Technische Universität München
Max-Planck-Institut für Quantenoptik

Quantum Variational Methods for Gaussian States and Beyond

Tommaso Guaita

Vollständiger Abdruck der von der Fakultät für Physik
der Technischen Universität München zur Erlangung
des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Jonathan Finley

Prüfer der Dissertation: Hon-Prof. Dr. J. Ignacio Cirac
Prof. Dr. Michael Knap

Die Dissertation wurde am 12.4.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 13.6.2022 angenommen.

Abstract

This thesis presents a complete, self-consistent and rigorous theory of variational methods for quantum many-body physics, with a specific focus on methods related to Gaussian states. This is achieved by describing in detail the differential geometry and group theory structures that lie behind these methods.

Firstly, a geometric theory of variational methods is outlined. Understanding variational ansätze as differentiable manifolds allows us to construct consistent variational principles to compute ground state approximations, spectral properties and time evolution for relevant Hamiltonians.

Then, the application of these methods to Gaussian *ansätze* is discussed, together with its limitations. To overcome the latter, a method is devised to construct useful variational families that go beyond Gaussian states, while preserving their favourable computational properties.

Finally, concrete examples are presented of the results that can be obtained with these methods for specific physical problems. In particular, the Bose-Hubbard and Sherrington-Kirkpatrick models are considered, where Gaussian methods or generalisations of Gaussian methods give excellent descriptions of the equilibrium and, in some cases, non-equilibrium physics.

Zusammenfassung

In dieser Dissertation wird eine vollständige, konsistente und rigorose Theorie der Variationsmethoden für die Quantenvielkörperphysik vorgestellt, mit einem besonderen Schwerpunkt auf Methoden im Zusammenhang mit Gaußschen Zuständen. Dies wird erreicht, indem die differentialgeometrischen und gruppentheoretischen Strukturen, die diesen Methoden zugrunde liegen, im Detail beschrieben werden.

Zunächst wird eine geometrische Theorie der Variationsmethoden dargestellt. Wenn man Variationsansätze als differenzierbare Mannigfaltigkeiten versteht, kann man konsistente Variationsprinzipien konstruieren, um Grundzustandsannäherungen, spektrale Eigenschaften und die zeitliche Entwicklung für relevante Hamiltonoperatoren zu berechnen.

Anschließend wird die Anwendung dieser Methoden auf Gaußsche Ansätze sowie ihre Grenzen diskutiert. Um letztere zu überwinden, wird eine Methode entwickelt, um nützliche Variationsfamilien zu konstruieren, die über Gaußsche Zustände hinausgehen und dabei ihre günstigen Berechnungseigenschaften beibehalten.

Schließlich werden konkrete Beispiele für die Ergebnisse vorgestellt, die mit diesen Methoden für spezifische physikalische Probleme erzielt werden können. Insbesondere werden die Bose-Hubbard- und Sherrington-Kirkpatrick-Modelle betrachtet, bei denen Gauß-Methoden oder Verallgemeinerungen von Gauß-Methoden ausgezeichnete Beschreibungen der Gleichgewichts- und in einigen Fällen der Nicht-Gleichgewichts-Physik liefern.

List of Publications

Articles included in this thesis:

1. T. Guaita, L. Hackl, T. Shi, C. Hubig, E. Demler, J.I. Cirac,
   Gaussian time-dependent variational principle for the Bose-Hubbard model,
   Physical Review B **100**, 094529 (2019)

2. L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, J.I. Cirac,
   Geometry of variational methods: dynamics of closed quantum systems,
   SciPost Physics 9, 48 (2020)

3. T. Guaita, L. Hackl, T. Shi, E. Demler, J.I. Cirac,
   Generalization of group-theoretic coherent states for variational calculations,
   Physical Review Research **3**, 023090 (2021)

4. P.M. Schindler*, T. Guaita*, T. Shi, E. Demler, J.I. Cirac,
   A variational ansatz for the ground state of the Quantum Sherrington-Kirkpatrick model, arXiv:2204.02923 (2022)
   * Co-first authors

Other articles not included in this thesis:

5. C. Jones, T. Guaita, A. Bassi,
   Impossibility of extending the Ghirardi-Rimini-Weber model to relativistic particles,
   Physical Review A **103**, 042216 (2021)

6. Z. Gong, T. Guaita,
   Topology of quantum Gaussian states and operations,
   arXiv:2106.05044 (2021)

Contents

Chapter 1

Introduction

1.1 Motivation

Calculating exactly the properties of a quantum many body system using classical computational resources is in general an intractable task [1]. Indeed, consider a system composed of N individual components (atoms, spins, lattice sites, available modes of a bosonic or fermionic field...). The corresponding many body Hilbert space will have dimension d^N, where d is the dimension associated to each local degree of freedom[1]. The exponential scaling of this dimension with respect to N means that it is sufficient to consider a relatively modest number of components to generate a problem unapproachable for any available or imaginable classical computer.

Just storing in a computer's memory a representation of a quantum state (represented as a Hilbert space vector) would require storing a d^N-dimensional complex vector. Assuming single floating point precision, this means that storing the state of a system of 40 elements of spin 1/2 would require a few terabytes of memory space. Storing the state of 65 spins would require hundreds of exabytes, close to the total storage space available to humanity as estimated in 2007 [2]. A few hundreds of spins would reach the limit of the estimated available storage capacity of the universe, $\sim 10^{90}$ bits of information [3].

Typically, on top of storing a single state we also want to consider the system's Hamiltonian $\hat{H}$. This then corresponds to a $d^N \times d^N$ matrix. Investigating the energy spectrum of the system, for example, requires diagonalising such a matrix. Simulating the time evolution of the system means constructing the unitary time

[1] In the case of bosonic components the local dimension is itself infinite. However in models of practical interest it is usually possible to impose the constraint of a maximum occupation number for each mode, thus reducing the problem to a finite dimension, without incurring in significant errors.

evolution operator $\hat{U} = e^{\mathrm{i}\hat{H}t}$ and applying it to an initial state vector. These are tasks that not only need an exponential memory space, but also require a number of elementary operations that scales exponentially in N.

All classical methods for the analysis of quantum many body systems must therefore rely on some level of approximation or on some intuition – more or less model-specific – which sufficiently simplifies the problem.

One may have the expectation that the situation will be changed by progress in quantum computation. Indeed, already in the 1980s, these observations led Feynman to famously state that it may be advisable to attempt to use other quantum mechanical systems – rather than classical ones – to simulate the quantum system of interest [4]. This and other similar statements are often credited with having provided the initial motivations for the ongoing quest of the development of a functioning quantum computer. For sure, it appears that the study of the physics of quantum many body systems represents an ideal playground to observe the so-called *quantum advantage* [5]. In other words, it may be possible in the near term future to build quantum devices that can clearly outperform the best available classical computers on the solution of some quantum many body problems [6].

Not everything will be automatically solved, however, even for the lucky owner of a quantum computer. It has been proven that finding the ground state energy of a local Hamiltonian is in the worst case a hard task, even when using a quantum computer [7]. Nonetheless, a wide range of heuristic quantum algorithms has been developed in the hope of solving the ground state problem in the average case, or at least in some interesting cases, including adiabatic state preparation [8], the Quantum Approximate Optimization Algorithm (QAOA) [9], variational quantum eigensolvers and other hybrid quantum-classical methods [10–12].

In simulating the time evolution of quantum many body systems the superiority of quantum computers over classical ones seems instead more obvious. Taking easily preparable initial states and evolving them with simple local Hamiltonians for a fixed time produces highly entangled quantum states. Sampling accurately from the outcome distribution of local measurements on these states is believed to be a difficult problem for a classical computer [13]. It is no surprise then that the first attempts to demonstrate quantum advantage have focused on sampling from a state produced through a random quantum circuit, *i.e.*, a discretised quantum evolution [14, 15]. Although these results are certainly remarkable, the current error rates per gate are still too high to hope to simulate a physical evolution by discretising it into a quantum circuit (*digital* quantum simulation). In the shorter term more relevant results may come from *analog* quantum simulation, where a quantum system is engineered to emulate directly the Hamiltonian of interest [16].

Given this context, we believe that it is of great importance to continue to develop approximate *classical* methods for quantum many body physics, both for

the computation of ground states and for the simulation of time evolution. This is the case for at least two important reasons:

1. It is important to characterise carefully what results are achievable classically in order to have a better perspective of the potential of quantum computing. This will guide research in quantum computation towards the most useful directions and provide fundamental physical knowledge in those cases where quantum computation will not be of help in the short term.

2. In the development of quantum computers it will be fundamental to have reliable benchmarks to verify the performance of quantum algorithms. A quantum computer is a device specifically designed to output results difficult to reproduce classically. It is therefore intrinsically complex to verify whether the quantum computer is functioning as expected. It will be vital to have at least some non-trivial test cases where a reliable classical prediction is available to convince oneself of the correct running of the quantum device. High quality classical input may possibly also improve performance of quantum algorithms.

These reasons provide the main motivations why the substance of this thesis will be the characterisation, development and application of classical computational methods for quantum many body physics. We will focus on a class of classical methods known as *variational methods* and in particular on the ones that can be implemented using techniques derived from Gaussian states.

In the rest of this introduction we will briefly review the current status of classical computational methods in general. This will then allow us to put into context and summarise the more specific variational methods and introduce the new results which are discussed in detail in the main body of the thesis.

1.2 Classical computational methods in quantum many body physics: an overview

In the almost one hundred years that have passed since the formulation of quantum mechanics a huge variety of classical methods has been developed and applied to compute – analytically or numerically – the predictions of the theory. Giving a comprehensive review of all the available methods is thus an overwhelming task, and it is by no means the objective of this section. Instead, we would like to give here a general feeling for the physical insights that lie behind some of the most successful of these methods. This will allow us to put into a broader context the

specific methods that are the subject of this thesis, *i.e.*, variational methods, and discuss their usefulness, their limitations and the alternatives that exist to them.

Exactly solvable models

In some cases the exact solutions to many body models can be expressed analytically. Most of these models fall into the class of *free* or *non-interacting* models. The physical intuition behind such definition is that these models, although formally written as many body models, can be seen as composed of individual degrees of freedom that do not affect each other. The difficulty of solving these problems is therefore equivalent to the one of solving a single particle problem and there is no many body exponential scaling of the complexity. The individual degrees of freedom will satisfy either bosonic or fermionic statistics and the Hamiltonian of an arbitrary number of copies of them can be written in second-quantised form as

$$\hat{H} = \sum_{nm} \hat{a}_n^\dagger \, h_{nm} \, \hat{a}_m \, . \tag{1.1}$$

Here $\hat{a}_n^\dagger$, $\hat{a}_n$ are the creation and annihilation operators of an orthonormal set of modes that span the single particle Hilbert space and h_{nm} is the single particle Hamiltonian matrix on such space. The operators $\hat{a}_n^\dagger$, $\hat{a}_n$ satisfy canonical bosonic ($[\hat{a}_n^\dagger, \hat{a}_m] = \delta_{nm}$) or fermionic ($\{\hat{a}_n^\dagger, \hat{a}_m\} = \delta_{nm}$) (anti)commutation relations.

A very powerful tool to describe the eigenstates of such a Hamiltonian is the theory of bosonic and fermionic Gaussian states. This language allows to encode the simple non-interacting nature of the states without giving up on the many body, second-quantised formalism. This method also allows to slightly enlarge the class of exactly solvable Hamiltonians to models with pairing terms, *i.e.*,

$$\hat{H} = \sum_{nm} \left(\hat{a}_n^\dagger \, h_{nm} \, \hat{a}_m + \hat{a}_n^\dagger \, \Delta_{nm} \, \hat{a}_m^\dagger + h.c. \right) \, . \tag{1.2}$$

Although this Hamiltonian cannot be written as a first-quantised single particle Hamiltonian (it is not particle number preserving), its ground state is nonetheless a Gaussian state.

Furthermore, there exist systems, not necessarily containing physical fermions or bosons, that can be mapped, through a suitable unitary transformation, to a bosonic/fermionic model of the form 1.2, thus making them solvable. For instance, the Jordan-Wigner transformation maps some one dimensional spin models, such as the Ising model, to a free fermion Hamiltonian.

On top of this, also among *interacting* systems there exist classes of models for which exact solutions exist. In 1D for some interacting models an exact solution can be constructed using *Bethe Ansatz* techniques [17], although it is not fully clear

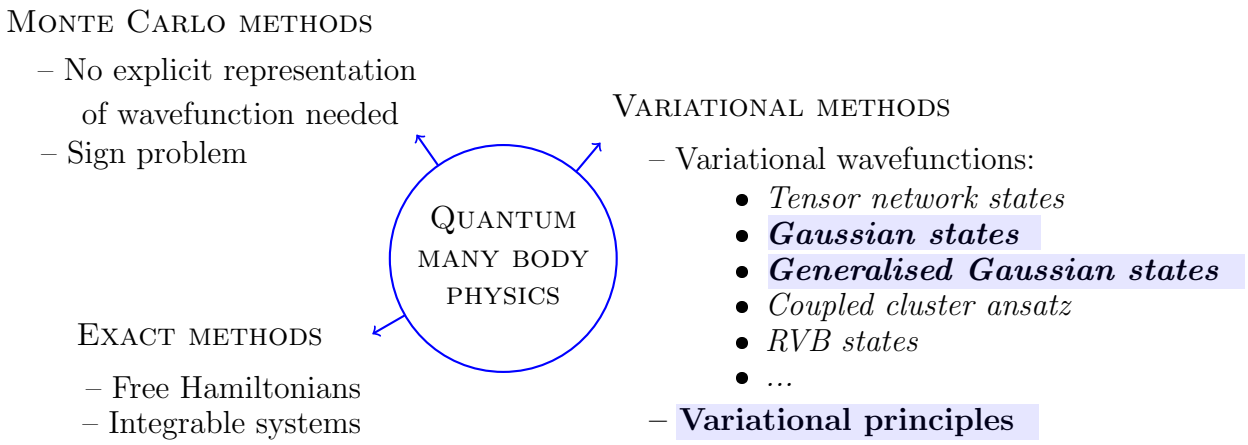


Figure 1.1: Map of computational methods for many body physics. The topics that will be considered in this thesis are highlighted.

what features distinguish this measure zero set of *integrable models* from generic quantum models. In 2D, some Conformal Field Theories are exactly solvable [18].

For all other models, for which no exact solution is known, we have to accept some kind of approximation or limitation in our treatment.

Monte Carlo methods

One possible approach could be to observe that, if we accept that we will not be able to describe the wavefunction of our system exactly, then we may as well give up writing down the wavefunction at all. After all, quantum mechanics is a probabilistic theory and the wavefunction is not directly observable. If our computational algorithm can probabilistically sample from the space of possible outcomes of a given measurement with the same probability distribution as the one predicted by quantum theory, then we can claim to have effectively simulated the theory [4], even without having stored a full representation of the wavefunction. The question at this point is in which cases this can be done efficiently with classical resources.

The basic idea of Quantum Monte Carlo is to compute the quantum thermal expectation value

$$\langle \hat{A} \rangle = \mathrm{Tr} \hat{A}\, e^{\beta \hat{H}} \tag{1.3}$$

of some observable $\hat{A}$ at inverse temperature β by classical statistical sampling. To achieve this one needs first to recast the D-dimensional quantum average (1.3)

as a classical average in $D+1$ dimensions

$$\langle \hat{A} \rangle = \frac{\sum_C A(C) W(C)}{\sum_C W(C)}, \tag{1.4}$$

where the sums run over all possible configurations of some classical variables, for example configurations $C = \{S_{\mathbf{r}_i,\tau_j}\}$ of a set classical spin variables associated to points in discretised $(D+1)$-dimensional space-time. Here, $W(C)$ are some weights associated to each configuration. This mapping can be achieved in various ways, including expanding $e^{\beta \hat{H}}$ into a Taylor series or into a Trotter decomposition, or discretising a path integral formulation in the imaginary time direction.

One can now use a classical Monte Carlo algorithm to sample spin configurations with probabilities distributed like the weights $W(C)$ and estimate the average value of the observable function A. Two types of difficulties can arise in this kind of scheme:

1. It may be difficult to accurately sample configurations according to the probability distribution defined by W. For example, if the system has a large correlation length and one draws samples through local modifications of a previous configuration, one may have to propose a large number of such updates before one is accepted by the sampling algorithm. This would lead to having to propose a number of configurations that is exponentially large in the system size to compute $\langle \hat{A} \rangle$ to a desired precision.

2. The weights W may be negative and therefore may not be interpreted as a probability distribution. In this case one would have to rewrite (1.4) as

$$\langle \hat{A} \rangle = \frac{\sum_C A(C) s(C) \, |W(C)|}{\sum_C |W(C)|} \left[\frac{\sum_C s(C) \, |W(C)|}{\sum_C |W(C)|} \right]^{-1}, \tag{1.5}$$

   where $s(C) = \mathrm{sign}[W(C)]$, and sample according to the well-defined probability distribution given by $|W(C)|$. However, in some cases the *average sign* function in the square brackets will be exponentially small in the system size, so one will need again to sample exponentially many configurations to achieve the desired error on $\langle \hat{A} \rangle$. This is generally referred to as the *sign problem*.

The first difficulty may be thought of as a "technical" difficulty, in the sense that a plethora of methods has been developed to improve sampling efficiency in many situations [19, 20]. However, developing the appropriate algorithm to sample efficiently in a given model often requires a lot of fine-tuned handcrafting. The sign problem, on the other hand, may be a more "fundamental" problem, in the sense that it is believed to be intrinsic to the nature of many models involving

spins in frustrated geometries or an odd number of fermions. In these cases it is likely impossible to recast the problem in a way that is sign-problem-free. Indeed, it has been proven that a generic solution to the sign problem would violate some commonly accepted assumptions of complexity theory [21].

Ansatz-based methods

A complementary approach to the Monte Carlo one is to insist that we do want to be able to write down some description of the state of the system, and not just sample from it. This can in general give us some deeper physical insights into the nature of the solution that we find and might be fundamental to treat those problems where Monte Carlo methods fail because of the sign problem or because of difficulties in sampling from complicated distributions.

As writing down the full wavefunction for a large enough system is impossible, we have to find ways to compress the information contained in the state. In other words, we have to be able to give an approximate description of the state which depends on sufficiently few parameters, that we can efficiently save and manipulate.

The simplest approach one can take is to focus on a parametrised *ansatz* in Hilbert space. That is, essentially a function that takes a set of M parameters, which we will collectively indicate as $x = (x_1, x_2, \ldots, x_M)$, and maps it to a state $|\psi(x)\rangle \in \mathcal{H}$. The image of this function defines a subset $\mathcal{M}$ of the many body Hilbert space which can be efficiently addressed by simply manipulating the parameters x.

For example, if we are interested in approximating the ground state of a Hamiltonian $\hat{H}$, we can compute the energy function

$$E(x) = \frac{\langle\psi(x)|\hat{H}|\psi(x)\rangle}{\langle\psi(x)|\psi(x)\rangle}\,. \tag{1.6}$$

We shall then try to find the optimal parameter set x_0 for which the function $E(x)$ attains its minimum value E_0. According to the Rayleigh-Ritz variational principle [22], the value E_0 will give an upper bound to the ground state energy of $\hat{H}$ and the state $|\psi(x_0)\rangle$ can be seen as an approximation of the ground state. The closer this state is to the actual ground state, the tighter the bound on the energy will be.

If, instead, we are interested in time evolution, we can similarly define methods – often referred to as *Time Dependent Variational Principles (TDVP)* – which seek to identify, through a suitable optimisation scheme, a trajectory $x(t)$ within the parameter space such that $|\psi(x(t))\rangle$ best approximates the real time evolution of a given initial state.

We will generically refer to this kind of methods as *Variational Methods*. In this thesis we will discuss variational methods in detail, starting from the next section

where we will give a more precise definition and an overview of the available techniques. Already from what we have said here, however, it should be clear that the quality of their predictions will depend crucially on the choice of the set $\mathcal{M}$ of states $|\psi(x)\rangle$. To achieve computational efficiency, the set $\mathcal{M}$ has to be parametrised by a relatively small number of parameters, thus it will only cover a small corner of the high dimensional Hilbert space. Only if the states of interest – *e.g.*, the ground states or the time evolved states – are very close to precisely this corner, then the variational methods can give good results. In other words, the difficulty of variational methods is to use some physical intuition to choose states $|\psi(x)\rangle$ suitable for the problem faced.

Besides proper variational methods, there exists also other methods that rely on a similar *ansatz*-based philosophy. For instance, in quantum chemistry the so-called *Coupled Cluster* methods [23] also rely on fitting a certain wavefunction *ansatz* to the problem's ground state. However, this is not done by minimising the the energy function (1.6), which in this case is too difficult to compute, thus it does not lead to variational bounds on the ground state energy.

In *Density Functional Theory (DFT)* [24], the information contained in a many body electronic wavefunction is compressed by only considering the electron density function

$$n(\mathbf{r}) = \int d\mathbf{r}_2 \cdots d\mathbf{r}_2 |\Psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)|^2 . \tag{1.7}$$

It can be proven that the energy of an atomic system only depends on this function, which then in principle contains all the relevant information [25]. However, the precise analytical form of this dependence is unknown. DFT therefore relies on developing approximate functionals which estimate the energy of the state given its density function. Thus, in some sense, the *ansatz* is at the level of the energy function rather than at the level of the state.

Finally, there is also a large region of overlap between *ansatz*-based methods and Monte Carlo methods. In *Variational Monte Carlo* methods, wavefunction *ansätze* are considered which are too complex to evaluate the energy function (1.6) in a closed form, similarly to the case of Coupled Cluster methods. However, due to the positive definiteness of their wavefunctions, it is still possible to estimate $E(x)$ and perform optimisations thanks to stochastic Monte Carlo methods. These variational *ansätze* include Jastrow-Slater wavefunctions for correlated electron systems [26], Resonating Valence Bond (RVB) [27] states for high temperature superconductors and Restricted Boltzmann Machines (RBM) [28] for spin systems.

1.3 Variational Methods in Quantum Mechanics

Of all the methods outlined in the previous section, this thesis will focus on variational methods for pure states. These are methods that aim at providing a closed expression – albeit approximate – for the wavefunction of interest. They achieve this by assuming that these wavefunctions are well approximated by states of a set $\mathcal{M} \subset \mathcal{H}$. This leads to an advantage because $\mathcal{M}$ is described by a number M of parameters that is taken to be small compared to the dimension of the full Hilbert space $\mathcal{H}$.

It follows that, when we want to apply variational methods to quantum mechanical problems, two separate issues are of fundamental importance. Firstly, how do we choose the set of states $\mathcal{M}$, which we will refer to as *variational states*, over which we perform the variational optimisations? Secondly, once this subset has been defined, what methods do we use to extract interesting information about the physical system while referring only to these states and not to the full, exponentially large, Hilbert space? What are the quantities that can be computed in this way? In this section we will introduce more in detail both these questions. Answering them in detail will be the core task of the first part of this thesis.

Variational states

A set $\mathcal{M}$ of parametrised variational states $|\psi(x)\rangle$, in order to be a useful set of variational states has to fulfil the following three general requirements [29]:

1. There exists an efficient algorithm to compute the expectation value $\langle\psi(x)|\hat{A}|\psi(x)\rangle$ of some observables $\hat{A}$ on any given variational state $|\psi(x)\rangle$. Efficient means that the computational cost of the algorithm scales at most polynomially with respect to the number M of variational parameters and that this number of parameters in turn scales polynomially with respect to the system size N.
   Usually, the most important observable is the Hamiltonian $\hat{H}$, but it may be useful to be able to compute also other observables relevant for the system.

2. There exists some reason to expect that at least some states in $\mathcal{M}$ have a large overlap with the *target* states of the variational problem. By target states we generally mean the ground state of the Hamiltonian or the states resulting from time-evolving under $\hat{H}$ a given initial state.

3. It is possible to find a solution to the variational optimisation problem. In other words, there should exist a procedure that efficiently finds the optimal

state in $\mathcal{M}$ that minimises the energy function $E(x)$ (in case of a ground state problem) or that solves the equations for $x(t)$ in case of a Time Dependent Variational Principle.

Finding classes of states that fulfil all three requirements is far from a trivial task. Nonetheless, over the decades some sets of variational states have found wide application, showcasing the power of variational methods. Let us illustrate how this can be the case with two such examples.

Perhaps some of the most impressive results in this field have been achieved by *Tensor Network States*. These comprise a variety of *ansätze* that represent a wavefunction using as building blocks some tensor quantities contracted into a specific network geometry. In particular, a one dimensional variant of Tensor Network State, namely *Matrix Product States (MPS)*, has affirmed itself as an unrivaled numerical method for $1D$ many body systems [30].

Example 1 (Matrix Product States)**.** *MPS give a very clear example of how the requirements above can be fully satisfied:*

1. *For a one dimensional system of N spin degrees of freedom, an MPS is parametrised by N tensors whose dimensions are fixed by the local spin dimension and a so-called bond dimension. The number of variational parameters thus scales linearly with N. Computing the expectation value of a sum of local observables can be achieved in a time also linear in N [31]. This satisfies point 1.*

2. *The intuition why MPS approximate so well the ground states of many relevant Hamiltonians comes from Quantum Information arguments. MPS with finite bond dimension[2] are able to approximate well states whose entanglement satisfies a so-called area law [32, 33]: that is, in one dimension, when the entanglement entropy across any bipartition of the system is bounded by a quantity independent of the bipartition [34]. It is proven that the ground state of any local gapped Hamiltonian will fulfil such an area law and can thus be represented by an MPS [35]. We thus have point 2.*

3. *Finally, there exists an algorithm, known for historical reasons as Density Matrix Renormalisation Group (DMRG), that allows to optimise an MPS for a given local Hamiltonian in a remarkably robust way [36]. This ensures point 3.*

[2]By finite bond dimension here we mean that it does not grow with the the total system size N. In principle, any state can be represented exactly as an MPS, provided that one takes an arbitrarily high bond dimension. What we care about here is that, in order to represent the states of interest, we are not forced to choose a bond dimension that grows exponentially as N gets larger, as this would violate the previous point on the polynomial scaling of the number of parameters.

These observations also shed light onto the limitations of MPS. Firstly, the power of MPS as a variational *ansatz* starts to decrease as soon as the target states can no longer be expected to satisfy the entanglement area law. In particular, this is the case for ground states of non-local or gapless Hamiltonians or for time-evolved states [37]. Secondly, the high numerical efficiency of MPS methods relies strongly on the one dimensional geometry of the system. It is indeed considerably harder to construct efficient algorithms for *Projected Entangled Pair States (PEPS)*, which can be understood the natural generalisation of MPS to higher dimensional systems, and which would be ideal for representing states satisfying the higher dimensional area laws [38].

Other variational *ansätze* based on tensor network principles have also been developed and share some of these favourable properties. For example, the *Multiscale Entanglement Renormalisation Ansatz (MERA)* [39] (thought for critical systems), tree tensor networks or generalisations of Tensor Network States for systems in continuous space [40, 41].

Besides Tensor Network States another highly successful class of variational families is the one based on bosonic and fermionic Gaussian states. We have already discussed in Section 1.2 how Gaussian states are the exact ground states of non interacting systems. However, it is also possible to see them as a set of states on which to apply variational methods.

Example 2 (Gaussian states)**.** *The reason why Gaussian states can be potentially very useful as a variational ansatz can again be understood in the terms discussed above:*

1. *A Gaussian state of a system of N bosonic or fermionic modes is parametrised by a number of parameters that scales as N^2. Computing expectation values of polynomials of creation and annihilation operators on Gaussian states is made highly efficient thanks to a well-known theorem by Wick [42].*

2. *Gaussian states are known to include exact ground states of non interacting systems. We may expect them to also do a good job at describing interacting systems where the interactions are weak enough.*

3. *Optimisation algorithms on the space of Gaussian states can be easily implemented [43].*

In this thesis, methods based on Gaussian states will be taken into particular consideration. We will introduce Gaussian states more detail in Section 1.4, exploring also the possibilities of going beyond them.

Variational principles

In the previous sections we have discussed several examples of strategies that can be applied once we have selected a reasonable set of variational states, for instance optimising the energy for an approximate ground state or applying a Time Dependent Variational Principle. But what is exactly the full range of *variational principles* that can be applied in these cases? In other words, what are the available abstract methods and principles that can be used, given any variational class of states, to compute physically relevant quantities?

Ideally, what we want to do is to perform any standard quantum mechanical calculation – evaluation of energy spectra and eigenstates, time evolution, computation of measurement outcome distributions *etc.* – restricting ourselves to considering as relevant only the states in our chosen variational class $\mathcal{M}$, escaping in this way the prohibitive dimensionality of the full space of allowed states $\mathcal{H}$. However, this objective immediately presents a difficulty. The Hilbert space $\mathcal{H}$ is a complex *linear* space and standard quantum mechanics is formulated in a way that makes explicit use of this linear structure. The chosen space $\mathcal{M}$ on the other hand, may not necessarily be a complex linear subspace of $\mathcal{H}$. In fact, the most promising choices of $\mathcal{M}$ discussed above (Tensor Network States, Gaussian states) are by no means linear spaces. The essence of variational methods can therefore be understood as formulating approximate versions of quantum mechanics on non-linear spaces $\mathcal{M}$ in such a way that they resemble as much as possible the true linear quantum mechanics of the linear space $\mathcal{H}$ in which $\mathcal{M}$ is embedded.

Having put things in this perspective, we realise that the geometric structure of $\mathcal{M}$ plays here a fundamental role: not only its linearity or non-linearity, but also its curvature, its tangent spaces, its intrinsically complex or real structures. In short, it has been realised that the language of differential geometry is an exceptional tool to better describe and understand variational methods. First, this language allows us to correctly describe the nature of the chosen set of states $\mathcal{M}$ and its relationship with $\mathcal{H}$ by viewing it as geometric manifold embedded in the metric space $\mathcal{H}$. Second, all variational methods can be formulated as the differential geometric equivalent of quantum mechanical concepts normally defined on the linear Hilbert space.

The first part of this thesis will be dedicated to following this perspective to all of its consequences, in the belief that it will best allow us to understand the full power and possible limitations of variational methods and how to best apply them. To conclude this section we will list the most important variational methods, highlighting their connection to differential geometry. Hopefully, this will give the reader a sufficient intuition of the potential benefits of this approach. In the main body of the thesis, we will explore each one in detail, deriving step-by-step their interpretation in the geometric language.

Ground state problem – The problem of finding the ground state of a given Hamiltonian operator $\hat{H}$ is normally formulated as finding the eigenstate of $\hat{H}$ with the lowest possible eigenvalue. Restricting ourselves to a non-linear space $\mathcal{M}$, where linear operators and spectral decompositions cease to be well-defined, the best possible reformulation consists of interpreting $\hat{H}$ as defining a scalar energy function on $\mathcal{M}$

$$E(\psi) = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle}\,, \quad \forall\,|\psi\rangle \in \mathcal{M}\,. \tag{1.8}$$

The ground state is then naturally interpreted as the point in $\mathcal{M}$ coinciding with the global minimum of E. Indeed, the Ritz principle [22] assures us that if we extend $\mathcal{M}$ to include the whole Hilbert space the two definitions will coincide. Finding the approximate ground state on $\mathcal{M}$ thus reduces to a problem of optimisation on a non-linear manifold. The preferred methods employed in this case are ones based on gradient descent. We will show that these methods can be improved by including information on the intrinsic geometry of $\mathcal{M}$, leading to what is sometimes referred to as *natural* gradient descent and which, from the more physical point of view, can be understood as projected imaginary time evolution.

Excitation spectrum – On top of the ground state, one is sometimes interested in the other eigenstates lying close above the ground state in the energy spectrum. These low energy excitations again can hardly be described in terms of the spectrum of a linear operator, if we restrict ourselves to only considering a non-linear manifold of states $\mathcal{M}$. However, something similar can be defined by looking at the small oscillations that are possible on the manifold close to its minimal energy state. In other words, a linear response theory can be defined on $\mathcal{M}$, allowing us to identify well-defined excitations that fulfil the properties of having low energy and lying close to the ground state on the manifold of considered states.

Time evolution – Finally, an important application of variational families is to compute the time evolution of an initial state chosen on $\mathcal{M}$. Here one faces two possible paths to defining an evolution equation – *i.e.*, a Schrödinger-like equation – for states on $\mathcal{M}$. One possibility is to look for a global description: a Lagrangian that is defined on the whole manifold $\mathcal{M}$ and which, through its Euler-Lagrange equations, governs motion on it. The second possible choice is to look at things from a more local point of view: at each infinitesimal time step of evolution the true linear Schrödinger equation will in general move our state out of the chosen manifold. What is the best way to approximate this motion while remaining in $\mathcal{M}$, *i.e.*, making the least error at each such step? These two approaches lead to two apparently different Time Dependent Variational Principles, which we will refer to as the *Lagrange* and the *McLachlan* variational principles respectively.

The issue of the equivalence or difference of the Lagrange and McLachlan vari-

ational principles has given rise to some debate in the past. However, it is now well understood that the answer depends crucially on the chosen set of states $\mathcal{M}$. We will show in this thesis how the relevant property of $\mathcal{M}$ in this case can again be best understood in terms of differential geometric structures. We will point out that the two Time Dependent Variational Principles are equivalent only if $\mathcal{M}$ is what in differential geometry is known as a *Kähler* manifold. This property is ultimately related to the extent to which $\mathcal{M}$ preserves the complex nature of the Hilbert space $\mathcal{H}$. We will also discuss what happens if this is not fulfilled and how to best apply the variational principles in that case.

1.4 Gaussian methods

Discussing variational methods from a general perspective, we have mentioned some examples of commonly used variational states. In the second part of this thesis we will focus more in detail on one specific class of variational states, namely Gaussian states and related constructions. We have already anticipated in Example 2 how all these states are exceptionally favourable to deal with computationally. We have further mentioned how Gaussian states are naturally well-suited to represent the states commonly occurring in weakly interacting quantum systems. These properties make them a natural and highly successful testing ground for the application of variational techniques.

There exists indeed a long list of widely applied variational methods that rely on Gaussian states. The Hartree-Fock method, fundamental in Quantum Chemistry, can be understood as the variational optimisation over the set of particle number preserving fermionic Gaussian states [44]. The celebrated Bardeen-Cooper-Schrieffer (BCS) *ansatz* for superconducting states is itself a subset of fermionic Gaussian states (in this case not particle number preserving) [45]. The Quantum Hall states and Topological Insulator states are all represented as fermionic Gaussian states [46]. In the study of Bose-Einstein Condensation (BEC) bosonic Gaussian states are widely applied. For example, the Bogoliubov theory for the ground state and excitations of such condensates relies heavily on the Gaussian state formalism [47]. The Gross-Pitaevskij equations describe the dynamics of these systems through a Time Dependent Variational Principle based on a Gaussian state *ansatz* [48, 49].

Many attempts have also been made to extend Gaussian methods or combine them with other techniques to enlarge their range of application to more strongly interacting systems. For example, linear combinations of fermionic Gaussian states are commonly used in chemistry, under the name of *Configuration Interaction (CI) method* [50], as well as in other contexts [51]. More recently, the combination of Gaussian states with some fixed canonical transformation has been shown to give

very good results in the description of many impurity problems [52–54].

In the second part of the thesis we will give a general overview of the Gaussian formalism and introduce a new construction that allows to generate useful variational families that extend Gaussian states.

Gaussian states

In the literature there exist a variety of approaches to defining and understanding bosonic and fermionic Gaussian states [55–59]. Often these approaches are influenced by the specific needs of the various communities that make use of them, such as Quantum Optics and Communications, Quantum Chemistry and Condensed Matter physics. In our analysis, we will instead try to keep a more abstract perspective, focusing on understanding the underlying structures that make Gaussian states as effective as they are.

From this point of view, it becomes clear that a fundamental property is the fact that Gaussian unitaries (*i.e.*, unitary operators that map Gaussian states into each other) provide a representation on Fock space of certain group theory concepts. It is these group structures, together with the related algebraic structures, that ultimately encode many of the useful properties of Gaussian states.

More precisely, it is possible to define groups of unitary operators $\mathcal{U}_{\mathrm{G}}(g)$, which we will refer to as Gaussian unitaries. In the case of bosons, they can be either displacement operators (generated by linear combinations of creation and annihilation operators) or squeezing operators (generated by quadratic combinations of creation and annihilation operators). In the fermionic case they are generated by quadratic combinations of creation and annihilation operators. Importantly, in all cases these sets of unitaries can be parametrised by an element g belonging to a certain matrix group. The unitaries $\mathcal{U}_{\mathrm{G}}$ then inherit (or rather *represent*) many of the properties of these groups. Indeed, the products of two different operators $\mathcal{U}_{\mathrm{G}}(g)$, their action on creation and annihilation operators and their effects on certain states can all be described and encoded in the group properties of the corresponding matrices g.

Gaussian states are then defined as all the states

$$|\psi_{\mathrm{G}}(g)\rangle = \mathcal{U}_{\mathrm{G}}(g)\,|0\rangle \tag{1.9}$$

that can be generated out of the vacuum by Gaussian unitaries. It becomes clear that exploiting the known group properties of g to manipulate $\mathcal{U}_{\mathrm{G}}(g)$ gives in turn the power of manipulating the states $|\psi_{\mathrm{G}}(g)\rangle$ equally efficiently. This lies at the core of the applicability of Gaussian states, as it allows to compute expectation values, overlaps, take derivatives with respect to parameters and much more.

There are, of course, also other ways to frame the special features of Gaussian states. But we believe that the group-theoretic perspective sketched here has one

further advantage. It allows to understand which other families of states, if any, share the same fundamental features of Gaussian states. Indeed, several other classes of useful states can be written as (1.9), replacing the Gaussian unitaries with the unitary representations of some other arbitrary group $\mathcal{G}$ and the Fock vacuum with a reference state $|\phi\rangle$ in some other Hilbert space. We obtain states of the form $\mathcal{U}(g)\,|\phi\rangle$, which then share many of the fundamental properties of Gaussian states. Namely, one can exploit the group structures of $\mathcal{U}(g)$ to facilitate computations.

This broad class of states $\mathcal{U}(g)\,|\phi\rangle$, for some unitary representation $\mathcal{U}(g)$ of a Lie group $\mathcal{G}$, are known as *group-theoretic coherent states* [60, 61] or *Gilmore-Perelemov coherent states* [62, 63]. They contain fermionic and bosonic Gaussian states as specific instances, but they also include several other interesting examples, such as atomic coherent states [64] or more complex states based on SU(N) groups [65, 66].

In general, it is possible to show that all group-theoretic coherent states can be potentially successful variational manifolds. If the system's relevant observables can be written in terms of the operators $\hat{Z}_i$ representing the Lie algebra $\mathfrak{g}$ associated to $\mathcal{G}$, then expectation values can be computed efficiently. The differential geometric structures of the variational manifold can be also evaluated handily. Under some simple to verify conditions, the states $\mathcal{U}(g)\,|\phi\rangle$ form a Kähler manifold, which we have anticipated above makes variational methods particularly well-behaved.

Beyond Gaussian states

Once we have seen how the group-theoretic structures of Gaussian and coherent states greatly facilitate their applications as variational families, it comes natural to ask: are we exploiting these structures to their fullest? Can the group-theoretic structures be exploited to manipulate more general families of states that go beyond the purely Gaussian or coherent paradigm?

This is a question of great importance. Indeed, although Gaussian states have favourable computational properties, their ability to represent physical systems is limited. By their nature, Gaussian states are only suited to describe systems near their non interacting point or close to some classical limit. Although Gaussian states do contain entanglement [67], the quantum correlations they can express are constrained: by construction, they can only exhibit correlations that satisfy Wick's theorem. Some of the other examples of group-theoretic coherent states suffer from even lower expressivity. For instance, atomic coherent states are essentially product states.

Families of states that can go beyond these limitations would therefore greatly increase the available expressive power of variational methods. The difficulty of

this search lies in balancing the need for an expressivity going beyond the Gaussian one, with the requirement that classical computations with the resulting states should be efficient, in order to obtain a usable variational family. Some progress has been made in these directions. Some generalisations of bosonic and fermionic Gaussian states were introduced in reference [52], where their computational properties were proven to be favourable, although their fully variational applications have been so far limited. Similarly, there exist examples of sets of states, such as spin squeezed states [68] or weighted graph states [29], that generalise other classes of group-theoretic coherent states.

In one of the main contributions of this thesis we will show how these generalisations can be understood as stemming directly from the group-theoretic framework introduced above. Indeed, we will introduce the concept of *generalised group-theoretic coherent states*. That is, we present a construction such that, for every existing class of group-theoretic coherent states, one can define a generalised version which contains extended non-Gaussian correlations, while maintaining favourable computational properties. This formalism allows us to understand in a unified way what are all the generalisation that can be constructed. But it also sheds light on where these states draw their computational powers from and in what sense they exploit to the fullest the group-theoretic structures present in the Hilbert space.

Given a set of group-theoretic coherent state defined as $\mathcal{U}(g)\,|\phi\rangle$ for g in the Lie group $\mathcal{G}$, the corresponding generalised states have the form

$$|\psi(g, g', M)\rangle = \mathcal{U}(g')\mathcal{V}(M)\mathcal{U}(g)\,|\phi\rangle\ . \tag{1.10}$$

We have thus added two further parametrised unitaries compared to the plain coherent states. One is a second group unitary $\mathcal{U}(g')$. The other is a new special unitary that takes the form

$$\mathcal{V}(M) = \exp\left(\frac{\mathrm{i}}{2} M^{ab} \hat{H}_a \hat{H}_b\right)\ , \tag{1.11}$$

for any real symmetric matrix M. Here, $\hat{H}_a$, for $a = 1, \ldots, \ell$, are a special set of algebra operators known as *Cartan operators*. They are a commuting subset of the operators $\hat{Z}_i$ which represent on $\mathcal{H}$ the Lie algebra $\mathfrak{g}$ associated to $\mathcal{G}$. Concretely they might be the particle number operators $\hat{n} = \hat{a}^\dagger \hat{a}$ in case of bosonic or fermionic Gaussian states or the Pauli operators $\hat{\sigma}^z$ for spin systems. The unitary $\mathcal{V}$ is responsible for introducing on the states (1.10) a much more elaborate entanglement and correlation structure than the one of plain group-theoretic coherent states.

Furthermore, there always exists a procedure to compute expectation values of operators $\hat{\mathcal{O}}$ on the states $|\psi(g, g', M)\rangle$, provided that $\hat{\mathcal{O}}$ can be written as a low degree polynomial in the algebra operators $\hat{Z}_i$. This procedure may in some

cases appear elaborate, but ultimately relies only on the algebraic structures of the problem. In fact, all necessary operations are performed in terms of objects (matrices, vectors...) whose dimension is at most the one of the group $\mathcal{G}$. In most physical situations, the relevant Lie groups have dimensions that scale polynomially with the systems size N, thus guaranteeing the computational efficiency of our construction.

The group-theoretic formalism makes it also simpler to see in what sense the construction (1.10) is optimal. Reasonable attempts to go beyond this structure will critically impact the efficiency of calculations: the introduction of further unitaries in the definition of $|\psi(g, g', M)\rangle$, the extension of $\mathcal{V}(M)$ to include operators from outside the Cartan subalgebra, the introduction in (1.11) of cubic or higher order terms in $\hat{H}_a$ will all break some step of the procedure for computing expectation values. We can therefore feel convinced we are exploiting the structures of group-theoretic coherent states to the fullest to define the most expressive variational manifold possible with the mathematical tools we have available.

In conclusion, we have seen that the Gaussian formalism gives us tools that can greatly facilitate the computations necessary for variational methods. Indeed, simple Gaussian states are of utmost importance in many long-standing computational methods. However, the range of applicability of these methods is often limited by a certain lack of expressivity of Gaussian states. And yet, it is possible to construct extensions of Gaussian states which exploit the same favourable computational properties while presenting an increased expressivity. There is therefore hope that Gaussian-based methods and extensions will establish themselves as an important variational approach that can complement other highly successful methods, such as Tensor Network based methods, filling the gaps where these other methods are weakest.

1.5 Outline of the Thesis

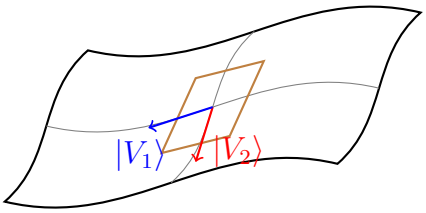
Let us now briefly summarise the structure of this thesis. It is divided into three main parts. In the first part (Chapters 2 and 3) we will present general results about variational methods. In the second part (Chapters 4 and 5) we will then focus on more specifically on Gaussian methods and generalisations thereof. In the third part (Chapters 6 and 7) we will finally present some applications of the methods discussed previously to concrete examples of physical models.

More specifically, in Chapter 2, we will define the notion of variational manifold of states. We will discuss what makes such a manifold well-defined and useful for variational computations. We will especially explain how it can be described from the point of view of differential geometry. We will introduce the interesting geometric structures that can be defined on such a manifold and discuss their

Part I: Variational methods

Section 2: Variational states as differential manifolds

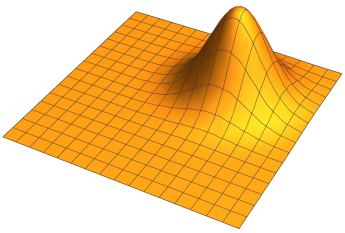
Section 3: Variational principles as geometric methods



Part II: Gaussian methods

Section 4: Gaussian states and coherent states

Section 5: Extensions of Gaussian states



Part III: Applications

Section 6: Bose-Hubbard model

Section 7: Sherrington -Kirkpatrick model

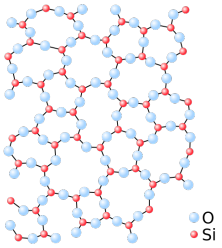


Figure 1.2: Strucure of the thesis, with its three main parts.

meaning from the perspective of variational methods. We will finally establish the notion of Kähler manifold, which plays a key role in time dependent variational methods.

In Chapter 3, we will then review all the available variational methods which can be implemented once one considers a manifold like the ones defined in the previous chapter. We will discuss real time evolution, the computation of excitation spectra, linear response theory and the use of imaginary time evolution to optimise energy functions. In all these cases our focus will be on deriving results that allow to fully understand the interplay between these methods and the geometry of the considered manifolds. This concludes the first part on abstract variational methods.

In Chapter 4 we will then shift our focus on a more specific class of variational manifolds, namely Gaussian manifolds. We will introduce the general notion of group-theoretic coherent states and show how Gaussian states can be understood as a specific example of it. We will present and discuss separately bosonic and

fermionic Gaussian states, highlighting how their key structures can be understood from a group theory perspective. We will also mention some other relevant examples of group-theoretic coherent states, namely spin coherent states. In Chapter 5, we will move on to consider possible generalisations of the Gaussian and coherent state concepts. We will in particular introduce *generalised* group-theoretic coherent states and show how they fulfil the criterion of efficient classical variational computations.

In Chapter 6 we will introduce the first example of an application of variational methods to many body physics problems. We will consider the Bose-Hubbard model and analyse it using bosonic Gaussian states as a variational manifold. We will show that, thanks to all the techniques discussed previously, even this relatively simple manifold can be used to describe many non-trivial features of the model in its superfluid phase. In Chapter 7, we will then present an example where a more elaborate *ansatz* is used, belonging to the family of generalised group-theoretic coherent states. We will discuss the Quantum Sherrington-Kirkpatrick model, showing how the manifold of generalised spin coherent states gives a good variational description of its ground state properties.

Each chapter will start with an *Overview* section where we will summarise the main results presented the chapter and explain the structure of the chapter itself. The appendix sections A.1–A.5 contain some extra material that completes the discussion of the main chapters, but is not necessary for a clear understanding of the thesis.

Finally, let us mention a couple of notation conventions which might facilitate the reader's appreciation of the mathematical formulae. All objects marked with a hat $\hat{}$ are Hilbert space operators. All objects without a hat should be understood as *c-numbers*. We will use objects with indices (x^μ, M_{ij}...) to indicate the entries of vectors, matrices *etc.* We will use the following conventions to help the reader navigate through them. Greek indices μ, ν, ρ... will label variational parameters and latin indices m, n, l... will indicate lattice sites, modes or individual degrees of freedom of a many body system. In contexts where we discuss notions of group theory we will indicate Lie groups as $\mathcal{G}$, SU, Sp... and the corresponding Lie algebras as $\mathfrak{g}$, $\mathfrak{su}$, $\mathfrak{sp}$... We will label algebra elements by the indices i, j, k... and Cartan subalgebra elements by a, b, c... Limited to Chapter 6, we will use the indices k, p, q... to label momentum values in the reciprocal lattice. With $M^\intercal$ we will indicate the transpose of a matrix M.

Part I

Chapter 2

Variational states and their geometry

In this chapter we will define the most general notion of variational states: a parametrised subset of a many-body Hilbert space. We will introduce some simple and yet very powerful geometric structures that can be constructed for any such set. Two different types of geometries can appear, namely *Kähler* and *non-Kähler* geometries. This, as we will see in later chapters, plays a key role in defining variational principles.

Most of the material presented in this chapter, as well as in the subsequent chapter 3, was published in reference [69]:
L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, J.I. Cirac,
Geometry of variational methods: dynamics of closed quantum systems,
SciPost Physics 9, 48 (2020), used under .

2.1 Overview

We consider variational sets of states as parametrised subsets of the Hilbert space $\mathcal{H}$. That is, we consider a set of states defined by the function $|\psi(x)\rangle$ where $x \in \mathbb{R}^M$ is a collection of M real parameters. We will indicate the set of all states that can be written this way as $\mathcal{M}$.

The objective of this chapter is to interpret $\mathcal{M}$ as a differential manifold of dimension M embedded in a Hilbert space of much higher dimension. Following this interpretation we will define the geometric notion of *tangent space* to the manifold. This is a linear space, that we indicate as $\mathcal{T}_\psi\mathcal{M}$, which can be defined for each point $|\psi\rangle$ of the manifold. Thanks to the tangent space we can define a range of further geometric structures, such as tangent projectors, metrics, vector fields and so on, which will be of great use in the next chapters to construct and

understand variational principles.

It is important to stress that we have defined the parameters x to be real. In principle, one could restrict oneself to variational families that admit a complex parametrisation, *i.e.*, defined by a function $|\psi(z)\rangle \in \mathcal{H}$ holomorphic in $z \in \mathbb{C}^M$ and thus independent of z^*. As we will see, this leads to enormous simplifications, as in the geometric language we are then dealing with so-called Kähler manifolds, which have very friendly properties. However, in general, we want to use real parametrisations, which cover the complex case (taking the real and imaginary part of z as independent real parameters), but apply to more general situations.

