \frac{1}{2}\left(\frac{\lambda}{4!\hbar}\right)^2 72\int_0^{\hbar\beta}d\tau\int_0^{\hbar\beta}d\tau'\; G(\tau-\tau)G^2(\tau-\tau')G(\tau'-\tau') \\ &+ \frac{1}{2}\left(\frac{\lambda}{4!\hbar}\right)^2 24\int_0^{\hbar\beta}d\tau\int_0^{\hbar\beta}d\tau'\; G^4(\tau-\tau') \\ &+ \frac{1}{2}\left(\frac{\lambda}{4!\hbar}\right)^2 9\left[\int_0^{\hbar\beta}d\tau G^2(\tau-\tau)\right]^2 . \end{aligned} \tag{191}$$

The corresponding Feynman diagrams are shown in Fig.9. They come in two types: the first two are *connected*, while the third is *disconnected*. The factor of 1/2 that has been separated off is associated with second order perturbation theory.

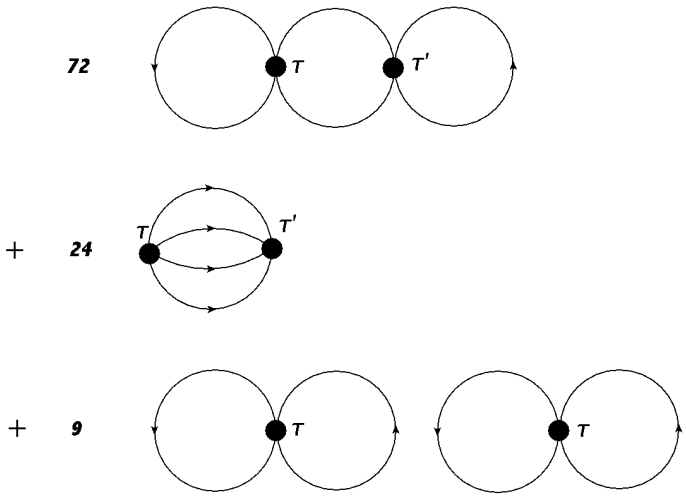


Figure 9: Feynman diagram for the 2nd order perturbative contribution to the partition function.

The point about the Feynman diagrams is that rather than carrying out functional derivatives and then representing various contributions in diagrammatic form, in practice we do the calculation by writing down the diagrams *first* and then working out the corresponding integrals! How do we know what diagrams to

draw? As we are dealing with the partition function, we can never produce a diagram with a line sticking out: all (imaginary) times must be integrated over. Such diagrams are sometimes called *vacuum diagrams*.

An important part of the process is to work out the correct combinatorial factors. For the second order vacuum diagrams this is shown in Figs 10, 11 and 12.

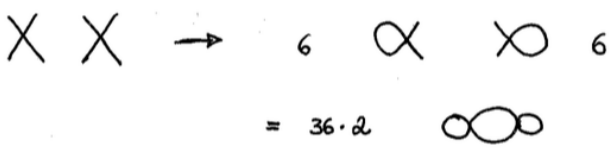


Figure 10: To produce the first diagram in Fig. 9 we need to make one loop at each vertex (there are six ways to do this), and then to connect the two resulting pieces (there are two ways of doing this).

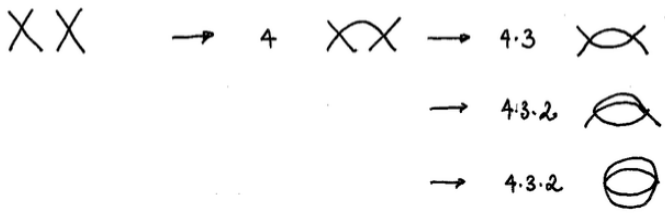


Figure 11: To produce the second diagram in Fig. 9 we need to connect all lines of the first vertex with the second vertex. There are 4 ways of connecting the first line, 3 ways of connecting the second, and two ways of connecting the third.



Figure 12: To produce last second diagram in Fig. 9 we need to connect all lines at each individual vertex. As discussed before there are 3 ways of doing this.

3: Aside: Symmetry Factors.

You might have wondered why we have the $\frac{1}{4!}$ normalization in front of the λx^4 term. This is designed so that if we had to, e.g. consider a process with 4 external legs (a 4-point correlation function) then the leading connected contribution which appears at $O(\lambda)$ comes with a coefficient of 1. This follows because there are 4! different ways in which we can connect an external leg to the vertex. However, when we have loops, then two different contractions could give the same diagrammatic result. There are two ways to account for this. The most *straightforward* is the route taken in these notes, which is to simply compute the number of ways a particular diagram can appear and keep all the factors of 1/4! in front, and cancel off against the combinatorics as needed. The more *elegant* approach, adopted by expert practitioners, is to work with *symmetry factors*: it turns out that the ratio between the combinatorics (the number of ways of getting a particular diagram) and the factorials is related to the symmetries of the Feynman diagram. So one can 'read off' this ratio without going through tedious combinatorics.

Put simply: one can either compute a large combinatorial factor and divide with various powers of the 4! terms, or else one can just compute the symmetry factor S and divide the diagram by it. For example, the symmetry factor for the 'figure of 8' diagram in the partition function is given by $S^{-1} = \frac{1}{8} = \frac{3}{4!}$.
Familiarity with symmetry factor technology is often crucial to doing high order diagrammatic perturbation theory, but not terribly important for our present purposes.

3.6 Homework Question 7

Question 7. Anharmonic Oscillator. Consider the anharmonic oscillator

$$H(\lambda_1, \lambda_2) = \frac{\hat{p}^2}{2m} + \frac{\kappa}{2}\hat{x}^2 + \frac{\lambda_1}{3!}\hat{x}^3 + \frac{\lambda_2}{4!}\hat{x}^4. \tag{192}$$

where $\kappa, \lambda_{1,2} > 0$ and $\lambda_1^2 - 3\kappa\lambda_2 < 0$. Define a generating functional by

$$W_{\lambda_1,\lambda_2}[J] = \mathcal{N}\int \mathcal{D}x(\tau)\; e^{\left\{\left[-\frac{1}{\hbar}S_E[x(\tau)] + \int_0^{\hbar\beta} d\tau J(\tau)x(\tau)\right] + U\left(x(\tau)\right)\right\}}, \tag{193}$$

where

$$U\big(x(\tau)\big) = -\frac{1}{\hbar}\int_0^{\hbar\beta} d\tau \left[\frac{\lambda_1}{3!}x^3(\tau) + \frac{\lambda_2}{4!}x^4(\tau)\right], \quad \hat{D} = -\frac{m}{\hbar}\frac{d^2}{d\tau^2} + \frac{\kappa}{\hbar}. \tag{194}$$

a) Show that the partition function is equal to

$$Z_{\lambda_1,\lambda_2}(\beta) = W_{\lambda_1,\lambda_2}[0]. \tag{195}$$

b) Show that the generating functional can be expressed in the form

$$W_{\lambda_1,\lambda_2}[J] = \exp\left(U\Big(\frac{\delta}{\delta J(\tau)}\Big)\right) W_{0,0}[J]. \tag{196}$$

c) Determine the first order perturbative corrections in λ_1 and λ_2 to the partition function. Draw the corresponding Feynman diagrams.
d) Determine the perturbative correction to the partition function proportional to λ_1^2. Draw the corresponding Feynman diagrams. Are there corrections of order $\lambda_1\lambda_2$?
e)* Determine the first order corrections to the two-point function

$$\langle T_\tau \bar{x}(\tau_1)\bar{x}(\tau_2)\rangle_\beta. \tag{197}$$

Draw the corresponding Feynman diagrams. What diagrams to you get in second order in perturbation theory?

4: Aside: Response Functions

To see why correlation functions are important for applications to experiments let us briefly summarize the main facts on *linear response theory*. To ease notations let us introduce the *thermal density matrix*

$$\rho = \frac{1}{Z(\beta)}e^{-\beta H} = \frac{1}{Z(\beta)}\sum_n e^{-\beta E_n}|n\rangle\langle n|, \tag{198}$$

where $H|n\rangle = E_n|n\rangle$. Then

$$\langle \mathcal{O}\rangle_\beta = \mathrm{Tr}\left(\rho\mathcal{O}\right). \tag{199}$$

The time evolution of ρ follows from that of the energy eigenstates

$$|n(t)\rangle = U(t,t_0)|n(t_0)\rangle \Rightarrow \rho(t) = U(t,t_0)\rho(t_0)U^\dagger(t,t_0) \tag{200}$$

Let us now consider a time-dependent Hamiltonian

$$H = H_0 + \lambda(t)\mathcal{O}_1 \ , \tag{201}$$

where we take $\lambda(t)$ to be very small at all times. In practice the perturbation $\lambda(t)\mathcal{O}_1$ will correspond to some external experimental probe. Time-dependent perturbation theory tells us that the time-evolution operator in this case can be written to leading order in $\lambda(t)$ as

$$U(t,t_0) = U_0(t,t_0)\left[1 - \frac{i}{\hbar}\int_{t_0}^{t} dt' \ \lambda(t')U_0^\dagger(t,t_0)\mathcal{O}_1 U_0(t,t_0)\right] + \dots \ , \tag{202}$$

where

$$U_0(t,t_0) = e^{-\frac{i}{\hbar}H_0(t-t_0)}. \tag{203}$$

Let us denote Heisenberg-picture operators in the unperturbed theory ($\lambda(t) = 0$) by

$$\mathcal{O}^{(0)}(t) = U_0^\dagger(t,t_0)\mathcal{O}(t_0)U_0(t,t_0). \tag{204}$$

Let us now look at the thermal average of an observable $\mathcal{O}$

$$\mathrm{Tr}\left(\rho(t)\mathcal{O}\right) = \mathrm{Tr}\left(\rho\mathcal{O}(t)\right), \tag{205}$$

where

$$\mathcal{O}(t) = U^\dagger(t,t_0)\mathcal{O}U(t,t_0) \approx \mathcal{O}^{(0)}(t) - \frac{i}{\hbar}\int_{t_0}^{t} dt'\lambda(t')[\mathcal{O}^{(0)}(t), \mathcal{O}_1^{(0)}(t')] \tag{206}$$

Let us now imagine that we switch on the perturbation at $t_0 = -\infty$ and calculate the change in the thermal average of an observable $\mathcal{O}$ induced by the perturbation

$$\begin{aligned} \mathrm{Tr}\left(\rho(t)\mathcal{O}\right) - \mathrm{Tr}\left(\rho(-\infty)\mathcal{O}\right) &\approx \frac{1}{\hbar}\int_{-\infty}^{\infty} dt' \ \lambda(t') \ \chi_{\mathcal{O}\mathcal{O}_1}(t,t') \ , \\ \chi_{\mathcal{O}\mathcal{O}_1}(t,t') &= -i\Theta(t-t')\langle[\mathcal{O}^{(0)}(t), \mathcal{O}_1^{(0)}(t')]\rangle_\beta. \end{aligned} \tag{207}$$

Here $\chi_{\mathcal{O}\mathcal{O}_1}(t,t')$ is called a generalized *susceptibility*. Next comes a bit of magic. First, observe that the susceptibility in fact depends only on the time difference, i.e. $\chi_{\mathcal{O}\mathcal{O}_1}(t,t') = \chi_{\mathcal{O}\mathcal{O}_1}(t-t')$. Considering the Fourier transform then gives

$$\widetilde{\chi}_{\mathcal{O}\mathcal{O}_1}(\omega) = \int_{-\infty}^{\infty} dt \ e^{i\omega t}\chi_{\mathcal{O}\mathcal{O}_1}(t) = -i\int_0^{\infty} dt \ e^{i\omega t} \ \langle[\mathcal{O}^{(0)}(t), \mathcal{O}_1^{(0)}(0)]\rangle_\beta \tag{208}$$

Inserting a resolution of the identity in terms of the eigenstates of H_0 gives the following *Lehmann representation*

$$\begin{aligned} \widetilde{\chi}_{\mathcal{O}\mathcal{O}_1}(\omega) &= -i\int_0^{\infty} dt \ e^{i\omega t}\sum_{n,m}\frac{e^{-\beta E_n}}{Z(\beta)}\left[\langle n|\mathcal{O}^{(0)}(t)|m\rangle\langle m|\mathcal{O}_1^{(0)}(0)|n\rangle - \langle n|\mathcal{O}_1^{(0)}(0)|m\rangle\langle m|\mathcal{O}^{(0)}(t)|n\rangle\right] \\ &= \sum_{n,m}\frac{\left(e^{-\beta E_n} - e^{-\beta E_m}\right)}{Z(\beta)}\langle n|\mathcal{O}^{(0)}(0)|m\rangle\langle m|\mathcal{O}_1^{(0)}|n\rangle(-i)\int_0^{\infty} dt \ e^{i(\omega - E_m + E_n)t} \\ &= \frac{1}{Z(\beta)}\sum_{n,m}\frac{\left(e^{-\beta E_n} - e^{-\beta E_m}\right)}{\omega + i0 - E_m + E_n}\langle n|\mathcal{O}^{(0)}(0)|m\rangle\langle m|\mathcal{O}_1^{(0)}|n\rangle. \end{aligned} \tag{209}$$

Now consider the imaginary-time correlator

$$-\langle T_\tau \mathcal{O}(\tau)\mathcal{O}_1(0)\rangle_\beta \tag{210}$$

This is a periodic function of τ and its Fourier transform is

$$\begin{aligned}\bar{\chi}_{\mathcal{O}\mathcal{O}_1}(\omega_\ell) &= -\int_0^{\hbar\beta} d\tau\ e^{i\omega_\ell\tau}\langle T_\tau \mathcal{O}(\tau)\mathcal{O}_1(0)\rangle_\beta \\ &= \frac{1}{Z(\beta)}\sum_{n,m}\frac{\left(e^{-\beta E_n}-e^{-\beta E_m}\right)}{i\omega_\ell - E_m + E_n}\langle n|\mathcal{O}^{(0)}(0)|m\rangle\langle m|\mathcal{O}_1^{(0)}|n\rangle. \end{aligned}\tag{211}$$

Comparing (211) and (209) we see that the Fourier transform of the dynamical susceptibility is equal to the Fourier transform of the imaginary-time correlator, if we *analytically continue* from Matsubara frequencies to real frequencies

$$i\omega_\ell \longrightarrow \omega + i0. \tag{212}$$

5: Aside: Response Functions for the Harmonic Oscillator

Let us now consider the explicit example on a harmonic oscillator that is perturbed by an infinitesimal electric field $\mathcal{E}(t)$

$$H(t) = \frac{\hat{p}^2}{2m} + \frac{\kappa}{2}\hat{x}^2 - \mathcal{E}(t)\hat{x} \tag{213}$$

Then according to our general result (207) the change in the average position is

$$\begin{aligned}\mathrm{Tr}\,(\rho(t)\hat{x}) - \mathrm{Tr}\,(\rho(-\infty)\hat{x}) &\approx \frac{1}{\hbar}\int_{-\infty}^{\infty} dt'\ \mathcal{E}(t')\ \chi_{\hat{x}\hat{x}}(t-t')\ , \\ &= \frac{1}{\hbar}\int_{-\infty}^{\infty} dt'\ \mathcal{E}(t')\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega(t-t')}\widetilde{\chi}_{\hat{x}\hat{x}}(\omega). \end{aligned}\tag{214}$$

According to (167), (138) we have

$$\langle T_\tau \bar{x}(\tau)\bar{x}(0)\rangle_\beta = G(\tau) = \frac{1}{\beta\kappa}\sum_{n=-\infty}^{\infty}\frac{\omega^2}{\omega^2+\omega_n^2}e^{i\omega_n\tau}\ , \tag{215}$$

which results in

$$\bar{\chi}_{\hat{x}\hat{x}}(\omega_\ell) = -\frac{\hbar}{\kappa}\frac{\omega^2}{\omega^2+\omega_\ell^2}\ . \tag{216}$$

Analytically continuing this to real frequencies we obtain

$$\widetilde{\chi}_{\hat{x}\hat{x}}(z) = \frac{\hbar}{2\kappa}\left(\frac{\omega}{z+i0-\omega} - \frac{\omega}{z+i0+\omega}\right), \tag{217}$$

and Fourier transforming back gives

$$\chi_{\hat{x}\hat{x}}(t) = -\theta(t)\frac{\hbar\omega}{\kappa}\sin(\omega t)\ . \tag{218}$$

This susceptibility equals the *retarded correlation function* of the position operator and, interestingly, is independent of temperature.

Part II
Path Integrals and Transfer Matrices

4 Relation of D dimensional quantum systems to $D+1$ dimensional classical ones

Let's start by defining what we mean by the spatial "dimension" D of a system. Let us do this by considering a (quantum) field theory. There the basic objects are fields, that depend on time and are defined at all points of a D-dimensional space. This value of D defines what we mean by the spatial dimension. For example, in electromagnetism we have $D = 3$. In this terminology a single quantum mechanical particle or spin are zero-dimensional systems. On the other hand, a linear chain of spins is a one-dimensional system, while a bcc lattice of spins has $D = 3$. Interestingly, there is a representation of D dimensional quantum systems in terms of $D+1$ dimensional classical ones. We will now establish this for the particular case of the simple quantum mechanical harmonic oscillator.

4.1 Some Facts from Statistical Physics

Consider a classical many-particle system coupled to a heat bath at temperature T. The *partition function* is defined as

$$Z = \sum_{\text{configurations C}} e^{-\beta E(C)} , \qquad \beta = \frac{1}{k_B T}. \tag{219}$$

Here the sum is over all possible configurations C, and $E(C)$ is the corresponding energy. Thermal averages of observables are given by

$$\langle \mathcal{O} \rangle_\beta = \frac{1}{Z} \sum_{\text{configurations C}} \mathcal{O}(C) e^{-\beta E(C)} , \tag{220}$$

where $\mathcal{O}(C)$ is the value of the observable $\mathcal{O}$ in configuration C. The average energy is

$$E = \frac{1}{Z} \sum_{\text{configurations C}} E(C) e^{-\beta E(C)} = -\frac{\partial}{\partial \beta} \ln(Z). \tag{221}$$

The *free energy* is

$$F = -k_B T \ln(Z). \tag{222}$$

The *entropy* is

$$S = \frac{E - F}{T} = k_B \ln(Z) - k_B \beta \frac{\partial}{\partial \beta} \ln(Z) = k_B \frac{\partial}{\partial T} \left[T \ln(Z) \right]. \tag{223}$$

4.2 Quantum Mechanical Particle

Let us revisit the path-integral representation (108) for the partition function of our QM particle at temperature β

$$Z(\beta) = \lim_{N \to \infty} \int dx \left[\frac{m}{2\pi\hbar\epsilon} \right]^{\frac{N}{2}} \int dx_1 \dots dx_{N-1} \exp\left(-\frac{\epsilon}{\hbar} \sum_{n=0}^{N-1} \frac{m}{2} \left(\frac{x_{n+1} - x_n}{\epsilon} \right)^2 + V(x_n) \right), \tag{224}$$

where we have set $x_0 = x_N = x$. For a given value of N, this can be interpreted as the partition function of N classical degrees of freedom x_j, that can be thought of as deviations of classical particles from their equilibrium positions, cf. Fig. 13. In this interpretation $V(x_j)$ is simply a potential energy associated with moving the j^{th} particle a distance x_j away from its equilibrium position, while $\frac{m}{2}(x_{n+1} - x_n)^2/\epsilon^2$ describes



Figure 13: Periodic array of classical particles.

an interaction energy that favours equal displacements, i.e. $x_n = x_{n+1}$. Importantly, the temperature T_{cl} of this one-dimensional classical model equals

$$k_B T_{cl} = \frac{\hbar}{\epsilon} = N k_B T. \tag{225}$$

So for large values of N (and fixed T) this temperature is very large. A convenient way for working out partition functions in classical statistical mechanics is by using *transfer matrices*. In the case at hand, this is defined as an integral operator $\hat{T}$ with kernel

$$\begin{aligned} T(x,x') &= \sqrt{\frac{m}{2\pi\epsilon\hbar}} e^{-\frac{\beta}{N} E_{\rm cl}(x,x')} \ , \\ E_{\rm cl}(x,x') &= \frac{m}{2}\left(\frac{x-x'}{\epsilon}\right)^2 + \frac{V(x)+V(x')}{2}. \end{aligned} \tag{226}$$

The integral operator $\hat{T}$ acts on functions $f(x)$ as

$$(\hat{T} * f)(x) = \int dx' \ T(x,x') f(x'). \tag{227}$$

In terms of this transfer matrix the partition function can be written as

$$\begin{aligned} Z(\beta) &= \lim_{N\to\infty} \int dx dx_1 \dots dx_{N-1} \ T(x,x_1) T(x_1,x_2) \dots T(x_{N-1},x) \\ &= \lim_{N\to\infty} \int dx \ T^N(x,x) = \lim_{N\to\infty} {\rm Tr}(\hat{T}^N). \end{aligned} \tag{228}$$

By construction $\hat{T}$ is a real, symmetric operator and can therefore be diagonalized. Hence the partition function can be expressed in terms of the eigenvalues of $\hat{T}$ using

$${\rm Tr}(\hat{T}^N) = \sum_n \lambda_n^N. \tag{229}$$

In order to get a clearer idea how to use transfer matrices in statistical mechanics problems we now turn to a simpler example, the celebrated *Ising model*. This is in fact *the* key paradigm in the theory of phase transitions.

6: Aside: Recap of the quantum-classical correspondence

It is important to realise that we have used the path integral formulation to make links between the transition amplitude and two *different* partition functions. It's useful at this stage to review the logical development to see what these are, working with the quantum-mechanical particle as an example, so that everything is collected in one place.
The fundamental object is the Feynman propagator for a real time interval T_R, written for specificity in the basis of position eigenstates:

$$F(x_0, x_1, T_R) = \langle x_1 | e^{-iHT_R/\hbar} | x_0 \rangle \tag{230}$$

We can write this as a path integral following the treatment so far:

$$F(x_0, x_1, T_R) = \int_{\substack{x(0)=x_0 \\ x(T_R)=x_1}} \mathcal{D}x(t) \exp\left\{ \frac{i}{\hbar} \int_0^{T_R} dt \left(\frac{m}{2} \left(\frac{dx}{dt} \right)^2 - V(x(t)) \right) \right\} \tag{231}$$

Let us imagine propagating the particle for an *imaginary time* T_I. The relevant propagator is obtained by replacing $T_R \to -iT_I$ in F:

$$\begin{aligned} Z(x_0, x_1, T_I) &= \langle x_1 | e^{-HT_I/\hbar} | x_0 \rangle \\ &= F(x, x', -iT_I) \\ &= \int_{\substack{x(0)=x_0 \\ x(-iT_I)=x_1}} \mathcal{D}x(t) \exp\left\{ \frac{i}{\hbar} \int_0^{-iT_I} dt \left(\frac{m}{2} \left(\frac{dx}{dt} \right)^2 - V(x(t)) \right) \right\}. \end{aligned} \tag{232}$$

Now, let's change variables and define $\tau = it$. Then, we see that

$$Z(x_0, x_1, T_I) = \int_{\substack{x(0)=x_0 \\ x(T_I)=x_1}} \mathcal{D}x(\tau) \exp\left\{ -\frac{1}{\hbar} \int_0^{T_I} d\tau \left(\frac{m}{2} \left(\frac{dx}{d\tau} \right)^2 + V(x(\tau)) \right) \right\}. \tag{233}$$

Note that the conditions on x in the path integral are now written for $x(\tau)$ rather than $x(t)$. The expression in the exponential is the Euclidean action defined previously. We can now use (233) in two ways.

Statement I: We can link F (via Z) to a classical statistical mechanical problem. Define a classical energy $E_{cl}[x(t)] = \frac{1}{\delta} \int_0^{T_I} d\tau \left(\frac{m}{2} \left(\frac{dx}{d\tau} \right)^2 + V(x(\tau)) \right)$. This is the energy of an elastic string of length T_I and tension m, vibrating in one dimension are captured by coordinate $x(\tau)$, and which has its ends fixed at $x(0) = x_0$ and $x(T_I) = x_1$. Here δ is an arbitrary constant with units of time, since the integral manifestly has units of action=energy×time. Thus, we have (writing $T_I = T$)

$$F(x_0, x_1, -iT) = Z(x_0, x_1, T) = \int_{\substack{x(0)=x_0 \\ x(T)=x_1}} e^{-\frac{\delta}{\hbar} E_{\rm cl}[x(\tau)]} \tag{234}$$

Notice that the choice of δ is arbitrary – the physical quantity is the dimensionless ratio $\frac{\delta}{\hbar} E_{\rm cl}[x(\tau)]$. However *given* some choice of 'energy scale', one can interpret the RHS as the partition function of a classical statistical mechanical problem at temperature $T_{cl} = \hbar/\delta k_B$). [In the main part of the notes, we took a slightly different route in which we also discretized the integrals. This introduced a similar small time scale ϵ (see (225)), with the arbitrariness being the number of slices N, but we were able to take a sensible $N \to \infty$ limit. Note also that it is very tempting to just take $\delta = T_I$ so that there is proportionality between β_{cl} and the total imaginary time T_I. However, this choice would challenge the interpretation of E_{cl} as a classical energy, since it would then no longer be extensive in the 'length'

parameter T_I. So, it is better to let δ remain finite as we take $T_I \to \infty$ so that the classical statistical mechanical analogy really makes sense.]
We have thus demonstrated that *the transition amplitude for a quantum particle over an imaginary time $-iT_I$ is equal to the classical partition function for a string of length T_I at ends fixed at x_0 and x_1, at a classical temperature $\propto 1/\hbar$.* Note that the classical problem has $D+1$ dimensions, since the time direction of the quantum problem has become a classical 'spatial' coordinate. This is OK, because classical statistical mechanics does not have any notion of time — we average over ensembles since we have assumed (by the ergodic hypothesis) that this is the same as a time average — and once you average over a quantity, the theory does not depend on it anymore.

Statement II: We can also link F to the partition function of a quantum statistical mechanical problem (this was the thing we did *first* in the lecture notes). Since the partition function of a quantum statistical mechanical prooblem with Hamiltonian H and inverse temperature β is $Z_{\rm qu}(\beta) = {\rm Tr} e^{-\beta H}$, we see that

$$Z_{\rm qu}(\beta) = \int dx \langle x|e^{-\beta H}|x\rangle = \int dx F(x,x,-i\beta\hbar). \tag{235}$$

We have already seen how this leads to a path integral over periodic paths. So, we have argued that the *quantum mechanical partition function at temperature β is given by the probability amplitude for the particle to return to its original position in an imaginary time $\beta\hbar$.*

The above discussion should make it clear that there are really two separate statements being made about the transition amplitude. To summarize: $F(x,x',-iT)$ captures two different statistical mechanical problems depending on how we choose to interpret the imaginary time parameter T. We can view T as a 'length' of classical system with fixed end points at some effective classical temperature $T_{cl} \propto \beta^{-1}$, and thereby identify a classical statistical mechanical problem with one extra dimension, or we can set $T = \beta\hbar$ take the trace of F to obtain the quantum partition function.
What we have done in effect in the preceding discussion in these notes is to *combine* these two pictures. From (234) and (235) we have

$$Z_{\rm qu}(\beta) = \int dx F(x,x,-i\beta\hbar) = \int dx Z(x,x,\beta\hbar) = Z_{cl}^{\beta\hbar}(T_{cl} \propto \hbar^{-1}) \tag{236}$$

where on the RHS because of the trace we now have the classical partition function for a string that obeys periodic boundary conditions and finite extent $\beta\hbar$ in the 'extra' spatial direction corresponding to imaginary time. Since none of our results relied on starting with a single QM particle, we can extend them to arbitrary D. We have therefore formally related a quantum statistical mechanical partition function for a D dimensional quantum system to the classical statistical mechanical partition function for a $D+1$ dimensional problem.
The quantum classical correspondence is subtle, and at this point it is best to view as a beautiful formal step. To quote Alexandre Polyakov (arguably the greatest living quantum field theorist): *Our derivation of these analogies was technical. I feel that there are deeper reasons for them, connected with the properties of space-time. [...] At the moment our aims are more modest – we are going to exploit these analogies in concrete problems.* — A.M. Polyakov, in *Gauge Fields and Strings.*

5 The Ising Model

Ferromagnetism is an interesting phenomenon in solids. Some metals (like Fe or Ni) are observed to acquire a finite magnetization below a certain temperature. Ferromagnetism is a fundamentally quantum mechanical effect, and arises when electron spins spontaneously align along a certain direction. The Ising model is a

very crude attempt to model this phenomenon. It is defined as follows. We have a lattice in D dimensions with N sites. On each site j of this lattice sits a "spin" variable σ_j, which can take the two values ± 1. These are referred to as "spin-up" and "spin-down" respectively. A given set $\{\sigma_1, \sigma_2, \ldots, \sigma_N\}$ specifies a *configuration*. The corresponding energy is taken to be of the form

$$E(\{\sigma_j\}) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_{j=1}^{N} \sigma_j \ , \tag{237}$$

where $\langle ij \rangle$ denote nearest-neighbour bonds on our lattice and $J > 0$. The first term favours alignment on neighbouring spins, while h is like an applied magnetic field. Clearly, when $h = 0$ the lowest energy states are obtained by choosing all spins to be either up or down. The question of interest is whether the Ising model displays a finite temperature phase transition between a ferromagnetically ordered phase at low temperatures, and a paramagnetic phase at high temperatures.

5.1 Statistical mechanics of the Ising model

The partition function of the model is

$$Z = \sum_{\sigma_1 = \pm 1} \sum_{\sigma_2 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} e^{-\beta E(\{\sigma_j\})}. \tag{238}$$

The *magnetization per site* is given by

$$m(h) = \frac{1}{N} \langle \sum_{j=1}^{N} \sigma_j \rangle_\beta = \frac{1}{N\beta} \frac{\partial}{\partial h} \ln(Z). \tag{239}$$

The *magnetic susceptibility* is defined as

$$\chi(h) = \frac{\partial m(h)}{\partial h} = \frac{1}{N\beta} \frac{\partial^2}{\partial h^2} \ln(Z). \tag{240}$$

Substituting the expression (238) for the partition function and then carrying out the derivatives it can be expressed in the form

$$\chi(h) = \frac{\beta}{N} \sum_{l,m=1}^{N} \langle \sigma_l \sigma_m \rangle_\beta - \langle \sigma_l \rangle_\beta \langle \sigma_m \rangle_\beta. \tag{241}$$

5.2 The One-Dimensional Ising Model

The simplest case is when our lattice is one-dimensional, and we impose periodic boundary conditions. The energy then reads

$$E = \sum_{j=1}^{N} \left[-J \sigma_j \sigma_{j+1} - \frac{h}{2} (\sigma_j + \sigma_{j+1}) \right] \equiv \sum_{j=1}^{N} E(\sigma_j, \sigma_{j+1}), \tag{242}$$

where we have defined

$$\sigma_{N+1} = \sigma_1. \tag{243}$$

The partition function can be expressed in the form

$$Z = \sum_{\sigma_1, \ldots, \sigma_N} \prod_{j=1}^{N} e^{-\beta E(\sigma_j, \sigma_{j+1})}. \tag{244}$$

It can be evaluated exactly by means of the *transfer matrix method*.

5.2.1 Transfer matrix approach

The general idea is to rewrite Z as a product of matrices. The transfer matrix T is taken to be a 2×2 matrix with elements

$$T_{\sigma\sigma'} = e^{-\beta E(\sigma,\sigma')}. \tag{245}$$

Its explicit form is

$$T = \begin{pmatrix} T_{++} & T_{+-} \\ T_{-+} & T_{--} \end{pmatrix} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix}. \tag{246}$$

The partition function can be expressed in terms of the transfer matrix as follows

$$Z = \sum_{\sigma_1,\ldots,\sigma_N} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \ldots T_{\sigma_{N-1}\sigma_N} T_{\sigma_N\sigma_1} \tag{247}$$

As desired, this has the structure of a matrix multiplication

$$\boxed{Z = \mathrm{Tr}\left(T^N\right).} \tag{248}$$

The trace arises because we have imposed periodic boundary conditions. As T is a real symmetric matrix, it can be diagonalized, i.e.

$$U^\dagger T U = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}, \tag{249}$$

where U is a unitary matrix and

$$\lambda_\pm = e^{\beta J}\cosh(\beta h) \pm \sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}}. \tag{250}$$

Using the cyclicity of the trace and $UU^\dagger = I$, we have

$$Z = \mathrm{Tr}\left(UU^\dagger T^N\right) = \mathrm{Tr}\left(U^\dagger T^N U\right) = \mathrm{Tr}\left([U^\dagger T U]^N\right) = \mathrm{Tr}\begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} = \lambda_+^N + \lambda_-^N. \tag{251}$$

But as $\lambda_- < \lambda_+$ we have

$$Z = \lambda_+^N \left(1 + \left[\frac{\lambda_-}{\lambda_+}\right]^N\right) = \lambda_+^N \left(1 + e^{-N\ln(\lambda_+/\lambda_-)}\right). \tag{252}$$

So for large N, which is the case we are interested in, we have with exponential accuracy

$$\boxed{Z \simeq \lambda_+^N.} \tag{253}$$

Given the partition function, we can now easily calculate the magnetization per site

$$m(h) = \frac{1}{N\beta}\frac{\partial}{\partial h}\ln(Z). \tag{254}$$

In Fig. 14 we plot $m(h)$ as a function of inverse temperature $\beta = 1/k_B T$ for two values of magnetic field h. We see that for non-zero h, the magnetization per site takes its maximum value $m = 1$ at low temperatures. At high temperatures it goes to zero. This is as expected, as at low T the spins align along the direction of the applied field. However, as we decrease the field, the temperature below which $m(h)$ approaches unity decreases. In the limit $h \to 0$, the magnetization per site vanishes at all finite temperatures. Hence there is no phase transition to a ferromagnetically ordered state in the one dimensional Ising model.

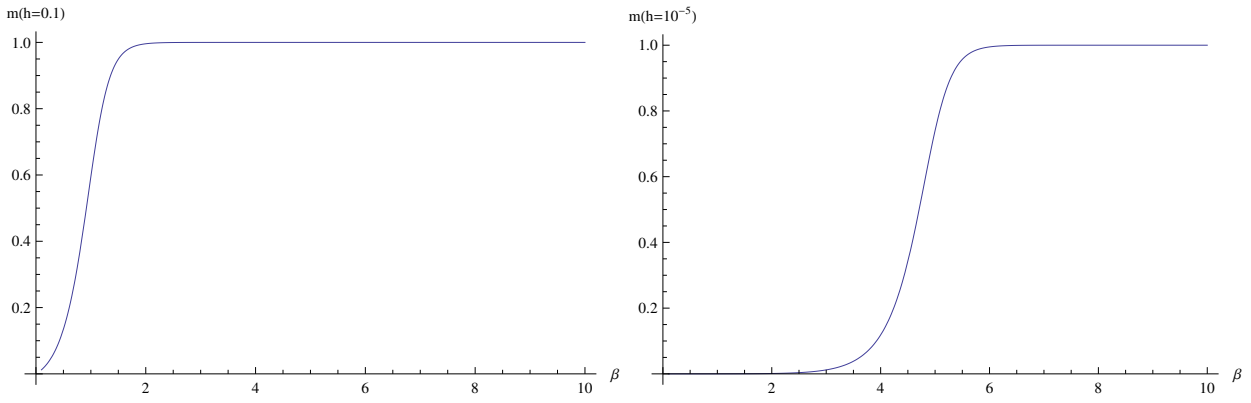


Figure 14: Magnetization per site as a function of inverse temperature for two values of applied magnetic field. We see that when we reduce the magnetic field, the temperature region in which the magnetization is essentially zero grows.

5.2.2 Averages of observables in the transfer matrix formalism

The average magnetization at site j is

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \sum_{\sigma_1,\ldots,\sigma_N} \sigma_j e^{-\beta E(\{\sigma_j\})}. \tag{255}$$

We can express this in terms of the transfer matrix as

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \sum_{\sigma_1,\ldots,\sigma_N} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \ldots T_{\sigma_{j-1}\sigma_j} \sigma_j T_{\sigma_j\sigma_{j+1}} \ldots T_{\sigma_N\sigma_1}. \tag{256}$$

Using that

$$\left(T\sigma^z\right)_{\sigma_{j-1}\sigma_j} = T_{\sigma_{j-1}\sigma_j}\sigma_j \ , \tag{257}$$

where $\sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the Pauli matrix, we obtain

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z} \mathrm{Tr}\left[T^{j-1}\sigma^z T^{N-j+1}\right] = \frac{1}{Z}\mathrm{Tr}\left[T^N \sigma^z\right]. \tag{258}$$

Diagonalizing T by means of a unitary transformation as before, this becomes

$$\langle \sigma_j \rangle_\beta = \frac{1}{Z}\mathrm{Tr}\left[U^\dagger T^N U U^\dagger \sigma^z U\right] = \frac{1}{Z}\mathrm{Tr}\left[\begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} U^\dagger \sigma^z U\right] \ . \tag{259}$$

The matrix U is given in terms of the normalized eigenvectors of T

$$T|\pm\rangle = \lambda_\pm |\pm\rangle \tag{260}$$

as

$$U = (|+\rangle, |-\rangle). \tag{261}$$

For $h = 0$ we have

$$U|_{h=0} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{262}$$

This gives

$$\langle \sigma_j \rangle_\beta \Big|_{h=0} = 0. \tag{263}$$

For general h the expression is more complicated

$$U = \begin{pmatrix} \frac{\alpha_+}{\sqrt{1+\alpha_+^2}} & \frac{\alpha_-}{\sqrt{1+\alpha_-^2}} \\ \frac{1}{\sqrt{1+\alpha_+^2}} & \frac{1}{\sqrt{1+\alpha_-^2}} \end{pmatrix} \quad , \quad \alpha_\pm = e^{2\beta J}\sinh(\beta h) \pm \sqrt{1+e^{4\beta J}\sinh^2(\beta h)} \; . \tag{264}$$

The magnetization per site in the thermodynamic limit is then

$$\lim_{N\to\infty}\langle\sigma_j\rangle_\beta = \lim_{N\to\infty}\frac{\left(\frac{\alpha_+^2-1}{\alpha_+^2+1}\right)\lambda_+^N + \left(\frac{\alpha_-^2-1}{\alpha_-^2+1}\right)\lambda_-^N}{\lambda_+^N+\lambda_-^N} = \left(\frac{\alpha_+^2-1}{\alpha_+^2+1}\right). \tag{265}$$

This now allows us to prove, that in the one dimensional Ising model there is no phase transition at any finite temperature:

$$\boxed{\lim_{h\to 0}\lim_{N\to\infty}\langle\sigma_j\rangle_\beta = 0 \; , \quad \beta<\infty.} \tag{266}$$

Note the order of the limits here: we first take the infinite volume limit at finite h, and only afterwards take h to zero. This procedure allows for spontaneous symmetry breaking to occur, but the outcome of our calculation is that the spin reversal symmetry remains unbroken at any finite temperature.

Similarly, we find

$$\langle\sigma_j\sigma_{j+r}\rangle_\beta = \frac{1}{Z}\mathrm{Tr}\left[T^{j-1}\sigma^z T^r \sigma^z T^{N+1-j-r}\right] = \frac{1}{Z}\mathrm{Tr}\left[U^\dagger\sigma^z U\begin{pmatrix}\lambda_+^r & 0\\ 0 & \lambda_-^r\end{pmatrix}U^\dagger\sigma^z U\begin{pmatrix}\lambda_+^{N-r} & 0\\ 0 & \lambda_-^{N-r}\end{pmatrix}\right]. \tag{267}$$

We can evaluate this for zero field $h=0$

$$\langle\sigma_j\sigma_{j+r}\rangle_\beta\Big|_{h=0} = \frac{\lambda_+^{N-r}\lambda_-^r + \lambda_-^{N-r}\lambda_+^r}{\lambda_+^N+\lambda_-^N} \approx \left[\frac{\lambda_-}{\lambda_+}\right]^r = e^{-r/\xi}. \tag{268}$$

So in zero field the two-point function decays exponentially with *correlation length*

$$\xi = \frac{1}{\ln\coth(\beta J)}. \tag{269}$$

5.2.3 The related zero-dimensional quantum model

The 1D classical Ising model is related to a 0D quantum mechanical system as follows. Given the discussion leading to (226), we are looking to write the transfer matrix of the Ising model in the form

$$T = \sqrt{\frac{c}{\epsilon}}\; e^{-\epsilon H_Q/\hbar}, \tag{270}$$

where c is a constant with dimension of time. Our transfer matrix can be written as

$$T = e^{\beta J}\cosh(\beta h)I + e^{\beta J}\sinh(\beta h)\sigma^z + e^{-\beta J}\sigma^x \tag{271}$$

We see that if we tune h to zero such that for large β

$$e^{-2\beta J} = \frac{\epsilon K}{\hbar} \; , \quad \beta h e^{2\beta J} = \lambda = \text{fixed}, \tag{272}$$

then

$$T = \sqrt{\frac{\hbar}{\epsilon K}}\left(I + \frac{\epsilon K}{\hbar}\left[\lambda\sigma^z + \sigma^x\right] + \mathcal{O}(\epsilon^2)\right). \tag{273}$$

We conclude that the one dimensional classical Ising model is related to the quantum mechanics of a single spin-1/2 with Hamiltonian

$$H_Q = -K\sigma^x - K\lambda\sigma^z. \tag{274}$$

5.3 The Two-Dimensional Ising Model

We now turn to the 2D Ising model on a square lattice with periodic boundary conditions. The spin variables have now two indices corresponding to rows and columns of the square lattice respectively

$$\sigma_{j,k} = \pm 1 \ , \quad j,k = 1,\ldots,N. \tag{275}$$

The boundary conditions are $\sigma_{k,N+1} = \sigma_{k,1}$ and $\sigma_{N+1,j} = \sigma_{1,j}$, which correspond to the lattice "living" on

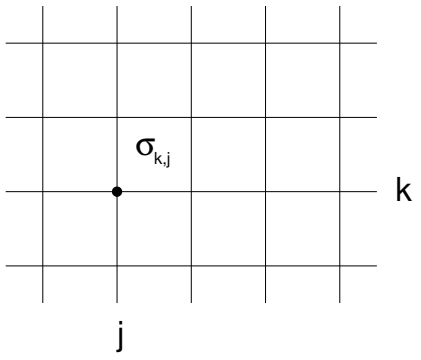


Figure 15: Ising model on the square lattice.

the surface of a torus. The energy in zero field is

$$E(\{\sigma_{k,j}\}) = -J \sum_{j,k} \sigma_{k,j}\sigma_{k,j+1} + \sigma_{k,j}\sigma_{k+1,j}. \tag{276}$$

5.3.1 Transfer Matrix Method

The partition function is given by

$$Z = \sum_{\{\sigma_{j,k}\}} e^{-\beta E(\{\sigma_{k,j}\})}. \tag{277}$$

The idea of the transfer matrix method is again to write this in terms of matrix multiplications. The difference to the one dimensional case is that the transfer matrix will now be much larger. We start by expressing the partition function in the form

$$Z = \sum_{\{\sigma_{j,k}\}} e^{-\beta \sum_{k=1}^{N} E(k;k+1)}, \tag{278}$$

where

$$E(k;k+1) = -J \sum_{j=1}^{N} \sigma_{k,j}\sigma_{k+1,j} + \frac{1}{2}\left[\sigma_{k,j}\sigma_{k,j+1} + \sigma_{k+1,j}\sigma_{k+1,j+1}\right]. \tag{279}$$

This energy depends only on the configurations of spins on rows k and $k+1$, i.e. on spins $\sigma_{k,1},\ldots,\sigma_{k,N}$ and $\sigma_{k+1,1},\ldots,\sigma_{k+1,N}$. Each configuration of spins on a given row specifies a sequence $s_1, s_2, \ldots, s_N$ with $s_j = \pm 1$. Let us associate a vector

$$|\boldsymbol{s}\rangle \tag{280}$$

with each such sequence. By construction there 2^N such vectors. We then define a scalar product on the space spanned by these vectors by

$$\langle \boldsymbol{t}|\boldsymbol{s}\rangle = \prod_{j=1}^{N} \delta_{t_j,s_j}. \tag{281}$$

With this definition, the vectors $\{|\boldsymbol{s}\rangle\}$ form an orthonormal basis of a 2^N dimensional linear vector space. In particular we have

$$I = \sum_{\boldsymbol{s}} |\boldsymbol{s}\rangle\langle\boldsymbol{s}|. \tag{282}$$

Finally, we define a $2^N \times 2^N$ *transfer matrix* T by

$$\langle\boldsymbol{\sigma}_k|T|\boldsymbol{\sigma}_{k+1}\rangle = e^{-\beta E(k;k+1)}. \tag{283}$$

The point of this construction is that the partition function can now be written in the form

$$Z = \sum_{\boldsymbol{\sigma}_1}\sum_{\boldsymbol{\sigma}_2}\cdots\sum_{\boldsymbol{\sigma}_N}\langle\boldsymbol{\sigma}_1|T|\boldsymbol{\sigma}_2\rangle\langle\boldsymbol{\sigma}_2|T|\boldsymbol{\sigma}_3\rangle\ldots\langle\boldsymbol{\sigma}_{N-1}|T|\boldsymbol{\sigma}_N\rangle\langle\boldsymbol{\sigma}_N|T|\boldsymbol{\sigma}_1\rangle \tag{284}$$

We now may use (282) to carry out the sums over spins, which gives

$$\boxed{Z = \mathrm{Tr}\left[T^N\right]}, \tag{285}$$

where the trace is over our basis $\{|\boldsymbol{s}\rangle|s_j = \pm 1\}$ of our 2^N dimensional vector space. Like in the 1D case, thermodynamic properties involve only the largest eigenvalues of T. Indeed, we have

$$Z = \sum_{j=1}^{2^N} \lambda_j^N \ , \tag{286}$$

where λ_j are the eigenvalues of T. The free energy is then

$$F = -k_B T \ln(Z) = -k_B T \ln\left[\lambda_{\max}^N \sum_{j=1}^{2^N}\left(\frac{\lambda_j}{\lambda_{\max}}\right)^N\right] = -k_B T N \ln(\lambda_{\max}) - k_B T \ln\left[\sum_j \left(\frac{\lambda_j}{\lambda_{\max}}\right)^N\right], \tag{287}$$

where $\lambda_{\max}$ is the largest eigenvalue of T, which we assume to be unique. As $|\lambda_j/\lambda_{\max}| < 1$, the second contribution in (287) is bounded by $-k_B T N \ln(2)$, and we see that in the thermodynamic limit the free energy per site is

$$\boxed{f = \lim_{N\to\infty}\frac{F}{N^2} = \lim_{N\to\infty} -\frac{k_B T}{N}\ln(\lambda_{\max}).} \tag{288}$$

Thermodynamic quantities are obtained by taking derivatives of f and hence only involve the largest eigenvalue of T. The main complication we have to deal with is that T is still a very large matrix. This poses the question, why we should bother to use a transfer matrix description anyway? Calculating Z from its basic definition (277) involves a sum with 2^{N^2} terms, i.e. at least 2^{N^2} operations on a computer. Finding the largest eigenvalue of a $M \times M$ matrix involves $\mathcal{O}(M^2)$ operations, which in our case amounts to $\mathcal{O}(2^{2N})$. For large values of N this amounts to an enormous simplification.

5.3.2 Spontaneous Symmetry Breaking

Surprisingly, the transfer matrix of the 2D Ising model can be diagonalized exactly. Unfortunately we don't have the time do go through the somewhat complicated procedure here (but see the Aside below), but the upshot is that the 2D Ising model can be solved *exactly*. Perhaps the most important result is that in the thermodynamic limit the square lattice Ising model has a *finite temperature phase transition* between a paramagnetic and a ferromagnetic phase. The magnetization per site behaves as shown in Fig.16. At low temperatures $T < T_c$ there is a non-zero magnetization per site, even though we did not apply a magnetic field. This is surprising, because our energy (276) is unchanged if we flip all spins

$$\sigma_{j,k} \to -\sigma_{j,k}. \tag{289}$$

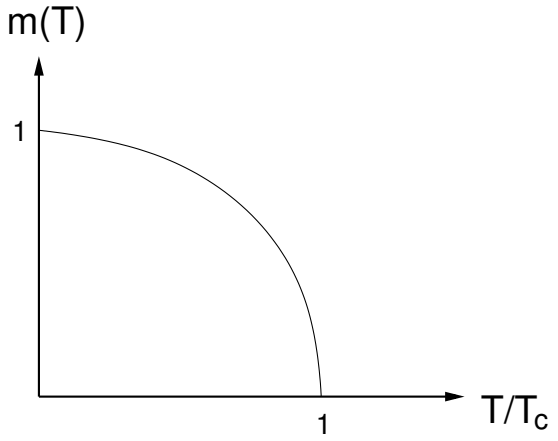


Figure 16: Phase Transition in the square lattice Ising model.

The operation (289) is a discrete (two-fold) symmetry of the Ising model. Because we have translational invariance, the magnetization per site is

$$m = \langle \sigma_{j,k} \rangle_\beta. \tag{290}$$

Hence a non-zero value of m signifies the *spontaneous breaking* of the discrete symmetry (289). In order to describe this effect mathematically, we have to invoke a bit of trickery. Let us consider zero temperature. Then there are exactly two degenerate lowest energy states: the one with all spins $\sigma_{j,k} = +1$ and the one with all spins $\sigma_{j,k} = -1$. We now apply a very small magnetic field to the system, i.e. add a term

$$\delta E = -\epsilon \sum_{j,k} \sigma_{j,k} \tag{291}$$

to the energy. This splits the two states, which now have energies

$$E_\pm = -JN_B \mp \epsilon N \ , \tag{292}$$

where N_B is the number of bonds. The next step is key: we now *define* the thermodynamic limit of the free energy per site as

$$f(T) \equiv \lim_{\epsilon \to 0} \lim_{N \to \infty} \frac{-k_B T \ln(Z)}{N}. \tag{293}$$

The point is that the contributions $Z_\pm = e^{-\beta E_\pm}$ of the two states to Z are such that

$$\frac{Z_-}{Z_+} = e^{-2\epsilon N/k_B T} \ . \tag{294}$$

This goes to zero when we take N to infinity! So in the above sequence of limits, only the state with all spins up contributes to the partition function, and this provides a way of describing spontaneous symmetry breaking! The key to this procedure is that

$$\boxed{\lim_{\epsilon \to 0} \lim_{N \to \infty} Z \neq \lim_{N \to \infty} \lim_{\epsilon \to 0} Z.} \tag{295}$$

The procedure we have outlined above, i.e. introducing a symmetry breaking field, then taking the infinite volume limit, and finally removing the field, is very general and applies to all instances where spontaneous symmetry breaking occurs.

5.3.3 The Related One-Dimensional Quantum Problem

As we have seen, the transfer matrix approach allows up to map between classical systems in $D+1$ dimensions and quantum systems in D dimensions, with the 'extra' dimension in the classical language mapping to the time variable. This mapping is implemented via the transfer matrix. For the 1D classical Ising model the related 0+1D quantum problem was of a single spin. What about the 2D classical Ising model? Recall that T is a $2^N \times 2^N$ matrix, whose matrix elements are given by

$$\langle \boldsymbol{s}|T|\boldsymbol{\sigma}\rangle = e^{-L(\boldsymbol{s},\sigma)} \tag{296}$$

Notice that we have changed our notation very slightly, to emphasize that the energies are specified by the 2^N-vectors $\boldsymbol{s}, \boldsymbol{\sigma}$, each element of which can be seen as a spin, viz. $s(j), \sigma_j \in \{-1, 1\}$. We have also replaced the k label previously used, by letting $\boldsymbol{s}, \boldsymbol{\sigma}$ stand in for $\boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}$. As we have noted above, the set of vectors $\{|\boldsymbol{\sigma}\rangle\}$ with $\sigma_j = \pm 1$ is a basis for a system of 2^N spins-1/2 labeled by index j. If we introduce Pauli matrices $\hat{X}_j, \hat{Y}_j, \hat{Z}_j$ that act on the spin-1/2 labeled by j, then our convention is that $\hat{Z}_j|\boldsymbol{\sigma}\rangle = \sigma_j|\boldsymbol{\sigma}\rangle$, and $\hat{X}_j|\boldsymbol{\sigma}\rangle = |\boldsymbol{\sigma}'^{(j)}\rangle$, where $\boldsymbol{\sigma}'^{(j)}$ is the vector obtained from $\boldsymbol{\sigma}$ by flipping the jth spin: $\sigma_i'^{(j)} = \sigma_i(1-\delta_{ij}) - \sigma_i\delta_{ij}$. (We use the $\hat{X}, \hat{Y}, \hat{Z}$ notation to avoid having too many confusing σs, but you can replace them with the more familiar $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ if it makes you uncomfortable.)

Our aim is to write $T = e^{-\Delta\tau \hat{H}}$, where $\hat{H}$ is a quantum Hamiltonian, and $\Delta\tau$ is a small time interval. (296) specifies the matrix elements of $\hat{H}$ (via $e^{-\Delta\tau\hat{H}}$). We can already see that $\hat{H}$ acts on the 2^N-dimensional Hilbert space of N spins-1/2. Comparing (296) and (283), we have

$$\begin{aligned} L(\boldsymbol{s},\boldsymbol{\sigma}) &= -\beta J \sum_{j=1}^{N} s_j\sigma_j - \frac{\beta J_x}{2}\sum_j s_j s_{j+1} + \sigma_j\sigma_{j+1} \\ &\equiv \text{const.} + \frac{\beta J}{2}(s_j - \sigma_j)^2 - \frac{\beta J_x}{2}\sum_j s_j s_{j+1} + \sigma_j\sigma_{j+1} \end{aligned} \tag{297}$$

Notice that we have also changed the problem very slightly, to allow for different couplings J, J_x in the horizontal and vertical directions (where x is the direction that connects the N spins; the other direction plays the role of time) ; the reasons for this will become clear below.

In order to determine $\hat{H}$, observe that its matrix elements are given by

$$e^{-L(\boldsymbol{s},\sigma)} = T_{\boldsymbol{s}\boldsymbol{\sigma}} = \langle \boldsymbol{s}|T|\boldsymbol{\sigma}\rangle \simeq \langle \boldsymbol{s}|\mathbf{1} - \Delta\tau\hat{H}|\boldsymbol{\sigma}\rangle, \tag{298}$$

where we will always work to leading order in $\Delta\tau$. For this to make sense will adjust parameters J, J_x so that $\Delta\tau$ is small. Since the space of states is very large (after all T is a *big* matrix), we should think about how to organize our calculation of matrix elements. It turns out a particularly convenient approach *count the number of spin flips needed to go from $\boldsymbol{\sigma}$ to $\boldsymbol{s}$.*

First of all, let's consider the case when $\boldsymbol{s} = \boldsymbol{\sigma}$. This means that the two configurations are 0 spin flips apart. Clearly,

$$L(0 \text{ flips}) = -\frac{\beta J_x}{2}\sum_j s_j s_{j+1} + \sigma_j\sigma_{j+1} = -\beta J_x \sum_j \sigma_j\sigma_{j+1}. \tag{299}$$

Next, consider the case where the two configurations are identical except for one spin, which is flipped (let's say it's the i^{th} spin, so that in the notation above, $\boldsymbol{s} = \boldsymbol{\sigma}'^{(i)}$. It follows that

$$L(1 \text{ flip}) = \frac{\beta J}{2}(1-(-1))^2 - \frac{\beta J_x}{2}\sum_j s_j s_{j+1} + \sigma_j\sigma_{j+1} = 2\beta J - \frac{\beta J_x}{2}\sum_j s_j s_{j+1} + \sigma_j\sigma_{j+1}. \tag{300}$$

It is not too hard to see that this generalizes, so that

$$L(n \text{ flips}) = 2n\beta J - \frac{\beta J_x}{2}\sum_j s_j s_{j+1} + \sigma_j\sigma_{j+1}. \tag{301}$$

Returning to (298) but now organising things in terms of spin flips, we have

$$e_{\boldsymbol{s}\boldsymbol{\sigma}}^{-L(0\text{ flips})} = \delta_{\boldsymbol{s}\boldsymbol{\sigma}} e^{\beta J_x \sum_j \sigma_j \sigma_{j+1}} \simeq \langle \boldsymbol{s}|\mathbf{1} - \Delta\tau \hat{H}|\boldsymbol{\sigma}\rangle\Big|_{0\text{ flips}}, \tag{302}$$

$$e_{\boldsymbol{s}\boldsymbol{\sigma}}^{-L(1\text{ flip})} = e^{-2\beta J} e^{\frac{\beta J_x}{2}\sum_j \sigma_j\sigma_{j+1}+s_j s_{j+1}} \simeq -\Delta\tau\ \langle \boldsymbol{s}|\hat{H}|\boldsymbol{\sigma}\rangle\Big|_{1\text{ flip}}, \tag{303}$$

$$e_{\boldsymbol{s}\boldsymbol{\sigma}}^{-L(n\text{ flips})} = e^{-2n\beta J} e^{\frac{\beta J_x}{2}\sum_j \sigma_j\sigma_{j+1}+s_j s_{j+1}} \simeq -\Delta\tau\ \langle \boldsymbol{s}|\hat{H}|\boldsymbol{\sigma}\rangle\Big|_{n\text{ flips}}. \tag{304}$$

where in the second and third line we have recognized that if $\boldsymbol{s}$ and $\boldsymbol{\sigma}$ differ in any way the identity cannot make a contribution. Expanding exponentials and taking matrix elements, while recognizing that $\hat{H}\Big|_{0\text{ flips}}$ must be diagonal, we see that

$$\delta_{\boldsymbol{s}\boldsymbol{\sigma}}(1+\beta J_x \sum_j \sigma_j\sigma_{j+1}) \simeq \delta_{\boldsymbol{s}\boldsymbol{\sigma}}(1-\Delta\tau\langle\boldsymbol{\sigma}|\hat{H}|\boldsymbol{\sigma}\rangle) \leftrightarrow \beta J_x \sim \Delta\tau \text{ and } \langle\boldsymbol{\sigma}|\hat{H}|\boldsymbol{\sigma}\rangle = -\sum_j \sigma_j\sigma_{j+1}. \tag{305}$$

A similar expansion of the 1 flip term though suggests that $e^{-\beta J} \sim \Delta\tau$, and we also have that $\Delta\tau$ should be small. The above are consistent if we make $\Delta\tau$ small by taking $\beta J, \beta J_x \to \infty$ while holding their ratio $g = (\beta J_x)^{-1} e^{-2\beta J}$ fixed. Then, we see that all matrix elements of H between states that differ by n spin flips scale like $(\Delta\tau)^n$ and so we can ignore all those with $n > 1$ since we are working to leading order in $\Delta\tau$. Notice that the scaling limit makes the Ising model very anisotropic – this is the price one pays for a convenient mapping to the classical problem but the power of universality means that this anisotropy will not matter.

The upshot of the above argument is that all we need to do is guess a $\hat{H}$ which is good in the zero and one-flip sectors, and then in the appropriate scaling limit introduced above we will have found the right quantum Hamiltonian. Now, here is a clue: in the one-flip sector, from our discussion above it follows that we must have $|\boldsymbol{s}\rangle = \hat{X}_i|\boldsymbol{\sigma}\rangle$ for some i; allowing i to range over $1,\ldots,N$ for a fixed $\boldsymbol{\sigma}$ we clearly will range over all states connected to $\boldsymbol{\sigma}$ by a single spinflip. Consider the following quantum Hamiltonian:

$$\hat{H}_{\text{TFIM}} = -\bar{J}\sum_{j=1}^{N}\left(g\hat{X}_j + \hat{Z}_j\hat{Z}_{j+1}\right). \tag{306}$$

where $\bar{J}$ is an overall constant (*not* simply related to J) with dimensions of energy (that we can pull out of $\Delta\tau$ by making it dimensionless – this is the freedom implied by the '$\sim$' in our expressions for $\Delta\tau$). Clearly, the diagonal elements of $\hat{H}_{\text{TFIM}}$ satisfy the condition (302) (or equivalently, (305)). What about the off-diagonal ones? Let us take $|\boldsymbol{s}\rangle = |\boldsymbol{\sigma}'^{(i)}\rangle$, so that we are in the 1-flip sector. We then have (using the fact that we are working to leading order in $\Delta\tau$ to expand and/or re-exponentiate as desired)

$$\begin{aligned}
-\Delta\tau\ \langle \boldsymbol{s}|\hat{H}|\boldsymbol{\sigma}\rangle\Big|_{1\text{ flip}} &= \langle\boldsymbol{s}|\left(\bar{J}\Delta\tau\sum_j \hat{Z}_j\hat{Z}_{j+1} + \bar{J}g\Delta\tau\sum_j \hat{X}_j + \right)|\boldsymbol{\sigma}\rangle\Big|_{1\text{ flip}} \\
&= \bar{J}g\Delta\tau\ \langle\boldsymbol{s}|\sum_j \hat{X}_j|\boldsymbol{\sigma}\rangle\Big|_{1\text{ flip}} \\
&\simeq e^{-2\beta J}
\end{aligned} \tag{307}$$

where in the first step we used the fact that $\hat{Z}_j\hat{Z}_{j+1}$ is diagonal and in the last step we defined $\bar{J}\Delta\tau = \beta J_x, \bar{J}g\Delta\tau = e^{-2\beta J}$ (consistent with our requirements above), and used the fact that if $|\boldsymbol{s}\rangle = |\boldsymbol{\sigma}'^{(i)}\rangle$ then $\langle\boldsymbol{s}|\hat{X}_j|\boldsymbol{\sigma}\rangle$ is nonzero only for $j = i$, and every matrix element of H in the single-flip sector is of this form. Now, notice that the final expression we obtain differs from what we want (cf. (303)) by a factor of $e^{\frac{\beta J_x}{2}\sum_j \sigma_j\sigma_{j+1}+s_j s_{j+1}}$ but since $\beta J_x \sim \Delta\tau$ and the overall term goes as $e^{-2\beta J} \sim \Delta\tau$, the discrepancy is $O((\Delta\tau)^2)$ so we can ignore it.

$\hat{H}_{\rm TFIM}$ is known as the"quantum Ising chain" or "1D transverse-field Ising model" (whence TFIM). It is one of the best studied problems in all of many-body quantum physics. It is central to conformal field theory as well. It also has a long history at Oxford, dating to work of Elliot, Pfeuty, and Wood in the 1970s. In another aside, we will use an idea of Pfeuty's (that uses the so-called Jordan-Wigner mapping) to map (306) to a model of free fermions. Here 'free' means that the result is quadratic in the fermionic creation and annihilation operators, which makes the model exactly solvable.

5.4 Homework Questions 8-10

Question 8. A lattice model for non-ideal gas is defined as follows. The sites i of a lattice may be empty or occupied by at most one atom, and the variable n_i takes the values $n_i = 0$ and $n_i = 1$ in the two cases. There is an attractive interaction energy J between atoms that occupy neighbouring sites, and a chemical potential μ. The model Hamiltonian is

$$\mathcal{H} = -J \sum_{\langle ij \rangle} n_i n_j - \mu \sum_i n_i \,, \tag{308}$$

where $\sum_{\langle ij \rangle}$ is a sum over neighbouring pairs of sites.
(a) Describe briefly how the *transfer matrix method* may be used to calculate the statistical-mechanical properties of one-dimensional lattice models with short range interactions. Illustrate your answer by explaining how the partition function for a one-dimensional version of the lattice gas, Eq. (1), defined on a lattice of N sites with periodic boundary conditions, may be evaluated using the matrix

$$\mathbf{T} = \begin{pmatrix} 1 & \mathrm{e}^{\beta\mu/2} \\ \mathrm{e}^{\beta\mu/2} & \mathrm{e}^{\beta(J+\mu)} \end{pmatrix} .$$

(b) Derive an expression for $\langle n_i \rangle$ in the limit $N \to \infty$, in terms of elements of the eigenvectors of this matrix.
(c) Show that

$$\langle n_i \rangle = \frac{1}{1 + \mathrm{e}^{-2\theta}} \,,$$

where

$$\sinh(\theta) = \exp(\beta J/2) \sinh(\beta[J+\mu]/2) \,.$$

Sketch $\langle n_i \rangle$ as a function of μ for $\beta J \gg 1$, and comment on the physical significance of your result.

Question 9. The one-dimensional 3-state Potts model is defined as follows. At lattice sites $i = 0, 1, \ldots, L$ "spin" variables σ_i take integer values $\sigma_i = 1, 2, 3$. The Hamiltonian is then given by

$$H = -J \sum_{i=0}^{L-1} \delta_{\sigma_i, \sigma_{i+1}}, \tag{309}$$

where $\delta_{a,b}$ is the Kronecker delta, $J > 0$.
(a) What are the ground states and first excited states for this model?
(b) Write down the transfer matrix for (309). Derive an expression for the free energy per site f in the limit of large L in terms of the transfer matrix eigenvalues. Show that vectors of the form $(1, z, z^2)$ with $z^3 = 1$ are eigenvectors, and hence find the corresponding eigenvalues. Show that at temperature T (with $\beta = 1/k_B T$) and in the limit $L \to \infty$

$$f = -k_B T \ln\left(3 + e^{\beta J} - 1\right). \tag{310}$$

(c) The boundary variable σ_0 is fixed in the state $\sigma_0 = 1$. Derive an expression (for large L), that the variable at site $\ell \gg 1$ is in the same state, in terms of the transfer matrix eigenvalues and eigenvectors. Show that your result has the form

$$\langle \delta_{\sigma_\ell, 1} \rangle = \frac{1}{3} + \frac{2}{3} e^{-\ell/\xi}. \tag{311}$$

How does ξ behave in the low and high temperature limits?

Question 10. Consider a one dimensional Ising model on an open chain with N sites, where N is odd. On all even sites a magnetic field $2h$ is applied, see Fig. 17. The energy is

$$E = -J \sum_{j=1}^{N-1} \sigma_j \sigma_{j+1} + 2h \sum_{j=1}^{(N-1)/2} \sigma_{2j}. \tag{312}$$

(a) Show that the partition function can be written in the form

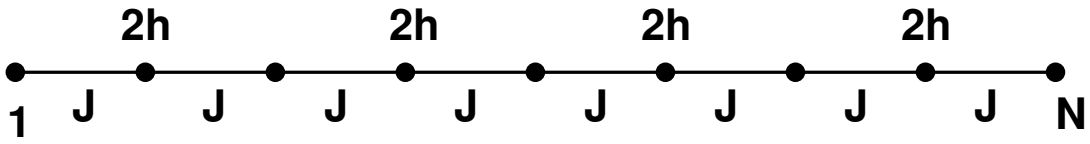


Figure 17: Open Ising chain with magnetic field applied to all even sites.

$$Z = \langle u | T^{(N-1)/2} | v \rangle \ , \tag{313}$$

where T is an appropriately constructed transfer matrix, and $|u\rangle$ and $|v\rangle$ two dimensional vectors. Give explicit expressions for T, $|u\rangle$ and $|v\rangle$.
(b) Calculate Z for the case $h = 0$.

Part III
Many-Particle Quantum Mechanics

In the basic QM course you encountered only quantum systems with *very small numbers* of particles. In the harmonic oscillator problem we are dealing with a single QM particle, when solving the hydrogen atom we had one electron and one nucleus. Perhaps the most important field of application of quantum physics is to systems of *many* particles. Examples are the electronic degrees of freedom in solids, superconductors, trapped ultra-cold atomic gases, magnets and so on. The methods you have encountered in the basic QM course are not suitable for studying such problems. In this part of the course we introduce a framework, that will allow us to study the QM of many-particle systems. This new way of looking at things will also reveal very interesting connections to *Quantum Field Theory*.

6 Second Quantization

The formalism we develop in the following is known as *second quantization*.

6.1 Systems of Independent Particles

You already know from second year QM how to solve problems involving *independent particles*

$$H = \sum_{j=1}^{N} H_j \tag{314}$$

where H_j is the Hamiltonian on the j'th particle, e.g.

$$H_j = \frac{\hat{\mathbf{p}}_j^2}{2m} + V(\hat{\mathbf{r}}_j) = -\frac{\hbar^2}{2m}\nabla_j^2 + V(\hat{\mathbf{r}}_j). \tag{315}$$

The key to solving such problems is that $[H_j, H_l] = 0$. We'll now briefly review the necessary steps, switching back and forth quite freely between using states and operators acting on them, and the position representation of the problem (i.e. looking at wave functions).

- Step 1. Solve the single-particle problem

  $$H_j|\phi_l\rangle = E_l|\phi_l\rangle \ . \tag{316}$$

  The corresponding wave functions are

  $$\phi_l(\mathbf{r}_j) = \langle \mathbf{r}_j|\phi_l\rangle. \tag{317}$$

  The eigenstates form an orthonormal set

  $$\langle\phi_l|\phi_m\rangle = \delta_{l,m} = \int d^D\mathbf{r}_j \ \phi_l^*(\mathbf{r}_j)\phi_m(\mathbf{r}_j). \tag{318}$$

- Step 2. Form N-particle eigenfunctions as *products*

  $$\left(\sum_{j=1}^{N} H_j\right) \phi_{l_1}(\mathbf{r}_1)\phi_{l_2}(\mathbf{r}_2)\dots\phi_{l_N}(\mathbf{r}_N) = \left(\sum_{j=1}^{N} E_{l_j}\right) \phi_{l_1}(\mathbf{r}_1)\phi_{l_2}(\mathbf{r}_2)\dots\phi_{l_N}(\mathbf{r}_N) \ . \tag{319}$$

  This follows from the fact that in the position representation H_j is a differential operator that acts only on the j'th position $\mathbf{r}_j$. The corresponding eigenstates are tensor products

  $$|l_1\rangle \otimes |l_2\rangle \otimes \dots \otimes |l_N\rangle. \tag{320}$$

- Step 3. Impose the appropriate *exchange symmetry* for indistinguishable particles, e.g.

$$\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \left[\phi_l(\mathbf{r}_1)\phi_m(\mathbf{r}_2) \pm \phi_l(\mathbf{r}_2)\phi_m(\mathbf{r}_1)\right] \ , \quad l \neq m. \tag{321}$$

Generally we require

$$\psi(\dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots) = \pm\psi(\dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots) \ , \tag{322}$$

where the + sign corresponds to bosons and the − sign to fermions. This is achieved by taking

$$\boxed{\psi_{l_1 \dots l_N}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathcal{N} \sum_{P \in S_N} (\pm 1)^{|P|} \phi_{l_{P_1}}(\mathbf{r}_1) \dots \phi_{l_{P_N}}(\mathbf{r}_N),} \tag{323}$$

where the sum is over all permutations of $(1, 2, \dots, N)$ and $|P|$ is the number of pair exchanges required to reduce $(P_1, \dots, P_N)$ to $(1, \dots, N)$. The normalization constant $\mathcal{N}$ is

$$\mathcal{N} = \frac{1}{\sqrt{N! n_1! n_2! \dots}} \ , \tag{324}$$

where n_j is the number of times j occurs in the set $\{l_1, \dots, l_N\}$. For fermions the wave functions can be written as *Slater determinants*

$$\psi_{l_1 \dots l_N}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \phi_{l_1}(\mathbf{r}_1) & \dots & \phi_{l_1}(\mathbf{r}_N) \\ \vdots & & \vdots \\ \phi_{l_N}(\mathbf{r}_1) & \dots & \phi_{l_N}(\mathbf{r}_N) \end{pmatrix} . \tag{325}$$

The states corresponding to (323) are

$$\boxed{|l_1, \dots, l_N\rangle = \mathcal{N} \sum_{P \in S_N} (\pm 1)^{|P|} |l_{P_1}\rangle \otimes \dots \otimes |l_{P_N}\rangle \ .} \tag{326}$$

6.1.1 Occupation Number Representation

By construction the states have the symmetry

$$|l_{Q_1} \dots l_{Q_N}\rangle = \pm |l_1 \dots l_N\rangle \ , \tag{327}$$

where Q is an arbitrary permutation of $(1, \dots, N)$. As the overall sign of state is irrelevant, we can therefore choose them without loss of generality as

$$|\underbrace{1 \dots 1}_{n_1} \underbrace{2 \dots 2}_{n_2} \underbrace{3 \dots 3}_{n_3} 4 \dots\rangle \equiv |n_1 n_2 n_3 \dots\rangle. \tag{328}$$

In (328) we have as many n_j's as there are single-particle eigenstates, i.e. $\dim H$ [4]. For fermions we have $n_j = 0, 1$ only as a consequence of the *Pauli principle*. The representation (328) is called *occupation number representation*. The n_j's tell us how many particles are in the single-particle state $|j\rangle$. By construction the states $\{|n_1 n_2 n_3 \dots\rangle | \sum_j n_j = N\}$ form an orthonormal basis of our N-particle problem

$$\langle m_1 m_2 m_3 \dots | n_1 n_2 n_3 \dots \rangle = \prod_j \delta_{n_j, m_j} \ , \tag{329}$$

where we have defined $\langle m_1 m_2 m_3 \dots | = | m_1 m_2 m_3 \dots \rangle^\dagger$.

[4] Note that this is different from the particle number N.

6.2 Fock Space

We now want to allow the particle number to vary. The main reason for doing this is that we will encounter physical problems where particle number is in fact not conserved. Another motivation is that experimental probes like photoemission change particle number, and we want to be able to describe these. The resulting space of states is called *Fock Space*.

1. The state with no particles is called the *vacuum state* and is denoted by $|0\rangle$.
2. N-particle states are $|n_1 n_2 n_3 \dots\rangle$ with $\sum_j n_j = N$.

6.2.1 Creation and Annihilation Operators

Given a basis of our space of states we can define operators by specifying their action on all basis states.

- particle creation operators with quantum number l

$$\boxed{c_l^\dagger |n_1 n_2 \dots\rangle = \begin{cases} 0 & \text{if } n_l = 1 \text{ for fermions} \\ \sqrt{n_l+1}(\pm 1)^{\sum_{j=1}^{l-1} n_j} |n_1 n_2 \dots n_l + 1 \dots\rangle & \text{else.} \end{cases}} \tag{330}$$

  Here the + (−) sign applies to bosons (fermions).

- particle annihilation operators with quantum number l

$$\boxed{c_l |n_1 n_2 \dots\rangle = \sqrt{n_l}(\pm 1)^{\sum_{j=1}^{l-1} n_j} |n_1 n_2 \dots n_l - 1 \dots\rangle \ .} \tag{331}$$

We note that (331) follows from (330) by

$$\langle m_1 m_2 \dots | c_l^\dagger | n_1 n_2 \dots\rangle^* = \langle n_1 n_2 \dots | c_l | m_1 m_2 \dots\rangle \ . \tag{332}$$

The creation and annihilation operators fulfil *canonical (anti)commutation relations*

$$\boxed{[c_l, c_m] = 0 = [c_l^\dagger, c_m^\dagger] \ , \quad [c_l, c_m^\dagger] = \delta_{l,m} \qquad \text{bosons,}} \tag{333}$$

$$\boxed{\{c_l, c_m\} = c_l c_m + c_m c_l = 0 = \{c_l^\dagger, c_m^\dagger\} \ , \quad \{c_l, c_m^\dagger\} = \delta_{l,m} \qquad \text{fermions.}} \tag{334}$$

7: Proof of the anticommutation relation

Let us see how to prove these in the fermionic case. For $l < m$ we have

$$\begin{aligned} c_l^\dagger c_m |\dots n_l \dots n_m \dots\rangle &= c_l^\dagger \sqrt{n_m}(-1)^{\sum_{j=1}^{m-1} n_j} |\dots n_l \dots n_m - 1 \dots\rangle \\ &= \sqrt{n_l+1}\sqrt{n_m}(-1)^{\sum_{j=l}^{m-1} n_j} |\dots n_l + 1 \dots n_m - 1 \dots\rangle. \end{aligned} \tag{335}$$

Similarly we have

$$c_m c_l^\dagger |\dots n_l \dots n_m \dots\rangle = \sqrt{n_l+1}\sqrt{n_m}(-1)^{1+\sum_{j=l}^{m-1} n_j} |\dots n_l + 1 \dots n_m - 1 \dots\rangle. \tag{336}$$

This means that for any basis state $|n_1 n_2 \dots\rangle$ we have

$$\{c_l^\dagger, c_m\} |n_1 n_2 \dots\rangle = 0 \ , \qquad \text{if } l > m. \tag{337}$$

This implies that

$$\{c_l^\dagger, c_m\} = 0 \ , \qquad \text{if } l > m. \tag{338}$$

The case $l < m$ works in the same way. This leaves us with the case $l = m$. Here we have

$$c_l^\dagger c_l |\dots n_l \dots n_m \dots\rangle = c_l^\dagger \sqrt{n_l} (-1)^{\sum_{j=1}^{l-1} n_j} |\dots n_l - 1 \dots\rangle = n_l |\dots n_l \dots\rangle. \tag{339}$$

$$\begin{aligned} c_l c_l^\dagger |\dots n_l \dots\rangle &= \begin{cases} c_l \sqrt{n_l + 1} (-1)^{\sum_{j=1}^{l-1} n_j} |\dots n_l + 1 \dots\rangle & \text{if } n_l = 0 \ , \\ 0 & \text{if } n_l = 1 \ , \end{cases} \\ &= \begin{cases} |\dots n_l \dots\rangle & \text{if } n_l = 0 \ , \\ 0 & \text{if } n_l = 1 \ , \end{cases} \end{aligned} \tag{340}$$

Combining these we find that

$$\{c_l^\dagger, c_l\} |\dots n_l \dots\rangle = |\dots n_l \dots\rangle \ , \tag{341}$$

and as the states $|\dots n_l \dots\rangle$ form a basis this implies

$$\{c_l^\dagger, c_l\} = 1. \tag{342}$$

Note that here 1 really means the identity operator $\mathbf{1}$.

6.2.2 Basis of the Fock Space

We are now in a position to write down our Fock space basis in a very convenient way.

- Fock vacuum (state without any particles)

$$|0\rangle. \tag{343}$$

- Single-particle states

$$|0 \dots 0 \underbrace{1}_{l} 0 \dots\rangle = c_l^\dagger |0\rangle \ . \tag{344}$$

- N-particle states

$$|n_1 n_2 \dots\rangle = \prod_j \frac{1}{\sqrt{n_j!}} \left(c_j^\dagger\right)^{n_j} |0\rangle \ . \tag{345}$$

6.3 Homework Questions 11-13

Question 11. Consider a fermion 'system' with just one single-particle orbital, so that the only states of the system are $|0\rangle$ (unoccupied) and $|1\rangle$ (occupied). Show that we can represent the operators a and $a^\dagger$ by the matrices

$$a^\dagger = \begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix}, \quad a = \begin{pmatrix} 0 & C^* \\ 0 & 0 \end{pmatrix}.$$

You can do this by checking the values of aa, $a^\dagger a^\dagger$ and $a^\dagger a + a a^\dagger$. What values may the constant C take?

Question 12. A quantum-mechanical Hamiltonian for a system of an even number N of point unit masses interacting by nearest-neighbour forces in one dimension is given by

$$H = \frac{1}{2} \sum_{r=1}^{N} \left(p_r^2 + (q_{r+1} - q_r)^2\right),$$

where the Hermitian operators q_r, p_r satisfy the commutation relations $[q_r, q_s] = [p_r, p_s] = 0$, $[q_r, p_s] = \mathrm{i}\delta_{rs}$, and where $q_{r+N} = q_r$. New operators Q_k, P_k are defined by

$$q_r = \frac{1}{\sqrt{N}} \sum_k Q_k \mathrm{e}^{\mathrm{i}kr} \qquad \text{and} \qquad p_r = \frac{1}{\sqrt{N}} \sum_k P_k \mathrm{e}^{-\mathrm{i}kr},$$

where $k = 2\pi n/N$ with $n = -N/2 + 1, \ldots, 0, \ldots, N/2$.

Show that:

(a) $Q_k = \dfrac{1}{\sqrt{N}} \displaystyle\sum_{s=1}^{N} q_s \mathrm{e}^{-\mathrm{i}ks}$ and $P_k = \frac{1}{\sqrt{N}} \sum_{s=1}^{N} p_s \mathrm{e}^{\mathrm{i}ks}$

(b) $[Q_k, P_{k'}] = \mathrm{i}\delta_{kk'}$

(c) $H = \frac{1}{2}\Big(\sum_k P_k P_{-k} + \omega^2 Q_k Q_{-k} \Big)$, where $\omega^2 = 2(1 - \cos k)$.

Similarly to the treatment of the simple harmonic oscillator in QM I we then define annihilation operators a_k by

$$a_k = \frac{1}{(2\omega_k)^{1/2}} (\omega_k Q_k + \mathrm{i} P_{-k}).$$

Show that the Hermitian conjugate operators are

$$a_k^\dagger = \frac{1}{(2\omega_k)^{1/2}} (\omega_k Q_{-k} - \mathrm{i} P_k),$$

and determine the canonical commutation relations for a_k and $a_p^\dagger$. Construct the Fock space of states and determine the eigenstates and eigenvalues of H.

Question 13. Bosonic creation operators are defined through their action on basis states in the occupation number representation as

$$c_l^\dagger |n_1 n_2 \ldots\rangle = \sqrt{n_l + 1} |n_1 n_2 \ldots n_l + 1 \ldots\rangle \ , \tag{346}$$

a) Deduce from this how bosonic annihilation operators act.
b) Show that the creation and annihilation operators fulfil *canonical commutation relations*

$$[c_l, c_m] = 0 = [c_l^\dagger, c_m^\dagger] \ , \quad [c_l, c_m^\dagger] = \delta_{l,m}. \tag{347}$$

6.3.1 Change of Basis

The Fock space is built from a given basis of single-particle states

$$\boxed{\text{single-particle states } |l\rangle} \longrightarrow \boxed{\text{N-particle states } |n_1 n_2 \ldots\rangle} \longrightarrow \boxed{\text{Fock Space}}. \tag{348}$$

You know from second year QM that it is often convenient to switch from one basis to another, e.g. from energy to momentum eigenstates. This is achieved by a unitary transformation

$$\{|l\rangle\} \longrightarrow \{|\alpha\rangle\} \ , \tag{349}$$

where

$$|\alpha\rangle = \sum_l \underbrace{\langle l|\alpha\rangle}_{U_{l\alpha}} |l\rangle. \tag{350}$$

By construction

$$\sum_\alpha U_{l\alpha} U_{\alpha m}^\dagger = \sum_\alpha \langle l|\alpha\rangle\langle\alpha|m\rangle = \langle l|m\rangle = \delta_{lm}. \tag{351}$$

We now want to "lift" this unitary transformation to the level of the Fock space. We know that

$$\begin{aligned} |l\rangle &= c_l^\dagger|0\rangle \ , \\ |\alpha\rangle &= d_\alpha^\dagger|0\rangle \ . \end{aligned} \tag{352}$$

On the other hand we have

$$|\alpha\rangle = \sum_l U_{l\alpha}|l\rangle = \sum_l U_{l\alpha} c_l^\dagger|0\rangle. \tag{353}$$

This suggests that we take

$$\boxed{d_\alpha^\dagger = \sum_l U_{l\alpha} c_l^\dagger,} \tag{354}$$

and this indeed reproduces the correct transformation for N-particle states. Taking the hermitian conjugate we obtain the transformation law for annihilation operators

$$\boxed{d_\alpha = \sum_l U_{\alpha l}^\dagger c_l.} \tag{355}$$

We emphasize that these transformation properties are compatible with the (anti)commutation relations (as they must be). For fermions

$$\{d_\alpha, d_\beta^\dagger\} = \sum_{l,m} U_{\alpha l}^\dagger U_{m\beta} \underbrace{\{c_l, c_m^\dagger\}}_{\delta_{l,m}} = \sum_l U_{\alpha l}^\dagger U_{l\beta} = (U^\dagger U)_{\alpha\beta} = \delta_{\alpha,\beta}. \tag{356}$$

6.4 Second Quantized Form of Operators

In the next step we want to know how observables such as H, P, X etc act on the Fock space.

6.4.1 Occupation number operators

These are the simplest hermitian operators we can build from c_l and $c_m^\dagger$. They are defined as

$$\hat{n}_l \equiv c_l^\dagger c_l. \tag{357}$$

From the definition of c_l and $c_l^\dagger$ it follows immediately that

$$\hat{n}_l|n_1 n_2 \dots\rangle = n_l|n_1 n_2 \dots\rangle. \tag{358}$$

6.4.2 Single-particle operators

Single-particle operators are of the form

$$\hat{O} = \sum_j \hat{o}_j \ , \tag{359}$$

where the operator $\hat{o}_j$ acts only on the j'th particle. Examples are kinetic and potential energy operators

$$\hat{T} = \sum_j \frac{\hat{p}_j^2}{2m} \ , \qquad \hat{V} = \sum_j V(\hat{x}_j). \tag{360}$$

We want to represent $\hat{O}$ on the Fock space built from single-particle eigenstates $|\alpha\rangle$. We do this in two steps:

- Step 1: We first represent $\hat{O}$ on the Fock space built from the eigenstates of $\hat{o}$

$$\hat{o}|l\rangle = \lambda_l |l\rangle = \lambda_l c_l^\dagger |0\rangle. \tag{361}$$

Then, when acting on an N-particle state (326), we have

$$\hat{O}|l_1, l_2, \ldots, l_N\rangle = \left[\sum_{j=1}^{N} \lambda_j\right] |l_1, l_2, \ldots, l_N\rangle. \tag{362}$$

This is readily translated into the occupation number representation

$$\hat{O}|n_1 n_2 \ldots\rangle = \left[\sum_k n_k \lambda_k\right] |n_1 n_2 \ldots\rangle. \tag{363}$$

As $|n_1 n_2 \ldots\rangle$ constitute a basis, this together with (358) imply that we can represent $\hat{O}$ in the form

$$\hat{O} = \sum_k \lambda_k \hat{n}_k = \sum_k \lambda_k c_k^\dagger c_k. \tag{364}$$

- Step 2: Now that we have a representation of $\hat{O}$ on the Fock space built from the single-particle states $|l\rangle$, we can use a basis transformation to the basis $\{|\alpha\rangle\}$ to obtain a representation on a general Fock space. Using that $\langle k|\hat{O}|k'\rangle = \delta_{k,k'}\lambda_k$ we can rewrite (364) in the form

$$\hat{O} = \sum_{k,k'} \langle k'|\hat{O}|k\rangle c_{k'}^\dagger c_k. \tag{365}$$

Then we apply our general rules for a change of single-particle basis of the Fock space

$$c_k^\dagger = \sum_\alpha U_{\alpha k}^\dagger d_\alpha^\dagger. \tag{366}$$

This gives

$$\hat{O} = \sum_{\alpha,\beta} \underbrace{\sum_{k'} \left(\langle k'|U_{\alpha k'}^\dagger\right)}_{\langle\alpha|} \hat{O} \underbrace{\sum_k \left(U_{k\beta}|k\rangle\right)}_{|\beta\rangle} d_\alpha^\dagger d_\beta. \tag{367}$$

where we have used that

$$|k\rangle = \sum_\alpha U_{\alpha k}^\dagger |\alpha\rangle \ . \tag{368}$$

This gives us the final result

$$\boxed{\hat{O} = \sum_{\alpha,\beta} \langle\alpha|\hat{O}|\beta\rangle \ d_\alpha^\dagger d_\beta.} \tag{369}$$

We now work out a number of explicit examples of Fock space representations for single-particle operators.