While in certain situations, it is easy to extend or map a real parametrisation to a complex one, this is not always the case. This applies, in particular, to parametrisations of the form

$$|\psi(x)\rangle = \mathcal{U}(x)\,|\phi\rangle\ , \tag{2.1}$$

where $|\phi\rangle$ is a suitably chosen reference state and $\mathcal{U}(x)$ is a unitary operator that depends on $x \in \mathbb{R}^M$. Such parametrisations are often used to describe various many body models [52–54, 70–73], and the fact that $\mathcal{U}(x)$ is unitary is crucial to compute physical properties efficiently. However, extending x analytically to complexify our parametrisations, would break the unitarity of $\mathcal{U}$ and often make computations inefficient, thereby limiting the applicability of the variational class. In conclusion, the focus on real parametrisations is of important physical interest.

The tangent space $\mathcal{T}_\psi\mathcal{M}$ is the space of all possible linear variations on the manifold around $|\psi(x)\rangle$. We can write them as $\sum_\mu \dot{x}^\mu\,\partial_\mu|\psi(x)\rangle$ and thus the tangent space can be defined as the span of the tangent vectors $|v_\mu\rangle = \partial_\mu\,|\psi(x)\rangle$. Importantly, as our parameters x are taken to be real to maintain generality, this span should only allow real coefficients. The tangent space should therefore be understood as a *real* linear space embedded in the complex Hilbert space $\mathcal{H}$.

This observation lies at the heart of all the structures that we will define for $\mathcal{M}$ in this chapter and plays a crucial role in defining the possible geometries of variational manifolds. To best explain this fact, we will take a step back in our exposition and begin the chapter by a more general discussion of real structures in Hilbert space and only gradually move to discuss the detail of variational manifolds.

In section 2.2 we will reinterpret the Hilbert space as a real linear space and discuss the specific structures that emerge if we take this perspective. This will allow us to introduce the concept of Kähler structures and real linear maps. In section 2.3 we then move on to discuss projective Hilbert space $\mathcal{P}(\mathcal{H})$, that is the space of state vectors defined up to arbitrary constant phase and normalisation factors. This is the space where physical quantum states live and it is the space in which it will be most natural to embed the manifold $\mathcal{M}$.

In section 2.4 we will be then ready to discuss the variational manifold $\mathcal{M}$ itself. We will define it as embedded in $\mathcal{P}(\mathcal{H})$ and we will introduce its tangent

space $\mathcal{T}_\psi\mathcal{M}$ and the Kähler structures that can be defined on it. In section 2.5 we will then come to the crucial point of this chapter, and that is the fact that the geometry of $\mathcal{M}$ that we have introduced can be of two types, defined as *Kähler* or *non-Kähler*. Finally, in section 2.6 we will introduce some further notions that will be useful in later discussions, including the differential geometric definition of observables and the related Poisson brackets.

Throughout the chapter we will accompany the discussion with simple examples to clarify the concepts that we introduce.

2.2 Kähler structures in Hilbert space

Given a separable Hilbert space $\mathcal{H}$ with inner product $\langle\cdot|\cdot\rangle$, we can always describe vectors by a set of complex number ψ_n with respect to a basis $\{|n\rangle\}$, *i.e.*,

$$|\psi\rangle = \sum_n \psi_n |n\rangle \ . \tag{2.2}$$

As stressed in the overview, we will be particularly interested in real subspaces of $\mathcal{H}$. Given a set of vectors $\{|n\rangle\}$, we thus distinguish the real and complex span

$$\begin{aligned} \mathrm{span}_{\mathbb{C}}\{|n\rangle\} &= \left\{\textstyle\sum_n \psi_n |n\rangle \ \middle|\ \psi_n \in \mathbb{C}\right\} \ , \\ \mathrm{span}_{\mathbb{R}}\{|n\rangle\} &= \left\{\textstyle\sum_n \psi_n |n\rangle \ \middle|\ \psi_n \in \mathbb{R}\right\} \ . \end{aligned} \tag{2.3}$$

On a real vector space, $|\psi\rangle \neq 0$ and $\mathrm{i}\,|\psi\rangle$ are linearly independent vectors, because one cannot be expressed as linear combination with real coefficients of the other. A real basis $\{|V_\mu\rangle\}$ of $\mathcal{H}$ has therefore twice as many elements as the complex basis $\{|n\rangle\}$, such as

$$\{|V_\mu\rangle\} \equiv \{|1\rangle\,, \mathrm{i}|1\rangle\,, |2\rangle\,, \mathrm{i}|2\rangle\,, \dots\} \ . \tag{2.4}$$

Given any real basis $\{|V_\mu\rangle\}$ of vectors, we can express every vector $|X\rangle$ as real linear combination

$$|X\rangle = X^\mu |V_\mu\rangle \ , \tag{2.5}$$

where we use Einstein's index summation convention, as we will do in the rest of the chapter.

A general real linear map is a map $\hat{A} : \mathcal{H} \to \mathcal{H}$ that satisfies $\hat{A}(\alpha\,|X\rangle) = \alpha\hat{A}\,|X\rangle$ only for real α. If it also holds for complex α, we refer to $\hat{A}$ as complex-linear. The imaginary unit i becomes itself a linear map, which only commutes with complex-linear maps.

The Hermitian inner product $\langle\cdot|\cdot\rangle$ can be decomposed into its real and imaginary parts given by

$$\langle V_\mu|V_\nu\rangle = \frac{\mathcal{N}}{2}\big(\boldsymbol{g}_{\mu\nu} + \mathrm{i}\,\boldsymbol{\omega}_{\mu\nu}\big) \tag{2.6}$$

with $\boldsymbol{g}_{\mu\nu} = \frac{2}{\mathcal{N}}\,\mathrm{Re}\,\langle V_\mu|V_\nu\rangle$, $\boldsymbol{\omega}_{\mu\nu} = \frac{2}{\mathcal{N}}\,\mathrm{Im}\,\langle V_\mu|V_\nu\rangle$ and $\mathcal{N}$ being an arbitrary normalisation which we will later fix in (2.25). This gives rise to the following set of structures, illustrated in Figure 2.1.

Definition 1 (Kähler space)**.** *A real vector space is called Kähler space if it is equipped with the following two bilinear forms*

- ***Metric***[1] $\boldsymbol{g}_{\mu\nu}$*: symmetric and positive-definite with inverse* $\boldsymbol{G}^{\mu\nu}$*, so that* $\boldsymbol{G}^{\mu\sigma}\boldsymbol{g}_{\sigma\mu} = \delta^\mu{}_\nu$*,*
- ***Symplectic form*** $\boldsymbol{\omega}_{\mu\nu}$*: antisymmetric and non-degenerate*[2] *with inverse* $\boldsymbol{\Omega}^{\mu\nu}$*, so that* $\boldsymbol{\Omega}^{\mu\sigma}\boldsymbol{\omega}_{\sigma\nu} = \delta^\mu{}_\nu$*,*

and such that the linear map $\boldsymbol{J}^\mu{}_\nu := -\boldsymbol{G}^{\mu\sigma}\boldsymbol{\omega}_{\sigma\nu}$ *is a*

- ***Complex structure*** $\boldsymbol{J}^\mu{}_\nu$*: satisfying* $\boldsymbol{J}^2 = -\mathbb{1}$*.*

The last condition is also called compatibility between $\boldsymbol{g}$ *and* $\boldsymbol{\omega}$*. We refer to* $(\boldsymbol{g}, \boldsymbol{\omega}, \boldsymbol{J})$ *as* ***Kähler structures****.*

In equation (2.6), clearly $\boldsymbol{g}$ is a metric and $\boldsymbol{\omega}$ is a symplectic form. Furthermore, we will see that they are indeed compatible and define a complex structure $\boldsymbol{J}$. For this, it is useful to introduce the real dual vectors $\mathrm{Re}\langle X|$ and $\mathrm{Im}\langle X|$ that act on a vector $|Y\rangle$ via

$$\mathrm{Re}\langle X|Y\rangle = \tfrac{\mathcal{N}}{2} X^\mu \boldsymbol{g}_{\mu\nu} Y^\nu\,,\ \mathrm{Im}\langle X|Y\rangle = \tfrac{\mathcal{N}}{2} X^\mu \boldsymbol{\omega}_{\mu\nu} Y^\nu\,, \tag{2.7}$$

as one may expect. The identity $\mathbb{1} = \sum_n |n\rangle\langle n|$ is then

$$\mathbb{1} = \tfrac{2}{\mathcal{N}}\,\boldsymbol{G}^{\mu\nu}\,|V_\mu\rangle\ \mathrm{Re}\langle V_\nu|\ . \tag{2.8}$$

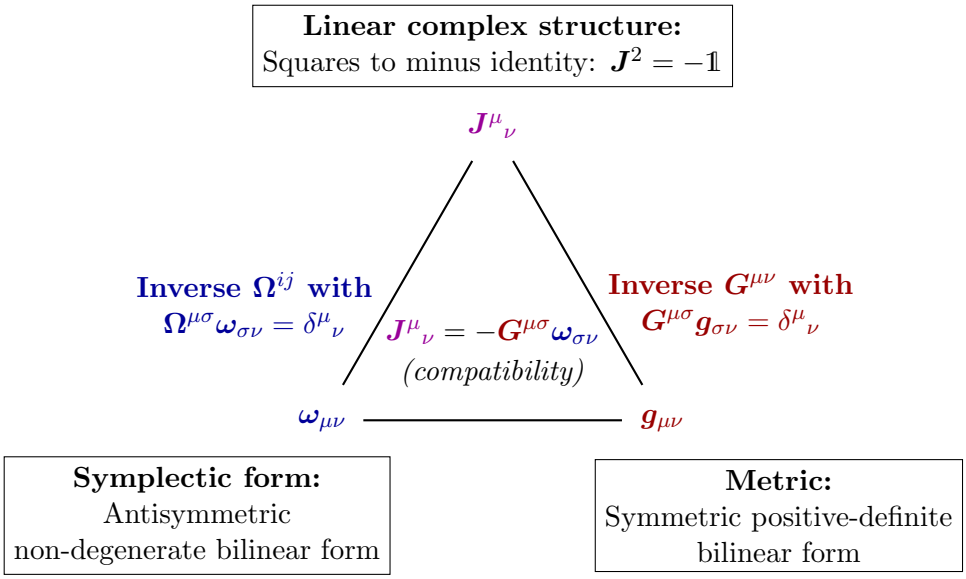
Similarly, the matrix representation of an operator $\hat{A}$ is

$$A^\mu{}_\nu = \tfrac{2}{\mathcal{N}}\,\boldsymbol{G}^{\mu\sigma}\mathrm{Re}\langle V_\sigma|\hat{A}|V_\nu\rangle\ . \tag{2.9}$$

[1] Here, "metric" refers to a metric tensor, *i.e.*, an inner product on a vector space. It should not be confused with the notion of metric spaces in analysis and topology.

[2] A bilinear form $b_{\mu\nu}$ is called non-degenerate, if it is invertible. For this, we can check $\det(b) \neq 0$ in any basis of our choice.

Figure 2.1: *Triangle of Kähler structures.* This sketch illustrates the triangle of Kähler structures, consisting of a symplectic form $\boldsymbol{\omega}$, a positive definite metric $\boldsymbol{g}$ and a linear complex structure $\boldsymbol{J}$. We also define the inverse symplectic form $\boldsymbol{\Omega}$ and the inverse metric $\boldsymbol{G}$.



In particular, we compute the matrix representation of the imaginary unit i to be given by

$$\boldsymbol{J}^\mu{}_\nu = \tfrac{2}{\mathcal{N}}\,\boldsymbol{G}^{\mu\sigma}\mathrm{Re}\langle V_\sigma|\mathrm{i}|V_\nu\rangle = -\boldsymbol{G}^{\mu\sigma}\boldsymbol{\omega}_{\sigma\nu} \tag{2.10}$$

as anticipated in our definition. From $\mathrm{i}^2 = -1$, we conclude that the so defined $\boldsymbol{J}$ indeed satisfies $\boldsymbol{J}^2 = -\mathbb{1}$ and is thus a complex structures. Therefore, $\boldsymbol{g}$ and $\boldsymbol{\omega}$ as defined in (2.6) are compatible.

Example 3 (Hilbert space of a Qubit)**.** *A qubit is described by the Hilbert space* $\mathcal{H} = \mathbb{C}^2$ *with complex basis* $\{|0\rangle, |1\rangle\}$ *and real basis*

$$|V_i\rangle \equiv \{|0\rangle, |1\rangle, \mathrm{i}\,|0\rangle, \mathrm{i}\,|1\rangle\}\,. \tag{2.11}$$

With respect to this real basis $\boldsymbol{g}_{\mu\nu}$, $\boldsymbol{\omega}_{\mu\nu}$ *and* $\boldsymbol{J}^\mu{}_\nu$ *are*

$$\boldsymbol{g}_{\mu\nu} \equiv \tfrac{2}{\mathcal{N}}\begin{pmatrix}\mathbb{1} & 0\\ 0 & \mathbb{1}\end{pmatrix}\,,\ \boldsymbol{\omega}_{\mu\nu} \equiv \tfrac{2}{\mathcal{N}}\begin{pmatrix}0 & \mathbb{1}\\ -\mathbb{1} & 0\end{pmatrix}\,,\ \boldsymbol{J}^\mu{}_\nu \equiv \begin{pmatrix}0 & -\mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\,, \tag{2.12}$$

where $\mathbb{1}$ *is the* 2×2 *identity matrix. We can represent a complex-linear map* $\hat{A} = \sum_{n,m} a_{nm} |n\rangle \langle m|$*, i.e., with* $[A, J] = 0$*, as the matrix*

$$A^{\mu}{}_{\nu} \equiv \begin{pmatrix} \mathbb{A} & -\mathbb{B} \\ \mathbb{B} & \mathbb{A} \end{pmatrix} , \tag{2.13}$$

where $\mathbb{A} = \mathrm{Re}(a)$ *and* $\mathbb{B} = \mathrm{Im}(a)$ *in the above basis.*

In summary, every Hilbert space is a real Kähler space with metric, symplectic form and complex structure.

2.3 Projective Hilbert space

Multiplying a Hilbert space vector $|\psi\rangle$ with a non-zero complex number does not change the quantum state it represents. Therefore, the manifold representing all physical states is given by the projective Hilbert space $\mathcal{P}(\mathcal{H})$, which we will define and analyze in this section. Variational families, which we will discuss in the following section, should then naturally be understood as submanifolds $\mathcal{M}$ of projective Hilbert space $\mathcal{P}(\mathcal{H})$.

The projective Hilbert space of $\mathcal{H}$

$$\mathcal{P}(\mathcal{H}) = (\mathcal{H} \backslash \{0\}) \,/ \sim \tag{2.14}$$

is given by the equivalence classes of non-zero Hilbert space vectors with respect to the equivalence relation

$$|\psi\rangle \sim |\tilde{\psi}\rangle \quad \Leftrightarrow \quad \exists\, c \in \mathbb{C} \;\; \text{with} \;\; |\tilde{\psi}\rangle = c\,|\psi\rangle \; . \tag{2.15}$$

Thus, a state $\psi \in \mathcal{P}(\mathcal{H})$ is a ray in Hilbert space consisting of all non-zero vectors that are related by multiplication with a non-zero complex number c.

Let us now consider the tangent space to the manifold $\mathcal{P}(\mathcal{H})$. For a rigorous definition of tangent spaces we refer the reader to any standard textbook of differential geometry. However, intuitively $\mathcal{T}_{\psi}\mathcal{P}(\mathcal{H})$ represents the space of linear changes $\delta\psi$ around an element $\psi \subset \mathcal{P}(\mathcal{H})$. Changing a representative $|\psi\rangle$ in the direction of itself, *i.e.*, $|\delta\psi\rangle \propto |\psi\rangle$, corresponds to changing $|\psi\rangle$ by a complex factor and thus does not change the underlying state ψ. Two Hilbert space vectors $|X\rangle, |\tilde{X}\rangle \in \mathcal{H}$ therefore represent the same change $|\delta\psi\rangle$ of the state $|\psi\rangle \in \psi$, if they only differ by some $\alpha\,|\psi\rangle$. We define tangent space as

$$\mathcal{T}_{\psi}\mathcal{P}(\mathcal{H}) = \mathcal{H} / \approx , \tag{2.16}$$

where we introduced the equivalence relation

$$|X\rangle \approx |\tilde{X}\rangle \;\; \Leftrightarrow \;\; \exists\, c \in \mathbb{C} \;\text{with}\; |X\rangle - |\tilde{X}\rangle = c\,|\psi\rangle \; , \tag{2.17}$$

leading to a regular (not projective) vector space.

We can pick a unique representative $|X\rangle$ of the class $[|\delta\psi\rangle]$ at the state $|\psi\rangle$ by requiring $\langle\psi|X\rangle = 0$. Viceversa, two vectors $|X\rangle \neq |\tilde{X}\rangle$ both satisfying $\langle\psi|X\rangle = \langle\psi|\tilde{X}\rangle = 0$ belong to different equivalence classes. We thus identify $\mathcal{T}_\psi\mathcal{P}(\mathcal{H})$ with

$$\mathcal{H}_\psi^\perp = \left\{ |X\rangle \in \mathcal{H} \,\middle|\, \langle\psi|X\rangle = 0 \right\} . \tag{2.18}$$

Given a general representative $|\delta\psi\rangle \in [|\delta\psi\rangle]$, we compute the unique representative mentioned above as $|X\rangle = \mathbb{Q}_\psi |\delta\psi\rangle$ with

$$\mathbb{Q}_\psi |\delta\psi\rangle = |\delta\psi\rangle - \frac{\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} |\psi\rangle . \tag{2.19}$$

There is a further subtlety: representing a change $\delta\psi$ of a state ψ as vector $|\delta\psi\rangle$ will always be with respect to a representative $|\psi\rangle$. If we choose a different representative $|\tilde{\psi}\rangle = c\,|\psi\rangle \in \psi$, the same change $\delta\psi$ would be represented by a different Hilbert space vector $|\delta\tilde{\psi}\rangle = c\,|\delta\psi\rangle$. It therefore does not suffice to specify a Hilbert space vector $|\delta\psi\rangle$, but we always need to say with respect to which representative $|\psi\rangle$ it was chosen. This could be avoided when moving to density operators[3].

The fact we can identify the tangent space at each point with a Hilbert space $\mathcal{H}_\psi^\perp$ enables us, given a *local* real basis $\{|V_\mu\rangle\}$ at ψ, such that $\mathcal{H}_\psi^\perp = \mathrm{span}_\mathbb{R}\{|V_\mu\rangle\}$, to induce a canonical metric $\boldsymbol{g}_{\mu\nu}$, symplectic form $\boldsymbol{\omega}_{\mu\nu}$ and $\boldsymbol{J}^\mu{}_\nu$ onto the tangent space, which thus is a Kähler space, as discussed previously. We see at this point that on the tangent space $\mathcal{T}_\psi\mathcal{P}(\mathcal{H})$, it is convenient to choose $\mathcal{N} = \langle\psi|\psi\rangle$ as normalisation for the Kähler structures. The rescaled metric $\frac{1}{2}\boldsymbol{g}_{\mu\nu}$ is well-known as the Fubini-Study metric [74, 75], while the symplectic form gives projective Hilbert space a natural phase space structure.

Manifolds such as $\mathcal{P}(\mathcal{H})$, whose tangent spaces are equipped with differentiable Kähler structures, are called almost-Hermitian manifolds. In appendix A.2, we show that $\mathcal{P}(\mathcal{H})$ satisfies even stronger conditions, which make it a so-called *Kähler* manifold.

Example 4 (Bloch sphere)**.** *The projective Hilbert space of a qubit is* $\mathcal{P}(\mathbb{C}^2) = S^2$*, equivalent to the Bloch sphere. Using spherical coordinates* $x \equiv (\theta, \phi)$ *and the complex Hilbert space basis* $\{|0\rangle, |1\rangle\}$*, we can parametrize the set of states as*

$$|\psi(x)\rangle = \cos\left(\tfrac{\theta}{2}\right)|0\rangle + e^{\mathrm{i}\phi}\sin\left(\tfrac{\theta}{2}\right)|1\rangle . \tag{2.20}$$

[3] We can equivalently define projective Hilbert space as the set of pure density operators, *i.e.*, Hermitian, positive operators ρ with $\mathrm{Tr}\rho = \mathrm{Tr}\rho^2 = 1$. The state ψ is then given by the density operator $\rho_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}$ and its change $\delta\psi$ by $\delta\rho_\psi = \frac{|\delta\psi\rangle\langle\psi| + |\psi\rangle\langle\delta\psi|}{\langle\psi|\psi\rangle}$.

The elements of $\mathcal{P}(\mathcal{H})$ *are the equivalence classes* $\psi(x) = \{c\,|\psi(x)\rangle \,|\, c \in \mathbb{C}, c \neq 0\}$. *Consequently, the tangent space* $\mathcal{T}_\psi\mathcal{P}(\mathbb{C}^2) = \mathcal{H}_\psi^\perp$ *of the Bloch sphere at* $x^\mu \equiv (\theta, \phi)$ *can be spanned by the basis* $|V_\mu\rangle = \mathbb{Q}_\psi\left(\frac{\partial}{\partial x^\mu}\right)|\psi(x)\rangle$ *with*

$$\begin{aligned} |V_1\rangle &= -\tfrac{1}{2}\sin\left(\tfrac{\theta}{2}\right)|0\rangle + \tfrac{e^{\mathrm{i}\phi}}{2}\cos\left(\tfrac{\theta}{2}\right)|1\rangle\ , \\ |V_2\rangle &= -\tfrac{\mathrm{i}}{2}\sin\left(\tfrac{\theta}{2}\right)\sin\theta\,|0\rangle + \tfrac{\mathrm{i}e^{\mathrm{i}\phi}}{2}\cos\left(\tfrac{\theta}{2}\right)\sin\theta\,|1\rangle\ . \end{aligned} \tag{2.21}$$

Using the definition (2.6) *of the metric and symplectic form from the Hilbert space inner product, we can compute the matrix representations*

$$\boldsymbol{g}_{\mu\nu} \equiv 2\begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \text{ and } \boldsymbol{\omega}_{\mu\nu} \equiv 2\begin{pmatrix} 0 & \sin\theta \\ -\sin\theta & 0 \end{pmatrix}. \tag{2.22}$$

We recognize $\boldsymbol{g}_{\mu\nu}dx^\mu dx^\nu = \frac{1}{2}(d\theta^2 + \sin^2(\theta)d\phi^2)$ *to be the standard metric of a sphere with radius* $1/\sqrt{2}$. *Similarly, we recognize* $\boldsymbol{\omega}_{\mu\nu}dx^\mu dx^\nu = \frac{1}{2}\sin\theta d\theta\wedge d\phi$ *to be the standard volume form on this sphere. Finally, it is easy to verify that* $\boldsymbol{J}^2 = -\mathbb{1}$ *everywhere.*

In summary, a given pure state can be represented by the equivalence class $\psi \in \mathcal{P}(\mathcal{H})$ of all states related by multiplication with a non-zero complex number. Similarly, a tangent vector $[|X\rangle] \in \mathcal{T}_\psi\mathcal{P}(\mathcal{H})$ at a state ψ is initially defined as the affine space $[|X\rangle]$ of all vectors $|X\rangle$ differing by a complex multiple of $|\psi\rangle$. A unique representative $|\tilde{X}\rangle$ can be chosen requiring $\langle\psi|\tilde{X}\rangle = 0$. This leads to the identification $\mathcal{T}_\psi\mathcal{P}(\mathcal{H}) \simeq \mathcal{H}_\psi^\perp$, such that the Hilbert space inner product $\langle\cdot|\cdot\rangle$ induces local Kähler structures onto $\mathcal{T}_\psi\mathcal{P}(\mathcal{H})$.

2.4 Generic variational manifolds

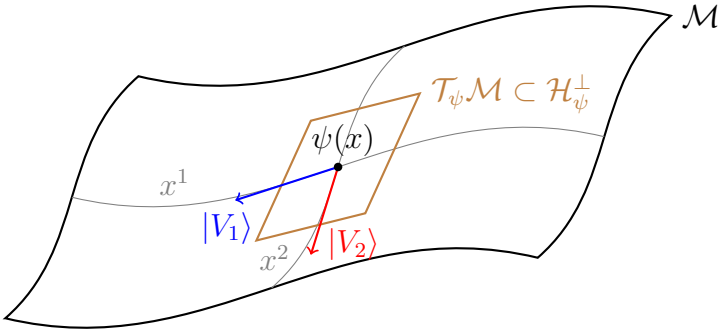
The most general variational family is a real differentiable submanifold $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$. Similarly to what we did for the full projective Hilbert space, at every point $\psi(x) \in \mathcal{M}$, we can define the tangent space $\mathcal{T}_\psi\mathcal{M}$. This is the space of tangent vectors $|X\rangle_\psi$ representing local linear variations of the state $\psi(x)$ that can be achieved by moving within the manifold $\mathcal{M}$. As $\mathcal{M}$ is embedded in $\mathcal{P}(\mathcal{H})$, so is $\mathcal{T}_\psi\mathcal{M}$ can be embedded into $\mathcal{T}_\psi\mathcal{P}(\mathcal{H}) \simeq \mathcal{H}_\psi^\perp$.

More specifically, we can define a local basis $|V_\mu\rangle_\psi \in \mathcal{H}_\psi^\perp$, such that

$$|X\rangle_\psi = X^\mu\,|V_\mu\rangle_\psi\ . \tag{2.23}$$

Note that in general the tangent space $\mathcal{T}_\psi\mathcal{M} = \mathrm{span}_{\mathbb{R}}\{|V_\mu\rangle\}$ is only a *real*, but not necessarily a complex subspace of $\mathcal{H}_\psi^\perp$. Thus, we will encounter families, for which $|X\rangle$ is a tangent vector, but not $\mathrm{i}|X\rangle$.

Figure 2.2: *Tangent vectors.* We sketch the basis vectors $|V_\mu\rangle$ of tangent space $\mathcal{T}_\psi\mathcal{M}$ for a manifold $\mathcal{M}$ parametrized by two coordinates (x^1, x^2).



In practice, we often parametrize $\psi(x) \in \mathcal{M}$ by choosing a representative $|\psi(x)\rangle \in \mathcal{H}$. This allows us to construct the local basis $|V_\mu(x)\rangle$ of tangent space $\mathcal{T}_\psi\mathcal{M}$ as

$$|V_\mu(x)\rangle = \mathbb{Q}_{\psi(x)}\, \partial_\mu\, |\psi(x)\rangle \ , \tag{2.24}$$

at the state $|\psi(x)\rangle$, where $\mathbb{Q}_\psi$ was defined[4] in (2.19). To simplify notation, we will usually drop the reference to $\psi(x)$ or x and only write $|V_\mu\rangle$, whenever it is clear at which state we are. The schematic idea behind tangent space is sketched in Figure 2.2.

Similar to projective Hilbert space, we define *restricted* Kähler structures on tangent space $\mathcal{T}_\psi\mathcal{M} \subset \mathcal{T}_\psi\mathcal{P}(\mathcal{H})$ as

$$\boldsymbol{g}_{\mu\nu} = \frac{2\,\mathrm{Re}\langle V_\mu|V_\nu\rangle}{\langle\psi|\psi\rangle} \quad \text{and} \quad \boldsymbol{\omega}_{\mu\nu} = \frac{2\,\mathrm{Im}\langle V_\mu|V_\nu\rangle}{\langle\psi|\psi\rangle} \ . \tag{2.25}$$

There are two important differences to the corresponding definition (2.6) in full Hilbert space. First, with a slight abuse of notation, $|V_\mu\rangle$ here does not span the Hilbert space, but rather the typically much smaller tangent space. Second, we set $\mathcal{N} = \langle\psi|\psi\rangle$ just like for $\mathcal{P}(\mathcal{H})$, such that

$$\langle V_\mu|V_\nu\rangle = \frac{\langle\psi|\psi\rangle}{2}\left(\boldsymbol{g}_{\mu\nu} + \mathrm{i}\boldsymbol{\omega}_{\mu\nu}\right) \ . \tag{2.26}$$

[4]The projector $\mathbb{Q}_\psi$ is important to ensure that $|V_\mu\rangle$ can be identified with an element of $\mathcal{H}_\psi^\perp \simeq \mathcal{T}_\psi\mathcal{P}(\mathcal{H})$ as discussed in Section 2.3, *i.e.*, $\langle\psi|V_\mu\rangle = 0$. For derivations, it can be useful to choose a local coordinate system of x, in which $|V_\mu\rangle = \partial_\mu\,|\psi\rangle$, *i.e.*, the action of $\mathbb{Q}_\psi$ can be ignored. This can always be achieved locally at a point and any invariant expressions derived this way, will be valid in any coordinate system.

This has the important consequence that the restricted Kähler structures are invariant under the change of representative $|\psi\rangle$ of the physical state. Namely, under the transformation $|\psi\rangle \to c\,|\tilde{\psi}\rangle$ with $|V_\mu\rangle \to c\,|V_\mu\rangle$, our Kähler structures will not change. This ensures that equations involving restricted Kähler structures are manifestly defined on projective Hilbert space and thus independent of the representative $|\psi(x)\rangle \in \mathcal{H}$, we use to represent the abstract state $\psi(x) \in \mathcal{M} \subset \mathcal{P}(\mathcal{H})$.

Projectors on tangent space

For every two Hilbert space vectors $|X\rangle, |Y\rangle \in \mathcal{H}$, we have the real inner product $\mathrm{Re}\langle X|Y\rangle$. This induces the norm $\||X\rangle\| = \sqrt{\mathrm{Re}\,\langle X|X\rangle} = \sqrt{\langle X|X\rangle}$, which is nothing more than the regular Hilbert space norm. We can then define the orthogonal projector $\mathbb{P}_\psi$ from $\mathcal{H}$ onto $\mathcal{T}_\psi\mathcal{M}$ with respect to this norm $\mathrm{Re}\langle\cdot|\cdot\rangle$, *i.e.*, for each vector $|X\rangle \in \mathcal{H}$ we define the vector $\mathbb{P}_\psi\,|X\rangle$ as the vector in $\mathcal{T}_\psi\mathcal{M}$ that is closest to $|X\rangle$ in this norm. That is, we have

$$\mathbb{P}_\psi\,|X\rangle = \underset{|\Phi\rangle\in\mathcal{T}_\psi\mathcal{M}}{\mathrm{argmin}} \;\; \||X\rangle - |\Phi\rangle\|\,. \tag{2.27}$$

We can write this orthogonal projector in two ways:

$$\mathbb{P}_\psi = \frac{2\,|V_\mu\rangle\,\boldsymbol{G}^{\mu\nu}\mathrm{Re}\langle V_\nu|}{\langle\psi|\psi\rangle}\,, \quad \mathbb{P}^\mu_\psi = \frac{2\boldsymbol{G}^{\mu\nu}\mathrm{Re}\langle V_\nu|}{\langle\psi|\psi\rangle}\,, \tag{2.28}$$

such that we have $\mathbb{P}_\psi = |V_\mu\rangle\,\mathbb{P}^\mu_\psi$. The difference lies in the co-domain: while $\mathbb{P}_\psi : \mathcal{H} \to \mathcal{H}$ maps back onto Hilbert space, *e.g.*, to compute $\mathbb{P}^2_\psi = \mathbb{P}_\psi$, we have that $\mathbb{P}^\mu_\psi : \mathcal{H} \to \mathcal{T}_\psi\mathcal{M}$ is a map from Hilbert space into tangent space. Due to $\mathcal{T}_\psi\mathcal{M} \subset \mathcal{T}_\psi\mathcal{P}(\mathcal{H})$, we have

$$\mathbb{P}_\psi = \mathbb{P}_\psi\mathbb{Q}_\psi = \mathbb{Q}_\psi\mathbb{P}_\psi \quad \text{and} \quad \mathbb{P}^\mu_\psi = \mathbb{P}^\mu_\psi\mathbb{Q}_\psi\,, \tag{2.29}$$

which follows from $\mathbb{Q}_\psi\,|V_\mu\rangle = |V_\mu\rangle$ and $\mathbb{Q}^\dagger_\psi = \mathbb{Q}_\psi$. In contrast to $\mathbb{Q}_\psi$, the projector $\mathbb{P}_\psi$ is in general not Hermitian.

Kähler structures

Provided that there are no redundancies or gauge directions (only changing phase or normalisation) in our choice of parameters, $\boldsymbol{g}_{\mu\nu}$ will still be positive-definite and invertible with inverse $\boldsymbol{G}^{\mu\nu}$. We find that

$$\boldsymbol{J}^\mu{}_\nu = -\boldsymbol{G}^{\mu\sigma}\boldsymbol{\omega}_{\sigma\nu} = \frac{2\boldsymbol{G}^{\mu\sigma}\mathrm{Re}\langle V_\sigma|\mathrm{i}|V_\nu\rangle}{\langle\psi|\psi\rangle} = \mathbb{P}^\mu_\psi\mathrm{i}\,|V_\nu\rangle \tag{2.30}$$

is the projection of the multiplication by the imaginary unit (as real-linear map) onto $\mathcal{T}_\psi\mathcal{M}$. It will not square to minus identity if multiplication by i in full Hilbert space does not preserve the tangent space.

If $\boldsymbol{g}_{\mu\nu}$ is not invertible, it means that there exists a set of coefficients X^μ such that $X^\mu \boldsymbol{g}_{\mu\nu} X^\nu = 0$, that is $\|X^\mu |V_\mu\rangle\| = 0$ and therefore $X^\mu |V_\mu\rangle = 0$. In other words, not all vectors $|V_\mu\rangle$ are linearly independent and thus also not all parameters are independent. If this is the case, it is not a real problem as the formalism introduced can still be used with little modifications. More precisely, the projectors (2.28), as well as all other objects we will introduce, are meaningfully defined if we indicate with $\boldsymbol{G}^{\mu\nu}$ the Moore-Penrose pseudo-inverse of $\boldsymbol{g}_{\mu\nu}$, *i.e.*, we invert $\boldsymbol{g}_{\mu\nu}$ only on the orthogonal complement to its kernel (orthogonal with respect of the flat metric $\delta_{\mu\nu}$ in our coordinates[5]). Indeed, all directions in the kernel correspond to a vanishing vector in the tangent space and therefore do not matter. In this case, also $\boldsymbol{\Omega}^{\mu\nu}$, should be defined as the inverse of $\boldsymbol{\omega}_{\mu\nu}$ on the orthogonal complement to the kernel of $\boldsymbol{g}_{\mu\nu}$.[6] However, it is still possible that $\boldsymbol{\omega}$ and $\boldsymbol{J}$ are not invertible even on this reduced subspace.

In this case, in order to define $\boldsymbol{\Omega}$ one has to reduce oneself to working on an even smaller subspace, that is one that does not contain the kernel of $\boldsymbol{\omega}$ and $\boldsymbol{J}$. Here, however, the way in which we reduce these extra dimensions is not equivalent, as these directions are not anymore just redundant gauge choices of our parametrisation. The reduction here effectively corresponds to working on a physically smaller manifold, as we will explain better in the next section. For what follows we will always suppose that $\boldsymbol{\Omega}$ is defined by inverting $\boldsymbol{\omega}$ on the tangent subspace orthogonal, with respect to the metric $\boldsymbol{g}_{\mu\nu}$, to the kernel of $\boldsymbol{J}$. That is, $\boldsymbol{\Omega}$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{\omega}$ with respect to $\boldsymbol{g}$, *i.e.*, the pseudo-inverse is evaluated in an orthonormal basis.

In conclusion, we see that we are able to define the *restricted* structures $(\boldsymbol{g}, \boldsymbol{\omega}, \boldsymbol{J})$ which, however, do not necessarily satisfy the Kähler property. This is due to the fact that the tangent space, as we have pointed out, is a real, but not necessarily complex subspace of $\mathcal{H}$. Note that these objects are locally defined in each tangent space $\mathcal{T}_\psi\mathcal{M}$ for $\psi \in \mathcal{M}$.

Example 5 (Two qubits)**.** *For the Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes 2}$ of two qubits, we can choose the variational manifold $\mathcal{M}$ of symmetric product states represented by*

$$|\varphi(x)\rangle = |\psi(x)\rangle \otimes |\psi(x)\rangle \ , \tag{2.31}$$

[5]In the specific case of the manifold of matrix product states, there exists a different, more natural definition of orthogonality [76].

[6]Note that the kernel of $\boldsymbol{\omega}_{\mu\nu}$ itself does not necessarily correspond to redundant directions of the parametrisation as $X^\mu \boldsymbol{\omega}_{\mu\nu} = 0$ does not imply $X^\mu |V_\mu\rangle = 0$.

with $x^\mu \equiv (\theta, \phi)$, where $|\psi(x)\rangle$ is a single qubit state as parametrized in (2.20)*. The tangent space is spanned by*

$$|W_\mu\rangle = |V_\mu\rangle \otimes |\psi(x)\rangle + |\psi(x)\rangle \otimes |V_\mu\rangle \ , \tag{2.32}$$

where $|V_\mu\rangle$ are the single qubit tangent vectors defined in (2.21)*. With this, we find according to* (2.25)

$$\boldsymbol{g}_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(\theta) \end{pmatrix} \text{ and } \boldsymbol{\omega}_{\mu\nu} \equiv \begin{pmatrix} 0 & \sin\theta \\ -\sin\theta & 0 \end{pmatrix} \tag{2.33}$$

leading to $\boldsymbol{J}^2 = -\mathbb{1}$ everywhere. We therefore conclude that the tangent space $\mathcal{T}_\psi\mathcal{M}$ satisfies the Kähler property everywhere.

Example 6 (Bloch sphere equator)**.** *For the single qubit Hilbert space $\mathcal{H} = \mathbb{C}^2$, we can choose the equator of the Bloch sphere as our variational manifold $\mathcal{M}$. This amounts to fixing $\theta = \pi/2$ in the single qubit state* (2.20) *leading to the representatives*

$$|\psi(\phi)\rangle = \tfrac{1}{\sqrt{2}}|0\rangle + \frac{e^{\mathrm{i}\phi}}{\sqrt{2}}|1\rangle \tag{2.34}$$

with a single variational parameter ϕ. We have the single tangent vector $|V\rangle = |V_1\rangle$ as defined in (2.21)*. From the inner product $\langle V|V\rangle = \frac{1}{4}$, we find $\boldsymbol{g} = \frac{1}{2}$ and $\boldsymbol{\omega} = 0$ implying $\boldsymbol{J} = 0$. Consequently, and not surprising due to the odd dimension, the tangent spaces of our variational manifold $\mathcal{M}$ are not Kähler spaces. Moreover, neither $\boldsymbol{\omega}$ nor $\boldsymbol{J}$ are invertible.*

Example 7 (Squeezed coherent state)**.** *We consider a bosonic system with two degrees of freedom associated with annihilation operators $\hat{a}_1$ and $\hat{a}_2$. The vacuum state $|0,0\rangle$ satisfies $\hat{a}_m|0,0\rangle = 0$, $\hat{a}_1^\dagger|0,0\rangle = |1,0\rangle$ and $\hat{a}_2^\dagger|0,0\rangle = |0,1\rangle$. We introduce*

$$\hat{b} = \cosh r\, \hat{a}_1 + \sinh r\, \hat{a}_2^\dagger \tag{2.35}$$

with canonical commutation relations $[\hat{b}, \hat{b}^\dagger] = 1$ and r being a fixed constant (not a variational parameter). We then define the states of our variational manifold as

$$|\psi(\alpha)\rangle = e^{\alpha\hat{b}^\dagger - \alpha^*\hat{b}}|0\rangle \ , \tag{2.36}$$

parametrized by a single complex number α. $|\psi(\alpha)\rangle$ is not the one-mode coherent state $|\alpha\rangle = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}|0\rangle$, because $\hat{b}|0\rangle \neq 0$. Our variational manifold has two

independent real parameters given by $x \equiv (\mathrm{Re}\alpha, \mathrm{Im}\alpha)$. *After some algebra taking* $[\hat{b}, \hat{b}^\dagger] = 1$ *into account, we find*

$$\begin{aligned} |V_1\rangle &= e^{\alpha\hat{b}^\dagger - \alpha^*\hat{b}} \left(\cosh r \ |1,0\rangle - \sinh r \, |0,1\rangle\right), \\ |V_2\rangle &= e^{\alpha\hat{b}^\dagger - \alpha^*\hat{b}} \, \mathrm{i}(\cosh r \, |1,0\rangle + \sinh r \, |0,1\rangle). \end{aligned} \tag{2.37}$$

Metric and symplectic form take the forms

$$\boldsymbol{g}_{\mu\nu} \equiv \cosh 2r \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \textit{and} \quad \boldsymbol{\omega}_{\mu\nu} \equiv \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}. \tag{2.38}$$

This gives rise to the restricted complex structure

$$\boldsymbol{J}^\mu{}_\nu \equiv \operatorname{sech} 2r \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{2.39}$$

which only satisfies $\boldsymbol{J}^2 = -\mathbb{1}$ *for* $r = 0$.

In summary, we introduced general variational manifolds as real differentiable submanifolds $\mathcal{M}$ of projective Hilbert space $\mathcal{P}(\mathcal{H})$. By embedding the tangent spaces $\mathcal{T}_\psi\mathcal{M}$ into Hilbert space, the Hilbert space inner product defines restricted Kähler structures on the tangent spaces, whose properties we will explore next.

2.5 Kähler and non-Kähler manifolds

We categorize variational manifolds depending on whether their tangent spaces are Kähler spaces or not. We will see in the following sections that this distinction has some important consequences for the application of variational methods on the given family.

Definition 2. *We classify general variational families* $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$ *based on their restricted Kähler structures. We refer to a variational family* $\mathcal{M}$ *as*

- ***Kähler***[7]*, if all tangent spaces* $\mathcal{T}_\psi\mathcal{M}$ *are a Kähler spaces, i.e.,* $\boldsymbol{J}^2 = -\mathbb{1}$ *everywhere on the manifold,*
- ***Non-Kähler****, if it is not Kähler. If* $\boldsymbol{\omega}$ *is degenerate, we define* $\boldsymbol{\Omega}$ *as the pseudo-inverse.*

[7] A general manifold $\mathcal{M}$, whose tangent spaces are equipped with compatible Kähler structures, is known as an almost Hermitian manifold. However, if an almost Hermitian manifold is the submanifold of a Kähler manifold (as defined in appendix A.2), then it is also a Kähler manifold itself. Thus, due to the fact that $\mathcal{P}(\mathcal{H})$ is a Kähler manifold, all almost Hermitian submanifolds $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$ are also Kähler manifolds, which is why we use the term Kähler in this context.

This classification refers to the manifold as a whole. In table 2.1 we summarize the properties of each class of manifolds.

Many well-known variational families, such as Gaussian states [55], coherent states [62, 63, 77], matrix product states [78] and projected entangled pair states [31], are Kähler. On the other hand, one naturally encounters non-Kähler manifolds when one parametrizes states through a family of general unitaries $\mathcal{U}(x)$ applied to a reference state $|\phi\rangle$, *i.e.*,

$$|\psi(x)\rangle = \mathcal{U}(x)\,|\phi\rangle\ . \tag{2.40}$$

For example, this issue arises for the classes of generalized Gaussian states introduced in [52], for the Multi-scale Entanglement Renormalisation Ansatz states [79] or if one applies Gaussian transformations $\mathcal{U}(x)$ to general non-Gaussian states.

Example 8 (Kähelr and non-Kähler manifolds)**.** *We already reviewed examples for these cases in the previous section. More precisely, example 5 is Kähler, example 6 is non-Kähler with degenerate $\boldsymbol{\omega}$ and example 7 is non-Kähler with non-degenerate $\boldsymbol{\omega}$.*

Given a submanifold $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$, we can use the embedding in the manifold $\mathcal{P}(\mathcal{H})$ to constrain the form that the restricted complex structure J can take on $\mathcal{M}$.

Proposition 1. *On a tangent space $\mathcal{T}_\psi\mathcal{M} \subset \mathcal{H}$ of a submanifold $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$ we can always find an orthonormal basis $\{|V_\mu\rangle\}$, such that $\boldsymbol{g}_{\mu\nu} \equiv \mathbb{1}$ and the restricted complex structure is represented by the block matrix*

$$\boldsymbol{J}^{\mu}{}_{\nu} \equiv \left(\begin{array}{c|c|c} \begin{array}{cc} & 1 \\ -1 & \\ & \ddots \end{array} & & \\ \hline & \begin{array}{cccc} c_1 & & & \\ & -c_1 & & \\ & & c_2 & \\ & & & -c_2 \end{array} \ddots & \\ \hline & & \begin{array}{cc} 0 & \\ & \ddots \end{array} \end{array}\right) \tag{2.41}$$

with $0 < c_i < 1$. This standard form induces the decomposition of $\mathcal{T}_\psi\mathcal{M}$ into the three orthogonal parts

$$\mathcal{T}_\psi\mathcal{M} = \underbrace{\overline{\overline{\mathcal{T}_\psi\mathcal{M}} \oplus \mathcal{I}_\psi\mathcal{M}}}_{\overline{\mathcal{T}_\psi\mathcal{M}}} \oplus \mathcal{D}_\psi\mathcal{M}\,, \tag{2.42}$$

where $\overline{\overline{\mathcal{T}_\psi\mathcal{M}}}$ *is the largest Kähler subspace and* $\overline{\mathcal{T}_\psi\mathcal{M}}$ *is the largest space on which* J *and* ω *are invertible.*

Proof. We present a constrictive proof in appendix A.1. □

Proposition 1 is also relevant for classifying real subspaces of complex Hilbert spaces. Interestingly, it is linked to the entanglement structure of fermionic Gaussian states, as made explicit in [80].

The manifold $\mathcal{M}$ is Kähler if there is only the first block in (2.41) everywhere. The symplectic form $\boldsymbol{\omega}$ is non-degenerate if we only have the first two diagonal blocks. The next proposition provides some further intuition for the non-Kähler case, which is also known in mathematics in the context of sub manifolds of Kähler manifolds [81].

Proposition 2. *The Kähler property is equivalent to requiring that* $\mathcal{T}_\psi\mathcal{M}$ *is not just a real, but also a complex subspace, i.e., for all* $|X\rangle \in \mathcal{T}_\psi\mathcal{M}$*, we also have* $\mathrm{i}|X\rangle \in \mathcal{T}_\psi\mathcal{M}$*. Therefore, the multiplication by* i *commutes with the projector* $\mathbb{P}_\psi$*, i.e.,* $\mathbb{P}_\psi\mathrm{i} = \mathrm{i}\mathbb{P}_\psi$ *and* $\mathbb{P}_\psi$ *is complex-linear.*

Proof. We present a proof in appendix A.1. □

If a manifold admits a complex holomorphic parametrisation, *i.e.*, a parametrisation that depends on the complex parameters z, but not on z^*, then the manifold will be Kähler. Indeed, taking $\mathrm{Re}\, z$ and $\mathrm{Im}\, z$ as real parameters gives the tangent vectors

$$|v_\mu\rangle = \frac{\partial}{\partial \mathrm{Re} z^\mu}|\psi(z)\rangle\,, \quad \mathrm{i}\,|v_\mu\rangle = \frac{\partial}{\partial \mathrm{Im} z^\mu}|\psi(z)\rangle\,. \tag{2.43}$$

It is actually also possible to show that, viceversa, a Kähler manifold is also a complex manifold, that is it admits, at least locally, a complex holomorphic parametrisation.

As mentioned before, in order to define the inverse of $\boldsymbol{\omega}$ it is necessary to restrict ourselves to work only in a subspace of $\mathcal{T}_\psi\mathcal{M}$. We now see that the definition we gave previously of always defining $\boldsymbol{\Omega}$ as the pseudo-inverse with respect to $\boldsymbol{g}$ coincides with always choosing to consider only the tangent directions in

$$\overline{\mathcal{T}_\psi\mathcal{M}} = \mathrm{span}_{\mathbb{R}}\{|\overline{V}_i\rangle\} \tag{2.44}$$

In order to apply variational methods as explained in the following chapters, it may be necessary to at least locally restore the Kähler property. We can achieve this by locally further restricting ourselves to

$$\overline{\overline{\mathcal{T}_\psi\mathcal{M}}} = \mathrm{span}_{\mathbb{R}}\{|\overline{\overline{V}}_i\rangle\}\,. \tag{2.45}$$

Using the bases $\{|\overline{\overline{V}}_\mu\rangle\}$ and $\{|\overline{V}_\mu\rangle\}$, we can define the restricted Kähler structures $(\overline{\overline{\boldsymbol{g}}}, \overline{\overline{\boldsymbol{\omega}}}, \overline{\overline{\boldsymbol{J}}})$, which are compatible, and $(\overline{\boldsymbol{g}}, \overline{\boldsymbol{\omega}}, \overline{\boldsymbol{J}})$, where $\overline{\boldsymbol{\omega}}$ and $\overline{\boldsymbol{J}}$ are non-degenerate.

Our assumption on the definition of $\boldsymbol{\Omega}$ can be understood as taking $\boldsymbol{\Omega}$ to be zero on the subspace $\mathcal{D}_\psi\mathcal{M}$, where $\boldsymbol{\omega}$ is not invertible, and equal to the inverse of $\overline{\boldsymbol{\omega}}$ on $\overline{\mathcal{T}_\psi\mathcal{M}}$. Note that this definition is only possible because the tangent space is also equipped with a metric $\boldsymbol{g}$, which makes the orthogonal decomposition $\mathcal{T}_\psi\mathcal{M} = \overline{\mathcal{T}_\psi\mathcal{M}} \oplus \mathcal{D}_\psi\mathcal{M}$ well-defined.

In summary, a general variational family $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$ is not necessarily a Kähler manifold. We can check locally, if the restricted Kähler structures fail to satisfy the Kähler condition. If this happens, we can always choose local subspaces

$$\overline{\overline{\mathcal{T}_\psi\mathcal{M}}} \subset \overline{\mathcal{T}_\psi\mathcal{M}} \subset \mathcal{T}_\psi\mathcal{M} \tag{2.46}$$

on which the restricted Kähler structures satisfy the Kähler properties or are at least invertible. Defining $\boldsymbol{\Omega}$ as the pseudo-inverse with respect to $\boldsymbol{g}$ is equivalent to inverting $\boldsymbol{\omega}$ only on $\overline{\mathcal{T}_\psi\mathcal{M}}$. In what follows, we therefore do not need to distinguish between the non-Kähler cases with degenerate or non-degenerate structures, as we will always be able to apply the same variational techniques based on $\boldsymbol{\Omega}$.

2.6 Observables and Poisson brackets

Any Hermitian operator $\hat{A}$ defines a real-valued function $\langle\hat{A}\rangle$ on the manifold $\mathcal{M}$ and in fact on the whole projective Hilbert space. The function is given by the expectation value

$$A(x) = \langle\hat{A}\rangle(x) = \frac{\langle\psi(x)|\hat{A}|\psi(x)\rangle}{\langle\psi(x)|\psi(x)\rangle}\,. \tag{2.47}$$

It is invariant under rescalings of $|\psi\rangle$ by complex factors and is thus a well-defined map on projective Hilbert space $\mathcal{P}(\mathcal{H})$. We will use the notation $\langle\hat{A}\rangle$ and $A(x)$ interchangeably. For the function deriving from the Hamiltonian operator $\hat{H}$, we use the symbol $E = \langle\hat{H}\rangle$ and call it the *energy*.

Given a Hermitian operator $\hat{A}$ and the representative $|\psi(x)\rangle$, we have the important relation

$$\mathbb{P}^\mu_\psi \hat{A}\,|\psi\rangle = \boldsymbol{G}^{\mu\nu}(\partial_\nu A)\,, \tag{2.48}$$

which is invariant under the change of representative $|\psi\rangle \to c\,|\psi\rangle$ and $|V_\mu\rangle \to c\,|V_\mu\rangle$. It follows from

$$\partial_\mu A = \frac{2\,\mathrm{Re}\langle V_\mu|\,\hat{A}\,|\psi\rangle}{\langle\psi|\psi\rangle} = \boldsymbol{g}_{\mu\nu}\mathbb{P}^\nu_\psi\hat{A}\,|\psi\rangle\ , \tag{2.49}$$

Table 2.1: *Comparison: Kähler vs. Non-Kähler.* We review the properties of restricted Kähler structures and their inverses in each case. See appendix A.2 for a review of the conditions for a general manifold to be Kähler.

	Kähler	**Non-Kähler**	
		(non-degenerate)	**(degenerate)**
Metric $\boldsymbol{g}$	symmetric, positive definite	symmetric, positive definite,	
inverse $\boldsymbol{G}$	invertible	invertible	
Symplectic form $\boldsymbol{\omega}$	antisymmetric, closed ($d\omega = 0$)	antisymmetric, may not be closed	
inverse $\boldsymbol{\Omega}$ (or pseudo inverse)	non-degenerate	non-degenerate	degenerate
Complex structure $\boldsymbol{J}$	$\boldsymbol{J}^2 = -\mathbb{1}$,	$0 \geq \boldsymbol{J}^2 \geq -\mathbb{1}$,	
inverse $\boldsymbol{J}^{-1} = -\boldsymbol{\Omega g}$ (or pseudo inverse)	invertible with $\boldsymbol{J}^{-1} = -\boldsymbol{J}$	invertible	pseudo-invertible

where we used product rule and (2.24).

The following definition will play an important role in the context of Poisson brackets and conserved quantities. Every operator $\hat{A}$ defines a vector field given by $\mathbb{Q}_\psi \hat{A} \left|\psi\right\rangle$. If this vector field is tangent to $\mathcal{M}$ for all $\psi \in \mathcal{M}$, the following definition applies.

Definition 3. *Given a general operator $\hat{A}$ and a variational family $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$, we say $\hat{A}$ **preserves** $\mathcal{M}$ if*

$$\mathbb{Q}_\psi \hat{A} \left|\psi\right\rangle = (\hat{A} - \langle \hat{A} \rangle) \left|\psi\right\rangle \quad \textit{for all} \quad \psi \in \mathcal{M} \tag{2.50}$$

lies in the tangent space $\mathcal{T}_\psi \mathcal{M}$, i.e., $\mathbb{Q}_\psi \hat{A} \left|\psi\right\rangle = \mathbb{P}_\psi \hat{A} \left|\psi\right\rangle$.

The symplectic structure of the manifold naturally induces a Poisson bracket on the space of differentiable functions, which is given by

$$\{A, B\} := (\partial_\mu A)\, \boldsymbol{\Omega}^{\mu\nu} (\partial_\nu B)\,. \tag{2.51}$$

In some special cases this can be related to the commutator of the related operators.

Proposition 3. *Given two Hermitian operators $\hat{A}$ and $\hat{B}$ of which one preserves the Kähler manifold $\mathcal{M}$, i.e.,*

$$(\hat{A} - \langle \hat{A} \rangle) \left|\psi\right\rangle \in \mathcal{T}_\psi \mathcal{M} \textit{ or } (\hat{B} - \langle \hat{B} \rangle) \left|\psi\right\rangle \in \mathcal{T}_\psi \mathcal{M}\,, \tag{2.52}$$

the Poisson bracket is related to the commutator via

$$\{A, B\} = \mathrm{i} \frac{\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle}{\langle \psi | \psi \rangle}\,. \tag{2.53}$$

Proof. We compute

$$\mathrm{i} \tfrac{\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle}{\langle \psi | \psi \rangle} = \tfrac{2\,\mathrm{Re} \langle \psi | (\hat{A} - \langle \hat{A} \rangle) \mathrm{i} (\hat{B} - \langle \hat{B} \rangle) | \psi \rangle}{\langle \psi | \psi \rangle}\,. \tag{2.54}$$

As one of the vectors $(\hat{A} - \langle \hat{A} \rangle) \left|\psi\right\rangle$ or $(\hat{B} - \langle \hat{B} \rangle) \left|\psi\right\rangle$ lies in the tangent space $\mathcal{T}_\psi \mathcal{M}$, (2.7) applies, giving

$$\begin{aligned} \mathrm{i} \tfrac{\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle}{\langle \psi | \psi \rangle} &= \mathbb{P}^\mu_\psi \hat{A} \left|\psi\right\rangle \, \boldsymbol{g}_{\mu\nu} \, \mathbb{P}^\nu_\psi \mathrm{i} \hat{B} \left|\psi\right\rangle \\ &= \partial_\nu A \; \boldsymbol{J}^\nu{}_\rho \boldsymbol{G}^{\rho\sigma} \partial_\sigma B = \partial_\nu A \; \boldsymbol{\Omega}^{\nu\sigma} \partial_\sigma B\,, \end{aligned} \tag{2.55}$$

where we used (2.49) and $\boldsymbol{J} = -\boldsymbol{J}^{-1} = \boldsymbol{\Omega} \boldsymbol{g}$ for a Kähler manifold. □

For $\mathcal{M} = \mathcal{P}(\mathcal{H})$, above conditions are clearly met for any Hermitian operators $\hat{A}$ and $\hat{B}$. For a general Kähler submanifold $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$, however, the validity of (2.53) depends on the choice of operators considered. On a submanifold which is not Kähler the statement is in general no longer valid.

Chapter 3

Variational principles

In the previous chapter we have defined variational manifolds. In this chapter we will now address how to apply these manifolds to the computation of properties of quantum mechanical systems. The approaches which underlie these calculations take the form of variational principles. We will show that they can be used to compute several static, dynamical and spectral properties of quantum Hamiltonians.

Similarly to the previous chapter, much of the content of this chapter is derived from reference [69]:
L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, J.I. Cirac,
Geometry of variational methods: dynamics of closed quantum systems,
SciPost Physics 9, 48 (2020), used under CC BY 4.0.

3.1 Overview

Given a variational manifold $\mathcal{M}$ defined according to the observations of Chapter 2, we can define several variational principles, which allow us to compute information about a model of interest, while relying only on the use of states contained in $\mathcal{M}$. As discussed in Section 1.3, this is to be understood essentially as restricting the laws of quantum mechanics to a non-linear subspace $\mathcal{M}$ of the Hilbert space. In this chapter we will discuss four sets of variational principles which are particularly useful for the study of quantum many body systems.

In Section 3.2 we will focus on variational principles for real time evolution. These are often also referred to as Time Dependent Variational Principles. The aim of these principles is to define an approximation of the Schrödinger time evolution that is restricted to take place only inside the manifold $\mathcal{M}$. For this reason we will refer to it as *projected real time evolution*.

As we will see, there exist two fundamentally different principles to define this

projected real time evolution. There is one based on defining a global Lagrangian action principle on the variational manifold. We will refer to this as the *Lagrange* principle. There is a second one based on minimising a local error measure, which we will refer to as the *McLachlan* principle. We will discuss how these two principles are in general inequivalent and how they coincide only in those cases in which the manifold $\mathcal{M}$ is a Kähler manifold, according to the definition given in the previous chapter. Finally, we will mention how there also exists a principle that in the literature often goes by the name of *Dirac-Frenkel* principle. This principle, however, is well-defined only in cases in which the other two are equivalent and then coincides with them.

At the end of the section, we will discuss two other related questions, namely how conserved quantities of the dynamics are affected by the projection of time evolution on $\mathcal{M}$ and how to take care correctly of the dynamics of the global phase and normalisation of a state.

In Section 3.3 we will then move on to the use of variational principles to estimate the spectrum of excitations of a many body model. We will present two alternative methods, both based on a variational manifold $\mathcal{M}$ and both with some successful applications. The *projection* method is based on constructing and diagonalising a restricted version of the system's Hamiltonian, defined only on the low-dimensional tangent plane of $\mathcal{M}$. The *linearisation* methods is based on finding the fundamental oscillation frequencies of the system around its ground state: we achieve this through a linearisation of the projected real time evolution discussed previously.

Ultimately, the two methods provide different approximations to the system's energy spectrum, with different characteristic properties. We will show the relationship between the two methods and discuss when each can be most advantageous. In summary, the projection methods provides truly variational bounds on the energy spectrum, while the linearisation method can reflect better the physically relevant properties of the system, especially when the natural symmetries of the manifold and the system match each other appropriately.

In Section 3.4 we then address the computation of a quantity that is of great importance in quantum many body physics, namely the *spectral function* of the system. This function is of great relevance because it represents an experimentally measurable susceptibility, in particular in experiments where the system is perturbed with a small pump field and then its response is probed. In this section we present a method for estimating this quantity based on performing linear response theory directly on the variational manifold $\mathcal{M}$.

Finally, in Section 3.5 we will present a method for finding within a manifold $\mathcal{M}$ the state with the minimal energy expectation value, *i.e.*, the state representing an approximate ground state. This method goes by the name of *projected imaginary*

time evolution and can be understood as a form of gradient descent method which also takes into account some information about the geometry and curvature of the manifold on which the optimisation is performed. Indeed, it is sometimes also referred to as *natural* gradient descent.

Throughout this chapter we will always stress the importance of the underlying geometric principles for the understanding of each variational principle. In particular, in each section we will highlight how the presented methods can lead to different outcomes depending on whether the manifold $\mathcal{M}$ has a Kähler or non-Kähler geometry.

3.2 Real time evolution

For what concerns real time evolution, we would like to approximate the Schrödinger equation

$$\mathrm{i}\frac{d}{dt}\left|\psi\right\rangle = \hat{H}\left|\psi\right\rangle \tag{3.1}$$

on our variational manifold $\mathcal{M}$. There are different principles, used extensively in the literature, according to which this approximation can be performed. We will see that only in the case of Kähler manifolds they are all equivalent.

3.2.1 Variational principles

Following the literature, we can define the following variational principles for $|\psi\rangle \equiv |\psi(t)\rangle$.

Lagrangian action principle [82]. The most commonly used variational principle relies on defining a Lagrangian action

$$\mathcal{S} = \int_{t_i}^{t_f} \mathcal{L}\, dt = \int_{t_i}^{t_f} dt\ \mathrm{Re}\,\frac{\langle\psi|(\mathrm{i}\frac{d}{dt} - \hat{H})|\psi\rangle}{\langle\psi|\psi\rangle}\,. \tag{3.2}$$

Requiring that $|\psi(t)\rangle$ extremises this action leads to

$$0 = \mathrm{Re}\,\langle\mathbb{Q}_\psi\delta\psi|(\mathrm{i}\tfrac{d}{dt} - \hat{H})|\psi\rangle \tag{3.3}$$

for all times and all allowed variations $|\delta\psi(t)\rangle$ with $\mathbb{Q}_\psi\left|\delta\psi\right\rangle = \left|\delta\psi\right\rangle - \frac{\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle}\left|\psi\right\rangle$ from (2.19). This is equivalent to Schrödinger's equation on projective Hilbert space[1].

[1]The fact that the projector $\mathbb{Q}_\psi$ onto projective tangent space $\mathcal{H}_\psi^\perp$ appears, shows that the resulting dynamics is defined on projective Hilbert space, while global phase and normalisation are left undetermined. We will explain how to recover the dynamics of phase and normalisation in Section 3.2.3.

If we restrict ourselves to having the dynamics only taking place on a variational manifold $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$, we have that the allowed variations in (3.3) must satisfy $\mathbb{Q}_\psi |\delta\psi(t)\rangle \in \mathcal{T}_{\psi(t)}\mathcal{M}$. So, we have instead

$$\mathbb{P}_\psi(\mathrm{i}\tfrac{d}{dt} - \hat{H})\,|\psi\rangle \implies \mathbb{P}_\psi \mathrm{i}\,\tfrac{d}{dt}\,|\psi\rangle = \mathbb{P}_\psi \hat{H}\,|\psi\rangle\ . \tag{3.4}$$

This gives rise to equations of motion which we will derive in Proposition 4. For a time-independent Hamiltonian, they always preserve the energy expectation value.

McLachlan minimal error principle [83]. Alternatively, we can try to minimise the error between the approximate trajectory and the true solution. As we do not know the latter, we cannot compute the total error, but at least we can quantify the local error in state norm

$$\left\| \tfrac{d}{dt}\,|\psi\rangle - (-\mathrm{i}\hat{H})\,|\psi\rangle \right\| , \tag{3.5}$$

made at each point in time due to imposing that $\frac{d}{dt}|\psi(x)\rangle$ represents a variation tangent to the manifold, *i.e.*, $\mathbb{Q}_\psi \frac{d}{dt}|\psi(x)\rangle \in \mathcal{T}_\psi\mathcal{M}$. According to (2.27), this error is minimised if

$$\mathbb{Q}_\psi \tfrac{d}{dt}\,|\psi\rangle = -\mathbb{P}_\psi \mathrm{i}\hat{H}\,|\psi\rangle\ . \tag{3.6}$$

This gives rise to equations of motion which we will derive in Proposition 5. The resulting equations of motion only agree with the Lagrangian action if $\mathcal{M}$ is a Kähler manifold. Otherwise, they may not preserve the energy expectation value.

Dirac-Frenkel variational principle [84, 85]. Another variational principle requires

$$\langle\delta\psi|(\mathrm{i}\tfrac{d}{dt} - \hat{H})|\psi\rangle = 0 \tag{3.7}$$

for all allowed variations $|\delta\psi(t)\rangle$. It is easy to see that the real and imaginary parts of (3.7) are equivalent to (3.4) and (3.6) respectively. Therefore, this principle is well-defined (and equivalent to the other two) only in the cases in which they are equivalent between themselves, that is, as we will see, if and only if $\mathcal{M}$ is a Kähler manifold. Otherwise, the resulting equations will be overdetermined.

Expressing equations (3.4) and (3.6) in coordinates leads to flow equations for the manifold parameters $x(t)$. We can then define a real time evolution vector field $\mathcal{X}$ everywhere on $\mathcal{M}$, such that

$$\frac{dx^\mu}{dt} = \mathcal{X}^\mu(x)\,. \tag{3.8}$$

The schematic idea is illustrated in Figure 3.1. Integrating such equations defines the flow map Φ_t that maps an initial set of coordinates $x(0)$ to the values $x(t)$ that they assume after evolving for a time t.

In the case of the Lagrangian action principle, the vector field $\mathcal{X}$ takes the form given in the following proposition. A similar derivation was also considered in [82]. Note that here, as well as elsewhere in this chapter, we will be using Einstein's summation notation for repeated indices.

Proposition 4. *The real time evolution projected according to the* ***Lagrangian action principle*** (3.4) *is*

$$\frac{dx^\mu}{dt} \equiv \mathcal{X}^\mu = -\boldsymbol{\Omega}^{\mu\nu}(\partial_\nu E)\,. \quad \textbf{(Lagrangian)} \tag{3.9}$$

where $E(x)$ is the energy function, defined in the context of equation (2.47)*. Such evolution always conserves the energy expectation value.*

Proof. From the definition (2.24) of the tangent space basis, we have

$$\tfrac{d}{dt}\,|\psi\rangle = \dot{x}^\mu\,\partial_\mu\,|\psi\rangle = \dot{x}^\mu\,|V_\mu\rangle + \frac{\langle\psi|\frac{d}{dt}\psi\rangle}{\langle\psi|\psi\rangle}\,|\psi\rangle\ . \tag{3.10}$$

We substitute this in (3.4) and then expand the projectors using the relations (2.28), (2.30) and $\mathbb{P}_\psi \mathrm{i}\,|\psi\rangle = 0$ to obtain

$$\boldsymbol{J}^\mu{}_\nu \mathcal{X}^\nu = \boldsymbol{G}^{\mu\rho}\,\frac{2\mathrm{Re}\langle V_\rho|\,\hat{H}\,|\psi\rangle}{\langle\psi|\psi\rangle}\,. \tag{3.11}$$

We further simplify the expression by using (2.49) and $(\boldsymbol{J}^{-1})^\mu{}_\nu = -\boldsymbol{\Omega}^{\mu\rho}\boldsymbol{g}_{\rho\nu}$ from (2.30). This leads to

$$\mathcal{X}^\mu = (\boldsymbol{J}^{-1})^\mu{}_\nu \boldsymbol{G}^{\nu\sigma}\partial_\sigma E = -\boldsymbol{\Omega}^{\mu\nu}\partial_\nu E\,. \tag{3.12}$$

To obtain the variation of the energy expectation value E we compute directly

$$\frac{dE}{dt} = (\partial_\mu E)\frac{dx^\mu}{dt} = -(\partial_\mu E)\boldsymbol{\Omega}^{\mu\nu}(\partial_\nu E) = 0\,, \tag{3.13}$$

where we used the antisymmetry of $\boldsymbol{\Omega}^{\mu\nu}$. If $\boldsymbol{J}$ (and thus also $\boldsymbol{\Omega}$) is not invertible, one needs to restrict to an appropriate subspace. □

The most important lesson of (3.9) is that projected time evolution on a Kähler manifold is equivalent to classical Hamiltonian evolution with respect to energy function $E(x)$. As was pointed out in [86], already the time evolution in full projective Hilbert space, *i.e.*, $\mathcal{M} = \mathcal{P}(\mathcal{H})$, follows the classical Hamilton equations if we use the natural symplectic form $\boldsymbol{\Omega}^{\mu\nu}$. Let us point out that the sign in equation (3.9) depends on the convention chosen for the symplectic form, which in classical mechanics differs from the one adopted here. One further consequence of equation (3.9) is that the real time evolution vector field $\mathcal{X}(x)$ vanishes in

stationary points of the energy, that is points x_0 such that $\partial_\mu E(x_0) = 0$. These points will therefore also be stationary points of the evolution governed by $\mathcal{X}$.

Let us here recall that, if $\boldsymbol{\omega}_{\mu\nu}$ is not invertible, $\boldsymbol{\Omega}^{\mu\nu}$ refers to the pseudo-inverse, as discussed in Sections 2.4 and 2.5. This convention means that in practice the Lagrangian evolution will always take place in the submanifold of $\mathcal{M}$ on which $\boldsymbol{\omega}$ is invertible. There may be pathological cases where $\boldsymbol{\omega}$ vanishes on the whole tangent space and therefore the Lagrangian principle does not lead to any evolution.

In the case of the McLachlan minimal error principle, the evolution equations take the form given in the following proposition, which cannot be simplified further. It is also in general not true that this evolution conserves the energy or that it has a stationary point in energy minima.

Proposition 5. *The real time evolution projected based on the* ***McLachlan minimal error principle*** (3.6) *is*

$$\frac{dx^\mu}{dt} \equiv \mathcal{X}^\mu = -\frac{2\boldsymbol{G}^{\mu\nu}\mathrm{Re}\langle V_\nu|\mathrm{i}\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle}\,. \quad \textbf{(McLachlan)} \tag{3.14}$$

Proof. By substituting (2.24) in (3.6), analogously to what was done in (3.10), we have

$$\dot{x}^\mu = \mathbb{P}^\mu_\psi(-\mathrm{i}\hat{H}\,|\psi\rangle)\,, \tag{3.15}$$

from which the proposition follows by expanding the projector according to (2.28). □

To perform real time evolution in practice, either based on (3.9) for Lagrangian evolution or based on (3.14) for McLachlan evolution, we will typically employ a numerical integration scheme [87, 88] to evolve individual steps. It is generally hard to get rigorous bounds on the resulting error that increases over time, but in certain settings there still exist meaningful estimates [89]. Let us now relate the different variational principles, which has also been discussed in [90].

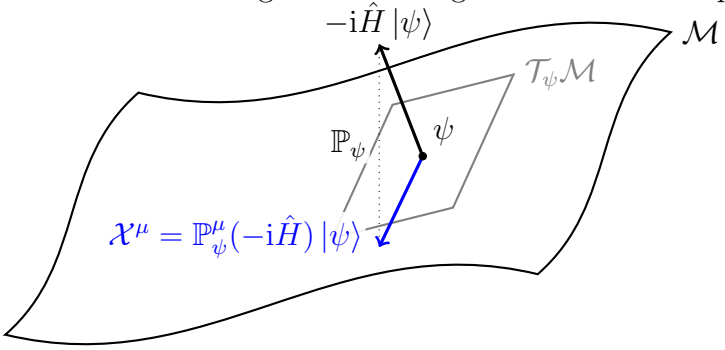
Proposition 6. *The Lagrangian, the McLachlan and the Dirac-Frenkel variational principle are equivalent if the variational family is Kähler.*

Proof. To prove the statement, it is sufficient to see that equations (3.4) and (3.6) can be written simply as applying the tangent space projector $\mathbb{P}_\psi$ to two different forms of the Schrödinger equation, *i.e.*,

$$\textbf{Lagrangian:} \qquad \mathbb{P}_\psi(\mathrm{i}\tfrac{d}{dt} - \hat{H})\,|\psi\rangle = 0 \tag{3.16}$$

$$\textbf{McLachlan:} \qquad \mathbb{P}_\psi(\tfrac{d}{dt} + \mathrm{i}\hat{H})\,|\psi\rangle = 0. \tag{3.17}$$

Figure 3.1: *Real time evolution.* We illustrate real time evolution on a variational manifold $\mathcal{M}$ according to the Dirac-Frenkel variational principle (when Lagrangian and McLachlan principles coincide). The time evolution vector $-\mathrm{i}\hat{H}\,|\psi\rangle$ at a state ψ is orthogonally projected through $\mathbb{P}_\psi$ onto the variational manifold $\mathcal{M}$ to define the vector field $\mathcal{X}^\mu$. The resulting evolution is guaranteed to take place on $\mathcal{M}$.



These two forms only differ by a factor of i. However, as we discussed in Proposition 2, one equivalent way to define the Kähler property of our manifold is that multiplication by i commutes with the projector $\mathbb{P}_\psi$. Therefore, if the manifold is Kähler, an imaginary unit can be factored out of equations (3.16) and (3.17) making them coincide. If, on the other hand, the manifold is non-Kähler, this operation is forbidden and they are in general not equivalent. □

As discussed in Section 2.5, if the chosen manifold does not respect the Kähler condition, we always have the choice to locally restrict ourselves to consider only a subset of tangent directions with respect to which the manifold is again Kähler, *i.e.*, $\overline{\overline{\mathcal{T}_\psi\mathcal{M}}}$. Then both principles will again give the same equation of motion, which will have the same form as (3.9) where we just replace $\mathbf{\Omega}^{\mu\nu}$ with $\overline{\overline{\mathbf{\Omega}}}^{\mu\nu}$, which will conserve the energy and have stationary points in the minima of the energy. We will refer to this procedure as *Kählerisation*.

We can compute explicitly how the vector fields of the Lagrangian and McLachlan variational principles differ. For this, we only consider the subspaces, defined in Proposition 1, in which the complex structure fails to be Kähler, *i.e.*, where we have

$$J \equiv \bigoplus_i \begin{pmatrix} & c_i \\ -c_i & \end{pmatrix}, \tag{3.18}$$

as in (2.41). If we enlarge the tangent space to include all vectors $\mathrm{i}\,|V_\mu\rangle$, the

enlarged complex structure will have the form

$$\check{J} \equiv \bigoplus_i \begin{pmatrix} & c_i & & \sqrt{1-c_i^2} \\ -c_i & & \sqrt{1-c_i^2} & \\ & -\sqrt{1-c_i^2} & & c_i \\ -\sqrt{1-c_i^2} & & -c_i & \end{pmatrix} \tag{3.19}$$

which clearly satisfies $\check{J}^2 = -\mathbb{1}$. As the enlarged space is Kähler, we can unabiguoulsy define a time evolution vector field $\check{\mathcal{X}} \equiv \oplus_i(a_i, b_i, \alpha_i, \beta_i)$ on the enlarged space, and we find the two distinct restrictions

$$\begin{aligned} \mathcal{X}_{\text{Lagrangian}} &= J^{-1}\mathbb{P}_\psi \check{J}\check{\mathcal{X}} \\ &\equiv \oplus_i \left(a_i - \tfrac{\sqrt{1-c_i^2}}{c_i}\alpha_i, b_i + \tfrac{\sqrt{1-c_i^2}}{c_i}\beta_i \right) \end{aligned} \tag{3.20}$$

$$\mathcal{X}_{\text{McLachlan}} = \mathbb{P}_\psi \check{\mathcal{X}} \equiv \oplus_i (a_i, b_i)\,, \tag{3.21}$$

associated to the Lagrangian and the McLachlan principle, respectively. We see explicitly that they agree for $c_i = 1$, but also when $\alpha_i = \beta_i = 0$.

Kähler vs. non-Kähler. On a Kähler manifold all three variational principles are well-defined and equivalent. They all give the same energy conserving equations of motion (3.9). On a non-Kähler manifold, only the Lagrangian and McLachlan variational principles are well-defined, but they give in general inequivalent equations of motion given by (3.9) and (3.14). Only the Lagrangian ones will manifestly conserve the energy and have stationary points in the minima of the energy. In table 3.1, we review advantages and drawbacks discussed in the following sections. While in most cases, the Lagrangian principle appears to be a natural choice, the McLachlan principle is often preferable if $\boldsymbol{\omega}$ is highly degenerate—in particular, if $\boldsymbol{\omega} = 0$ its pseudo-inverse is $\boldsymbol{\Omega} = 0$ and the evolution would vanish everywhere independent of $\hat{H}$, such that the McLachlan principle appears to be the better choice.

3.2.2 Conserved quantities

Given the generator $\hat{A}$ of a symmetry of the Hamiltonian $\hat{H}$, *i.e.*, $[\hat{H}, \hat{A}] = 0$, the expectation value $A(t) = \langle\psi_t| \hat{A} |\psi_t\rangle$ is necessarily preserved by unitary time evolution on the full Hilbert space

$$|\psi_t\rangle = U(t)\,|\psi_0\rangle = e^{-\mathrm{i}\hat{H}t}\,|\psi_0\rangle\ . \tag{3.22}$$

We now consider if this continues to be true for projected time evolution on a manifold.

Table 3.1: *Action principles.* We review the different action principles and how they relate to the respective manifolds.

	Lagrangian	**McLachlan**	**Dirac-Frenkel**
Definition	$\mathbb{P}_\psi(\mathrm{i}\frac{d}{dt} - \hat{H})\,\lvert\psi\rangle = 0$	$\mathbb{P}_\psi(\frac{d}{dt} + \mathrm{i}\hat{H})\,\lvert\psi\rangle = 0$	both
Kähler manifold	always defined and all equivalent		
Non-Kähler manifold	defined for chosen inverse $\boldsymbol{\Omega}$ (see proposition 4)	always defined (see proposition 5)	not defined
Advantage	energy conservation (see proposition 4)	conservation of symmetries (see proposition 7)	both
Linearisation around ground state	possible (see section 3.3.2)	not possible (see section 3.3.2)	possible

For a time-independent Hamiltonian $\hat{H}$, we have seen that the energy expectation value E is always conserved by Lagrangian projected real time evolution. However, projected time evolution will not in general preserve expectation values of an operator $\hat{A}$ with $[\hat{H}, \hat{A}] = 0$. To guarantee this, one has to further require that $\hat{A}$ preserves the manifold.

Proposition 7. *Given a variational manifold $\mathcal{M}$ and a Hermitian operator $\hat{A}$, such that $[\hat{H}, \hat{A}] = 0$ and $\hat{A}$ preserves the manifold in the sense of Definition 3, i.e.,*

$$\mathbb{Q}_\psi \hat{A}\,\lvert\psi\rangle = (\hat{A} - \langle\hat{A}\rangle)\,\lvert\psi\rangle \in \mathcal{T}_\psi\mathcal{M} \quad \forall\psi \in \mathcal{M}\,, \tag{3.23}$$

*the expectation value $A(x(t))$, defined as in equation (2.47), is preserved under real time evolution projected according to the **McLachlan** variational principle. It is also true for **Lagrangian** variational principle, if the two principles agree, i.e., if the manifold is Kähler.*

Proof. We compute

$$\begin{aligned}\frac{d}{dt}A(t) = (\partial_\mu A)\,\frac{dx^\mu}{dt} &= \mathbb{P}^\nu_\psi \hat{A}\,\lvert\psi\rangle\, \boldsymbol{g}_{\nu\mu}\, \mathbb{P}^\mu_\psi(-\mathrm{i}\hat{H}\,\lvert\psi\rangle) \\ &= \tfrac{2\mathrm{Re}\langle\psi\rvert(\hat{A}-\langle\hat{A}\rangle)(-\mathrm{i}\hat{H})\lvert\psi\rangle}{\langle\psi\rvert\psi\rangle} = \tfrac{\mathrm{i}\langle\psi\rvert[\hat{H},\hat{A}]\lvert\psi\rangle}{\langle\psi\rvert\psi\rangle} = 0,\end{aligned} \tag{3.24}$$

where in the first line we used relation (2.49) for the gradient of A, the definition of McLachlan evolution (3.6) and that $\mathbb{P}^{\mu}_{\psi} \langle \hat{A} \rangle |\psi\rangle = 0$, in the second step we used that, thanks to the condition (3.23), the restricted bilinear form $\boldsymbol{g}$ in the first line coincides with the full Hilbert space one in the second line of (3.24). □

This result only applies the McLachlan projected real time evolution, for which the equation of motion (3.15) holds. In the Lagrangian case, we would have

$$\dot{A} = (\partial_\mu A)\, \mathcal{X}^\mu = -(\partial_\mu A)\, \Omega^{\mu\nu} \partial_\nu E = \{E, A\}, \tag{3.25}$$

which is in general not equal to $\mathrm{i} \langle\psi|[\hat{H}, \hat{A}]|\psi\rangle \langle\psi|\psi\rangle^{-1}$ on a non-Kähler manifold[2] and thus not necessarily zero.

We see here the main advantage of the *Kählerisation* procedure described in the previous section. Indeed, through *Kählerisation* we are able to define, even for general non-Kähler manifolds, a projected real time evolution that shares the desirable properties of both, the Lagrangian and the McLachlan projections, *i.e.*, it is a symplectic, energy preserving evolution with stationary points in the energy minima and at the same time preserves the expectation value of symmetry generators satisfying (3.23). Note that Kählerisation may spoil the conservation laws of observables $\hat{A}$, for which $\mathbb{Q}_\psi \hat{A} |\psi\rangle$ does not lie in the Kähler subspace $\overline{\overline{\mathcal{T}_\psi \mathcal{M}}}$, in which case we will need to enforce conservation by hand, discussed next.

For operators $\hat{A}$ where (3.23) is not satisfied, we have two options to correct for this:

(a) Enlarge the variational manifold $\mathcal{M}$, such that condition (3.23) is satisfied.

(b) Enforce conservation by hand, for which we modify the projected time evolution vector field $\mathcal{X}^\mu$.

While option (a) is typically more desirable, it requires creativity to find a suitable extension of a given family $\mathcal{M}$. Of course, if $\hat{A}$ is an important physical observable that is relevant to the problem, a manifold that does not preserve it may not be a good choice to approximate the system's behavior. In practice, however, it may still be worthwhile to check the predictions of an approximated time-evolution adopting (b).

This is done by adding a further projection of the real time evolution flow onto the subspace of the tangent space orthogonal (with respect to $\boldsymbol{g}$) to the direction $\mathbb{P}^{\mu}_{\psi} \hat{A} |\psi\rangle = \boldsymbol{G}^{\mu\nu} \partial_\nu A$. This is equivalent to restricting ourselves to the submanifold

$$\widetilde{\mathcal{M}} = \left\{ |\psi\rangle \in \mathcal{M} \,\middle|\, \tfrac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} = A_0 \right\} \subset \mathcal{M}, \tag{3.26}$$

[2]For Kähler manifolds, as discussed in the context of Proposition 3, $\{F, G\} = \mathrm{i} \langle\psi|[\hat{F}, \hat{G}]|\psi\rangle \langle\psi|\psi\rangle^{-1}$ only holds if either $\hat{F}$ or $\hat{G}$ preserves the manifold.

where A_0 is the initial value $\langle\psi(0)|\hat{A}|\psi(0)\rangle \langle\psi(0)|\psi(0)\rangle^{-1}$. Note that this modified evolution may spoil other conservation laws (*e.g.*, energy) that were previously intact.

To preserve several quantities $\hat{A}_I$, we can project onto the subspace orthogonal to the span of $X_I = \mathbb{P}_\psi \hat{A}_I |\psi\rangle$. If we also want to preserve the Kähler property, we should choose $X_I = (\mathbb{P}_\psi \hat{A}_I |\psi\rangle, \mathrm{i}\mathbb{P}_\psi \hat{A}_I |\psi\rangle)$. We can then define $\widetilde{g}_{IJ} = X_I^\mu \boldsymbol{g}_{\mu\nu} X_J^\nu$ to define the projector

$$\widetilde{P}^\mu{}_\nu = \delta^\mu{}_\nu - X_I^\mu \widetilde{G}^{IJ} X_J^\rho \boldsymbol{g}_{\rho\nu}\,, \tag{3.27}$$

where $\widetilde{G}^{IJ}$ is the inverse (or pseudo inverse, if not all vectors X_I are linearly independent) of $\widetilde{g}_{IJ}$.

The modified Lagrangian evolution vector field $\widetilde{\mathcal{X}}^\mu$ is

$$\widetilde{\mathcal{X}}^\mu = -\widetilde{\boldsymbol{\Omega}}^{\mu\nu}(\partial_\nu E) \quad \text{with} \quad \widetilde{\boldsymbol{\Omega}}^{\mu\nu} = \widetilde{P}^\mu{}_\sigma \widetilde{P}^\nu{}_\rho \boldsymbol{\Omega}^{\sigma\rho}\,, \tag{3.28}$$

while for the McLachlan evolution, we find

$$\widetilde{\mathcal{X}}^\mu = \widetilde{P}^\mu{}_\nu \mathcal{X}^\nu\,, \tag{3.29}$$

where $\mathcal{X}^\mu$ represents the unmodified evolution vector field in the McLachlan case. It will conserve all expectation values $A_I(t)$ by construction. In the Lagrangian case also the energy will continue to be conserved by construction, which would need to be enforced by hand for the McLachlan case.

Kähler vs. non-Kähler. On a non-Kähler manifold, where we have two inequivalent definitions of the evolution, only the one coming from the McLachlan principle preserves the expectation value of symmetry generators satisfying (3.23). Thus a key reason to Kählerise a non-Kähler manifold is to conserve these expectation values also in the Lagrangian evolution.

3.2.3 Dynamics of global phase

Up to now we have always considered our variational manifolds $\mathcal{M}$ as submanifolds of projective Hilbert space and thus the tangent space $\mathcal{T}_\psi\mathcal{M}$ as a subspace of $\mathcal{H}_\psi^\perp$. This means all states are only defined up to a complex factor. In practice, our family $\psi(x) \in \mathcal{P}(\mathcal{H})$ will be described by a choice $|\psi(x)\rangle \in \mathcal{H}$, *i.e.*, for every set of parameters x, we will have a Hilbert space vector $|\psi(x)\rangle$ representing the quantum state $\psi(x) \in \mathcal{P}(\mathcal{H})$.

If the parametrisation x happens to contain the global phase or normalisation of the state as an independent parameter, we are overparametrising our family and the evolution equations (3.9) or (3.14) will keep the evolution of these parameters undetermined, leading to some gauge redundancy. This is due to the fact that

normalisation and phase do not change the quantum state that $|\psi(x)\rangle$ represents and our equations of motion only determine the physical evolution of the quantum state and not of its Hilbert space representative.

In those cases in which we actually are interested in computing the dynamics of the global phase and normalisation of a state, we can first of all ensure that they are included in the parametrisation by defining

$$|\Psi(x,\kappa,\varphi)\rangle = e^{\kappa+\mathrm{i}\varphi}\,|\psi(x)\rangle\ , \tag{3.30}$$

where κ and φ are two additional real parameters. If phase or normalisation were already contained in x this will lead to an overparametrisation, but we have already explained how to take care of this in Section 2.4.

Then, on top of the real time evolution equations (3.4) or (3.6), we can obtain equations for these extra parameters by projecting Schrödinger's equation on the corresponding tangent space directions, *i.e.*, $|V_\kappa\rangle = |\Psi\rangle$ and $|V_\varphi\rangle = \mathrm{i}\,|\Psi\rangle$, to find the two equations

$$\mathrm{Re}\,\langle\Psi|(-\mathrm{i}\tfrac{d}{dt}+\hat{H})|\Psi\rangle = 0\,,\quad \mathrm{Re}\,\langle\Psi|\tfrac{d}{dt}+\mathrm{i}\hat{H}|\Psi\rangle = 0\,. \tag{3.31}$$

Equivalently, we can use the Lagrangian action principle to find the same equations by extremizing the alternative action

$$S = \int_{t_i}^{t_f} dt\ \mathrm{Re}\,\langle\Psi(t)|(\mathrm{i}\tfrac{d}{dt}-\hat{H})|\Psi(t)\rangle \tag{3.32}$$

for the full set of parameters (x,κ,φ) rather than $\mathcal{S}$ from (3.2) for only x.

In both cases, the time evolution of $x(t)$ is unchanged, but we find the additional equations

$$\dot{\varphi} = \tfrac{\mathrm{Re}\langle\psi|\mathrm{i}\frac{d}{dt}|\psi\rangle}{\langle\psi|\psi\rangle} - E(t) \quad \text{and} \quad \dot{\kappa} = -\tfrac{\mathrm{Re}\langle\psi|\frac{d}{dt}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{3.33}$$

relating the evolution of phase and normalisation with $|\psi(x(t))\rangle$. Interestingly, the time evolution of κ will ensure that $|\Psi(x,\kappa,\varphi)\rangle$ does not change normalisation.

The procedure can be understood as follows. Global phase and normalisation are conjugate parameters when considering Hilbert space as Kähler space, as can be seen from $|V_\varphi\rangle = \mathrm{i}\,|V_\kappa\rangle$. When considering a variational manifold $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$, we have the following options:

1. When we are only interested in the time evolution of physical states ψ, we must project out the information about global phase and normalisation using $\mathbb{P}^{\mu}_{\psi}$. Consequently, our evolution equations will not determine how to change global phase or normalisation as this information is pure gauge. We followed this philosophy until the current section.

2. When we are also interested in the time evolution of global phase and normalisation, we can always extend $\mathcal{M}$ to include both phase and normalisation as independent parameters. Given a generic parametrisation $|\psi(x)\rangle$, we can extend it to $|\Psi(x,\kappa,\varphi)\rangle$ to ensure that it satisfies the Kähler property in the phase/normalisation subspace. Thus we find unambiguous evolution equations for φ and κ. This is what we explained in the current subsection.

Finally, let us emphasize that using equations (3.31) or extremizing action (3.32) without first ensuring both phase and normalisation are included as independent parameters may lead to unphysical results.

Example 9 (Unnormalised coherent states)**.** *We consider coherent states parametrised as* $|\psi(x)\rangle = e^{\mathrm{i}\varphi(x_1,x_2)}e^{(x_1+\mathrm{i}x_2)\hat{a}^\dagger}|0\rangle$*, where the states are not normalised due to* $\langle\psi(x)|\psi(x)\rangle = e^{x_1^2+x_2^2}$*. We chose intentionally a phase* $\varphi(x_1,x_2)$ *that is not an independent parameter but rather depends on* x*. We further consider the Hamiltonian* $\hat{H}=\omega\hat{a}^\dagger\hat{a}$*. The equation of motion on projective Hilbert space based on the action* (3.2) *are*

$$\dot{x}_1 = \omega x_2 \quad \textit{and} \quad \dot{x}_2 = -\omega x_1\,. \tag{3.34}$$

However, if we use (3.32)*, we find the action*

$$S = \int dt \left(\dot{x}_1 x_2 - \dot{x}_2 x_1 - \tfrac{\partial\varphi}{\partial x_1}\dot{x}_1 - \tfrac{\partial\varphi}{\partial x_2}\dot{x}_2\right) e^{x_1^2+y_1^2}\,, \tag{3.35}$$

which leads to the equations of motion given by

$$(1+x_1^2+x_2^2)(\omega x_1+\dot{x}_2) = (\tfrac{\partial\varphi}{\partial x_2}x_1 - \tfrac{\partial\varphi}{\partial x_1}x_2)\dot{x}_2\,, \tag{3.36}$$

$$(1+x_1^2+y_2^2)(\omega x_2-\dot{x}_1) = (\tfrac{\partial\varphi}{\partial x_1}x_2 - \tfrac{\partial\varphi}{\partial x_2}x_1)\dot{x}_1\,. \tag{3.37}$$

They only agree with (3.34) *in the special case* $\frac{\partial\varphi}{\partial x_2}x_1 - \frac{\partial\varphi}{\partial x_1}x_2 = 0$*.*

3.3 Excitation spectra

We would now like to use a variational family $\mathcal{M}$ to approximate the excitation energies E_i of some eigenstates $|E_i\rangle$ of the Hamiltonian. Typically, we are interested in low energy eigenstates, that is eigenstates close to the groundstate of the Hamiltonian. Suppose then that on $\mathcal{M}$ we are able to find an approximate ground state $|\psi_0\rangle$, that is the state with energy ω_0 that represents the global energy minimum on $\mathcal{M}$ (we will describe a method for finding such state in Section 3.5). Then there are two distinct approaches of deriving a spectrum: the projection of the Hamiltonian and the linearisation of the equations of motion.

3.3.1 Projected Hamiltonian

Given a tangent space $\mathcal{T}_{\psi_0}\mathcal{M}$ at a state $\psi_0 \in \mathcal{M}$, we can approximate the excitation spectrum of the Hamiltonian $\hat{H}$ from its projection onto $\mathcal{T}_{\psi_0}\mathcal{M}$. This is a matrix of dimension M which can be readily diagonalised.

Similarly to the case of real time evolution, we define can two different projections given by

$$\begin{aligned} \boldsymbol{H}^{\mu}{}_{\nu} &= \mathbb{P}^{\mu}_{\psi_0} \hat{H} \, |V_\nu\rangle \ , \qquad \textbf{(Lagrangian)} \\ \boldsymbol{R}^{\mu}{}_{\nu} &= -\mathbb{P}^{\mu}_{\psi_0} \mathrm{i}\hat{H} \, |V_\nu\rangle \ . \quad \textbf{(McLachlan)} \end{aligned} \tag{3.38}$$

On a Kähler manifold $\mathcal{M}$, we will have $\boldsymbol{R}^{\mu}{}_{\nu} = -\boldsymbol{J}^{\mu}{}_{\sigma}\boldsymbol{H}^{\sigma}{}_{\nu}$ and $[\boldsymbol{J}, \boldsymbol{H}] = [\boldsymbol{R}, \boldsymbol{H}] = 0$. In this case, $\boldsymbol{H}$ represents a Hermitian operator on tangent space (which is complex sub Hilbert space) and $\boldsymbol{R}$ is anti-Hermitian. The eigenvalues of $\boldsymbol{H}$ are real and come in pairs $(\omega_\ell, \omega_\ell)$, while the ones of $\boldsymbol{R}$ come are purely imaginary and come in conjugate pairs $(\mathrm{i}\omega_\ell, -\mathrm{i}\omega_\ell)$. The two associated eigenvectors of $\boldsymbol{R}$ are related by multiplication of $\boldsymbol{J}$ and also span the respective eigenspace of $\boldsymbol{H}$.

On a non-Kähler manifold $\mathcal{M}$, the relation between $\boldsymbol{H}$ and $\boldsymbol{R}$ as well as their respective spectra is non-trivial. $\boldsymbol{H}$ will now represent just a symmetric linear operator on the tangent space (understood simply as a linear vector space). Its eigenvalues ω_ℓ are therefore real, although not paired. $\boldsymbol{R}$, on the other hand, will represent an anti-symmetric linear operator and its eigenvalues will appear in conjugate pairs[3], but will not be related to the ones of $\boldsymbol{H}$.

The projected Hamiltonian $\boldsymbol{H}$ represents the full Hamiltonian restricted to the tangent space. The Courant–Fischer–Weyl min-max principle states that the eigenvalues E_ℓ of $\hat{H}$ and the eigenvalues ω_ℓ of $\boldsymbol{H}$ satisfy

$$E_\ell \leq \omega_\ell \leq E_{N-M+\ell} \, , \tag{3.39}$$

with $N = \dim_{\mathbb{R}} \mathcal{H}$ and $M = \dim_{\mathbb{R}} \mathcal{T}_{\psi_0}\mathcal{M}$, where we assume that all eigenvalues are sorted and appear with their multiplicity. Therefore, every approximate eigenvalue ω_ℓ bounds a corresponding true eigenvalue E_i from above. How good this approximation is will highly depend on the choice of variational manifold. Note that the energy differences $\omega_\ell - \omega_0$ instead do not necessarily bound $E_\ell - E_0$, because the ground state energy ω_0 might not be exact, *i.e.*, $\omega_0 > E_0$.

Furthermore, the eigenvalues ω_ℓ are variational in the sense that if X^{μ}_{ℓ} is an eigenvector of $\boldsymbol{H}$ such that $\boldsymbol{H}^{\mu}{}_{\nu} X^{\nu}_{\ell} = \omega_\ell \, X^{\mu}_{\ell}$, then the corresponding Hilbert space vector $|X_\ell\rangle = X^{\mu}_{\ell} \, |V_\mu\rangle$ satisfies

$$\frac{\langle X_\ell | \hat{H} | X_\ell \rangle}{\langle X_\ell | X_\ell \rangle} = \omega_\ell \, . \tag{3.40}$$

[3]This also implies that for an odd-dimensional manifold $\boldsymbol{R}^{\mu}{}_{\nu}$ must have a vanishing eigenvalue, which is a pure artefact of the projection.

Kähler vs. non-Kähler. On a Kähler manifold, $\boldsymbol{H}$ and $\boldsymbol{R}$ are related via $\boldsymbol{R} = -\boldsymbol{J}\boldsymbol{H}$ and they will be the representations of a complex Hermitian and anti-Hermitian operators, respectively. Real eigenvalue pairs $(\omega_\ell, \omega_\ell)$ of $\boldsymbol{H}$ will be related to imaginary eigenvalue pairs $(\mathrm{i}\omega_\ell, -\mathrm{i}\omega_\ell)$ of $\boldsymbol{R}$. On a non-Kähler manifold, the eigenvalues ω_ℓ of $\boldsymbol{H}$ could all be different and unrelated to the ones $\boldsymbol{R}$, which are still imaginary appearing in conjugate pairs.

3.3.2 Linearised equations of motion

A common alternative to projecting the Hamiltonian is to linearise the equations of motion around a fixed point x_0 such that $\mathcal{X}(x_0) = 0$

$$\frac{dx^\mu}{dt} = \mathcal{X}^\mu \quad \Rightarrow \quad \frac{d}{dt}\,\delta x^\mu \simeq \boldsymbol{K}^\mu{}_\nu\,\delta x^\nu \tag{3.41}$$

with $\delta x^\mu = x^\mu - x_0^\mu$ and $\boldsymbol{K}^\mu{}_\nu = \partial_\nu \mathcal{X}^\mu|_{x=x_0}$. Here, δx represents a small perturbation around an approximate ground state x_0. The frequencies ω_ℓ appearing in conjugate pairs $\pm\mathrm{i}\omega_\ell$ in the spectrum of $\boldsymbol{K}$ thus represent the frequencies with which such perturbations oscillate around the ground state and provide an approximation to the excitation energies $E_\ell - E_0$ of the Hamiltonian.

As pointed out in Section 3.2.1, the fixed point x_0 only coincides with the approximate ground state ψ_0 if the real time evolution is defined in terms of the Lagrangian action principle. We thus assume the equations of motion (3.9) based on Lagrangian action principle. In this case, we find

$$\boldsymbol{K}^\mu{}_\nu = \partial_\nu \mathcal{X}^\mu = -\partial_\nu(\boldsymbol{\Omega}^{\mu\rho}\partial_\rho E) = -\boldsymbol{\Omega}^{\mu\rho}(\partial_\rho\partial_\nu E)\,, \tag{3.42}$$

where everything is evaluated at x_0 after taking the derivatives. We used that $\partial_\rho E = 0$ at the fixed point[4].

Provided that ψ_0 is an energy minimum, the bilinear form $h_{\mu\nu} = \partial_\nu\partial_\mu E$ is positive definite. By Williamson's theorem [91], $\boldsymbol{K}$ is diagonalizable and the resulting eigenvalues appear in conjugate pairs $(\mathrm{i}\omega_\ell, -\mathrm{i}\omega_\ell)$. Furthermore, by construction, $\boldsymbol{K}$ is a symplectic generator, that is it satisfies $\boldsymbol{K}\boldsymbol{\Omega} + \boldsymbol{\Omega}\boldsymbol{K}^\intercal = 0$, which implies that $M = e^{\boldsymbol{K}}$ preserves the symplectic form, *i.e.*, $M\boldsymbol{\Omega}M^\intercal = \boldsymbol{\Omega}$.

The geometric meaning of (3.41), is that δx represents a tangent vector $|\delta\psi\rangle = \delta x^\mu\,|V_\mu\rangle$ living in the tangent space $\mathcal{T}_{\psi_0}\mathcal{M}$ at the approximate ground state $|\psi_0\rangle$.

[4]Usually, defining a derivative of a vector field $\mathcal{X}$ requires a way to relate tangent spaces at adjacent points via a so-called connection. The resulting covariant derivative $\nabla_\nu\mathcal{X}^\mu = \partial_\nu\mathcal{X}^\mu + \Gamma^\mu_{\nu\rho}\mathcal{X}^\rho$ will depend on $\Gamma^\mu_{\nu\rho}$ that encodes the connection. In our case $\mathcal{X}^\mu$ vanishes at the fixed point, so that the dependence of $\Gamma^\mu_{\nu\rho}$ drops out and the spectrum of $\boldsymbol{K}^\mu{}_\nu = \nabla_\nu\mathcal{X}^\mu$ at $|\psi_0\rangle$ is canonically defined.

The time evolution of such a tangent vector at a fixed point $|\psi_0\rangle$ is described by the linearised evolution flow[5]

$$d\Phi_t : \mathcal{T}_{\psi_0}\mathcal{M} \to \mathcal{T}_{\psi_0}\mathcal{M}\,. \tag{3.43}$$

$\boldsymbol{K}$ is the generator of the flow $d\Phi_t$ leading to the important relation

$$d\Phi_t = e^{t\boldsymbol{K}}\,, \tag{3.44}$$

which shows that $d\Phi_t$ is symplectic.

Unitary evolution on Hilbert space leads to a flow on projective Hilbert space that preserves all three Kähler structures. However, when we project this flow onto a variational manifold to find $\mathcal{X}$, we will project out the part of the vector field orthogonal to tangent space. When using the Lagrangian action principle, the projected flow will continue to be symplectic, *i.e.*, preserve $\boldsymbol{\Omega}$, but none of the other two Kähler structures[6]. Geometrically, this implies that the trajectories of states near the fixed point ψ_0 will be elliptic rather than circular, when measured with respect to $\boldsymbol{G}$.