1. Momentum Operators $\mathbf{P}$ in the infinite volume:

(i) Let us first consider $\mathbf{P}$ in the single-particle basis of momentum eigenstates

$$\hat{\mathbf{P}}|\mathbf{k}\rangle = \mathbf{k}|\mathbf{k}\rangle \ , \quad \langle\mathbf{p}|\mathbf{k}\rangle = (2\pi\hbar)^3 \delta^{(3)}(\mathbf{p} - \mathbf{k}). \tag{370}$$

8: Remark

These are shorthand notations for

$$\hat{P}_a|k_x,k_y,k_z\rangle = k_a|k_x,k_y,k_z\rangle \ , \quad a=x,y,z. \tag{371}$$

and

$$\langle p_x,p_y,p_z|k_x,k_y,k_z\rangle = (2\pi\hbar)^3\delta(k_x-p_x)\delta(k_y-p_y)\delta(k_z-p_z) \ . \tag{372}$$

Using our general result for representing single-particle operators in a Fock space built from their eigenstates (364) we have

$$\hat{\mathbf{P}} = \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\mathbf{p}c^\dagger(\mathbf{p})c(\mathbf{p}) \ , \qquad [c^\dagger(\mathbf{k}),c(\mathbf{p})\} = (2\pi\hbar)^3\delta^{(3)}(\mathbf{p}-\mathbf{k}). \tag{373}$$

Here we have introduced a notation

$$[c^\dagger(\mathbf{k}),c(\mathbf{p})\} = \begin{cases} c^\dagger(\mathbf{k})c(\mathbf{p}) - c(\mathbf{p})c^\dagger(\mathbf{k}) & \text{for bosons} \\ c^\dagger(\mathbf{k})c(\mathbf{p}) + c(\mathbf{p})c^\dagger(\mathbf{k}) & \text{for fermions.} \end{cases} \tag{374}$$

(ii) Next we want to represent $\hat{\mathbf{P}}$ in the single-particle basis of position eigenstates

$$\hat{\mathbf{X}}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle \ , \quad \langle\mathbf{x}|\mathbf{x}'\rangle = \delta^{(3)}(\mathbf{x}-\mathbf{x}'). \tag{375}$$

Our general formula (369) gives

$$\hat{\mathbf{P}} = \int d^3\mathbf{x}d^3\mathbf{x}' \ \langle\mathbf{x}'|\hat{\mathbf{P}}|\mathbf{x}\rangle c^\dagger(\mathbf{x}')c(\mathbf{x}) \ . \tag{376}$$

We can simplify this by noting that

$$\langle\mathbf{x}'|\hat{\mathbf{P}}|\mathbf{x}\rangle = -i\hbar\nabla_{\mathbf{x}'}\delta^{(3)}(\mathbf{x}-\mathbf{x}'), \tag{377}$$

which allows us to eliminate three of the integrals

$$\hat{\mathbf{P}} = \int d^3\mathbf{x}d^3\mathbf{x}' \left[-i\hbar\nabla_{\mathbf{x}'}\delta^{(3)}(\mathbf{x}-\mathbf{x}')\right] c^\dagger(\mathbf{x}')c(\mathbf{x}) = \int d^3\mathbf{x}c^\dagger(\mathbf{x})\left(-i\hbar\nabla_{\mathbf{x}}\right)c(\mathbf{x}). \tag{378}$$

2. Single-particle Hamiltonian:

$$H = \sum_{j=1}^{N}\frac{\hat{\mathbf{p}}_j^2}{2m} + V(\hat{\mathbf{x}}_j). \tag{379}$$

(i) Let us first consider H in the single-particle basis of energy eigenstates $H|l\rangle = E_l|l\rangle$, $|l\rangle = c_l^\dagger|0\rangle$. Our result (364) tells us that

$$H = \sum_l E_l c_l^\dagger c_l. \tag{380}$$

(ii) Next we consider the position representation, i.e. we take position eigenstates $|\mathbf{x}\rangle = c^\dagger(\mathbf{x})|0\rangle$ as a basis of single-particle states. Then by (369)

$$H = \int d^3\mathbf{x}d^3\mathbf{x}' \ \langle\mathbf{x}'|H|\mathbf{x}\rangle \ c^\dagger(\mathbf{x}')c(\mathbf{x}). \tag{381}$$

Substituting (379) into (381) and using

$$\langle\mathbf{x}'|V(\hat{\mathbf{x}})|\mathbf{x}\rangle = V(\mathbf{x})\delta^{(3)}(\mathbf{x}-\mathbf{x}') \ , \quad \langle\mathbf{x}'|\hat{\mathbf{p}}^2|\mathbf{x}\rangle = -\hbar^2\nabla^2\delta^{(3)}(\mathbf{x}-\mathbf{x}') \ , \tag{382}$$

we arrive at the position representation

$$\boxed{H = \int d^3\mathbf{x}\; c^\dagger(\mathbf{x}) \left[-\frac{\hbar^2\nabla^2}{2m} + V(\mathbf{x}) \right] c(\mathbf{x}).} \tag{383}$$

(iii) Finally we consider the momentum representation, i.e. we take momentum eigenstates $|\mathbf{p}\rangle = c^\dagger(\mathbf{p})|0\rangle$ as a basis of single-particle states. Then by (369)

$$H = \int \frac{d^3\mathbf{p} d^3\mathbf{p}'}{(2\pi\hbar)^6}\; \langle \mathbf{p}'|H|\mathbf{p}\rangle\; c^\dagger(\mathbf{p}')c(\mathbf{p}). \tag{384}$$

Matrix elements of the kinetic energy operator are simple

$$\langle \mathbf{p}'|\hat{\mathbf{p}}^2|\mathbf{p}\rangle = \mathbf{p}^2\langle \mathbf{p}'|\mathbf{p}\rangle = \mathbf{p}^2(2\pi\hbar)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}'). \tag{385}$$

Matrix elements of the potential can be calcuated as follows

$$\begin{aligned} \langle \mathbf{p}'|\hat{V}|\mathbf{p}\rangle &= \int d^3\mathbf{x} d^3\mathbf{x}'\; \langle \mathbf{p}'|\mathbf{x}'\rangle\langle \mathbf{x}'|\hat{V}|\mathbf{x}\rangle\langle \mathbf{x}|\mathbf{p}\rangle = \int d^3\mathbf{x} d^3\mathbf{x}' \quad \underbrace{\langle \mathbf{x}'|\hat{V}|\mathbf{x}\rangle}_{V(\mathbf{x})\delta^{(3)}(\mathbf{x}-\mathbf{x}')} \quad e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{x} - \frac{i}{\hbar}\mathbf{p}'\cdot\mathbf{x}'} \\ &= \int d^3\mathbf{x}\; V(\mathbf{x}) e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} = \tilde{V}(\mathbf{p}-\mathbf{p}'), \end{aligned} \tag{386}$$

where $\tilde{V}(\mathbf{p})$ is essentially the *three-dimensional Fourier transform* of the (ordinary) function $V(\mathbf{x})$. Hence

$$\boxed{H = \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\; \frac{\mathbf{p}^2}{2m} c^\dagger(\mathbf{p})c(\mathbf{p}) + \int \frac{d^3\mathbf{p} d^3\mathbf{p}'}{(2\pi\hbar)^6}\; \tilde{V}(\mathbf{p}-\mathbf{p}')c^\dagger(\mathbf{p}')c(\mathbf{p}).} \tag{387}$$

6.4.3 Two-particle operators

These are operators that act on two particles at a time. A good example is the interaction potential $V(\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2)$ between two particles at positions $\mathbf{r}_1$ and $\mathbf{r}_2$. For N particles we want to consider

$$\hat{V} = \sum_{i<j}^{N} V(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j). \tag{388}$$

On the Fock space built from single-particle position eigenstates this is represented as

$$\boxed{\hat{V} = \frac{1}{2}\int d^3\mathbf{r} d^3\mathbf{r}'\; c^\dagger(\mathbf{r})c^\dagger(\mathbf{r}')V(\mathbf{r},\mathbf{r}')c(\mathbf{r}')c(\mathbf{r}).} \tag{389}$$

Note that when writing down the first quantized expression (388), we assumed that the operators acts specifically on states with N particles. On the other hand, (389) acts on the Fock space, i.e. on states where the particle number can take any value. The action of (389) on N-particle states (where N is fixed but arbitrary) is equal to the action of (388).

9: Derivation of (389)

Let us concentrate on the fermionic case. The bosonic case can be dealt with analogously. We start with our original representation of N-particle states (326)

$$|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \mathcal{N}\sum_{P\in S_N}(-1)^{|P|}|\mathbf{r}_1\rangle\otimes\ldots|\mathbf{r}_N\rangle\ . \tag{390}$$

Then

$$\hat{V}|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \sum_{i<j}V(\mathbf{r}_i,\mathbf{r}_j)|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \frac{1}{2}\sum_{i\neq j}V(\mathbf{r}_i,\mathbf{r}_j)|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle\ . \tag{391}$$

On the other hand we know that

$$|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \prod_{j=1}^{N}c^\dagger(\mathbf{r}_j)|0\rangle. \tag{392}$$

Now consider

$$c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = c(\mathbf{r})\prod_{j=1}^{N}c^\dagger(\mathbf{r}_j)|0\rangle = [c(\mathbf{r}),\prod_{j=1}^{N}c^\dagger(\mathbf{r}_j)\}|0\rangle\ , \tag{393}$$

where is the last step we have used that $c(\mathbf{r})|0\rangle = 0$, and $[A,B\}$ is an anticommutator if both A and B involve an odd number of fermions and a commutator otherwise.
In our case we have a commutator for even N and an anticommutator for odd N.
By repeatedly adding and subtracting terms we find that

$$\begin{aligned}[c(\mathbf{r}),\prod_{j=1}^{N}c^\dagger(\mathbf{r}_j)\} \;&=\; \{c(\mathbf{r}),c^\dagger(\mathbf{r}_1)\}\prod_{j=2}^{N}c^\dagger(\mathbf{r}_j) - c^\dagger(\mathbf{r}_1)\{c(\mathbf{r}),c^\dagger(\mathbf{r}_2)\}\prod_{j=3}^{N}c^\dagger(\mathbf{r}_j)\\ &\quad +\ldots+\prod_{j=1}^{N-1}c^\dagger(\mathbf{r}_j)\{c(\mathbf{r}),c^\dagger(\mathbf{r}_N)\}.\end{aligned} \tag{394}$$

Using that $\{c(\mathbf{r}),c^\dagger(\mathbf{r}_j)\} = \delta^{(3)}(\mathbf{r}-\mathbf{r}_j)$ we then find

$$c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \sum_{n=1}^{N}(-1)^{n-1}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)\prod_{j\neq n}^{N}c^\dagger(\mathbf{r}_j)|0\rangle = \sum_{n=1}^{N}(-1)^{n-1}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)|\mathbf{r}_1\ldots\overbrace{\mathbf{r}_n}^{\text{missing}}\ldots\mathbf{r}_N\rangle. \tag{395}$$

Hence

$$\underbrace{c^\dagger(\mathbf{r}')c(\mathbf{r}')}_{\text{number op.}}c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \sum_{n=1}^{N}(-1)^{n-1}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)\sum_{m\neq n}^{N}\delta^{(3)}(\mathbf{r}'-\mathbf{r}_m)\ |\mathbf{r}_1\ldots\overbrace{\mathbf{r}_n}^{\text{missing}}\ldots\mathbf{r}_N\rangle, \tag{396}$$

and finally

$$c^\dagger(\mathbf{r})c^\dagger(\mathbf{r}')c(\mathbf{r}')c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \sum_{n=1}^{N}\delta^{(3)}(\mathbf{r}-\mathbf{r}_n)\sum_{m\neq n}^{N}\delta^{(3)}(\mathbf{r}'-\mathbf{r}_m)\ |\mathbf{r}_1\ldots\mathbf{r}_n\ldots\mathbf{r}_N\rangle. \tag{397}$$

This implies that

$$\frac{1}{2}\int d^3\mathbf{r}d^3\mathbf{r}'\ V(\mathbf{r},\mathbf{r}')\ c^\dagger(\mathbf{r})c^\dagger(\mathbf{r}')c(\mathbf{r}')c(\mathbf{r})|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle = \frac{1}{2}\sum_{n\neq m}V(\mathbf{r}_n,\mathbf{r}_m)|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle. \tag{398}$$

As $\{|\mathbf{r}_1,\ldots,\mathbf{r}_N\rangle\}$ form a basis, this establishes (389).

Using our formula for basis transformations (354)

$$c^\dagger(\mathbf{r}) = \sum_l \langle l|\mathbf{r}\rangle \; c_l^\dagger, \tag{399}$$

we can transform (389) into a general basis. We have

$$\hat{V} \;=\; \frac{1}{2}\sum_{ll'mm'} \int d^3\mathbf{r}d^3\mathbf{r}' \; V(\mathbf{r},\mathbf{r}')\langle l|\mathbf{r}\rangle\langle l'|\mathbf{r}'\rangle\langle \mathbf{r}'|m'\rangle\langle \mathbf{r}|m\rangle c_l^\dagger c_{l'}^\dagger c_{m'}c_m \; . \tag{400}$$

We can rewrite this by using that the action of $\hat{V}$ on two-particle states is obtained by taking $N=2$ in (388), which tells us that $\hat{V}|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle = V(\mathbf{r},\mathbf{r}')|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle$. This implies

$$\begin{aligned} V(\mathbf{r},\mathbf{r}')\langle l|\mathbf{r}\rangle\langle l'|\mathbf{r}'\rangle\langle \mathbf{r}'|m'\rangle\langle \mathbf{r}|m\rangle \;&=\; V(\mathbf{r},\mathbf{r}')\left[\langle l|\otimes\langle l'|\right]\left[|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle|\right]\left[\langle \mathbf{r}|\otimes\langle \mathbf{r}'|\right]\left[|m\rangle\otimes|m'\rangle|\right] \\ &=\; \left[\langle l|\otimes\langle l'|\right]\hat{V}\left[|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle|\right]\left[\langle \mathbf{r}|\otimes\langle \mathbf{r}'|\right]\left[|m\rangle\otimes|m'\rangle|\right] \end{aligned} \tag{401}$$

Now we use that

$$\int d^3\mathbf{r}d^3\mathbf{r}' \left[|\mathbf{r}\rangle\otimes|\mathbf{r}'\rangle\right]\left[\langle \mathbf{r}|\otimes\langle \mathbf{r}'|\right] = \mathbf{1} \tag{402}$$

to obtain

$$\hat{V} \;=\; \frac{1}{2}\sum_{l,l',m,m'} \left[\langle l|\otimes\langle l'|\right]|\hat{V}|\left[|m\rangle\otimes|m'\rangle\right] c_l^\dagger c_{l'}^\dagger c_{m'}c_m. \tag{403}$$

Finally we can express everything in terms of states with the correct exchange symmetry

$$|mm'\rangle = \frac{1}{\sqrt{2}}\left[|m\rangle\otimes|m'\rangle \pm |m'\rangle\otimes|m\rangle\right] \qquad (m\neq m'). \tag{404}$$

in the form

$$\boxed{\hat{V} = \sum_{(ll'),(mm')} \langle ll'|\hat{V}|mm'\rangle c_l^\dagger c_{l'}^\dagger c_{m'}c_m \; .} \tag{405}$$

Here the sums are over a basis of 2-particle states. In order to see that (403) is equal to (405) observe that

$$\sum_{m,m'}[|m\rangle\otimes|m'\rangle]c_{m'}c_m = \frac{1}{2}\sum_{m,m'}[|m\rangle\otimes|m'\rangle \pm |m'\rangle\otimes|m\rangle]c_{m'}c_m = \frac{1}{\sqrt{2}}\sum_{m,m'}|mm'\rangle \; c_{m'}c_m \tag{406}$$

Here the first equality follows from relabelling summation indices $m\leftrightarrow m'$ and using the (anti)commutation relations between c_m and $c_{m'}$ to bring them back in the right order. The second equality follows from the definition of 2-particle states $|mm'\rangle$. Finally we note that because $|mm'\rangle = \pm|m'm\rangle$ (the minus sign is for fermions) we have

$$\frac{1}{\sqrt{2}}\sum_{m,m'}|mm'\rangle \; c_{m'}c_m = \sqrt{2}\sum_{(mm')}|mm'\rangle \; c_{m'}c_m, \tag{407}$$

where the sum is now over a basis of 2-particle states with the appropriate exchange symmetry. The representation (405) generalizes to arbitrary two-particle operators $\mathcal{O}$.

6.5 Homework Question 14

Question 14. Consider the N-particle interaction potential

$$\hat{V} = \sum_{i<j}^{N} V(\hat{\mathbf{r}}_i,\hat{\mathbf{r}}_j),$$

where $V(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j) = V(\hat{\mathbf{r}}_j, \hat{\mathbf{r}}_i)$. Show that in second quantization it is expressed as

$$\hat{V} = \frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \; V(\mathbf{r}, \mathbf{r}') \; c^\dagger(\mathbf{r}) c^\dagger(\mathbf{r}') c(\mathbf{r}') c(\mathbf{r}).$$

To do so consider the action of $\hat{V}$ on a basis of N-particle position eigenstates

$$|\mathbf{r}_1 \dots \mathbf{r}_N\rangle = \frac{1}{\sqrt{N! n_1! n_2! \dots}} \sum_P (\pm 1)^{|P|} |\mathbf{r}_1\rangle \otimes |\mathbf{r}_2\rangle \otimes |\mathbf{r}_N\rangle = \frac{1}{\sqrt{n_1! n_2! \dots}} \prod_{j=1}^{N} c^\dagger(\mathbf{r}_j)|0\rangle \; ,$$

where n_j is the occupation number of the j^{th} single-particle state. Argue that in an arbitrary basis of single-particle eigenstates $|l\rangle$ $\hat{V}$ can be expressed in the form

$$\hat{V} = \sum_{ll'mm'} \langle ll'|\hat{V}|mm'\rangle c_l^\dagger c_{l'}^\dagger c_{m'} c_m.$$

7 Application I: The Ideal Fermi Gas

Consider an ideal gas of spin-1/2 fermions. The creation operators in the momentum representation (in the infinite volume) are

$$c_\sigma^\dagger(\mathbf{p}) \; , \quad \sigma = \uparrow, \downarrow \; . \tag{408}$$

They fulfil canonical anticommutation relations

$$\{c_\sigma(\mathbf{p}), c_\tau(\mathbf{k})\} = 0 = \{c_\sigma^\dagger(\mathbf{p}), c_\tau^\dagger(\mathbf{k})\} \; , \qquad \{c_\sigma(\mathbf{p}), c_\tau^\dagger(\mathbf{k})\} = \delta_{\sigma,\tau} (2\pi\hbar)^3 \delta^{(3)}(\mathbf{k} - \mathbf{p}). \tag{409}$$

The Hamiltonian, in the grand canonical ensemble, is

$$H - \mu \hat{N} = \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \underbrace{\left[\frac{\mathbf{p}^2}{2m} - \mu \right]}_{\epsilon(\mathbf{p})} \sum_{\sigma=\uparrow,\downarrow} c_\sigma^\dagger(\mathbf{p}) c_\sigma(\mathbf{p}). \tag{410}$$

Here $\mu > 0$ is the chemical potential. As $c_\sigma^\dagger(\mathbf{p}) c_\sigma(\mathbf{p}) = \hat{n}_\sigma(\mathbf{p})$ is the number operator for spin-σ fermions with momentum $\mathbf{p}$, we can easily deduce the action of the Hamiltonian on states in the Fock space:

$$\begin{aligned} \left[H - \mu\hat{N}\right] |0\rangle &= 0 \; , \\ \left[H - \mu\hat{N}\right] c_\sigma^\dagger(\mathbf{p})|0\rangle &= \epsilon(\mathbf{p}) \; c_\sigma^\dagger(\mathbf{p})|0\rangle \; , \\ \left[H - \mu\hat{N}\right] \prod_{j=1}^{n} c_{\sigma_j}^\dagger(\mathbf{p}_j)|0\rangle &= \left[\sum_{k=1}^{n} \epsilon(\mathbf{p}_k)\right] \prod_{j=1}^{n} c_{\sigma_j}^\dagger(\mathbf{p}_j)|0\rangle \; . \end{aligned} \tag{411}$$

7.1 Quantization in a large, finite volume

In order to construct the ground state and low-lying excitations, it is convenient to work with a discrete set of momenta. This is achieved by considering the gas in a large, periodic box of linear size L. Momentum eigenstates are obtained by solving the eigenvalue equation e.g. in the position representation

$$\hat{\mathbf{p}} \psi_{\mathbf{k}}(\mathbf{r}) = -i\hbar \nabla \psi_{\mathbf{k}}(\mathbf{r}) = \mathbf{k} \psi_{\mathbf{k}}(\mathbf{r}). \tag{412}$$

The solutions are plane waves

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{r}}. \tag{413}$$

Imposing periodic boundary conditions ($\mathbf{e}_a$ is the unit vector in a direction)

$$\psi_{\mathbf{k}}(\mathbf{r} + L\mathbf{e}_a) = \psi_{\mathbf{k}}(\mathbf{r}) \quad \text{for } a = x, y, z, \tag{414}$$

gives quantization conditions for the momenta $\mathbf{k}$

$$e^{\frac{i}{\hbar}Lk_a} = 1 \Rightarrow k_a = \frac{2\pi\hbar n_a}{L} \ , \quad a = x, y, z. \tag{415}$$

To summarize, in a large, periodic box the momenta are quantized as

$$\mathbf{k} = \frac{2\pi\hbar}{L}\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \tag{416}$$

Importantly, we can now normalize the eigenstates to 1, i.e.

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{L^{\frac{3}{2}}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{r}}. \tag{417}$$

Hence

$$\langle \mathbf{k}|\mathbf{k}'\rangle = \int d^3\mathbf{r}\psi^*_{\mathbf{k}}(\mathbf{r})\psi_{\mathbf{k}'}(\mathbf{r}) = \delta_{\mathbf{k},\mathbf{k}'}. \tag{418}$$

As a consequence of the different normalization of single-particle states, the anticommutation relations of creation/annihilation operators are changed and now read

$$\{c_\sigma(\mathbf{p}), c_\tau(\mathbf{k})\} = 0 = \{c^\dagger_\sigma(\mathbf{p}), c^\dagger_\tau(\mathbf{k})\} \ , \qquad \{c^\dagger_\sigma(\mathbf{p}), c_\tau(\mathbf{k})\} = \delta_{\sigma,\tau}\delta_{\mathbf{k},\mathbf{p}}. \tag{419}$$

The Hamiltonian is

$$\boxed{H - \mu\hat{N} = \sum_{\mathbf{p}} \epsilon(\mathbf{p}) \sum_{\sigma=\uparrow,\downarrow} c^\dagger_\sigma(\mathbf{p})c_\sigma(\mathbf{p}).} \tag{420}$$

We define a *Fermi momentum* by

$$\frac{p_F^2}{2m} = \mu. \tag{421}$$

7.1.1 Ground State

Then the lowest energy state is obtained by *filling all negative energy single-particle states*, i.e.

$$\boxed{|\mathrm{GS}\rangle = \prod_{|\mathbf{p}|<p_F,\sigma} c^\dagger_\sigma(\mathbf{p})|0\rangle.} \tag{422}$$

The ground state energy is

$$E_{\mathrm{GS}} = \sum_\sigma \sum_{|\mathbf{p}|<p_F} \epsilon(\mathbf{p}). \tag{423}$$

This is extensive (proportional to the volume) as expected. You can see the advantage of working in a finite volume: the product in (422) involves only a finite number of factors and the ground state energy is finite. The ground state momentum is

$$P_{\mathrm{GS}} = \sum_\sigma \sum_{|\mathbf{p}|<p_F} \mathbf{p} = 0. \tag{424}$$

The ground state momentum is zero, because is a state with momentum $\mathbf{p}$ contributes to the sum, then so does the state with momentum $-\mathbf{p}$.

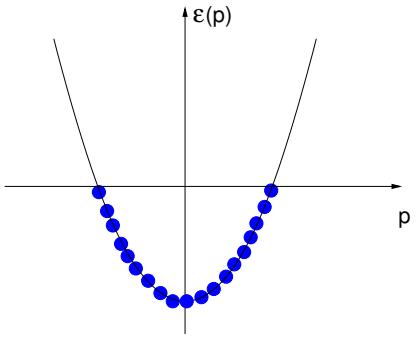


Figure 18: Ground state in the 1 dimensional case. Blue circles correspond to "filled" single-particle states.

7.1.2 Excitations

- Particle excitations

$$c_\sigma^\dagger(\mathbf{k})|\mathrm{GS}\rangle \qquad \text{with } |\mathbf{k}| > p_F. \tag{425}$$

  Their energies and momenta are

$$E = E_{\mathrm{GS}} + \epsilon(\mathbf{k}) > E_{\mathrm{GS}}\ , \quad \mathbf{P} = \mathbf{k}. \tag{426}$$

- Hole excitations

$$c_\sigma(\mathbf{k})|\mathrm{GS}\rangle \qquad \text{with } |\mathbf{k}| < p_F. \tag{427}$$

  Their energies and momenta are

$$E = E_{\mathrm{GS}} - \epsilon(\mathbf{k}) > E_{\mathrm{GS}}\ , \quad \mathbf{P} = -\mathbf{k}. \tag{428}$$

- Particle-hole excitations

$$c_\sigma^\dagger(\mathbf{k})c_\tau(\mathbf{p})|\mathrm{GS}\rangle \qquad \text{with } |\mathbf{k}| > p_F > |\mathbf{p}|. \tag{429}$$

  Their energies and momenta are

$$E = E_{\mathrm{GS}} + \epsilon(\mathbf{k}) - \epsilon(\mathbf{p}) > E_{\mathrm{GS}}\ , \quad \mathbf{P} = \mathbf{k} - \mathbf{p}. \tag{430}$$

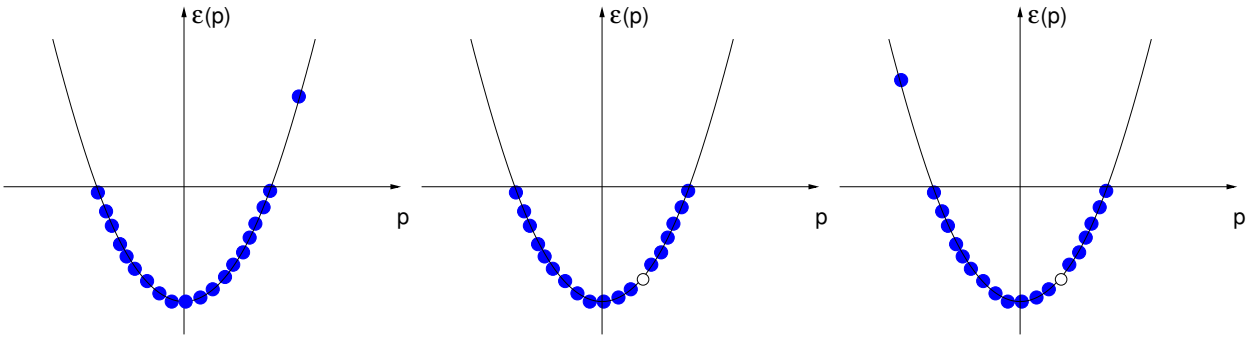


Figure 19: Some simple excited states: (a) particle (b) hole (c) particle-hole.

7.1.3 Density Correlations

Consider the single-particle operator

$$o = |\mathbf{r}\rangle\langle\mathbf{r}| \tag{431}$$

It represents the particle density at position $|\mathbf{r}\rangle$ as can be seen by acting on position eigenstates. In second quantization it is

$$\rho(\mathbf{r}) = \sum_\sigma \int d^3\mathbf{r}' d^3\mathbf{r}'' \ \langle \mathbf{r}'|o|\mathbf{r}''\rangle \ c_\sigma^\dagger(\mathbf{r}')c_\sigma(\mathbf{r}'') = \sum_\sigma c_\sigma^\dagger(\mathbf{r})c_\sigma(\mathbf{r}). \tag{432}$$

1. One-point function.

   We now want to determine the expectation value of this operator in the ground state

   $$\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle = \sum_\sigma \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{r})c_\sigma(\mathbf{r})|\mathrm{GS}\rangle. \tag{433}$$

   A crucial observation is that the ground state has a simple description in terms of the Fock space built from momentum eigenstates. Hence what we want to do is to work out the momentum representation of $\rho(\mathbf{r})$. We know from our general formula (355) that

   $$c_\sigma(\mathbf{r}) = \sum_{\mathbf{p}} \underbrace{\langle \mathbf{r}|\mathbf{p}\rangle}_{L^{-3/2}e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}} c_\sigma(\mathbf{p}). \tag{434}$$

   Substituting this as well as the analogous expression for the creation operator into (433), we obtain

   $$\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle = \sum_\sigma \frac{1}{L^3} \sum_{\mathbf{p},\mathbf{p}'} e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}} \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle. \tag{435}$$

   For the expectation value $\langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle$ to be non-zero, we must have that $c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle$ reproduces $|\mathrm{GS}\rangle$ itself. The only way this is possible is if $|\mathbf{p}| < p_F$ (so that the c pokes a hole in the Fermi sea) and $\mathbf{p}' = \mathbf{p}$ (so that the $c^\dagger$ precisely fills the hole made by the c). By this reasoning we have

   $$\langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\tau(\mathbf{p})|\mathrm{GS}\rangle = \delta_{\sigma,\tau}\delta_{\mathbf{p},\mathbf{p}'}\theta(p_F - |\mathbf{p}'|). \tag{436}$$

   Similarly we can show that

   $$\begin{aligned} \langle \mathrm{GS}|c_\sigma(\mathbf{p}')c_\tau^\dagger(\mathbf{p})|\mathrm{GS}\rangle &= \delta_{\sigma,\tau}\delta_{\mathbf{p},\mathbf{p}'}\theta(|\mathbf{p}| - p_F), \\ \langle \mathrm{GS}|c_\sigma(\mathbf{p}')c_\tau(\mathbf{p})|\mathrm{GS}\rangle &= 0 = \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\tau^\dagger(\mathbf{p})|\mathrm{GS}\rangle. \end{aligned} \tag{437}$$

   Substituting (436) back into (435) we find

   $$\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle = \sum_\sigma \frac{1}{L^3} \sum_{\mathbf{p},\mathbf{p}'} e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}} \delta_{\mathbf{p},\mathbf{p}'} \ \theta(p_F - |\mathbf{p}|) = \underbrace{2}_{\text{spin}} \frac{1}{L^3} \sum_{\mathbf{p}} \theta(p_F - |\mathbf{p}|) = \frac{N}{L^3}. \tag{438}$$

   So our expectation value gives precisely the particle density. This is expected because our system is translationally invariant and therefore $\langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle$ cannot depend on $\mathbf{r}$.

2. Two-point function.

   Next we want to determine the two-point function

   $$\langle \mathrm{GS}|\rho(\mathbf{r})\rho(\mathbf{r}')|\mathrm{GS}\rangle = \sum_{\sigma,\tau} \frac{1}{L^6} \sum_{\mathbf{p},\mathbf{p}'} \sum_{\mathbf{k},\mathbf{k}'} e^{\frac{i}{\hbar}(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}} e^{\frac{i}{\hbar}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}'} \langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})c_\tau^\dagger(\mathbf{k}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle. \tag{439}$$

   The expectation value $\langle \mathrm{GS}|c_\sigma^\dagger(\mathbf{p}')c_\sigma(\mathbf{p})c_\tau^\dagger(\mathbf{k}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle$ can be calculated by thinking about how the creation and annihilation operators act on the ground state, and then concentrating on the processes that reproduce the ground state itself in the end (see Fig. 20).

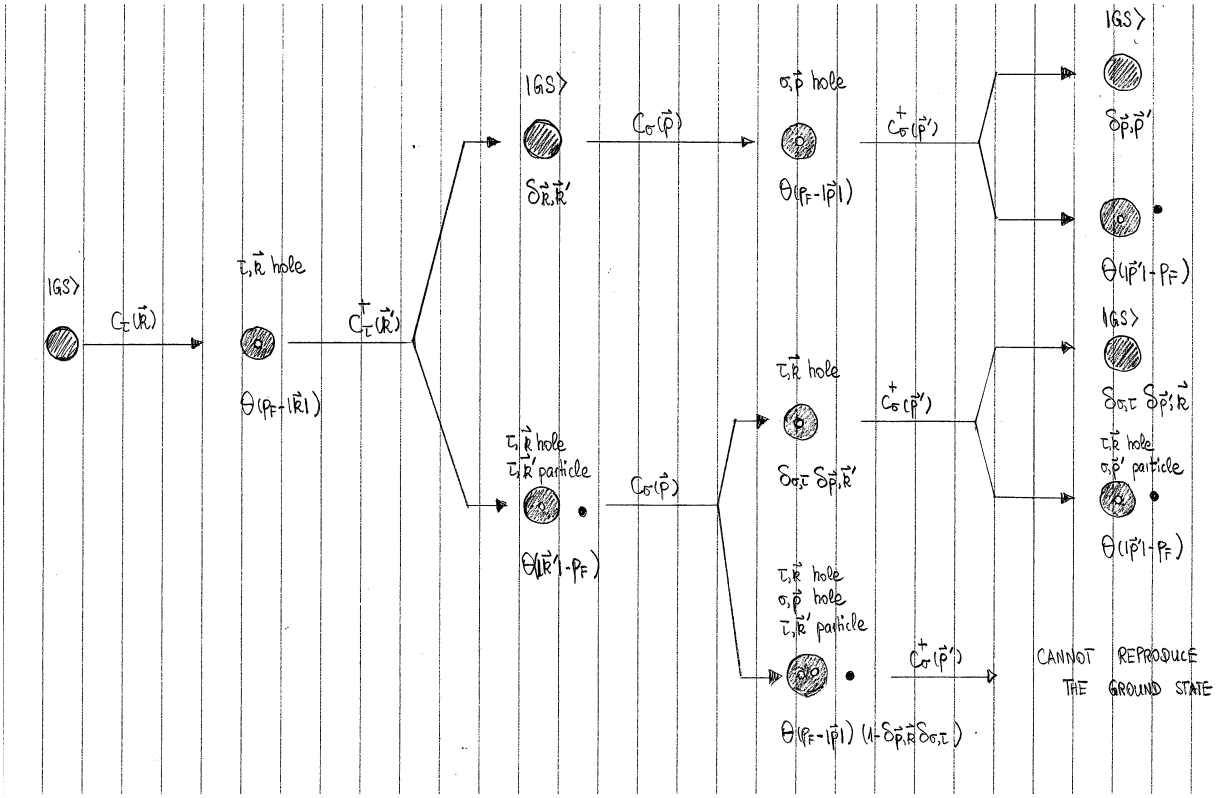


Figure 20:

The result is

$$\begin{aligned}\langle \mathrm{GS}|c^\dagger_\sigma(\mathbf{p}')c_\sigma(\mathbf{p})c^\dagger_\tau(\mathbf{k}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle &= \delta_{\mathbf{k},\mathbf{k}'}\delta_{\mathbf{p},\mathbf{p}'}\theta(p_F-|\mathbf{p}|)\theta(p_F-|\mathbf{k}|)\\ &\quad +\delta_{\sigma,\tau}\delta_{\mathbf{p},\mathbf{k}'}\delta_{\mathbf{k},\mathbf{p}'}\theta(|\mathbf{k}'|-p_F)\theta(p_F-|\mathbf{k}|). \end{aligned} \tag{440}$$

Observe that by virtue of (436) and (437) this can be rewritten in the form

$$\langle \mathrm{GS}|c^\dagger_\sigma(\mathbf{p}')c_\sigma(\mathbf{p})|\mathrm{GS}\rangle\langle \mathrm{GS}|c^\dagger_\tau(\mathbf{k}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle + \langle \mathrm{GS}|c^\dagger_\sigma(\mathbf{p}')c_\tau(\mathbf{k})|\mathrm{GS}\rangle\langle \mathrm{GS}|c_\sigma(\mathbf{p})c^\dagger_\tau(\mathbf{k}')|\mathrm{GS}\rangle\ . \tag{441}$$

The fact that the 4-point function (440) can be written as a sum over products of two-point functions is a reflection of *Wick's theorem* for noninteracting spin-1/2 fermions. This is not part of the syllabus and we won't dwell on it, but apart from extra minus signs, this says that $2n$-point functions are given by the sum over all possible "pairings", giving rise to a product of two-point functions. In our particular case this gives

$$\begin{aligned}\langle c^\dagger_\sigma(\mathbf{p}')c_\sigma(\mathbf{p})c^\dagger_\tau(\mathbf{k}')c_\tau(\mathbf{k})\rangle &= \langle c^\dagger_\sigma(\mathbf{p}')c_\sigma(\mathbf{p})\rangle\langle c^\dagger_\tau(\mathbf{k}')c_\tau(\mathbf{k})\rangle - \langle c^\dagger_\sigma(\mathbf{p}')c^\dagger_\tau(\mathbf{k}')\rangle\langle c_\sigma(\mathbf{p})c_\tau(\mathbf{k})\rangle\\ &+ \langle c^\dagger_\sigma(\mathbf{p}')c_\tau(\mathbf{k})\rangle\langle c_\sigma(\mathbf{p})c^\dagger_\tau(\mathbf{k}')\rangle, \end{aligned} \tag{442}$$

and using that the two point function of two creation or two annihilation operators is zero we obtain (441). Substituting (440) back in to (439) gives

$$\begin{aligned}\langle \mathrm{GS}|\rho(\mathbf{r})\rho(\mathbf{r}')|\mathrm{GS}\rangle &= \sum_{\sigma,\sigma'}\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{p}}\theta(p_F-|\mathbf{k}|)\theta(p_F-|\mathbf{p}|)\\ &+ \sum_{\sigma}\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}\theta(|\mathbf{k}|-p_F)\theta(p_F-|\mathbf{k}'|)e^{\frac{i}{\hbar}(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{r}-\mathbf{r}')}\\ &= \langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle\langle \mathrm{GS}|\rho(\mathbf{r}')|\mathrm{GS}\rangle + 2\frac{1}{L^3}\sum_{|\mathbf{k}|>p_F}e^{\frac{i}{\hbar}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\frac{1}{L^3}\sum_{|\mathbf{k}'|<p_F}e^{-\frac{i}{\hbar}\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}. \end{aligned} \tag{443}$$

10: Evaluating k-sums

Evaluting the $\mathbf{k}$ sums for large L: The idea is to turn sums into integrals

$$\frac{1}{L^3}\sum_{|\mathbf{k}|<p_F} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} \longrightarrow \int \frac{d^3\mathbf{k}}{(2\pi\hbar)^3}\theta(p_F-|\mathbf{k}|)e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} = \int_0^\infty dp p^2 \int_0^\pi d\vartheta \sin\vartheta \int_0^{2\pi} d\varphi \frac{\theta(p_F-\hbar p)}{(2\pi)^3} e^{ip|\mathbf{R}|\cos\vartheta}$$
$$= \int_0^{p_F/\hbar} \frac{dp}{(2\pi)^2}\frac{2p\sin(p|\mathbf{R}|)}{|\mathbf{R}|} = \frac{\sin(p_F|\mathbf{R}|)-p_F|\mathbf{R}|\cos(p_F|\mathbf{R}|)}{2\pi^2|\mathbf{R}|^3} \equiv h(|\mathbf{R}|). \tag{444}$$

Here we have introduced spherical polar coordinates such that the z-axis of our co-ordinate system is along the $\mathbf{R}$ direction, and

$$\begin{aligned} k_x &= \hbar p\sin\vartheta\cos\varphi\ , \\ k_y &= \hbar p\sin\vartheta\sin\varphi\ , \\ k_z &= \hbar p\cos\vartheta. \end{aligned} \tag{445}$$

The other sum works similarly

$$\frac{1}{L^3}\sum_{|\mathbf{k}|>p_F} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} = \frac{1}{L^3}\sum_{\mathbf{k}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} - \frac{1}{L^3}\sum_{|\mathbf{k}|<p_F} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}}. \tag{446}$$

The second part is evaluated above, while the first part is

$$\frac{1}{L^3}\sum_{\mathbf{k}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} = \delta^{(3)}(\mathbf{R}). \tag{447}$$

The equality can be proved by multiplying both sides by a test-function $f(\mathbf{R})$ and then integrating over $\mathbf{R}$:

$$\int d^3\mathbf{R}\frac{1}{L^3}\sum_{\mathbf{k}} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} f(\mathbf{R}) = \frac{1}{L^3}\sum_{\mathbf{k}}\int d^3\mathbf{R} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{R}} f(\mathbf{R}) = \frac{1}{L^3}\sum_{\mathbf{k}} \tilde{f}_{\mathbf{k}} = f(\mathbf{0}). \tag{448}$$

Here we have used standard definitions for Fourier series, cf Riley/Hobson/Bence 12.7.

Using these simplifications for large L we arrive at our final answer

$$\boxed{\langle \mathrm{GS}|\rho(\mathbf{r})\rho(\mathbf{r}')|\mathrm{GS}\rangle = \langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle^2 + \langle \mathrm{GS}|\rho(\mathbf{r})|\mathrm{GS}\rangle\delta^{(3)}(\mathbf{r}-\mathbf{r}') - 2\left[h(|\mathbf{r}-\mathbf{r}'|)\right]^2.} \tag{449}$$

The first two terms are the same as for a *classical* ideal gas, while the third contribution is due to the *fermionic statistics* (Pauli exclusion: "fermions don't like to be close to one another").

7.2 Homework Questions 15-16

Question 15. Consider a system of fermions moving freely on a one-dimensional ring of length L, i.e. periodic boundary conditions are applied between $x = 0$ and $x = L$. The fermions are all in the same spin state, so that spin quantum numbers may be omitted. Fermion creation and annihilation operators at the point x are denoted by $\psi^\dagger(x)$ and $\psi(x)$.
a) Write down the complete set of anticommutation relation satisfied by $\psi^\dagger(x_1)$ and $\psi(x_2)$.
b) Write down the wave-functions of single-particle momentum eigenstates (make sure to take the boundary conditions into account!). What are the allowed values of momentum? Using this result, derive an expression for the momentum space creation and annihilation operators $\Psi_p^\dagger$ and Ψ_p in terms of $\psi^\dagger(x)$ and $\psi(x)$ (hint: use the general result for basis transformation obtained in the lecture notes).

c) Starting with your expression for the anticommutator $\{\psi^\dagger(x_1), \psi(x_2)\}$, evaluate $\{\Psi_p^\dagger, \Psi_q\}$.
d) Derive an expression for $\psi(x)$ in terms of Ψ_k.
e) The density operator $\rho(x)$ is defined by $\rho(x) = \psi^\dagger(x)\,\psi(x)$. The number operator is

$$N = \int_0^L dx\, \rho(x)\,.$$

Express $\rho(x)$ in terms of $\Psi_p^\dagger$ and Ψ_q, and show from this that

$$N = \sum_k \Psi_k^\dagger \Psi_k\,.$$

Let $|0\rangle$ be the vacuum state (containing no particles) and define $|\phi\rangle$ by

$$|\phi\rangle = A \prod_k (u_k + v_k \Psi_k^\dagger)|0\rangle,$$

where u_k and v_k are complex numbers depending on the label k, and A is a normalisation constant.