Therefore, even if $\mathcal{M}$ is a Kähler manifold, $\boldsymbol{K}$ will in general neither commute with $\boldsymbol{J}$ nor be antisymmetric with respect to G, *i.e.*, satisfy $\boldsymbol{KG} = -\boldsymbol{GK}^\intercal$. This has the following consequences:

- Right-eigenvectors $\mathcal{E}(\lambda)$ with $\boldsymbol{K}^\mu{}_\nu\mathcal{E}^\nu(\lambda) = \lambda\mathcal{E}^\mu(\lambda)$ and left-eigenvectors $\widetilde{\mathcal{E}}(\lambda)$ with $\boldsymbol{K}^\mu{}_\nu\widetilde{\mathcal{E}}_\mu(\lambda) = \lambda\widetilde{\mathcal{E}}_\nu(\lambda)$ are not related via $\mathcal{E}^\mu(\lambda) = \boldsymbol{G}^{\mu\nu}\widetilde{\mathcal{E}}_\nu(\lambda)$, but need to be computed independently. This is important when computing spectral functions in section 3.4.

- There does in general not exist a Hilbert space operator $\hat{K}$, such that $\boldsymbol{K}$ is its restriction in the sense of $\boldsymbol{K}^\mu{}_\nu = \mathbb{P}^\mu_{\psi_0}\hat{K}\,|V_\nu\rangle$ or $\boldsymbol{K}^\mu{}_\nu = \mathbb{P}^\mu_{\psi_0}\mathrm{i}\hat{K}\,|V_\nu\rangle$. Thus, $\boldsymbol{K}$ is not a restriction of a Hamiltonian.

Kähler vs. non-Kähler. On a non-Kähler manifold, where we have two inequivalent definitions of the equations of motion, it only makes sense to linearise the ones coming from the Lagrangian action principle, as their fixed point coincides with the approximate ground state. The resulting generator $\boldsymbol{K}$ will in generally not commute with $\boldsymbol{J}$, even for Kähler manifolds, which has important consequences for its eigenvectors relevant for spectral functions.

[5]Mathematically, the linearised flow $d\Phi_t$ is defined as the differential (also known as push-forward) of the flow map Φ_t defined after equation (3.8). In general it is a map from the tangent space $\mathcal{T}_{\psi(0)}\mathcal{M}$ to the tangent space $\mathcal{T}_{\psi(t)}\mathcal{M}$. In the special case of $|\psi_0\rangle$ being a fixed point of the time evolution, it reduces to a linear map from $\mathcal{T}_{\psi_0}\mathcal{M}$ onto itself. One can then show that this map is generated by the linearisation $\boldsymbol{K}$ of the vector field $\mathcal{X}$ that defines the evolution flow.

[6]Note that due to the 2-out-of-3 principle, any linear map M satisfying two out of the three conditions $M\boldsymbol{\Omega}M^\intercal = \boldsymbol{\Omega}$, $M\boldsymbol{G}M^\intercal = \boldsymbol{G}$ and $M\boldsymbol{J}M^{-1} = \boldsymbol{J}$ will satisfy all three. Thus, any violation will necessarily affect at least two Kähler structures.

3.3.3 Comparison: projection vs. linearisation

In the following, we will compare the previously introduced approaches of approximating excitation energy. This comparison is particularly illuminating in the case of Kähler manifold.

At a stationary point, *i.e.*, $\partial_\mu E = 2\mathrm{Re}\,\langle V_\mu|\hat{H}|\psi\rangle = 0$, we consider the symplectic generator $\boldsymbol{K}$ defined as

$$\boldsymbol{K}^\mu{}_\nu = -\boldsymbol{\Omega}^{\mu\sigma}(\partial_\sigma\partial_\nu E) = (\boldsymbol{J}^{-1})^\mu{}_\sigma\,\partial_\nu\big(\mathbb{P}^\sigma_\psi\hat{H}\,|\psi\rangle\big)\,, \tag{3.45}$$

where we only have $\boldsymbol{J}^{-1} = -\boldsymbol{J}$ for Kähler manifolds. In the second step we used (2.49). Evaluating the derivative in (3.45) gives the two pieces

$$\partial_\nu\big(\mathbb{P}^\sigma_\psi\hat{H}\,|\psi\rangle\big) = \mathbb{P}^\sigma_\psi\hat{H}\,|V_\nu\rangle + (\partial_\nu\mathbb{P}^\sigma_\psi)\hat{H}\,|\psi\rangle\,, \tag{3.46}$$

where we evaluate everything at ψ_0 after computing the derivatives. We recognize $\boldsymbol{H}^\sigma{}_\nu = \mathbb{P}^\sigma_\psi\hat{H}\,|V_\nu\rangle$ and define $\boldsymbol{F}^\sigma{}_\nu = (\partial_\nu\mathbb{P}^\sigma_\psi)\hat{H}\,|\psi\rangle = \frac{2}{\langle\psi|\psi\rangle}\boldsymbol{G}^{\sigma\rho}\,\langle\partial_\nu V_\rho|\hat{H}|\psi_0\rangle$ leading to

$$\boldsymbol{K} = (\boldsymbol{J}^{-1})\Big(\boldsymbol{H} + \boldsymbol{F}\Big)\,. \tag{3.47}$$

In summary, we see that the linearisation $\boldsymbol{K}$ consists of the two pieces. The first is nothing other than the projected Hamiltonian $\boldsymbol{H}$. The second is the derivative of the projector. These terms are then multiplied with the inverse complex structure $\boldsymbol{J}^{-1}$.

In the case of a Kähler manifold there is a further deeper way to understand these two term that make up $\boldsymbol{K}$. We will discuss this now.

In the Kähler case, we can use $\boldsymbol{J}^2 = -\mathbb{1}$ to decompose any linear operator $\boldsymbol{K}$ on $\mathcal{T}_{\psi_0}\mathcal{M}$ as $\boldsymbol{K} = \boldsymbol{K}_+ + \boldsymbol{K}_-$ with

$$\boldsymbol{K}_\pm = \tfrac{1}{2}(\boldsymbol{K} \pm \boldsymbol{J}\boldsymbol{K}\boldsymbol{J})\,,\ \{\boldsymbol{K}_+, \boldsymbol{J}\} = 0\,,\ [\boldsymbol{K}_-, \boldsymbol{J}] = 0\,. \tag{3.48}$$

We will see that this decomposition coincides exactly with the one of $\boldsymbol{K}$ in (3.47). To do this, we use the fact that a Kähler manifold of dimension $2n$ always admits[7] a parametrisation $x = (x_1, \cdots, x_{2n})$ satisfying for $1 \le j \le n$

$$|V_j\rangle = \mathrm{i}\,|V_{n+j}\rangle\,, \tag{3.49}$$

i.e., the coordinate x_j is conjugate to x_{n+j}. In this basis, $\boldsymbol{J}$ and $\boldsymbol{\Omega}$ are

$$\boldsymbol{J} \equiv \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad \boldsymbol{\Omega} \equiv \frac{1}{2}\begin{pmatrix} -\mathrm{Im}\,\eta^{-1} & -\mathrm{Re}\,\eta^{-1} \\ \mathrm{Re}\,\eta^{-1} & -\mathrm{Im}\,\eta^{-1} \end{pmatrix}, \tag{3.50}$$

[7] This ultimately coincides with showing that a Kähler manifold is also a complex manifold, that is it admits a holomorphic parametrisation in terms of complex parameters $z^\alpha = \mathrm{x}^\alpha + \mathrm{i}\mathrm{y}^\alpha$.

where $\eta_{jk} = \langle V_j | V_k \rangle$. Then the structure of matrices that commute or anti-commute with $\boldsymbol{J}$ is

$$\boldsymbol{K}_- = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad \boldsymbol{K}_+ = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}. \tag{3.51}$$

We can evaluate $\boldsymbol{K}$ to find exactly this form

$$\boldsymbol{K}^\mu{}_\nu = -\boldsymbol{\Omega}^{\mu\rho}\partial_\rho\partial_\nu E = (\boldsymbol{K}_+)^\mu{}_\nu + (\boldsymbol{K}_-)^\mu{}_\nu \tag{3.52}$$

where its two pieces are explicitly given by

$$\boldsymbol{K}_+ \equiv \begin{pmatrix} \mathrm{Im}(\eta^{-1}h) & \mathrm{Re}(\eta^{-1}h) \\ -\mathrm{Re}(\eta^{-1}h) & \mathrm{Im}(\eta^{-1}h) \end{pmatrix}, \tag{3.53}$$

$$\boldsymbol{K}_- \equiv \begin{pmatrix} \mathrm{Im}(\eta^{-1}f) & \mathrm{Re}(\eta^{-1}f) \\ \mathrm{Re}(\eta^{-1}f) & -\mathrm{Im}(\eta^{-1}f) \end{pmatrix}, \tag{3.54}$$

where $h_{jk} = \langle V_j | \hat{H} | V_k \rangle$ and $f_{jk} = \langle \partial_j V_k | \hat{H} | \psi_0 \rangle$. This clearly shows that the two pieces are given by $\boldsymbol{K}_- = \boldsymbol{J}\boldsymbol{H}$ and $\boldsymbol{K}_+ = \boldsymbol{J}\boldsymbol{F}$ as defined before (3.47).

In conclusion, from the decomposition (3.52) we immediately see again that $\boldsymbol{K}$ has two contributions. One is related to the projected Hamiltonian $\boldsymbol{H}$ and commutes with $\boldsymbol{J}$. The other is related to the overlap of $\hat{H}\,|\psi_0\rangle$ with the *double tangent vectors* $|\partial_\alpha V_\beta\rangle$, which coincides with the one we previously described in terms of the derivative of the projector and anti-commutes with $\boldsymbol{J}$. Thus $\boldsymbol{K}_-$ is a complex linear map, while $\boldsymbol{K}_+$ is a contribution that makes $\boldsymbol{K}$ non-complex linear.

Finally, if we complexify tangent space, *i.e.*, treat complex linear combinations of $|V_\mu\rangle$ as linearly independent, there exists a basis transformation that makes $\boldsymbol{J}$ diagonal and brings $\boldsymbol{K}_+$ and $\boldsymbol{K}_-$ respectively, into block diagonal and block off-diagonal form, given by

$$\boldsymbol{J} \equiv \mathrm{i}\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \begin{aligned} \boldsymbol{K}_- &= \mathrm{i}\begin{pmatrix} -\eta^{-1}h & 0 \\ 0 & (\eta^{-1}h)^* \end{pmatrix}, \\ \boldsymbol{K}_+ &= \mathrm{i}\begin{pmatrix} 0 & (\eta^{-1}f)^* \\ -\eta^{-1}f & 0 \end{pmatrix}, \end{aligned} \tag{3.55}$$

i.e., for Kähler manifolds the terms in (3.47) decouple nicely. For a non-Kähler manifold, neither $\boldsymbol{H}$ nor $\boldsymbol{F}$ may commute with $\boldsymbol{J}$, but even the decomposition (3.48) will not work for $\boldsymbol{J}^2 \neq -\mathbb{1}$.

In the next section, we will see that the term $\boldsymbol{K}_+ = \boldsymbol{J}\boldsymbol{F}$ can be a blessing and a curse: on the one hand, it can ensure that in systems with spontaneously broken symmetry the eigenvalues of $\boldsymbol{K}$ contain a Goldstone mode. On the other hand, for unfortunate choices of the variational family $\mathcal{M}$ we may encounter such massless

modes even if there is no spontaneously broken symmetry (spurious Goldstone mode).

Kähler vs. non-Kähler. We can relate the linearisation $\boldsymbol{K}$ with the projected Hamiltonian $\boldsymbol{H}$ via (3.47). For Kähler manifolds, this decomposition becomes particularly geometric, as the two pieces correspond to its complex linear and complex anti-linear part.

3.3.4 Spurious Goldstone mode

The spectrum of $\boldsymbol{K}$ is not variational. In contrast to a variational approximation of an eigenstate, our eigenvector $|\mathcal{E}(\lambda)\rangle$ of $\boldsymbol{K}$ with

$$\boldsymbol{K}^{\mu}{}_{\nu}\,\mathcal{E}^{\nu}(\lambda) = \lambda\,\mathcal{E}^{\mu}(\lambda)\,, \tag{3.56}$$

and $|\mathcal{E}(\lambda)\rangle = \mathcal{E}^{\mu}(\lambda)\,|V_{\mu}\rangle$, does not satisfy

$$\lambda = \pm\,\langle\mathcal{E}(\lambda)|\mathrm{i}\hat{H}|\mathcal{E}(\lambda)\rangle\ . \tag{3.57}$$

The expectation value of the full Hamiltonian with respect to $|\mathcal{E}(\lambda)\rangle$ is in general not easily related to λ, as it would be for a variational state. It is also not true that for every eigenvalue pair $\pm\mathrm{i}\omega_\ell$, there exists a true eigenstate $|E_\ell\rangle$ of $\hat{H}$ with excitation energy $E_\ell - E_0 \leq \omega_\ell$.

In fact, there are situations, where the true ground state $|E_0\rangle$ is non-degenerate, but $\boldsymbol{K}$ still has a zero eigenvalue associated to a massless Goldstone mode. This typically occurs if we have a conserved quantity $\hat{A}$ with $[\hat{A},\hat{H}] = 0$, such that $-\mathrm{i}\hat{A}\,|\psi\rangle \in \mathcal{T}_{\psi}\mathcal{M}$ everywhere as discussed in the context of Proposition 7. At this point, the question is if the global energy minimum $|\psi_0\rangle$ on $\mathcal{M}$ is invariant under $e^{-\mathrm{i}\hat{A}}$ or not. Whenever the global minimum on ψ_0 on $\mathcal{M}$ is not invariant, *i.e.*, there is a whole family $|\psi_0(\varphi)\rangle = e^{-\mathrm{i}\varphi\hat{A}}\,|\psi_0\rangle$ of approximate ground states, the generator $\boldsymbol{K}$ will have a massless Goldstone mode

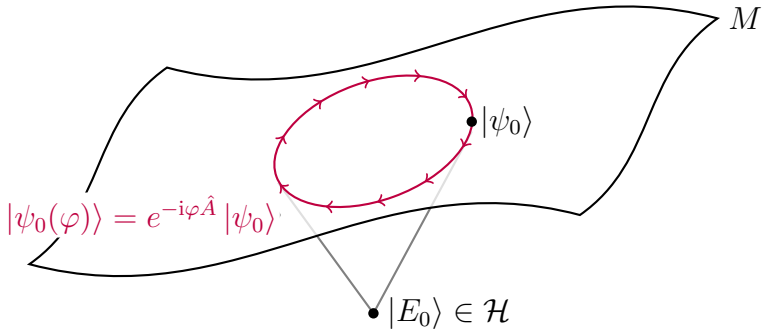
$$\mathcal{E}^{\mu}_{\mathrm{G}} = \mathbb{P}^{\mu}_{\psi_0}(-\mathrm{i}\hat{A})\,|\psi_0\rangle \quad \text{with} \quad \boldsymbol{K}^{\mu}{}_{\nu}\mathcal{E}^{\nu}_{\mathrm{G}} = 0\,. \tag{3.58}$$

Whenever the true ground state $|E_0\rangle$ of the system is invariant under $e^{-\mathrm{i}\hat{A}}$, this Goldstone mode is spurious and merely an artefact of a spontaneous symmetry breaking on $\mathcal{M}$, but not on full $\mathcal{P}(\mathcal{H})$.

We illustrate this issue in Figure 3.2, where the Hamiltonian is spontaneously broken only on the variational manifold $\mathcal{M}$, but not in the full Hilbert space, where it has a unique ground state $|E_0\rangle$.

This was pointed out in [92] as an important problem of approximating the spectrum via linearised equations of motion rather than using the projected Hamiltonian. However, we also found that sometimes this can also be desirable to capture

Figure 3.2: *Spurious Goldstone mode.* Even if the Hamiltonian has a unique ground state $|E_0\rangle$ without a spontaneously broken symmetry in full Hilbert space, on a chosen submanifold M there may be inequivalent states $|\psi_0(\varphi)\rangle = e^{-\mathrm{i}\varphi\hat{A}}\,|\psi_0\rangle$ that all minimise the energy. This leads to a spontaneous breaking of the symmetry generated by $\hat{A}$ and the appearance of a spurious Goldstone mode.



physical features of models with spontaneous symmetry breaking. In Chapter 6, we will discuss an application where the gapless Bogoliubov excitation spectrum of the Bose-Hubbard model can be shown to result from the diagonalization of the generator (3.42) on a suitable variational manifold.

3.4 Spectral functions

Next, we would like to use the variational manifold $\mathcal{M}$ to estimate the spectral function of a system with respect to the perturbation operator $\hat{V}$.

Given a Hermitian operator $\hat{V}$, the spectral function is

$$\mathcal{A}(\omega) = -\frac{1}{\pi}\,\mathrm{Im}\,G^R(\omega), \tag{3.59}$$

where G^R is the retarded Green's function

$$G^R(\omega) = -\mathrm{i}\int dt\, e^{\mathrm{i}\omega t}\Theta(t)\frac{\langle\psi_0|[\hat{V}(t),\hat{V}]|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} \tag{3.60}$$

with $\Theta(t)$ being the step function and $\hat{V}(t)$ the Heisenberg evolved operator under the system Hamiltonian $\hat{H}$.

The definition in terms of the retarded Green's function stems from linear response theory. Indeed, let us suppose that a small external perturbing probe field $\epsilon\varphi(t)$ couples to our system through the operator $\hat{V}$. That is, the system state $|\psi_\epsilon(t)\rangle$ evolves under the perturbed Hamiltonian $\hat{H} + \epsilon\varphi(t)\hat{V}$. Then, let

us measure the response of the system through the expectation value of the same observable $\hat{V}$. As the perturbation is ideally infinitesimally small, we only consider such response up to linear order in ϵ. Consequently, we define the time-domain linear response as

$$\delta V(t) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \frac{\langle\psi_\epsilon(t)|\hat{V}|\psi_\epsilon(t)\rangle}{\langle\psi_\epsilon(t)|\psi_\epsilon(t)\rangle}. \tag{3.61}$$

Then in frequency domain we have that

$$\delta\tilde{V}(\omega) \equiv \int dt\, e^{\mathrm{i}\omega t}\delta V(t) = \tilde{\varphi}(\omega)G^R(\omega). \tag{3.62}$$

That is, G^R is exactly the so-called linear susceptibility of the system, which is an experimentally accessible quantity.

Given a variational manifold there are two possible paths to trying to approximate $\mathcal{A}(\omega)$.

1. We can calculate the quantity (3.61) after having projected the evolution of $|\psi_\epsilon(t)\rangle$ on the manifold. In other words, we perform linear response theory directly on the variational manifold. This leads us to express $\mathcal{A}(\omega)$ in terms of the eigendecomposition of the generator of linearised real time evolution $\boldsymbol{K}$ introduced in (3.42).

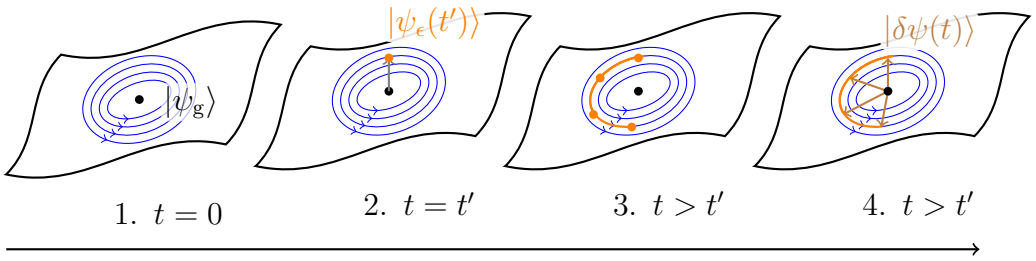
2. Alternatively, one can try to approximate on the manifold the quantity

$$e^{-\mathrm{i}\hat{H}t}\hat{V}\,|\psi_0\rangle\,, \tag{3.63}$$

that appears in equation (3.60). In this case one should note that in general $\hat{V}\,|\psi_0\rangle$ does not belong to the variational manifold, so one has to perform some truncation even before applying the time evolution operator $e^{-\mathrm{i}\hat{H}t}$. The other subtlety here is that one must make sure that the quantity (3.63) is calculated with the correct global phase, as we explained in Section 3.2.3.

It seems to us that method 2 captures less the spirit of variational manifolds. Indeed one has that the quantity $\hat{V}\,|\psi_0\rangle$ would morally represent a small perturbation around the groundstate $|\psi_0\rangle$ and would thus naturally live in the tangent space to the manifold of states at $|\psi_0\rangle$. Representing $\hat{V}\,|\psi_0\rangle$ as a vector of $\mathcal{M}$ therefore is only meaningful if the manifold itself is a good representation of its own tangent space. But this is not true for general manifolds and indeed there is no uniquely defined method for representing $\hat{V}\,|\psi_0\rangle$ on $\mathcal{M}$. The first method, on the other hand, can alternatively be thought of precisely as representing the perturbations generated by $\hat{V}$ on $\mathcal{T}_{\psi_0}\mathcal{M}$. Furthermore, we will show that method 1 leads to a closed expression for the spectral function from which it is immediate

Figure 3.3: *Linear response theory.* We consider an approximate ground state $\psi_0 \in \mathcal{M}$. While $|\psi_0\rangle$ does not evolve in time, certain trajectories of nearby states are approximately elliptic. A finite perturbation at time t' changes the state to $|\psi_\epsilon(t')\rangle = e^{\mathrm{i}\epsilon\hat{V}}|\psi_0\rangle$. This state will then evolve according to the equations of motion. We linearise by taking the limit $\epsilon \to 0$, where we find that the tangent vector $|\delta\psi(t)\rangle = \frac{d}{d\epsilon}|\psi_\epsilon(t)\rangle|_{\epsilon=0}$ rotates in a way defined by the eigenmodes of $\boldsymbol{K}$.



to see that $\mathrm{sgn}\mathcal{A}(\omega) = \mathrm{sgn}\,\omega$ (as it is in the full Hilbert space), while this cannot be shown in general for method 2.

For these reasons in the next subsections we will focus on the details of the first method, giving a final expression for the spectral function estimated in this way in Proposition 10.

3.4.1 Linear response theory

A possible way of calculating spectral functions is thus to perform linear response theory directly on the variational manifold. In this subsection we will then briefly explain how this can be done. The idea is illustrated in Figure 3.3.

Let us consider a possibly time-dependent perturbation $\hat{A}(t)$ of our unperturbed Hamiltonian $\hat{H}_0$, such that $\hat{H}_\epsilon(t) = \hat{H}_0 + \epsilon\hat{A}(t)$, and an observable $\hat{B}$, whose response we are interested in. For spectral functions, we will be interested in the particular case where $\hat{A}(t) = \varphi(t)\hat{V}$ for arbitrary functions $\varphi(t)$ and $\hat{B} = \hat{V}$, but we will for the moment keep our treatment general.

Our perturbed Hamiltonian gives rise to the time dependent real time evolution vector field $\mathcal{X}_\epsilon(t)$, which is

$$\mathcal{X}_\epsilon(t) = \mathcal{X}_0 + \epsilon\,\mathcal{X}_A(t)\,, \tag{3.64}$$

where $\mathcal{X}_0$ and $\mathcal{X}_A(t)$ are the evolution vector fields associated to the Hamiltonians $\hat{H}_0$ and $\hat{A}(t)$ respectively. The solution of this perturbed evolution is $|\psi_\epsilon(t)\rangle$ satisfying

$$\mathbb{Q}_\psi \tfrac{d}{dt}|\psi_\epsilon(t)\rangle = \mathcal{X}^\mu_\epsilon(t)\,|V_\mu\rangle\ . \tag{3.65}$$

For the following Proposition 8 it would not be important whether the evolution vector field is defined according to the Lagrangian or McLachlan variational principles, as long as it has the form (3.64). However, later we will be interested in the case in which the perturbed evolution happens around the approximate ground state $|\psi_0\rangle$ and it will be important that this state is also a fixed point of the time evolution. So, as was the case in Section 3.3.2, from now on we will suppose that the evolution vector fields are defined according to the Lagrangian evolution (3.9).

We are interested in the response in expectation value of the observable $\hat{B}$ at linear order in ϵ, that is

$$\begin{aligned} \delta B(t) &= \tfrac{d}{d\epsilon} \tfrac{\langle\psi_\epsilon(t)|\hat{B}|\psi_\epsilon(t)\rangle}{\langle\psi_\epsilon(t)|\psi_\epsilon(t)\rangle}\Big|_{\epsilon=0} \\ &= \delta x^\mu(t)\, \partial_\mu B(x(t))\,, \end{aligned} \tag{3.66}$$

where we defined the propagated perturbation

$$\delta x^\mu(t)\, |V_\mu\rangle = \mathbb{Q}_\psi \tfrac{d}{d\epsilon} |\psi_\epsilon(t)\rangle \Big|_{\epsilon=0} \in \mathcal{T}_{\psi(t)}\mathcal{M} \tag{3.67}$$

which can be evaluated as follows.

Proposition 8. *Given a variational manifold $\mathcal{M}$ we define (according to the Lagrangian action principle) the free projected real time evolution $|\psi(t)\rangle$ as governed by the free Hamiltonian $\hat{H}_0$ and the perturbed projected real time evolution $|\psi_\epsilon(t)\rangle$ as governed by the perturbed Hamiltonian $\hat{H}_\epsilon(t) = \hat{H}_0 + \epsilon\hat{A}(t)$, both with the same initial state $|\psi(0)\rangle$. Then, the propagated perturbation, defined according to* (3.67)*, is given by*

$$\delta x^\mu(t) = -\int_{-\infty}^{t} dt'\, (d\Phi_{t-t'})^\mu{}_\nu\, \boldsymbol{\Omega}^{\nu\rho}\, \partial_\rho A(t')\big|_{\psi(t')}\,, \tag{3.68}$$

where $d\Phi_t$ is the linearised free evolution flow[8]*.*

Proof. This can be shown in a standard way by using the interaction representation. We sketch a proof in Appendix A.1. □

Put simply, δx^μ is the superposition of all propagated perturbations, *i.e.*, a perturbation

$$-\boldsymbol{J}^\nu{}_\rho \mathbb{P}^\rho_{\psi(t')} \hat{A}(t')\, |\psi(t')\rangle = -\boldsymbol{\Omega}^{\nu\rho}\, \partial_\rho A\big|_{\psi(t')} \tag{3.69}$$

at time t' is evolved with the linearised free evolution $d\Phi_{t-t'}$ to time t where it contributes towards $\delta x^\mu(t)$.

[8] See footnote 5.

If we now take as initial state $|\psi(0)\rangle$ the approximate ground state $|\psi_0\rangle$, that is a fixed point of the projected evolution, we have that the free evolution is trivial $|\psi(t)\rangle = |\psi_0\rangle$. It also follows that $d\Phi_t$ is a linear map from $\mathcal{T}_{\psi_0}\mathcal{M}$ onto itself given by

$$d\Phi_t = e^{\boldsymbol{K}t}\,, \tag{3.70}$$

where $\boldsymbol{K}$ is the generator of the linearised flow introduced in (3.42). The map $d\Phi_t$ can therefore be evaluated in terms of the spectral decomposition of $\boldsymbol{K}$.

Proposition 9. *The linear response to a perturbation $\hat{A}(t)$, measured in terms of the observable $\hat{B}$, for a system initially in the state $\psi_0 \in \mathcal{M}$ is given by*

$$\begin{aligned}\delta B(t) = -\mathrm{i}\sum_\ell \operatorname{sgn}(\mathrm{i}\lambda_\ell)[\mathcal{E}^\mu(\lambda_\ell)\partial_\mu B] \\ \times \int_{-\infty}^{t} dt'\, e^{\lambda_\ell(t-t')}\,[\mathcal{E}^\nu(\lambda_\ell)\partial_\nu A(t')]^*\,,\end{aligned} \tag{3.71}$$

where all derivatives are evaluated at $|\psi_0\rangle$ and $\mathcal{E}^\mu(\lambda_\ell)$ is an eigenvector of $\boldsymbol{K}$ such that

$$\boldsymbol{K}^\mu{}_\nu \mathcal{E}^\nu(\lambda_\ell) = \lambda_\ell \mathcal{E}^\mu(\lambda_\ell)\,, \tag{3.72}$$

and normalised so that $\mathcal{E}^\mu(\lambda_\ell)\,\boldsymbol{\omega}_{\mu\nu}\mathcal{E}^\nu(\lambda_\ell)^* = \mathrm{i}\,\operatorname{sgn}(\mathrm{i}\lambda_\ell)$.

Proof. We can always decompose $\boldsymbol{K}$ in terms eigenvectors $\mathcal{E}(\lambda)$ with eigenvalues λ and dual eigenvectors [9] $\widetilde{\mathcal{E}}(\lambda)$, such that

$$\boldsymbol{K}^\mu{}_\nu = \sum_\ell \lambda_\ell\, \mathcal{E}^\mu(\lambda_\ell)\, \widetilde{\mathcal{E}}_\nu(\lambda_\ell)\,. \tag{3.73}$$

The eigenvalues λ_ℓ will come in conjugate pairs $\pm\mathrm{i}\omega_\ell$, which implies that the associated eigenvectors and dual eigenvectors are complex and mathematically speaking lie the complexified tangent space.
However, as $\boldsymbol{K}$ is a real map, we must have $\mathcal{E}^\mu(\mathrm{i}\omega) = \mathcal{E}^\mu(-\mathrm{i}\omega)^*$. We then notice that $\boldsymbol{\Omega}^{\mu\nu}\widetilde{\mathcal{E}}_\nu(-\mathrm{i}\omega)$ is an eigenvector of $\boldsymbol{K}$ with eigenvalue $\mathrm{i}\omega$. To see this it is sufficient to apply $\boldsymbol{K}$ to it and use the symplectic property $\boldsymbol{K}\boldsymbol{\Omega} = -\boldsymbol{\Omega}\boldsymbol{K}^\intercal$. It is then always possible to normalise the eigenvectors $\mathcal{E}^\mu$ such that the relation $\boldsymbol{\Omega}^{\mu\nu}\widetilde{\mathcal{E}}_\nu(-\mathrm{i}\omega) = -\mathrm{i}\,\operatorname{sgn}(\omega)\mathcal{E}^\mu(\mathrm{i}\omega) = -\mathrm{i}\,\operatorname{sgn}(\omega)\mathcal{E}^\mu(-\mathrm{i}\omega)^*$ holds. [10] From this, inverting $\boldsymbol{\Omega}$ and exploiting its antisymmetry, we have $\widetilde{\mathcal{E}}_\mu(-\mathrm{i}\omega) = \mathrm{i}\,\operatorname{sgn}(\omega)\mathcal{E}^\nu(-\mathrm{i}\omega)^*\boldsymbol{\omega}_{\nu\mu}$.

[9] The dual vector $\widetilde{\mathcal{E}}_\mu(\lambda)$ is defined by $\widetilde{\mathcal{E}}_\mu(\lambda)\,\mathcal{E}^\mu(\lambda') = \delta_{\lambda,\lambda'}$

[10] Doing this rescaling while maintaining the property $\mathcal{E}^\mu(\mathrm{i}\omega) = \mathcal{E}^\mu(-\mathrm{i}\omega)^*$ is actually only possible if $\mathcal{E}^\mu(-\mathrm{i}\omega)\boldsymbol{\omega}_{\mu\nu}\mathcal{E}^\nu(\mathrm{i}\omega) = \mathrm{i}a$ with $a > 0$, $\forall\omega > 0$. But this is always true because by definition $\boldsymbol{K} = -\boldsymbol{\Omega}h$, where $h_{\mu\nu} = \partial_\mu\partial_\nu E$ is positive definite (Hessian at a local minimum). It follows that $-\boldsymbol{\omega}\boldsymbol{K} > 0$ and therefore $0 < -\mathcal{E}^\mu(\mathrm{i}\omega)^*\boldsymbol{\omega}_{\mu\rho}\boldsymbol{K}^\rho{}_\nu\mathcal{E}^\nu(\mathrm{i}\omega) = -\mathrm{i}\omega\mathcal{E}^\mu(-\mathrm{i}\omega)\boldsymbol{\omega}_{\mu\nu}\mathcal{E}^\nu(\mathrm{i}\omega) = \omega a$.

Using this and (3.73), we can rewrite (3.70) as

$$(d\Phi_t)^\mu{}_\nu = \mathrm{i}\sum_\ell \mathrm{sign}(\mathrm{i}\lambda_\ell)\, e^{\lambda_\ell t}\, \mathcal{E}^\mu(\lambda_\ell)\, \mathcal{E}^\rho(\lambda_\ell)^*\, \boldsymbol{\omega}_{\rho\nu}. \tag{3.74}$$

Combining this with (3.66) and (3.68) we have (3.71). □

3.4.2 Spectral response

To calculate spectral functions we now just need to evaluate the result (3.71) for $\hat{A}(t) = \varphi(t)\hat{V}$ and $\hat{B} = \hat{V}$ and then take the Fourier transform.

Proposition 10. *The spectral function with respect to the perturbation operator $\hat{V}$, estimated by performing linear response theory on the variational manifold $\mathcal{M}$, is*

$$\mathcal{A}(\omega) = \mathrm{sgn}(\omega) \sum_\ell |\mathcal{E}^\mu(\mathrm{i}\omega_\ell)\, \partial_\mu V|^2\, \delta(\omega - \omega_\ell)\,, \tag{3.75}$$

where $\mathcal{E}^i(\mathrm{i}\omega_\ell)$ are the eigenvectors of $\boldsymbol{K}$, normalised such that $\mathcal{E}^\mu(\mathrm{i}\omega_\ell)^*\, \boldsymbol{\omega}_{\mu\nu} \mathcal{E}^\nu(\mathrm{i}\omega_\ell) = \mathrm{i}\,\mathrm{sgn}(\omega_\ell)$, and the sum runs over all possible values of ω_ℓ (appearing in pairs of opposite signs).

Proof. Evaluating the Fourier transform of (3.71) and comparing with (3.62) leads us to the estimate for the retarded Green's function

$$\begin{aligned} G^R(\omega) &= -\mathrm{i}\sum_\ell \mathrm{sgn}(\omega_\ell) |\mathcal{E}^\mu(\mathrm{i}\omega_\ell)\, \partial_\mu V|^2 \int dt\, e^{\mathrm{i}(\omega-\omega_\ell)t}\Theta(t) \\ &= \sum_\ell \mathrm{sgn}(\omega_\ell) |\mathcal{E}^\mu(\mathrm{i}\omega_\ell)\, \partial_\mu V|^2 \\ &\qquad \times \left[P\frac{1}{\omega - \omega_\ell} - \mathrm{i}\pi\delta(\omega - \omega_\ell) \right], \end{aligned} \tag{3.76}$$

where the Sokhotski-Plemelj formula has been used. The imaginary part of this expression can be then be inserted into the definition of the spectral function (3.59), leading to the result (3.75). □

Spectral functions calculated in this way have the desirable property $\mathrm{sgn}\mathcal{A}(\omega) = \mathrm{sgn}\,\omega$.

Kähler vs. non-Kähler. On a non-Kähler manifold, where we have two inequivalent definitions of the equations of motion, it only makes sense to perform linear response theory with the ones coming from the Lagrangian action principle, as their fixed point coincides with the approximate ground state.

3.5 Imaginary time evolution

In the previous sections we have assumed we knew the state $|\psi_0\rangle$ that minimises the energy on the variational manifold $\mathcal{M}$. Solving this optimisation problem is often non-trivial and different methods may be appropriate in different situations. However, we would like here to present a method, known as *projected imaginary time evolution*, that makes use of the same geometric notions introduced in Section 3.2 for real time evolution.

On full Hilbert space, imaginary time evolution is

$$\frac{d}{d\tau}|\psi(\tau)\rangle = -(\hat{H} - E(\tau))|\psi(\tau)\rangle \ , \tag{3.77}$$

which can be integrated to the solution

$$|\psi(\tau)\rangle = \frac{e^{-\hat{H}\tau}|\psi(0)\rangle}{\sqrt{\langle\psi(0)|e^{-2\hat{H}\tau}|\psi(0)\rangle}} \, . \tag{3.78}$$

This will converge in the limit $\tau \to \infty$ to a true ground state if and only if the initial state $|\psi(0)\rangle$ had some non-zero overlap with it.

Given a variational manifold $\mathcal{M}$, we can approximate imaginary time evolution on it and hope that it will also converge to the approximate ground state $|\psi_0\rangle$. This can be done by projecting (3.77) onto tangent space. Contrary to real time evolution, there does not exist a formulation of imaginary time evolution in terms of an action principle, so the projection can only be done according to the McLachlan minimal error principle.

We would like to minimise the local projection error

$$\left\| \tfrac{d}{d\tau}|\psi(\tau)\rangle - (E - \hat{H})|\psi(\tau)\rangle \right\| \ , \tag{3.79}$$

imposing that $\frac{d}{d\tau}|\psi(\tau)\rangle \in \mathcal{T}_\psi\mathcal{M}$, which leads to

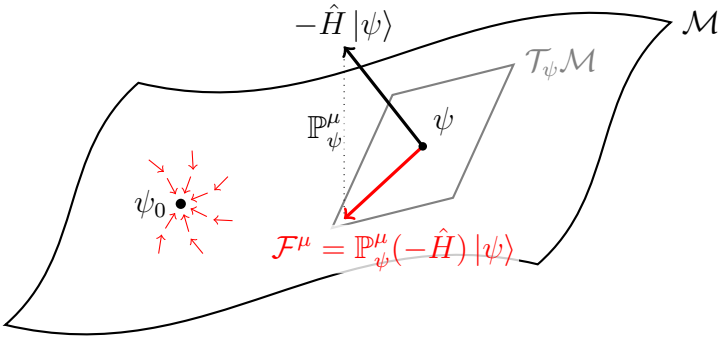
$$\begin{aligned} \tfrac{d}{d\tau}|\psi(\tau)\rangle &= \mathbb{P}_{\psi(\tau)}(E - \hat{H})|\psi(\tau)\rangle \\ &= -\mathbb{P}_{\psi(\tau)}\hat{H}|\psi(\tau)\rangle \ , \end{aligned} \tag{3.80}$$

where we used $\mathbb{P}_\psi|\psi\rangle = 0$.

This leads to the projected evolution equation

$$\frac{dx^\mu}{d\tau}|V_\mu\rangle = -\mathbb{P}_{\psi(\tau)}\hat{H}|\psi(\tau)\rangle \ , \tag{3.81}$$

Figure 3.4: *Imaginary time evolution.* We illustrate, analogously to Figure 3.1, the imaginary time evolution vector field $\mathcal{F}^\mu$ on the variational family $\mathcal{M}$ which is given by the orthogonal projection of $-\hat{H}\,|\psi\rangle$ through $\mathbb{P}_\psi$ onto $\mathcal{M}$. This vector field flows towards the global minimum ψ_0 of the energy function.



from which we can define the imaginary time evolution vector field $\mathcal{F}$ everywhere on $\mathcal{M}$, such that

$$\frac{dx^\mu}{d\tau} = \mathcal{F}^\mu(x) = -\mathbb{P}^\mu_{\psi(x)}\hat{H}\,|\psi(x)\rangle\ . \tag{3.82}$$

This vector field can be understood as follows.

Proposition 11. *Given a manifold $\mathcal{M}$, the projected imaginary time evolution is given by*

$$\frac{dx^\mu}{d\tau} = \mathcal{F}^\mu(x) = -\boldsymbol{G}^{\mu\nu}(\partial_\nu E)\,, \tag{3.83}$$

where $E(x)$ is the energy function, defined in the context of equation (2.47). Its solution $x(\tau)$ monotonically decreases the energy.

Proof. We apply the projector $\mathbb{P}^\mu_\psi$ in (3.82) to find

$$\mathcal{F}^\mu = -\mathbb{P}^\mu\hat{H}\,|\psi\rangle = -\frac{2}{\langle\psi|\psi\rangle}\boldsymbol{G}^{\mu\nu}\,\mathrm{Re}\,\langle V_\nu|\,\hat{H}\,|\psi\rangle\ . \tag{3.84}$$

We simplify this by using (2.49). Plugging this back into the previous equations, we arrive at (3.83). To show that the energy monotonically decreases, we find

$$\frac{dE}{d\tau} = (\partial_\mu E)\frac{dx^\mu}{d\tau} = -(\partial_\mu E)\,\boldsymbol{G}^{\mu\nu}\,(\partial_\nu E) \leq 0\,, \tag{3.85}$$

which follows from the positivity of $\boldsymbol{G}^{\mu\nu}$. □

We thus recognize projected imaginary time evolution (3.83) as gradient descent of the energy function $E(x)$ with respect to the natural geometry encoded in the metric $\boldsymbol{G}$ on the manifold $\mathcal{M}$, as illustrated in Figure 3.4. It is our experience that solving (3.83) numerically has better convergence properties than performing a naive gradient descent, where we just try to minimise the energy $E(x)$ as a function of x assuming a flat metric.

When replacing the fixed time step by a line search, imaginary time evolution becomes equivalent to Riemannian gradient descent. More specifically, the literature on Riemannian optimisation [93–96] describes how the Riemannian geometry (*i.e.*, the metric) of a manifold can be taken into account in each of the standard optimisation algorithms such as the gradient descent method, Newton's method, the conjugate gradient method, and quasi-Newton methods such as the (limited memory) Broyden–Fletcher–Goldfarb –Shanno scheme [97–100].

Kähler vs. non-Kähler. The results discussed in this section do not rely on the manifold $\mathcal{M}$ being a Kähler manifold. The McLachlan projection principle is the only one that can be resonably defined for imaginary time evolution and leads to the desirable gradient descent result for any real differentiable manifold, independently of the Kähler property.

3.5.1 Conserved quantities

In many situations, one would like to further constrain our variational manifold by requiring that certain operators $\hat{A}_I$ have fixed expectation values A_I. Geometrically, this amounts to restricting the search to the submanifold

$$\widetilde{\mathcal{M}} = \left\{ \psi \in \mathcal{M} \,\middle|\, \tfrac{\langle\psi|\hat{A}_I|\psi\rangle}{\langle\psi|\psi\rangle} = A_I \,\forall\, I \right\} \subset \mathcal{M}\,. \tag{3.86}$$

For example, for Hamiltonians commuting with the total particle number operator $\hat{N}$, one often wants to find lowest energy state within an eigenspace of $\hat{N}$ with $\hat{N}\,|\psi\rangle = N\,|\psi\rangle$. To approximate such a state on a non-linear variational manifold $\mathcal{M}$, we can search for minimal energy state on the submanifold of states with $\langle\hat{N}\rangle = N$.

In general, this manifold $\widetilde{\mathcal{M}}$ will not satisfy the Kähler property anymore. In particular, if we only fix a single expectation value, we will generically reduce the dimension of a Kähler $\mathcal{M}$ to an odd dimension, which cannot be again a Kähler manifold. However, we have seen that for the purpose of finding the state of minimal energy, we can apply formula (3.77) on the reduced manifold, regardless of the Kähler property.

Instead of finding a new parametrisation of the reduced variational manifold, as long as we choose an initial state for the imaginary time evolution that satisfies the desired constraints, we can also just implement them locally. We can indeed

modify the imaginary time evolution vector field $\mathcal{F}$ by further projecting it onto the restricted tangent space $\mathcal{T}\widetilde{\mathcal{M}}$. In this way the respective expectation values are preserved by construction.

If there are several quantities $\hat{A}_I$ that we wish to fix, $\mathcal{T}_\psi\widetilde{\mathcal{M}}$ is given by the sub tangent space orthogonal to the span of $X_I^\mu = \mathbb{P}_\psi^\mu \hat{A}_I \,|\psi\rangle$. To project onto it, we define

$$\widetilde{g}_{IJ} = X_I^\mu \, \boldsymbol{g}_{\mu\nu} \, X_J^\nu \,, \tag{3.87}$$

which gives rise to the projector

$$\widetilde{P}^\mu{}_\nu = \delta^\mu{}_\nu - X_I^\mu \, \widetilde{G}^{IJ} X_J^\rho \, \boldsymbol{g}_{\rho\nu} \,, \tag{3.88}$$

where $\widetilde{G}^{IJ}$ is the inverse of $\widetilde{g}_{IJ}$ (or pseudo inverse, if not all constraints are independent). The modified imaginary time evolution vector field is then

$$\widetilde{\mathcal{F}}^\mu = \widetilde{P}^\mu{}_\nu \mathcal{F}^\nu \,, \tag{3.89}$$

which will conserve all the expectation values $A_I(\tau)$. In analogy to (3.28), this is equivalent to

$$\widetilde{\mathcal{F}}^\mu = -\widetilde{\boldsymbol{G}}^{\mu\nu}(\partial_\nu E) \quad \text{with} \quad \widetilde{\boldsymbol{G}}^{\mu\nu} = \widetilde{P}^\mu{}_\sigma \widetilde{P}^\nu{}_\rho \boldsymbol{G}^{\sigma\rho} \,. \tag{3.90}$$

If we want to fix the expectation value of the number operator $\hat{N}$, we have the scalar function $N(x) = \langle \hat{N} \rangle$ with $X^\mu = \mathbb{P}_\psi^\mu \hat{N} \,|\psi\rangle = \boldsymbol{G}^{\mu\nu}\partial_\nu N$, such that

$$\widetilde{\mathcal{F}}^\mu = \mathcal{F}^\mu - \frac{\boldsymbol{G}^{\mu\nu}(\partial_\nu N)}{(\partial_\sigma N)\boldsymbol{G}^{\sigma\rho}(\partial_\rho N)} \, (\partial_\lambda N)\mathcal{F}^\lambda \,, \tag{3.91}$$

which clearly satisfies $\frac{dN}{d\tau} = (\partial_\mu N)\widetilde{\mathcal{F}}^\mu = 0$.

Part II

Chapter 4

Gaussian states

In this Chapter we will introduce a very important family of variational states, namely Gaussian states. We will see that both fermionic and bosonic Gaussian states can be understood as examples of a much larger class of states, known as *group-theoretic coherent states*, and we will derive some of their most important properties.

Part of the material presented in this chapter was published in reference [101]:

T. Guaita, L. Hackl, T. Shi, E. Demler, J.I. Cirac,
Generalization of group-theoretic coherent states for variational calculations,
Phys. Rev. Research **3**, 023090 (2021), used under CC BY 4.0.

4.1 Overview

As already mentioned in the Introduction in Section 1.4, Gaussian states represent a very successful variational set of states. They can be defined for both bosonic and fermionic systems and in both cases represent exact ground states of non-interacting Hamiltonians. From the perspective of variational methods, they can be readily used as *ansätze* for an approximate description of the physical states of systems which are weakly interacting. This approach lies at the heart of many widely applied methods such as the Hartree-Fock methods for fermionic systems or the Bogoliubov theory for bosonic systems. Furthermore, Gaussian states are exceptionally easy to handle computationally, as most relevant quantities can be expressed in very simple forms through the use of Wick's theorem.

In this chapter we will introduce the necessary formalism to define and manipulate bosonic and fermionic Gaussian states. However, to set these states into a broader perspective we will first take a step back and introduce a more general class of states, known as *group-theoretic coherent states*. These states were introduced independently by Gilmore [62, 102] and Perelomov [63, 103] and can be

constructed for a wide range of different quantum systems. We will see that Gaussian states represent one particular example of group-theoretic coherent states in systems made up of bosonic or fermionic modes. We will also see that many of the advantageous computational properties of Gaussian state derive in fact from more general properties of group-theoretic coherent states. Finally, understanding the structure of group-theoretic coherent states will prove itself very useful when exploring the possibilities of defining variational manifolds that go beyond Gaussian or coherent states, which will be the subject of Chapter 5.

Therefore, in Section 4.2, we will first focus on defining group-theoretic coherent states. The only requirement to construct a set of these states is to have a Hilbert space on which acts a unitary representation of a Lie group. Given such a space and such a representation, we will show a construction which ultimately defines a well-behaved submanifold of the Hilbert space. For states of this manifold, computing expectation values of operators and doing other manipulations is particularly efficient thanks to the group-theoretic structures with which they are endowed.

In Section 4.3, we will then move on to define bosonic Gaussian states. Their understanding will be made easier by using the language of group-theoretic coherent states introduced in the previous section. We will collect a series of results that, combined appropriately, allow to make an efficient use of Gaussian states as a variational manifold. In Section 4.4, we will then repeat the analogue construction for fermionic Gaussian states, highlighting the parallelism with their bosonic counterparts.

Additionally to bosonic and fermionic Gaussian states, in Example 10 we will introduce another important example of group-theoretic coherent states, namely *spin-$\frac{1}{2}$ coherent states*. These states have many useful applications and we will elaborate on them further in Chapters 5 and 7.

4.2 Group-theoretic coherent states

In this section, we review the basic definition and properties of group-theoretic coherent states, basing ourselves mostly on [60].

We consider a semi-simple Lie group $\mathcal{G}$ with Lie algebra $\mathfrak{g}$. Let $\mathcal{U}$ be a unitary representation of $\mathcal{G}$ on the Hilbert space $\mathcal{H}$, *i.e.*, $\mathcal{U}(g)$ is a unitary operator on $\mathcal{H}$ for every group element $g \in \mathcal{G}$, such that

$$\mathcal{U}(g_1)\mathcal{U}(g_2) = \mathcal{U}(g_1 g_2) \quad \forall g_1, g_2 \in \mathcal{G}\,. \tag{4.1}$$

The representation of the group induces a corresponding representation of the algebra. Indeed, for group elements g sufficiently close to the identity, it is possible to write $\mathcal{U}(g) = \exp(K^i \hat{Z}_i)$, where $\hat{Z}_i$ is a set of anti-Hermitian operators

representing a basis of the algebra $\mathfrak{g}$ and K^i are real coefficients. We have the commutation relations

$$[\hat{Z}_i, \hat{Z}_j] = c_{ij}^k \hat{Z}_k \,, \tag{4.2}$$

fixed by the structure constants c_{ij}^k of the algebra. Note that here, as in the rest of the chapter, we use Einstein's convention of summing implicitly over all repeated indices.

The action of $\mathcal{U}(g)$ on the operators $\hat{Z}_i$ follows the adjoint representation of the group. More precisely, we have

$$\mathcal{U}^{-1}(g)\, \hat{Z}_i \mathcal{U}(g) = \mathrm{Ad}(g)_i^j\, \hat{Z}_j \,, \tag{4.3}$$

i.e., $\mathcal{U}^{-1}(g)\, \hat{Z}_i \mathcal{U}(g)$ is just a linear combination of operators $\hat{Z}_i$ with the coefficients given by the adjoint matrix $\mathrm{Ad}(g)_i^j$, which is a fixed property of the group[1].

The set $\mathcal{M}_\phi$ of group-theoretic coherent states is then defined as the set of states obtained by acting with all possible $\mathcal{U}(g)$ on a fixed reference state $|\phi\rangle \in \mathcal{H}$, *i.e.*,

$$\mathcal{M}_\phi = \{\, \mathcal{U}(g)\, |\phi\rangle \;:\; g \in \mathcal{G} \} \subset \mathcal{H} \,. \tag{4.4}$$

$\mathcal{M}_\phi$ is determined by the choice of the group $\mathcal{G}$, of its representation $\mathcal{U}$ and of the reference state $|\phi\rangle$. The elements of $\mathcal{M}_\phi$ are parametrized by group elements g. This parametrization may entail some redundancies, as there might exist in $\mathcal{G}$ a stabilizer subgroup for $|\phi\rangle$

$$S_\phi = \{\, g \;:\; \mathcal{U}(g)\, |\phi\rangle = e^{\mathrm{i}\theta}\, |\phi\rangle \} \,, \tag{4.5}$$

i.e., a set of group transformations that leave $|\phi\rangle$ unchanged up to an overall phase, which is irrelevant for what concerns the definition of quantum states. The set of inequivalent group-theoretic coherent states is then isomorphic to the quotient G/S_ϕ.

For our purposes, it is necessary to restrict the possible choices for the reference state $|\phi\rangle$. We will indeed assume that $|\phi\rangle$ is a so-called lowest weight state of the representation $\mathcal{U}$. To understand what is meant by this it is necessary to give some more details about the structure of the algebra operators [104, 105]. We will explain this in the rest of the section.

It is always possible to pick a set of ℓ linearly independent mutually commuting anti-Hermitian operators $\hat{H}_a = H_a^i \hat{Z}_i$, defined by $H_a^i \in \mathbb{R}$ for $a = 1, \ldots, \ell$, such that $[\hat{H}_a, \hat{H}_b] = 0$. In the standard theory of Lie algebras, the space spanned by real linear combinations of $\hat{H}_a$, which we will indicate with $\mathfrak{h}$, is known as a *Cartan subalgebra* of $\mathfrak{g}$. The choice of $\mathfrak{h}$ is not unique, however all possible choices

[1] In particular, if we can write $\mathcal{U}(g) = \exp(K^i \hat{Z}_i)$, then it is straightforward to see that $\mathrm{Ad}(g)_i^j = [\exp \mathrm{ad}(K)]_i^j$ where the matrix $\mathrm{ad}(K)$ is given by $\mathrm{ad}(K)_i^j = K^k c_{ki}^j$.

are isomorphic and will therefore have the same dimension ℓ, known as the *rank* of the algebra. A given a choice of Cartan subalgebra identifies the following structures:

- There exist real vectors $\eta = (\eta_1, \ldots, \eta_\ell) \in \mathbb{R}^\ell$ and corresponding operators $\hat{E}_\eta$ such that
$$[\hat{H}_a, \hat{E}_\eta] = \mathrm{i}\eta_a \hat{E}_\eta \,. \tag{4.6}$$
The operators $\hat{E}_\eta$ will be linear combinations of $\hat{Z}_i$, however they will in general be complex linear combinations and therefore will not be anti-Hermitian operators.

- The vectors η are known as *roots* of the algebra and the operators $\hat{E}_\eta$ as *root space operators*. There is a finite set of non-zero roots which we indicate as Δ. The roots always come in pairs $(\eta, -\eta)$. One can choose a conventional ordering of the roots such that they split into the two disjoint sets of positive roots Δ_+ and negative roots Δ_-, with $\Delta = \Delta_+ \cup \Delta_-$ and $-\eta \in \Delta_-$ for every $\eta \in \Delta_+$.

- Let us indicate with $\mathfrak{g}^{\mathbb{C}}$ the space of all complex linear combinations of algebra elements $\hat{Z}_i$, which is known as the *complexified* Lie algebra. The operators $\hat{H}_a$ together with the operators $\hat{E}_\eta$ span $\mathfrak{g}^{\mathbb{C}}$ under complex linear combinations.

A Hilbert space vector $|\mu\rangle \in \mathcal{H}$ is called a *weight* vector of the representation if it is a common eigenstate of all Cartan subalgebra operators $\hat{H}_a$, *i.e.*, $\hat{H}_a |\mu\rangle = \mathrm{i}\mu_a |\mu\rangle$ for some number $\mu_a \in \mathbb{R}\ \forall a$. Among the weight vectors $|\mu\rangle$ there is a unique one, called the *lowest* weight vector, such that $\hat{E}_\eta |\mu\rangle = 0$ for all negative roots $\eta \in \Delta_-$. From now on we assume that the reference state $|\phi\rangle$ that appears in the definition (4.4) of group-theoretic coherent states is a lowest weight vector $|\mu\rangle$ for a given choice of Cartan subalgebra and root ordering.

Example 10 (Spin-$\frac{1}{2}$ coherent states)**.** *Spin-$\frac{1}{2}$ coherent states are defined with respect to the group* $\mathrm{SU}(2)$ *and algebra* $\mathfrak{su}(2)$*, represented as complex* 2*-by-*2 *matrices. For the algebra, we choose the basis* $\hat{Z}_i = \mathrm{i}\hat{\sigma}_i$ *with* $\hat{\sigma}_i$ *being the well-known Pauli operators. The rank of* $\mathfrak{su}(2)$ *is* 1 *and, as conventional, we choose* $\hat{H} = \frac{\mathrm{i}}{2}\hat{\sigma}_3$ *as basis of the Cartan subalgebra* $\mathfrak{h}$*. For this choice, we have the roots* $\pm\eta = \pm 1$*, with the respective root space operators*

$$\hat{E}_{\pm\eta} = \hat{\sigma}_\pm = \frac{1}{2\sqrt{2}}(\hat{\sigma}_1 \pm \mathrm{i}\hat{\sigma}_2) = \frac{1}{2\sqrt{2}}(-\mathrm{i}\hat{Z}_1 \pm \hat{Z}_2) \,. \tag{4.7}$$

The resulting weight vectors are $|\uparrow\rangle$ *and* $|\downarrow\rangle$ *because they are the eigenvectors of* $\hat{H} = \frac{\mathrm{i}}{2}\hat{\sigma}_3$*. Due to* $\hat{E}_{-\eta} |\downarrow\rangle = 0$*, the state* $|\downarrow\rangle$ *is the lowest weight vector, which we*

thus choose as reference state. The family of group-theoretic coherent states results then from applying all possible group elements $\mathcal{U} \in \mathrm{SU}(2)$ *and is given by*

$$\mathcal{M}_{\mathrm{SU}(2)} = \{e^{\mathrm{i}K^i\hat{\sigma}_i} |{\downarrow}\rangle \ : \ K \in \mathbb{R}^3\} . \tag{4.8}$$

This construction can be readily extended to a system of N *spin-*$\frac{1}{2}$*, in which case the Cartan algebra will be composed of* N *operators* $\hat{H}_n = \frac{\mathrm{i}}{2}\hat{\sigma}_3^n$*, one for each spin* n*, and the lowest weight vector will be* $|\mu\rangle = |{\downarrow}\ldots{\downarrow}\rangle$*. In such case, spin coherent states take the form*

$$e^{\mathrm{i}K^{i,n}\hat{\sigma}_i^n} |{\downarrow}\ldots{\downarrow}\rangle = \prod_{n=1}^{N} \left(e^{\mathrm{i}K^{i,n}\hat{\sigma}_i} |{\downarrow}\rangle_n\right) , \tag{4.9}$$

with the coefficients $K^{i,n}$ *taking values for* $i = 1, 2, 3$ *and for each spin* $n = 1, \ldots, N$ *and* $\hat{\sigma}_i^n$ *representing the* i*-th Pauli operator acting on the* n*-th spin. In other terms, the set of spin-*$\frac{1}{2}$ *coherent states for a system of multiple spins is simply the set of all normalised product states.*

4.2.1 Group-theoretic coherent states as variational manifolds

We now enquire on the use of the manifold of group-theoretic coherent states $\mathcal{M}_\phi$ as a variational submanifold for the Hilbert space $\mathcal{H}$ according to the methods discussed in Chapters 2 and 3. As discussed in the Introduction in Section 1.3, the practical usefulness of any set of variational states depends on the possibility of manipulating them in a way that is computationally inexpensive. This depends on two considerations.

First, the dimension of the manifolds of group-theoretic coherent states is equal to the dimension of the corresponding Lie algebra $\mathfrak{g}$. Indeed, the states are parametrised by Lie group elements, and the Lie group is itself by definition a manifold of dimension $\dim \mathfrak{g}$. Although there are in general no constraints on the dimensions that the algebras can have, in most cases of practical interest the Lie algebras defined on $\mathcal{H}$ have a dimension that scales polynomially in the system size N, guaranteeing an efficient parametrisation of the manifold. In Example 10, for instance, the algebra has dimension $3N$.

Second, we must evaluate the practicability of computing expectation values of relevant observables. Again, in most cases of practical interest, the system's most relevant observables (*e.g.*, the Hamiltonian) can be expressed as polynomials in the algebra operators $\hat{Z}_i$. This means that relation (4.3) can be used to compute expectation values efficiently. Indeed, the expectation value on $\mathcal{U}(g)\,|\mu\rangle$ of any monomial of order d of algebra operators can be expressed as

$$\langle\mu|\mathcal{U}^\dagger(g)\,\hat{Z}_{i_1}\hat{Z}_{i_2}\cdots\hat{Z}_{i_d}\mathcal{U}(g)|\mu\rangle = \mathrm{Ad}(g)^{j_1}_{i_1}\cdots\mathrm{Ad}(g)^{j_d}_{i_d}\ \langle\mu|\hat{Z}_{j_1}\cdots\hat{Z}_{j_d}|\mu\rangle \ , \tag{4.10}$$

where $\langle\mu|\hat{Z}_{j_1}\cdots\hat{Z}_{j_d}|\mu\rangle$ is a d dimensional tensor that can be computed using standard algebra properties. The sum (4.10) then requires evaluating order M^d terms, where M is the dimension of the Lie algebra. In other words, provided that the observables of interest can be expressed as polynomials of a fixed – and possibly small – degree d, then evaluating expectation values for group-theoretic coherent states scales polynomially in the manifold dimensions.

4.2.2 Tangent space and Kähler structure

In the previous chapters we have also explained how the geometric properties of a variational manifold play an important role in its application within variational methods. In particular, the tangent spaces to the manifold and their Kähler properties are a key aspect. We will therefore conclude our discussion of group-theoretic coherent states by analysing their tangent spaces.

Let us consider the tangent space $\mathcal{T}_{|\psi(g_0)\rangle}\mathcal{M}$ at the state $|\psi(g_0)\rangle \equiv \mathcal{U}(g_0)\,|\mu\rangle$, parametrised by the group element $g_0 \in \mathcal{G}$. To better understand this space it is convenient to introduce a new *local* parametrisation of the manifold of group-theoretic coherent states

$$|\psi(x)\rangle = \mathcal{U}(g_0)\exp(x^i\hat{Z}_i)\,|\mu\rangle\ . \tag{4.11}$$

Because of the group property (4.1) this parametrisation locally gives rise to the same manifold as (4.4). However, it is easier to take derivatives with respect to the parameters x, allowing us to compute the corresponding tangent space basis of $\mathcal{T}_{|\psi(g_0)\rangle}\mathcal{M}$ according to (2.24), that is

$$|V_i\rangle = \mathbb{Q}_{\psi(g_0)}\tfrac{\partial}{\partial x^i}\,|\psi(x)\rangle\,\big|_{x=0} \tag{4.12}$$

$$= \mathbb{Q}_{\psi(g_0)}\mathcal{U}(g_0)\,\hat{Z}_i\,|\mu\rangle\ . \tag{4.13}$$

We are thus able to compute a simple closed expression for the tangent vectors at a given point $|\psi(g_0)\rangle$. Of course, if we wanted to compute tangent vectors also at other points, *i.e.*, at $x \neq 0$, then the parametrisation (4.11) would no longer be very helpful. However, in practice such a local description is often already enough, for two reasons. First, for many applications it is enough to focus on one tangent space – usually the tangent space at the approximate ground state. This is for instance the case when computing excitation spectra or spectral functions with the methods discussed in Sections 3.3 and 3.4. Second, even when computing time evolution, in numerical applications this is normally done through a series of small discrete time steps δt. At each time step the local tangent space $\mathcal{T}_{|\psi(g(t_0))\rangle}\mathcal{M}$ can be used to compute $\dot{x}$, which determines the increment of the local parameters $\delta t\,\dot{x}$. The group multiplication rule can then be used to define a new group theoretic coherent

state $|\psi(g(t_0+\delta t))\rangle$ through $g(t_0+\delta t) = g(t_0)\exp\left(\delta t\, \dot{x}^i \hat{Z}_i\right)$. The procedure can then be repeated at each step only involving the local tangent space.

Concerning the Kähler structures and geometric properties of the manifold of group theoretic coherent states, it is possible to also obtain simple closed formulas for the metric and symplectic forms.

Proposition 12. *The restricted Kähler structures of the manifold $\mathcal{M}_\mu$ at the point $|\psi(g)\rangle \equiv \mathcal{U}(g)\,|\mu\rangle$, expressed with respect to the tangent space basis* (4.13)*, are*

$$\boldsymbol{g}_{ij} = -\tfrac{\langle\mu|\hat{Z}_i\mathbb{Q}_\mu\hat{Z}_j+\hat{Z}_j\mathbb{Q}_\mu\hat{Z}_i|\mu\rangle}{\langle\mu|\mu\rangle}\,, \tag{4.14}$$

$$\boldsymbol{\omega}_{ij} = \tfrac{\langle\mu|\hat{Z}_j\mathbb{Q}_\mu\hat{Z}_i-\hat{Z}_i\mathbb{Q}_\mu\hat{Z}_j|\mu\rangle}{\langle\mu|\mu\rangle}\,, \tag{4.15}$$

which are independent of g and thus everywhere the same.

Proof. We can straightforwardly compute

$$\langle V_i|V_j\rangle = \langle\mu|\hat{Z}_i^\dagger\mathcal{U}^\dagger(g)\mathbb{Q}_{\psi(g)}\mathcal{U}(g)\hat{Z}_j|\mu\rangle \tag{4.16}$$

$$= -\langle\mu|\hat{Z}_i\mathbb{Q}_\mu\hat{Z}_j|\mu\rangle\,, \tag{4.17}$$

where we used $\mathcal{U}^\dagger(g)\mathbb{Q}_{\psi(g)}\mathcal{U}(g) = \mathbb{Q}_\mu$ and $\hat{Z}_i^\dagger = -\hat{Z}_i$. □

Whether these structures give rise to a Kähler manifold or not is more complicated to assess. In reference [69] we have shown the following:

- **Semi-simple compact algebra.** If the group $\mathcal{G}$ is compact (and so also the Lie algebra $\mathfrak{g}$ is compact), then the manifold $\mathcal{M}_\mu$ is always Kähler, if $|\mu\rangle$ is the lowest weight state of the representation.

- **Semi-simple non-compact algebra.** If the group is non-compact, not all choices of Cartan subalgebra and lowest weight vector give rise to Kähler manifolds. This has to be assessed by analysing the root structure of the algebra and in particular the *imaginary* or *non-imaginary* nature of roots. See Chapter 5B of reference [69].

4.3 Bosonic Gaussian states

We consider a system of N bosonic modes with creation and annihilation operators $\hat{b}_n^\dagger$ and $\hat{b}_n$. The index $n = 1,\ldots,N$ labels the modes. For each mode we define the following Hermitian operators, known as quadrature operators,

$$\hat{x}_n = \frac{1}{\sqrt{2}}(\hat{b}_n^\dagger + \hat{b}_n)\,, \qquad \hat{p}_n = \frac{\mathrm{i}}{\sqrt{2}}(\hat{b}_n^\dagger - \hat{b}_n)\,, \tag{4.18}$$

and collect them in a single vector

$$\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_N, \hat{p}_1, \ldots, \hat{p}_N)^\intercal \,. \tag{4.19}$$

They satisfy the commutation relations

$$[\hat{\mathbf{x}}^i, \hat{\mathbf{x}}^j] = \mathrm{i}\,\Omega_{ij} \,, \tag{4.20}$$

where Ω is the anti-symmetric matrix

$$\Omega = \begin{pmatrix} 0 & \mathbb{1}_N \\ -\mathbb{1}_N & 0 \end{pmatrix} . \tag{4.21}$$

We identify the *vacuum* state $|0\rangle$ as the state for which $\hat{b}_n |0\rangle = 0$ for all n.

On the bosonic Fock space defined by these modes it is possible to identify two different group representations. We can consider the group of unitary operators $\mathcal{S} = e^{\hat{Q}}$, where $\hat{Q}$ is any anti-Hermitian homogeneous order 2 polynomial in the operators $\hat{\mathbf{x}}$. We will refer to them as *squeezing* operators. Or we can consider the group of unitary operators $\mathcal{D} = e^{\hat{L}}$, where $\hat{L}$ is any anti-Hermitian homogeneous order 1 polynomial in the operators $\hat{\mathbf{x}}$. We will refer to them as *displacement* operators.

Both squeezing and displacement operators give a unitary representation of a Lie group. Squeezing operators $\mathcal{S}$ represent the real symplectic group, while displacement operators $\mathcal{D}$ represent the Heisenberg group. Thus in both cases we can use them to construct a corresponding manifold of group-theoretic coherent states, according to the principles set out in the previous section. As for both representations the lowest weight state is the vacuum $|0\rangle$, we can define the following sets of states:

$$\mathcal{M}_{\text{Coherent}} = \{\mathcal{D}\,|0\rangle \quad \forall \mathcal{D}\} \,, \tag{4.22}$$

$$\mathcal{M}_{\text{Squeezed}} = \{\mathcal{S}\,|0\rangle \quad \forall \mathcal{S}\} \,, \tag{4.23}$$

$$\mathcal{M}_{\text{Gaussian}} = \{\mathcal{D}\mathcal{S}\,|0\rangle \quad \forall \mathcal{D},\, \forall \mathcal{S}\} \,. \tag{4.24}$$

We will refer to the states (4.22), generated by acting on the vacuum with any possible displacement operator, as *bosonic coherent states* and to the states (4.23), generated by acting on the vacuum with any possible squeezing operator, as *bosonic squeezed states*. Finally, we will also consider the states (4.24), generated by acting on the vacuum with any possible combination of displacement *and* squeezing operators. We will refer to these as the full set of *bosonic Gaussian states*.

On a more technical side, note that only the real symplectic group, represented by squeezing operators, is a semi-simple Lie group. The Heisenberg group represented by the displacement operators is not semi-simple. Therefore, strictly

speaking, only squeezed states are truly group-theoretic coherent states according to the definition of Section 4.2. Nonetheless, we will see that for the non semi-simple Heisenberg group it is still possible to make constructions almost fully analogous to the ones that underpin group-theoretic coherent states[2]. To what extent this procedure can be generalised and whether there exists a systematic way to define group-theoretic coherent states also for generic non semi-simple Lie groups is a very interesting question, which however goes beyond the scope of the present work.

In the next sections we will now study more in detail at the two sets of operators – squeezings and displacements – that we have just defined.

4.3.1 Squeezing operators

The squeezing unitary operators $\mathcal{S}$ that we have defined give a unitary representation of the Lie group of real symplectic matrices[3]

$$\mathrm{Sp}(2N,\mathbb{R}) = \{S \in \mathrm{GL}(2N,\mathbb{R}) \ : \ S^\intercal \Omega S = \Omega\}\,, \tag{4.25}$$

with Ω defined in (4.21). Similarly, the set of anti-Hermitian operators $\hat{Q}$ give a representation of the symplectic Lie algebra

$$\mathfrak{sp}(2N,\mathbb{R}) = \{K \in \mathfrak{gl}(2N,\mathbb{R}) \ : \ \Omega K + K^\intercal \Omega = 0\}\,. \tag{4.26}$$

Indeed, for each matrix $K \in \mathfrak{sp}(2N,\mathbb{R})$, one can construct a symmetric matrix $h = \Omega K$ and the corresponding anti-Hermitian Hilbert space operator

$$\hat{Q}(K) = \frac{\mathrm{i}}{2}\,\hat{\mathbf{x}}^\intercal h \hat{\mathbf{x}} = \frac{\mathrm{i}}{2}\,\hat{\mathbf{x}}^\intercal \Omega K \hat{\mathbf{x}}\,. \tag{4.27}$$

Similarly, for any matrix $S \in \mathrm{Sp}(2N,\mathbb{R})$ that can be written as $S = e^K$ for some $K \in \mathfrak{sp}(2N,\mathbb{R})$, one can define the corresponding unitary

$$\mathcal{S}(S) = \mathcal{S}(e^K) = e^{\hat{Q}(K)}\,. \tag{4.28}$$

The operators $\mathcal{S}(S)$ constitute a group representation, in the sense that one can show that[4]

$$\mathcal{S}(S)\,\mathcal{S}(\tilde{S}) = \mathcal{S}(S\tilde{S})\,. \tag{4.29}$$

[2]Indeed, the manifold (4.22) of bosonic coherent states is a very well-behaved set of states that shares all the desirable properties of group-theoretic coherent states and, as the name suggests, represents the original source of inspiration that led to the definition of group-theoretic coherent states.

[3]To be completely precise they are a unitary representation of the double cover of the group $\mathrm{Sp}(2N,\mathbb{R})$, known as the *metaplectic group* $\mathrm{Mp}(2N,\mathbb{R})$.

[4]As discussed in footnote 3 they rigorously constitute a representation only of the double cover of the group. In practice this means that relation (4.29) may be valid only up to a sign. For more detail on how to compute such sign see [106].

Furthermore, the adjoint action of squeezing operators on the quadratures is given by

$$\mathcal{S}^\dagger(S)\,\hat{\mathbf{x}}^i\mathcal{S}(S) = S^i{}_j\,\hat{\mathbf{x}}^j\,. \tag{4.30}$$

The algebraic structure of the quadratic operators $\hat{Q}$ can be further highlighted by noticing that among them we can choose the Cartan operators

$$\hat{H}_n = \mathrm{i}(\hat{b}_n^\dagger\hat{b}_n + \frac{1}{2})\,, \tag{4.31}$$

and root space operators

$$\begin{aligned} &\hat{E}_{+\eta^{(n,m)}} = \mathrm{i}\hat{b}_n^\dagger\hat{b}_m^\dagger, && \hat{E}_{-\eta^{(n,m)}} = \mathrm{i}\hat{b}_n\hat{b}_m, && n\le m \\ &\hat{E}_{+\tilde{\eta}^{(n,m)}} = \hat{b}_n^\dagger\hat{b}_m, && \hat{E}_{-\tilde{\eta}^{(n,m)}} = \hat{b}_n\hat{b}_m^\dagger, && n<m \end{aligned} \tag{4.32}$$

corresponding to the root vectors $\eta_a^{(n,m)} = (\delta_{an}+\delta_{am})$ and $\tilde{\eta}_a^{(n,m)} = (\delta_{an}-\delta_{am})$. All algebra operators $\hat{Q}$ can be expressed as complex linear combinations of the Cartan and root space operators. The lowest weight vector of this representation is indeed the Fock vacuum $|0\rangle$ as it is an eigenstate of all $\hat{H}_n$ and is annihilated by all $\hat{E}_-$.

Finally, a generic squeezing operator $\mathcal{S}(S)$ can always be split as

$$\mathcal{S}(S) = \mathcal{S}(Tu) = \mathcal{S}(T)\,\mathcal{S}(u)\,, \tag{4.33}$$

with u and T satisfying

$$\mathcal{S}(u)\,|0\rangle = e^{\mathrm{i}\theta}\,|0\rangle \quad \text{and} \quad \Omega T = T^{-1}\Omega\,. \tag{4.34}$$

These requirements actually fix a unique solution given[5] by $T=\sqrt{SS^\intercal}$ and $u = T^{-1}S$. Supposing that $T = e^K$, for some symplectic generator K, the condition (4.34) on T is then equivalent to $\{K,\Omega\}=0$. Considering that K is also symplectic, it must then have the form

$$K = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \tag{4.37}$$

[5]Indeed, considering that T should also be symplectic, *i.e.*, $T^\intercal\Omega T=\Omega$, we have that $\Omega T = T^{-1}\Omega$ implies $T=T^\intercal$. The condition $\mathcal{S}(u)\,|0\rangle = e^{\mathrm{i}\theta}\,|0\rangle$ on the other hand implies $uu^\intercal = \mathbb{1}$, as can be seen by considering

$$\begin{aligned} \mathbb{1} &= 2\mathrm{Re}\,\langle 0|\hat{\mathbf{x}}\,\hat{\mathbf{x}}^\intercal|0\rangle = 2\mathrm{Re}\,\langle 0|\mathcal{S}^\dagger(u)\,\hat{\mathbf{x}}\,\hat{\mathbf{x}}^\intercal\,\mathcal{S}(u)|0\rangle && (4.35)\\ &= u\;(2\mathrm{Re}\,\langle 0|\hat{\mathbf{x}}\,\hat{\mathbf{x}}^\intercal|0\rangle)\;u^\intercal = uu^\intercal\,. && (4.36) \end{aligned}$$

Using these two properties one immediately has $SS^\intercal = T^2$.

with A and B real symmetric $N \times N$ matrices. Therefore

$$\mathcal{S}(T) = \exp\left(\frac{\mathrm{i}}{2}\,\Omega_{ij}\,K^{j}{}_{k}\,\hat{\mathbf{x}}^{i}\hat{\mathbf{x}}^{k}\right) \tag{4.38}$$

$$= \exp\left((K_{+})_{nm}\,\hat{b}_{n}^{\dagger}\,\hat{b}_{m}^{\dagger} - (K_{+}^{*})_{nm}\;\hat{b}_{n}\hat{b}_{m}\right)\,, \tag{4.39}$$

with $K_{+} = \frac{1}{2}(A + \mathrm{i}B)$. This implies in particular that a squeezed state $\mathcal{S}(S)\,|0\rangle = e^{\mathrm{i}\theta}\mathcal{S}(T)\,|0\rangle$ does not actually depend on all the information contained in S but rather only on T (or equivalently K_{+}). It is nonetheless for many applications convenient to keep in mind the formalism involving the full symplectic group.

4.3.2 Displacement operators

The displacement operators can be parametrised as

$$\mathcal{D}(\beta) = \exp(\mathrm{i}\beta_{i}\hat{\mathbf{x}}^{i})\,, \tag{4.40}$$

by the set of real numbers $\beta \in \mathbb{R}^{2N}$. They fulfil the group relation

$$\mathcal{D}(\beta)\mathcal{D}(\beta') = \mathcal{D}(\beta+\beta')\exp\left(-\frac{\mathrm{i}}{2}\beta_{i}\Omega_{ij}\beta'_{j}\right)\,, \tag{4.41}$$

and, on the quadrature operators, they have the adjoint action

$$\mathcal{D}^{\dagger}(\beta)\,\hat{\mathbf{x}}^{i}\,\mathcal{D}(\beta) = \hat{\mathbf{x}}^{i} - \Omega_{ij}\beta_{j}\,. \tag{4.42}$$

It may also be useful to split $\mathcal{D}(\beta)$ into terms that contain only creation or only annihilation operators. For this, it follows from (4.41) that

$$\mathcal{D}(\beta) = \mathcal{D}_{+}\mathcal{D}_{-}\mathcal{D}_{0}\,, \tag{4.43}$$

with

$$\mathcal{D}_{+} = \exp\left(\mathrm{i}\beta_{i}\,\mathbb{P}_{+}{}^{i}{}_{j}\,\hat{\mathbf{x}}^{j}\right) = \exp\left(\mathrm{i}\beta_{+}^{n}\,\hat{b}_{n}^{\dagger}\right)\,, \tag{4.44}$$

$$\mathcal{D}_{-} = \exp\left(\mathrm{i}\beta_{i}\,\mathbb{P}_{-}{}^{i}{}_{j}\,\hat{\mathbf{x}}^{j}\right) = \exp\left(\mathrm{i}\beta_{-}^{n}\,\hat{b}_{n}\right)\,, \tag{4.45}$$

$$\mathcal{D}_{0} = \exp\left(\frac{\mathrm{i}}{2}\beta_{i}\,\mathbb{P}_{+}{}^{i}{}_{j}\,\Omega_{jk}\,\mathbb{P}_{-}{}^{l}{}_{k}\,\beta_{l}\right) = \exp\left(-\frac{1}{2}\beta_{+}^{\intercal}\beta_{-}\right)\,, \tag{4.46}$$

where we have introduced the projectors

$$\mathbb{P}_{\pm} = \frac{1}{2}\begin{pmatrix} \mathbb{1}_{N} & \mp\mathrm{i}\mathbb{1}_{N} \\ \pm\mathrm{i}\mathbb{1}_{N} & \mathbb{1}_{N} \end{pmatrix} = \frac{1}{2}(\mathbb{1} \mp \mathrm{i}\Omega)\,, \tag{4.47}$$

which project $\hat{\mathbf{x}}$ on the subspace containing only $\hat{b}^{\dagger}$ or $\hat{b}$ respectively. That is, $\beta_{i}\,\mathbb{P}_{+}{}^{i}{}_{j}\,\hat{\mathbf{x}}^{j} = \beta_{+}^{n}\,\hat{b}_{n}^{\dagger}$ and $\beta_{i}\,\mathbb{P}_{-}{}^{i}{}_{j}\,\hat{\mathbf{x}}^{j} = \beta_{-}^{n}\,\hat{b}_{n}$.

4.3.3 Expectation values and overlaps

Computations with Gaussian states are particularly convenient. It can be shown [59] that the full manifold $\mathcal{M}_{\text{Gaussian}}$ of bosonic Gaussian states can be parametrised by $M = N(N+3)$ real parameters. That is, the number of parameters is polynomial (quadratic) in the number N of bosonic modes of the system.

Furthermore, as discussed in Section 4.2.1 for generic group-theoretic coherent states, computing the expectation value on a Gaussian state of a product of d quadrature operators requires order M^d operations. However, in the case of Gaussian states this can be made even easier by a result known as Wick's theorem [42]. Thanks to it, expressions like (4.10) can be further simplified and decomposed into factors involving only up to two quadrature operators at a time.

Proposition 13 (Wick's theorem)**.** *Consider a product of d operators*

$$\hat{A}_1 \hat{A}_2 \cdots \hat{A}_d \,, \tag{4.48}$$

where each term is a linear combination of quadrature operators, i.e., $\hat{A} = A^{(0)} + A^{(1)}_i \hat{\mathbf{x}}^i$. Let us then take the expectation value of this product on a bosonic Gaussian state $|S, \beta\rangle \equiv \mathcal{D}(\beta)\mathcal{S}(S)\,|0\rangle$. Such expectation value can be expressed as

$$\begin{aligned} \langle S,\beta|\hat{A}_1\hat{A}_2\cdots\hat{A}_d|S,\beta\rangle = {} & \langle\hat{A}_1\rangle\,\langle\hat{A}_2\rangle\cdots\langle\hat{A}_d\rangle \\ & + \langle\hat{A}_1\hat{A}_2\rangle\cdots\langle\hat{A}_d\rangle \\ & + \ldots \textit{(Sum over all possible contractions)}\,. \end{aligned} \tag{4.49}$$

Here, by contractions we mean all possible ways to subdivide the operators in product (4.48) *into groups of one or two operators, and then defining*

$$\langle\hat{A}_i\rangle = \langle S,\beta|\hat{A}_i|S,\beta\rangle \;, \tag{4.50}$$

$$\langle\hat{A}_{i_1}\hat{A}_{i_2}\rangle = \langle S,\beta|\hat{A}_{i_1}\hat{A}_{i_2}|S,\beta\rangle - \langle\hat{A}_{i_1}\rangle\,\langle\hat{A}_{i_2}\rangle \;. \tag{4.51}$$

Proof. For a full proof see, for example, Chapter 8 of reference [107]. The proof essentially relies on decomposing the product (4.48) into terms involving products of creation and annihilation operators and then using their canonical commutation relations to show that the result (4.49) holds for these terms individually. □

This result does not directly improve the scaling of the computation with respect to the degree d of the operator (4.48). Indeed the total number of possible contractions scales factorially in d. However, if d is small, as is usually the case, the number of terms in the sum (4.49) is limited. A great reduction of the overall computational cost is then achieved due to the fact that each term in this sum can

be evaluated just by knowing the one and two point functions

$$\langle \hat{\mathbf{x}}^i \rangle = -\Omega_{ij}\beta_j \,, \tag{4.52}$$

$$\langle \hat{\mathbf{x}}^i \hat{\mathbf{x}}^j \rangle = \frac{1}{2} S^i{}_k S^j{}_l \left(\delta_{kl} + \mathrm{i}\Omega_{kl}\right), \tag{4.53}$$

where we have again assumed that the Gaussian state is parametrised as $\mathcal{D}(\beta)\mathcal{S}(S)\,|0\rangle$. Notice that the one point function always vanishes if the state contains no displacement operator.

Calculations with Gaussian states can often be performed easily also because, on top of expectation values, it is possible to compute overlaps between different Gaussian states efficiently. This is in fact a feature Gaussian states share with all group-theoretic coherent states and which, we will see in Chapter 5, is crucial for defining useful generalisations of Gaussian and group-theoretic coherent states. The procedure that we will sketch now can be indeed seen as a special case of the general one for group-theoretic coherent states contained in Section 5.4.

Proposition 14 (Overlap of bosonic Gaussian states)**.** *Given two bosonic Gaussian states $|S,\beta\rangle \equiv \mathcal{D}(\beta)\mathcal{S}(S)\,|0\rangle$ and $|S',\beta'\rangle \equiv \mathcal{D}(\beta')\mathcal{S}(S')\,|0\rangle$, their overlap is given by*

$$\begin{aligned} r_0 &= \langle S,\beta | S',\beta' \rangle \\ &= \exp\left[-\frac{\mathrm{i}}{4}\mathrm{tr}(\Omega \log\sqrt{\Delta^\intercal\Delta}\Delta^{-1}) + \frac{\mathrm{i}}{2}\beta^\intercal\Omega\beta' \right. \\ &\qquad \left. -\frac{1}{4}(\beta'-\beta)^\intercal S'\mathbf{R}\mathbf{R}^\intercal S'^\intercal(\beta'-\beta)\right] \\ &\qquad \times \left[\det\frac{2\sqrt{\Delta^\intercal\Delta}}{\Delta^\intercal\Delta+\mathbb{1}}\right]^{\frac{1}{4}}, \end{aligned} \tag{4.54}$$

where $\Delta = S^{-1}S'$ and $\mathbf{R} = \mathbb{1} - \mathbb{P}_+(\Delta^\intercal\Delta - \mathbb{1})(\Delta^\intercal\Delta+\mathbb{1})^{-1}\mathbb{P}_-$, with $\mathbb{P}_\pm$ are defined as in (4.47)*.*

Furthermore, the overlap can be computed efficiently even if a product of quadrature operators is inserted between the two states:

$$\langle S,\beta|\hat{\mathbf{x}}^{i_1}\cdots\hat{\mathbf{x}}^{i_d}|S',\beta'\rangle = r_0\,\langle 0|\hat{\tilde{\mathbf{x}}}^{i_1}\cdots\hat{\tilde{\mathbf{x}}}^{i_d}|0\rangle \,. \tag{4.55}$$

Here, the operators $\hat{\tilde{\mathbf{x}}}$ are specific linear combinations of the regular quadrature operators therefore their expectation value can be computed using Wick's theorem based on their one and two point functions

$$\langle 0|\hat{\tilde{\mathbf{x}}}^\intercal\hat{\tilde{\mathbf{x}}}|0\rangle = \frac{1}{2}S'\mathbf{R}(\mathbb{1}+\mathrm{i}\Omega)\mathbf{R}^\intercal S'^\intercal \tag{4.56}$$

$$\langle 0|\hat{\tilde{\mathbf{x}}}|0\rangle = S'\mathbf{R}\,\Omega\,\mathbb{P}_-^\intercal\,\mathbf{R}^\intercal\,S'^\intercal(\beta'-\beta) - \Omega\beta' \tag{4.57}$$

Proof. These quantities can be computed by making use of the fact that any squeezing operator of the form (4.39) can be further decomposed as

$$\mathcal{S}(T) = \mathcal{S}_+ \mathcal{S}_0 \mathcal{S}_- \,, \tag{4.58}$$

where

$$\begin{aligned}
\mathcal{S}_+ &= e^{(A_+)_{nm}\, \hat{a}_n^\dagger\, \hat{a}_m^\dagger} && (4.59)\\
\mathcal{S}_0 &= e^{(A_0)_{nm}\, \hat{a}_n^\dagger\, \hat{a}_m + (A_0^\intercal)_{nm}\, \hat{a}_n\, \hat{a}_m^\dagger} && (4.60)\\
&= e^{2(A_0)_{nm}\, \hat{a}_n^\dagger\, \hat{a}_m + \mathrm{tr} A_0} && (4.61)\\
\mathcal{S}_- &= e^{-(A_+^*)_{nm}\, \hat{a}_n \hat{a}_m} \,. && (4.62)\\
& && (4.63)
\end{aligned}$$

We have that A_+ is defined by the relation

$$2 \begin{pmatrix} \mathrm{Re}A_+ & \mathrm{Im}A_+ \\ \mathrm{Im}A_+ & -\mathrm{Re}A_+ \end{pmatrix} = \tanh \log T = (T^2 - \mathbb{1})(T^2 + \mathbb{1})^{-1} \tag{4.64}$$

and A_0 is calculated as

$$A_0 = \frac{1}{4} \log(\mathbb{1} - 4 A_+ A_+^*) \quad \Rightarrow \quad \mathrm{tr} A_0 = \frac{1}{8} \mathrm{tr} \log \left[2T \left(T^2 + \mathbb{1} \right)^{-1} \right] \,. \tag{4.65}$$

We further have that $\mathcal{S}_- \, \hat{\mathbf{x}}^i = \mathbf{R}^i{}_j \, \hat{\mathbf{x}}^j \, \mathcal{S}_-$. Using there relations repeatedly together with (4.33), (4.43) and (4.41) we can first decompose the displacement and squeezing operators in (4.55) and then commute their components left and right in order to exploit the fact that $\langle 0| \, \mathcal{S}_+ \mathcal{D}_+ = \langle 0|$ and $\mathcal{D}_- \mathcal{S}_- \, |0\rangle = |0\rangle$, leading to the final result. □

4.3.4 Tangent space and Kähler property

Comparing the general discussion on the tangent space of group-theoretic coherent states with the particular form of Gaussian states and especially with the decomposition (4.32), one finds the following structure for the Gaussian tangent space. The tangent space of the full manifold of bosonic Gaussian states at a point given by the state $|S, \beta\rangle \equiv \mathcal{D}(\beta)\mathcal{S}(S) \, |0\rangle$ is given by the real span of the following tangent vectors

$$\begin{aligned}
\mathcal{T}_{|S,\beta\rangle} \mathcal{M}_{\text{Gaussian}} = \mathrm{span}_{\mathbb{R}} \Big\{ & \mathcal{D}(\beta)\mathcal{S}(S) \hat{b}_n^\dagger \, |0\rangle \,, \quad \mathcal{D}(\beta)\mathcal{S}(S) \hat{b}_n^\dagger \hat{b}_m^\dagger \, |0\rangle \,, \\
& \mathrm{i}\mathcal{D}(\beta)\mathcal{S}(S) \hat{b}_n^\dagger \, |0\rangle \,, \quad \mathrm{i}\mathcal{D}(\beta)\mathcal{S}(S) \hat{b}_n^\dagger \hat{b}_m^\dagger \, |0\rangle \,, \quad \forall n \leq m \Big\}
\end{aligned} \tag{4.66}$$

Firstly, notice that this tangent space can be interpreted as containing essentially one and two particle excitations on top of the base Gaussian state. The one particle excitations are due to the displacement operators while the two particle excitations are due to the squeezing terms.

Second, we have highlighted that, for each tangent vector $|V\rangle$, also $\mathrm{i}\,|V\rangle$ is in the tangent space. In other words, $\mathcal{T}_{|S,\beta\rangle}\mathcal{M}_{\text{Gaussian}}$ is a complex linear space. According to Proposition 2, this ensures that bosonic Gaussian states *are a Kähler manifold*.

4.4 Fermionic Gaussian states

We consider a system of N fermionic modes with creation and annihilation operators $\hat{c}_n^\dagger$ and $\hat{c}_n$. The index $n = 1, \ldots, N$ labels the modes. For each mode we define the following Hermitian operators, known as Majorana operators,

$$\hat{\gamma}_n = \frac{1}{\sqrt{2}}(\hat{c}_n^\dagger + \hat{c}_n)\,, \qquad \hat{\tilde{\gamma}}_n = \frac{\mathrm{i}}{\sqrt{2}}(\hat{c}_n^\dagger - \hat{c}_n)\,. \tag{4.67}$$

They play a role analogous to the quadratures in the bosonic case. We collect them in a single vector

$$\hat{\mathbf{x}} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_N, \hat{\tilde{\gamma}}_1, \ldots, \hat{\tilde{\gamma}}_N)^\intercal\,. \tag{4.68}$$

They satisfy the anti-commutation relations

$$\{\hat{\mathbf{x}}^i, \hat{\mathbf{x}}^j\} = \delta_{ij}\,. \tag{4.69}$$

We identify the *vacuum* state $|0\rangle$ as the state for which $\hat{c}_n\,|0\rangle = 0$ for all n.

On the fermionic Fock space defined by these modes it is possible to identify a group representation given by the unitary operators $\mathcal{U} = e^{\hat{Q}}$, where $\hat{Q}$ is any anti-Hermitian homogeneous order 2 polynomial in the operators $\hat{\mathbf{x}}$. We will refer to them as fermionic Gaussian unitaries.

These unitaries belong to a representation of the real orthogonal group. Thus we can use them to construct a corresponding manifold of group-theoretic coherent states, according to the principles set out in Section 4.2. For the representation the lowest weight state is the vacuum $|0\rangle$, so we can define the following set of states:

$$\mathcal{M}_{\text{f.Gaussian}} = \{\mathcal{U}\,|0\rangle\ , \quad \forall \mathcal{U}\}\,. \tag{4.70}$$

We will refer to these states as *fermionic Gaussian states*.

In the next sections we will now study more in detail the fermionic Gaussian unitaries that we have just defined.

4.4.1 Fermionic Gaussian unitaries

The fermionic Gaussian unitaries $\mathcal{U}$ belong to a unitary representation of the Lie group of real orthogonal matrices

$$\mathrm{O}(2N,\mathbb{R}) = \{G \in \mathrm{GL}(2N,\mathbb{R}) \ : \ G^\intercal G = \mathbb{1}\}\,. \tag{4.71}$$

Similarly, the set of anti-Hermitian operators $\hat{Q}$ gives a representation of the special orthogonal Lie algebra

$$\mathfrak{so}(2N,\mathbb{R}) = \{K \in \mathfrak{gl}(2N,\mathbb{R}) \ : \ K + K^\intercal = 0\}\,. \tag{4.72}$$

Indeed, for each matrix $K \in \mathfrak{so}(2N,\mathbb{R})$, one can construct the corresponding anti-Hermitian Hilbert space operator

$$\hat{Q}(K) = \frac{1}{2}\,\hat{\mathbf{x}}^\intercal K \hat{\mathbf{x}}\,. \tag{4.73}$$

Similarly, for any matrix $G \in \mathrm{O}(2N,\mathbb{R})$ that can be written as $G = e^K$ for some $K \in \mathfrak{so}(2N,\mathbb{R})$, one can define the corresponding unitary

$$\mathcal{U}(G) = \mathcal{U}(e^K) = e^{\hat{Q}(K)}\,. \tag{4.74}$$

Here, it is important to notice that, contrarily to the bosonic case, the group $\mathrm{O}(2N,\mathbb{R})$ is not fully connected. There exist two disconnected components, corresponding to matrices with $\det G = \pm 1$. It turns out that they generate two disconnected components of fermionic Gaussian states, corresponding to states with even and odd fermion number parity. Only the matrices in the component connected to the identity, *i.e.*, the subgroup $\mathrm{SO}(2N,\mathbb{R})$, can actually be written as $G = e^K$. In order to be able to represent also the elements with $\det G = -1$, we can choose any $2N$-dimensional vector v, normalised such that $v^\intercal v = 2$ and introduce the unitary $\mathcal{U}(G_v) = v_i \hat{\mathbf{x}}^i$, representing the matrix $G_v = vv^\intercal - \mathbb{1}$. We see that $\det G_v = -1$ and that, for any $G \in \mathrm{O}(2N,\mathbb{R})$ with $\det G = -1$, we can write $\mathcal{U}(G) = \mathcal{U}(G_v)\mathcal{U}(e^K)$ for some $K \in \mathfrak{so}(2N,\mathbb{R})$.

The operators $\mathcal{U}(G)$ constructed this way constitute a group representation, in the sense that one can show that

$$\mathcal{U}(G)\,\mathcal{U}(\tilde{G}) = \mathcal{U}(G\tilde{G})\,. \tag{4.75}$$

Furthermore, the adjoint action of fermionic Gaussian unitaries on the Majorana operators is given by

$$\mathcal{U}^\dagger(G)\,\hat{\mathbf{x}}^i \mathcal{U}(G) = G^i{}_j\,\hat{\mathbf{x}}^j\,. \tag{4.76}$$

The algebraic structure of the quadratic operators $\hat{Q}$ can be further highlighted by noticing that among them we can choose the Cartan operators

$$\hat{H}_n = \mathrm{i}(\hat{c}_n^\dagger \hat{c}_n - \frac{1}{2})\,, \tag{4.77}$$

and root space operators

$$\begin{aligned} \hat{E}_{+\eta^{(n,m)}} &= \mathrm{i}\hat{c}_n^\dagger \hat{c}_m^\dagger, & \hat{E}_{-\eta^{(n,m)}} &= \mathrm{i}\hat{c}_n \hat{c}_m, & n<m \\ \hat{E}_{+\tilde{\eta}^{(n,m)}} &= \hat{c}_n^\dagger \hat{c}_m, & \hat{E}_{-\tilde{\eta}^{(n,m)}} &= \hat{c}_n \hat{c}_m^\dagger, & n<m \end{aligned} \tag{4.78}$$

corresponding to the root vectors $\eta_a^{(n,m)} = (\delta_{an} + \delta_{am})$ and $\tilde{\eta}_a^{(n,m)} = (\delta_{an} - \delta_{am})$. All algebra operators $\hat{Q}$ can be expressed as complex linear combinations of the Cartan and root space operators. The lowest weight vector of this representation is indeed the Fock vacuum $|0\rangle$ as it is an eigenstate of all $\hat{H}_n$ and is annihilated by all $\hat{E}_-$.

Finally, a generic squeezing operator $\mathcal{U}(G)$ can always be split as

$$\mathcal{U}(G) = \mathcal{U}(Tu) = \mathcal{U}(T)\,\mathcal{U}(u)\,, \tag{4.79}$$

with u and T satisfying

$$\mathcal{U}(u)\,|0\rangle = e^{\mathrm{i}\theta}\,|0\rangle \quad \text{and} \quad \Omega T = T^{-1}\Omega\,. \tag{4.80}$$

These requirements actually fix a unique solution given[6] by $T = \sqrt{-G\Omega G^\intercal \Omega}$ and $u = T^{-1}G$, with Ω defined as in the bosonic case (4.21). Supposing that $T = e^K$, for some anti-symmetric generator K, the condition (4.80) on T is then equivalent to $\{K, \Omega\} = 0$. Considering that K is also anti-symmetric, it must then have the form

$$K = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \tag{4.83}$$

with A and B real anti-symmetric $N \times N$ matrices. Therefore

$$\mathcal{U}(T) = \exp\left(\frac{1}{2}\,K_{ij}\,\hat{\mathbf{x}}^i \hat{\mathbf{x}}^j\right) \tag{4.84}$$

$$= \exp\left((K_+)_{nm}\,\hat{c}_n^\dagger\,\hat{c}_m^\dagger + (K_+^*)_{nm}\;\hat{c}_n \hat{c}_m\right)\,, \tag{4.85}$$

[6]Indeed, considering that T should also be orthogonal, *i.e.*, $T^\intercal T = \mathbb{1}$, we have that $\Omega T = T^{-1}\Omega$ implies $T\Omega = \Omega T^\intercal$. The condition $\mathcal{S}(u)\,|0\rangle = e^{\mathrm{i}\theta}\,|0\rangle$ on the other hand implies $u\Omega u^\intercal = \Omega$, as can be seen by considering

$$\mathbb{1} = 2\mathrm{Im}\,\langle 0|\hat{\mathbf{x}}\,\hat{\mathbf{x}}^\intercal|0\rangle = 2\mathrm{Im}\,\langle 0|\mathcal{U}^\dagger(u)\,\hat{\mathbf{x}}\,\hat{\mathbf{x}}^\intercal\,\mathcal{U}(u)|0\rangle \tag{4.81}$$

$$= u\;(2\mathrm{Im}\,\langle 0|\hat{\mathbf{x}}\,\hat{\mathbf{x}}^\intercal|0\rangle)\;u^\intercal = u\Omega u^\intercal\,. \tag{4.82}$$

Using these two properties one immediately has $-G\Omega G^\intercal \Omega = T^2$.

with $K_+ = \frac{1}{2}(A + \mathrm{i}B)$. This implies in particular that a gaussian state $\mathcal{U}(G)\,|0\rangle = e^{\mathrm{i}\theta}\mathcal{U}(T)\,|0\rangle$ does not actually depend on all the information contained in G but rather only on T (or equivalently K_+). It is nonetheless for many applications convenient to keep in mind the formalism involving the full orthogonal group.

4.4.2 Expectation values and overlaps

Computations with fermionic Gaussian states are as convenient as in the bosonic case. The manifold $\mathcal{M}_{\text{f.Gaussian}}$ of fermionic Gaussian states has dimension $M = N(N-1)$ [59]. That is, the number of parameters is polynomial (quadratic) in the number N of fermionic modes of the system. Similarly to the case of bosonic Gaussian states, the evaluation of the expectation value of products of d Majorana operators requires order M^d operations are greatly simplified by Wick's theorem

Proposition 15 (Wick's theorem)**.** *Consider a product of d operators*

$$\hat{A}_1\hat{A}_2\cdots\hat{A}_d\,, \tag{4.86}$$

where each term is a linear combination of Majorana operators, i.e., $\hat{A} = A^{(0)} + A^{(1)}_i\hat{\mathbf{x}}^i$. Let us then take the expectation value of this product on a fermionic Gaussian state $|G\rangle \equiv \mathcal{U}(G)\,|0\rangle$. Such expectation value can be expressed as

$$\begin{aligned}\langle S,\beta|\hat{A}_1\hat{A}_2\cdots\hat{A}_d|S,\beta\rangle =&(-1)^s\,\langle\hat{A}_1\hat{A}_2\rangle\cdots\langle\hat{A}_{d-1}\hat{A}_d\rangle\\ &+(-1)^{s'}\,\langle\hat{A}_1\hat{A}_3\rangle\,\langle\hat{A}_2\hat{A}_4\rangle\cdots\langle\hat{A}_{d-1}\hat{A}_d\rangle\\ &+\ldots\,\textit{(Sum over all possible contractions)}\,.\end{aligned} \tag{4.87}$$

Here, by contractions we mean all possible ways to subdivide the operators in product (4.86) *into groups of two operators, and then defining*

$$\langle\hat{A}_{i_1}\hat{A}_{i_2}\rangle = \langle G|\hat{A}_{i_1}\hat{A}_{i_2}|G\rangle\ . \tag{4.88}$$

The signs s are given by the sign of the permutation that is necessary to reorder the initial product (4.86) *into the contractions.*

Proof. The proof is based on the same ideas as in the bosonic case. See, for example, Chapter 8 of reference [107]. □

The total number of possible contractions scales factorially in d. However, if d is small, as is usually the case, the number of terms in the sum (4.87) is limited. A great reduction of the overall computational cost is then achieved due to the fact that each term in this sum can be evaluated just by knowing the two point functions

$$\langle\hat{\mathbf{x}}^i\hat{\mathbf{x}}^j\rangle = \frac{1}{2}G^i{}_kG^j{}_l\,(\delta_{kl} + \mathrm{i}\Omega_{kl})\,, \tag{4.89}$$

where we have again assumed that the Gaussian state is parametrised as $\mathcal{U}(G)\,|0\rangle$. Notice that the one point function always vanishes on fermionic Gaussian states.