Evaluate (i) $|A|^2$, (ii) $\langle\phi|N|\phi\rangle$, and (iii) $\langle\phi|N^2|\phi\rangle$. Under what conditions is $|\phi\rangle$ an eigenstate of particle number?

Question 16. Consider a system of fermions in which the functions $\varphi_\ell(x)$, $\ell = 1, 2 \ldots N$, form a complete orthonormal basis for single particle wavefunctions.
a) Explain how Slater determinants may be used to construct a complete orthonormal basis for n-particle states with $n = 2, 3 \ldots N$. Calculate the normalisation constant for such a Slater determinant at a general value of n. How many independent n-particle states are there for each n?
b) Let $C_\ell^\dagger$ and C_ℓ be fermion creation and destruction operators which satisfy the usual anticommutation relations. The quantities a_k are defined by

$$a_k = \sum_{\ell=1}^{N} U_{k\ell} C_\ell,$$

where $U_{k\ell}$ are elements of an $N \times N$ matrix, U. Write down an expression for $a_k^\dagger$. Find the condition which must be satisfied by the matrix U in order that the operators $a_k^\dagger$ and a_k also satisfy fermion anticommutation relations.
c) Non-interacting spinless fermions move in one dimension in an infinite square-well potential, with position coordinate $0 \le x \le L$. The normalised single particle energy eigenstates are

$$\varphi_\ell(x) = \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{\ell\pi x}{L}\right)\,,$$

and the corresponding fermion creation operator is $C_\ell^\dagger$.

Write down expressions for $C^\dagger(x)$, the fermion creation operator at the point x, and for $\rho(x)$, the particle density operator, in terms of $C_\ell^\dagger$, C_ℓ and $\varphi_\ell(x)$.
d) What is the ground state expectation value $\langle\rho(x)\rangle$ in a system of n fermions?

In the limit $n \to \infty$, $L \to \infty$, taken at fixed average density $\rho_0 = n/L$, show that

$$\langle\rho(x)\rangle = \rho_0 \left[1 - \frac{\sin 2\pi\rho_0 x}{2\pi\rho_0 x}\right]\,.$$

Sketch this function and comment briefly on its behaviour for $x \to 0$ and $x \to \infty$.

8 Application II: Weakly Interacting Bosons

As you know from Statistical Mechanics, the ideal Bose gas displays the very interesting phenomenon of *Bose condensation*. This has been observed in systems of trapped Rb atoms and led to the award of the Nobel prize in 2001 to Ketterle, Cornell and Wiemann. The atoms in these experiments are bosonic, but the atom-atom interactions are not zero. We now want to understand the effects of interactions in the framework of a microscopic theory. The kinetic energy operator is expressed in terms of creation/annihilation operators single-particle momentum eigenstates as

$$\hat{T} = \sum_{\mathbf{p}} \frac{\mathbf{p}^2}{2m} c^\dagger(\mathbf{p}) c(\mathbf{p}). \tag{450}$$

Here we have assumed that our system in enclosed in a large, periodic box of linear dimension L. The boson-boson interaction is most easily expressed in position space

$$\hat{V} = \frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \; c^\dagger(\mathbf{r}) c^\dagger(\mathbf{r}') V(\mathbf{r}, \mathbf{r}') c(\mathbf{r}') c(\mathbf{r}) \tag{451}$$

A good model for the potential $V(\mathbf{r}, \mathbf{r}')$ is to take it of the form

$$V(\mathbf{r}, \mathbf{r}') = U \delta^{(3)}(\mathbf{r} - \mathbf{r}'), \tag{452}$$

i.e. bosons interact only if they occupy the same point in space. Changing to the momentum space description

$$c(\mathbf{r}) = \frac{1}{L^{3/2}} \sum_{\mathbf{p}} e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} c(\mathbf{p}), \tag{453}$$

we have

$$\hat{V} = \frac{U}{2L^3} \sum_{\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}} c^\dagger(\mathbf{p_1}) c^\dagger(\mathbf{p_2}) c(\mathbf{p_3}) c(\mathbf{p_1} + \mathbf{p_2} - \mathbf{p_3}). \tag{454}$$

8.1 Ideal Bose Gas

For $U = 0$ we are dealing with an ideal Bose gas and we know that the ground state is a *condensate*: all particles occupy the lowest-energy single-particle state, i.e. the zero-momentum state

$$|\mathrm{GS}\rangle_0 = \frac{1}{\sqrt{N!}} \Big(c^\dagger(\mathbf{p} = \mathbf{0}) \Big)^N |0\rangle. \tag{455}$$

So $\mathbf{p} = 0$ is special, and in particular we have

$${}_0\langle \mathrm{GS} | c^\dagger(\mathbf{p} = \mathbf{0}) c(\mathbf{p} = \mathbf{0}) | \mathrm{GS} \rangle_0 = N. \tag{456}$$

8.2 Bogoliubov Approximation

For small $U > 0$ we expect the Bose-Einstein condensate to persist, i.e. we expect

$$\langle \mathrm{GS} | c^\dagger(\mathbf{p} = \mathbf{0}) c(\mathbf{p} = \mathbf{0}) | \mathrm{GS} \rangle = N_0 \gg 1. \tag{457}$$

However,

$$[c^\dagger(\mathbf{0}) c(\mathbf{0}), \hat{V}] \neq 0, \tag{458}$$

so that the number of $\mathbf{p} = \mathbf{0}$ bosons is not conserved, and the ground state $|\mathrm{GS}\rangle$ will be a superposition of states with different numbers of $\mathbf{p} = \mathbf{0}$ bosons. However, for the ground state and low-lying excited states we will have

$$\langle \Psi | c^\dagger(\mathbf{0}) c(\mathbf{0}) | \Psi \rangle \simeq N_0 \ , \tag{459}$$

where N_0, crucially, is a very large number. The *Bogoliubov approximation* states that, *when acting on the ground state or low-lying excited states*, we in fact have

$$\boxed{c^\dagger(\mathbf{0}) \simeq \sqrt{N_0}\ , \qquad c(\mathbf{0}) \simeq \sqrt{N_0}\ ,} \tag{460}$$

i.e. creation and annihilation operators are approximately diagonal. This is a much stronger statement than (459), and at first sight looks rather strange. It amounts to making an ansatz for low-energy states $|\psi\rangle$ that fulfils

$$\langle\psi'|c(\mathbf{0})|\psi\rangle = \sqrt{N_0}\langle\psi'|\psi\rangle + \dots \tag{461}$$

where the dots denote terms that are small compared to $\sqrt{N_0}$. We'll return to what this implies for the structure of $|\psi\rangle$ a little later. Using (460) we may expand H in inverse powers of N_0

$$\begin{aligned} H &= \sum_{\mathbf{p}} \frac{\mathbf{p}^2}{2m} c^\dagger(\mathbf{p})c(\mathbf{p}) \\ &\quad + \frac{U}{2L^3}N_0^2 + \frac{UN_0}{2L^3}\sum_{\mathbf{k}\neq\mathbf{0}} 2c^\dagger(\mathbf{k})c(\mathbf{k}) + 2c^\dagger(-\mathbf{k})c(-\mathbf{k}) + c^\dagger(\mathbf{k})c^\dagger(-\mathbf{k}) + c(-\mathbf{k})c(\mathbf{k}) \\ &\quad + \dots \end{aligned} \tag{462}$$

Now use that

$$N_0 = c^\dagger(\mathbf{0})c(\mathbf{0}) = N - \sum_{\mathbf{p}\neq\mathbf{0}} c^\dagger(\mathbf{p})c(\mathbf{p}), \tag{463}$$

where N is the (conserved) total number of bosons, and define

$$\rho = \frac{N}{L^3} = \text{density of particles.} \tag{464}$$

Then our Hamiltonian becomes

$$\boxed{H = \frac{U\rho}{2}N + \sum_{\mathbf{p}\neq\mathbf{0}} \underbrace{\left[\frac{\mathbf{p}^2}{2m} + U\rho\right]}_{\epsilon(\mathbf{p})} c^\dagger(\mathbf{p})c(\mathbf{p}) + \frac{U\rho}{2}\left[c^\dagger(\mathbf{p})c^\dagger(-\mathbf{p}) + c(-\mathbf{p})c(\mathbf{p})\right] + \dots} \tag{465}$$

The Bogoliubov approximation has reduced the complicated four-boson interaction to two-boson terms. The price we pay is that we have to deal with the "pairing"-terms quadratic in creation/annihilation operators.

8.3 Bogoliubov Transformation

Consider the creation/annihilation operators defined by

$$\begin{pmatrix} b(\mathbf{p}) \\ b^\dagger(-\mathbf{p}) \end{pmatrix} = \begin{pmatrix} \cosh(\theta_\mathbf{p}) & \sinh(\theta_\mathbf{p}) \\ \sinh(\theta_\mathbf{p}) & \cosh(\theta_\mathbf{p}) \end{pmatrix} \begin{pmatrix} c(\mathbf{p}) \\ c^\dagger(-\mathbf{p}). \end{pmatrix} \tag{466}$$

It is easily checked that for any choice of Bogoliubov angle $\theta_\mathbf{p}$

$$[b(\mathbf{p}), b(\mathbf{q})] = 0 = [b^\dagger(\mathbf{p}), b^\dagger(\mathbf{q})]\ , \qquad [b(\mathbf{p}), b^\dagger(\mathbf{q})] = \delta_{\mathbf{p},\mathbf{q}}. \tag{467}$$

Exercise 5

Let $f(\boldsymbol{p})$, $f^\dagger(\boldsymbol{p})$ be *fermionic* annihilation and creation operators

$$\{f(\boldsymbol{p}), f^\dagger(\boldsymbol{q})\} = \delta_{\boldsymbol{p},\boldsymbol{q}}\ , \qquad \{f(\boldsymbol{p}), f(\boldsymbol{q})\} = 0. \tag{468}$$

Show that the operators defined by

$$\begin{pmatrix} g(\mathbf{p}) \\ g^\dagger(-\mathbf{p}) \end{pmatrix} = \begin{pmatrix} \cos(\theta_\mathbf{p}) & i\sin(\theta_\mathbf{p}) \\ i\sin(\theta_\mathbf{p}) & \cos(\theta_\mathbf{p}) \end{pmatrix} \begin{pmatrix} f(\mathbf{p}) \\ f^\dagger(-\mathbf{p}) \end{pmatrix} , \quad \theta_{-\boldsymbol{p}} = -\theta_{\boldsymbol{p}} , \tag{469}$$

fulfil canonical anticommutation relations for any value of the angle $\theta_\mathbf{p}$. Eqn (469) is a Bogoliubov transformation for fermions.

In terms of the Bogoliubov bosons the Hamiltonian becomes

$$\begin{aligned} H &= \text{const} + \frac{1}{2}\sum_{\mathbf{p}\neq\mathbf{0}} \left[\left(\frac{\mathbf{p}^2}{2m} + U\rho\right)\cosh(2\theta_\mathbf{p}) - U\rho\sinh(2\theta_\mathbf{p})\right]\left[b^\dagger(\mathbf{p})b(\mathbf{p}) + b^\dagger(-\mathbf{p})b(-\mathbf{p})\right] \\ &\quad - \left[\left(\frac{\mathbf{p}^2}{2m} + U\rho\right)\sinh(2\theta_\mathbf{p}) - U\rho\cosh(2\theta_\mathbf{p})\right]\left[b^\dagger(\mathbf{p})b^\dagger(-\mathbf{p}) + b(-\mathbf{p})b(\mathbf{p})\right] + \ldots \end{aligned} \tag{470}$$

Now we choose

$$\tanh(2\theta_\mathbf{p}) = \frac{U\rho}{\frac{\mathbf{p}^2}{2m} + U\rho}, \tag{471}$$

as this removes the $b^\dagger b^\dagger + bb$ terms, and leaves us with a *diagonal* Hamiltonian

$$\boxed{H = \text{const} + \sum_{\mathbf{p}\neq\mathbf{0}} E(\mathbf{p})b^\dagger(\mathbf{p})b(\mathbf{p}) + \ldots} \tag{472}$$

where

$$E(\mathbf{p}) = \sqrt{\left(\frac{\mathbf{p}^2}{2m} + U\rho\right)^2 - (U\rho)^2} \ . \tag{473}$$

We note that

$$E(\mathbf{p}) \longrightarrow \frac{\mathbf{p}^2}{2m} \qquad \text{for } |\mathbf{p}| \to \infty, \tag{474}$$

which tells us that at high momenta (and hence high energies) we recover the quadratic dispersion. In this limit $\theta_\mathbf{p} \to 0$, so that the Bogoliubov bosons reduce to the "physical" bosons we started with. On the other hand

$$E(\mathbf{p}) \longrightarrow \sqrt{\frac{U\rho}{m}}|\mathbf{p}| \qquad \text{for } |\mathbf{p}| \to 0. \tag{475}$$

So here we have a *linear* dispersion.

8.4 Ground State and Low-lying Excitations

We note that the Hamiltonian (472) involves only creation/annihilation operators with $\mathbf{p} \neq 0$. Formally, we can define zero-momentum Bogoliubov bosons as simply being equal to the original ones

$$b(\mathbf{0}) = c(\mathbf{0}) \ . \tag{476}$$

Let us now define the Bogoliubov vacuum state $|\tilde{0}\rangle$ by

$$b(\mathbf{p})|\tilde{0}\rangle = 0 \ . \tag{477}$$

Clearly, for $\mathbf{p} \neq 0$ we have $E(\mathbf{p}) > 0$, and hence no Bogoliubov quasiparticles will be present in the ground state. On the other hand, a basic assumption we made was that

$$\langle \text{GS}|b(\mathbf{0})|\text{GS}\rangle \simeq \sqrt{N_0}. \tag{478}$$

In order to get an idea what this implies for the structure of the ground state, let us express it in the general form

$$|\mathrm{GS}\rangle = \sum_{n=0}^{\infty} \alpha_n \big(b^\dagger(\mathbf{0})\big)^n |\tilde{0}\rangle \ . \tag{479}$$

Eqn (478) then implies that

$$\alpha_{n+1} \simeq \frac{\sqrt{N_0}}{n+1}\alpha_n \ . \tag{480}$$

Replacing this approximate relation by an equality leads to a *coherent state*

$$|\mathrm{GS}\rangle = e^{-N_0/2} e^{\sqrt{N_0} b^\dagger(\mathbf{0})}|\tilde{0}\rangle. \tag{481}$$

Low-lying excited states can now be obtained by creating Bogoliubov quasipartices, e.g.

$$b^\dagger(\mathbf{q})|\mathrm{GS}\rangle, \tag{482}$$

is a particle-excitation with energy $E(\mathbf{q}) > 0$.

8.5 Ground state correlation functions

We are now in a position to work out correlation functions in the ground state such as

$$\langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle \ , \quad \mathbf{p},\mathbf{q} \neq \mathbf{0}. \tag{483}$$

Inverting the Bogoliubov transformation (466) we have

$$\begin{aligned} c^\dagger(\mathbf{p}) &= \cosh(\theta_{\mathbf{p}}) b^\dagger(\mathbf{p}) - \sinh(\theta_{\mathbf{p}}) b(-\mathbf{p}) \ , \\ c(\mathbf{q}) &= \cosh(\theta_{\mathbf{q}}) b(\mathbf{q}) - \sinh(\theta_{\mathbf{q}}) b^\dagger(-\mathbf{q}) \ . \end{aligned} \tag{484}$$

Using that

$$\langle \mathrm{GS}|b^\dagger(\mathbf{p}) = 0 = b(\mathbf{q})|\mathrm{GS}\rangle, \tag{485}$$

we find that

$$\begin{aligned} \langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle &= \sinh(\theta_{\mathbf{q}})\sinh(\theta_{\mathbf{q}})\langle \mathrm{GS}|b(-\mathbf{p})b^\dagger(-\mathbf{q})|\mathrm{GS}\rangle \\ &= \sinh^2(\theta_{\mathbf{p}})\delta_{\mathbf{q},\mathbf{p}} \quad (\mathbf{p},\mathbf{q} \neq \mathbf{0}). \end{aligned} \tag{486}$$

This tells us that, in contrast to the ideal Bose gas, in the ground state of the interacting Bose gas we have a finite density of bosons with non-zero momentum

$$\langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{p})|\mathrm{GS}\rangle = \sinh^2(\theta_{\mathbf{p}}) \ . \tag{487}$$

Another feature of the ground state is that the two-point function of two annihilation/creation operators is non-zero

$$\langle \mathrm{GS}|c(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle = \langle \mathrm{GS}|c^\dagger(\mathbf{q})c^\dagger(\mathbf{p})|\mathrm{GS}\rangle = -\cosh(\theta_{\mathbf{p}})\sinh(\theta_{\mathbf{q}})\delta_{\mathbf{p},-\mathbf{q}}. \tag{488}$$

These imply that boson number is not a good quantum number in the ground state. More formally, we say that *the ground state spontaneously breaks the U(1) symmetry of the Hamiltonian* $H = \hat{T} + \hat{V}$. Let us explain that statement. The Hamiltonian is invariant under the symmetry operation

$$\begin{aligned} \hat{U}c(\mathbf{p})\hat{U}^\dagger &= e^{i\phi}c(\mathbf{p}) \ , \quad \phi \in \mathbb{R} \ , \\ \hat{U}c^\dagger(\mathbf{p})\hat{U}^\dagger &= e^{-i\phi}c^\dagger(\mathbf{p}) \ , \end{aligned} \tag{489}$$

i.e.

$$\hat{U}H\hat{U}^\dagger = H. \tag{490}$$

The reason for this is that all terms in H involve the same number of creation as annihilation operators, and the total particle number is therefore conserved. This is referred to as a global U(1) symmetry (as the transformations (489) form a group called U(1)). Let us now investigate how ground state expectation values transform. We have

$$\langle \mathrm{GS}|c(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle = \langle \mathrm{GS}|\hat{U}^\dagger \hat{U} c(\mathbf{p}) \hat{U}^\dagger \hat{U} c(\mathbf{q}) \hat{U}^\dagger \hat{U}|\mathrm{GS}\rangle = e^{2i\phi}\langle \mathrm{GS}|\hat{U}^\dagger c(\mathbf{p})c(\mathbf{q})\hat{U}|\mathrm{GS}\rangle \ . \tag{491}$$

If the ground state were invariant under the symmetry, we would have $\hat{U}|\mathrm{GS}\rangle = |\mathrm{GS}\rangle$. Eqn (491) would then imply that $\langle \mathrm{GS}|c(\mathbf{p})c(\mathbf{q})|\mathrm{GS}\rangle = 0$. Reversing the argument, we see that a non-zero value of the expectation value (488) implies that the ground state *cannot be* invariant under the U(1) symmetry, and in fact *"breaks it spontaneously"*.

8.6 Depletion of the Condensate

We started out by asserting that for small interactions $U > 0$ we retain a Bose-Einstein condensate, i.e. the consensate fraction N_0/N remains large. We can now check that this assumption is *self-consistent*. We have

$$N_0 = N - \sum_{\mathbf{p}\neq\mathbf{0}} c^\dagger(\mathbf{p})c(\mathbf{p}). \tag{492}$$

Thus in the ground state

$$\frac{N_0}{N} = 1 - \frac{1}{N}\sum_{\mathbf{p}\neq\mathbf{0}} \langle \mathrm{GS}|c^\dagger(\mathbf{p})c(\mathbf{p})|\mathrm{GS}\rangle = 1 - \frac{1}{N}\sum_{\mathbf{p}\neq\mathbf{0}} \sinh^2(\theta_{\mathbf{p}}), \tag{493}$$

where we have used (486). This equals

$$\frac{N_0}{N} \ = \ 1 - \frac{1}{2N}\sum_{\mathbf{p}\neq\mathbf{0}} \left[\frac{1}{\sqrt{1-\tanh^2(2\theta_{\mathbf{p}})}} - 1 \right] = 1 - \frac{1}{2N}\sum_{\mathbf{p}\neq\mathbf{0}} \left[\frac{1}{\sqrt{1-\left[\frac{U\rho}{\epsilon(\mathbf{p})}\right]^2}} - 1 \right]. \tag{494}$$

We again turn this into an integral and evaluate it in spherical polar coordinates, which gives

$$\frac{N_0}{N} \approx 1 - \frac{2\pi}{\rho}\int_0^\infty \frac{dp}{(2\pi\hbar)^3} p^2 \left[\frac{1}{\sqrt{1-\left[\frac{U\rho}{\epsilon(\mathbf{p})}\right]^2}} - 1 \right]. \tag{495}$$

By means of the substitution $p = \sqrt{2mU\rho}z$ we can see that the integral is proportional to $U^{3/2}$ and thus indeed small for small U.

9 Application III: Spinwaves in a Ferromagnet

Consider the following model of a magnetic insulator: at each site $\mathbf{r}$ of a D-dimensional with N sites lattice we have a magnetic moment. In QM such magnetic moments are described by three spin-operators

$$S^\alpha_{\mathbf{r}} \ , \quad \alpha = x, y, z \ , \tag{496}$$

which fulfil the angular momentum commutation relations

$$[S^\alpha_{\mathbf{r}}, S^\beta_{\mathbf{r}'}] = \delta_{\mathbf{r},\mathbf{r}'} i\epsilon_{\alpha\beta\gamma} S^\gamma_{\mathbf{r}}. \tag{497}$$

We will assume that the spin are large in the sense that

$$\mathbf{S}^2_{\mathbf{r}} = \sum_\alpha \left(S^\alpha_{\mathbf{r}}\right)^2 = s(s+1) \gg 1. \tag{498}$$

Let us begin by constructing a basis of quantum mechanical states. At each site we have $2s+1$ eigenstates of $S^z_{\mathbf{r}}$

$$S^z_{\mathbf{r}}|m\rangle_{\mathbf{r}} = m|m\rangle_{\mathbf{r}}, \quad m = s, s-1, \ldots, -s. \tag{499}$$

They can be constructed from $|s\rangle_{\mathbf{r}}$ using spin lowering operators $S^-_{\mathbf{r}} = S^x_{\mathbf{r}} - iS^y_{\mathbf{r}}$

$$|s-n\rangle_{\mathbf{r}} = \frac{1}{\mathcal{N}_n}\big(S^-_{\mathbf{r}}\big)^n|s\rangle_{\mathbf{r}}\ , \quad n = 0, 1, \ldots, 2s, \tag{500}$$

where $\mathcal{N}_n$ are normalization constants. A basis of states is then given by

$$\prod_{\mathbf{r}}|s_{\mathbf{r}}\rangle_{\mathbf{r}}\ , \quad -s \le s_{\mathbf{r}} \le s \text{ spin on site } \mathbf{r}. \tag{501}$$

9.1 Heisenberg model and spin-rotational SU(2) symmetry

An appropriate Hamiltonian for a ferromagnetic insulator was derived by Heisenberg

$$\boxed{H = -J\sum_{\langle \mathbf{r},\mathbf{r}'\rangle}\mathbf{S}_{\mathbf{r}}\cdot\mathbf{S}_{\mathbf{r}'}.} \tag{502}$$

Here $\langle \mathbf{r}, \mathbf{r}'\rangle$ denote nearest-neighbour pairs of spins and we will assume that $J > 0$. The model (502) is known as the ferromagnetic *Heisenberg model*. You can check that the Hamiltonian (502) commutes with the three total spin operators

$$[H, S^\alpha] = 0\ , \quad S^\alpha = \sum_{\mathbf{r}} S^\alpha_{\mathbf{r}}. \tag{503}$$

These imply that the Hamiltonian is invariant under general rotations (in spin space)

$$e^{i\boldsymbol{\alpha}\cdot\mathbf{S}}He^{-i\boldsymbol{\alpha}\cdot\mathbf{S}} = H. \tag{504}$$

The transformations (504) form a group known as SU(2), and the Heisenberg Hamiltonian (502) is invariant under them.

9.2 Exact ground states

One ground state of H is given by

$$\boxed{|\mathrm{GS}\rangle = \prod_{\mathbf{r}}|s\rangle_{\mathbf{r}}.} \tag{505}$$

Its energy is

$$H|\mathrm{GS}\rangle = -J\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} s^2|\mathrm{GS}\rangle = -Js^2N_B|\mathrm{GS}\rangle, \tag{506}$$

where N_B is the total number of bonds in our lattice. The total spin lowering operator $S^- = \sum_{\mathbf{r}} S^-_{\mathbf{r}}$ commutes with H by virtue of (503) and hence

$$|\mathrm{GS}, n\rangle = \frac{1}{N_n}\left(S^-\right)^n|\mathrm{GS}\rangle\ , \quad 0 \le n \le 2sN \tag{507}$$

are ground states as well (as they have the same energy). Here N_n is a normalization.

11: Proof that $|\mathrm{GS}\rangle$ is a ground state

We note that the spin-spin interaction can be written in the form

$$2\mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'} = (\mathbf{S}_{\mathbf{r}} + \mathbf{S}_{\mathbf{r}'})^2 - \mathbf{S}_{\mathbf{r}}^2 - \mathbf{S}_{\mathbf{r}'}^2 = \mathbf{J}^2 - 2s(s+1). \tag{508}$$

Here $\mathbf{J}^2$ is the total angular momentum squared. Its eigenvalues follow from the theory of adding angular momenta to be

$$\mathbf{J}^2|j,m\rangle = j(j+1)|j,m\rangle \ , \quad j = 2s, 2s-1, \ldots, 1, 0. \tag{509}$$

This tells us that the maximal eigenvalue of $\mathbf{J}^2$ is $2s(2s+1)$, and by expanding $|\psi\rangle$ in a basis of eigenstates of $\mathbf{J}^2$ we can easily show that

$$\begin{aligned} \langle\psi|\mathbf{J}^2|\psi\rangle &= \sum_{j,m,j',m'} \langle\psi|j,m\rangle\langle j,m|\mathbf{J}^2|j',m'\rangle\langle j',m'|\psi\rangle \\ &= \sum_{j,m} |\langle\psi|j,m\rangle|^2 j(j+1) \leq 2s(2s+1)\sum_{j,m} |\langle\psi|j,m\rangle|^2 = 2s(2s+1). \end{aligned} \tag{510}$$

This tells us that

$$\langle\psi|\mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'}|\psi\rangle \leq s^2. \tag{511}$$

This provides us with a bound on the eigenvalues of the Hamiltonian, as

$$\langle\psi|H|\psi|\rangle \geq -J\sum_{\langle\mathbf{r},\mathbf{r}'\rangle} s^2 = -Js^2Nz. \tag{512}$$

The state we have constructed saturates this bound, so must be a ground state.

Let us now see how the SU(2) symmetry is reflected in expectation values of operators $\mathcal{O}$. At finite temperature we have

$$\langle\mathcal{O}\rangle_\beta = \frac{1}{Z(\beta)}\mathrm{Tr}\left[e^{-\beta H}\mathcal{O}\right], \tag{513}$$

where $Z(\beta) = \mathrm{Tr}[e^{-\beta H}]$ is the partition function and $\beta = 1/k_BT$. In the $T \to 0$ limit we have

$$\langle\mathcal{O}\rangle_\infty = \frac{1}{2sN+1}\sum_{n=0}^{2sN}\langle\mathrm{GS},\mathrm{n}|\mathcal{O}|\mathrm{GS},\mathrm{n}\rangle, \tag{514}$$

i.e. we average over all ground states. The thermal average, as well as its $T = 0$ limit, are invariant under rotations in spin space. Indeed, under a rotation in spin space we have

$$\langle e^{i\boldsymbol{\alpha}\cdot\mathbf{S}}\mathcal{O}e^{-i\boldsymbol{\alpha}\cdot\mathbf{S}}\rangle_\beta = \frac{1}{Z(\beta)}\mathrm{Tr}\left[e^{-\beta H}e^{i\boldsymbol{\alpha}\cdot\mathbf{S}}\mathcal{O}e^{-i\boldsymbol{\alpha}\cdot\mathbf{S}}\right] \tag{515}$$

where $\mathbf{S} = \sum_{\mathbf{r}}\mathbf{S}_{\mathbf{r}}$ are the global spin operators. Using the cyclicity of the trace and the fact that H commutes with the global spin operators, we see that this equals $\langle\mathcal{O}\rangle_\beta$. If we choose as our operator $\mathcal{O}$ any of the global spin operators, and consider a rotation by π around one of the orthogonal axes, we see that the magnetization always vanishes

$$\langle S^\alpha\rangle_\beta = 0 \ , \quad \alpha = x, y, z. \tag{516}$$

Physically this is what one would expect for a system that is spin rotationally invariant, i.e. looks the same in any direction in spin space.

9.3 Spontaneous Symmetry Breaking

In a real system, the $2sN+1$-fold ground state degeneracy is usually broken through imperfections. In practice the details of these imperfections are not important, the only thing that matters is that the symmetry gets broken. To keep things simple, one retains the spin-rotationally symmetric Hamiltonian, and says that the ground state breaks the symmetry "spontaneously".

A convenient mathematical description of this effect is as follows. Imagine adding an infinitesimal magnetic field $-\epsilon \sum_{\mathbf{r}} S^z_{\mathbf{r}}$ to the Hamiltonian. This will break the symmetry and hence the degeneracy of the ground states, which now will have energies

$$E_{\mathrm{GS,n}} = -Js^2 N_B - \epsilon(sN - n). \tag{517}$$

Now consider the sequence of limits

$$\lim_{\epsilon\to 0}\lim_{N\to\infty}\frac{1}{N}\left[E_{\mathrm{GS,n}} - E_{\mathrm{GS},0}\right] = \begin{cases} 0 & \text{if } \lim_{N\to\infty}\frac{n}{N} = 0\ , \\ \infty & \text{else.} \end{cases} \tag{518}$$

This means that if we *define* the thermodynamic limit in the above way, then the only surviving ground states (defined as states whose energy *density* is minimum in the thermodynamic limit) will have magnetization per site s, i.e. contain only a non-extensive number of spin flips. In all of these remaining ground states the spin rotational symmetry has been broken. As we have taken $\epsilon \to 0$ our Hamiltonian is again SU(2) symmetric, but the remaining ground states "spontaneously" break this symmetry.

9.4 Holstein-Primakoff Transformation

We succeeded in finding the ground states of H because of their simple structure. For more general spin Hamiltonians, or even the Hamiltonian (502) with negative value of J, this will no longer work and we need a more general, but approximate way of dealing with such problems. This is provided by *(linear) spinwave theory.*

As shown by Holstein and Primakoff, spin operators can be represented in terms of bosonic creation and annihilation operators as follows:

$$S^z_{\mathbf{r}} = s - a^\dagger_{\mathbf{r}} a_{\mathbf{r}}\ , \quad S^+_{\mathbf{r}} = S^x_{\mathbf{r}} + iS^y_{\mathbf{r}} = \sqrt{2s}\sqrt{1 - \frac{a^\dagger_{\mathbf{r}} a_{\mathbf{r}}}{2s}}\, a_{\mathbf{r}}\ . \tag{519}$$

You can check that the bosonic commutation relations

$$[a_{\mathbf{r}}, a^\dagger_{\mathbf{r}'}] = \delta_{\mathbf{r},\mathbf{r}'} \tag{520}$$

imply that

$$[S^\alpha_{\mathbf{r}}, S^\beta_{\mathbf{r}'}] = \delta_{\mathbf{r},\mathbf{r}'} i\epsilon_{\alpha\beta\gamma} S^\gamma_{\mathbf{r}}. \tag{521}$$

However, there is a caveat: the spaces of QM states are different! At site $\mathbf{r}$ we have

$$\left(S_{\mathbf{r}}\right)^n |s\rangle_{\mathbf{r}}\ , \quad n = 0, \ldots, 2s \tag{522}$$

for spins, but for bosons there are infinitely many states

$$\left(a^\dagger_{\mathbf{r}}\right)^n |0\rangle_{\mathbf{r}}\ , \quad n = 0, \ldots, \infty. \tag{523}$$

To make things match, we must impose a *constraint*, that there are at most 2s bosons per site. Now we take advantage of the fact that we have assumed s to be large: in the ground state there are no bosons present, because

$$\langle \mathrm{GS}|s - a^\dagger_{\mathbf{r}} a_{\mathbf{r}}|\mathrm{GS}\rangle = \langle \mathrm{GS}|S^z_{\mathbf{r}}|\mathrm{GS}\rangle = s = \tag{524}$$

Low-lying excited states will only have a few bosons, so for large enough s we don't have to worry about the constraint. Using the Holstein-Primakoff transformation, we can rewrite H in a $1/s$ *expansion*

$$\boxed{H = -J \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} s^2 - s \left[a^\dagger_{\mathbf{r}} a_{\mathbf{r}} + a^\dagger_{\mathbf{r}'} a_{\mathbf{r}'} - a^\dagger_{\mathbf{r}} a_{\mathbf{r}'} - a^\dagger_{\mathbf{r}'} a_{\mathbf{r}} \right] + \dots} \tag{525}$$

Here the dots indicate terms proportional to s^0, s^{-1}, etc. Once again using that s is large, we drop these terms (for the time being). We then can diagonalize H by going to momentum space

$$a_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} a(\mathbf{k}) \ , \qquad [a(\mathbf{k}), a^\dagger(\mathbf{p})] = \delta_{\mathbf{k},\mathbf{p}} \ , \tag{526}$$

which gives

$$\boxed{H = -Js^2 N z + \sum_{\mathbf{q}} \epsilon(\mathbf{q}) a^\dagger(\mathbf{q}) a(\mathbf{q}) + \dots} \tag{527}$$

For a simple cubic lattice the energy is

$$\epsilon(\mathbf{q}) = 2Js \left[3 - \cos q_x - \cos q_x - \cos q_z \right]. \tag{528}$$

For small wave numbers this is quadratic

$$\epsilon(\mathbf{q}) \approx Js|\mathbf{q}|^2 \qquad \text{for } |\mathbf{q}| \to 0. \tag{529}$$

In the context of spontaneous symmetry breaking these gapless excitations are known as *Goldstone modes*.

Let us now revisit the logic underlying our $1/s$ expansion. For things to be consistent, we require that the terms of order s in (527) provide only a small correction to the leading contribution proportional to s^2. This will be the case as long as we are interested only is states $|\Psi\rangle$ such that

$$\langle \Psi | a^\dagger(\mathbf{q}) a(\mathbf{q}) | \Psi \rangle \ll s. \tag{530}$$

This condition is certainly fulfilled for the ground state and low-lying excited states.

9.4.1 Heisenberg Antiferromagnet

Another example to which spinwave theory can be fruitfully applied is the model

$$H = J \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} \mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'}, \tag{531}$$

where $\langle \mathbf{r}, \mathbf{r}' \rangle$ denote nearest-neighbour pairs of spins on a simple cubic lattice and $J > 0$. Compared to (502) all we have done is to switch the overall sign of H, but this has important consequences. In particular, it is no longer possible to obtain an exact ground state for the model. Instead, we start by considering our spins to be *classical*. This is justified if we are interested only in states with large spin quantum numbers. We will assume this to be the case and check the self-consistency of our assumption later. In the classical limit we can think of the spins as three-dimensional vectors. The lowest energy configuration is then one, where all neighbouring spins point in opposite directions, i.e. along the three cystal axes the spin configuration looks like $\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow$ .... This is known as a *Néel state*. It is convenient to subdivide our lattice into two sublattices: on sublattice A all spins point in the same direction, while on sublattice B all spins point in the opposite direction. Like the ferromagnet, the model (531) has a global spin-rotational symmetry, that will be spontaneously broken in the ground state. By choosing our spin quantization axis appropriately, the classical ground state can then be written in the form

$$\prod_{\mathbf{r} \in A} |s\rangle_{\mathbf{r}} \prod_{\mathbf{r}' \in B} |-s\rangle_{\mathbf{r}'} \tag{532}$$

The idea is now to map this state to a ferromagnetic one, by inverting the spin quantization axis in the B sublattice. After that we can employ the Holstein-Primakoff transformation to carry out a $1/s$ expansion. As a result of the rotation of spin quatization axis on the B sublattice, the part of the Hamiltonian of order s now contains terms involving two annihilation or two creation operators. Diagonalizing the Hamiltonian then requires a Bogoliubov transformation.

9.5 Homework Questions 17-18

Question 17. (optional) Consider the ideal Fermi gas at finite density N/V in a periodic 3-dimensional box of length L.
(a) Give an expression of the ground state in terms of creation operators for momentum eigenstates.
(b) Calculate the **single-particle Green's function**

$$\begin{aligned} G_{\sigma\tau}(\omega,\mathbf{q}) &= \int dt\; e^{i\omega(t-t')} \int d^3\mathbf{r}\; e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} G_{\sigma\tau}(t,\mathbf{r};t',\mathbf{r}')\;, \\ G_{\sigma\tau}(t,\mathbf{r};t',\mathbf{r}') &= -i\langle GS|T c_\sigma(\mathbf{r},t)\; c^\dagger_\tau(\mathbf{r}',t')|GS\rangle, \end{aligned} \tag{533}$$

where T denotes time-ordering (i.e. $T\mathcal{O}(t_1)\mathcal{O}(t_2) = \theta(t_1-t_2)\mathcal{O}(t_1)\mathcal{O}(t_2) - \theta(t_2-t_1)\mathcal{O}(t_2)\mathcal{O}(t_1)$ for fermionic operators), and

$$c_\sigma(\mathbf{r},t) \equiv e^{\frac{i}{\hbar}Ht} c_\sigma(\mathbf{r}) e^{-\frac{i}{\hbar}Ht}. \tag{534}$$

First express the creation/annihilation operators $c^\dagger_\sigma(\mathbf{r},t)$, $c_\sigma(\mathbf{r},t)$ in terms of creation/annihilation operators in momentum space $c^\dagger_\sigma(\mathbf{p},t)$, $c_\sigma(\mathbf{p},t)$. Then show that for annihilation operators in momentum space we have

$$c_\sigma(\mathbf{p},t) \equiv e^{\frac{i}{\hbar}Ht} c_\sigma(\mathbf{p}) e^{-\frac{i}{\hbar}Ht} = c_\sigma(\mathbf{p}) e^{-\frac{i}{\hbar}t\epsilon(\mathbf{p})}\;, \tag{535}$$

where $\epsilon(\mathbf{p}) = \mathbf{p}^2/2m - \mu$. Use this to show that

$$c_\sigma(\mathbf{r},t) = \frac{1}{L^{3/2}} \sum_{\mathbf{p}} e^{-\frac{i}{\hbar}t\epsilon(\mathbf{p}) + i\mathbf{p}\cdot\mathbf{r}}\; c_\sigma(\mathbf{p}). \tag{536}$$

Now insert (536) into (533) and evaluate the ground state expectation value to obtain an integral representation for $G_{\sigma\tau}(t,\mathbf{r};t',\mathbf{r}')$. Why does the Green's function only depend on $t-t'$ and $\mathbf{r}-\mathbf{r}'$? Finally, calculate $G_{\sigma\tau}(\omega,\mathbf{q})$.

Note: the imaginary part of the single-particle Green's function is (approximately) measured by angle resolved photoemission (ARPES) experiments.

Question 18. A magnetic system consists of two types of Heisenberg spin $\mathbf{S}^A$ and $\mathbf{S}^B$ located respectively on the two inter-penetrating sublattices of an NaCl crystal structure (i.e. a simple cubic structure with alternate A and B in any Cartesian direction). Its Hamiltonian is

$$H = J \sum_{i,j} \mathbf{S}^A_i \cdot \mathbf{S}^B_j$$

where the i,j are nearest neighbours, respectively on the A and B sublattices. J is positive. Show that the classical ground state has all the A spins ferromagnetically aligned in one direction and all the B spins ferromagnetically aligned in the opposite direction. Assume the quantum mechanical ground state is well approximated by the classical one. To a first approximation the spin operators are given in terms of boson operators a, b by

A sublattice	B sublattice
$S^z_i = S^A - a^\dagger_i a_i$	$S^z_j = -S^B + b^\dagger_j b_j$
$S^+_i \equiv S^x_i + \mathrm{i}S^y_i \simeq (2S^A)^{1/2} a_i$	$S^+_j \equiv S^x_j + \mathrm{i}S^y_j \simeq (2S^B)^{1/2} b^\dagger_j$
$S^-_i \equiv S^x_i - \mathrm{i}S^y_i \simeq (2S^A)^{1/2} a^\dagger_i$	$S^-_j \equiv S^x_j - \mathrm{i}S^y_j \simeq (2S^B)^{1/2} b_j$

Discuss the validity of this approximation. Use these relations to express the Hamiltonian in terms of the boson operators to quadratic order.

Transforming to crystal momentum space using (with N the number of sites on one sublattice)

$$a_i = N^{-1/2} \sum_{\mathbf{k}} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}_i} a_{\mathbf{k}}, \quad b_j = N^{-1/2} \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_j} b_{\mathbf{k}}$$

show that your result can be expressed in the form

$$H = E_0 + \sum_{\mathbf{k}} \left[A_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + B_{\mathbf{k}} b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + C_{\mathbf{k}} (a^\dagger_{\mathbf{k}} b^\dagger_{\mathbf{k}} + b_{\mathbf{k}} a_{\mathbf{k}}) \right]$$

and determine the coefficients. Hence calculate the spectrum of excitations at low momenta. Consider both the cases with $S^A = S^B$ and $S^A \neq S^B$ and comment on your results.

10 Path Integral for interacting Bose systems

A key ingredient in our construction of the path integral for a single particle was the resolution of the identity

$$\mathbf{1} = \int dx \; |x\rangle\langle x| \; , \tag{537}$$

which allowed us to break up $e^{-\beta H}$ into a product over infinitesimal propagators.