Calculations with Gaussian states can often be performed easily also because, on top of expectation values, it is possible to compute overlaps between different Gaussian states efficiently. This is in fact a feature Gaussian states share with all group-theoretic coherent states and which, we will see in Chapter 5, is crucial for defining useful generalisations of Gaussian and group-theoretic coherent states.

As in the bosonic case, there exists a procedure to compute efficiently the overlaps of different fermionic Gaussian states. This is a special case of the general one for group-theoretic coherent states contained in Section 5.4.

Proposition 16 (Overlap of fermionic Gaussian states)**.** *Given two fermionic Gaussian states parametrised as* $|G\rangle \equiv \mathcal{U}(G)\,|0\rangle$ *and* $|G'\rangle \equiv \mathcal{U}(G')\,|0\rangle$*, their overlap is given by*

$$r_0 = \langle G|G'\rangle = \exp\left[\frac{\mathrm{i}}{4}\mathrm{tr}(\Omega \log\sqrt{-\Omega\Delta^{\intercal}\Omega\Delta}\Delta^{-1})\right]\left[\det\frac{2\sqrt{-\Omega\Delta^{\intercal}\Omega\Delta}}{\mathbb{1}-\Omega\Delta^{\intercal}\Omega\Delta}\right]^{\frac{1}{4}}, \tag{4.90}$$

where $\Delta = G^{-1}G'$ *and* $\mathbf{R} = \mathbb{1} - \mathbb{P}_+(\Omega\Delta^{\intercal}\Omega\Delta + \mathbb{1})(\Omega\Delta^{\intercal}\Omega\Delta - \mathbb{1})^{-1}\mathbb{P}_-$*, with* $\mathbb{P}_\pm$ *are defined as in the bosonic case* (4.47)*.*

Furthermore, the overlap can be computed efficiently even if a product of Majorana operators is inserted between the two states:

$$\langle G|\hat{\mathbf{x}}^{i_1}\cdots\hat{\mathbf{x}}^{i_d}|G'\rangle = r_0\,\langle 0|\hat{\tilde{\mathbf{x}}}^{i_1}\cdots\hat{\tilde{\mathbf{x}}}^{i_d}|0\rangle\ . \tag{4.91}$$

Here, the operators $\hat{\tilde{\mathbf{x}}}$ *are specific linear combinations of the regular quadrature operators therefore their expectation value can be computed using Wick's theorem based on their two point functions*

$$\langle 0|\hat{\tilde{\mathbf{x}}}^{\intercal}\hat{\tilde{\mathbf{x}}}|0\rangle = \frac{1}{2}G'\mathbf{R}(\mathbb{1}+\mathrm{i}\Omega)\mathbf{R}^{\intercal}G'^{\intercal} \tag{4.92}$$

Proof. These quantities can be computed by making use of the fact that any unitary operator of the form (4.85) can be further decomposed as

$$\mathcal{U}(T) = \mathcal{U}_+\,\mathcal{U}_0\,\mathcal{U}_-\,, \tag{4.93}$$

where

$$\mathcal{U}_+ = e^{(A_+)_{nm}\,\hat{c}_n^\dagger\,\hat{c}_m^\dagger} \tag{4.94}$$

$$\mathcal{U}_0 = e^{(A_0)_{nm}\,\hat{c}_n^\dagger\,\hat{c}_m - (A_0^{\intercal})_{nm}\,\hat{c}_n\,\hat{c}_m^\dagger} \tag{4.95}$$

$$= e^{2(A_0)_{nm}\,\hat{c}_n^\dagger\,\hat{c}_m - \mathrm{tr}A_0} \tag{4.96}$$

$$\mathcal{U}_- = e^{(A_+^*)_{nm}\,\hat{c}_n\hat{c}_m}\,. \tag{4.97}$$

$$\tag{4.98}$$

We have that A_+ is defined by the relation

$$2\begin{pmatrix} \mathrm{Re}A_+ & \mathrm{Im}A_+ \\ \mathrm{Im}A_+ & -\mathrm{Re}A_+ \end{pmatrix} = \tanh\log T = (T^2-\mathbb{1})(T^2+\mathbb{1})^{-1} \tag{4.99}$$

and A_0 is calculated as

$$A_0 = \frac{1}{4}\log(\mathbb{1}-4A_+A_+^*) \quad \Rightarrow \quad \mathrm{tr}A_0 = \frac{1}{8}\mathrm{tr}\log\left[(2T\,(T^2+\mathbb{1})^{-1}\right]. \tag{4.100}$$

We further have that $\mathcal{U}_-\,\hat{\mathbf{x}}^i = \mathbf{R}^i{}_j\,\hat{\mathbf{x}}^j\,\mathcal{U}_-$. Using there relations repeatedly together with (4.79), we can first decompose the unitary operators in (4.91) and then commute their components left and right in order to exploit the fact that $\langle 0|\,\mathcal{U}_+ = \langle 0|$ and $\mathcal{U}_-\,|0\rangle = |0\rangle$, leading to the final result. □

4.4.3 Tangent space and Kähler property

Comparing the general discussion on the tangent space of group-theoretic coherent states with the particular form of Gaussian states and especially with the decomposition (4.78), one finds the following structure for the Gaussian tangent space. The tangent space of the manifold of fermionic Gaussian states at a point given by the state $|G\rangle \equiv \mathcal{U}(G)\,|0\rangle$ is given by the real span of the following tangent vectors

$$\mathcal{T}_{|S,\beta\rangle}\mathcal{M}_{\text{f.Gaussian}} = \mathrm{span}_{\mathbb{R}}\left\{\mathcal{U}(G)\hat{c}_n^\dagger\hat{c}_m^\dagger\,|0\rangle\ ,\ \mathrm{i}\mathcal{U}(G)\hat{c}_n^\dagger\hat{c}_m^\dagger\,|0\rangle\ ,\quad \forall n<m\right\} \tag{4.101}$$

Firstly, notice that this tangent space can be interpreted as containing essentially two particle excitations on top of the base Gaussian state. Second, we have highlighted that, for each tangent vector $|V\rangle$, also $\mathrm{i}\,|V\rangle$ is in the tangent space. In other words, $\mathcal{T}_{|S,\beta\rangle}\mathcal{M}_{\text{f.Gaussian}}$ is a complex linear space. According to Proposition 2, this ensures that fermionic Gaussian states *are a Kähler manifold*.

Chapter 5

Beyond Gaussian states

In this chapter we will discuss the possibility of defining classes of variational states that go beyond Gaussian states, while maintaining some of their most useful properties, in particular the easiness of computations.

Most of the material presented in this chapter was published in reference [101]:

T. Guaita, L. Hackl, T. Shi, E. Demler, J.I. Cirac,
Generalization of group-theoretic coherent states for variational calculations,
Phys. Rev. Research **3**, 023090 (2021), used under CC BY 4.0.

5.1 Overview

In Chapters 2 and 3 we have discussed the important role that sets of variational states play in the classical simulation of quantum many body systems. In Chapter 4 we have discussed the example of bosonic and fermionic Gaussian states as variational states, stressing particularly how they allow to perform variational calculations at a low computational cost. However, Gaussian states also have limitations. For example, Gaussian states have only limited forms of correlations, similar to the ones of free systems. Also other commonly used variational *ansätze* present drawbacks, for instance the use of MPS is limited to 1D geometries.

The goal of this chapter is to define extensions of some existing variational *ansätze*, that continue to satisfy the property of efficient computations, but contain more correlations or can be used for higher dimensional systems. To do this we will be inspired by two observations: (i) there exist known classes of states that extend Gaussian states [52] or spin product states [68, 108] to contain more correlations while continuing to admit easy computations of expectation values; (ii) Gaussian states, bosonic coherent states and some classes of product states can all be understood within a unified framework, namely the framework of group-theoretic coherent states, that we have anticipated in the previous chapter in

Section 4.2.

The group-theoretic perspective provides some very powerful tools that allow us to show how all the properties of the constructed families of states are actually encoded purely in the algebraic structures of the chosen groups and representations. Conveniently, several frequently used families of quantum states can be understood as instances of group-theoretic coherent states resulting from different choices of Lie groups. Standard bosonic coherent states arise from the Heisenberg group, bosonic and fermionic Gaussian states arise from representations of the groups $\mathrm{Sp}(2N, \mathbb{R})$ and $\mathrm{O}(2N, \mathbb{R})$, while atomic coherent states [64] arise from the two dimensional representation of SU(2).

Exploiting these available group-theoretical structures, in Section 5.2 we thus consistently define an extension for all families of group-theoretic coherent states, which we will refer to as *generalised group-theoretic coherent states*. We achieve this by applying to the coherent states a single unitary transformation $\mathcal{V}(M) = \exp(-\frac{\mathrm{i}}{2} M^{ab} \hat{H}_a \hat{H}_b)$, where $\hat{H}_b$ represents Cartan subalgebra operators and the matrix M contains additional variational parameters.

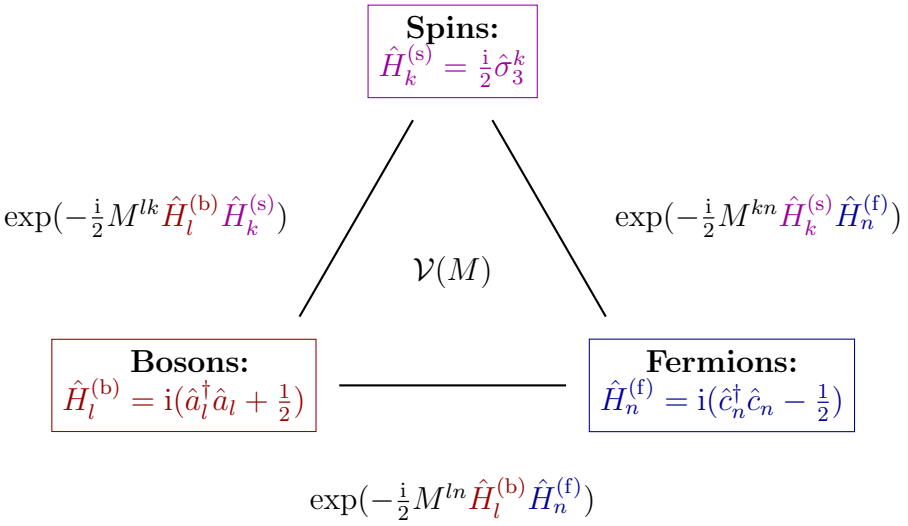
In Sections 5.3 and 5.4, we will then show how the specific form of this extension is designed to preserve the desirable feature of being able to compute expectation values efficiently. In fact, all necessary operations are performed in terms of objects (matrices and vectors) whose dimension is at most the one of the Lie group. In most examples, this dimension scales polynomially with the size of the considered system, making our methods feasible even for studying large systems and exploring the thermodynamic limit.

While satisfying this constraint, the extension also enlarges the range of available types of quantum correlations, going thus beyond mean field treatments. Indeed, the exponent of $\mathcal{V}(M)$, which is quadratic in algebra operators, can represent structures not present in coherent states. For example, it can be used to introduce non trivial density-density correlations in Gaussian states or spin-spin correlations in spin systems. Furthermore, in composite systems it can produce entanglement between different types of degrees of freedom (spins, bosons, fermions) as it can contain products of Cartan subalgebra operators from the different sectors, as sketched in figure 5.1.

The proposed construction is very general, in the sense that it can be applied to group-theoretic coherent states associated to any choice of Lie group. For this reason, we will give all definitions in a sufficiently general language that does not refer to a specific Lie group and algebra, but rather makes use of the general notation introduced in Section 4.2 of the previous chapter. To make the rather formal construction more concrete, we will illustrate each step for two paradigmatic examples, namely spin-$\frac{1}{2}$ coherent states and bosonic Gaussian states. In Section 5.5, we then give a more complete discussion of the applications where we believe this

construction could reveal itself most promising.

Figure 5.1: We show schematically how a unitary operator $\mathcal{V}(M)$ (edges of the triangle) can generate entanglement in composite systems between different sectors, *e.g.*, bosonic, fermionic or spin sectors (vertices of the triangle).



5.2 Generalised group-theoretic coherent states

In this section, we will define new families of states, which we refer to as *generalised group-theoretic coherent states*, that extend the families of group-theoretic coherent states described in Section 4.2.

Definition

We choose a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$, spanned by the operators $\hat{H}_a$ as defined in Section 4.2. Let us then consider the unitary operator

$$\mathcal{V}(M) = \exp\left(\frac{\mathrm{i}}{2}M^{ab}\hat{H}_a\hat{H}_b\right), \tag{5.1}$$

where, as we will do in the rest of this chapter, we used Einstein's convention of implicitly summing over repeated indices. The real symmetric matrix M^{ab} defines a bilinear form on $\mathfrak{h}$ and contains $\ell(\ell+1)/2$ real parameters that define the operator.

The exponent of (5.1) is not an element of the Lie algebra $\mathfrak{g}$, as it is quadratic in the basis operators $\hat{Z}_i$. Consequently, $\mathcal{V}(M)$ is not a group transformation and the product of more operators of this type does not follow a group multiplication rule. Furthermore, the action of a transformation $\mathcal{V}(M)$ will in general take an element of $\mathcal{M}_\phi$ out of the set of group-theoretic coherent states.

We now define the class of generalised group-theoretic coherent states as the set of states of the form

$$|\psi(g_1, g_2, M)\rangle = \mathcal{U}(g_1)\,\mathcal{V}(M)\,\mathcal{U}(g_2)\,|\mu\rangle\ . \tag{5.2}$$

The states are conveniently parametrized by two group elements g_1 and g_2 and one bilinear form M, although this parametrization will contain several redundancies. Similarly to group-theoretic coherent states, this class of states is determined by the choice of the group $\mathcal{G}$ and of its representation $\mathcal{U}$ on Hilbert space. In the case of compact Lie groups any choice of Cartan subalgebra and lowest weight state $|\mu\rangle$ will define the same family of states[1].

Example 11 (generalised spin-$\frac{1}{2}$ coherent states)**.** *Based on Example 10, we consider a system of N spin-$\frac{1}{2}$ degrees of freedom with Cartan algebra spanned by $\hat{H}_n = \frac{\mathrm{i}}{2}\sigma_3^n$. The unitary operator* (5.1) *takes the form*

$$\mathcal{V}(M) = \exp\left(-\frac{\mathrm{i}}{8} M_{nm}\,\hat{\sigma}_3^n \hat{\sigma}_3^m\right)\ , \tag{5.3}$$

for any given $N \times N$ real symmetric matrix M. The generalised *spin-$\frac{1}{2}$ coherent states take the form*

$$|\psi(K_1, K_2, M)\rangle = \mathcal{U}(K_1)\,\mathcal{V}(M)\,\mathcal{U}(K_2)\,|\downarrow\cdots\downarrow\rangle\ , \tag{5.4}$$

where, similarly to Example 10, the group unitaries are defined as

$$\mathcal{U}(K) = \exp\left(\mathrm{i} K^{i,n}\hat{\sigma}_i^n\right)\ , \tag{5.5}$$

with the coefficients $K^{i,n}$ taking values for $i = 1, 2, 3$ and for each spin $n = 1, \ldots, N$.

[1]This is because in this case all Cartan subalgebras and lowest weight states are equivalent up to group unitary transformations, which can be absorbed in to the parameters g_1 and g_2. In the case of non-compact Lie groups there may instead exist unitarily inequivalent classes of Cartan subalgebras. Their choice is therefore relevant. Note that the Cartan subalgebra choice with respect to which operator (5.1) is defined may even be different from the one with respect to which the lowest weight state $|\mu\rangle$ is defined.

Example 12 (Generalised bosonic Gaussian states)**.** *We consider a system of N bosonic modes with Cartan algebra spanned by $\hat{H}_n = \mathrm{i}(\hat{b}_n^\dagger \hat{b}_n + \frac{1}{2})$, as discussed in Section 4.3. The unitary operator* (5.1) *takes the form*

$$\mathcal{V}(M) = \exp\left(-\frac{i}{2} M^{nm} (\hat{b}_n^\dagger \hat{b}_n + \tfrac{1}{2})(\hat{b}_m^\dagger \hat{b}_m + \tfrac{1}{2})\right) , \tag{5.6}$$

for any given $N \times N$ real symmetric matrix M. The generalised *bosonic squeezed states take the form*

$$|\psi(S_1, S_2, M)\rangle = \mathcal{S}(S_1)\, \mathcal{V}(M)\, \mathcal{S}(S_2)\, |0\rangle \ , \tag{5.7}$$

where $\mathcal{S}(S)$ are the squeezing unitaries discussed in Section 4.3.1. We recognize that these states constitute one of the classes of non-Gaussian states previously introduced in [52].

Entangling degrees of freedom in composite systems

The construction of group-theoretic coherent states is possible also in the case in which different groups act on different sectors of a composite system. In this case the construction of *generalised* group-theoretic coherent states is particularly useful, because, as mentioned in the overview, it enables us to entangle and correlate the different types of degrees of freedom in the system, such as spins, bosons and fermions. This provides a distinct advantage over coherent states alone, which are always product states over the different system components, described by the different groups (special unitary group for spin, symplectic group for bosons, orthogonal group for fermions).

More precisely, let us assume that we have two semi-simple Lie groups $\mathcal{G}_1$ and $\mathcal{G}_2$, such that the respective representations act on a tensor product of Hilbert spaces $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and thus commute with each other, *i.e.*, we have a representation of the product group $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2$ with Lie algebra $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$. By applying the construction of group-theoretic coherent states, we will find that the Cartan subalgebra $\mathfrak{h} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$ is the direct sum of the respective Cartan subalgebras. Following our definition of generalised coherent states, the transformation $\mathcal{V}(M)$ will contain three terms, *i.e.*,

$$\frac{\mathrm{i}}{2} M^{ab} \hat{H}_a \hat{H}_b = \frac{\mathrm{i}}{2} \big(M^{ab}_{(1)} \hat{H}^{(1)}_a \hat{H}^{(1)}_b + M^{ab}_{(2)} \hat{H}^{(2)}_a \hat{H}^{(2)}_b + 2 M^{ab}_{(12)} \hat{H}^{(1)}_a \hat{H}^{(2)}_b \big) , \tag{5.8}$$

where $H_a^{(i)} \in \mathfrak{h}_i$. We thus see explicitly that the last term is a product of Cartan generators associated to the two different original groups. As our representation acts on a tensor product, this last term in $\mathcal{V}(M)$ will be responsible for entangling degrees of freedom associated to different parts of a composite system. This is

particularly relevant when $\mathcal{G}_1$ and $\mathcal{G}_2$ are associated to different types of physical degrees of freedom, such as spins, bosons and fermions.

Example 13 (Entangling spin-$\frac{1}{2}$ and bosonic systems)**.** *Let us consider a system composed of N spin-$\frac{1}{2}$ degrees of freedom, as described in Example 10, and $\tilde{N}$ bosonic modes, as described in Section 4.3. The total Lie group acting on it will be given by $\mathcal{G} = \mathrm{SU}(2)^N \times \mathrm{Sp}(2\tilde{N}, \mathbb{R})$. The corresponding Cartan subalgebra is given by the span of all the operators*

$$\hat{H}^{(1)}_n = \frac{\mathrm{i}}{2}\hat{\sigma}^n_3, \qquad \hat{H}^{(2)}_{\tilde{n}} = \mathrm{i}(\hat{b}^\dagger_{\tilde{n}}\hat{b}_{\tilde{n}} + \frac{1}{2})\,. \tag{5.9}$$

Consequently the unitary $\mathcal{V}(M)$ takes the form

$$\begin{aligned}\mathcal{V}(M) = \exp\Bigg[-\frac{\mathrm{i}}{8}M^{(1)}_{nm}\,\hat{\sigma}^n_3\hat{\sigma}^m_3 \\ - \frac{i}{2}M^{\tilde{n}\tilde{m}}_{(2)}(\hat{b}^\dagger_{\tilde{n}}\hat{b}_{\tilde{n}} + \tfrac{1}{2})(\hat{b}^\dagger_{\tilde{m}}\hat{b}_{\tilde{m}} + \tfrac{1}{2}) \\ -\frac{i}{2}M^{(12)}_{n\tilde{m}}\,\hat{\sigma}^n_3(\hat{b}^\dagger_{\tilde{m}}\hat{b}_{\tilde{m}} + \tfrac{1}{2})\Bigg]\,.\end{aligned} \tag{5.10}$$

In particular we see that the last term generates entanglement between the spin and bosonic degrees of freedom.

5.3 Standard form of expectation values

Our definition was carefully chosen, such that we can efficiently compute the expectation value of physical observables $\hat{\mathcal{O}}$ of interest (*e.g.*, Hamiltonians). To achive this, we first show that these expectation values can be brought into a *standard form*, which can then be evaluated efficiently.

Here, we assume that the group was chosen, such that $\hat{\mathcal{O}}$ can be expressed as a polynomial in the operators $\hat{Z}_i$, which can be accomplished in most physical systems. Then, any such expectation value can be brought into the standard form

$$\langle\psi|\hat{\mathcal{O}}|\psi\rangle = \sum_{d,\{i\}} C^{i_1\ldots i_d}\,\langle\mu|\,\mathcal{U}(g_d)\,\hat{Z}_{i_1}\cdots\hat{Z}_{i_d}|\mu\rangle\ . \tag{5.11}$$

To reach this standard form, we need to commute $\mathcal{U}_1 \equiv \mathcal{U}(g_1)$, $\mathcal{V}$ and $\mathcal{U}_2 \equiv \mathcal{U}(g_2)$ through the operators $\hat{Z}_i$ that appear in $\hat{\mathcal{O}}$ according to

$$\langle\psi|Z_{i_1}\ldots Z_{i_d}|\psi\rangle = \langle\mu|\mathcal{U}_2^\dagger\mathcal{V}^\dagger\mathcal{U}_1^\dagger\,\underbrace{Z_{i_1}\ldots Z_{i_d}\mathcal{U}_1\mathcal{V}\mathcal{U}_2}|\mu\rangle\,. \tag{5.12}$$

This will only transform the operators $\hat{Z}_i$ or generate additional group unitaries $\mathcal{U}(g_i)$, which can all be collected to the left to form the single unitary $\mathcal{U}(g_d)$. To do this, we need the following two commutation rules:

- **Commuting $\mathcal{U}$ with $\hat{Z}_i$:**
  From (4.3) we have that commuting group transformations with algebra operators only gives rise to linear combinations according to

$$\hat{Z}_i \mathcal{U}(g) = \mathrm{Ad}(g)_i^j \mathcal{U}(g)\, \hat{Z}_j \,. \tag{5.13}$$

- **Commuting $\mathcal{V}$ with $\hat{Z}_i$:**
  Even though $\mathcal{V}(M)$ is not a group transformation, its action on algebra elements has a simple form. Indeed, from relation (4.6) it follows that

$$\hat{E}_\eta\, \mathcal{V}(M) = \mathcal{V}(M)\, e^{\eta_a M^{ab} \hat{H}_b - \frac{\mathrm{i}}{2}\eta_a M^{ab}\eta_b} \hat{E}_\eta \tag{5.14}$$

$$= \mathcal{V}(M)\, e^{\mathrm{i}\theta_\eta}\, \mathcal{U}(e^{K_\eta}) \hat{E}_\eta \,, \tag{5.15}$$

  where in the second line we have recognised that the exponential can be decomposed into a complex phase factor $\theta_\eta = -\frac{1}{2}\eta_a M^{ab}\eta_b$ and the exponential of a real linear combination of algebra operators $K_\eta = \eta_a M^{ab} H_b$. Furthermore we have that

$$\hat{H}_a \mathcal{V}(M) = \mathcal{V}(M) \hat{H}_a \,, \tag{5.16}$$

  as $\mathcal{V}(M)$ is a function exclusively of Cartan subalgebra operators and therefore commutes with $\hat{H}_a$. As all algebra operators $\hat{Z}_i$ can be expressed as complex linear combinations of operators of the types $\hat{H}_a$ or $\hat{E}_\eta$, it follows that the commutation of $\mathcal{V}(M)$ through $\hat{Z}_i$ will be a linear combination of (5.16) and (5.15).

By combining a series of operations of these kinds, we can always commute the unitaries $\mathcal{U}_1$, $\mathcal{V}$ and $\mathcal{U}_2$ in (5.12) through any monomial of operators $\hat{Z}_i$. They will then combine with the corresponding $\mathcal{U}_1^\dagger$, $\mathcal{V}^\dagger$ and $\mathcal{U}_2^\dagger$ coming from the bra vector $\langle\psi|$ yielding identities and leaving a linear combination of terms of the form $\mathcal{U}(g)\, \hat{Z}_{j_1} \cdots \hat{Z}_{j_d}$.

More specifically, the unitaries $\mathcal{V}(M)$ will give rise to a series of group transformations $e^{\theta_{\eta_i}} \mathcal{U}(e^{K_{\eta_i}})$ according to (5.15). Then one has to commute all $\mathcal{U}(e^{K_{\eta_i}})$ to the left using using (5.13), which will produce linear combinations of $\mathcal{U}(e^{K_{\eta_i}})\, \hat{Z}_{j_1} \cdots \hat{Z}_{j_d}$. Once all the group transformations are on the left side, they combine with $\mathcal{U}(g) = \mathcal{U}(e^{K_{\eta_{i_1}}}) \ldots \mathcal{U}(e^{K_{\eta_{i_d}}})$. Thus, the action of $\mathcal{V}(M)$ on a monomial of algebra operators $\hat{Z}_i$ will give rise to a polynomial of the same order multiplied with a single group transformation $\mathcal{U}(g)$ from the left.

In summary, any expectation value of an observable $\hat{\mathcal{O}}$ can be brought into the standard form (5.11), whose efficient evaluation will be subject of the next section. This enables the application of the full range of variational methods

when using generalised group-theoretic coherent states as an approximation of the true state of the system[2]. The specific form of definition (5.1) – which at first sight may appear somewhat arbitrary – was fundamental for achieving this. Indeed, the inclusion in the exponent of (5.1) of algebra elements outside of the Cartan subalgebra or of non-quadratic terms would make it impossible to express the transformations (5.13) and (5.15) exclusively in terms of algebra and group operators, and thus would prevent the subsequent calculations.

Example 14 (Commutation rules for generalised spin-$\frac{1}{2}$ coherent states)**.** *The operators $\mathcal{U}(K)$ and $\mathcal{V}(M)$, defined in Example 11 satisfy the following relations:*

$$\begin{pmatrix} \hat{\sigma}_1^n \\ \hat{\sigma}_2^n \\ \hat{\sigma}_3^n \end{pmatrix} \mathcal{U}(K) = \mathcal{U}(K)\, e^{-2K^{i,n}\mathbf{L}_i} \begin{pmatrix} \hat{\sigma}_1^n \\ \hat{\sigma}_2^n \\ \hat{\sigma}_3^n \end{pmatrix}, \tag{5.17}$$

corresponding to (5.13)*, where we have the* 3*-by-*3 *matrices $(\mathbf{L}_i)_{kl} = \epsilon_{ikl}$, with ϵ_{ikl} being the totally antisymmetric tensor; and*

$$\hat{\sigma}_3^n\, \mathcal{V}(M) = \mathcal{V}(M)\, \hat{\sigma}_3^n\,, \tag{5.18}$$

$$\hat{\sigma}_\pm^n\, \mathcal{V}(M) = \mathcal{V}(M)\, e^{-\frac{\mathrm{i}}{2}M_{nn}} e^{\pm\frac{\mathrm{i}}{2}M_{nm}\hat{\sigma}_3^m}\, \hat{\sigma}_\pm^n\,, \tag{5.19}$$

corresponding to (5.16) *and* (5.15)*.*

Example 15 (Commutation rules for generalised Gaussian states)**.** *The commutation of the operators $\mathcal{S}(S)$, discussed in Section 4.3 and Example 12, with any creation or annihilation operator can be achieved through*

$$\mathcal{S}^\dagger(S)\, \hat{\mathbf{x}}\, \mathcal{S}(S) = S\hat{\mathbf{x}}\,, \tag{5.20}$$

where $\hat{\mathbf{x}} = (\hat{q}_1, \ldots, \hat{q}_N, \hat{p}_1, \ldots, \hat{p}_N)^\intercal$ and $\hat{q}_n = (\hat{b}_n^\dagger + \hat{b}_n)/\sqrt{2}$ and $\hat{p}_n = \mathrm{i}(\hat{b}_n^\dagger - \hat{b}_n)/\sqrt{2}$ are canonical quadrature operators. The commutation of $\mathcal{V}(M)$, discussed in Examples 12, with creation or annihilation operators can be achieved through

$$\begin{aligned} \mathcal{V}^\dagger(M)\hat{b}_n\hat{b}_m\mathcal{V}(M) = {} & e^{-\frac{\mathrm{i}}{2}(M^{nn}+M^{nm}+M^{mn}+M^{mm})} \\ & \times e^{-\mathrm{i}(M^{nl}+M^{ml})(\hat{b}_l^\dagger\hat{b}_l+\frac{1}{2})}\, \hat{b}_n\hat{b}_m\,, \end{aligned} \tag{5.21}$$

$$\begin{aligned} \mathcal{V}^\dagger(M)\hat{b}_n^\dagger\hat{b}_m\mathcal{V}(M) = {} & e^{-\frac{\mathrm{i}}{2}(M^{nn}-M^{nm}-M^{mn}+M^{mm})} \\ & \times e^{\mathrm{i}(M^{nl}-M^{ml})(\hat{b}_l^\dagger\hat{b}_l+\frac{1}{2})}\, \hat{b}_n^\dagger\hat{b}_m\,, \end{aligned} \tag{5.22}$$

[2]The careful reader will know by now that to apply the full range of known variational methods to a given family of quantum states (*e.g.*, as described in Chapters 2 and 3), it is not always sufficient to be able to compute the expectation values of the Hamiltonian. It is also necessary to compute quantities involving so-called *tangent vectors*. In appendix A.3, however, we show that for generalised group-theoretic coherent states also these quantities can be simply brought to the standard form (5.11).

and the corresponding conjugate relations, which follow from (5.15)*. Combining transformations of these types, the expectation value on the states* (5.7) *of any polynomial of creation and annihilation operators can be brought to the standard form of linear combinations of*

$$\langle 0|\mathcal{S}(S)\hat{\mathbf{x}}_{i_1}\cdots\hat{\mathbf{x}}_{i_d}|0\rangle \ . \tag{5.23}$$

5.4 Efficient evaluation of expectation values in standard form

Generalised group-theoretic coherent states will only be useful as variational families if we can efficiently evaluate expectation values $\langle\psi|\hat{\mathcal{O}}|\psi\rangle$. In the previous section, we have shown that any such expectation value can be reduced to the standard form (5.11). To evaluate this standard form, we need to be able to compute its building blocks of the form

$$\langle\mu|\,\mathcal{U}(g)\,\hat{Z}_{i_1}\cdots\hat{Z}_{i_d}|\mu\rangle \ . \tag{5.24}$$

In this section, we will discuss how to compute (5.24) efficiently and thereby evaluate arbitrary expectation values from the standard form (5.11).

BCH decomposition

Computing (5.24) can be achieved by performing a normal ordered Baker-Campbell-Hausdorff decomposition, also known as Gauss decomposition, of the group unitary $\mathcal{U}(g)$ that appears in it. Let us assume that $\mathcal{U}(g)$ can be written as an exponential of algebra elements. We therefore have

$$\mathcal{U}(g) = \exp\left(\sum_{\eta\in\Delta_+} K_+^\eta\hat{E}_\eta + K_0^a\hat{H}_a + \sum_{\eta\in\Delta_+} K_-^\eta\hat{E}_{-\eta}\right) \ , \tag{5.25}$$

where we have used that the algebra operators $\hat{Z}_i$ can be decomposed on the basis $\hat{H}_a, \hat{E}_\eta$ and we have introduced the corresponding complex coefficients K_0^a, $K_\pm^\eta$. We would like to split the exponential appearing in (5.25) into the product of three terms and rewrite $\mathcal{U}(g)$ as

$$\mathcal{U}(g) = \hat{T}_+\,\hat{T}_0\,\hat{T}_- \ , \tag{5.26}$$

where $\hat{T}_\pm$ and $\hat{T}_0$ are operators of the forms

$$\hat{T}_\pm = \exp\left(\sum_{\eta\in\Delta_+} A_\pm^\eta\,\hat{E}_{\pm\eta}\right) , \qquad \hat{T}_0 = \exp\left(A_0^a\,\hat{H}_a\right) , \tag{5.27}$$

for some appropriate choice of the coefficients A_0^a, $A_\pm^\eta$.

The specific functional dependence of A_0^a and $A_\pm^\eta$ on K_0^a and $K_\pm^\eta$ and the extent to which it can be calculated analytically will depend on the given choice of the group $\mathcal{G}$. However, let us point out that the decomposition (5.26) only depends on the abstract group and algebra properties and not on the specific choice of representation. It may therefore be convenient to perform such decomposition working in a smaller representation than the one of the physical system, *e.g.*, the fundamental or adjoint representation.

Once the decomposition (5.26) of $\mathcal{U}(g)$ has been performed the computation of the expectation value (5.24) becomes relatively straightforward. Indeed, one can commute $\hat{T}_-$ to the right of the algebra operators $\hat{Z}_{i_1}\cdots\hat{Z}_{i_d}$ just giving rise to new linear combinations of algebra operators. To do this one needs a relation analogous[3] to equation (5.13), *i.e.*,

$$\hat{T}_-\,\hat{Z}_i = \mathbf{R}_i^j\,\hat{Z}_j\,\hat{T}_-\,. \tag{5.28}$$

In this way, one reduces (5.24) to the form

$$\begin{aligned}\mathbf{R}_{i_1}^{j_1}\cdots\mathbf{R}_{i_d}^{j_d}\,\langle\mu|\hat{T}_+\hat{T}_0\,\hat{Z}_{j_1}\cdots\hat{Z}_{j_d}\,\hat{T}_-|\mu\rangle \\ = e^{\mathrm{i}A_0^a\mu_a}\,\mathbf{R}_{i_1}^{j_1}\cdots\mathbf{R}_{i_d}^{j_d}\,\langle\mu|\hat{Z}_{i_1}\cdots\hat{Z}_{i_d}|\mu\rangle\ ,\end{aligned} \tag{5.29}$$

where we used that the lowest weight vector $|\mu\rangle$ is left-invariant by $\hat{T}_-$ on the right, right-invariant by $\hat{T}_+$ on the left and is an eigenstate with eigenvalue $\mathrm{i}\mu_a$ of the operators $\hat{H}_a$ that appear in $\hat{T}_0$. Let us stress again that the eigenvalues μ_a are the only object in this derivation that depends on the choice of representation that we are using.

The information on the group element g appearing in the original expression (5.24) is contained in the linear coefficients $\mathbf{R}_i^j$ (which will depend on A_-^η) and in the coefficients A_0^a that appear in the first factor of (5.29). The factor $\langle\mu|\hat{Z}_{i_1}\cdots\hat{Z}_{i_d}|\mu\rangle$ is instead independent of g and thus needs to be computed only once. This can be done using the standard algebra commutation relations.

Example 16 (BCH for spin-$\frac{1}{2}$ coherent states)**.** *As $\mathcal{U}(g)$ is always a tensor product over individual spin degrees of freedom, we can evaluate the standard form of the expectation value for each one individually. We thus consider*

$$\langle\downarrow|e^{\mathrm{i}K^i\sigma_i}\hat{\sigma}_{i_1}\ldots\hat{\sigma}_{i_d}|\downarrow\rangle\ . \tag{5.30}$$

[3]Formula (5.28) and the form of matrix $\mathbf{R}$ can be derived in the same way as (4.3) and (5.13) as explained in footnote 1. Note that we have here the quantity $\hat{T}_-$ (instead of $\mathcal{U}(g)$) which is not a unitary operator, but is still the exponential of complex combinations of algebra elements.

The BCH decomposition of $e^{\mathrm{i}K^i\sigma_i}$ *is well-known [64, 109] and explicitly given by*

$$e^{K_+\hat{\sigma}_+ + \mathrm{i}\frac{K_0}{2}\hat{\sigma}_3 - K_+^*\hat{\sigma}_-} = e^{A_+\hat{\sigma}_+} e^{\frac{A_0}{2}\hat{\sigma}_3} e^{A_-\hat{\sigma}_-}\,, \tag{5.31}$$

where the respective coefficients are given by

$$A_0 = -2\log\left(\cos\varphi - \tfrac{1}{2}K_0\tfrac{\sin\varphi}{\varphi}\right) \tag{5.32}$$

$$A_+ = A_-^* = -\mathrm{i}K_+\tfrac{\sin\varphi}{\varphi}\left(\cos\varphi - \tfrac{1}{2}K_0\tfrac{\sin\varphi}{\varphi}\right)^{-1}, \tag{5.33}$$

with $\varphi = \sqrt{|K_+|^2 + \frac{1}{4}K_0^2}$*. To find the equivalent of* (5.29)*, we can use* (5.17) *to deduce* $e^{A_-\hat{\sigma}_-}\ \hat{\sigma}_i = \mathbf{R}_{ij}\,\hat{\sigma}_j\ e^{A_-\hat{\sigma}_-}$ *with*

$$\mathbf{R} = \begin{pmatrix} 1 - \frac{1}{4}A_-^2 & \frac{\mathrm{i}}{4}A_-^2 & \frac{1}{\sqrt{2}}A_- \\ \frac{\mathrm{i}}{4}A_-^2 & 1 + \frac{1}{4}A_-^2 & -\frac{\mathrm{i}}{\sqrt{2}}A_- \\ -\frac{1}{\sqrt{2}}A_- & \frac{\mathrm{i}}{\sqrt{2}}A_- & 1 \end{pmatrix}. \tag{5.34}$$

Combining these results, we thus find

$$\langle\downarrow|e^{\mathrm{i}K^i\hat{\sigma}_i}\hat{\sigma}_{i_1}\cdots\hat{\sigma}_{i_d}|\downarrow\rangle = e^{sA_0}\ \mathbf{R}_{i_1j_1}\cdots\mathbf{R}_{i_dj_d}\,\langle\downarrow|\hat{\sigma}_{j_1}\cdots\hat{\sigma}_{j_d}|\downarrow\rangle \tag{5.35}$$

with $s = -\frac{1}{2}$ *for spin-*$\frac{1}{2}$*, which generalises easily to larger spin.*

Example 17 (BCH for bosonic Gaussian states)**.** *To evaluate* (5.23) *via BCH we first can decompose the unitary as* $\mathcal{S}(S) = \mathcal{S}(u)\mathcal{S}(T)$*, where* $\langle 0|\,\mathcal{S}(u) = e^{-\mathrm{i}\theta}\,\langle 0|$ *and*

$$\mathcal{S}(T) = \exp\left((K_+)_{nm}\,\mathrm{i}\,\hat{b}_n^\dagger\,\hat{b}_m^\dagger + (K_+^*)_{nm}\ \mathrm{i}\,\hat{b}_n\hat{b}_m\right), \tag{5.36}$$

for a suitable K_+*. For this type of unitary the decomposition* $\mathcal{U}(T) = \hat{T}_+\hat{T}_0\hat{T}_-$ *is known analytically [43]. Using this decomposition one can obtain the final result*

$$\langle 0|\,\mathcal{S}(S)\hat{\mathbf{x}}_{i_1}\cdots\hat{\mathbf{x}}_{i_d}|0\rangle = r_0\mathbf{R}_{i_1j_1}\cdots\mathbf{R}_{i_dj_d}\langle 0|\hat{\mathbf{x}}_{j_1}\cdots\hat{\mathbf{x}}_{j_d}|0\rangle\,, \tag{5.37}$$

where r_0 *is given by*

$$r_0 = e^{-\frac{\mathrm{i}}{4}\mathrm{tr}(\Omega\log\sqrt{S^\intercal S}S^{-1})}\det(\mathbb{1} - 4A_+A_+^*)^{\frac{1}{4}}\,, \tag{5.38}$$

and $\mathbf{R}$ *is the* $2N\times 2N$ *matrix*

$$\mathbf{R} = \begin{pmatrix} \mathbb{1} - A_+^* & -\mathrm{i}A_+^* \\ -\mathrm{i}A_+^* & \mathbb{1} + A_+^* \end{pmatrix}. \tag{5.39}$$

The matrix A_+ *can be derived analytically from* S *according to* (4.64)*. See Section 4.3 for more details.*

Time evolution of the BCH decomposition

In the previous section, we showed how to compute (5.24) which required a normal ordered Baker-Campbell-Hausdorff decomposition of $\mathcal{U}(g)$ for every g. For many standard Lie groups, the needed formulas already exist in the literature. However, this decomposition can also be computed by solving a corresponding set of differential equations. This approach can be used if the respective closed analytical formulas are not known or difficult to implement and is especially convenient in settings where one performs time evolution.

Time evolution is an important application of generalised group-theoretic coherent states, where one uses them to simulate the dynamics of quantum systems, either in real time or imaginary time. A similar setting is the one where one applies gradient descent methods to our family of states. In all these applications one has the need to compute a certain set of expectation values at each time step of the evolution, then update the state to a new one which is (theoretically) infinitesimally close and repeat the procedure. Therefore, it is required to calculate the decomposition (5.26) at a series of subsequent time steps as g evolves as a function of time (more precisely, g is a function of the variational parameters which in turn evolve as functions of time). In these settings, it would be useful if one could compute the BCH decomposition for $\mathcal{U}(g(t+dt))$ based on the decomposition of $\mathcal{U}(g(t))$ at the previous time step, instead of having to compute it from scratch at each step. We will now show how this can be done. As already mentioned above, this will also lead to a general method for computing (5.26), that, although not always the most efficient, can be useful in cases where a closed formula is not available.

Let us assume that $\mathcal{U}(g(t))$ can be written as

$$\mathcal{U}(g(t)) = e^{K^i(t)\,\hat{Z}_i} \tag{5.40}$$

and that we want to decompose it as

$$\mathcal{U}(g(t)) = \hat{T}_+(t)\,\hat{T}_0(t)\,\hat{T}_-(t)\,, \tag{5.41}$$

where $\hat{T}_-(t)$, $\hat{T}_0(t)$ and $\hat{T}_+(t)$ are operators of the forms

$$\hat{T}_-(t) = e^{\sum_{\eta\in\Delta_+} A_-^\eta(t)\,\hat{E}_{-\eta}}\,, \tag{5.42a}$$

$$\hat{T}_0(t) = e^{A_0^a(t)\,\hat{H}_a}\,, \tag{5.42b}$$

$$\hat{T}_+(t) = e^{\sum_{\eta\in\Delta_+} A_+^\eta(t)\,\hat{E}_{\eta}}\,. \tag{5.42c}$$

We now take the time derivative of $\mathcal{U}(g(t))$ and multiply it by $\mathcal{U}^{-1}(g(t))$.

From (5.40), we have

$$\mathcal{U}^{-1}(g(t))\frac{d}{dt}\mathcal{U}(g(t)) = \int_0^1 d\tau\, e^{-\tau K^j(t)\,\hat{Z}_j}\left[\frac{d}{dt}K^i(t)\,\hat{Z}_i\right] e^{\tau K^j(t)\,\hat{Z}_j} \tag{5.43}$$

$$= \left[\int_0^1 d\tau\, e^{\tau\,\mathrm{ad}(K(t))}\right]^i_j \frac{d}{dt}K^j(t)\,\hat{Z}_i \tag{5.44}$$

$$= \left[\mathrm{ad}(K(t))^{-1}\left(e^{\mathrm{ad}(K(t))} - \mathbb{1}\right)\right]^i_j \frac{d}{dt}K^j(t)\,\hat{Z}_i \tag{5.45}$$

where $\mathrm{ad}(K(t))$ represents the matrix

$$[\mathrm{ad}(K(t))]^i_j = K^k(t)\,c^i_{kj}\,, \tag{5.46}$$

similarly to what explained in footnote 1. For the expression used in (5.43) see, *e.g.*, the appendix of [52]. From (5.41), we instead have

$$\begin{aligned}\mathcal{U}^{-1}(g(t))\frac{d}{dt}\mathcal{U}(g(t)) = \hat{T}_-(t)^{-1}\hat{T}_0(t)^{-1}\Big[\sum_{\eta\in\Delta_+} d^\eta_+(t)\hat{E}_\eta\Big]\hat{T}_0(t)\,\hat{T}_-(t)\\ + \hat{T}_-(t)^{-1}\left[d^a_0(t)\hat{H}_a\right]\hat{T}_-(t) + \Big[\sum_{\eta\in\Delta_+} d^\eta_-(t)\hat{E}_{-\eta}\Big]\,.\end{aligned} \tag{5.47}$$

The coefficients $d^a_0(t)$ and $d^\eta_\pm(t)$ are defined by[4]

$$\frac{d}{dt}\hat{T}_-(t) = \hat{T}_-(t)\Big[\sum_{\eta\in\Delta_+} d^\eta_-(t)\hat{E}_{-\eta}\Big]\,, \tag{5.48a}$$

$$\frac{d}{dt}\hat{T}_0(t) = \hat{T}_0(t)\left[d^a_0(t)\hat{H}_a\right]\,, \tag{5.48b}$$

$$\frac{d}{dt}\hat{T}_+(t) = \hat{T}_+(t)\Big[\sum_{\eta\in\Delta_+} d^\eta_+(t)\hat{E}_\eta\Big]\,. \tag{5.48c}$$

By applying relations analogous to (5.28), equation (5.47) can be brought to the form of a linear combination of the algebra basis operators $\hat{Z}_i$, similarly to (5.45).

Finally, comparing these algebra elements, one can write $d^a_0(t)$ and $d^\eta_\pm(t)$ as functions of $\frac{d}{dt}K^a_0$ and $\frac{d}{dt}K^\eta_\pm$ and of $A^a_0(t)$ and $A^\eta_\pm(t)$. More precisely, equating (5.47) and (5.45) leads to

$$\begin{aligned}\mathbf{M}[A_0(t), A_-(t)]\begin{pmatrix} d^\eta_-(t)\\ d^a_0(t)\\ d^\eta_+(t)\end{pmatrix}\\ = \left[\mathrm{ad}(K(t))^{-1}\left(e^{\mathrm{ad}(K(t))} - \mathbb{1}\right)\right]\frac{d}{dt}\begin{pmatrix} K^\eta_-(t)\\ K^a_0(t)\\ K^\eta_+(t)\end{pmatrix}\,,\end{aligned} \tag{5.49}$$

[4]Note that in general $d^\eta_\pm(t) \neq \frac{d}{dt}A^\eta_\pm(t)$, because not all $\hat{E}_\eta$ commute among themselves.

where $\mathbf{M}[A_0(t), A_-(t)]$ is a matrix of the dimension of the algebra, that depends on $A_0(t)$ and $A_-(t)$ through the adjoint representation of the corresponding group elements, and which we need to invert.

Note that here the derivatives $\frac{d}{dt}K(t)$ depend only on how we update the variational parameters at the given time step and how this update influences $g(t)$. We therefore assume them to be known. Similarly, the quantities $K(t)$, $A_0(t)$ and $A_-(t)$ depend only on the group element $g(t)$ and on its BCH decomposition at the current time step. Having found $d_0^a(t)$, $d_\pm^\eta(t)$ from equation (5.49), we can then integrate equations (5.48a) to (5.48c) for one time step to obtain the BCH decomposition (5.41) at time $t + dt$.

If instead we just want to compute the Baker-Campbell-Hausdorff decomposition for a fixed group transformation of the form (5.25), we can write $K^i(t) = tK^i$ and integrate from $t = 0$ to $t = 1$ the corresponding differential equations (5.48a) to (5.48c) as described in this section to obtain the desired decomposition (5.26).

5.5 Applications

In this manuscript, we have introduced *generalised group-theoretic coherent states* as a new family of pure quantum states. This family is defined on top of the well-known *group-theoretic coherent states* by applying an additional unitary $\mathcal{V}(M)$. There exist many examples of group-theoretic coherent states, defined by different choices of Lie groups and representations, and this makes our construction quite general and applicable in various contexts.

The transformation $\mathcal{V}(M)$ is defined as the exponential of a quadratic expression in the so-called Cartan subalgebra operators $\hat{H}_a$. This introduces quantum correlations not contained in traditional group-theoretic coherent states, thus allowing the treatment of problems beyond mean-field. The dynamics of regular group-theoretic coherent states correspond to the group-theoretic version of semi-classical Landau-Lifshitz (LL) equations for SU(2) spin models [110]. Our new class of wavefunctions allows in this sense to go beyond semi-classical dynamics. In particular, we expect generalised coherent states to be suitable for systems with interacting Hamiltonians containing terms also quadratic in Cartan operators. For these, it will be interesting to explore whether the many exact theoretical results that have been proven for the Landau-Lifshitz equations, such as existence of solitons in one dimension, will be be robust to going beyond the LL factorizable wavefunction ansatz. We further emphasized that generalised group theoretic states are particularly powerful when we want to correlate different types of degrees of freedom (*e.g.*, spins, bosons, fermions) in composite systems, as the transformation $\mathcal{V}(M)$ can be used to entangle them by including Cartan generators of different types.

While going beyond coherent states, we showed in section 5.4 that generalised coherent states still allow for an efficient evaluation of generic expectation values. We stress, however, that computing the overlap $\langle\psi|\tilde{\psi}\rangle$ between two arbitrary generalised group-theoretic coherent states $|\psi\rangle$ and $|\tilde{\psi}\rangle$ remains in general a hard task.

We gave two key examples of how our construction can be applied in different settings, namely for spin-$\frac{1}{2}$ coherent states and bosonic Gaussian states. However, the range of applications of our proposal is by no means limited to these examples: they can be extended, combined or complemented in many ways. The SU(2) construction can, for instance, be extended to higher spin representations, for example to atomic coherent states [64] obtaining so-called *spin squeezed states* [68]. The Gaussian state construction can be repeated for *fermionic* Gaussian states, defined in Section 4.4.