10.1 Coherent States

In order to deal with many-boson systems, we require a convenient analog on the Fock space. This is provided by *coherent states*

$$\boxed{|\boldsymbol{\phi}\rangle = \exp\left(\sum_\ell \phi_\ell a^\dagger_\ell\right)|0\rangle \; , \quad \phi_\ell \in \mathbb{C} \; ,} \tag{538}$$

where a_ℓ denotes the bosonic annihilation operator for the single-particle state labeled by ℓ and $|0\rangle$ is the Fock vacuum. If N_{SP} is the number of single-particle states, then $\boldsymbol{\phi}$ is a N_{SP}-dimensional complex vector. The states (538) are simultaneous eigenstates of all annihilation operators

$$\boxed{a_j|\boldsymbol{\phi}\rangle = \phi_j|\boldsymbol{\phi}\rangle.} \tag{539}$$

In order to prove (539) consider

$$a_j|\boldsymbol{\phi}\rangle = a_j \exp\left(\sum_\ell \phi_\ell a^\dagger_\ell\right)|0\rangle = \left[a_j, \exp\left(\sum_\ell \phi_\ell a^\dagger_\ell\right)\right]|0\rangle = \exp\left(\sum_{\ell\neq j} \phi_\ell a^\dagger_\ell\right)\left[a_j, \exp\left(\phi_j a^\dagger_j\right)\right]|0\rangle. \tag{540}$$

The commutator is easily calculated by expanding the exponential in its power series

$$\left[a_j, \exp\left(\phi_j a^\dagger_j\right)\right] = \phi_j \exp\left(\phi_j a^\dagger_j\right), \tag{541}$$

and substituting this back into (540) establishes (539). Coherent states are *not* mutually orthogonal. In fact, they fulfil

$$\boxed{\langle\boldsymbol{\psi}|\boldsymbol{\phi}\rangle = e^{\sum_\ell \psi^*_\ell \phi_\ell}.} \tag{542}$$

This result for the scalar product can be obtained by applying the Baker-Campbell-Hausdorff (BCH) formula, which states that for two operators such that $[A,[A,B]] = 0 = [B,[A,B]]$ we have

$$e^A e^B = e^{A+B} e^{\frac{1}{2}[A,B]} = e^B e^A e^{[A,B]} \; . \tag{543}$$

Setting $A = \sum_\ell \psi_\ell^* a_\ell$, $B = \sum_j \phi_j a_j^\dagger$, using the BCH formula, and then noting that $A|0\rangle = 0 = \langle 0|B$, we obtain (542). While coherent states do not form an orthogonal set, they nevertheless provide a resolution of the identity on the Fock space

$$\boxed{\mathbf{1} = \int \underbrace{\prod_j \frac{d^2\phi_j}{\pi}}_{d(\boldsymbol{\phi},\boldsymbol{\phi}^*)} e^{-\sum_\ell |\phi_\ell|^2} \; |\boldsymbol{\phi}\rangle\langle\boldsymbol{\phi}|.} \tag{544}$$

Here $d^2\phi_\ell$ denotes the integration over the complex variable ϕ_ℓ, e.g. in polar co-ordinates we have

$$\int d^2\phi_j = \int_0^\infty dr_j r_j \int_0^{2\pi} d\varphi_j \;, \quad \phi_j = r_j e^{i\varphi_j}. \tag{545}$$

To prove (544) we note that

$$\begin{aligned} |\boldsymbol{\phi}\rangle &= e^{\phi_1 a_1^\dagger}\, e^{\phi_2 a_2^\dagger}\, e^{\phi_3 a_3^\dagger} \ldots |0\rangle = \sum_{n_1=0}^\infty \frac{(\phi_1)^{n_1}}{n_1!} \sum_{n_2=0}^\infty \frac{(\phi_2)^{n_2}}{n_2!} \ldots (a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} (a_3^\dagger)^{n_3} \ldots |0\rangle \\ &= \sum_{n_1=0}^\infty \frac{(\phi_1)^{n_1}}{\sqrt{n_1!}} \sum_{n_2=0}^\infty \frac{(\phi_2)^{n_2}}{\sqrt{n_2!}} \sum_{n_3=0}^\infty \frac{(\phi_3)^{n_3}}{\sqrt{n_3!}} \ldots |n_1 n_2 n_3 \ldots\rangle \;, \end{aligned} \tag{546}$$

where $|n_1 n_2 \ldots\rangle$ is a state in the occupation number representation. Hence

$$|\boldsymbol{\phi}\rangle\langle\boldsymbol{\phi}| = \sum_{n_1,n_2,\ldots} \sum_{m_1,m_2,\ldots} \frac{\phi_1^{n_1}(\phi_1^*)^{m_1}\phi_2^{n_2}(\phi_2^*)^{m_2}\cdots}{\sqrt{n_1!m_1!n_2!m_2!\ldots}} |n_1 n_2 n_3 \ldots\rangle\langle m_1 m_2 m_3 \ldots| \;. \tag{547}$$

Inspection of (544) and (547) shows that the integral over ϕ_j and ϕ_j^* is

$$\int_0^\infty dr_j \int_0^{2\pi} d\varphi_j \; r^{n_j+m_j+1} e^{-r_j^2} e^{i\varphi(n_j-m_j)} = n_j!\, \delta_{n_j,m_j} \;. \tag{548}$$

Carrying out all integrals we obtain

$$\int d(\boldsymbol{\phi},\boldsymbol{\phi}^*) e^{-\sum_\ell |\phi_\ell|^2} \; |\boldsymbol{\phi}\rangle\langle\boldsymbol{\phi}| = \sum_{n_1,n_2,\ldots} |n_1 n_2 n_3 \ldots\rangle\langle n_1 n_2 n_3 \ldots| \;. \tag{549}$$

The right hand side is a resolution of the identity in the occupation number representation.

10.2 Partition Function

Let us now consider a general many-boson Hamiltonian of the form

$$\boxed{\hat{H} = \sum_{i,j} h_{ij} a_i^\dagger a_j + \sum_{i,j,k,l} V_{ijkl} a_i^\dagger a_j^\dagger a_k a_l.} \tag{550}$$

We first want to derive a path integral representation for the partition function

$$Z(\beta) = \mathrm{Tr}\left[e^{-\beta\hat{H}}\right] = \sum_n \langle n|e^{-\beta\hat{H}}|n\rangle. \tag{551}$$

Inserting a resolution of the identity in terms of coherent states this can rewritten as

$$\begin{aligned} \int d(\boldsymbol{\psi},\boldsymbol{\psi}^*) e^{-\sum_\ell |\psi_\ell|^2} \sum_n \langle n|\boldsymbol{\psi}\rangle \; \langle\boldsymbol{\psi}|e^{-\beta\hat{H}}|n\rangle &= \int d(\boldsymbol{\psi},\boldsymbol{\psi}^*) e^{-\sum_\ell |\psi_\ell|^2} \sum_n \langle\boldsymbol{\psi}|e^{-\beta\hat{H}}|n\rangle\langle n|\boldsymbol{\psi}\rangle \\ &= \int d(\boldsymbol{\psi},\boldsymbol{\psi}^*) e^{-\sum_\ell |\psi_\ell|^2} \langle\boldsymbol{\psi}|e^{-\beta\hat{H}}|\boldsymbol{\psi}\rangle. \end{aligned} \tag{552}$$

In the next step we break up $e^{-\beta H}$

$$e^{-\beta H} = e^{-\epsilon H} e^{-\epsilon H} \dots e^{-\epsilon H} \ , \quad \epsilon = \frac{\beta}{N}, \tag{553}$$

and then insert resolutions of the identity between each of the factors. This leaves us with matrix elements of the form

$$\begin{aligned} \langle \boldsymbol{\psi}^{(n+1)} | e^{-\epsilon \hat{H}} | \boldsymbol{\psi}^{(n)} \rangle &= \langle \boldsymbol{\psi}^{(n+1)} | \mathbf{1} - \epsilon \hat{H} | \boldsymbol{\psi}^{(n)} \rangle + \mathcal{O}(\epsilon^2) \\ &= \langle \boldsymbol{\psi}^{(n+1)} | \boldsymbol{\psi}^{(n)} \rangle \left[1 - \epsilon H\big(\boldsymbol{\psi}^{(n+1)^*}, \boldsymbol{\psi}^{(n)} \big) \right] + \mathcal{O}(\epsilon^2) \\ &= \langle \boldsymbol{\psi}^{(n+1)} | \boldsymbol{\psi}^{(n)} \rangle \ e^{-\epsilon H\left(\boldsymbol{\psi}^{(n+1)^*}, \boldsymbol{\psi}^{(n)} \right)} + \mathcal{O}(\epsilon^2) \ , \end{aligned} \tag{554}$$

where

$$H\big(\boldsymbol{\psi}^*, \boldsymbol{\psi}'\big) = \sum_{i,j} h_{ij} \psi_i^* \psi_j' + \sum_{i,j,k,l} V_{ijkl} \psi_i^* \psi_j^* \psi_k' \psi_l'. \tag{555}$$

In going from the first to the second line in (554) we have used that

$$a_j | \boldsymbol{\psi'} \rangle = \psi_j' | \boldsymbol{\psi'} \rangle \ , \quad \langle \boldsymbol{\psi} | a_i^\dagger = \psi_i^* \langle \boldsymbol{\psi} | \ . \tag{556}$$

After these steps we end up with a representation of the form

$$Z(\beta) = \lim_{N\to\infty} \int \prod_{m=1}^{N} \int d(\boldsymbol{\psi}^{(m)}, \boldsymbol{\psi}^{(m)^*}) \ \exp\left[-\epsilon \sum_{n=0}^{N-1} \left(\frac{\boldsymbol{\psi}^{(n)^*} - \boldsymbol{\psi}^{(n+1)^*}}{\epsilon} \right) \cdot \boldsymbol{\psi}^{(n)} + H\big(\boldsymbol{\psi}^{(n+1)^*}, \boldsymbol{\psi}^{(n)} \big) \right]. \tag{557}$$

In complete analogy to what we did in the single-particle case, we can now interpret the sequence $\boldsymbol{\psi}^{(1)}$, $\boldsymbol{\psi}^{(2)}$,$\dots$ $\boldsymbol{\psi}^{(N-1)}$ as a discretization of a path on the space of N_{SP} dimensional complex vectors

$$\boldsymbol{\psi}(\tau_n) = \boldsymbol{\psi}^{(n)} \ , \quad \tau_n = n\epsilon. \tag{558}$$

In the limit $N \to \infty$ this goes over into a vector-valued function of imaginary time $\boldsymbol{\psi}(\tau)$, and the partition function acquires the following formal expression

$$Z(\beta) = \int \mathcal{D}\big(\boldsymbol{\psi}^*(\tau), \boldsymbol{\psi}(\tau) \big) \ e^{-S[\boldsymbol{\psi}^*(\tau), \boldsymbol{\psi}(\tau)]} \ , \tag{559}$$

where the action S is given by

$$S[\boldsymbol{\psi}^*(\tau), \boldsymbol{\psi}(\tau)] = \int_0^\beta d\tau \left[\sum_{i,j} \psi_i^*(\tau) \left[\delta_{ij} \partial_\tau + h_{ij} \right] \psi_j(\tau) + \sum_{i,j,k,l} V_{ijkl} \psi_i^*(\tau) \psi_j^*(\tau) \psi_k(\tau) \psi_l(\tau) \right]. \tag{560}$$

Part IV
Landau Theory of Phase Transitions

11 Phase Transitions

Physically a phase transition is a point in parameter space, where the physical properties of a many-particle system undergo a sudden change. An example is the paramagnet to ferromagnet transition in Fe or Ni shown in Fig. 21.

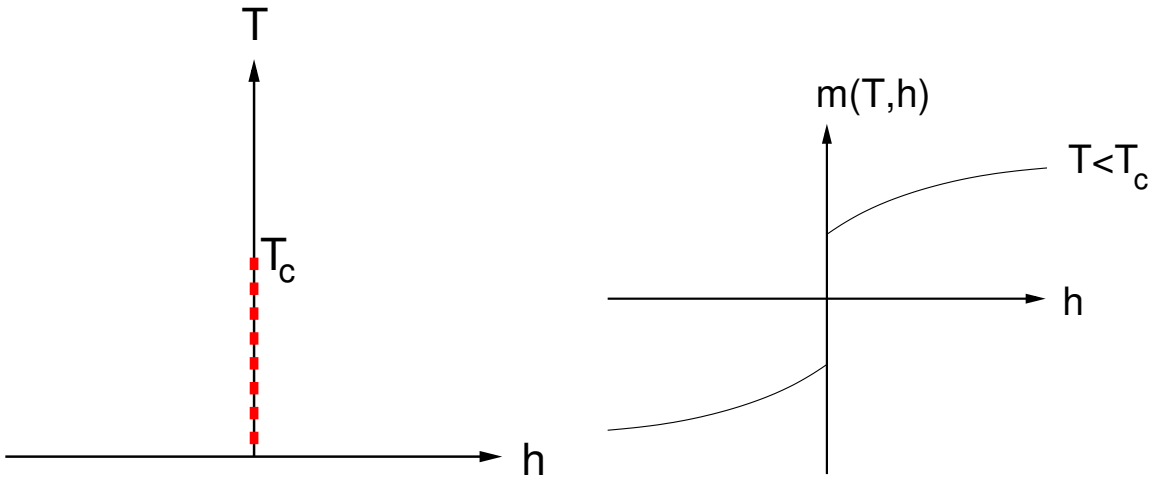


Figure 21: Phase diagram for the paramagnetic to ferromagnetic transition. The *magnetization per site* $m(T,h)$ jumps when crossing zero for $T < T_c$.

Mathematically a phase transition is a point in parameter space, where the free energy $F = -k_B T \ln(Z)$ becomes a nonanalytic function of one of its parameters (i.e. F or some of its derivatives becomes singular or discontinuous) in the *thermodynamic limit*.

For a finite system this can never happen, because

$$Z = \sum_{\text{configurations C}} e^{-E(C)/k_B T} \tag{561}$$

is a finite sum over finite, positive terms. Hence all derivatives are finite and well defined as well.

Phase transitions are usually divided into two categories:

1. *First Order Phase Transitions.*

   Here the free energy is continuous, but a first derivative is discontinuous. At the transition there is *phase coexistence*. The magnetization per site is a first order derivative of the free energy with respect to the magnetic field h. Therefore the phase transition at $h = 0$ and $T < T_c$ in Fig. 21 is first order.

2. *Second Order Phase Transitions.*

   These are characterized by a divergence in one of the higher order derivatives ("susceptibilities") of the free energy. The phase transition as a function of T for $h = 0$ in Fig. 21 is second order.

12 The Ising Model

The two-dimensional square-lattice Ising model is one of the simplest statistical models to show a phase transition and as a result has played a hugely important role in the development of the theory of critical

phenomena. The model can in fact be solved exactly by the transfer matrix method introduced above. Interestingly, it is possible to establish the existence of a finite-temperature phase transition in the model without solving it. This was pioneered by Sir Rudolf Peierls, who was the Wykeham Professor of Theoretical Physics at Oxford for many years and during his career made numerous ground breaking contributions in particular to condensed matter physics.

12.1 Peierls Argument

The Peierls argument is a nice way of establishing that the 2D square lattice Ising model has magnetic long-range order at sufficiently low temperatures *without actually solving the model*. Given that at very high temperature there is no magnetic order, this shows that there must be at least one phase transition at a finite temperature.

Consider the Ising model on the square lattice with boundary conditions such that all spins on the boundary are up, i.e. take the value $+1$. You can think of these boundary conditions as a symmetry breaking field. The bulk magnetic field is taken to be zero. Configurations look like the one shown in Fig. 22, and can be characterized by *domain walls*. These are lines separating $+$ and $-$ spins such that

1. The $+$ ($-$) spins lie always to the left (right) of the wall.
2. Where ambiguities remain, the wall is taken to bend to the right.
3. The *length* of the wall is defined as the number of lattice spacings it traverses.

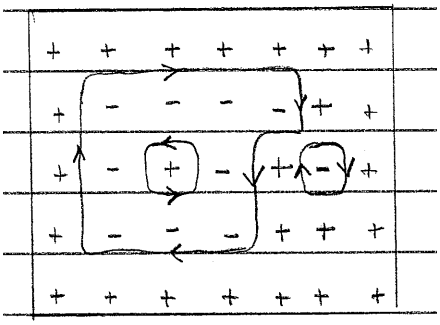


Figure 22: A configuration of spins and the corresponding domain walls.

A wall of length b encloses at most $b^2/16$ spins. The total number of domain walls of length b, $m(b)$, is bounded by

$$\boxed{m(b) \leq 4N_t 3^{b-1},} \tag{562}$$

where N_t is the total number of sites. This can be seens as follows:

- the first link can go into less than $4N_t$ positions (starting at an arbitrary site and going in any of the four possible directions).
- subsequent links have at most 3 possible directions each.

Let us denote the i'th domain wall of length b by (b, i). Next consider a particular configuration $\boldsymbol{\sigma} = \{\sigma_{j,k}\}$ of spins on the lattice, and define

$$X_{\boldsymbol{\sigma}}(b,i) = \begin{cases} 1 & \text{if (b,i) occurs in } \boldsymbol{\sigma} \\ 0 & \text{else} \end{cases} \tag{563}$$

Then the total number of $-$ spins in $\boldsymbol{\sigma}$ is bounded by

$$N_- \leq \sum_b \frac{b^2}{16} \sum_{i=1}^{m(b)} X_{\boldsymbol{\sigma}}(b,i), \tag{564}$$

because each spin is enclosed by at least one domain wall due to our choice of boundary conditions. Taking thermal averages, we have

$$\langle N_- \rangle_\beta \leq \sum_b \frac{b^2}{16} \sum_{i=1}^{m(b)} \langle X_{\boldsymbol{\sigma}}(b,i) \rangle_\beta. \tag{565}$$

Here the thermal average of $X_{\boldsymbol{\sigma}}(b,i)$ can be written as

$$\langle X_{\boldsymbol{\sigma}}(b,i) \rangle_\beta = \frac{1}{Z} {\sum_\sigma}' e^{-\beta E(\boldsymbol{\sigma})}, \tag{566}$$

where the sum is only over configurations, in which (b,i) occurs. Now consider the configuration $\boldsymbol{\sigma}'$ obtained from $\boldsymbol{\sigma}$ by reversing the spins inside the domain wall (b,i). Clearly the energies of the two configurations

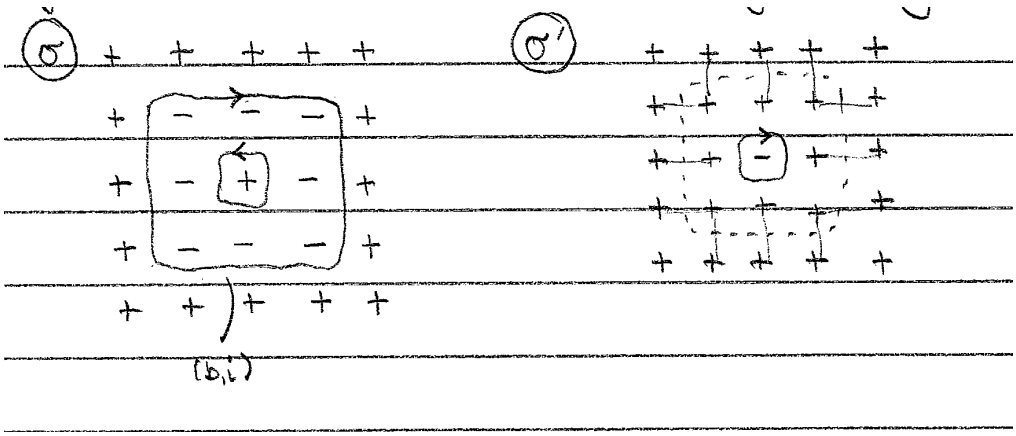


Figure 23: Configurations $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ related by reversing all spins inside the domain wall (b,i). Shown are all the bonds whose energies have been changed from $-J$ to J.

are related by

$$E(\boldsymbol{\sigma}) = E(\boldsymbol{\sigma}') + 2bJ. \tag{567}$$

This can be used to obtain a bound on Z

$$Z \geq {\sum_{\boldsymbol{\sigma}'}}' e^{-\beta E(\boldsymbol{\sigma}')} \geq {\sum_{\boldsymbol{\sigma}}}' e^{-\beta E(\boldsymbol{\sigma})} e^{2bJ\beta}, \tag{568}$$

where the first sum is only over configurations in which (b,i) occurs, and where we then have flipped all spins inside the domain wall. This gives us a bound on

$$\langle X_{\boldsymbol{\sigma}}(b,i)\rangle_\beta = \frac{1}{Z}{\sum_\sigma}' e^{-\beta E(\boldsymbol{\sigma})} \leq e^{-2bJ\beta}. \tag{569}$$

Now we put everything together

$$\langle N_-\rangle_\beta \;\leq\; \sum_b \frac{b^2}{16}\sum_{i=1}^{m(b)} e^{-2\beta Jb} \leq \sum_b \frac{b^2}{16} 4N_t 3^{b-1} e^{-2\beta Jb} = \frac{N_t}{12}\sum_{b=4,6,8,\ldots} b^2\left[3e^{-2\beta J}\right]^b. \tag{570}$$

The sum over b can now be easily carried out, and the results at small T (large β) is

$$\boxed{\langle N_-\rangle_\beta \leq 108 N_t e^{-8\beta J}.} \tag{571}$$

So, at low temperatures we have

$$\boxed{\frac{\langle N_-\rangle_\beta}{N_t} \ll \frac{1}{2}.} \tag{572}$$

This proves the existence of a spontaneous magnetization at low temperatures.

12.2 Mean Field Theory

Consider the Ising model on a D-dimensional lattice with coordination number (number of nearest neighbour sites) z

$$E = -J\sum_{\langle ij\rangle}\sigma_i\sigma_j - h\sum_j \sigma_j. \tag{573}$$

Here $\langle ij\rangle$ denotes nearest neighbour bonds, and each bond is counted once. The magnetization per site is

$$m = \frac{1}{N}\sum_{j=1}^N \langle\sigma_j\rangle_\beta. \tag{574}$$

We now rewrite the energy using

$$\sigma_i = m + (\sigma_i - m). \tag{575}$$

In particular we have

$$\sigma_i\sigma_j = m^2 + m(\sigma_j - m) + m(\sigma_i - m) + (\sigma_i - m)(\sigma_j - m). \tag{576}$$

The idea of the *mean-field approximation* is to assume that the deviations $\sigma_j - m$ of the spins from their average values are small, and to neglect the terms quadratic in these fluctuations. This gives

$$\boxed{E_{\rm MF} = -J\sum_{\langle ij\rangle} -m^2 + m(\sigma_i + \sigma_j) - h\sum_j \sigma_j.} \tag{577}$$

Physically, what we have done is to replace the interaction of a given spin with its neighbours by *an average magnetic field*. We can simplify (577) further by noting that

$$\begin{aligned} -J\sum_{\langle ij\rangle} -m^2 &= Jm^2\frac{Nz}{2}\,, \\ \sum_{\langle ij\rangle}\sigma_i + \sigma_j &= z\sum_j \sigma_j. \end{aligned} \tag{578}$$

The mean-field energy then becomes

$$\boxed{E_{\rm MF} = \frac{JNz}{2}m^2 - (Jzm+h)\sum_{j=1}^{N}\sigma_j.} \tag{579}$$

The partition function in the mean-field approximation is

$$\begin{aligned} Z_{\rm MF} = \sum_{\{\sigma_j\}} e^{-\beta E_{\rm MF}} &= e^{-\frac{NJz\beta m^2}{2}} \sum_{\sigma_1}\cdots\sum_{\sigma_N}\prod_{j=1}^{N} e^{\beta(Jzm+h)\sigma_j} \\ &= e^{-\frac{NJz\beta m^2}{2}} \left[\sum_{\sigma_1} e^{\beta(Jzm+h)\sigma_1}\right]\cdots\left[\sum_{\sigma_N} e^{\beta(Jzm+h)\sigma_N}\right] \\ &= e^{-\frac{NJz\beta m^2}{2}} \left[2\cosh(Jzm\beta+h\beta)\right]^N. \end{aligned} \tag{580}$$

The magnetization per site is

$$m = \frac{1}{N}\sum_{j=1}^{N}\langle\sigma_j\rangle_\beta = \langle\sigma_N\rangle_\beta, \tag{581}$$

where we have used translational invariance in the last step. In mean field theory we have

$$\boxed{m = \langle\sigma_j\rangle_\beta = \frac{1}{Z_{\rm MF}} e^{-\frac{NJz\beta m^2}{2}} \sum_{\sigma_1}\cdots\sum_{\sigma_N}\sigma_N\prod_{j=1}^{N} e^{\beta(Jzm+h)\sigma_j} = \tanh(Jzm\beta+\beta h).} \tag{582}$$

This is a *self-consistency equation* for m.

12.3 Solution of the self-consistency equation for $h = 0$

For zero field the self-consistency equation reads

$$m = \tanh(Jzm\beta). \tag{583}$$

This can be solved graphically by looking for intersections of the functions $g_1(m) = \tanh(Jzm\beta)$ and $g_2(m) = m$. There are either one or three solutions

$$m = \begin{cases} 0 & \text{if } Jz\beta < 1 \\ \pm m_0, 0 & \text{if } Jz\beta > 1 \end{cases}. \tag{584}$$

We still have to check whether these solutions correspond to mimima of the free energy per site. The latter is

$$f_{\rm MF} = -\frac{1}{\beta N}\ln(Z_{\rm MF}) = \frac{Jzm^2}{2} - \frac{1}{\beta}\ln[2\cosh(Jzm\beta)]. \tag{585}$$

We have

$$\frac{\partial^2 f_{\rm MF}}{\partial m^2} = Jz\left[1 - \frac{Jz\beta}{\cosh(Jzm\beta)}\right]. \tag{586}$$

This is negative for $m = 0$ and $Jz\beta > 1$, and hence this solution corresponds to a *maximum* of the free energy and hence must be discarded. This leaves us with

$$\boxed{m = \begin{cases} 0 & \text{if } T > T_c \\ \pm m_0 & \text{if } T < T_c \end{cases},} \tag{587}$$

where the transition temperature is

$$T_c = \frac{Jz}{k_B}. \tag{588}$$

12.4 Vicinity of the Phase Transition

Let us define a dimensionless variable, that measures the distance in temperature to the phase transition

$$t = \frac{T - T_c}{T_c}. \tag{589}$$

For $|t| \ll 1$ we obtain the following results

1. Magnetization per site

$$\boxed{m\big|_{h=0} \simeq \begin{cases} 0 & \text{if } T > T_c, \\ \sqrt{-3t} & \text{if } T < T_c. \end{cases}} \tag{590}$$

2. Magnetic susceptibility in zero field

$$\boxed{\chi = \frac{\partial m}{\partial h}\Big|_{h=0} \simeq \begin{cases} \frac{1}{k_B T_c} t^{-1} & \text{if } T > T_c, \\ \frac{1}{2k_B T_c}(-t)^{-1} & \text{if } T < T_c, \end{cases}} \tag{591}$$

3. Free energy per site and heat capacity per volume

$$\boxed{f_{\mathrm{MF}}\big|_{h=0} \simeq -k_B T \ln 2 + \begin{cases} 0 & \text{if } T > T_c \\ -\frac{3k_B T_c}{4} t^2 & \text{if } T < T_c. \end{cases}} \tag{592}$$

$$\boxed{\frac{C}{V} = -T \frac{\partial^2 f_{\mathrm{MF}}}{\partial T^2}\Big|_{h=0} \simeq \begin{cases} 0 & \text{if } T > T_c \\ \frac{3k_B}{2} & \text{if } T < T_c. \end{cases}} \tag{593}$$

13 Critical Behaviour and Universality

Close to a critical point thermodynamic functions display power-law behaviours characterized by *critical exponents*. We will now discuss various such exponents, using as a specific example the paramagnet to ferromagnet transition.

1. Order Parameter

   This is a quantity that is different in the various phases and can be used to characterize the phase transition. For the paramagnet to ferromagnet transition in zero magnetic field the appropriate order parameter is the magnetization per site

$$\boxed{m(T) = \lim_{h \to +0} \lim_{V \to \infty} \frac{1}{V} M(h, T).} \tag{594}$$

   Here $M(h, T)$ is the magnetization. Where $T \approx T_c$, one has

$$\boxed{m(T) \sim \begin{cases} 0 & \text{if } T > T_c \\ |t|^\beta & \text{if } T < T_c \end{cases} \qquad t = \frac{T - T_c}{T_c}.} \tag{595}$$

   β is a *critical exponent*.

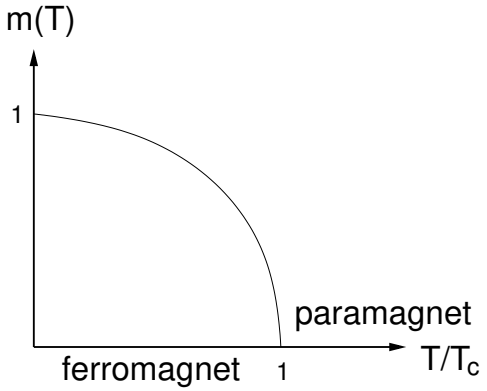


Figure 24: Order parameter for the paramagnet to ferromagnet transition in zero field.

2. Susceptibilities

At the critical point the system is very sensitive to external perturbations. The singularity in the response of the order parameter to a field "conjugate" to it is characterized by critical exponents $\gamma_\pm$. For our magnet

$$\boxed{\chi_\pm(T) = \frac{\partial}{\partial h}\Big|_{h=0} \lim_{V\to\infty} \frac{1}{V} M(h,T) \sim |t|^{-\gamma_\pm}.} \tag{596}$$

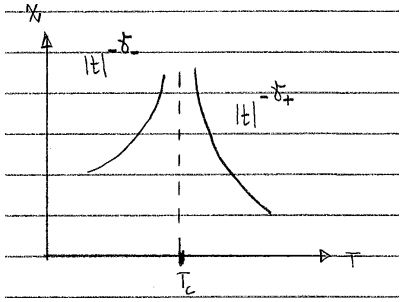


Figure 25: Critical behaviour of the magnetic susceptibility.

3. Heat Capacity

A third critical exponent is associated with the *heat capacity*

$$\boxed{C(T) = -T\frac{\partial^2 F}{\partial T^2} \sim \begin{cases} A_+|t|^{-\alpha_+} & \text{if } T > T_c \\ A_-|t|^{-\alpha_-} & \text{if } T < T_c. \end{cases}} \tag{597}$$

Depending on the signs of $\alpha_\pm$ this may or may not be singular, see Fig. 26.

13.1 Universality

The critical exponents are *insensitive* to microscopic details of the system under consideration and are characteristic of the critical point. A consequence of this is that completely different systems can exhibit

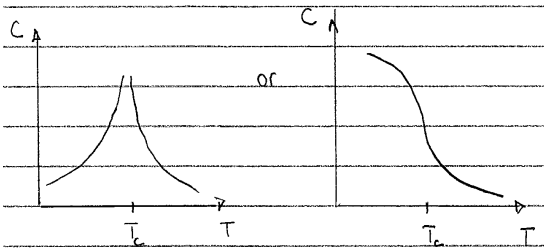


Figure 26: Critical behaviour of the heat capacity.

the same critical behaviour!

14 Landau Theory

Landau Theory is a general approach to phase transitions that

- is phenomenological in nature and deals only with macroscopic quantities;
- applies only to the neighbourhood of a critical point, where the order parameter is small.

Landau theory is constructed as follows.

1. Identify the order parameter(s) $M(\mathbf{r})$ characterizing the phase transition. Depending on which phase transition we are dealing with M can be a real scalar, a complex scalar, a real or complex vector or something more complicated. For the paramagnet to ferromagnet transition the order parameter is the magnetization per site, i.e. a real number.

2. Form a "coarse-grained" order parameter density $\Phi(\mathbf{r})$. Think of this as the microscopic order parameter averaged over atomic distances. This is a *continuum field*. For the example of the paramagnet to ferromagnet transition we are dealing with a real scalar field.

3. Consider the free energy density to be a *functional* of the order parameter field $\Phi(\mathbf{r})$. The free energy is then

$$\boxed{\beta F = \int d^D\mathbf{r} f[\Phi(\mathbf{r})].} \tag{598}$$

4. By construction of the order parameter(s), the latter is (are) *small* close to our critical point. This allows us to expand $f[\Phi(\mathbf{r})]$ is a power series around $\Phi = 0$. From now on we will focus on the simplest case of a real scalar order parameter $\phi(\mathbf{r})$. Then the series expansion is

$$\boxed{f[\phi] \simeq \text{const} - h\phi + \alpha_2\phi^2 + \frac{1}{2}|\nabla\phi|^2 + \alpha_3\phi^3 + \alpha_4\phi^4 + \dots} \tag{599}$$

where the coefficient of the gradient term is fixed by convention to be $1/2$. This makes ϕ in general dimensionful

$$\dim[\phi(\mathbf{r})] = (\text{length})^{1-D/2}. \tag{600}$$

The only linear term that is not a total derivative is $-h\phi$, where h is an external field (a "source") coupling to the order parameter. Total derivative terms can be dropped, because the only give boundary contributions to F. *The coefficients α_n are a priori all functions of temperature.*

5. In translationally invariant systems the free energy is minimized by **r**-independent order parameters (i.e. $\nabla\phi(\mathbf{r}) = 0$). The reason is that $\frac{1}{2}|\nabla\phi|^2 \geq 0$, and hence this contribution to F is minimized by constant solutions. For constant field h the potential $V(\phi(\mathbf{r})) = -h\phi + \alpha_2\phi^2 + \alpha_4\phi^4$ is also minimized by constant solutions. In order to understand the nature of the phase transition, we therefore can simply look at the minima of the potential $V(\phi)$.

6. Finally, we use symmetries and the fact that we are interested in the vicinity of a critial point to constrain the α_j.

   - If we truncate our expansion at order ϕ^4, then thermodynamic stability requires

   $$\boxed{\alpha_4 > 0.} \tag{601}$$

   If $\alpha_4 < 0$ the free energy density would be unbounded from below and could become infinitely negative, which is forbidden.

   - If we know that the system is invariant under certain symmetry operations, e.g.

   $$\phi \to -\phi, \tag{602}$$

   then the free energy must respect this symmetry. A ferromagnet has the symmetry (602) in absence of a magnetic field because of *time-reversal invariance*. Hence we must have $\alpha_3 = 0$ in this case.

   - In the case $h = \alpha_3 = 0$, for a translationally invariant system, we can obtain the temperature dependence of α_2 as follows. As discussed above, the nature of the phase transition can be inferred from the minima of the potential $V(\phi) = \alpha_2\phi^2 + \alpha_4\phi^4$. This is done in Fig. 27. We see that the phase transition corresponds to α_2 changing sign at $T = T_c$. So in the vicinity of the transition we have (by Taylor expanding α_2 in $T - T_c$)

   $$\boxed{\alpha_2(t) = At + \mathcal{O}(t^2)\ , \qquad t = \frac{T - T_c}{T_c}\ ,\ A > 0.} \tag{603}$$

   The parameter α_4 is also temperature dependent, but this dependence is subleading

   $$\alpha_4(t) = \alpha_4(0) + \mathcal{O}(t). \tag{604}$$

7. If we have $\alpha_3 < 0$ the transition is generically first order. To see this we again use that in a translationally invariant system the minima of the free energy density will be **r**-independent, so that we merely need to scrutinize the potential $V(\phi)$ to understand the nature of the phase transition. In Fig. 28 we plot $V(\phi)$ when α_2 is decreased at fixed α_3, α_4. We see that initially the minimum occurs at $\phi = 0$ (no order), and at some critical value of α_2 then jumps from zero to a finite value ϕ_0. This is characteristic of a first order transition.

14.1 Thermodynamic Equilibrium

The state of thermodynamic equilibrium is obtained by minimizing the free energy (we assume that $\alpha_3 = 0$)

$$\beta F = \int d^D\mathbf{r} \left[-h\phi + \alpha_2\phi^2 + \frac{1}{2}|\nabla\phi|^2 + \alpha_4\phi^4 \right]. \tag{605}$$

In our case we are searching for the order parameter configuration $\phi(\mathbf{r})$ that gives the smallest contribution to βF. This is found by functional extremization

$$\frac{\delta F}{\delta\phi(\mathbf{r})} = 0. \tag{606}$$

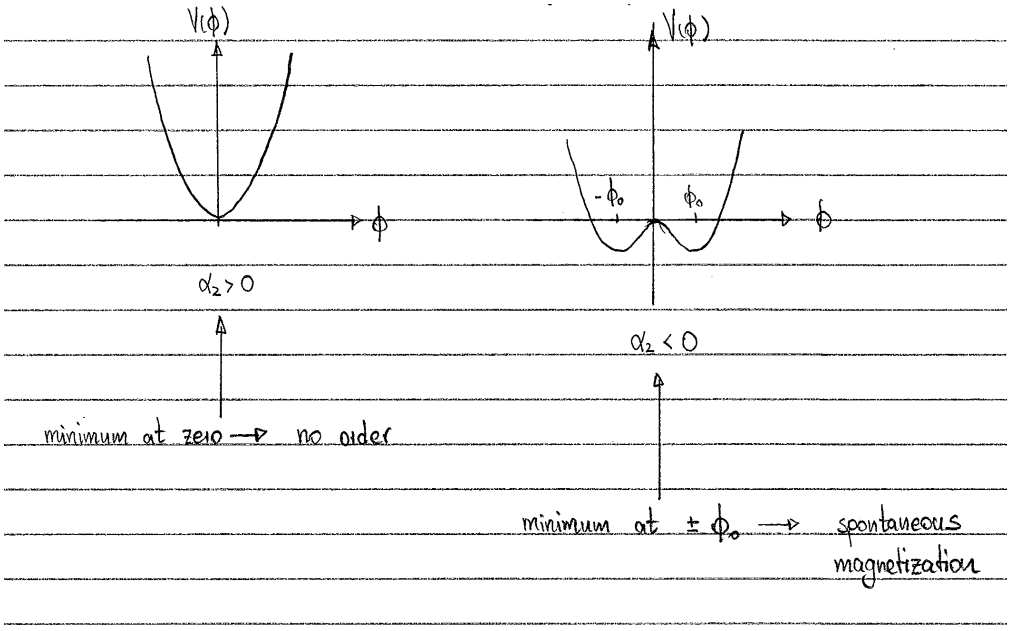


Figure 27: Minima of the potential $V(\phi)$. For $\alpha_2 > 0$ the minima occurs at $\phi = 0$, so there is no ferromagnetic order. For $\alpha_2 < 0$ there are two minima at $\phi = \pm\phi_0$, corresponding to the emergence of a spontaneous magnetization.

The resulting nonlinear differential equation is

$$\boxed{-\nabla^2\phi(\mathbf{r}) + 2\alpha_2\phi(\mathbf{r}) + 4\alpha_4\phi^3(\mathbf{r}) - h = 0.} \tag{607}$$

It is easy to see (see the discussion above) that the solutions to (607) that minimize the free energy are in fact $\mathbf{r}$-independent (if we ignore boundary conditions). Hence for zero field $h = 0$ the order parameter configuration that minimizes the free energy is

$$\boxed{\phi(\mathbf{r}) = \begin{cases} 0 & \text{if } \alpha_2 > 0 \leftrightarrow T > T_c \\ \pm\phi_0 = \pm\sqrt{-\frac{\alpha_2}{2\alpha_4}} & \text{if } \alpha_2 < 0 \leftrightarrow T < T_c \end{cases}.} \tag{608}$$

We observe that we are dealing with a second order phase transition (because ϕ_0 vanishes at the critical point) from a paramagnetic to a ferromagnetic phase.

14.2 Beyond the Landau Free Energy

So far we have focussed entirely on the state of thermodynamic equilibrium. We now want to extend Landau theory to take into account fluctuations as well. Let us go back to the microscopic model underlying our Landau free energy. The partition function for this microscopic theory is

$$Z_{\text{micro}} = \sum_{\text{configurations C}} e^{-\beta E(C)} \ . \tag{609}$$

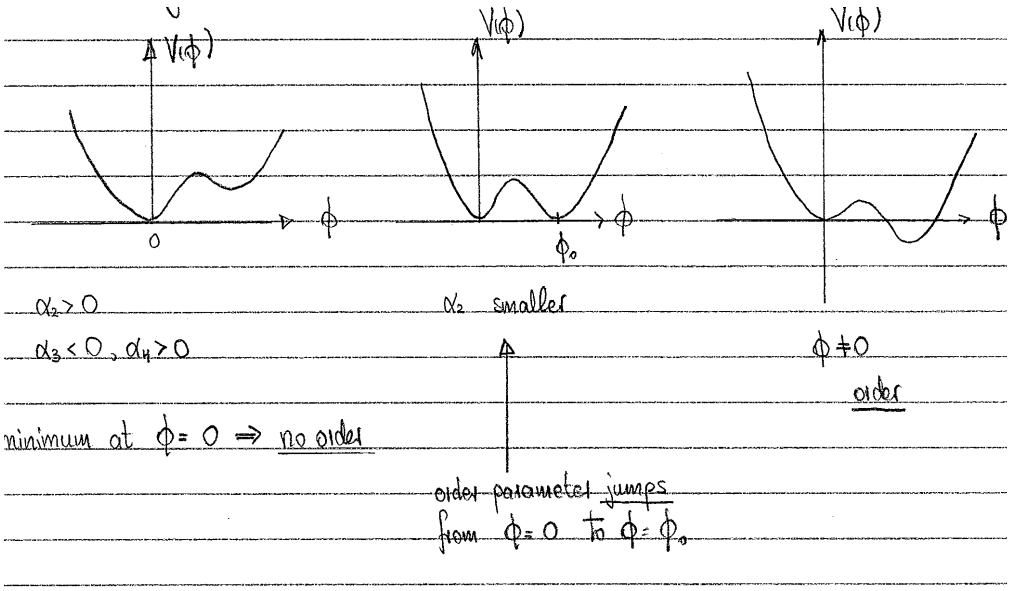


Figure 28: Minima of the potential $V(\phi)$ for $\alpha_3 < 0$. Decreasing the value of α_2 leads to a discontinuous jump in the order parameter at some critical value. The transition is therefore first order.

In order to define our order-parameter field, we used a coarse-graining procedure. Hence after coarse-graining Z should become

$$Z \longrightarrow \sum_{\substack{\text{order parameter} \\ \text{configurations}}} e^{-\beta\mathcal{H}} \ . \tag{610}$$

Because under the coarse-graining procedure many microscopic configurations C map onto the same order parameter configuration $\phi(\mathbf{r})$, the "Landau-Ginzburg Hamiltonian" $\beta\mathcal{H}$ incorporates certain entropic effects. By employing the same logic as before, we can argue that

$$\beta\mathcal{H} = \int d^D\mathbf{r} \ f[\phi(\mathbf{r})] \ , \tag{611}$$

where $f[\phi(\mathbf{r})]$ is the same functional we constructed when considering the Landau free energy. As the order parameter is really a continuous field, what we mean by the sum in (610) is really the *functional integral*

$$Z = \int \mathcal{D}\phi(\mathbf{r}) \ e^{-\beta\mathcal{H}}. \tag{612}$$

The latter is defined as follows:

- We first discretize our D-dimensional spatial variable

$$\mathbf{r} \longrightarrow a_0\mathbf{n} = a_0(n_1, \ldots, n_D), \tag{613}$$

  where a_0 is a lattice spacing and the total number of points on our discrete grid is N^D.

- We then discretize the order parameter field and its derivatives

$$\begin{aligned} \phi(\mathbf{r}) &\longrightarrow \phi_{\mathbf{n}} \ , \\ \partial_{r_j}\phi(\mathbf{r}) &\longrightarrow \frac{\phi_{\mathbf{n}+\mathbf{e}_j} - \phi_{\mathbf{n}}}{a_0}, \end{aligned} \tag{614}$$

  where $\mathbf{e}_j$ are unit vectors in the j-direction.

- The Landau-Ginzburg Hamiltonian is discretized as

$$\beta\mathcal{H} \longrightarrow \sum_{\mathbf{n}} f[\phi_{\mathbf{n}}]a_0^D. \tag{615}$$

- The functional integral is then defined as follows

$$\boxed{Z = \int \mathcal{D}\phi(\mathbf{r})\ e^{-\beta\mathcal{H}} \equiv \lim_{N\to\infty} \int \prod_{\mathbf{n}} d\phi_{\mathbf{n}} e^{-\sum_{\mathbf{m}} f[\phi_{\mathbf{m}}]a_0^D}.} \tag{616}$$

Crucially, the functional integral defined in this way can be manipulated according to the same rules we derived for path-integrals in Quantum Mechanics. In this new way of looking at things we now can analyze properties that are not directly related to the free energy. For example, we may ask about properties of correlation functions like

$$\langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta \equiv \frac{1}{Z}\int \mathcal{D}\phi\ \phi(\mathbf{r})\phi(\mathbf{0})\ e^{-\beta\mathcal{H}}. \tag{617}$$

14.3 Saddle Point Approximation

The Landau-Ginzburg field theory

$$Z = \int \mathcal{D}\phi(\mathbf{r})\ e^{-\int d^D\mathbf{r} f[\phi(\mathbf{r})]}, \tag{618}$$

is still difficult to analyze. For the example we have discussed, it reduces to the Euclidean space version of the $\lambda\phi^4$ theory you have encountered in the field theory part of the course. In order to proceed we therefore resort to further approximations. The *saddle-point approximation* takes into account the thermodynamically most likely configuration $\phi(\mathbf{r})$, i.e. the configuration that minimizes

$$\beta\mathcal{H} = \int d^D\mathbf{r} f[\phi(\mathbf{r})]. \tag{619}$$

We see that the saddle-point approximation precisely recovers the results of the Landau free energy approach! However, using our new formulation we now go beyond this approximation and take into account fluctuations. We will see below how to do this.

14.4 Mean Field Exponents

Using the saddle point solution we can determine the corresponding approximation for the critical exponents.

- Order parameter.

  Using that $\alpha_2 = At$ for $t = (T - T_c)/T_c < 0$, we have

$$\boxed{\phi_0 = \pm\sqrt{\frac{A}{2\alpha_4}}|t|^{\frac{1}{2}}.} \tag{620}$$

  This gives the critical exponent

$$\boxed{\beta = \frac{1}{2}.} \tag{621}$$

- Magnetic susceptibility.

  Differentiating (607) with respect to h gives for $\mathbf{r}$-independent solutions

$$2\alpha_2 \frac{\partial \phi}{\partial h} + 12\alpha_4 \frac{\partial \phi}{\partial h}\phi^2 = 1. \tag{622}$$

  The zero-field susceptibility is thus

$$\chi = \frac{\partial \phi}{\partial h}\Big|_{h=0} = \frac{1}{2\alpha_2 + 12\alpha_4 \phi^2}. \tag{623}$$

  Using that $\alpha_2 = At$ and $\phi_0^2 = A|t|/2\alpha_4$ this becomes

$$\boxed{\chi = \begin{cases} \frac{1}{2At} & \text{if } t > 0, \\ \frac{1}{4A|t|} & \text{if } t < 0. \end{cases}} \tag{624}$$

  This gives the critical exponents

$$\boxed{\gamma_\pm = 1.} \tag{625}$$

- Heat capacity.

  The heat capacity is defined by

$$C(T, h = 0) = -T \frac{\partial^2 F}{\partial T^2}. \tag{626}$$

  The saddle point contribution to the free energy is (letting $\beta \approx \beta_c$ to obtain F):

$$\frac{F}{V} \sim \begin{cases} 0 & \text{if } t > 0, \\ -\frac{A^2 k_B T_c t^2}{4\alpha_4} & \text{if } t < 0, \end{cases} \tag{627}$$

  giving

$$\boxed{C(T, h = 0) \sim \begin{cases} 0 & \text{if } t > 0, \\ \frac{A^2 k_B V}{2\alpha_4} & \text{if } t < 0, \end{cases}} \tag{628}$$

  Note that we have set $T \sim T_c$ in the prefactor. This has a finite jump at T_c, which corresponds to the critical exponent

$$\boxed{\alpha = 0.} \tag{629}$$

- Correlation length exponent.

  The exponents described above can all be obtained from the saddle point solution, or equivalently the Landau free energy. This is not the case for the *correlation length exponent* ν, which is related to fluctuations around the saddle point.

  Away from the critical point the (connected) order-parameter two-point function decays exponentially with distance

$$\langle \phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta - \langle \phi(\mathbf{r})\rangle_\beta \langle \phi(\mathbf{0})\rangle_\beta = e^{-|\mathbf{r}|/\xi}\ , \qquad |\mathbf{r}| \to \infty. \tag{630}$$

  The correlation length ξ diverges when the critical point is approached

$$\boxed{\xi \sim |t|^{-\nu}.} \tag{631}$$

The relation (631) defines the exponent ν. We now determine ν in what is known as the *Gaussian approximation*. In the disordered phase this amounts to simply dropping the ϕ^4 term in the free energy density, i.e. by setting

$$\langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta \approx \frac{1}{Z}\int \mathcal{D}\phi\ \phi(\mathbf{r})\phi(\mathbf{0})\exp\left(-\int d^D\mathbf{r}'\left[\frac{1}{2}|\nabla\phi(\mathbf{r}')|^2+\alpha_2\phi^2(\mathbf{r}')\right]\right). \tag{632}$$

It is not a priori clear that the Gaussian approximation will give a good account of th two point function. It turns out to be good if the spatial dimensionality D is sufficiently high.

The two-point function (632) can be calculated using a *generating functional*

$$Z[h]=\int \mathcal{D}\phi\ \exp\left(-\int d^D\mathbf{r}'\left[\frac{1}{2}|\nabla\phi(\mathbf{r}')|^2+\alpha_2\phi^2(\mathbf{r}')-h(\mathbf{r}')\phi(\mathbf{r}')\right]\right). \tag{633}$$

Indeed, we have

$$\langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta = \frac{\delta}{\delta h(\mathbf{r})}\frac{\delta}{\delta h(\mathbf{0})}\Big|_{h=0}\ln(Z[h]). \tag{634}$$

We calculate the generating functional by going to Fourier space

$$\phi(\mathbf{r})=\int\frac{d^D\mathbf{p}}{(2\pi)^D}e^{-i\mathbf{p}\cdot\mathbf{r}}\tilde{\phi}(\mathbf{p}),\quad h(\mathbf{r})=\int\frac{d^D\mathbf{p}}{(2\pi)^D}e^{-i\mathbf{p}\cdot\mathbf{r}}\tilde{h}(\mathbf{p}). \tag{635}$$

This gives

$$\beta\mathcal{H}=\int\frac{d^D\mathbf{p}}{(2\pi)^D}\left[\left(\frac{\mathbf{p}^2}{2}+\alpha_2\right)\tilde{\phi}(\mathbf{p})\tilde{\phi}(-\mathbf{p})-h(\mathbf{p})\tilde{\phi}(-\mathbf{p})\right]. \tag{636}$$

Next we "complete the square" by changing variables to

$$\tilde{\varphi}(\mathbf{p})=\tilde{\phi}(\mathbf{p})-\frac{\tilde{h}(\mathbf{p})}{\mathbf{p}^2+2\alpha_2}. \tag{637}$$

As the Jacobian of the change of variables is 1, this gives

$$Z[h]=\int\mathcal{D}\tilde{\varphi}\exp\left(-\int\frac{d^D\mathbf{p}}{(2\pi)^D}\left[\frac{\mathbf{p}^2}{2}+\alpha_2\right]\tilde{\varphi}(\mathbf{p})\tilde{\varphi}(-\mathbf{p})\right)\exp\left(\frac{1}{2}\int\frac{d^D\mathbf{p}}{(2\pi)^D}\frac{\tilde{h}(\mathbf{p})\tilde{h}(-\mathbf{p})}{\mathbf{p}^2+2\alpha_2}\right). \tag{638}$$

The first factor is merely a constant, which we will denote by $\mathcal{N}$, while the second factor is rewritten as

$$\begin{aligned} Z[h] &= \mathcal{N}\exp\left(\frac{1}{2}\int d^D\mathbf{r}d^D\mathbf{r}'h(\mathbf{r})G(\mathbf{r}-\mathbf{r}')h(\mathbf{r}')\right)\ , \\ G(\mathbf{r}) &= \int\frac{d^D\mathbf{p}}{(2\pi)^D}\frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{\mathbf{p}^2+2\alpha_2}. \end{aligned} \tag{639}$$

Taking functional derivatives we have

$$\langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta = G(\mathbf{r})\sim\frac{e^{-|\mathbf{r}|\sqrt{2\alpha_2}}}{|\mathbf{r}|^{D-2}}\ ,\qquad |\mathbf{r}|\to\infty. \tag{640}$$

This gives the correlation length

$$\xi=\frac{1}{\sqrt{2\alpha_2}}\sim\frac{1}{t^{\frac{1}{2}}}, \tag{641}$$

and thus the critical exponent

$$\boxed{\nu=\frac{1}{2}.} \tag{642}$$

Given the explicit calculation we have just done, we are now in the position to introduce a shortcut for obtaining the two point function in the Gaussian approximation in similar situations. In absence of a source $h(\mathbf{r})$ the Landau-Ginzburg Hamiltonian (and the Landau free energy βF for that matter) is written as

$$\beta\mathcal{H} \;=\; \frac{1}{2}\int \frac{d^D\mathbf{p}}{(2\pi)^D}\left[\left(\mathbf{p}^2+2\alpha_2\right)\tilde{\phi}(\mathbf{p})\tilde{\phi}(-\mathbf{p})+\text{quartic}\right]. \tag{643}$$

From this expression we can simply *read off* the result for the two point function in the Gaussian approximation

$$\langle\tilde{\phi}(\mathbf{p})\tilde{\phi}(\mathbf{q})\rangle_\beta = \frac{(2\pi)^D\delta^D(\mathbf{p}+\mathbf{q})}{\mathbf{p}^2+2\alpha_2}. \tag{644}$$

Here the denominator is given by whatever factor multiplies the quadratic term in $\beta\mathcal{H}$. The delta-function expresses momentum conservation.

In the ordered phase $t<0$ we expand $f[\phi(\mathbf{r})]$ around one of the minima at $\pm\phi_0$. The choice of minimum implements *spontaneous symmetry breaking*. We have

$$V(\phi)=\alpha_2\phi^2+\alpha_4\phi^4 \simeq \alpha_2\phi_0^2+\alpha_4\phi_0^4+(\alpha_2+6\alpha_4\phi_0^2)(\phi-\phi_0)^2+\ldots \tag{645}$$

We may drop the constant and retain only the contribution quadratic in $\delta\phi=\phi-\phi_0$ (this is the Gaussian approximation in the ordered phase), which gives

$$f[\delta\phi(\mathbf{r})]\simeq\frac{1}{2}|\nabla\delta\phi|^2+\tilde{\alpha}_2\delta\phi^2. \tag{646}$$

Here $\tilde{\alpha}_2=-2\alpha_2>0$. We may now copy the calculation in the disordered phase and obtain for the *connected correlation function*

$$\langle\delta\phi(\mathbf{r})\delta\phi(\mathbf{0})\rangle_\beta=\langle\phi(\mathbf{r})\phi(\mathbf{0})\rangle_\beta-\langle\phi(\mathbf{r})\rangle_\beta\langle\phi(\mathbf{0})\rangle_\beta\sim\frac{e^{-|\mathbf{r}|\sqrt{-4\alpha_2}}}{|\mathbf{r}|^{D-2}}\;,\quad |\mathbf{r}|\to\infty. \tag{647}$$

We see that the correlation length scales as $\xi\propto|t|^{-1/2}$, giving again $\nu=1/2$.

14.5 Homework Questions 19-21

Question 19. Consider a Landau expansion of the free energy of the form

$$F=\frac{a}{2}m^2+\frac{b}{4}m^4+\frac{c}{6}m^6$$

with $c>0$. Examine the phase diagram in the $a-b$ plane, and show that there is a line of critical transitions $a=0$, $b>0$ which joins a line of first order transitions $b=-4(ca/3)^{1/2}$ at a point $a=b=0$ known as a tricritical point.

Supposing that a varies linearly with temperature and that b is independent of temperature, compare the value of the exponent β at the tricritical point with its value on the critical line.

From Yeomans, *Statistical Mechanics of Phase Transitions*

Question 20.
(a) Discuss how an order parameter may be used to characterise symmetry breaking at a phase transition.

(b) Argue that the uniaxial ferromagnet-paramagnet transition can be described by a Landau free energy of the form

$$F = \int d^3\mathbf{r} \left[\frac{1}{2}|\nabla\phi(\mathbf{r})|^2 - h\phi(\mathbf{r}) + \alpha_2\phi^2(\mathbf{r}) + \alpha_3\phi^3(\mathbf{r}) + \alpha_4\phi^4(\mathbf{r}) \right]. \tag{648}$$

What can you say about α_4?
(c) What is the nature of the transition for $h = 0$ if $\alpha_3 \neq 0$? Explain your answer.
(d) Now assume that $\alpha_3 = h = 0$. Argue that close to the critical point

$$\alpha_2 = At \ , \quad t = \frac{T - T_c}{T_c} \text{ and } A > 0. \tag{649}$$

(e) Derive the equation characterizing the saddle point solution for $\alpha_3 = h = 0$. What are the configurations ϕ with the lowest free energy for $h = 0$, at $T > T_c$ and at $T < T_c$? Why are these $\mathbf{r}$ independent?
(f) Now consider more general solutions to the saddle point equation in the low-temperature phase. With suitable boundary conditions the saddle point solutions for the order parameter are functions of x only, i.e. $\phi = \phi(x)$. Show that in this case

$$E = \frac{1}{2}\left[\frac{d\phi(x)}{dx}\right]^2 - \alpha_2\phi^2 - \alpha_4\phi^4 \tag{650}$$

is independent of x. Construct a solution $\phi(x)$ such that

$$\lim_{x\to\infty} \phi(x) = \phi_1 \ , \quad \lim_{x\to-\infty} \phi(x) = \phi_2, \tag{651}$$

where $\phi_{1,2}$ are the solutions found in (d). Hint: determine E for such solutions first.

Question 21. A system with a real, two-component order parameter $(\phi_1(\mathbf{r}), \phi_2(\mathbf{r}))$ has a free energy

$$F = \int \mathrm{d}^d\mathbf{r} \left[\frac{1}{2}|\nabla\phi_1(\mathbf{r})|^2 + \frac{1}{2}|\nabla\phi_2(\mathbf{r})|^2 - \frac{1}{2}\left(\phi_1^2(\mathbf{r}) + \phi_2^2(\mathbf{r})\right) + \frac{1}{4}\left(\phi_1^2(\mathbf{r}) + \phi_2^2(\mathbf{r})\right)^2 \right] .$$

Find the order-parameter values Φ_1, Φ_2 that minimise this free energy. Consider small fluctuations around such state, with $(\phi_1(\mathbf{r}), \phi_2(\mathbf{r})) = (\Phi_1 + \varphi_1(\mathbf{r}), \Phi_2 + \varphi_2(\mathbf{r}))$ and expand F to second order in φ.

Assuming that the statistical weight of thermal fluctuations is proportional to $\exp(-F)$, calculate approximately the correlation function

$$\langle \varphi_1(\mathbf{r})\varphi_1(\mathbf{0}) + \varphi_2(\mathbf{r})\varphi_2(\mathbf{0}) \rangle$$

by evaluating a Gaussian functional integral. How does your result depend on the dimensionality d of the system?

15 Other Examples of Phase Transitions

15.1 Isotropic-Nematic Transition in Liquid Crystals

Liquid crystals are fluids of rod-like molecules. At high temperatures their centres of mass are randomly distributed and the rods are randomly oriented. At low temperatures, in the *nematic phase* the rods spontaneously align along a common axis, see Fig. 29. What is the order parameter characterizing this transition?

Let us associate a unit vector $\vec{n}(\mathbf{r})$ with a molecule at position $\mathbf{r}$. The first guess one may have is to take $\langle \vec{n}(\mathbf{r}) \rangle_\beta$ as the order parameter. This will not work, because the two vectors $\vec{n}(\mathbf{r})$ and $-\vec{n}(\mathbf{r})$ describe the same orientation of the molecule. Hence the order parameter must be invariant under

$$\vec{n}(\mathbf{r}) \longrightarrow -\vec{n}(\mathbf{r}). \tag{652}$$

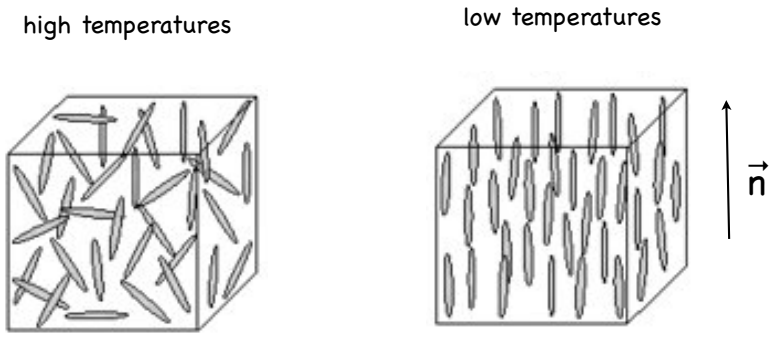


Figure 29: At low temperatures the rod-like molecules spontaneously align along an axis $\vec{n}$.

So how about something quadratic like

$$\langle n_i(\mathbf{r}) n_j(\mathbf{r}) \rangle_\beta. \tag{653}$$

The problem with this expression is that it is different from zero even for randomly oriented molecules (which is what one has at very high temperatures). Indeed, using a parametrization of the unit vector for a single molecule in terms of polar coordinates we have

$$\vec{n} = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}. \tag{654}$$

Then averaging over all possible orientations gives

$$\langle\langle n_i n_j \rangle\rangle = \frac{1}{4\pi}\int_0^\pi d\theta \int_0^{2\pi} d\varphi \sin\theta n_i n_j = \frac{1}{3}\delta_{i,j} \neq 0. \tag{655}$$

This consideration suggests to try

$$\boxed{Q_{ij} = \langle n_i n_j \rangle_\beta - \frac{1}{3}\delta_{i,j}} \tag{656}$$

as our order parameter. At very high temperatures, when molecules have random orientations, this is zero. On the other hand, if the molecules are aligned in the z-direction, i.e. $\vec{n} = \vec{e}_z$, we have

$$Q = \begin{pmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{2}{3} \end{pmatrix}. \tag{657}$$

So this seems to work. Given the order parameter, how do we then determine the free energy? In the high temperature phase the free energy must be invariant under rotations of the molecules, i.e. under

$$Q(\mathbf{r}) \longrightarrow RQ(\mathbf{r})R^T. \tag{658}$$

This suggests the following expansion for the free energy density

$$\boxed{f[Q(\mathbf{r})] = \frac{1}{2}|\nabla Q|^2 + \alpha_2 \mathrm{Tr}[Q^2] + \alpha_3 \mathrm{Tr}[Q^3] + \alpha_4 \left(\mathrm{Tr}[Q^2]\right)^2 + \dots} \tag{659}$$

Here $|\nabla Q|^2 = \sum_{i,j,k=1}^{3} (\partial_k Q_{ij})^2$. The presence of a cubic term suggests that the transition is first order, which is indeed correct.

15.2 Superfluid Transition in Weakly Interacting Bosons

Let us recall the second quantized Hamiltonian for weakly repulsive bosons

$$H = \int d^4\mathbf{r} \left[c^\dagger(\mathbf{r}) \left(-\frac{\nabla^2}{2m} \right) c(\mathbf{r}) + \frac{U}{2} c^\dagger(\mathbf{r}) c^\dagger(\mathbf{r}) c(\mathbf{r}) c(\mathbf{r}) \right]. \tag{660}$$

In the superfluid phase we have macroscopic occupation of the zero momentum single-particle state

$$\langle c(\mathbf{p}=0) \rangle_\beta = \sqrt{N_0}. \tag{661}$$

This implies that

$$\psi(\mathbf{r}) = \langle c(\mathbf{r}) \rangle_\beta \neq 0 \ , \quad T < T_c. \tag{662}$$

At high temperatures the expectation value vanishes, so that we may take $\psi(\mathbf{r})$ as a *complex valued order parameter*. The Landau free energy describing the transition can then be constructed by noting that it should be invariant under $U(1)$ transformations

$$\psi(\mathbf{r}) \longrightarrow e^{i\theta} \psi(\mathbf{r}). \tag{663}$$

This gives

$$\boxed{F[\psi] = \int d^D\mathbf{r} \left[\frac{1}{2m} |\nabla \psi|^2 + \alpha_2 |\psi(\mathbf{r})|^2 + \alpha_4 |\psi|^4 \right].} \tag{664}$$

16 Ising Model and Φ^4 Theory

The 3-dimensional Ising model has a paramagnet to ferromagnet phase transition that can be described by Landau theory. An obvious question is then whether it is possible to *derive* the appropriate Landau theory starting from the lattice model. Let us therefore consider a general Ising model with Hamiltonian

$$H = -\sum_{i,j} \sigma_i \widetilde{C}_{ij} \sigma_j - \tilde{h} \sum_i \sigma_i. \tag{665}$$

Let's work on a hypercubic lattice in D dimensions, and take

$$\widetilde{C}_{ij} = \widetilde{C}(|\mathbf{r}_i - \mathbf{r}_j|) \tag{666}$$

to be translationally invariant. The patition function is

$$Z = \sum_{\{\sigma_j\}} e^{\sum_{j,\ell} \sigma_j C_{j\ell} \sigma_\ell + h \sum_\ell \sigma_\ell}, \tag{667}$$

where we have defined $h = \beta \tilde{h}$ and $C_{ij} = \beta \widetilde{C}_{ij}$. Recalling the Gaussian integration identity

$$1 = \underbrace{\sqrt{\det(4\pi C^{-1})}}_{\mathcal{N}} \int \underbrace{\prod_\ell d\psi'_\ell}_{\mathcal{D}\psi'} \ e^{-\frac{1}{4}\sum_{j,k} \psi'_j C^{-1}_{jk} \psi'_k}, \tag{668}$$

and then shifting the integration variables

$$\psi_j' = \psi_j - 2\sum_k C_{jk}\sigma_k, \tag{669}$$

we arrive at the following identity

$$1 = \mathcal{N}\int \mathcal{D}\psi\; e^{-\frac{1}{4}\sum_{j,k}\psi_j C^{-1}_{jk}\psi_k + \sum_j \sigma_j\psi_j - \sum_{j,k}\sigma_j C_{jk}\sigma_k}. \tag{670}$$

Substituting this into our expression for the partition function gives

$$\boxed{Z = \mathcal{N}\int \mathcal{D}\psi \sum_{\{\sigma_j\}} e^{-\frac{1}{4}\sum_{j,k}\psi_j C^{-1}_{jk}\psi_k + \sum_j \sigma_j(\psi_j + h)}.} \tag{671}$$

What we have done is to remove the interaction between the spins by introducing auxiliary variables ψ_j at every site. This is an example of a *Hubbard-Stratonovich transformation*. Now we can perform the summations over spin variables

$$\sum_{\{\sigma_j\}} e^{\sum_k \sigma_k(\psi_k + h)} = \prod_k \left[2\cosh(\psi_k + h)\right]. \tag{672}$$

Thus

$$Z = \mathcal{N}\int \mathcal{D}\psi\; e^{-\frac{1}{4}\sum_{j,k}\psi_j C^{-1}_{jk}\psi_k + \sum_j \ln(2\cosh(h+\psi_j))} \tag{673}$$

Finally we change variables to

$$\phi_j = \frac{1}{2}\sum_k C^{-1}_{jk}(\psi_k + h), \tag{674}$$

which leads to the following expression for the partition function

$$Z = \mathcal{N}'\int \mathcal{D}\phi\; e^{-\sum_{j,k}\phi_j C_{jk}\phi_k + h\sum_j \phi_j + \sum_j \ln(\cosh(2\sum_k C_{jk}\phi_k))}. \tag{675}$$

This is still at the level of the lattice model. We now want to focus on large distances/small energies, in order to obtain a field theory description. We furthermore will now assume that the matrix C, which depends on temperature, is such that large fluctuations of ϕ_j are supressed. We then can take ϕ_j to be small and expand

$$\ln[\cosh(\underbrace{2\sum_k C_{jk}\phi_k}_{X})] \simeq \frac{X^2}{2} - \frac{X^4}{12} + \dots. \tag{676}$$

Finally, we go to Fourier space using

$$\phi_j = \frac{1}{N}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}}\phi(\mathbf{k})\;, \quad C_{jk} = \frac{1}{N}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_k)}\; C(\mathbf{k}), \tag{677}$$

and focus on the small-$|\mathbf{k}|$ regime. This gives

$$Z = \mathcal{N}''\int \mathcal{D}\phi\; e^{-\frac{1}{N}\sum_{\mathbf{k}}\phi(\mathbf{k})\left[c_1 + c_2\mathbf{k}^2 + \dots\right]\phi(-\mathbf{k}) + c_3 h\phi(\mathbf{0}) + \frac{c_4}{N^3}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3}\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\phi(-\mathbf{k}_1-\mathbf{k}_2-\mathbf{k}_3) + \dots}, \tag{678}$$

where the coefficients c_l depend on the matrix $\widetilde{C}_{jk}$ specifying the interactions of Ising spins. Going back to real space we recover Euclidean Φ^4 theory

$$Z = \mathcal{N}''\int \mathcal{D}\phi\; e^{-\int d^D\mathbf{x}\left[c_1(\nabla\phi(\mathbf{x}))^2 + c_2\phi^2(\mathbf{x}) + c_3 h\phi(\mathbf{x}) + c_4\phi^4(\mathbf{x})\right]}. \tag{679}$$

Voilà. Note however, that the relation of the field $\phi(\mathbf{x})$ to the original Ising spin is rather indirect. In order to get a better feeling of "what happens to the spin" under all the above transformations, we could repeat the above steps for the generating functional rather than the partiton function. This will show us that Φ is indeed a good order parameter for the Ising transition.

17 Transverse-Field Ising Model

We now use the methods of many-particle quantum mechanics to study aspects of the transverse-field Ising model (TFIM). As a reminder this model was obtained by interpreting the transfer matrix of the 2D classical Ising model as the transition amplitude for a quantum 1-dimensional Hamiltonian (see (306)). The TFIM is given by the following Hamiltonian for spin-1/2 particles arranged on a 1D lattice: The degrees of freedom of this model are quantum spins-$\frac{1}{2}$ on a lattice, with dynamics described by the following quantum Hamiltonian,

$$H_{\text{TFIM}} = -J\left[\sum_{i=1}^{L} \boldsymbol{\sigma}_i^z \boldsymbol{\sigma}_{i+1}^z + g\sum_i \boldsymbol{\sigma}_i^x\right]. \tag{680}$$

where $\boldsymbol{\sigma}_i^\mu$ are Pauli matrices associated to sites i (we have gone back to using $\boldsymbol{\sigma}$ instead of X, Y, Z; also, to unclutter notation, we have dropped the bar on J introduced in (680) and also the hats on operators.) Pauli matrices on sites $i \neq j$ commute. We may or may not choose to impose periodic boundary conditions (in which case we take $Z_{L+1} \equiv Z_1$, etc. Also, we we will often just write $\sum_i$ with the understanding that the limits are as abov. In this section we will discuss various properties of this model.

Model (680) describes a set of spins-$\frac{1}{2}$ on a line, in the presence of an external transverse magnetic field (we will explain the meaning of 'transverse' in a moment.) The 'spins' can arise in a variety of different contexts, and it's best to think of these as 'generic' two-level systems — for instance, in a condensed-matter context these could arise due to some conspiracy of interactions of electrons with ions and each other, that forces the electrons to remain tied to individual lattice sites, and interact only through their spin degrees of freedom. Another example is provided by spinless fermions or hard-core[5] bosons but now free to hop around; the presence or absence of a particle on a site is our 'Ising' variable. For our purposes, the origin of the spins is not terribly important; we will just work with the simple model that we will introduce in the next section. In this section[6], we will study the transverse field Ising model in one spatial dimension, often also called the quantum Ising chain.

Note that J simply sets the overall energy scale and hence plays a rather minor role; the physics is controlled by the dimensionless parameter g. The reason we refer to the second term as a *transverse* field is that it 'points' in a direction different from that favored by the first term, and as such *competes* with the first term. The essentially quantum mechanical nature of the model stems from the fact that the two terms don't commute: as such, the uncertainty principle guarantees for us that if we are in a limit where one term is 'pinned' to a definite value, we have infinite quantum-mechanical uncertainty in the other term. However, one might ask why the first term takes such a complex form; after all, a simpler Hamiltonian, of the form $H_{\text{boring}} = -J\left[\sum_i \boldsymbol{\sigma}_i^z + g\sum_i \boldsymbol{\sigma}_i^x\right]$ *also* has such competition. The difference lies in the evolution of the eigenstates as a function of g: it is easy to see that in H_{boring}, we may simply rotate our spin coordinate system about the y-axis by an angle $\theta = \tan^{-1} g$, so that in the new coordinate system, $H_{\text{boring}} \to -J\sqrt{1+g^2}\sum_i \boldsymbol{\sigma}_i^z$, and there is no competition — thus, as we change g, we do not expect to see anything but smooth behavior of various quantities computed from H_{boring}. In contrast, we cannot 'rotate away' the competition in H_{TFIM} in this way, and (as we will see) the evolution of the physics as a function of g must become singular somewhere between $g \ll 1$ and $g \gg 1$ (and in fact, occurs at precisely $g = 1$.) The key distinction between H_{boring} and H_{TFIM} is one of *symmetry*, that we now explain.

17.1 Symmetries

H_{TFIM} possesses an *Ising symmetry* (also called a $\mathbb{Z}_2$ symmetry) of taking up spins to down spins, implemented via the mapping

$$U : (\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z) \to (\boldsymbol{\sigma}_x, -\boldsymbol{\sigma}_y, -\boldsymbol{\sigma}_z) \tag{681}$$

Clearly, doing this twice is equivalent to the identity — hence the symmetry group is $\mathbb{Z}_2$, and its elements are $(1, U)$ where $U^2 = 1$. Note that the transformation (681) is equivalent to a π rotation of all the spins

[5] The term 'hard-core' here simply means no more than a single boson can sit at a site.

[6] The treatment here closely parallels John McGreevy's excellent notes, but will differ on some points.

about the x-axis, generated by the operator

$$U = e^{i\pi \sum_j \frac{1}{2}\boldsymbol{\sigma}_j^x} = \prod_j e^{i\frac{\pi}{2}\boldsymbol{\sigma}_j^x} = \prod_j \left(\cos\frac{\pi}{2} + i\boldsymbol{\sigma}_j^x \sin\frac{\pi}{2}\right) = i^N \prod_j \boldsymbol{\sigma}_j^x \sim \prod_j \boldsymbol{\sigma}_j^x \tag{682}$$

where we have used elementary results from the quantum theory of spin, as well as the fact that spin operators on different sites commute, and finally ignored an overall constant prefactor of i^N. Clearly, $U^\dagger H_{\rm TFIM} U = H_{\rm TFIM}$, hence U is a symmetry. We now analyze the manner in which this symmetry is realized by the ground state of $H_{\rm TFIM}$ in the limits $g \gg 1$ and $g \ll 1$.

17.2 Phase Diagram

17.2.1 Disordered Phase

Consider $H_{\rm TFIM}$ in the limit $g \gg 1$. As a first step in this limit, we may ignore the first 'ferromagnetic' term, and just satisfy the second 'transverse field' term. This gives a unique ground state, where all the spins point along the x direction, that we represent as $|0\rangle = |\rightarrow\rightarrow\rightarrow\ldots\rightarrow\rangle$. Clearly, this ground state preserves the symmetry: $U|0\rangle = \prod \boldsymbol{\sigma}_i^x |\rightarrow\rightarrow\rightarrow\ldots\rightarrow\rangle = (1)^N |\rightarrow\rightarrow\rightarrow\ldots\rightarrow\rangle = |0\rangle$. The first excited states in this limit are obtained by reversing a single spin at point i:

$$|1\rangle = |\leftarrow\rightarrow\rightarrow\ldots\rightarrow\rangle, |2\rangle = |\rightarrow\leftarrow\rightarrow\ldots\rightarrow\rangle, \ldots |i\rangle = |1\rangle = |\rightarrow\rightarrow\ldots\leftarrow_i\ldots\rightarrow\rangle, \text{etc.} \tag{683}$$

There are clearly N such 'spin flip' states in an N-site system, all of which are degenerate in the absence of the ferromagnetic term, and have energy $\Delta = 2gJ$ above the ground state (whose energy we label E_0.) Adding the ferromagnetic term as a perturbation gives the spin flips a dispersion: let us for the moment only work in the subspace of a single spin flip. It is easy to see that the term $\boldsymbol{\sigma}_i^z \boldsymbol{\sigma}_{i+1}^z$ just moves a spin flip from site i to site $i+1$ (or vice-versa). Thus, to leading order in perturbation theory in the ferromagnetic term in this sector, we may write the (approximate) eigenvalue equation in terms of 'hopping' of spin flips:

$$H|i\rangle = -J\left[|i+1\rangle + |i-1\rangle\right] + (E_0 + 2gJ)|i\rangle. \tag{684}$$

It is easiest to consider the system to live on a circle with N sites, so that when we may identify $0 \equiv N$, and thus interpret the hopping terms in this manner. Then, we may readily solve by means of a Fourier transform: let us write

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_j e^{ikr_j} |j\rangle \tag{685}$$

where the locations of the spins are at sites $r_j = ja$. Periodic boundary conditions require $k = 2\pi\frac{m}{Na}$ with m an integer, for $|k\rangle$ to be single-valued; also, only $m = 1, 2, \ldots, N$ are distinct (which is good, since we can't have more than N states arising from any linear combination of spin flips!). In the thermodynamic limit of $N \to \infty$, k becomes a continuous variable in the first Brillouin zone, *i.e.* the segment $k \in [-\pi/a, \pi/a]$.

Inserting this into our effective hopping problem for spin flips, we find

$$(H - E_0)|k\rangle = (2gJ - 2J\cos k)|k\rangle \tag{686}$$

In other words, we find the gap to the first excited state at k follows the pattern $\epsilon(k) = 2J[g - \cos k]$. A spin flip thus behaves like a *particle* excitation: for a given momentum, it has a fixed energy $\epsilon(k)$. Hence, the lowest-lying excitations above the symmetric $g \gg 1$ ground state are gapped particles, with an energy gap $\Delta = 2J(g-1)$ and a dispersion $\epsilon(k) \sim \Delta + Jk^2$ for small momenta. $\boldsymbol{\sigma}_i^x$ behaves like a particle 'number operator' as it measures the number of spin flips, while $\boldsymbol{\sigma}_i^z$ inserts a single spin flip at site i and is associated with creating a particle. Note that creating two particles at a single site is equivalent to doing nothing — suggesting that the particle is its own antiparticle, a point we will return to below.

In the above discussion, we assumed $g \gg 1$. Suppose we simply extend the calculation well beyond its regime of validity, to $g \sim 1$; we anticipate a transition at $g = 1$, where the gap to excitations closes, and the spin-flips 'condense'. Serendipitously, it turns out that this is the exact value at which the ground state changes its fundamental nature — a fact which is generally never true (one should not be able to predict critical parameters from such a theory in most cases) but lies on the fact that the Ising model in $d = 1$ is *self-dual*. Before discussing duality, we will first try to understand the ordered phase

17.2.2 Ordered Phase

We now turn to the opposite limit of $H_{\rm TFIM}$, namely the 'weak coupling' limit when $g \ll 1$. Let us begin as before, by first setting $g = 0$ and then perturbatively considering what hapens at finite but small g. At $g = 0$. we retain only the ferromagnetic term, which is clearly minimized by either of the *two* states $|+\rangle = |\uparrow\uparrow \dots \uparrow\rangle$ or $|-\rangle = |\downarrow\downarrow \dots \downarrow\rangle$. Clearly, these ground states break symmetry: $U|+\rangle = |-\rangle$, since the action of $\boldsymbol{\sigma}^x$ is to flip a spin pointing up along the z-axis to one pointing down. Note however that linear combinations of the two states of the form $|+\rangle \pm |-\rangle$ actually do preserve the symmetry; for any non-zero g, there is a non-zero matrix elements to mix these states, since the transverse field terms can flip spins. Let us now estimate this matrix element in perturbation theory. The perturbing term here is the transverse field; this can flip the spin on a *single* site. Thus, in order to obtain a non-vanishing matrix element between the state $|+\rangle$, with all spins up, and the state $|-\rangle$ with all spins down, we need to go to $N^{\rm th}$ order in perturbation theory, where N is the size of the system. The resulting effective Hamiltonian takes the form (here $s, s' = +, -$),

$$\langle s|H_{\rm eff}|s'\rangle = \begin{pmatrix} E_0 & g^N \\ g^N & E_0 \end{pmatrix} \tag{687}$$

from which we see that the eigenstates are $|\psi_\pm\rangle \propto |+\rangle \pm |-\rangle$, with a small splitting $\delta = |E_+ - E_-| \sim O(g^N) = O(e^{-N \log \frac{1}{g}})$ (remember, $g \ll 1$) — in other words, the splitting between the states vanishes exponentially as $N \to \infty$. We will provide a more physical description of this in a moment, but let us assume that this splitting does indeed vanish; then, we can see how spontaneous symmetry breaking appears. While at any finite system size the ground states are indeed of the 'Schrödinger cat' form $|+\rangle \pm |-\rangle$ and hence actually respect the symmetry, if one initially *prepares* the system in (say) all up, it takes an extremely long time (exponential in the size of the system) to evolve out of this state to demonstrate that it is not an eigenstate. So for all practical purposes, we may consider $|+\rangle$ and $|-\rangle$ to be eigenstates, and can improve the approximation by going to larger and larger systems while keeping other parameters fixed. (Note that we could also impose a symmetry-breaking longitudinal field, take the thermodynamic limit, and then take the field to zero, just as we did when diagnosing SSB in a classical system.) We will see later that this degeneracy is mirrored by a certain topological robustness in a different description of the TFIM in terms of fermions.

To understand the excited-state spectrum in the $g \ll 1$ limit, let us consider the state $|+\rangle = |\uparrow\uparrow \dots \uparrow\rangle$. What is the lowest-energy excitation here? For simplicity, w econsider an open chain, and we might naively answer that this is again a spin flip, where at a single site we send $\uparrow \to \downarrow$. This costs energy $\Delta = 4J$. However, we can do better. Consider a *domain wall* in the spin configuration, where between site $i, i+1$ the spin goes from $\uparrow$ to $\downarrow$: $|\overline{\dots \uparrow\uparrow \dots \uparrow_i}|\underline{\downarrow_{i+1}\downarrow \dots}\rangle$. This has a lower energy than a spin flip: $\Delta = 2J$. With periodic boundary conditions, we have to make a pair of domain walls (can you see why?), but these are in fact independent excitations — it turns out that single spin flip in the Ising model is *not* a 'particle' in that it does not in general have a sharp energy-momentum relation. Indeed, the domain walls are the true 'particle' excitations, as we will show in a moment; so a spin flip of a given momentum will decay by partitioning its energy and momentum into a pair of domains walls, *i.e.* it merges with the 'two-particle continuum'. The domain walls are however non-local objects: one needs to change the state of a macroscopic number of spins to create a single domain wall far from any others.

Let us analyze the spectrum of the single domain wall sector. To avoid the complications of disentangling the behavior of a single domain wall from the two-domain-wall sector that is the minimal accessible one with periodic boundary conditions, we work directly in the $N \to \infty$ limit. It is most convenient to think of the domain walls as living on links of the lattice, and label a state with a domain wall between sites i, $i+1$ as $|\bar{i} = i + \frac{1}{2}\rangle$. As before, in the absence of the perturbation (in this case, the $g \sum_i \boldsymbol{\sigma}_i^x$ term) there are N degenerate, localized domain wall states[7]. These states acquire a dispersion for $g \neq 0$: within the single-domain-wall sector, all that the $g \sum_i \boldsymbol{\sigma}_i^x$ term is 'hop' the domain wall one spacing to the left or right.

[7] We are glossing over the fact of boundary conditions in making this kind of counting argument; it is in fact possible to consider single domain walls on a ring by also allowing boundary conditions that are antiperiodic, rather than periodic.

Hence, we arrive at a very similar problem in the one-domain-wall sector as we did in the single-spin-flip sector:

$$(H - E_0)|\bar{i}\rangle = -gJ\left[|\bar{i}+1\rangle + |\bar{i}-1\rangle\right] + 2J|\bar{i}\rangle \tag{688}$$

Fourier transforming again (though now observe that the spatial degrees of freedom lie on the mid-points of the links, rather than on lattice sites) we obtain the dispersion $\epsilon(k) = 2J[1 - g\cos k]$ for the domain walls. Once again, we see that the $\epsilon(k)$ becomes negative at $k = 0$ when $g = 1$, and we may expect a transition, now driven by the condensation of domain walls.

This argument also allows us to see why the broken-symmetry state is stable: one needs to create a pair of domain walls and make them move around the system 'unzipping' up spins to down spins, and then annihilate when they span the system. However, these are gapped excitations: it first costs energy $4J$ to nucleate the domain wall pair, and then each step that the domain wall moves costs an energy of order gJ. Computing the action for this process, we find $S \sim e^{-N \log \frac{1}{g}}$ as we computed earlier, which vanishes in the thermodynamic limit $N \to \infty$.

Note that in the $g \to 0$ limit, the system has *long range order* in the z-basis, since all the spins point in the same direction. We may also think of the ordered phase as a condensate of spin flips out of the paramagnetic (disordered) phase that forms at large g: consider the operator $\boldsymbol{\sigma}_i^z$, that creates a single spin flip; we have for its expectation value at $g = 0$

$$\langle \boldsymbol{\sigma}_j^z \rangle = \langle \pm | \boldsymbol{\sigma}_j^z | \pm \rangle = \pm 1 \text{ for any } j \tag{689}$$

In the opposite limit of $g \to \infty$, we see that $\langle \boldsymbol{\sigma}_j^z \rangle = 0$, and the value does not deviate from this until we start condensing spin flips, which occurs in the vicinity of $g = 1$. We can thus think of $\langle \boldsymbol{\sigma}_j^z \rangle$ as an *order parameter* for the Ising transition, as it is nonzero only in the broken-symmetry phase.

17.3 Duality

We have thus far studied the TFIM in two limits: the disordered or 'paramagnetic' phase for $g \gg 1$, and the Ising-ordered broken symmetry ferromagnet for $g \ll 1$. We found that in both limits there is a gap to excitations, with low-lying particle-like excitations: spin flips and domain walls, in the two limits. We found that the dispersion of these excitations have a simple form:

$$\begin{aligned} \epsilon_{\text{spin flip}}(k) &= 2J[g - \cos k] \\ \epsilon_{\text{domain wall}}(k) &= 2J[1 - g\cos k] \end{aligned} \tag{690}$$

There is a suggestive symmetry between these expressions, that seems to require

$$g \to \frac{1}{g}, J \to Jg. \tag{691}$$

It turns out there is a rigorous duality between strong ($g \gg 1$) and weak ($g \ll 1$) coupling in the quantum Ising model, that maps spin flips to domain walls and vice versa.