It is also straightforward to apply the described generalisation to more elaborate Lie groups and algebras [111, 112]. This is particularly useful as many lattice systems can be described as an SU(N) problem, where N is the dimension of the Hilbert space at a site [65, 66]. Our approach can thus be used to study dynamics with variational states that have non-trivial entanglement utilizing this SU(N) perspective. Finally, a further interesting possibility is that of defining $\mathcal{V}(M)$ in terms of a choice of Cartan subalgebra different from the one with respect to which the reference state $|\mu\rangle$ is a lowest weight state, which can be done for non-compact Lie groups, such as $\mathrm{Sp}(2N, \mathbb{R})$ for bosonic Gaussian states.

We currently restricted ourselves to semi-simple Lie groups, as those are the ones studied systematically in mathematical physics and for which the construction of Cartan subalgebra and root system is fully understood. While this enabled us to present a systematic framework of generalised group-theoretic coherent states, we know that in special cases we can follow the same philosophy also for Lie groups that are not semi-simple. The most prominent example is the Heisenberg group associated to displacement operators for bosonic degrees of freedom, which plays the key role in the definition of regular bosonic coherent states. It will be an interesting exercise to explore the full extent to which this group can be incorporated in our formalism and consider whether the same can be done for other non-semi-simple groups.

Some of the examples discussed above have already been proposed and studied [52]. A few of them already have a history of successful applications. For example, by choosing a fermionic number operator $\hat{n}_\mathrm{f}$ and a bosonic quadrature operator $\hat{p} = \frac{\mathrm{i}}{\sqrt{2}}(\hat{b}^\dagger - \hat{b})$ as Cartan-type generators we obtain a $\mathcal{V}(M)$ that corresponds to the well-known Lang-Firsov Polaron transformation [113], often used for correlated boson-fermion systems. However, the presented framework can lead to a whole spectrum of new generalisations which we believe can be of great interest.

In terms of concrete applications, we believe that interesting developments can come from two directions. First, as our states are particularly amenable to being produced in common experimental implementations *and* their expectation values can be computed efficiently by classical computation, they provide an ideal setting for benchmarking experimental set-ups and quantum computer prototypes. Second, they can be applied as variational states to describe and understand ground state and dynamical properties of many quantum many-body systems. Some families of states that can be understood as generalised coherent states have already been successfully employed to perform both exact and variational calculations [54, 71, 108, 114], testifying to the large spectrum of potential applications of the construction. In particular, they include systems that contain bosons or fermions or both, for which our construction allows to go beyond a Gaussian approach and also caters for the necessity of entangling the bosonic and fermionic sectors. One can also consider systems where a spin impurity is coupled to a bosonic, fermionic or spin bath, such as the paradigmatic Kondo [115, 116] and Bose polaron models [117, 118]. We can finally take in consideration pure spin problems for which tensor network methods do not give satisfactory results, *e.g.*, in higher dimensions.

Some specific systems of the types above for which we believe generalised coherent states would represent an interesting novelty include the case of fermions with bi-phonon coupling [119], where the interaction is given by $\hat{H}_{\mathrm{e-ph}} = \sum_i \hat{Q}_i^{(\mathrm{f})} \hat{Q}_i^{(\mathrm{b})}$, where $\hat{Q}_i^{(\mathrm{f})}$ and $\hat{Q}_i^{(\mathrm{b})}$ are respectively fermionic and bosonic quadratic operators. Of interest is also the case of the Jahn-Teller polaron [120] where, after a Lee-Low-Pines transformation [121], the Hamiltonian takes the form $\hat{H}_{\mathrm{e-ph}} = \sum_i \hat{F}_i \hat{q}_i$. Here, the $\hat{q}_i$ are quadratures of a bosonic bath and the $\hat{F}_i$ are a set of fermionic operators realizing an $\mathfrak{su}(2)$ algebra, which could be described by generalised spin-$\frac{1}{2}$ coherent states.

Part III

Chapter 6

Application: the Bose-Hubbard model

Here, we will present a concrete application of the methods discussed in the previous chapters. We will show how the variational methods of Chapters 2 and 3 can be applied to the manifold of bosonic Gaussian states, defined in Chapter 4, to study the well-known Bose-Hubbard model. This will highlight how, using the full power of variational methods, even a relatively simple manifold like Gaussian states is enough to reveal many non-trivial properties of the model in its superfluid phase.

Most of the material of this chapter was published in reference [122]:

T. Guaita, L. Hackl, T. Shi, C. Hubig, E. Demler, J.I. Cirac,
Gaussian time-dependent variational principle for the Bose-Hubbard model,
Phys. Rev. B **100**, 094529,

6.1 Overview

The Bose-Hubbard model provides a theoretical description of interacting cold atoms in optical lattices [123], which in the last years have proven to be a promising experimental platform. Its Hamiltonian is given by

$$\hat{H} = -\sum_{\langle n,m\rangle} \hat{b}_n^\dagger \hat{b}_m + \frac{U}{2}\sum_n \hat{b}_n^\dagger \hat{b}_n^\dagger \hat{b}_n \hat{b}_n - \mu \sum_n \hat{b}_n^\dagger \hat{b}_n \,, \tag{6.1}$$

where $\hat{b}_n^\dagger$ and $\hat{b}_n$ are the bosonic creation and annihilation operators for a particle on site n of a square lattice. The model has been analysed theoretically with a several different methods, ranging from the historical Bogoliubov theory [47] to later approaches based on the Gutzwiller ansatz [124].

For different choices of the model parameters U and μ, the system exhibits two different phases in the thermodynamic limit: a superfluid phase (small U) and a Mott insulator phase (large U). One characterization of the superfluid phase is that the U(1) symmetry generated by the particle number operator $\hat{N} = \sum_n \hat{b}_n^\dagger \hat{b}_n$ [125] is spontaneously broken for $N \to \infty$. This leads to both a gapless Goldstone mode and a massive Higgs amplitude mode in the excitation spectrum around the transition. The properties of these have both been described theoretically, *e.g.*, with methods based on the Gutzwiller ansatz [126–128], strong coupling [129, 130], the variational cluster approach [131], the random phase approximation [132] or the ladder diagram approximation for the continuum theory [133], and observed in experimental realizations of the model [134–136].

The aim of this chapter is to describe a systematic study of the superfluid phase based on variational methods using the variational family of bosonic Gaussian states – including both squeezing and displacement (see Section 4.3). This can be seen in some way as a generalisation of the Bogoliubov mean field theory, which can be understood as based on the smaller variational family of bosonic coherent states. Bogoliubov theory describes the model by suitably truncating the Hamiltonian to a quadratic non-interacting mean field Hamiltonian. This mean field Hamiltonian can be diagonalized using Bogoliubov transformations and its spectrum describes the dispersion relation of the gapless Goldstone mode of the model. This last step is in fact equivalent to applying the linearised time dependent variational principle to coherent states (coherent TDVP).

The minimal energy of the Bogoliubov Hamiltonian approximates remarkably well the exact ground state energy and the spectrum captures the gaplessness of the model. Bogoliubov theory, however, also presents several drawbacks. First, the Bogoliubov ground state energy approximation is not variational, *i.e.*, the mean field ground state does not minimise the expectation value with respect to the full Hamiltonian. Second, it does not capture other excitations beyond the Goldstone one, such as the Higgs amplitude mode or bound doublon states. Third, the Goldstone quasiparticles are non-interacting and thus, the decay of quasiparticles excitations can only be studied by re-including the initially discarded Hamiltonian terms as a perturbation [137]. By instead applying linearised TDVP to an extended variational manifold, *i.e.*, the larger class of Gaussian states in place of just coherent states, we overcome all of these drawbacks.

First, in Section 6.2, we compute a variational ground state approximation given by the Gaussian state $|\psi_\mathrm{g}\rangle$ with minimal energy expectation value. We show that $|\psi_\mathrm{g}\rangle$ can be efficiently computed in any dimension from two self-consistent equations.

Second, in Section 6.3, we study the linearisation of the projected real time evolution on the variational manifold to obtain an expression for the system's

excitation spectrum. The approximate spectrum captures both 1- and 2-particle states, which include the gapless Goldstone mode, a doublon mode and a gapped mode which may be interpreted as a Higgs amplitude mode.

Third, in Section 6.4, we apply geometric linear response theory to capture how linear perturbations couple to different parts of the spectrum. As the Gaussian tangent space naturally captures the interaction of quasi-particle excitations in the 1- and 2-particle sector, this allows us to compute the decay and time evolution of excitations.

We conclude in Section 6.5 with a summary of the methods we have discussed. In particular, we point out how the Gaussian variational method is related to and generalises simpler methods such as Bogoliubov theory or variational methods based on coherent states.

6.2 Gaussian ground state approximation

As first step of applying our variational methods, we compute the best Gaussian state $|\psi_{\mathrm{g}}\rangle$, *i.e.*, the normalised Gaussian states whose energy expectation value $E_{|\psi_{\mathrm{g}}\rangle} = \langle\psi_{\mathrm{g}}| \hat{H} |\psi_{\mathrm{g}}\rangle$ on the full Hamiltonian is minimal.

6.2.1 Variational manifold

We generalise the Bogoliubov theory of the Bose-Hubbard model by extending the variational manifold for the system state to the full manifold $\mathcal{M}_{\mathrm{Gaussian}}$ of bosonic Gaussian states. This is in contrast to regular Bogoliubov theory, where the variation is only done with respect to bosonic coherent states. The manifold of Gaussian states can be conveniently defined by first squeezing and then displacing the reference vacuum $|0\rangle$, as discussed in Section 4.3.

In practice, we use a variant of the parametrisations (4.39) and (4.40), namely

$$\mathcal{M}_{\mathrm{Gaussian}} = \left\{ |\beta, \lambda\rangle = \mathcal{U}(\beta, \lambda) |0\rangle \right\}, \tag{6.2}$$

with unitaries $\mathcal{U}(\beta, \lambda) = \mathcal{D}(\beta)\mathcal{S}(\lambda)$ defined by

$$\mathcal{D}(\beta) = \exp\left[\frac{1}{2}\sum_k \left(\beta_k \hat{b}_k^\dagger - \beta_k^* \hat{b}_k\right)\right], \tag{6.3}$$

$$\mathcal{S}(\lambda) = \exp\left[\frac{1}{2}\sum_{kq} \left(\lambda_{k,q} \hat{b}_{k-q}^\dagger \hat{b}_q^\dagger - \lambda_{k,q}^* \hat{b}_{k-q} \hat{b}_q\right)\right], \tag{6.4}$$

where $\hat{b}_k = \frac{1}{\sqrt{N}} \sum_n e^{-\mathrm{i}\, k \cdot x_n} \hat{b}_n$ are the momentum space annihilation operators. Here, β_k is a complex vector and $\lambda_{k,q}$ is a complex matrix invariant under $q \to k-q$.

The indices k and q run in the reciprocal lattice. The only redundancy contained in this parametrization is the symmetry of λ. For a system with N bosonic degrees of freedom, we count $N(N+3)/2$ complex coordinates $(\beta_k, \lambda_{k,q})$ or $N(N+3)$ real coordinates

$$x = \big(\mathrm{Re}(\beta_k), \mathrm{Re}(\lambda_{k,q}), \mathrm{Im}(\beta_k), \mathrm{Im}(\lambda_{k,q})\big)\,. \tag{6.5}$$

We will use the shorthand notation $\mathcal{U}(x_\mathrm{g})$ for the choices of x, such that $\mathcal{U}(x_\mathrm{g})\,|0\rangle = |\psi_\mathrm{g}\rangle$.

The manifold is closed under the action of any subgroup generated by any operators that are linear and quadratic in creation/annihilation operators. In particular, this applies to the U(1) symmetry group generated by the total number operator $\hat{N} = \sum_n \hat{b}_n^\dagger \hat{b}_n$. Here, for any Gaussian state $|\psi\rangle$ other than the vacuum $|0\rangle$ we find a whole ring of inequivalent states $e^{\mathrm{i}\theta\hat{N}}\,|\psi\rangle$ with the same energy expectation value. Therefore, we expect this variational manifold to be well suited to capture spontaneously broken U(1) symmetry phase of the system, *i.e.*, the superfluid phase, and its features, such as the massless Goldstone mode.

While the symmetry of the Bose-Hubbard model is known to be only spontaneously broken in the thermodynamic limit ($N \to \infty$), our ansatz already gives rise to a family of non-symmetric approximate ground states at finite N. We furthermore point out that, while finite temperature spontaneous breaking of a continuous symmetry at zero temperature is ruled out in one dimension, there might still be quasi-long range order, therefore a broken symmetry ansatz can turn out to be a reasonable choice also in one dimension.

The manifold contains a submanifold of states which are translationally invariant, namely the set of states $|\psi(\beta_k, \lambda_{k,q})\rangle$ with $\beta_k = \delta_{k,0}\beta_0$ and $\lambda_{k,q} = \delta_{k,0}\lambda_{0,q}$. For the ground state search, it is sufficient to restrict ourselves to this submanifold as we expect the ground state to preserve the translational symmetry of the problem. For the study of excitations around the translationally invariant ground state, we will then use the full manifold in order to capture also excitations with non-zero momentum.

The tangent space $\mathcal{T}_{|\psi\rangle}\mathcal{M}_\mathrm{Gaussian}$ of the variational manifold at the point $|\psi\rangle$ is spanned by the tangent vectors $|V_\mu\rangle$, which naturally include states with 1- and 2-particle excitations (compare with Section 4.3.4):

$$\mathcal{T}_{|\psi(x)\rangle}\mathcal{M}_\mathrm{Gaussian} = \mathrm{span}_\mathbb{C}\Big\{\,\mathcal{U}(x)\hat{b}_k^\dagger\,|0\rangle\,,\ \mathcal{U}(x)\hat{b}_{k-q}^\dagger\hat{b}_q^\dagger\,|0\rangle\,\Big\}_{k,q}\,. \tag{6.6}$$

Put differently, the variational class of *all* Gaussian states captures accurately the 1- and 2-particle quasiparticle excitation sector of our model.

6.2.2 Ground state search

The first step of our procedure to exploit the given choice of variational manifold is to find within it the best approximation of the ground state, that is the state with the lowest energy expectation value.

Our strategy to do this is simply to look for the stationary points of the energy function. That is to find states $|\psi(x)\rangle$ in the variational manifold such that

$$0 = \frac{\partial}{\partial x^\mu} E(x) = \sum_\nu \boldsymbol{g}_{\mu\nu} \mathbb{P}^\nu_{\psi(x)} \hat{H} \, |\psi(x)\rangle \ , \tag{6.7}$$

where $E(x) \equiv E(\beta, \lambda) = \langle \beta, \lambda | \hat{H} | \beta, \lambda \rangle$ is simply the energy function and we have applied (2.49). Notice that it is not necessary to actually compute $\boldsymbol{g}_{\mu\nu} = 2\mathrm{Re}\, \langle V_\mu | V_\nu \rangle$, since its positive definiteness guarantees that (6.7) is equivalent to

$$\mathbb{P}_{|\psi(x)\rangle} \hat{H} \, |\psi(x)\rangle = 0 \, , \tag{6.8}$$

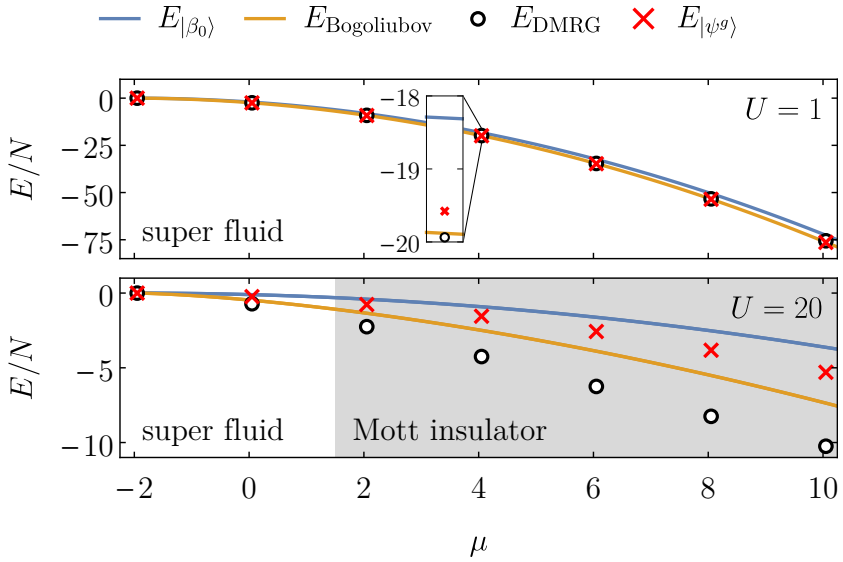
where $\mathbb{P}_{|\psi(x)\rangle}$ is the orthogonal projector onto the tangent space to the manifold at $|\psi(x)\rangle$. For Gaussian states, we find simple equations for this stationary point and see that they only admit a single solution up to the redundancy generated by $e^{\mathrm{i}\alpha \hat{N}}$.

This solution $|\psi_\mathrm{g}\rangle$ can be characterized analytically, independently of the system size or dimensionality, in terms of two parameters A and B, which can be efficiently computed numerically as the fixed point of two coupled self-consistent equations. For more details on this calculation and on how to parametrize the approximate ground state see Appendix A.4.2.

6.2.3 Ground state properties

Having obtained an analytical expression for the approximate ground state, it is then possible to calculate the predictions of our model for ground state properties such as the energy and particle densities. The quality of our method can be benchmarked by comparing these quantities with the ones obtained through other methods, such as Bogoliubov theory or, at least in one dimension, with a numerical DMRG calculation [36] (see Figure 6.1). Our variational energy $E_{|\psi_\mathrm{g}\rangle} = \langle \psi_\mathrm{g} | \, \hat{H} \, | \psi_\mathrm{g} \rangle$ is higher than the DMRG one, as expected, but lower than the one obtained by other variational choices, such as the coherent state $|\beta_0^\mathrm{c}\rangle$ with minimal energy $E_{|\beta_0^\mathrm{c}\rangle}$. The energy obtained as the ground state energy of the Bogoliubov mean field Hamiltonian is generally lower than ours and remarkably close to the DMRG result. However, it is important to emphasize that this energy $E_\mathrm{Bogoliubov}$ is not variational as it is computed with respect to the truncated mean field Hamiltonian, which actually does not admit a well defined ground state in the zero momentum

Figure 6.1: *Comparison of ground state energies in 1D.* We compare the following approaches: (a) Minimal energy on manifold of coherent states $E_{|\beta_0^c\rangle}$, (b) Bogoliubov ground state energy $E_{\text{Bogoliubov}} = E_{|\beta_0^c\rangle} - \Delta^c$ from (6.16), (c) DMRG energy E_{DMRG} and (d) minimal energy $E_{|\psi^g\rangle}$ of all Gaussian states, in the limit $N \to \infty$. The DMRG results were computed for finite systems with open boundary conditions and then extrapolated to the thermodynamic limit $N \to \infty$. The Gaussian state energy $E_{|\psi^g\rangle}$ appears to have saturated to the thermodynamic limit value by $N \approx 500$.



mode. More precisely, the state minimising the mean field energy is infinitely squeezed, which would lead to a diverging energy expectation value with respect to the full Bose-Hubbard Hamiltonian.

Thus, the Gaussian variational family provides a consistent class to approximate the ground state of the Bose-Hubbard model in the superfluid phase. Its ground state energy estimate is reasonable, although worse than the one obtained from Bogoliubov theory. The strength of our extended variational family, however, lies in its prediction of quasiparticle excitations and their properties, such as life time and linear response.

6.3 Quasi particle excitations

We can derive an approximate excitation spectrum from the perspective of our Gaussian variational manifold by using the method of linearised projected real time evolution (see Section 3.3.2). As our variational class generalises the coherent state

manifold used in standard Bogoliubov theory, we will be able to capture higher excitation modes of the model.

The projected real time evolution is computed as prescribed by the time dependent variational principle (TDVP)

$$\frac{d}{dt}\left|\psi(t)\right\rangle = \mathbb{P}_{|\psi(t)\rangle}(-\mathrm{i}\hat{H})\left|\psi(t)\right\rangle \ . \tag{6.9}$$

Note that, as bosonic Gaussian states are a Kähler manifold, this is equivalent to both the Lagrangian and McLachlan TDVP (compare with Section 3.2). The equation of motion (6.9) generates a Hamiltonian time evolution flow $\Phi_t : \mathcal{M}_{\text{Gaussian}} \to \mathcal{M}_{\text{Gaussian}}$ that, linearised around the stationary state $|\psi_{\text{g}}\rangle$, reduces to a sum of phase rotations. From the perspective of our variational manifold, the frequencies of these rotations provide a natural approximation of the lowest excitation energies.

6.3.1 Linearised TDVP

We calculate the excitation energies, shown in Figure 6.2, as the eigenvalues of the linearisation of the equations of motion (6.9), that can be understood as a generalisation of the well-known Gross-Pitaevskij equations [48, 49].

In particular, we consider the linearisation of the projected real time evolution around the stationary point $|\psi_{\text{g}}\rangle$. Given our real variational parameters x from (6.5), the projected Schrödinger equation (6.9) takes the form

$$\sum_{\mu} \dot{x}^{\mu} \frac{\partial}{\partial x^{\mu}}\left|\psi(x)\right\rangle = \mathbb{P}_{|\psi(x)\rangle}(-\mathrm{i}\hat{H})\left|\psi(x)\right\rangle \ . \tag{6.10}$$

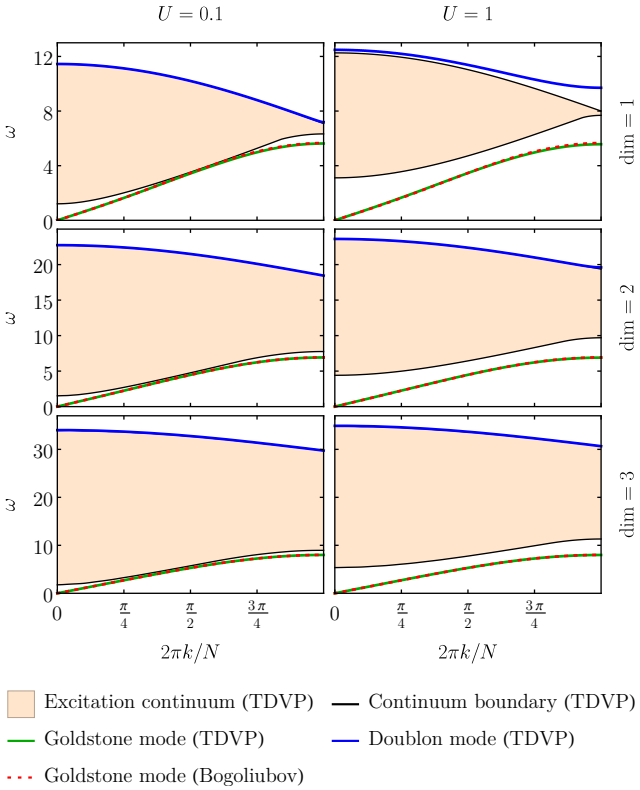
The projector $\mathbb{P}_{|\psi(x)\rangle}$ projects onto the tangent plane $\mathcal{T}_{|\psi(x)\rangle}\mathcal{M}_{\text{Gaussian}}$ spanned by the vectors $\frac{\partial}{\partial x^{\mu}}\left|\psi(x)\right\rangle$. We can therefore in particular introduce the component $\mathbb{P}^{\mu}_{|\psi(x)\rangle}$ referring to the projection onto the specific direction corresponding to the coordinate x^{μ}. This leads to the coordinate time evolution equation

$$\dot{x}^{\mu} = \mathcal{X}^{\mu}(x) = \mathbb{P}^{\mu}_{|\psi(x)\rangle}(-\mathrm{i}\hat{H})\left|\psi(x)\right\rangle \ , \tag{6.11}$$

which can be compared to (3.14). We then linearise the equations of motion around the stationary point x_{g}, *i.e.*, where $\mathcal{X}^{\mu}(x_{\text{g}}) = 0$. The linearisation is based on taking $x^{\mu} = x^{\mu}_{\text{g}} + \delta x^{\mu}$ and expanding (6.11) to first order in δx^{μ} leading to

$$\delta\dot{x}^{\mu} = \sum_{\nu} \frac{\partial \mathcal{X}^{\mu}(x_{\text{g}})}{\partial x^{\nu}}\,\delta x^{\nu} = \sum_{\nu} \boldsymbol{K}^{\mu}{}_{\nu}\,\delta x^{\nu} \ , \tag{6.12}$$

Figure 6.2: *Excitation spectra for* $\mu = 0$. We compare the quasiparticle excitation spectrum computed from Gaussian TDVP with Bogoliubov theory. The results are shown for $\mu = 0$ and two different values of the interaction strength ($U = 0.1$ and $U = 1$) in 1, 2 and 3 dimensions. The spectrum was computed as eigenvalues of $\boldsymbol{K}^{\mu}{}_{\nu}$ from (6.12), where we interpolated the continuum part of the spectrum. The computations were performed for $N = (501, 101^2, 41^3)$ for dim $= (1, 2, 3)$ respectively.



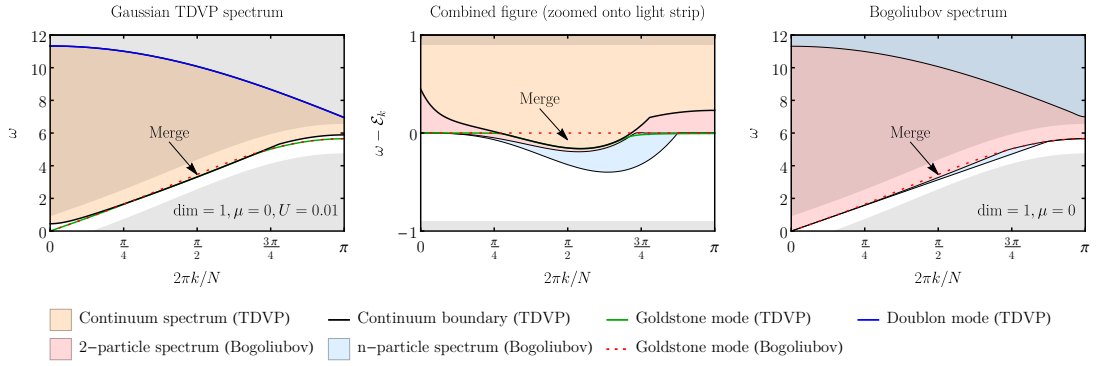
where we introduced the $N(N+3) \times N(N+3)$ linearisation matrix $\boldsymbol{K}^{\mu}{}_{\nu}$. The approximate excitation spectrum is then found as the eigenvalues of $\boldsymbol{K}^{\mu}{}_{\nu}$. We construct $\boldsymbol{K}^{\mu}{}_{\nu}$ explicitly in Appendix A.4.2.

Another formal expression for the matrix $\boldsymbol{K}$ can be shown to be

$$\boldsymbol{K}^{\mu}{}_{\nu} = -\sum_{\rho} \Omega^{\mu\rho} \frac{\partial}{\partial x^{\rho}} \frac{\partial}{\partial x^{\nu}} E(x), \tag{6.13}$$

where $E(x)$ is the energy expectation value of the state $|\psi(x)\rangle$, and the matrix Ω

Figure 6.3: *Comparison with Bogoliubov spectrum.* We compare the spectra from Gaussian TDVP (left figure) with the one from Bogoliubov theory (right figure). For this, we overlap both figures (middle figure) and zoom into the narrow light strip around the Bogoliubov dispersion relation $\mathcal{E}_k$ (red dotted line). We see that the Goldstone mode merges into the TDVP continuum spectrum in the same region, where free Bogoliubov theory predicts $\mathcal{E}_k$ to lie inside the 2-particle continuum (indicated by arrows). The TDVP was performed in 1 dimension for $N = 501$, $\mu = 0$ and $U = 0.01$.



is the antisymmetric symplectic form defined as

$$\Omega = \begin{pmatrix} 0 & \mathbb{1}_N \\ -\mathbb{1}_N & 0 \end{pmatrix} . \tag{6.14}$$

The evaluation of the matrix $\boldsymbol{K}^{\mu}{}_{\nu} = \partial_\nu \mathcal{X}^\mu(x_\mathrm{g})$ reduces to calculating expectation values using Wick's theorem and taking derivatives, therefore it can be calculated analytically in terms of the ground state parameters obtained in the previous section. More details on the form of $\boldsymbol{K}$ can be found in Appendix A.4.2. $\boldsymbol{K}$ is a symplectic matrix whose eigenvalues come in complex conjugate pairs $\pm \mathrm{i}\omega$. The values of ω are our estimates of the excitation energies of the model.

Due to the translational invariance of $\hat{H}$ the matrix $\boldsymbol{K}$ is block diagonal, with each block acting on the span of tangent vectors with fixed total momentum, which we labeled by k in (6.6). The approximate excitation energies ω can therefore also be labeled by the total momentum k their respective eigenvector. The size of each block grows linearly in N, and therefore in the thermodynamic limit $N \to \infty$ there is an infinity of eigenvalues ω_k for each k, which can arrange themselves in a continuum plus possibly some discrete excitations that represent bound states.

6.3.2 Excitation spectrum

In Figure 6.2, we show the dispersion relations obtained by diagonalizing the matrix $\boldsymbol{K}$ numerically. For momenta close to zero, we always find a gapless isolated mode that agrees well with the Bogoliubov dispersion relation $\mathcal{E}_k$. However, we also find a continuum of states that have energies above this Goldstone mode and that always shows a gap around $k = 0$. Finally, for certain parameter choices, *e.g.*, for strong interactions, our spectrum also contains another isolated state above the continuum, which can be interpreted as a doublon state.

We point out that the fact that our method gives a gapless mode was to be expected. Indeed, $\hat{N}$ commutes with the Hamiltonian and the vector $\hat{N}\,|\psi\rangle$ is part of the tangent space for all $|\psi\rangle \in \mathcal{M}_{\text{Gaussian}}$, because $\hat{N}$ is quadratic in the bosonic creation and annihilation operators. Therefore there exists a direction in the manifold along which the energy is constant. In this direction, the Hessian $\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}E(x)$ has a vanishing eigenvalue and thus, because of equation (6.13), also $\boldsymbol{K}$ does.

Our method captures the tangent space generated by displacements and squeezing, *i.e.*, it is spanned by 1- and 2-particle excitations. A generic eigenvector $|E_k\rangle$ of $\boldsymbol{K}^\mu{}_\nu$ with momentum k is

$$|E_k\rangle = \mathcal{U}(x^{\text{g}})\left[C\,\hat{b}_k^\dagger + \sum_q C_q \hat{b}_{k+q}^\dagger \hat{b}_{-q}^\dagger\right]|0\rangle\ , \tag{6.15}$$

where $C, C_q \in \mathbb{C}$. We should therefore compare our results with the 1- and 2-particle excitation spectrum obtained from Bogoliubov theory.

Traditional Bogoliubov theory constructs the excitation spectrum from the 1-particle dispersion relation $\mathcal{E}_k$ of the mean field Hamiltonian

$$[\hat{H}]_{|\beta_0^{\text{c}}\rangle} = E_{|\beta_0^{\text{c}}\rangle} - \Delta^{\text{c}} + \sum_k \mathcal{E}_k^{\text{c}}\,(\delta\hat{b}_k^{\text{c}})^\dagger \delta\hat{b}_k^{\text{c}}\,, \tag{6.16}$$

as reviewed in Appendix A.4.1. The dispersion relation $\mathcal{E}_k$ is independent of the interaction strength U and becomes exact in the limit $U \to 0^+$. General eigenstates of $[\hat{H}]_{|\beta_0^{\text{c}}\rangle}$ consist of non-interacting excitations created by $(\delta\hat{b}_k^{\text{c}})^\dagger$. A general 2-particle excitation with momentum k is therefore given by $(\delta\hat{b}_{k+q}^{\text{c}})^\dagger(\delta\hat{b}_{k-q}^{\text{c}})^\dagger\,|\beta_0^{\text{c}}\rangle$ and has energy $\mathcal{E}_{k+q} + \mathcal{E}_{k-q}$.

Because of the gapless nature of the 1-particle Bogoliubov dispersion relation, the continuum of non-interacting 2-particle excitations is never separated in energy from the 1-particle dispersion, as seen in Figure 6.3 (right) and discussed in Appendix A.4.1. The gap between the isolated bound state (Goldstone mode) and the continuum of higher excitations is therefore a new feature of Gaussian TDVP due to the fact that it implements the interaction within the 1- and 2-particle sectors.

While the Goldstone mode continues to be well-described by the Bogoliubov dispersion relation $\mathcal{E}_k$, the spectrum of 2-particle excitations from Gaussian TDVP starts to divert as we increase U. In particular, we see that for sufficiently large U both, the Goldstone mode and the doublon mode are completely separate from the continuum.

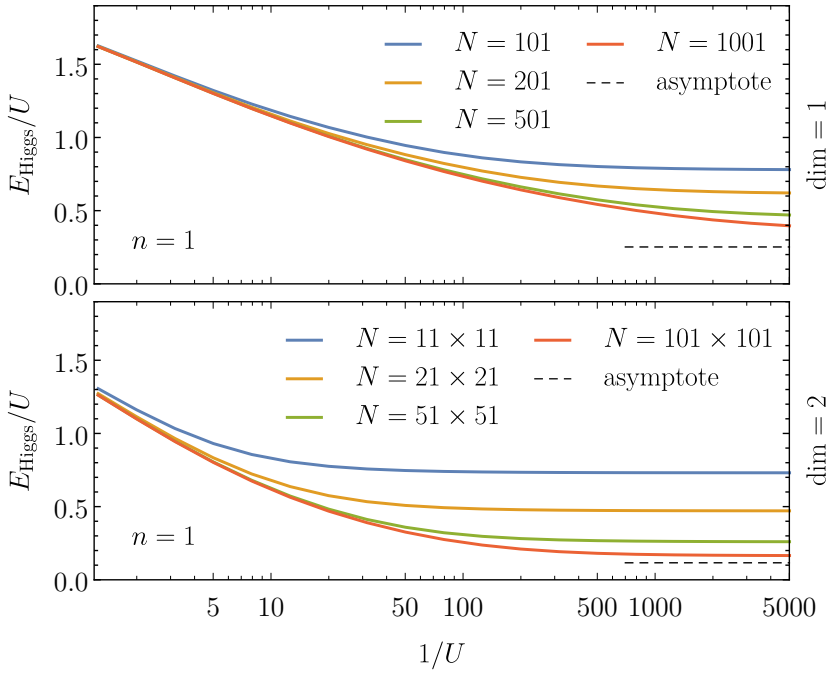
In figure 6.3, we compare Gaussian TDVP and Bogoliubov theory in the regime where the Goldstone mode partially intersects with the continuum. We observe that this intersection appears for small U in the Gaussian TDVP results only in those regimes where also the in Bogoliubov theory the 1-particle mode lies partially above the bottom of the many particle continuum. This phenomenon occurs for choices of μ and system dimension such that the dispersion relation $\mathcal{E}_k$ is not convex, *i.e.*, there exist q, k, such that $\mathcal{E}_k + \mathcal{E}_q < \mathcal{E}_{q+k}$. In Appendix A.4.1, we show that this can only happen for $\mu < 6 - 2\,\mathrm{dim}$. When going to the full Gaussian TDVP, these simple kinematic considerations are no longer sufficient, due to the interaction between 1- and 2-particle excitations. The Gaussian TDVP continuum (light orange) agrees well with the Bogoliubov 2-particle spectrum (light red), where it intersects with the Goldstone mode, *i.e.*, roughly for $2\pi k/N \in (\pi/4, 3\pi/4)$. Outside of this region, the two disagree: While the Gaussian TDVP gives rise to a finite gap between continuum and isolated Goldstone mode, 2-particle continuum and 1-particle dispersion relation necessarily touch for the non-interacting mean field Hamiltonian from Bogoliubov theory.

Gaussian TDVP can describe the decay of single particle excitations into a continuum of higher excitations for those momenta k, where the excitation continuum and the Goldstone mode intersect. The absence of isolated bound states for those momenta k leads to an excitation spectrum composed only of finite width peaks. This in turn means that every excitation that couples to the continuum will completely decay and thus have a finite lifetime. This phenomenon is known as Beliaev damping [137] and is not captured by the standard Bogoliubov theory, but so far has been typically obtained from perturbative expansions by re-including higher order terms of the Hamiltonian. We will further investigate this decay behaviour of excitations into the continuum in Section 6.4.2.

6.3.3 Higgs mode

Another suggestive observation can be made on the physical interpretation of the gapped mode at the bottom of the continuum. A possible interpretation is that it is a remnant of what, near the superfluid to Mott insulator transition, becomes known as the Higgs mode. It corresponds to oscillations of the amplitude of the order parameter $\langle \hat{b}_0 \rangle$ (while the Goldstone mode is interpreted as oscillations of the order parameter phase) and it has been observed experimentally by coupling to it through modulation of the tunneling amplitude [134].

Figure 6.4: *Continuum/Higgs gap as function of* $1/U$. This figure shows the gap between the gapless Goldstone mode and the continuum of excitations as a function of $1/U$ in 1D and 2D and for different system sizes. The asymptotic value for large N and small U obtained in equation (6.17) is also indicated for $N = 1001$ in 1d and $N = 101^2$ in 2d.



The prediction of our model for the continuum gap, which because of this possible interpretation we will label as E_{Higgs}, can be studied numerically through the diagonalization of the matrix $\boldsymbol{K}$ described in the previous paragraphs. At fixed non-zero U, E_{Higgs} converges to a finite non-zero value in the thermodynamic limit.

We are also able to give an analytical asymptotic result for the limit in which $U \to 0$ while μ varies so as to keep a constant particle number density of the ground state $n = \langle \hat{N} \rangle / N$ (see Appendix A.4.2). In this limit, we have that the gap goes to zero linearly in the interaction strength U, namely

$$\lim_{U \to 0} \frac{E_{\text{Higgs}}}{U} = \alpha(N, n) \sim 2\sqrt[3]{2} n^{\frac{2}{3}} N^{-\frac{1}{3}} \quad \text{as} \quad N \to \infty \,. \tag{6.17}$$

Note that it is instrumental that we took here first the limit $U \to 0$, before studying the large N asymptotics.

In Figure 6.4 one can see the numerical results for the behaviour of the ratio between the Higgs gap and U and notice how it indeed approaches a constant

asymptotic value for small U. In the large U region, it has instead an unexpected divergent behaviour (the gap should close at the SF/MI transition [127]), however this can be understood as a breaking down of our model at the transition where Gaussian states are no longer a good description of the system's ground state. It is instead interesting to see how the constant small U behaviour matches the experimentally measured value of the Higgs mode gap [134] even better than the previous theoretical results obtained with Gutzwiller theory.

6.4 Linear Response

We use our variational manifold and the real time evolution projected onto it to study the response of the system to small perturbations. We achieve this by applying the linear response framework discussed in Section 3.4. This is significant, as it provides possible connections to actual experiments, where certain system responses can be probed and measured.

6.4.1 Spectral Functions

We model an external perturbation by considering the time dependent perturbed Hamiltonian

$$\hat{H}_\epsilon(t) = \hat{H} + \epsilon\,\varphi(t)\,\hat{V}\,, \tag{6.18}$$

where $\hat{H}$ is the unperturbed Bose-Hubbard Hamiltonian (6.1), $\varphi(t)$ is a classical external field that couples to the system through the Hermitian operator $\hat{V}$ and ϵ is a small real parameter. We shall then consider the projected real time evolution $|\psi_\epsilon(t)\rangle$ of the system under such perturbed Hamiltonian and evaluate its response in terms of the expectation value of the same coupling operator $\hat{V}$. In particular, we consider this response in the limit of small perturbations, *i.e.*, we compute quantities only up to first order in the parameter ϵ. Thus, we consider the response

$$\delta V(t) = \frac{d}{d\epsilon}\,\langle\psi_\epsilon(t)|\hat{V}|\psi_\epsilon(t)\rangle_{\epsilon=0} \tag{6.19}$$

to the perturbation $\hat{V}$.

As discussed in Proposition 10, the Fourier space response $\delta V(\omega)$, calculated on the variational manifold as explained above, takes the form $\delta V(\omega) = \tilde{\varphi}(\omega)\,\chi(\omega)$, where $\tilde{\varphi}$ is the Fourier transform of the perturbing field $\varphi(t)$ and $\mathcal{A}_V(\omega) \equiv -\frac{1}{\pi}\mathrm{Im}\chi(\omega)$ is the response function of the system with respect to the perturbation $\hat{V}$. Such response functions are expressed in terms of the spectral decomposition of the

Figure 6.5: *Spectral functions for the density variation* $\hat{V}^{(k)}_{\text{density}}$ *from* (6.22). We show, as colour plots, the values of the spectral response function $\mathcal{A}_k(\omega)$ in the relevant range of values of k and ω. In the first column, we show more logarithmic graphs of $\mathcal{A}_k(\omega)$ for fixed slice of k (indicated by vertical lines of the respective color in the second column). The computations were performed for $N = (501, 101^2, 41^3)$ for dim $= (1, 2, 3)$ respectively. To extract a continuous response functions, we performed a binning in energy intervals of $\Delta\omega = (0.13, 0.2, 0.37)$ for dim $= (1, 2, 3)$ respectively.

(a) Density variation

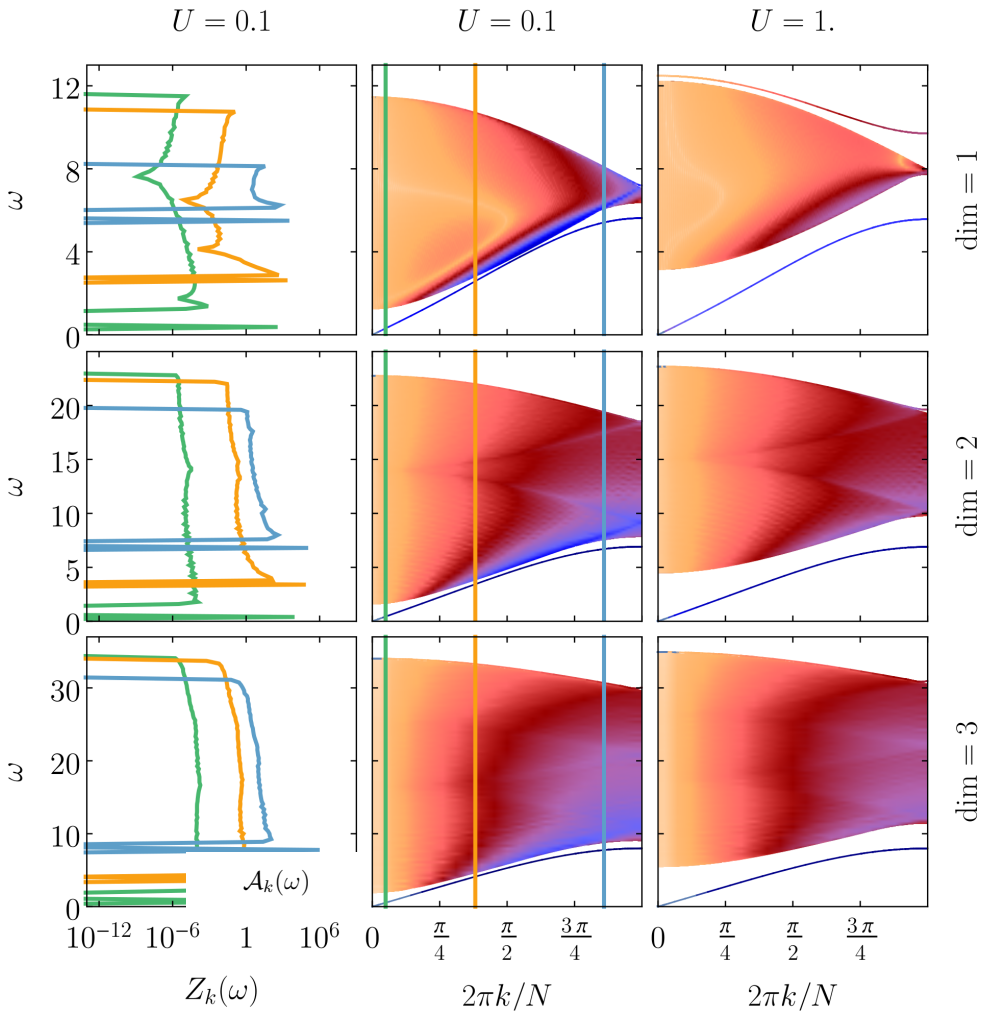
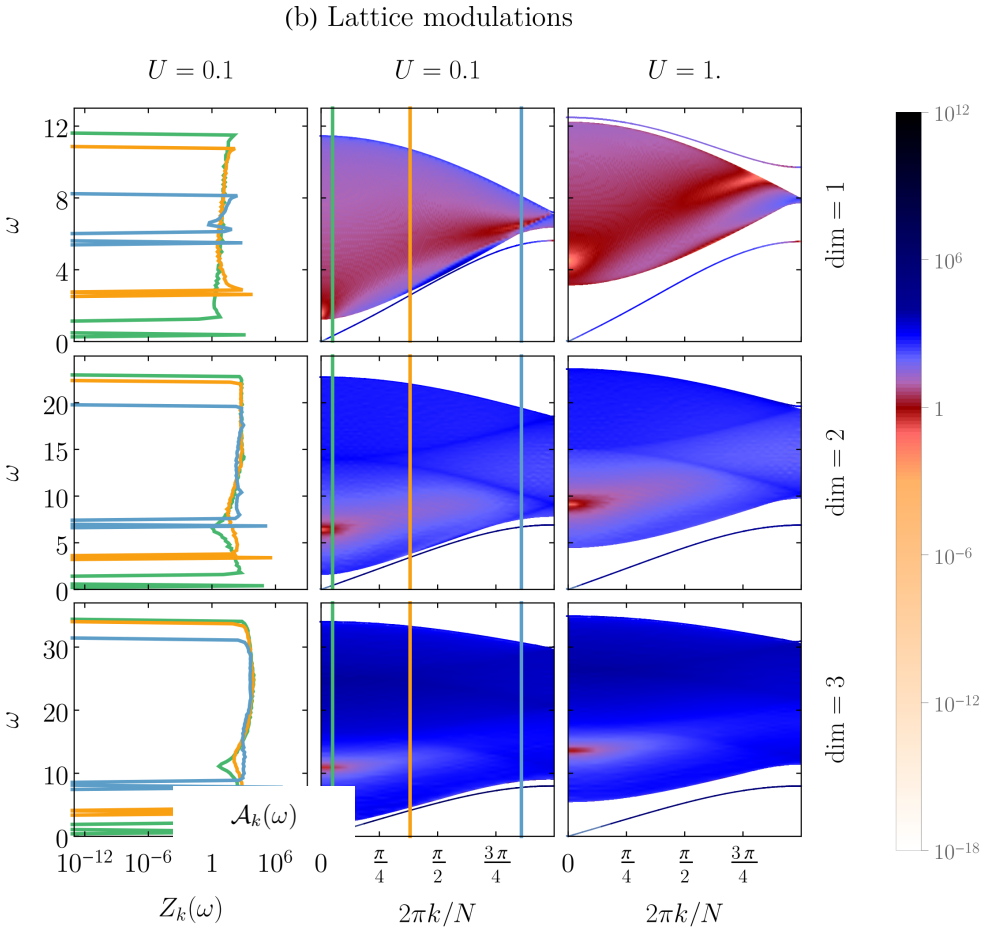


Figure 6.6: *Spectral functions for the lattice modulation* $\hat{V}^{(k)}_{\text{lattice}}$ *from* (6.23). We show, as colour plots, the values of the spectral response function $\mathcal{A}_k(\omega)$ in the relevant range of values of k and ω. In the first column, we show more logarithmic graphs of $\mathcal{A}_k(\omega)$ for fixed slice of k (indicated by vertical lines of the respective color in the second column). The computations were performed for $N = (501, 101^2, 41^3)$ for dim $= (1, 2, 3)$ respectively. To extract a continuous response functions, we performed a binning in energy intervals of $\Delta\omega = (0.13, 0.2, 0.37)$ for dim $= (1, 2, 3)$ respectively.



linearised real time evolution $\boldsymbol{K}$ defined in Section 6.3 as

$$\mathcal{A}_V(\omega) = \frac{1}{2}\,\mathrm{sign}(\omega)\,|e^{\mu}(\omega)\,dV_{\mu}|^2\,\delta(|\omega|)\,, \tag{6.20}$$

where dV is the gradient differential form of the real valued function on the manifold $\langle\psi(\beta,\lambda)|\hat{V}|\psi(\beta,\lambda)\rangle$, $e(\omega)$ are the eigenvectors of $\boldsymbol{K}$ (as defined in Appendix A.4.2) with eigenvalue ω and $\delta(|\omega|)$ is a normalisation of the eigenvectors such that the conditions of Proposition 9 are satisfied.

For the Bose-Hubbard model, we consider the following types of perturbations:

$$\hat{V}^{(k)}_{1-\mathrm{particle}} = \sum_k \mathcal{U}(x^{\mathrm{g}})(\mathrm{i}\hat{b}^{\dagger}_k - \mathrm{i}\hat{b}_k)\,\mathcal{U}^{\dagger}(x^{\mathrm{g}})\,. \tag{6.21}$$

$$\hat{V}^{(k)}_{\mathrm{density}} = \sum_n \hat{b}^{\dagger}_n\hat{b}_n\,\cos(kx_n) \tag{6.22}$$

$$\hat{V}^{(k)}_{\mathrm{lattice}} = \sum_{\langle n,m\rangle} (\hat{b}^{\dagger}_n\hat{b}_m + \hat{b}^{\dagger}_m\hat{b}_n)\,\cos(kx_n)\,. \tag{6.23}$$

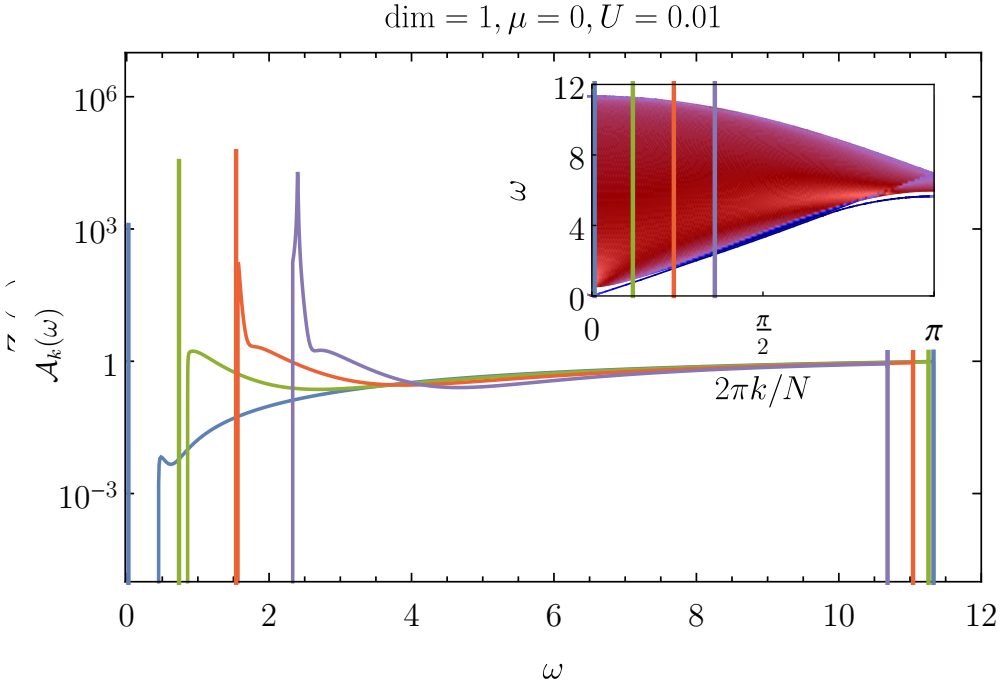
In (6.21), we use a linear operator to create a single particle perturbation of momentum k. The other two perturbations are quadratic in creation and annihilation operators, such that the excitation consists in general of both single and 2-particle excitations. In (6.22), we consider a spatial *density variation* by modulating the chemical potential with momentum k, which couples directly to the local particle density. In (6.23), we consider a spatial modulation of momentum k of the hopping constant. This can be achieved through a modulation of the lattice depth [138]. Such perturbation naturally couples to the kinetic energy operator.