First, assume open boundary conditions, so that the system just stops at $j = 1$ and $j = N$, so we can avoid the even-domain-wall constraint. Any state in the $\boldsymbol{\sigma}^z$ basis can be specified by the direction ($s^z = \uparrow, \downarrow$) of the first spin along the z axis, and the locations of the domain walls. Consider the operator[8], diagonal in this basis, that measures whether there is a domain wall at $\bar{j} = j + \frac{1}{2}$, *i.e.* between sites $j, j+1$:

$$\boldsymbol{\tau}_{\bar{j}}^x \equiv \boldsymbol{\sigma}_{\bar{j}-\frac{1}{2}}^z \boldsymbol{\sigma}_{\bar{j}+\frac{1}{2}}^z = \begin{cases} +1 & \text{if } s_{\bar{j}-\frac{1}{2}}^z = s_{\bar{j}+\frac{1}{2}}^z \\ -1 & \text{if } s_{\bar{j}-\frac{1}{2}}^z = -s_{\bar{j}+\frac{1}{2}}^z \end{cases} = (-1)^{\text{is there a domain wall at } \bar{j}?}. \tag{692}$$

[8] When discussing lattice dualities, we must be careful to pay attention to site variables, since dual variables live on links and 'direct' variables live on sites.

Note that $\boldsymbol{\tau}_{\bar{j}}^2 = 1$, and $\boldsymbol{\tau}_{\bar{j}}^\dagger = \boldsymbol{\tau}_{\bar{j}}$. Now, consider the operator that inserts a domain wall at $\bar{j}$: we implement this by flipping all spins to the right of $\bar{j}$, *i.e.*

$$\boldsymbol{\tau}_{\bar{j}}^z \equiv \boldsymbol{\sigma}_{\bar{j}+\frac{1}{2}}^x \boldsymbol{\sigma}_{\bar{j}+\frac{3}{2}}^x \cdots = \prod_{j>\bar{j}} \boldsymbol{\sigma}_j^x. \tag{693}$$

This operator is also Hermitian, and squares to the identity. Now, we observe the following neat identity:

$$\boldsymbol{\tau}_{\bar{j}}^z \boldsymbol{\tau}_{\bar{j}'}^x = (-1)^{\delta_{\bar{j}\bar{j}'}} \boldsymbol{\tau}_{\bar{j}'}^x \boldsymbol{\tau}_{\bar{j}}^z \tag{694}$$

— just like the original Pauli matrices, $\boldsymbol{\sigma}_j^z$, $\boldsymbol{\sigma}_{j'}^x$! (To see this, note that if $\bar{j}' < \bar{j}$, $\boldsymbol{\tau}_{\bar{j}}^z$ and $\boldsymbol{\tau}_{\bar{j}'}^x$ don't share any sites and hence commute. If $\bar{j}' = \bar{j}$, then $\boldsymbol{\tau}_{\bar{j}}^z$ and $\boldsymbol{\tau}_{\bar{j}'}^x$ share exactly one site, namely $\bar{j}+\frac{1}{2}$; as $\boldsymbol{\tau}_{\bar{j}}^z$ contains $\boldsymbol{\sigma}_{\bar{j}+\frac{1}{2}}^x$ and $\boldsymbol{\tau}_{\bar{j}}^x$ contains $\boldsymbol{\sigma}_{\bar{j}+\frac{1}{2}}^z$, commuting these past each other 'inherits' a factor of (-1) from the anticommutation of two Pauli matrices. Finally, if $\bar{j}' > \bar{j}$, then $\boldsymbol{\tau}_{\bar{j}}^z$ and $\boldsymbol{\tau}_{\bar{j}'}^x$ have two sites in common, and the overall sign after commuting the relevant Pauli matrices $\boldsymbol{\sigma}_i^\mu$ is $(-1)^2$ and thus cancels out.) This anticommutation relation tells us that domain walls can be represented in terms of two-state systems living on the links, suggesting that their character is quite similar to the original Ising spins.

Now, the inverse map from $\boldsymbol{\sigma}^x, \boldsymbol{\sigma}^z$ to $\boldsymbol{\tau}^x, \boldsymbol{\tau}^z$ is simply given by

$$\boldsymbol{\sigma}_j^x = \boldsymbol{\tau}_{j-\frac{1}{2}}^z \boldsymbol{\tau}_{j+\frac{1}{2}}^z \tag{695}$$

reflecting the fact that a single spin flip is equivalent to a pair of adjacent domain walls.

So, in the bulk, we have

$$H_{\text{TFIM}} = -J \sum_j \left(\boldsymbol{\sigma}_j^z \boldsymbol{\sigma}_{j+1}^z + g \boldsymbol{\sigma}_j^x \right) = -J \sum_{\bar{j}} \left(\boldsymbol{\tau}_{\bar{j}}^x + g \boldsymbol{\tau}_{\bar{j}}^z \boldsymbol{\tau}_{\bar{j}+1}^z \right) \tag{696}$$

Note that the second expression is a new TFIM, with couplings mapped via (691). Thus, we have 'dualized': the new degrees of freedom are the domain walls in the original Ising spins. It turns out that there is a similar mapping in the $D = 2$ classical Ising model due to Kramers and Wannier; due to the quantum to classical mapping, we will see that this is the same duality as discussed here. One important thing to note is the following: suppose the model has a single phase transition. Since the transition occurs as we move from strong to weak coupling, and (691) interchanges the role of these, if there is a single phase transition, then it *must* occur at the so-called 'self-dual point' $g = 1$.

Note that the paramagnetic (disordered) phase of the original variables is a *condensate* of domain walls in the following sense: the operator that creates a domain wall has an expectation value

$$\langle \boldsymbol{\tau}_{\bar{j}}^z \rangle = \langle \text{gs}_{g=\infty} | \boldsymbol{\tau}_{\bar{j}}^z | \text{gs}_{g=\infty} \rangle = \langle \text{gs}_{g=\infty} | \prod_{j>\bar{j}} \boldsymbol{\sigma}_{\bar{j}}^z | \text{gs}_{g=\infty} \rangle = 1 \text{ for any } \bar{j}. \tag{697}$$

For $1 < g < \infty$, this expectation value is between 0 and 1 — mirroring the behavior of $\boldsymbol{\sigma}_j^z$ but in reverse. While there is a condensate, there is no order: a nonzero expectation value of $\boldsymbol{\sigma}^x$ does not break any symmetry of H_{TFIM}. (Note that there is another state with the opposite sign, one with all spins pointing 'left', but for large g it's a very high-energy state.) The domain wall creation operator is thus often referred to as a *disorder* parameter, or a disorder operator.

An important point to note (and why you should always treat the word 'duality' with some suspicion!) is that the two sides of the phase diagram are *not* the same despite our mapping. On one side, there are two ground states related by the broken symmetry, on the other there is a unique symmetric ground state – indeed, it is this fact that requires the existence of an intervening phase transition!

It seems that since the duality transformation exchanges ordered and disordered phases, it also exchanges the ground state degeneracies. The disordered phase of the dual theory corresponds to the broken-symmetry phase of the Ising variables, and as such should have a twofold ground state degeneracy; similarly, the ordered

phase of the dual theory naively has two distinct ground states, but as it maps to the Ising paramagnet, there is again an apparent mismatch. To see that the dual theory actually does represent the phase diagram and the ground state degeneracies correctly, we need to be more precise than we have been about what happens at the boundaries. Consider N sites in a row, where the $j = 1$ and $j = N$ sites have only one neighbor. We have

$$H_{\text{TFIM}} = -J \sum_{j=}^{N-1} \left[\left(\boldsymbol{\sigma}_j^z \boldsymbol{\sigma}_{j+1}^z + g\boldsymbol{\sigma}_j^x \right) + g\boldsymbol{\sigma}_N^x \right] \tag{698}$$

Let's now perform the duality map: we have (not using our 'barred' link notation for clarity)

$$\begin{aligned} \boldsymbol{\sigma}_j^z \boldsymbol{\sigma}_{j+1}^z &= \boldsymbol{\tau}_{j+\frac{1}{2}}^x, \quad j = 1, 2, \ldots, N-1 \\ \boldsymbol{\sigma}_j^x &= \boldsymbol{\tau}_{j-\frac{1}{2}}^z \boldsymbol{\tau}_{j+\frac{1}{2}}^z, \quad j = 1, 2, \ldots, N \end{aligned} \tag{699}$$

so that in terms of domain-wall variables, we have

$$H_{\text{TFIM}} = -Jg \left[\sum_{j=1}^{N-1} \left(\boldsymbol{\tau}_{j-\frac{1}{2}}^z \boldsymbol{\tau}_{j+\frac{1}{2}}^z + \frac{1}{g} \boldsymbol{\tau}_{j+\frac{1}{2}}^x \right) + \boldsymbol{\tau}_{N-\frac{1}{2}}^z \boldsymbol{\tau}_{N+\frac{1}{2}}^z \right] \tag{700}$$

We have two special cases:

- $\boldsymbol{\tau}_{N+\frac{1}{2}}^z = 1$: this flips all spins with $j > N$; as there are no such spins, it must be the identity operator.
- $\boldsymbol{\tau}_{\frac{1}{2}}^x = \boldsymbol{\sigma}_0^z \boldsymbol{\sigma}_1^z$: this measures whether or not there is a domain wall between $j = 1$ and $j = 0$; but there is no spin at $j = 0$. One way to resolve this is to put a fiducial 'ghost spin' at $j = 0$ that always points up. Then, $\boldsymbol{\tau}_{\frac{1}{2}}^x = \boldsymbol{\sigma}_1^z$: it measures the value of the 'reference spin' at the origin, with respect to which we define the domain wall.

At $g = 0$, we have $H_{g=0} = -J \sum_{j=1}^{N-1} \boldsymbol{\tau}_{j+\frac{1}{2}}^x$, and the ground state has $\boldsymbol{\tau}_{j+\frac{1}{2}}^x$ for $j = 1, 2 \ldots, N-1$. Missing in this list is $\boldsymbol{\tau}_{\frac{1}{2}}^x$, so there are two degenerate ground states, eigenstates of $\boldsymbol{\tau}_{\frac{1}{2}}^x$ with eigenvalues ± 1 — these are just the 'no domain wall' states $|\pm\rangle$: since we have $\boldsymbol{\tau}_{\frac{1}{2}}^x$, all the spins to the right agree with the first one.

At $g = \infty$, we have $H_{g=\infty} = -Jg \sum_{j=1}^{N-1} \boldsymbol{\tau}_{j-\frac{1}{2}}^z \boldsymbol{\tau}_{j+\frac{1}{2}}^z + \boldsymbol{\tau}_{N-\frac{1}{2}}^z$ (where we have set $\boldsymbol{\tau}_{N+\frac{1}{2}}^z = 1$). The first term requires that $\boldsymbol{\tau}_{j-\frac{1}{2}}^z = \boldsymbol{\tau}_{j+\frac{1}{2}}^z$ for $j = 1, 2, \ldots, N-1$ but the second term forces $\boldsymbol{\tau}_{N-\frac{1}{2}}^z = 1$, thus fixing the sign. So, we obtain the unique paramagnetic state.

In other words, in the dual language, the freedom to choose $\boldsymbol{\tau}_{\frac{1}{2}}^x$ maps into the twofold degeneracy of the ordered phase of the original ising spins, and the fixed value of $\boldsymbol{\tau}_{N+\frac{1}{2}}^z = 1$ ensures that the paramagnetic phase of the original model has a unique ground state.

12: Aside: Duality with Periodic Boundary Conditions

How do we carry out the duality transformation with periodic boundary conditions? The specific problem is to understand how to define the creation operator for a single domain wall with periodic boundary conditions on the spins. It turns out this is impossible, since the fact we can only make an even number means that the *sign* of the operator is ill-defined — a single domain wall operator cannot be measured (we will see this also in a different sense, in that a single domain wall is a Majorana fermion that does not have a sense in which we can think of it as being occupied or empty.)

In spite of the sign ambiguity, we can 'fix' this by fiat as follows: and define the domain wall operators on a ring as follows:

$$\begin{aligned} \boldsymbol{\tau}^x_{\bar{j}} &= \boldsymbol{\sigma}^z_{\bar{j}-\frac{1}{2}}\boldsymbol{\sigma}^z_{\bar{j}+\frac{1}{2}} \\ \boldsymbol{\tau}^z_{\bar{j}} &= \prod_{N\geq j>\bar{j}} \boldsymbol{\sigma}^x_j \end{aligned} \tag{701}$$

In the second line, observe that we have made a choice: even though $\boldsymbol{\sigma}^z_{N+1} = \boldsymbol{\sigma}^z_1$ (as per periodic boundary conditions), we've decided to 'mark' the link $\bar{j} = N + \frac{1}{2}$ to stop the strong of $\boldsymbol{\sigma}^x$s. This is arbitrary; we could have marked any other link, the important thing is that we have some prescription for terminating the string. As in the case of open boundary conditions, this means that $\boldsymbol{\tau}^z_{\frac{1}{2}}$ and $\boldsymbol{\tau}^z_{N+\frac{1}{2}}$ are special: we have

$$\boldsymbol{\tau}^z_{N+\frac{1}{2}} = 1 \tag{702}$$

since the product is over the empty set, while

$$\boldsymbol{\tau}^z_{\frac{1}{2}} = \prod_{j=1}^{N} \boldsymbol{\sigma}^z_j = U, \tag{703}$$

the generator of the global Ising symmetry. This is significant, since $[H_{\text{TFIM}}, U] = 0$, so that eigenstates of $\boldsymbol{\tau}^z_{\frac{1}{2}}$ are energy eigenstates. Also, we see that it determines the boundary conditions on the $\boldsymbol{\tau}^z_j$ as follows: we observe that

$$1 = \boldsymbol{\tau}^z_{N+\frac{1}{2}} = \begin{cases} \boldsymbol{\tau}^z_{\frac{1}{2}} & \text{if } \prod_j \boldsymbol{\sigma}^x_j = U = 1 \\ -\boldsymbol{\tau}^z_{\frac{1}{2}} & \text{if } \prod_j \boldsymbol{\sigma}^x_j = U = -1 \end{cases} \tag{704}$$

Thus, the distinction between the two symmetry sectors of the Ising model on a ring gets mapped into the distinction between periodic and antiperiodic boundary conditions for he domain-wall operator in the dual description; the operator that says whether the domain wall insertion operator is periodic or antiperiodic is now a dynamical variable.

17.4 Mean Field Theory

So far, we have studied the phase structure of the $d = 1$ TFIM in detail in two distinct limits, $g \gg 1$ and $g \ll 1$. Let us now try and guess the phase diagram over all g, using a variational approximation[9] known as 'mean field theory', in which we completely ignore the entanglement[10] between different sites, and assume that the ground state is a product state,

$$|\text{MFT}\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_N\rangle. \tag{705}$$

Let us further assume a translationally invariant solution, so that each site is identical. Then, the state at each site is parametrized by a unit vector $\vec{n}$, *i.e.* a point on the Bloch sphere, and our uniform variational

[9] Note that the energy expectation value calculated within any variational state is an upper bound for the ground state energy: since the true ground state minimizes the energy, it had better improve on our naive guesswork wavefunction!

[10] You might wonder if it's possible to come up with variational wavefunctions that systematically build in 'just enough' entanglement to capture correlations beyond mean-field theory. Indeed, it turns out that the density-matrix renormalization group can be viewed as a variational optimization over a class of states that do just that: so-called *matrix product states*, that provide a remarkably efficient representation of one dimensional systems. Generalizations of this to higher dimensions using *tensor networks* is a line of research that is being vigorously pursued.

state is given by the product state

$$|\vec{n}\rangle = \otimes_j |\vec{n_j}\rangle = \otimes_j \left[\cos\frac{\theta}{2}|\rightarrow\rangle_j + \sin\frac{\theta}{2}e^{-i\varphi}|\leftarrow\rangle_j\right] \tag{706}$$

Here, we have parametrized the angles with reference to the x-axis in spin space: θ measures the deviation of the spin from the x-axis, and φ represents its azimuthal angle in the yz plane. To compute the variational energy of our mean-field solution, we must compute expectation values, which (owing to the lack of entanglement) follow straightforwardly from those for a single site[11]:

$$\langle\vec{n_j}|\boldsymbol{\sigma}_j^x|\vec{n_j}\rangle = \cos\theta, \quad \langle\vec{n_j}|\boldsymbol{\sigma}_j^z|\vec{n_j}\rangle = \sin\theta\cos\varphi \tag{707}$$

So, we find the variational energy

$$E(\theta,\varphi) = -NJ\left(\sin^2\theta\cos^2\varphi + g\cos\theta\right). \tag{708}$$

Clearly, this is extremized when $\varphi = 0, \pi$, independent of θ; the value of θ that optimizes the energy is then given by $0 = \partial_\theta E(\theta, \varphi = 0, \pi) = NJ\sin\theta(2\cos\theta - g)$.

We see that when $\theta = 0$, the two solutions for φ are the *same* since the azimuthal angle loses its meaning. This makes sense: looking back, $\theta = 0$ corresponds to the paramagnet, and we expect a unique solution, so having two different φ solutions makes no sense. The $\theta = 0$ state clearly corresponds to the solution where the spins are all polarized along the x-axis. Now, this is indeed the unique solution for $g > 2$, since there is no real solution for $\cos\theta = g/2$ in this regime; we may verify that it is a minimum of the energy by computing its second derivative. For $g < 2$, however, we find that in fact the $\theta = 0$ solution is a *maximum* (you should verify this) and instead two new minima appear at $\theta = \cos^{-1} g/2 \neq 0$, and we have $\langle\boldsymbol{\sigma}_i^z\rangle = \sin\theta \neq 0$. Thus, in the mean field approximation, the critical coupling, $g_c = 2$. We see a hallmark of mean-field theory: the ordered state is *overestimated* as we ignore fluctuations that tend to suppress order. (Recall from the duality argument that order is lost at $g = 1$, assuming a single transition, as we have done in the present instance.)

Close to the critical point, we anticipate that θ is small, and so we can Taylor expand: we up to quartic order, we have (setting $\varphi = 0, \pi$):

$$E(\theta) \simeq -NJ\left[\left(\theta - \frac{\theta^3}{6} + \ldots\right)^2 + g\left(1 - \frac{\theta^2}{2} + \frac{\theta^4}{24} + \ldots\right)\right] \simeq -2NJ + \left[\frac{g-2}{2}\theta^2 + \frac{\theta^4}{4}\right] \tag{709}$$

Note that in writing the above expression, we have set $g = 2$ everywhere except in the quadratic term, since the corrections are small except in this case.

It is easy to see that the minima of (709) occur when

$$\langle\boldsymbol{\sigma}_i^z\rangle = \sin\theta \simeq \theta = \pm\sqrt{g_c - g} \tag{710}$$

and the energy near the critical point and in the ordered phase behaves as

$$E_{\mathrm{MFT}}(g) = \begin{cases} \frac{3}{4}(g_c - g)^2, & g \leq g_c \\ 0, & , g \geq g_c \end{cases} \tag{711}$$

Note that $\partial_g E$ is continuous at the transition, indicating that the transition is second-order. This is perhaps a good place to pause and comment on the nature of phase transitions, as we will be spending a *lot* of time with second-order transitions.

What is going on in a phase transition? Note that clearly $E(g)$ has a singular point: we certainly cannot perform a Taylor series expansion near $g = g_c$, since the behavior is strictly zero on one side, and rising on the other, and a moment's thought will convince you that this is never going to be duplicated in any Taylor

[11] The state we have put each site into is known as a *spin coherent state*; it has the neat property that the expectation of the spin components along any axis are the same as the projection of a *classical* spin along the same axis.

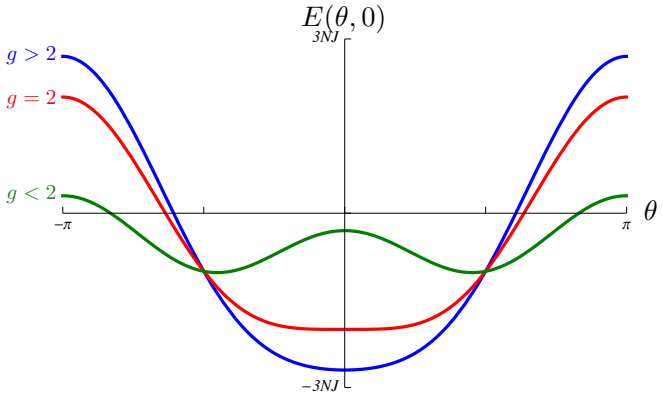


Figure 30: **Mean-Field Energy Function.** The energy functional goes smoothly from having a single minimum ($g > 2$) to having a double minimum ($g < 2$), indicating broken symmetry. At the critical point ($g_c = 2$ in mean-field theory) a careful analysis reveals that a Taylor expansion of the free energy has a vanishing quadratic term.

series. Such a *non-analyticity* in the energy (for classical transition, in the free energy) is the hallmark of a dramatic change in the organization of the ground state (for a classical transition, in the nature of the Boltzmann weight across the statistical ensemble.) How can such non-analyticities arise? There are two common situations:

1. A **first-order quantum phase transition** is simply a level crossing: two eigenstates with totally different character cross in energy as a parameter (call it g, in sympathy with our present example) is varied, and the ground state simply follows the lower branch. If the states truly are totally different in character (*e.g.*, have different quantum numbers), then the states go right through each other as there are no matrix elements between them to provide 'level repulsion'. It is clear that in this case, the first derivative of the energy is discontinuous. Near a first order phase transition, the two states don't need to 'know anything about each other, nor do the excited states above the crossing levels have any indication about the dramatic change in the ground state: the transition comes as a surprise to the ground states, which suddenly find themselves no longer the optimal solution of the Hamiltonian, and the excited states have no idea that anything happened!

2. A **second-order quantum phase transition** (often called a *continuous* phase transition, though this term could apply just as well to higher-order transitions) is one in which the derivative of the energy is continuous, as in the mean-field Ising problem discussed here. Loosely speaking, these are transitions at which the gap above the ground state also vanishes as the system is tuned to its critical point, *i.e.* as $g \to g_c$. Put differently, rather than having a level crossing as in the first-order case, the non-analyticity in the energy as a function of the tuning parameter is because of the spectrum 'collapsing' so that we have an accumulation point of energy eigenvalues near g_c. In this case, both the excitations and the ground states on either side are 'aware' of the phase transition, and we might expect to see significant experimental signatures of the transition. It also turns out that the vanishing of the gap implies that correlation length for local observables *diverges* near the transition.

Returning to the Ising model, let's pause and see where we stand. We have performed a mean field analysis, and found a phase transition as a function of g, at $g_c = 2$. We observe that the transition is continuous, and furthermore, we find that the magnetization scales with exponent $\frac{1}{2}$ near the transition (recall $\langle \boldsymbol{\sigma}_i^z \rangle \propto (g_c - g)^{1/2}$). How much can we trust these results?

It turns out that the seriousness of the neglect of fluctuations in mean field theory (in other words, assuming a product state that is uniform) is highly dependent on dimension. For sufficiently high dimensions ($d \geq d_u$, the *upper critical dimension*), mean field theory becomes exact, and we can trust its predictions for exponents, etc. For sufficiently low dimensions ($d \leq d_l$, the *lower critical dimension*), the neglect of

fluctuations is extremely problematic: the fluctuations are so strong that they destroy the order in the system. For the quantum Ising model[12], it turns out that $d_u = 3$ and $d_c = 0$. As we will show, for $d = 1$ (our present case) mean field theory correctly predicts a transition, but gets g_c wrong, and, more seriously[13] gets the wrong exponent for the rise in the magnetization, which an exact solution reveals is really $\frac{1}{8}$.

17.5 Jordan-Wigner 'Fermionization' of the TFIM

We now introduce a very useful tool in the study of a one-dimensional systems: the so-called Jordan-Wigner transformation, that allows us to map back and forth between spin systems and lattice models of fermions.

First, observe that we can define for each j raising and lowering operators $\boldsymbol{\sigma}_j^{\pm} = \frac{\boldsymbol{\sigma}_j^y \pm i\boldsymbol{\sigma}_j^z}{2}$ which satisfy the commutation and anticommutation relations

$$[\frac{1}{2}\boldsymbol{\sigma}_i^x, \boldsymbol{\sigma}_j^{\pm}] = \pm\boldsymbol{\sigma}_j^{\pm}\delta_{ij} \text{ and } \{\boldsymbol{\sigma}_i^x, \boldsymbol{\sigma}_j^{\pm}\} = 0 \tag{712}$$

and also satisfy

$$\boldsymbol{\sigma}_j^+\boldsymbol{\sigma}_j^- = \frac{1+\boldsymbol{\sigma}_j^x}{2} \text{ and } \boldsymbol{\sigma}_j^-\boldsymbol{\sigma}_j^+ = \frac{1-\boldsymbol{\sigma}_j^x}{2} \implies [\boldsymbol{\sigma}^+, \boldsymbol{\sigma}^-] = \boldsymbol{\sigma}_j^x \text{ and } \{\boldsymbol{\sigma}^+, \boldsymbol{\sigma}^-\} = 1 \tag{713}$$

Observe that $\sigma^{\pm}$ are related to spin flips in the x-basis. Now, let us define the following operators

$$\boldsymbol{c}_i = \boldsymbol{\sigma}_i^- \prod_{l>i} \boldsymbol{\sigma}_l^x \text{ and } \boldsymbol{c}_i^\dagger = \boldsymbol{\sigma}_i^+ \prod_{l>i} \boldsymbol{\sigma}_l^x \tag{714}$$

These can be viewed as an operator that creates a spin flip in the paramagnet, attached to a 'string' of operators attached that together create a domain wall in the ferromagnet (the string is non-local as it should be, since domain walls are nonlocal excitations). So the new object created/annihilated by $\boldsymbol{c}_i^\dagger, \boldsymbol{c}_i$ is something like a bound state of a spin flip and a domain wall, so in a sense combines properties of the excitations that are 'natural' in either phase. So it is perhaps not so surprising that they are linked to thinking about the critical fluctuations

We will now show that these new operators are fermionic, i.e., they obey the canonical anticommutation relations

$$\{\boldsymbol{c}_i, \boldsymbol{c}_j\} = 0 = \{\boldsymbol{c}_i^\dagger, \boldsymbol{c}_j^\dagger\} \text{ and } \{\boldsymbol{c}_i, \boldsymbol{c}_j^\dagger\} = \delta_{ij} \tag{715}$$

For the first case, let us assume without loss of generality that $i < j$. Then,

$$\begin{aligned} \{\boldsymbol{c}_i, \boldsymbol{c}_j\} &= \left(\boldsymbol{\sigma}_i^- \prod_{l>i} \boldsymbol{\sigma}_l^x\right)\left(\boldsymbol{\sigma}_j^- \prod_{l'>j} \boldsymbol{\sigma}_{l'}^x\right) + \left(\boldsymbol{\sigma}_j^- \prod_{l'>j} \boldsymbol{\sigma}_{l'}^x\right)\left(\boldsymbol{\sigma}_i^- \prod_{l>i} \boldsymbol{\sigma}_l^x\right) \\ &= -\boldsymbol{\sigma}_i^-\boldsymbol{\sigma}_j^- \prod_{l>j} \boldsymbol{\sigma}_l^x + \boldsymbol{\sigma}_j^-\boldsymbol{\sigma}_i^- \prod_{l>j} \boldsymbol{\sigma}_l^x = 0 \end{aligned} \tag{716}$$

where the different signs for the two terms follows from the fact that because $i < j$ we had to commute a $\boldsymbol{\sigma}_j^-$ past a $\boldsymbol{\sigma}_j^x$ only in the first term (which we did using the anticommutation relation (712)), and we have used the fact that $(\boldsymbol{\sigma}_j^x)^2 = 1$ to simplify terms where possible. A similar argument gives $\{\boldsymbol{c}_i^\dagger, \boldsymbol{c}_j^\dagger\} = 0$. For the last commutator, let us first consider $i < j$. Using the same logic, we find

$$\begin{aligned} \{\boldsymbol{c}_i, \boldsymbol{c}_j^\dagger\} &= \left(\boldsymbol{\sigma}_i^- \prod_{l>i} \boldsymbol{\sigma}_l^x\right)\left(\boldsymbol{\sigma}_j^+ \prod_{l'>j} \boldsymbol{\sigma}_{l'}^x\right) + \left(\boldsymbol{\sigma}_j^+ \prod_{l'>j} \boldsymbol{\sigma}_{l'}^x\right)\left(\boldsymbol{\sigma}_i^- \prod_{l>i} \boldsymbol{\sigma}_l^x\right) \\ &= -\boldsymbol{\sigma}_i^-\boldsymbol{\sigma}_j^+ \prod_{l>j} \boldsymbol{\sigma}_l^x + \boldsymbol{\sigma}_j^+\boldsymbol{\sigma}_i^- \prod_{l>j} \boldsymbol{\sigma}_l^x = 0, \end{aligned} \tag{717}$$

[12] Actually, we can understand this more simply from the quantum-classical mapping by studying D_u, D_c for the classical Ising model in dimension $D = d + 1$; we find $D_c = 1, D_u = 4$, in accord with our discussion.

[13] Thinking in terms of universality, we care more about exponents than precise values of the critical parameters; nevertheless, overall scales are often worth worrying about if you want to access the properties within some limited parameter space, *e.g.* in a solid state setting we'd be somewhat dismayed if we needed temperatures in the nanoKelvin!

where we have used the fact that spin operators commute when $i \neq j$. However the $i = j$ case is special, since in that case

$$\{\boldsymbol{c}_i, \boldsymbol{c}_i^\dagger\} = \left(\boldsymbol{\sigma}_i^- \prod_{l>i} \boldsymbol{\sigma}_l^x\right)\left(\boldsymbol{\sigma}_j^+ \prod_{l'>i} \boldsymbol{\sigma}_{l'}^x\right) + \left(\boldsymbol{\sigma}_j^+ \prod_{l'>j} \boldsymbol{\sigma}_{l'}^x\right)\left(\boldsymbol{\sigma}_i^- \prod_{l>i} \boldsymbol{\sigma}_l^x\right) = \boldsymbol{\sigma}_i^-\boldsymbol{\sigma}_j^+ + \boldsymbol{\sigma}_i^+\boldsymbol{\sigma}_i^- = 1, \tag{718}$$

since the string does not include $\boldsymbol{\sigma}_i^x$ and so we don't need to worry about the associated -1.

The Jordan-Wigner transformation is invertible — you can try to see how to reverse this starting from the fermion language (hint: you will need a fermionic version of the 'string'). The lesson is that spins and fermions can be mapped into each other by a *nonlocal* transformation in 1D.

Let us now use this to rewrite the TFIM as a model of fermions. Using (713) we see that the transverse field term is

$$\boldsymbol{\sigma}_j^x = 1 - 2\boldsymbol{\sigma}_j^+\boldsymbol{\sigma}_j^- = 1 - 2\boldsymbol{c}_j^\dagger\boldsymbol{c}_j. \tag{719}$$

Now, consider

$$-(\boldsymbol{c}_{j+1}^\dagger + \boldsymbol{c}_{j+1})(\boldsymbol{c}_j^\dagger - \boldsymbol{c}_j) = -(\boldsymbol{\sigma}_{j+1}^- + \boldsymbol{\sigma}_{j+1}^+)\prod_{l>j+1}\boldsymbol{\sigma}_l^x \times (\boldsymbol{\sigma}_j^- - \boldsymbol{\sigma}_j^+)\prod_{l>j}\boldsymbol{\sigma}_l^x = -(\boldsymbol{\sigma}_{j+1}^y)\boldsymbol{\sigma}_j^x(-i\boldsymbol{\sigma}_j^z) = \boldsymbol{\sigma}_{j+1}^z\boldsymbol{\sigma}_j^z$$

Putting these together, we find:

$$\begin{aligned} H_{\text{TFIM}} &= -J\sum_j (\boldsymbol{c}_{j+1}^\dagger - \boldsymbol{c}_{j+1})(\boldsymbol{c}_j^\dagger + \boldsymbol{c}_j) + g(1 - 2\boldsymbol{c}_j^\dagger\boldsymbol{c}_j) \\ &= -NgJ + J\sum_j \left(2g\boldsymbol{c}_j^\dagger\boldsymbol{c}_j - \boldsymbol{c}_{j+1}^\dagger\boldsymbol{c}_j - \boldsymbol{c}_j^\dagger\boldsymbol{c}_{j+1} - \boldsymbol{c}_{j+1}^\dagger\boldsymbol{c}_j^\dagger - \boldsymbol{c}_j\boldsymbol{c}_{j+1}\right). \end{aligned} \tag{720}$$

To simplify further, we introduce Fourier modes of the fields, $\boldsymbol{c}_k = \frac{1}{\sqrt{N}}\sum_j \boldsymbol{c}_j e^{-ikx_j}$, and from the results above, we may rewrite (you should check that you can do this!) the TFIM Hamiltonian as

$$H_{\text{TFIM}} = J\sum_k \left[2\left(g - \cos ka\right)\boldsymbol{c}_k^\dagger\boldsymbol{c}_k - i\sin ka\left(\boldsymbol{c}_{-k}^\dagger\boldsymbol{c}_k^\dagger + \boldsymbol{c}_{-k}\boldsymbol{c}_k\right) - g\right]. \tag{721}$$

This Hamiltonian does *not* conserve fermion number: the second ('pairing'[14]) term violates it in units of two. It turns out any physical fermion Hamiltonian *must* conserve fermion number modulo two: the corresponding variable, *fermion parity* is always a good quantum number. Note that we may ignore the constant term in the above expression for most purposes unless we wish to compute the ground state energy.

There are subtleties to do with boundary conditions that we have been a little bit sloppy about above. Technically speaking, our construction is a bit ambiguous with periodic boundary conditions (since we made a special choice when defining the 'string'). This is closely related to the issues around implementing dualities, and we will not dwell on it, though we have implicitly assumed that things go through correctly with periodic bcs, as we need those to be able to go to Fourier space. Also, a more elegant solution introduces so-called Majorana fermions. We will not discuss these either but they are a very active area of research, as there are interesting properties associated to these fermions on open chains (when the fermionic chain is realised directly in a condensed matter setting.)

17.6 Solution of the Ising Model via Bogoliubov Transformation

The fermionized version of H_{TFIM} written above is quadratic in the $\boldsymbol{c}_k$s but is not quite diagonal, as the pairing terms mix k and $-k$. It turns out this can be solved quite easily. We perform a *fermionic* Bogoliubov transformation by introducing new operators

$$\boldsymbol{\gamma}_k \equiv u_k\boldsymbol{c}_k - iv_k\boldsymbol{c}_{-k}^\dagger \tag{722}$$

[14] The reason for the terminology is that such terms typically arise from superconducting attraction between the electrons.

and demand that the transformation is canonical, *i.e.* preserves the fermionic anticommutators $\{\gamma_k, \gamma_{k'}^\dagger\} = \delta_{k,k'}$. This restricts $u_k = \cos\phi_k/2, v_k = \sin\phi_k/2$. (Note that the difference between preserving commutators and anticommutators are why fermionic Bogoliubov transformations involve ordinary rather than hhyprbolic trigonometric functions — and can be viewed as a linear transformation that mixes creation and annihilation operators by a *unitary* rather than *antiunitary* matrix.) The remaining ambiguity is in the angle ϕ_k which we fix by asking that the the Hamiltonian in terms of γ_k is diagonal, and contains no pairing terms. The resulting condition is $\tan\phi_k = \frac{\epsilon_2(k)}{\epsilon_1(k)}$ where $\epsilon_1(k) = 2J(g - \cos ka), \epsilon_2(k) = -2J\sin ka$ and $H_{\text{TFIM}} = \sum_k \epsilon_k \left(\gamma_k^\dagger \gamma_k - \frac{1}{2}\right)$ with $\epsilon_k = \sqrt{\epsilon_1^2 + \epsilon_2^2}$.

At the end of the procedure, we see that the exact single-particle (here, 'particle' = 'quanta created by $\gamma_k^\dagger$') dispersion is

$$\epsilon_k = 2J\sqrt{1 + g^2 - 2g\cos ka} \tag{723}$$

Note that for any $g \geq 0$, the argument of the square root is positive, so we get a real energy; furthermore the minimum of the energy is at $k = 0$, which gives us the exact gap at all g:

$$\epsilon_k \geq \epsilon_0 = 2J|1 - g| \equiv \Delta(g) \tag{724}$$

— amazingly, exactly what we got from first order perturbation theory[15]. Note that the gap has a nonanalyticity at $g = g_c = 1$; this gets transmitted to other quantities, such as the energy[16].

An important observation we can make is that at criticality, the system is scale invariant: at $g = 1$, the fermions at $k = 0$ become gapless:

$$\epsilon_k \overset{g=g_c}{=} c|k| \tag{725}$$

where the speed of propagation is $c = 2Ja$. There is an emergent Lorentz symmetry (at least in the spectrum) at the critical point; furthermore, the fact that the long-wavelength ($k \sim 0$) modes have the lowest energy allows a continuum description that forgets the details of the lattice. Near the critical point ($|g - g_c| \ll g_c$) we gave

$$\epsilon_k \overset{g\sim g_c}{=} c\sqrt{k^2 + \underbrace{\left(\frac{g - g_c}{a}\right)^2}_{m^2 \to 0}} \tag{726}$$

We identify the second term with a 'mass', or alternatively, with a divergent length scale, $\xi = m^{-1} = a|g - g_c|^{-1}$. We may now introduce some (standard) terminology for critical exponents:

- the **correlation length exponent**, or ν in $\xi \sim |g - g_c|^{-\nu}$ controls the rate of divergence of the correlation length near the critical point. For the $d = 1$ TFIM, $\nu = 1$.
- the **dynamical critical exponent** is the power-law in the dispersion *at* criticality, namely $\epsilon(k, g_c) \propto k^z$. This controls 'length-energy scaling', as the transformation $t \to \lambda x$, $t \to \lambda^z t$ preserves the dispersion, and if we wish to perform RG near the critical point, this tells us the 'correct' RG transformations to use. We expect the continuum critical theory will also be invariant under this scale transformation. For the 1D Ising model, $z = 1$ as in a relativistic (Lorentz-invariant) theory, where space and time are equivalent and there are operations — boosts — that rotate between them.

17.7 Continuum Limit and Critical Theory

So far, we have seen how to exactly rewrite the TFIM in terms of fermionic degrees of freedom, and obtained a nice understanding of how the ferromagnetic and paramagnetic limits work. We already had nice variables to describe those, though — in a sense, the fermions are overkill in those limits, when we can think in terms

[15] Don't expect such yeomanlike service from first order perturbation theory outside of this very special example.

[16] It is clear that for any g we can compute the ground state energy of the fermions; by the quantum-to-classical mapping, we can relate this to the free energy of the classical Ising model as a function of temperature, the famous Onsager solution!

of domain walls or spins. Where the fermionic description comes into its own, though, is in addressing the critical properties.

In this section, we will show how to take a quantum lattice model (with continuous time but discrete space) into a continuum field theory, in which *both* space and time are continuous variables. This is known as the *continuum scaling limit* of the model, and is the most convenient way to access universal properties of the transition using analytical tools. We have actually already has some experience with the subtleties of continuum scaling limits hen we discussed the quantum-classical correspondence (e.g. for the 2D classical Ising model), but for now it's perhaps best to jump right into an example.

A continuum field theory emerges near the critical point where the correlation length is large. For $\xi \gg a$, define the continuum fermion field

$$\boldsymbol{\Psi}(x_j) \equiv \frac{1}{\sqrt{a}}\boldsymbol{c}_j \tag{727}$$

where the factor of $\sqrt{a}$ is designed to convert Kronecker deltas to Dirac deltas: we want

$$\{\boldsymbol{c}_j, \boldsymbol{c}_{j'}^\dagger\} = \delta_{jj'} \leftrightarrow \{\boldsymbol{\Psi}(x), \boldsymbol{\Psi}^\dagger(x')\} = \delta(x-x') \tag{728}$$

Now, we have

$$\boldsymbol{c}_k = \frac{1}{\sqrt{N}}\sum_j \boldsymbol{c}_j e^{-ikx_j} = \sqrt{\frac{a}{L}}\sum_j \sqrt{a}\boldsymbol{\Psi}(x_j)e^{-ikx_j} = \sum_{j|x_j=ja} \frac{a}{\sqrt{L}}\boldsymbol{\Psi}(x_j)e^{-ikx_j} \underset{a\to 0}{=} \int dx \frac{e^{-ikx}}{\sqrt{L}}\boldsymbol{\Psi}(x) \tag{729}$$

We can now use this in H_{TFIM}, and work in the $ka \ll 1$ limit (as these are the lightest modes and hence dominate the continuum limit). We find

$$J\sum_k (g - \cos ka)\boldsymbol{c}_k^\dagger \boldsymbol{c}_k \to (g - g_c)\int dx\, \boldsymbol{\Psi}^\dagger(x)\boldsymbol{\Psi}(x) \tag{730}$$

where we have dropped higher order terms in an expansion[17] in terms of $a\partial_x$. The second term gives us, again to leading order

$$-iJ\sum_k \sin ka \boldsymbol{c}_{-k}\boldsymbol{c}_k \to \frac{c}{2}\int dx\, \boldsymbol{\Psi}^\dagger(x)\partial_x \boldsymbol{\Psi}^\dagger(x). \tag{731}$$

Putting it all together, we have at leading order in our *gradient expansion* the continuum field theory

$$H_{\text{TFIM}} \xrightarrow{\xi \gg a} \frac{c}{2}\int dx \left[\boldsymbol{\Psi}^\dagger(x)\partial_x\boldsymbol{\Psi}^\dagger(x) - \boldsymbol{\Psi}(x)\partial_x\boldsymbol{\Psi}(x)\right] + \Delta\int dx\, \boldsymbol{\Psi}^\dagger\boldsymbol{\Psi} \tag{732}$$

with $c = 2Ja$ and $\Delta = 2J|g-1|$. In order to proceed, we need to build some technology to construct fermionic path integrals.

13: Aside: Grassmann Variables and Fermionic Coherent States

We start somewhat obliquely, by asking the question: what does it mean to have a coherent state for a fermionic operator? By a coherent state we mean a state with the property that

$$\boldsymbol{c}|\psi\rangle = \psi|\psi\rangle \tag{733}$$

where $\boldsymbol{c}$ is a (complex) fermion operator and ψ is a c-number, *i.e.* is not an operator. However, ψ cannot be your run-of-the-mill number, since we must have $\boldsymbol{c}^2 = 0$, it must follow that $\psi^2 = 0$. Such a number is called a *Grassmann number*. They have many peculiar properties, but it turns out that they are invaluable (and quite simple) to work with in theories that involve fermions. In particular, they are absolutely essential in order to build a path-integral description.

[17] A useful trick to remember is that $k \leftrightarrow -i\partial_x$.

The first rule to remember is that the Grassmann numbers anticommute with each other and with fermionic operators

$$\psi_1\psi_2 = -\psi_2\psi_1, \psi\boldsymbol{c} = -\boldsymbol{c}\psi \tag{734}$$

while commuting with bosonic operators and ordinary numbers. With Grassmanns, it's easy to define a fermionic coherent state as

$$|\psi\rangle = |0\rangle - \psi|1\rangle \tag{735}$$

where $|0\rangle$ and $|1\rangle = \boldsymbol{c}^\dagger|0\rangle$ are the empty and filled states. You may verify quite easily that this works. Similarly, if we take $\langle\bar{\psi}| = \langle 0| + \bar{\psi}\langle 1|$, it is easy to show that this is a left eigenvector of the creation operator: $\langle\bar{\psi}|\boldsymbol{c}^\dagger = \langle\bar{\psi}|\bar{\psi}$. One should note that the fermionic coherent states are somewhat subtle: we have enlarged the Hilbert space to allow Grassmann coefficients rather than ordinary numbers. Also, $\bar{\psi}$ is *not* the complex conjugate of ψ and $\langle\bar{\psi}|$ is *not* the adjoint of $|\psi\rangle$: indeed, their overlap is

$$\langle\bar{\psi}|\psi\rangle = 1 + \bar{\psi}\psi = e^{\bar{\psi}\psi}. \tag{736}$$

We will also need to do calculus with Grassmann variables; while this might seem scary, were we allowed to use Grassmans instead of ordinary variables, you could summarize your entire A-level calculus textbook in less than a page! The previous calculation of an overlap has already given a nice example of how simple Taylor's theorem is for Grassman variables:

$$f(\psi) = f_0 + f_1\psi \tag{737}$$

since higher powers all vanish! Integration is as easy as taking derivatives, since

$$\int \psi d\psi = 1, \int 1 d\psi = 0, \tag{738}$$

though with more than one Grassmann we have to worry about the order:

$$1 = \int \bar{\psi}\psi d\psi d\bar{\psi} = -\int \bar{\psi}\psi d\bar{\psi} d\psi \tag{739}$$

(You can remember this by noting that the positive sign is associated with the order of Grassmanns in the integrand being exactly the reverse of the order of integration.) Note that the definition of the integral is counterintuitive. More properly, we should view the integral as analogous to an indefinite integral for ordinary numbers, with limits $\pm\infty$. For this analogy to be sensible, we require 'shift invariance': namely, that $\int f(\psi+\eta)d\psi = \int f(\psi)d\psi$ for a constant (Grassmann) η. However, taking $f(\psi+\eta) = \psi + \eta$ and assuming η is a constant, this requires $\int d\psi = 0$. Then, given that the most general Grassmann function is $f(\psi) = f_0 + f_1\psi$, the only way to have any nonzero integrals at all is to assert that $\int \psi d\psi \neq 0$. This leaves freedom of normalization; we choose the simplest consistent choice, $\int \psi d\psi = 1$.
Really the only integral we need to care about is the Gaussian integral,

$$\int e^{-a\bar{\psi}\psi} d\bar{\psi} d\psi = a \tag{740}$$

and its multivariate generalization

$$\int e^{\bar{\psi}\cdot A\cdot\psi} d\bar{\psi} d\psi = \det A \tag{741}$$

where

$$\bar{\psi}\cdot A\cdot\psi = (\bar{\psi}_1, \ldots \bar{\psi}_M) \begin{pmatrix} A_{11} & A_{22} & \cdots \\ A_{21} & \ddots & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \psi_1 \\ . \\ \psi_M \end{pmatrix} \tag{742}$$

and we solve by performing a change of variable that brings A to diagonal form. Note that the results of Gaussian integration over Grassmann variables is exactly the inverse of that for ordinary numbers. The crucial result we will need is the resolution of the identity in fermionic coherent states:

$$\mathbf{1} = \int d\bar{\psi} d\psi \, e^{-\bar{\psi}\psi} |\psi\rangle\langle\bar{\psi}| \tag{743}$$

and, finally, for any operator $\mathbf{A}$ that is bosonic (made up of an even number of fermionic operators and Grassmanns) we have

$$\mathrm{tr}\mathbf{A} = \int d\bar{\psi} d\psi \, e^{-\bar{\psi}\psi} \langle -\bar{\psi}|\mathbf{A}|\psi\rangle \tag{744}$$

The sign is important — we'll come back to it in a bit.

17.7.1 Path Integral and Critical Theory for the TFIM

The reason for all this rigmarole is the following: we wish to understand the quantum statistical mechanics of the many-body problem of the Ising chain in its fermionic language. The derivation of the fermionic path integral closely parallels that for bosons, except that it uses Grassmannn variables.

We would like to compute the partition function ($\beta = 1/T$)

$$Z(\beta) = \mathrm{tr}\, e^{-\beta H} = \int \prod_{j=1}^{L} d\bar{\psi}_0(x_j) d\psi_0(x_j) \, e^{-\bar{\psi}_0\psi_0} \langle -\{\bar{\psi}_0(x)\}|e^{-\beta H}|\{\psi_0(x)\}\rangle \tag{745}$$

where we have introduced a composite resolution of the identity for all the sites in the system: $|\psi_0(x)\rangle = \otimes_{j=1}^{L} |\psi_0(x_j)\rangle$, i.e. we have introduced a Grassman variable parametrizing a coherent state for each fermionic site x_j. How are we to compute this partition function? The key tool is to use the path integral approach adapted to many-fermion problems, which we now develop.

As a warm-up, let us compute a simpler example, the partition function for a *single* fermionic mode, with Hamiltonian $H = \boldsymbol{c}^\dagger \boldsymbol{c}(\omega_0 - \mu)$. A particularly useful trick is to *Trotterize*: we write

$$e^{\beta H} = \underbrace{(1 - \Delta\tau H)(1 - \Delta\tau H)\ldots(1 - \Delta\tau H)}_{M \text{ times}} \tag{746}$$

where $\Delta\tau = \frac{\beta}{M}$. You will recognise that since $\lim_{\Delta\tau\to 0}(1 - \Delta\tau H)^{\beta/\Delta\tau} = e^{-\beta H}$, this is a reasonable approximation as long as $M \to \infty$.

The reason for Trotterizing is that we can now insert resolutions of the identity of the form (743) between the different $(1 - \Delta\tau_H)$: we see that there are $M - 1$ such points of insertion, so that

$$Z(\beta) = \int \prod_{l=0}^{M-1} d\bar{\psi}_l d\psi_l \, e^{-\bar{\psi}_l\psi_l} \langle \bar{\psi}_{l+1}|(1 - \Delta\tau H)|\psi_l\rangle. \tag{747}$$

where we have defined

$$\bar{\psi}_M = -\bar{\psi}_0, \psi_M = -\psi_0 \tag{748}$$

so that we can replace $\langle -\bar{\psi}_0| = \langle \bar{\psi}_N|$. We can now use the coherent state property to turn matrix elements into Grassman-valued functions:

$$\langle \bar{\psi}_{l+1}|(1 - \Delta\tau H(\boldsymbol{c}^\dagger, \boldsymbol{c}))|\psi_l\rangle = \langle \bar{\psi}_{l+1}|(1 - \Delta\tau H(\bar{\psi}_{l+1}, \psi_l))|\psi_l\rangle \overset{\Delta\tau\to 0}{=} e^{\bar{\psi}_{l+1}\psi_l} e^{-\Delta\tau H(\bar{\psi}_{l+1},\psi_l)} \tag{749}$$

Note that it was crucial that H was *normal ordered*, *i.e.* had all its $\boldsymbol{c}^\dagger$s to the right of its $\boldsymbol{c}$s.

$$\begin{aligned} Z(\beta) &= \int \prod_{l=0}^{M-1} d\bar{\psi}_l d\psi_l \, e^{-\bar{\psi}_l \psi_l} e^{\bar{\psi}_{l+1}\psi_l} e^{-\Delta\tau H(\bar{\psi}_{l+1},\psi_l)} \\ &= \int \prod_{l=0}^{M-1} d\bar{\psi}_l d\psi_l \, \exp\left[\Delta\tau \left(\underbrace{\frac{\bar{\psi}_{l+1}-\bar{\psi}_l}{\Delta\tau}}_{=\partial_\tau\bar{\psi}} \psi_l - H(\bar{\psi}_{l+1},\psi_l)\right)\right] \\ &\simeq \int [D\bar{\psi}D\psi] \exp\left[\int_0^\beta d\tau \, \bar{\psi}(\tau)(-\partial_\tau - \omega_0 + \mu)\psi(\tau)\right] = \int [D\bar{\psi}D\psi] e^{-S[\bar{\psi},\psi]} \end{aligned} \tag{750}$$

We should note that:

- In the penultimate step, we defined continuum (Grassmann) fields

$$\psi(\tau_l = \Delta\tau l) \equiv \psi_l, \bar{\psi}(\tau_l = \Delta\tau l) \equiv \bar{\psi}_l \tag{751}$$

- We also wrote $H(\bar{\psi}_{l+1},\psi_l) = H(\bar{\psi}_l,\psi_l) + \mathcal{O}(\Delta\tau)$. While this is usually a negligible distinction, it does sometimes matter (we will see an example in the homework).

- The antiperiodic boundary conditions (cf. (748)) that require $\psi(\tau+\beta) = -\psi(\tau)$ mean that in the Fourier representation[18]

$$\psi(\tau) = \frac{1}{\beta}\sum_n \psi(\omega_n) e^{-i\omega_n\tau}, \bar{\psi}(\tau) = \frac{1}{\beta}\sum_n \bar{\psi}(\omega_n) e^{i\omega_n\tau}, \tag{752}$$

  the *Matsubara frequencies* $\omega_n = (2n+1)\pi/\beta, n \in \mathbb{Z}$ are half-integer multiples of π/β.

- The measure $[D\bar{\psi}D\psi]$ is implicitly defined. We should view this the same way as we viewed the continuum limit of the scalar field theory: with mathematical suspicion but physical confidence. We'll encounter several examples in the future.

- The derivative of a Grassmann function is also implicitly defined; however, there is no sense in which one can think of Grassmann numbers as 'big' or 'small', so the numerator is heuristic at best.

- In the final step, we integrated by parts *i.e.* relabeled terms in the sum, as follows

$$\sum_l (\bar{\psi}_{l+1} - \bar{\psi}_l)\psi_l = \sum_l \bar{\psi}_{l+1}\psi_l - \sum_l \bar{\psi}_l\psi_l = \sum_{l'=l-1} \bar{\psi}_{l'}\psi_{l'-1} - \sum_l \bar{\psi}_{l1}\psi_l = -\sum_l \bar{\psi}_l(\psi_{l+1}-\psi_l) \tag{753}$$

  without moving Grassmanns past each other, and hence the sign makes sense.

The most important thing to take away is that the quantum partition function can be computed using a path integral formulation where we use the Euclidean action[19]

$$S[\bar{\psi},\psi] = \int d\tau \, \left(\bar{\psi}\partial_\tau\psi + H(\bar{\psi},\psi)\right). \tag{754}$$

It's straightforward to extend this to higher dimensional fermionic integrals: we would just insert a resolution of identity at each spatial point at each time slice (as in the definition of $|\{\psi_0\}\rangle$, above), and then

[18] Though $\bar{\psi} \neq \psi^*$, the relative sign in the Fourier transformation is still convenient.

[19] The reason this is 'Euclidean' is because τ plays the role of time, but it's *imaginary time* — and as a result, we measure spacetime distance using the Euclidean metric, $ds^2 = -d\tau^2 - d\mathbf{x}^2$ rather than the Minkowski metric $ds^2 = dt^2 - d\mathbf{x}^2$.

it's straightforward to show that we must now integrate over all space-time paths (since we may write, generalizing the previous example, $\prod_j [D\bar{\psi}(x_j) D\psi(x_j)] = [D\bar{\psi} D\psi]$) so that $Z = \int [D\bar{\psi} D\psi] e^{-S[\bar{\psi},\psi]}$, with

$$S[\bar{\psi},\psi] = \int d\tau d^d x \, \left(\bar{\mathbf{\Psi}}(x) \partial_\tau \mathbf{\Psi}(x) + \mathcal{H}(\bar{\mathbf{\Psi}}(x), \mathbf{\Psi}(x)) \right) \equiv \int d\tau d^d x \, \mathcal{L} \tag{755}$$

where we have introduced the Hamiltonian *density* $\mathcal{H}$, which when integrated over spatial coordinates, yields the Hamiltonian H, and implicitly defined the (Euclidean) Lagrangean density $\mathcal{L}$

In both cases, it's important to note that the kinetic term is first-order in time; this is sometimes called a 'Berry phase term'. We note that this term is *actually*

$$\sum_{l=0}^{M-1} \bar{\psi}_{l+1} (\psi_{l+1} - \psi_l) = \frac{1}{\beta} \sum_n \bar{\psi}(\omega_n) [1 - e^{i\omega_n \Delta\tau}] \psi(\omega_n) \tag{756}$$

but in writing it as a derivative, we have only kept the leading nonzero term, as we substituted $1 - e^{i\omega_n \Delta\tau} \to -i\omega_n \Delta\tau$. This is cleary innocuous if $\omega_n \Delta\tau \ll 1$ for all ω_n that matter. It turns out this is indeed the case for reasonable Hamiltonians and reasonable quantities (and like good physicists, we are cagey on the precise definition of 'reasonable', by claiming 'we'll know it when we see it.')

Armed with these results, we are ready to return to our study of the Ising chain. Recall that we had, in the continuum limit, that

$$H_{\text{TFIM}} \xrightarrow{\xi \gg a} \frac{c}{2} \int dx \, \left[\mathbf{\Psi}^\dagger(x) \partial_x \mathbf{\Psi}^\dagger(x) - \mathbf{\Psi}(x) \partial_x \mathbf{\Psi}(x) \right] + \Delta \int dx \, \mathbf{\Psi}^\dagger \mathbf{\Psi} \tag{757}$$

with $c = 2Ja$ and $\Delta = 2J|g-1|$. The Euclidean action is $S[\bar{\Psi}, \Psi] = \int d\tau \int dx \, \mathcal{L}$, where

$$\mathcal{L} = \bar{\Psi} \partial_\tau \Psi + \mathcal{H} = \bar{\Psi} \partial_\tau \Psi + \frac{c}{2} \left(\bar{\Psi} \partial_x \bar{\Psi} - \Psi \partial_x \Psi \right) + \Delta \bar{\Psi} \Psi \tag{758}$$

obtained by simply replacing field operators $\mathbf{\Psi}^\dagger(x), \mathbf{\Psi}(x)$ with their Grassmann eigenvalues $\bar{\Psi}(x), \Psi(x)$. It's actually much nicer to change variables again, by defining new variables $\psi_\pm$: we write

$$\Psi = \psi_+ + \psi_- + i(\psi_+ - \psi_-), \bar{\Psi} = \psi_+ + \psi_- - i(\psi_+ - \psi_-) \tag{759}$$

in terms of which we have

$$\mathcal{L} = \sum_\pm \psi_\pm \left(i\partial_\tau \pm \partial_x \right) \psi_\pm + \Delta \psi_+ \psi_-. \tag{760}$$

Note that the $\psi_\pm$ are counterpropagating modes; if we think of periodic boundary conditions in space, each has a definite 'handedness'. They are sometimes referred to as *chiral* modes. The mass term serves to couple the counterpropagating modes — colloquially, 'backscattering' them into each other.

We observe that the partition function now looks like a classical statistical mechanics problem (except that the classical fields are Grassman variables!) If you've taken (or plan to take) a course on RG [20], you know (will soon know) what to do now: we simply ask what operators are relevant, irrelevant or marginal under the RG. Recall that at its heart, the RG is scaling momenta to smaller and smaller scales, and distances to longer and longer scales. So we should study how the action behaves under scale transformations. In order to keep S invariant, we had better (remembering that $z = 1$) scale space and time the same way, i.e., $x \to \lambda x, t \to \lambda t$. How should we scale the field? The action has two pieces, and they would give us different answers. Which is correct? Here's a useful way to think about this: we could always set the mass term to zero — then we'd be at the critical point. This is the scale-invariant theory we're looking for, and it is only scale invariant if we specify the correct transformation for the $\mathbf{\Psi}$s. This leads us to the scaling rule

$$x \to \lambda x, t \to \lambda t, \mathbf{\Psi} \to \lambda^{-\frac{1}{2}} \mathbf{\Psi} \tag{761}$$

[20] In the MMathPhys, it's taught in Hilary Term.

So, just by scaling, we see that[21]

$$\langle \Psi^\dagger(x)\Psi(0)\rangle \sim \frac{1}{x}. \tag{762}$$

Indeed, as the theory is Gaussian — has a quadratic action — one could even calculate the precise coefficient.

More generally the scaling dimension δ of an operator $\mathcal{O}(x)$ in $\mathcal{O}?\lambda^{-\delta}\mathcal{O}$ in a scale-invariant quantum field theory determines its vacuum autocorrelation functions to be

$$\langle \mathcal{O}^\dagger(x)\mathcal{O}(0)\rangle \sim \frac{1}{x^{2\delta}}. \tag{763}$$

As we mentioned before, the mass perturbation has a different scaling than the first (derivative) piece. Given the scaling in (761), we see that

$$\int d\tau dx\, \bar{\Psi}\Psi \to \lambda \int d\tau dx\, \bar{\Psi}\Psi, \tag{764}$$

or in other words, that as we makes length bigger and bigger, the mass term *grows* — this is what we call a relevant perturbation, and it will actually dominate the behavior at long distances. It turns out that we knew this: if we tip the Ising model on either side of its critical point, it is described by a theory that at the longest length scales eventually looks 'boring', as all the action lies at very high energies (of the order the gap). It is only *at* the critical point that we have gapless excitations, and therefore interesting behavior at *all* length scales. The scaling dimension of the relevant perturbation determines the correlation length exponent: again, this should make sense, since the correlation length must diverge as the critical point is approached, and the proximity to the critical point is controlled by the relevant perturbation. To see this more formally recall that by definition, the correlation length is the length scale $\xi = a\lambda$ at which the relevant coupling $\Delta = \lambda^\nu \Delta_0$ has evolved to order one (Δ_0 is the 'bare' value at lattice scale a, and ν its scaling dimension.) From this, we find

$$1 = (\xi/a)^\nu \Delta_0 \implies \xi \sim \frac{1}{\Delta_0^\nu} \tag{765}$$

and since ν is the scaling dimension of the relevant operator Δ, we have confirmed that it is the correlation length exponent.

Are there any other *relevant* operators? If there were, the Ising transition would actually be *multicritical*: we'd have to tune more than one parameter to zero in order to stay at criticality. We demand that Ising symmetry is preserved; in the fermion language, the Ising symmetry is captured by $\Psi(x) \to -\Psi(x)$. So we should only consider operators containing an even number of fermions. All other operators that can be made of an even number of fermions are *irrelevant*: they shrink as we go to larger distances. A candidate marginal operator is $\bar{\Psi}\Psi\bar{\Psi}\Psi$, that vanishes due to Fermi statistics. Operators with higher derivatives have no chance to survive:

$$\int d\tau dx\, \bar{\Psi}\partial_x^2\Psi \sim \lambda^{-1}, \qquad \int d\tau dx\, \bar{\Psi}\partial_x\Psi\bar{\Psi}\partial_x\Psi \sim \lambda^{-2} \tag{766}$$

So, we only need to tune one parameter to access the transition, so we have verified that the transition in the TFIM is really critical. Note that this is more powerful than just cooking up a model where we only tune one parameter; such models are always approximate and you might worry in principle that other terms that you ignored may push you off the critical point. However, we have analyzed the continuum scaling limit and concluded that the transition is a universality class with exactly one relevant perturbation, and once we find it (g in our case) that's it: we have exhausted the list. Note also that having no classically marginal operators ('classical' here being shorthand for 'just do naive dimensional analysis without worrying about higher order terms in the RG') we need not worry about higher terms coming in to change things — small quantum corrections cannot beat the order-1 irrelevancy we have computed. Were there exactly marginal operators at the classical level, this higher level analysis would be absolutely necessary.

[21] In a $z = 1$ theory, it's enough to specify the scaling of spacelike-separated correlators as we can boost to get the full x, t dependence.

There *are* operators of dimension[22] less than 2, namely the operator $\boldsymbol{\sigma}^z$ itself. This operator creates a *branch cut* in the fermion fields, since they are made from domain wall operators. It turns out that the scaling dimension of this operator is 1/8, and this is important as it tells us the scaling of the magnetization as $g \to g_c$ from below (since σ^z is the quantity that measures the magnetization.) However, it is *not* an allowed perturbation to the Ising critical theory as it violates the Ising symmetry, and breaking this symmetry destroys the transition!

Notice that the critical theory is *completely different* from the naive guess we could have made from mean field analysis: we might have said that we should just 'add fluctuations to mean field theory' and thus written a theory in terms of the (bosonic) mean field variable θ. At $g = g_c$ we'd have said that the theory should be Gaussian with a fixed-point action

$$S_0[\theta] = \int d\tau dx \left((\partial_\tau \theta)^2 + (\partial_x \theta)^2 \right) \tag{767}$$

that preserves the $\theta \to -\theta$ symmetry. However, in this case, *many* relevant operators preserve the symmetry: $\theta^2, \theta^4, \ldots$ and indeed *all even functions* are potentially relevant. This would be a grave error in $d = 1$, though for $d > 1$ we have the correct variables (though not the right fixed point), and is perfectly OK for $d \geq 3$: we have identified both the variables and the fixed point correctly!

14: Experiments

We close our discussion of the Ising model by giving a brief discussion of experimental studies of the various quantum-critical properties that we have studied. In order to do this, we need to introduce some technology.

17.7.2 Correlation Functions, Structure Factors, and Long-Range Order

A useful set of physical observables to study are the correlations functions of the spins:

$$C(r, r') \equiv \langle 0 | \boldsymbol{\sigma}^z_r \boldsymbol{\sigma}^z_{r'} | 0 \rangle. \tag{768}$$

Notice that this is an *equal-time* correlator. For a translation-invariant ground state, this is only a function of the separation of the spins, $C(r, r') = C(r - r')$ (we are using the convention that $r = ja$ to go from lattice units to actual distances. The translation invariance allows us to take a Fourier transform of the correlation function, called the *static structure factor*:

$$S(q) \equiv \sum_r e^{-iqr} C(r) \tag{769}$$

where the sum is over the discrete positions $r \in a\{1, 2, \ldots, N\}$; assuming periodic boundary conditions, we have $q \in \frac{2\pi}{Na}\{1, 2, \ldots, N\}$. How do the correlation functions behave across the phase diagram?

- At $g = \infty$, deep into the paramagnetic phase, we have

$$\langle 0 | \boldsymbol{\sigma}^z_j \boldsymbol{\sigma}^z_{j'} | 0 \rangle \big|_{g \to \infty} = \delta_{jj'} \tag{770}$$

  which is the zero correlation length limit of a stable phase. Different sites are completely uncorrelated, and at finite but large $g \gg 1$, the correlations decay exponentially,

$$\langle 0 | \boldsymbol{\sigma}^z_j \boldsymbol{\sigma}^z_{j'} | 0 \rangle \big|_{g \gg 1} \overset{|x_j - x_{j'}| \gg a}{\simeq} e^{-|x_j - x_{j'}|/\xi}. \tag{771}$$

  In other words, fluctuations of spin flips communicate between sites, but only over a short range as the excitations are massive.

[22] The dimension usually counts the operator without the two dimensions in the integrand, so a dimension of 2 marks the dividing line between 'blow up' (dimension less than 2) and 'shrink' (dimension greater than 2), *i.e.* marginal operators have dimension 2. Note the quirk of terminology that we count the *energy* dimension rather than the *time* dimension!

- In the ferromagnetic phase, for $g \ll 1$, we have

$$\langle 0|\boldsymbol{\sigma}_j^z \boldsymbol{\sigma}_{j'}^z|0\rangle\big|_{g\gg 1} \overset{|x_j - x_{j'}|\gg a}{\simeq} N_0^2(g) \tag{772}$$

  where $N_0^2(g) = 1$ for $g = 0$, and vanishes as $g \to g_c$. More simply, if we *train* the system (in the sense if having an infinitesimal field that tells the order parameter which way to point, we have $\langle 0|\boldsymbol{\sigma}_j^z|0\rangle = \pm N_0(g)$. Here N_0 is the order parameter, the magnetization. Note that the 'training procedure' entails a sensitive order of limits: in any *finite* system N_0 must vanish, so we must have

$$\lim_{N\to\infty}\lim_{h\to 0}\langle \boldsymbol{\sigma}_i^z\rangle = 0, \text{ but } \lim_{h\to 0}\lim_{N\to\infty}\langle \boldsymbol{\sigma}_i^z\rangle = N_0. \tag{773}$$

  It is in this sense that we say that the ferromagnet has long-range order.

One thing should bother you about the ferromagnet: it looks like it has a divergent correlation length, and we earlier claimed that this is a property only of *critical* theories. This is indeed true, but it turns out the *correct* correlation function is the so called *connected* correlation function, that measures the extent to which two spins are correlated relative to the fact that their two-point function would have an expectation value just by virtue of each having a nonzero mean: for the Ising case, this is the quantity

$$G_{jl}^{zz} = \langle \boldsymbol{\sigma}_j^z \boldsymbol{\sigma}_l^z\rangle - \langle \boldsymbol{\sigma}_j^z\rangle\langle \boldsymbol{\sigma}_l^z\rangle \tag{774}$$

which you will recognize as the second cumulant of the distribution (also known as the variance). You may argue quite easily that the connected correlator indeed has a vanishing length scale as $g \to 0$, and a divergent one as $g \to g_c$, showing the expected behavior. In other words, the correct thing to measure are the *fluctuations* about the mean — it is these that diverge at criticality. The most important thing to take away, though, is that the long-range order gives a sharp distinction between phases: *no smooth function of g can interpolate between the exponential decay and long-range ordered behavior, so something singular must occur at $g = g_c$.*

How does the structure factor reflect this physics? It turns out that when there is long-range order (and for simplicity, assuming $N \to \infty$) so we can use continuous Fourier transforms, the structure factor has a delta function singularity at the origin, whose weight is propertional to the square of the magnetization:

$$S(q) = 2\pi\delta(q)N_0^2 + \text{regular} \tag{775}$$

where 'regular' means terms smooth as $q \to 0$. This reflects the broken symmetry.

If we write[a] $S(q) = \langle 0|\boldsymbol{\sigma}_q^z \boldsymbol{\sigma}_{-q}^z|0\rangle = \sum_n \left|\langle n|\boldsymbol{\sigma}_q^z|0\rangle\right|^2$ (using $\boldsymbol{\sigma}_q^z = \boldsymbol{\sigma}_{-q}^{z\dagger}$) then we have a *sum rule*

$$\int_{-\pi}^{\pi}\frac{dq}{2\pi}S(q) = C(r - r' = 0) = \langle 0|\boldsymbol{\sigma}_r^z \boldsymbol{\sigma}_r^z|0\rangle = 1. \tag{777}$$

which basically remembers that we have a normalization condition in somewhere. A consequence of a sum rule is that spectral weight can only *rearrange* — the total weight is conserved, and so the delta function at the origin has to come from somewhere.

More information can be obtained from the *dynamical structure factor*, that measures correlations *both* in space and in time. We begin with the correlation function

$$C(r,t) \equiv \langle 0|\boldsymbol{\sigma}_r^z(t)\boldsymbol{\sigma}_0^z(0)|0\rangle \tag{778}$$

which reduces to the equal-time correlator when $t = 0$, and is a temporal autocorrelator when $r = 0$. It is useful to think of $\boldsymbol{\sigma}_r^z(t)$ in the Heisenberg picture: $\boldsymbol{\sigma}_r^z(t) = e^{iHt}\boldsymbol{\sigma}_r^z e^{-iHt}$. The dynamic structure

factor is then[b]

$$\begin{aligned} S(q,\omega) &= \sum_r e^{-iqr} \int_{-\infty}^{\infty} dt\, e^{i\omega t} C(r,t) \\ &= \sum_r e^{-iqr} \int_{-\infty}^{\infty} dt\, e^{i\omega t} \langle 0|\boldsymbol{\sigma}_r^z(t)\boldsymbol{\sigma}_0^z(0)|0\rangle \\ &= \sum_r e^{-iqr} \int_{-\infty}^{\infty} dt\, e^{i\omega t} \langle 0|e^{iHt}\boldsymbol{\sigma}_r^z e^{-iHt}\boldsymbol{\sigma}_0^z|0\rangle \\ &= \sum_r e^{-iqr} \int_{-\infty}^{\infty} dt\, e^{i\omega t} \sum_n \langle 0|e^{iHt}\boldsymbol{\sigma}_r^z e^{-iHt}|n\rangle\langle n|\boldsymbol{\sigma}_0^z|0\rangle \\ &= \sum_r e^{-iqr} \int_{-\infty}^{\infty} dt\, e^{i(\omega+E_0-E_n)t} \langle 0|\boldsymbol{\sigma}_r^z|n\rangle\langle n|\boldsymbol{\sigma}_0^z|0\rangle \\ &= 2\pi \sum_n \left|\langle n|\boldsymbol{\sigma}_q^z|0\rangle\right|^2 \delta(\omega + E_0 - E_n) \end{aligned} \tag{779}$$

where in the last step we recognized the result from the calculation of $S(q)$ previously. $S(q,\omega)$ is the number of states that differ by momentum q and energy ω, weighted by (the square of) their overlap with the state created by acting on the ground state with $\boldsymbol{\sigma}_j^z$ — this piece is often called the 'matrix element', for obvious reasons. It turns out (and I won't go into detail here but you should remember a version of it from courses on condensed matter and scattering theory) that an application of Fermi's golden rule tells us that this is the transition rate in an (inelastic) scattering experiment; the obvious tool is scattering of neutrons as they will have a nonzero matrix elements for exciting the spins, as a neutron couples to spin via its magnetic dipole moment.

17.7.3 Ising Symmetry in $CoNb_2O_6$

Information about the excitations may be obtained by studying the inleastic neutron data: inelastic scattering of a neutron can create a spin flip, as we have argued above spin flips behave very differently in the two phases (the corresponding evolution of $S(q,\omega)$ is shown in Fig. 32.)
From our knowledge of the spectrum, we can infer the behavior of $S(q,\omega)$ in different limits of the phase diagram. At large g, where the low-energy spectrum is mostly built from spin flips As these are created by $\boldsymbol{\sigma}^z$ directly, we should see a contribution to $S(q,\omega)$ from spin flips — but *only* if there is the correct energy momentum relation $\omega = \epsilon_q$ where ϵ_q is the spin flip dispersion. So we should see a sharp peak in $S(q,\omega)$ along this curve for ω in the right range; for large ω, we merge into the 'multiparticle' continuum where we have many spin flips with the energy distributed haphazardly among them. It turns out this continuum begins near 3Δ, as the matrix element vanishes for an *even* number of spin flips. At large g, if we create a spin flip by scattering neutrons, it is usntable — the elementary excitations here are domain walls, and so a spin flip decays into a pair of domain walls which then slosh around. (Recall that the domain wall operator is non-local and is made up of a string of $\boldsymbol{\sigma}^x$s, so can't be accessed by $\boldsymbol{\sigma}_j^z$). So all we see is a continuous mess at small g – a multiparticle continuum of domain walls with no sharp peaks.

This behavior has been observed in a system well-described by the TFIM, the material $CoNb_2O_6$ (cobalt niobate). This material is a quasi-one-dimensional compound: it can be approximated as a collection of decoupled chains of spins with nearest- neighbor ferromagnetic Ising interactions. The transverse field is produced by, well, applying an external transverse field, allowing g to be varied. All the data and figures in this section have been reproduced from R. Coldea *et. al.*, *Science* **327**, 177 (2010).

First, we see that the system displays a phase transition as a function of transverse field. This is accomplished by tracking how the magnetic Bragg peak intensity (measured via elastic neutron scattering) evolves with a transverse field (Fig. 31).

Next, the system is analyzed via inelastic neutron scattering, which measures $S(q,\omega)$. (Fig. 32). As we have argued, this should show distinct behavior in the paramagnetic and ferromagnetic regimes. Thus, we experimentally confirm that spin excitations change character from pairs of domain walls in the ordered phase to spin-flips in the paramagnetic phase.

So far, this is in accord with our picture of the Ising model. If we turn off the field, it turns out that there is a weak effect owing to the fact that model is not 'truly' 1D: the coupling between the zigzag Co^{2+} chains generates a weak, longitudinal field (i.e., a field parallel to the direction of spin alignment.) One can further examine the model by examining the behavior in *zero external field* as a function of temperature. Note that a longitudinal field actually acts as a linear confining potential for domain walls. Intuitively, a field pointing along the $+z$ direction prefers domains of $+z$. A pair of domain walls in the $+$ phase are separated by a $-$ region, with costs an energy growing linearly with its length (Fig. 33, left top.).

This leads to the formation of domain wall bound states; this is a simple quantum-mechanical model that can be solved quite easily, leading to the theoretical predictions in Fig. 33 for temperatures above (C) and below (D) the experimentally determined transition temperature into the ordered phase $T_c \sim 2.95\,\mathrm{K}$ (recall that the Ising model has a thermally tuned transition only for $d \geq 2$; by including the weak interchain couplings, this is allowed as the model is effectively three-dimensional!). The experimental data Fig. 33(A,B) are in good agreement with the theory, including the theoretical simulation of the neutron result Fig. 33(F). Reading off the 'spectrum' of the bound state energies (called 'masses' in accord with the field-theoretic terminology) we see excellent match to experiment Fig. 33(F, and left bottom).

Finally, one can study the fine structure of the inelastic scattering data near the critical point. Owing to the small longitudinal field, the critical point is no longer gapless (Fig. 34, left top), and instead has a massive spectrum. It turns out that there is an emergent E_8 symmetry that allows one to solve for the 'particle' spectrum (the set of discrete excitations). It turns out there are 8 'particles' (i.e., distinct collective excitation modes, see Fig. 34, left bottom). While in principle these could all give sharp peaks in $S(q,\omega)$, it turns out that the mass of the third particle is just a bit bigger than twice the mass of the second, $m_3 \gtrsim 2m_1$; so particles of mass $M \geq m_3$ are unstable to decay into the $2m_1$ continuum. As a result, we can only see m_1, m_2 as sharp neutron peaks. The theory predicts, somewhat amazingly, that *at the critical point* the ration m_2/m_1 is precisely the *golden mean*, $\varphi = \frac{1+\sqrt{5}}{2}$. As shown in Fig. 34(A-D), the experiment is in remarkable agreement with theory. In fact it turns out that to match fine details of the structure, one needs to do a bit more work, and really the model has some slightly more subtle hidden symmetry that makes the story very slightly different from the simplified Ising picture we have given above [M. Fava *et al*, *Proc. Nat. Acad. Sci.* USA **117**, 25219 (2020)].

---

[a] The first equality can be a bit tricky but here's the easiest way to see it: recall that $\boldsymbol{\sigma}^z_q = \frac{1}{\sqrt{N}} \sum_r e^{-iqr} \sigma^z_r$. Then, we have

$$\langle 0|\boldsymbol{\sigma}^z_q \boldsymbol{\sigma}^z_{-q}|0\rangle = \frac{1}{N} \sum_{r_1,r_2} e^{-iq(r_1-r_2)} \langle 0|\boldsymbol{\sigma}^z_{r_1} \boldsymbol{\sigma}^z_{r_2}|0\rangle = \frac{1}{N} \sum_{r_1,r_2} e^{-iq(r_1-r_2)} C(r_1-r_2) = \sum_r e^{-iqr} C(r) = S(q) \tag{776}$$

where we used translational invariance and relabeled the summation.

[b] Observe that $S(q) = \int \frac{d\omega}{2\pi} S(q,\omega)$, **not** $S(q,0)$.

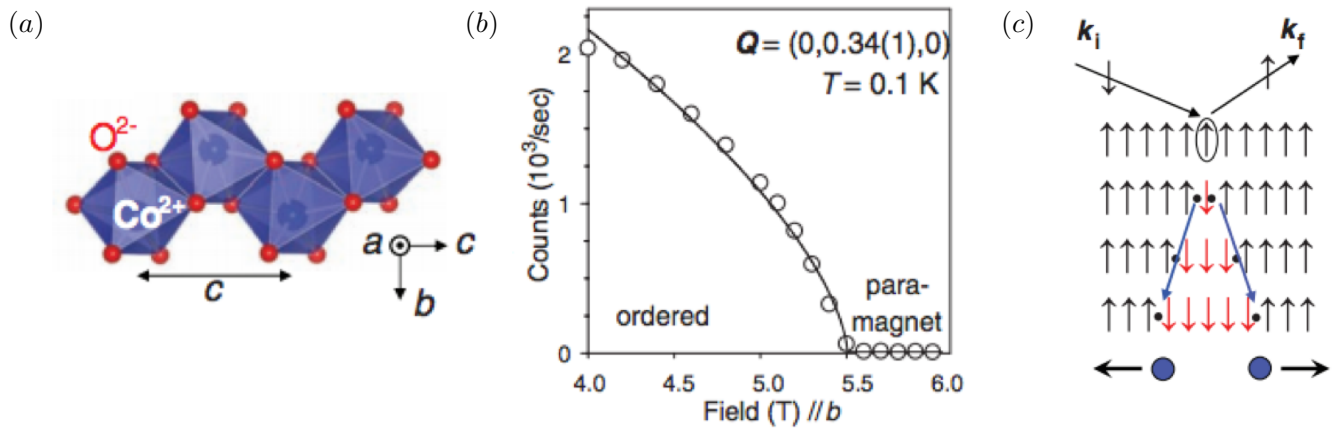


Figure 31: **Ising Symmetry in $CoNb_2O_6$.** (a) Lattice structure of $CoNb_2O_6$.; a single zig-zag chain of Co^{2+} is well-described by H_{TFIM} (b) Elastic neutron scattering data allows us to track the magnetization via Bragg peaks, showing classic TFIM behavior (c) Inelastic neutron scattering allows us to excite single spin flips and study their dynamics.

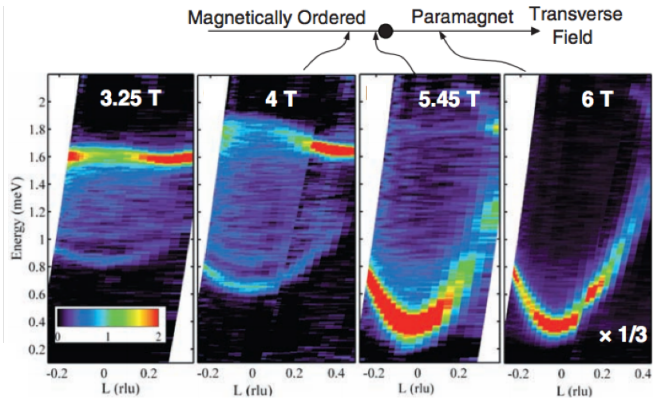


Figure 32: **Inelastic Structure Factor $S(q,\omega)$ across the Ising transition** as measured by neutron scattering. Note the sharp structure seen in a 6 T field, reflecting the fact that the spin flip is a sharp excitation in the paramagnetic phase; in the ferromagnet, $S(q\omega)$ is incoherent, as spin flips decay into pairs of domain walls.

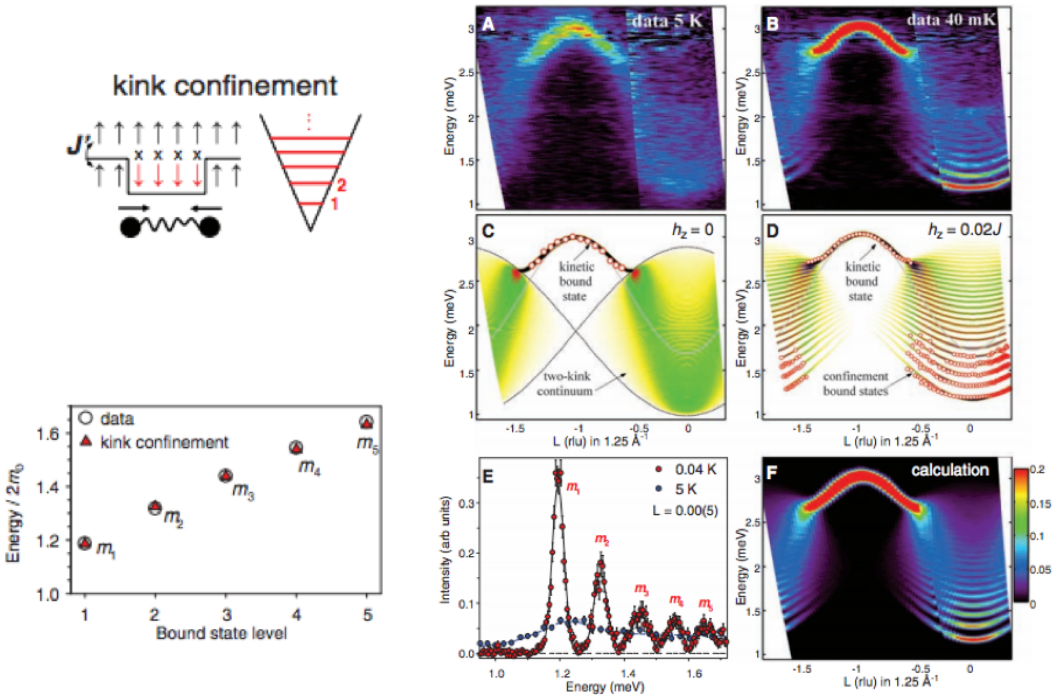


Figure 33: **Domain-wall bound states in the absence of external transverse field.** Sketch of the bound state formation ('kink confinement') and theoretical spectrum of bound states is shown alongside experimental data.

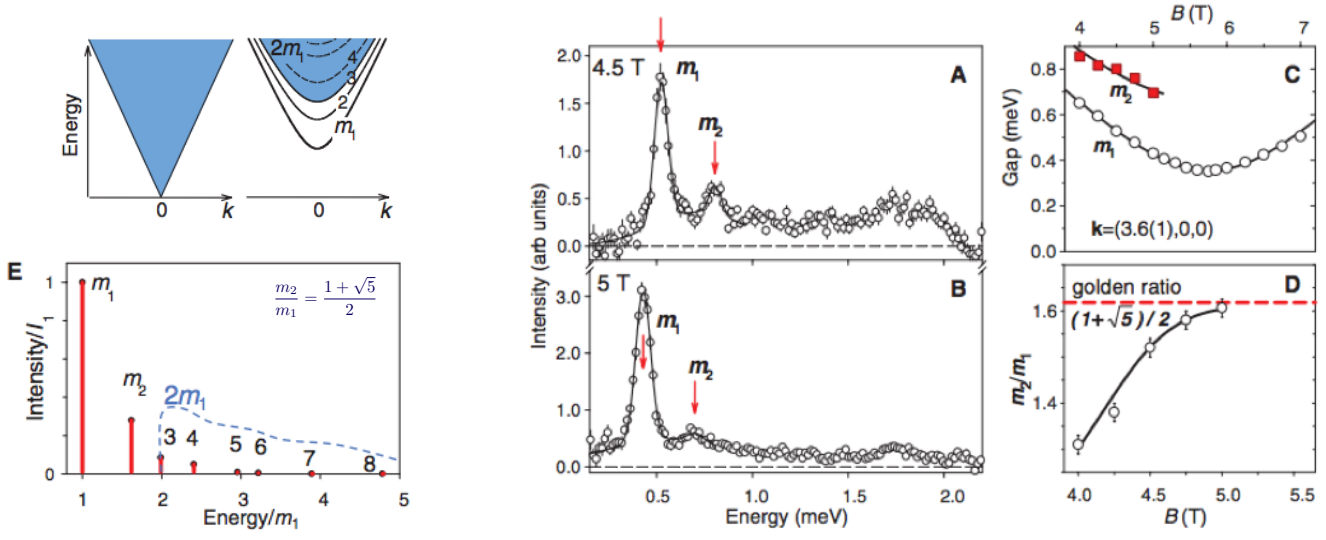


Figure 34: **Emergent E_8 Symmetry.** Particle spectrum computed from theory and corresponding experimental signatures of the emergent symmetry in the Ising model with both longitudinal and transverse fields.