The different response functions $\mathcal{A}_k(\omega)$, obtained by evaluating (6.20) for different types of perturbation operators of momentum k and at energy ω, give us an indication of how strongly each type of perturbation couples to different regions of the spectrum. Some of the resulting response functions are plotted in Figure 6.5 and 6.6.

A first observation we can make is on the behaviour of the isolated Goldstone mode in those situations when it merges with the continuum part of the spectrum. In Figure 6.7, we see how the isolated peak of the response function broadens into a wider feature inside the continuum. This indicates how, even when the Goldstone mode is not an isolated eigenstate, it still survives as a finite lifetime excitation of the system.

We can then also compare how the different perturbations considered couple to the system. In Figure 6.8, we see how the perturbation that couples the strongest to the continuum modes at $k = 0$ is the lattice modulation operator. Although it has to be mentioned that the definition of the normalisations of the perturbations (6.22) and (6.23) is not free of some arbitrariness, the large difference in

Figure 6.7: *Peak merging into continuum.* We plot the spectral response function $\mathcal{A}_k(\omega)$ relative to a density perturbation for a set of different momenta k. The function is calculated for a set of parameters (dim $= 1$, $U = 0.01$ and $\mu = 0$) such that for some values of k the Goldstone mode merges into the continuum spectrum. The plot shows how the delta-like peak of the Goldstone mode transforms into a finite width feature when this merge occurs (purple line).

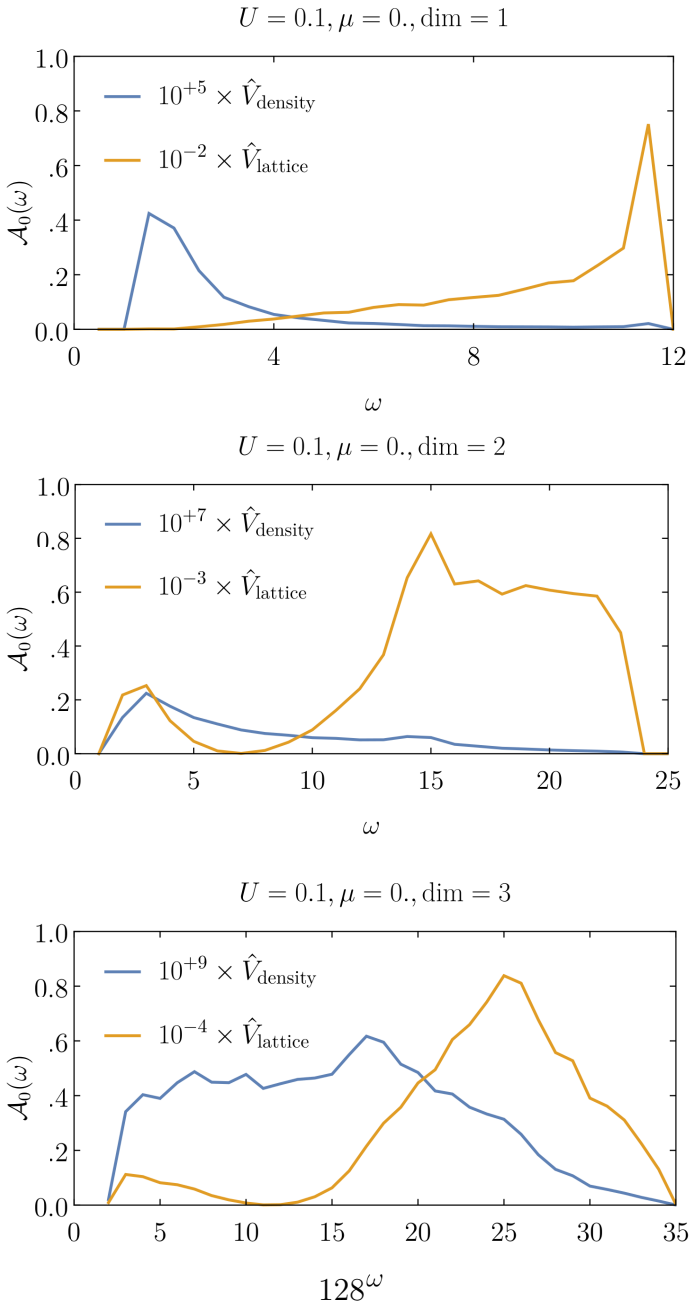


these coupling strengths provides a further element of support to the identification of the lower continuum modes as the Higgs excitation. Indeed it is known that the Higgs mode should be excited most easily through perturbations of the kinetic energy term of the Hamiltonian, while the Goldstone mode through perturbations in the particle density [127].

6.4.2 Real time evolution

The analysis of the response function can also give indications on the real time evolution of perturbations of the system. Indeed, we can interpret the operator $\hat{V}$ as creating a perturbation described by the tangent vector $|\delta\psi_{\mathrm{g}}(0)\rangle = \mathbb{P}_{|\psi_{\mathrm{g}}\rangle}(-\mathrm{i}\hat{V})\,|\psi_{\mathrm{g}}\rangle$ at $t = 0$, which is equivalent to giving the system a kick by choosing $\varphi(t) = \delta(t)$. The evolution in the tangent plane of this perturbation vector is then given by $d\Phi_t$ at $|\psi_{\mathrm{g}}\rangle$, *i.e.*, the push-forward of the real time evolution flow Φ_t around the stationary point. $d\Phi_t$ is a linear map on the tangent space at $|\psi_{\mathrm{g}}\rangle$, explicitly it is

Figure 6.8: *Comparison: Density variation* (6.22) *and lattice modulation* (6.23) *at* $k = 0$. We compare the response functions from density variation with the one from lattice modulations. The response to the lattice modulation is by several orders of magnitude stronger. Note that we rescaled the data by factors of $10^{\pm x}$ to fit into the same range. The computations were performed for $N = (501, 101^2, 41^3)$ for dim $= (1, 2, 3)$ respectively. To extract a continuous response functions, we performed a binning in energy intervals of $\Delta\omega = (0.5, 1, 1)$ for dim $= (1, 2, 3)$ respectively.



given by the matrix $e^{\boldsymbol{K}t}$ with respect to the basis (6.6).

We consider an initial perturbation $|\delta\psi_{\mathrm{g}}(0)\rangle$ in the 1-particle sector of the tangent plane, *i.e.*, we require $|\delta\psi_{\mathrm{g}}(0)\rangle$ to be only spanned by 1-particle states in basis (6.6). In particular, this is accomplished by the perturbation created by $\hat{V}^{(k)}_{1-\mathrm{particle}}$. The time evolution of $|\delta\psi_{\mathrm{g}}(0)\rangle$ under the map $d\Phi_t$ will either describe the full decay with finite life time or a partial decay leading to a remaining excitation with infinite lifetime. This was already mentioned in Section 6.3.2, where we argued that full decay only occurs if there is no bound state, *i.e.*, for those momenta where the Goldstone mode intersects with the 2-particle continuum.

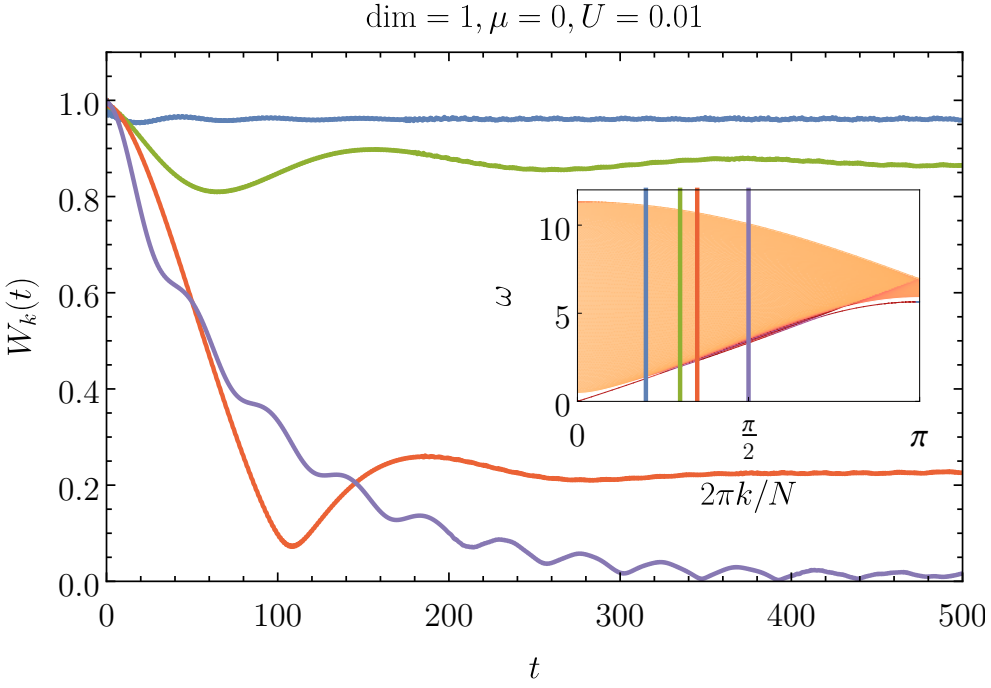
Indeed we observe that if the perturbation is created at a momentum value where there exists an isolated Goldstone state with a strong coupling to the 1-particle sector (quantified by the 1-particle response function), the perturbation will persist indefinitely. If, instead, the excitation has a momentum at which the 1-particle perturbation couples sufficiently strongly to the continuum, a part of it will decay into the continuum modes, disappearing in a time proportional to the inverse of the width of the response function. Finally, if there is no isolated Goldstone state at the chosen momentum, but only the continuum, the perturbation will have a finite lifetime and decay completely into continuum excitations.

In Figure 6.9, we show the overlap of the time evolved perturbation with the 1-particle sector of the tangent plane, for different total momenta of the initial perturbation. A perturbation with momentum k corresponding to an isolated Goldstone state will maintain a large overlap with the 1-particle sector. For perturbations with momentum k closer to the region where the Goldstone mode merges with the continuum, a larger part of the overlap with the 1-particle sector will decay in time. Finally, if the perturbation has a total momentum k, for which no isolated Goldstone state exists in the spectrum, the single particle overlap will decay completely to zero after a finite lifetime. Such decay behavior is similar to what can also be seen in quantum optical systems coupled to unconventional photon baths [139].

This behaviour of the evolution of perturbations can be interpreted as a remnant in lattice systems of what in continuum Bose-Einstein condensates (BEC) is known as the Beliaev damping of excitations, *i.e.*, the decay of 1-particle excitations into the continuum of many particle excitations due to scattering interactions. Our variational scheme successfully captures at least part of this behaviour, namely the one associated to the 1- and 2-particle sector that are fully included in our tangent space. This is in contrast to the traditional Bogoliubov theory that is restricted to the non-interacting 1-particle sector. In particular, standard Bogoliubov theory cannot describe the interaction with the continuum consistently, which can only be incorporated by re-including the previously neglected terms as perturbations [137].

Let us emphasize that our method does not allow us to infer the properties of excitations three or more particles. In particular, it is possible that by including three and more particle excitations, the gap between 1-particle Goldstone mode and 2-particle continuum may become filled. If this were the case, the respective 1-particle excitations would have a finite lifetime as they now could decay into the continuum of higher excitations.

Figure 6.9: *Time evolution of 1-particle weight.* We show, for different momenta k, the real time evolution of the 1-particle weight $W_k(t) = |\langle\psi_k(0)|\psi_k(t)\rangle|^2$, where $|\psi_k\rangle = \hat{V}^{(k)}_{1\text{ particle}}|\psi_g\rangle$ is a perturbation vector in the one particle sector of the tangent plane.



6.5 Comparison between methods

Our study is based on the time dependent variational principle, where we project the equations of motion on a given variational class and linearise them around the stationary state that provides the best approximation of the ground state. While we focused on the class of all Gaussian states, the method can be applied to any suitable family of states, so it is natural to compare the results between different variational classes. In the context of Bogoliubov theory, it is natural to compare

our larger manifold of *all* bosonic Gaussian states $\mathcal{D}(\beta)\mathcal{S}(\lambda)\,|0\rangle$ with the smaller sub manifold consisting only of coherent (or displaced) states $\mathcal{D}(\beta)\,|0\rangle$.

(i) Coherent TDVP around $|\beta_0^{\mathrm{c}}\rangle$. If we apply linearised TDVP to the manifold of coherent states, we obtain the same excitation spectrum as the single-particle spectrum $\mathcal{E}_k$ of Bogoliubov mean field theory (see Appendix A.4.1). However, we should point out that $[\hat{H}]_{|\beta_0^{\mathrm{c}}\rangle}$ contains more information than the linearised TDVP, as it gives us a Hamiltonian operator whose minimal energy $E_{\mathrm{Bogoliubov}}$ is a better estimate of the system's ground state energy than just $E_{|\beta_0^{\mathrm{c}}\rangle}$. On the other hand, this energy is not variational, *i.e.*, it cannot be expressed as the expectation value of an ansatz state on the full system Hamiltonian. Furthermore, the truncation of $[\hat{H}]_{|\beta_0^{\mathrm{c}}\rangle}$ is not self-consistent, because $|\beta_0^{\mathrm{c}}\rangle$ is not its ground state.

(ii) Coherent TDVP around $|\psi_{\mathrm{g}}\rangle$. After finding the best Gaussian ground state approximation $|\psi_{\mathrm{g}}\rangle$, we can linearise around it the equations of motion restricted to the space of displaced states $\mathcal{D}(\beta)\,|\psi_{\mathrm{g}}\rangle$. The 1-particle spectrum $\mathcal{E}_k^{\mathrm{g}}$ obtained this way is gapped and consequently not a good approximation to the Goldstone mode. However, we can use $\mathcal{E}_k^{\mathrm{g}}$ to construct the 2-particle continuum of the quadratic Hamiltonian $[\hat{H}]_{|\psi_{\mathrm{g}}\rangle}$. Interestingly, the resulting 2-particle spectrum provides a good approximation to the continuum with Gaussian TDVP. In this way, we can understand the gap $\mathcal{E}_0^{\mathrm{g}}$ as already encoding the interaction energies between two particle excitations that is required to approximate the interacting 2-particle spectrum.

(iii) Gaussian TDVP around $|\beta_{\mathrm{g}}\rangle$. In order to obtain a self-consistent ground state, we enlarge the manifold of states and introduce general Gaussian states, which also allow for squeezing. Indeed, the Gaussian state of minimal energy $|\psi_{\mathrm{g}}\rangle$ can also be identified as the state that fulfills the property of being the ground state of the corresponding mean field Hamiltonian $[\hat{H}]_{|\psi_{\mathrm{g}}\rangle}$, *i.e.*, the quadratic truncation of the full Hamiltonian when normal-ordered with respect to $|\psi_{\mathrm{g}}\rangle$. If we apply linearised TDVP to the extended manifold of Gaussian states, we obtain the spectrum which was the main object of this chapter, which naturally contains both 1- and 2-particle excitations (see Section 6.3).

Projected Hamiltonian. Finally, there is a well-known alternative based on the projected Hamiltonian (see Section 3.3) to compute excitation spectra from a tangent space. Instead of linearising the equations of motion, we can directly take the tangent space as variational ansatz for eigenstates by projecting the full Hamiltonian onto it, *i.e.*, $\boldsymbol{H} = \mathbb{P}_{|\psi\rangle}\hat{H}\mathbb{P}_{|\psi\rangle}$, and then computing its spectrum. The eigenstates $|E_i\rangle$ with energy E_i of the projected Hamiltonian $\boldsymbol{H}$ are manifestly variational, *i.e.*, their expectation value with respect to the full Hamiltonian is equal to E_i and there exists a true eigenstate of the full Hamiltonian with smaller energy. This is not necessarily the case for the eigenvectors of $\boldsymbol{K}$ in the linearised TDVP. In [92], it has been further pointed out that – in contrast to the projected

Hamiltonian method – the linearised TDVP may incorrectly predict massless excitation modes. This occurs whenever the approximate ground state within the chosen variational family spontaneously breaks a symmetry which is not spontaneously broken in the exact ground state. In the case of the Bose-Hubbard model, this is actually a desirable feature: while the true ground state only breaks the U(1) symmetry in the limit $N \to \infty$, the family of Gaussian states already breaks this symmetry for finite N and is thus well-suited to study the superfluid phase in the thermodynamic limit.

In conclusion, our Gaussian TDVP method naturally generalises Bogoliubov theory to describe the superfluid phase of the Bose-Hubbard model. The presented methods provide systematic framework to compute (a) approximate ground state energies, (b) excitation spectra and (c) linear response functions for general variational families.

The one of our predictions that calls most for further inquiry is the gapped 2-particle continuum above the Goldstone mode. It will be interesting to further explore with other methods whether the identification of the lowest continuum mode as the Higgs mode is correct and whether the gap that separates it from the Goldstone mode survives, once one also considers excitations of three or more particles.

The presented scheme is self-consistent and requires no other assumptions than the choice of variational manifold. It can be easily applied to other variational families, such as non-Gaussian states [52] or Gutzwiller states. In particular, it would be interesting to find a simple variational family with which it is possible to perform a similar study also in the Mott phase of the Bose-Hubbard model. Even more interesting would be a single variational set which can capture both the Mott and superfluid phases.

Chapter 7

Application: the Sherrington-Kirkpatrick model

In this chapter we will present another application of the methods discussed in this thesis to a physical model. We will construct an explicit example of the generalised group-theoretic coherent states introduced in Chapter 5 and show how they can be used to successfully describe the ground state of the Quantum Sherrington-Kirkpatrick model, a paradigmatic model for quantum spin glasses.

Most of the material of this chapter was published in reference [140]:

P.M. Schindler*, T. Guaita*, T. Shi, E. Demler, J.I. Cirac,
A Variational Ansatz for the Ground State of the Quantum Sherrington-Kirkpatrick Model, arXiv:2204.02923 (2022)
* Co-first authors

7.1 Overview

Spin glasses are an important paradigm in statistical physics. Besides their relevance in describing disordered classical magnets [141, 142], it was shown that optimisation tasks, such as the travelling salesman problem, can be mapped to solving for the ground states of spin glass systems [141, 143, 144]. Classical spin glasses can be promoted to quantum models by introducing a transverse field. The resulting quantum spin glasses form by themselves an important playground to study the interplay of disorder and frustration with quantum effects [145]. Moreover, there is evidence that the quantumness can be exploited to shortcut optimisation tasks, for instance through quantum annealing [146–150].

The textbook example of a quantum spin glass model is the Quantum Sherrington-Kirkpatrick (QSK) model, a generalisation of the classical Sherrington-Kirkpatrick (SK) model [151, 152]. The QSK model has been studied extensively in the lit-

erature both analytically [152–156] and numerically [157–167]. While the famous Parisi solution [168, 169] provides a full solution to the classical SK model, many open questions remain for the quantum SK model. Since the QSK model is an all-to-all coupled model one might assume that a mean-field product state *ansatz* well describes the ground state. However, this *ansatz* predicts a quantum phase transition (QPT) from a quantum spin glass phase to a paramagnetic phase at a critical transverse field $g_{\mathrm{C}} \approx 2\,J$ [163]. Field theory approaches [155, 156] using the replica method suggest instead a phase transition at $g_{\mathrm{C}} \approx 1.5\,J$. Numerical calculations at small system sizes [161, 162] or obtained at finite temperature [157, 158, 162, 166] confirm the latter. So far no good *ansätze* have been found which can describe the zero temperature regime for large system sizes, preventing the study of further properties of the ground state, such as entanglement.

To tackle this problem, in this chapter we propose a variational family which extends the product state *ansatz*, introducing a richer entanglement structure. This *ansatz* is constructed according to the generalised group-theoretic coherent states formalism defined in Chapter 5. Indeed, we are dealing with a spin system, where product states represent the notion of group-theoretic coherent states. Our generalised construction therefore provides a natural way to go beyond that. As previously discussed, the special structure of the generalised states allows us to introduce non-trivial quantum correlations while at the same time being able to efficiently find variational ground states for large systems, in this case of up to $N = 200$ spins.

In Section 7.2, we will introduce the QSK model in detail and present the generalised coherent state *ansatz* that we will use to study it. Thanks to these variational ground states, we obtain a significantly improved description of the model compared to plain product states. Indeed, we observe an extensive improvement in terms of ground state energies with respect to product states. More importantly, generalised coherent states predict the quantitatively correct critical transverse field of the spin glass phase transition, while product states fail to do so.

In Section 7.3, we further exploit our ground state *ansatz* to study some previously unexplored properties of the model. In particular, we develop a method to study the entanglement structure of the ground states. We numerically compute the entanglement entropy of subsystems of varying sizes with the rest of system and compare the results with some related models. Our results show a volume law of entanglement, which indicates that entanglement monogamy does not provide a scaling constraint despite the fact that the QSK model involves all-to-all spin interactions. Finally, we discuss how the observed entanglement structure admits an even simpler description. Indeed, a qualitatively analogous entanglement structure can be observed in an ensemble of states that have been introduced in the

Quantum Information Theory context, namely random weighted graph states [170] with normally distributed phase gates.

In Section 7.4, we address a further unresolved aspect of the model, namely the existence of a spin glass phase transition at non-vanishing longitudinal fields. It has been conjectured that such transitions exist, but the conjecture is debated and decisive evidence has yet to be collected in its favour. The power of our *ansatz* helps to provide a contribution to answering this question. Indeed, we show that, according to our analysis of both the order parameter and of the entropy data, there is no evidence of a phase transition at finite longitudinal fields.

7.2 Model and variational *ansatz*

Concretely, we consider a system of N spin-$\frac{1}{2}$ degrees of freedom. The QSK model corresponds to a mixed field Ising model with all-to-all couplings and quenched disorder in the couplings and longitudinal field,

$$\hat{H}_{\mathrm{QSK}} = -\frac{1}{2}\sum_{n,\,m=1}^{N} J_{nm}\hat{\sigma}_z^n\hat{\sigma}_z^m - g\sum_{n=1}^{N}\hat{\sigma}_x^n - \sum_{n=1}^{N} h_n\hat{\sigma}_z^n\,, \tag{7.1}$$

where $\hat{\sigma}_i^n$ is the i-th Pauli-matrix acting on the n-th spin. The longitudinal field h_n and the couplings J_{nm} are independently normally distributed numbers with zero mean and variance $\overline{h_n^2} = h^2$ and $\overline{J_{nm}^2} = J/N$, respectively. Here and in the following we use the convention that an overbar $\overline{\cdot}$ indicates disorder average.

For such a system, the natural group-theoretic coherent states are atomic coherent states (CS) [64], which we already introduced in Example 10. They can be written as

$$|\phi(x)\rangle = \mathcal{U}(x)\,|\uparrow,\,\ldots,\,\uparrow\rangle\,, \tag{7.2}$$

where $\hat{\sigma}_z\,|\uparrow\rangle = +\,|\uparrow\rangle$ and $\mathcal{U}(x) = \exp(-i\sum_{n,i} x_n^i\hat{\sigma}_i^n)$ rotates each of the N spins individually on the Bloch sphere. This construction makes explicit use of the $\mathrm{SU}(2)^N$ group representation that exists on the Hilbert space. The simple CS *ansatz* is parametrized by the real parameters $x = \{x_n^i\}_{i=x,y,z;\,n=1,\ldots,N}$ and corresponds to the set of normalised product states.

To these states we can apply the generalisation procedure defined in Chapter 5, which leads to the set of generalised atomic coherent states (GCS),

$$|\Psi(x,\,y,\,M)\rangle = \mathcal{U}(y)\mathcal{V}(M)\,|\phi(x)\rangle\,, \tag{7.3}$$

where x_n^k, y_n^k and M_{nm} $(n < m)$ are the variational parameters. $\mathcal{U}$ and $|\phi\rangle$ are defined as in equation (7.2) and the entangling unitary $\mathcal{V}(M)$ is given by

$$\mathcal{V}(M) = \exp(-\frac{i}{4}\sum_{n<m} M_{nm}\hat{\sigma}_z^n\hat{\sigma}_z^m)\,, \tag{7.4}$$

for any real symmetric matrix M. The states (7.3) have previously been studied in References [29, 171] and can be seen as a one-layer QAOA-*ansatz* [9, 172].

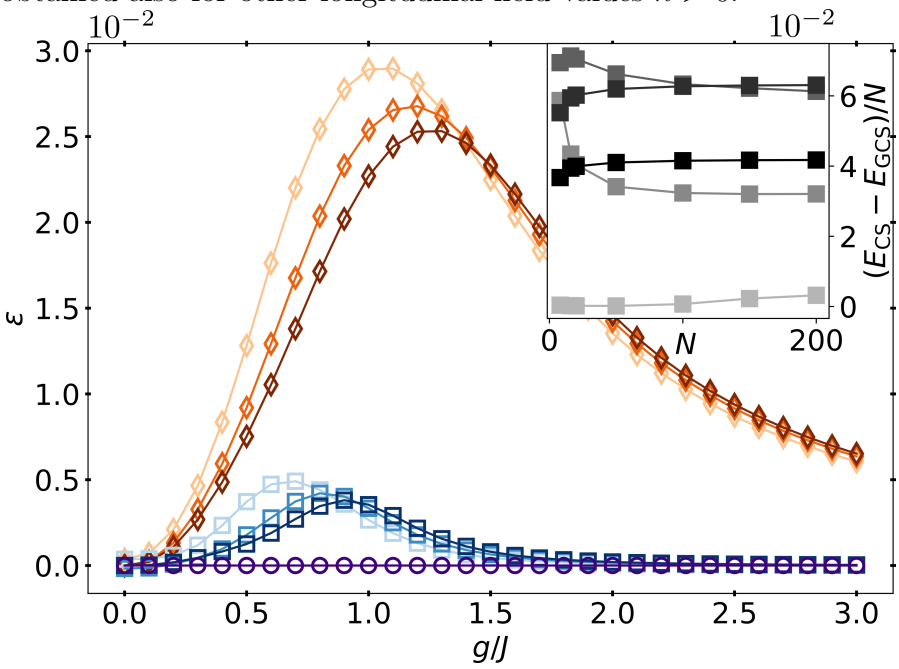
The entangling unitaries $\mathcal{V}(M)$ contain two-spin terms which give the states (7.3) a non-trivial correlation structure. However, as extensively discussed in Chapter 5, this structure is chosen so that it is still possible to efficiently compute the expectation values on $|\psi(x, y, M)\rangle$ of any polynomial of Pauli operators. In the present case, when computing expectation values of Pauli operators we have $\mathcal{V}(M)^\dagger \hat{\sigma}_\pm^n \mathcal{V}(M) = \hat{\sigma}_\pm^n \exp(\pm i/4 \sum_m M_{nm} \hat{\sigma}_z^m)$, that is the two-spin terms cancel and we are left just with products of single spin operators [101, 108]. This crucial property allows us to find analytical expressions for the energy and the gradient of the energy with respect to the variational parameters [173](see Appendix A.5.1 for more details). Thanks to this, we can apply the imaginary time evolution algorithm (see Section 3.5) to efficiently obtain the variational ground states of individual disorder realisations for large system sizes of up to $N = 200$ spins.

In order to demonstrate the expressivity of the GCS *ansatz*, we first consider the approximate ground state energy. For small system sizes we can compare the variational energies with numerically exact results, obtained via a Lanczos Exact Diagonalization method (ED) [174], see Figure 7.1. We find good quantitative agreement of the variational ground state energy with the exact ground state energy over a broad range of transverse and longitudinal field values. In particular a notable improvement of the GCS *ansatz* upon the CS *ansatz* becomes visible. For larger systems it is no longer possible to compare to an exact solution. However, we observe an extensive improvement in energy upon the CS *ansatz*, suggesting that the GCS *ansatz* gives a non-vanishing improvement even in the thermodynamic limit, see inset of Figure 7.1.

Quantum phase transition

Our variational *ansatz* also allows us to study the QPT on the $h = 0$ line of the model's parameter space. In Figure 7.2 we depict the spin glass susceptibility $\chi_{\text{sg}} = N^{-1} \sum_{n,m} \overline{\langle \hat{\sigma}_z^n \hat{\sigma}_z^m \rangle^2}$, which is independent of the system size in the paramagnetic phase (large g) and scales with the system size in the ordered spin glass phase (small g) [141, 161, 164, 175, 176]. For small system sizes we find good quantitative agreement of the variational results with numerically exact (ED) results, see left panel of Figure 7.2. However, importantly the variational *ansatz* also enables us to study the phase transition at much larger system sizes, see right panels of Figure 7.2. Strikingly for the larger system sizes $N \geq 100$ finite size effects are almost absent allowing us to read off the critical field value directly. Notice, that both variational *ansätze* clearly indicate the existence of a phase transition. However, in agreement with the literature [163] we find that the CS underestimate the

Figure 7.1: *Average error in energy density* $\epsilon = \overline{\Delta E/W}$ *as a function of the transverse field.* ΔE is the difference between the variational energy and the exact ground state energy for different methods (CS in orange, GCS in blue and ED in purple) and for different system sizes $N = 8$, 12, 16 (light to dark). It is normalised by the spectral bandwidth W, *i.e.* the difference between the highest and lowest energies in the exact spectrum. *Inset:* Difference between CS and GCS energies per site $(E_{\mathrm{CS}} - E_{\mathrm{GCS}})/N$ as a function of the system size N for different transverse field values $g/J = 0.1$, 1.0, 1.5, 2.0, 3.0 (light to dark grey). All data is for $h = 0$ and averaged over $n_{\mathrm{samples}} = 1000$ disorder realisations. Results of similar quality can be obtained also for other longitudinal field values $h > 0$.



quantum fluctuations leading to a phase transition at roughly $g_C \approx 2\,J$. In contrast, the GCS capture the true critical point at $g_C \approx 1.5\,J$. Thus, the additional entanglement structure introduced in the GCS not only leads to an improvement in energy but also seems crucial in capturing the physics of the QSK model in the thermodynamic limit.

7.3 Entanglement structure of the ground state

The findings above suggest that the GCS *ansatz* describes the ground state of the QSK model very well for all system sizes up to the thermodynamic limit. Having such an explicit expression for the ground state wavefunction allows us to study in detail its entanglement properties. Before looking into the numerical results, we will consider some hypotheses about the expected entanglement behaviour.

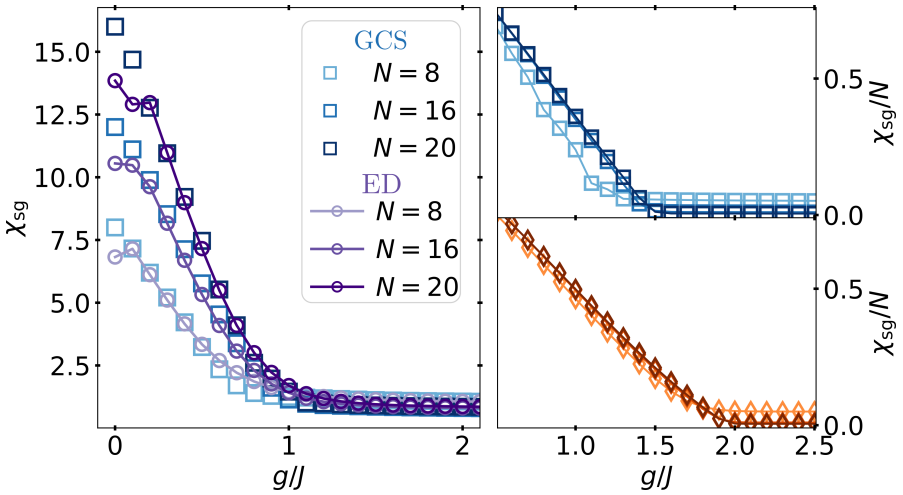
First, let us try to understand the role of the additional two-spin entangling gates contained in $\mathcal{V}(M)$ by taking a closer look at the matrix elements M_{nm}. Considering the probability distribution $p(M_{nm})$ over many disorder realisations, we observe that it resembles a Gaussian distribution with zero mean and variance scaling as $1/N$. In addition, we find that the mean level spacing ratio averaged over many realisations yields $\overline{\langle r \rangle} \approx 0.53$ roughly independent of the transverse field value $g > 0$ and system size N, which is in agreement with the result of the Gaussian Orthogonal Ensemble (GOE) [177].

This implies that most two-spin entangling gates approach the identity as $N \to \infty$. This may seem compatible with the naive hypothesis that, due to the mean-field nature of the model, product states should well describe the ground state, at least in the thermodynamic limit. This assumption would predict the entanglement entropy between any two subsystems going to zero as $N \to \infty$.

However, we emphasize that the number of entangling gates acting on each individual spin diverges in the thermodynamic limit, suggesting that a non-trivial entanglement structure is still possible also in this limit. Indeed, let us consider a subsystem A composed of the first L spins. We quantify the entanglement between these L spins and the rest of the system by computing the second Rényi entropy $S_2(L)$ of the subsystem's reduced density matrix. Given the all-to-all connectivity of our *ansatz*, there exist $L(N-L)$ two-spin entangling gates acting between spins in A and in its complement A^c. Each of these gates individually generates a two-spin state with an average entanglement entropy proportional to $\overline{M_{nm}^2} \sim 1/N$. The cancellation of these two scalings could lead to a second hypothesis, i.e. that the entanglement entropy between A and A^c is proportional to L in the thermodynamic limit $N \to \infty$. This expectation can also be made more rigorous with an argument based on the Central Limit Theorem (see Appendix A.5.2).

As a third alternative, we may compare the model to a related but analytically

Figure 7.2: *Spin glass quantum phase transition. Left Panel:* Spin Glass susceptibility χ_{sg} as a function of the transverse field g for ED (purple circles) and GCS (blue squares). *Right Panels:* Spin Glass susceptibility per site χ_{sg}/N for GCS (top, blue) and CS (bottom, orange) for different system sizes $N = 20, 100, 200$ (from light to dark). All data is for $h = 0$ and averaged over $n_{\text{samples}} = 1000$ disorder realisations.



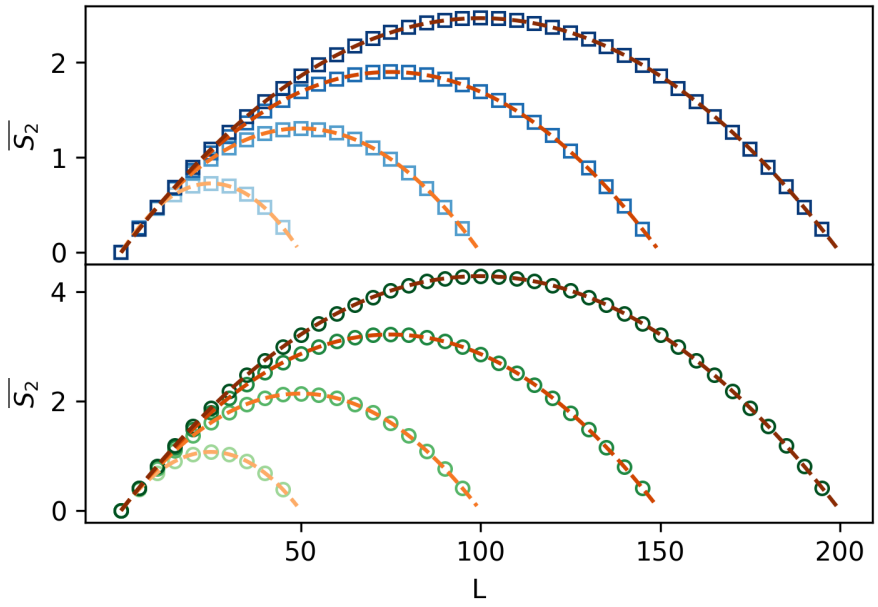
solvable model, namely a model with all-to-all interactions and invariant under any permutation of the spins. Notice that in our case, due to the disordered nature of the QSK model, individual realisations of the couplings J_{nm} and h_n are not permutationally invariant. However, invariance is present upon disorder averaging, so the permutationally invariant case may still provide a useful comparison. In such case the ground state $|\Psi\rangle$ must possess a Schmidt decomposition

$$|\Psi\rangle = \sum_k \lambda_k |\varphi_k\rangle |\eta_k\rangle \ , \tag{7.5}$$

where $|\varphi_k\rangle$ and $|\eta_k\rangle$ are orthonormal states of A and A^c respectively. Due to the permutational invariance of the system, the states $|\varphi_k\rangle$ must in particular belong to the subspace of permutationally invariant states of A. Such subspace has dimension $L + 1$, so there can be at most $L + 1$ terms in the sum (7.5). It follows that the entanglement entropy of A is bounded by $S_2(L) \leq \log(1 + L)$. This scaling of the entanglement can be viewed as a consequence of entanglement monogamy [178, 179].

We would like now to compare our results with these hypotheses. To do this we have developed an efficient method to numerically compute $S_2(L)$ for the states (7.3), reducing the problem to the one of computing averaged properties of a related classical model using Monte Carlo methods (see Appendix A.5.2 for

Figure 7.3: *Ground state entanglement entropy. Top panel:* Average Renyi-2 entanglement entropy of the QSK ground state as a function of the subsystem size L, computed for $g = 1\,J$, $h = 0$ and various total system sizes ($N = 50, 100, 150, 200$ from light to dark blue markers). The data is fitted with the function (7.6) (orange dashed lines). Fits of similar quality can be obtained also for other values of the fields g and h. *Bottom panel:* Average Renyi-2 entanglement entropy of an ensemble of weighted graph states (7.8) as a function of the subsystem size L, computed for various total system sizes ($N = 50, 100, 150, 200$ from light to dark green markers). Also in this case the data is well fitted by the function (7.6) (orange dashed lines).



details). The results for the average entropy, see top panel of Figure 7.3, are well fitted by the empirical functional form

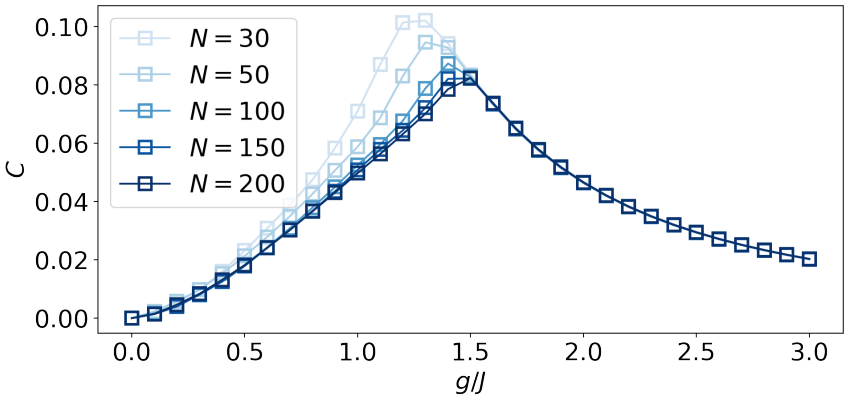
$$\overline{S_2(L;N)} = A(N)\log\left[1 + \frac{B(N)}{\pi}\sin(\frac{\pi L}{N})\right] . \tag{7.6}$$

Notice that, in the large N limit, this functional form may alternatively represent a $\overline{S_2(L)} \sim L$ scaling, a $\overline{S_2(L)} \to 0$ scaling or a $\overline{S_2(L)} \sim \log L$ scaling of the entropy, depending on the behaviour of the fit parameters $A(N)$ and $B(N)$.

In the range of system sizes that we were able to explore ($N \leq 200$), we observe that the parameter $B(N)$ converges to a finite constant as $N \to \infty$. Similarly, the product $C(N) \equiv A(N)B(N)/N$ also converges to a constant C. This suggests the asymptotic behaviour

$$\overline{S_2(L;N)} = CL + \mathcal{O}(1/N)\,, \tag{7.7}$$

Figure 7.4: *Entropy coefficient* $C(N)$. The coefficient $C(N) = A(N)B(N)/N$, extrapolated from the Rényi entanglement entropy fit, is plotted as a function of the transverse field g at $h = 0$ for different system sizes N (different shades of blue).



in the thermodynamic limit. In other words, we observe an entanglement scaling larger than both the one of a product state description and the one of a permutationally invariant model. It instead scales proportionally to the volume L of the considered subsystem.

Finally, we point out that the entanglement structure of the ground states appears to encode very clearly some information about the phase transition of the model. More specifically, if we compute the fit coefficient $C(N)$ defined above as a function of the transverse field g at $h = 0$, we will see that this function develops, in the thermodynamic limit, a discontinuity in its derivative at the critical value $g_C \approx 1.5J$, as shown in Figure 7.4.

Comparison to random weighted graph states

The uncomplicated form of the matrix M, which appears to be distributed according to a GOE, suggests that the entanglement structure of the QSK ground states could be described in an even simpler way. Consider indeed the set of states parametrized as

$$|\Psi(M)\rangle = \mathcal{V}(M)\,|+,\ldots,+\rangle\ , \tag{7.8}$$

where $|+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$. These are a subset of the full variational class (7.3) and, in the context of Quantum Information Theory, are referred to as *weighted graph states* [170]. Let us then consider a random ensemble of such states constructed by drawing the matrix M from a GOE with variance $\overline{M_{nm}^2} = \frac{1}{N}$.

We can compute the average subsystem entanglement entropy $\overline{S_2(L)}$ for this ensemble of states, similarly to what we did for the ground states. We find that this entropy is fitted by the same functional form (7.6) and that the fit parameters $A(N)$ and $B(N)$ obey the same large N scalings as in the ground state case. It is also possible to show analytically that the entanglement of these states must scale according to a volume law, as confirmed by these fits (see Appendix A.5.2).

Let us stress that the actual ground states contain more elaborate features than the states (7.8). The state $|\phi(x)\rangle$ appearing in the variational *ansatz* (7.3) is in general not equal to $|+,\ldots,+\rangle$. Rather, we observe that $|\phi(x)\rangle$ transitions from being polarized in the z-direction at small transverse fields g to being almost fully polarized in the xy-plane of the Bloch sphere for large g. Furthermore, the proportionality constant between $\overline{M_{nm}^2}$ and $\frac{1}{N}$ also shows a non-trivial dependence on g and h.

Nonetheless, the qualitative entanglement structure that we observe in the ground states appears to be present already in the simplified form (7.8) if one samples M from a GOE. In view of this, we conclude that the ground state of the QSK model can be seen as a weighted graph state, where the entangling weights are random variables taken from a Gaussian distribution, while the underlying product state encodes information about the model's phase.

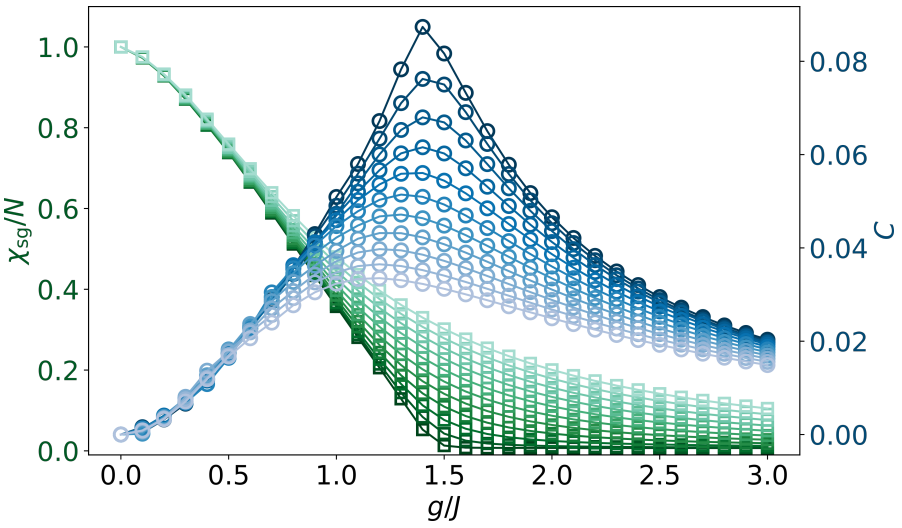
7.4 Phase transition at finite longitudinal fields

Another non-trivial feature of the QSK model which can be studied thanks to our method is the presence of a phase transition at $h > 0$. It has been conjectured that the model's spin glass phase survives also for non-vanishing longitudinal fields h. This would suggest the existence of a line of quantum phase transitions between the spin glass and paramagnetic phases that extends from the $g = g_C$, $h = 0$ critical point into the $h > 0$ plane (often referred to as the quantum de Almeida-Thouless line) [166, 180]. This conjecture is however debated and decisive evidence has yet to be collected in its favour.

Our analysis can provide a new contribution towards answering this question. Indeed, we can extend our analysis to variational ground states in the whole parameter space of the model, including $h > 0$. What we observe is that all indicators of a phase transition vanish as soon as $h > 0$.

More specifically, the spin glass susceptibility χ_{sg} becomes a smooth function of g whenever $h > 0$, no longer presenting the discontinuity in its derivative typical of a phase transition, even at large N. Similarly, the coefficient C characterizing the entropy behaviour of the ground states clearly shows a singular behaviour at $h = 0$ but not for finite h. These results are illustrated in Figure 7.5. In conclusion, our analysis was not able to identify any sign of the presence of a phase transition in

Figure 7.5: *Analysis at finite longitudinal fields.* Behaviour of the spin glass susceptibility per site χ_{sg}/N (green squares) and entropy coefficient C (blue circles) as functions of g for different values of $h/J = 0,\ 0.1,\ \ldots 1$ (dark to light) at $N = 150$. We observe that both functions develop a singularity typical of a phase transition only in the $h = 0$ limit.



the $h > 0$ region of the model's parameter space.

Chapter 8

Conclusion

In this thesis, we have made contributions to the development and understanding of variational methods for many body physics. These contributions have been directed both at understanding variational methods from a general perspective and at the development of specific new tools, especially in the context of Gaussian methods. Taken all together, they provide a robust set of techniques to circumvent the curse of dimensionality of the many body Hilbert space and derive predictions about quantum systems using classical computational resources efficiently.

In the first part of the thesis, we have presented a geometric perspective of variational methods. In Chapter 2, we have shown how a variational set of pure states $\mathcal{M}$ can be understood as a differential geometric manifold embedded in the Hilbert space $\mathcal{H}$. This point of view facilitates the definition of a series of geometric structures associated to $\mathcal{M}$. These structures encode information about the manifold's curvature, its linearity, its complex structure and provide important insights on the correct application of variational methods to it. The most important notion that one can consider is whether the manifold $\mathcal{M}$ is a Kähler manifold or not. This has significant impact on the definition of variational principles, especially for time evolution.

We have then looked more closely at these variational principles, in Chapter 3, defining them on a manifold $\mathcal{M}$ which we assumed to have been characterised in the way described above. We have derived a geometric equivalent for all the most relevant quantum mechanical calculations. We have shown how ground state search can be implemented as a natural gradient descent method on a Riemannian manifold. We have given geometric principles for defining time evolution on a curved variational surface and discussed how the geometry of this surface crucially impacts the inequivalence of different time dependent variational principles. We have described how the excitation energies of a Hamiltonian are encoded in the linear response theory of these time evolution equations on the manifold. In summary, we hope to have convinced the reader of how the toolbox of differential

geometry allows us to construct a complete, self-consistent and rigorous theory of variational methods on any given set of variational states, including ones with exotic but nonetheless useful geometric structures.

In the second part of the thesis, we have shown how one particular class of states – Gaussian states – fits into the variational paradigm. In Chapter 4, we have discussed how the characteristics of Gaussian states are well-suited for an efficient numerical implementation of variational methods. We have in particular focused on how this property stems from a special group-theoretic structure encoded in Gaussian unitaries. This structure is not unique to Gaussian states, indeed we have identified group-theoretic coherent states as the larger category of states in which these properties can be identified.

These observations have led us to realise that these group-theoretic structures can be exploited further than what is normally done in plain Gaussian or coherent states. In particular we have shown how to use them to define sets states which, thanks to their group structures, allow for efficient numerical computations, while at the same time extending the variational manifolds beyond purely Gaussian states. We have provided a rigorous construction of these states, which we called *generalised* group-theoretic coherent states, in Chapter 5, proving the efficiency of the related computational methods.

In the third part of the thesis we have then given some examples of the results which can be obtained using the variational methods described above. In Chapter 6 we have shown how the superfluid phase of the Bose-Hubbard model can be analysed using pure Gaussian states as a variational *ansatz*. Many non-trivial features of this phase can be observed even with such a relatively simple approach, provided one uses the full range of variational principles available. In Chapter 7 we have illustrated the power of the generalised group-theoretic coherent states. Indeed, we have shown how this construction in the context of spin systems provides a very reliable *ansatz* for the ground states of the paradigmatic Quantum Sherrington-Kirkpatrick model.

We believe that using the approaches described in this thesis many more such applications could be found, testifying the usefulness of variational methods. A full understanding of the geometry of variational methods will pave the way to applications involving more elaborate manifolds. Indeed, many variational *ansätze* can be though of that could potentially provide interesting solutions to relevant problems – for example, applying various kinds of parametrised unitaries to fixed reference states. These manifolds will, however, in general have complicated non-Kähler geometric structures. Their naive application will then lead to inconsistent results that only a precise understanding of the geometric nature of variational methods can resolve.

Generalised group-theoretic coherent states can themselves represent a prolific

source of new variational *ansätze* and replicate in other contexts their success at describing the Quantum Sherrington-Kirkpatrick model. We believe in particular that their most unexplored potential lies in the application to systems which contain mixtures of different types of degrees of freedom – for example, spins, bosons and fermions. The generalised group-theoretic construction allows to define states which present non-trivial correlations between these different degrees of freedom which are not normally found in common variational *ansätze*.

Mixed quantum states have not been addressed in this thesis, even though they represent a fundamental tool for the description of *open quantum systems*. A comprehensive description of the geometry and the variational principles for families of mixed states is most likely a non-trivial task, however it would provide a valuable extension and completion of the present work, which we hope will one day be addressed.

Appendix

A.1 Proofs

In this appendix, we present several technical proofs of selected propositions from the main text, whose proof would have interrupted the reading flow.

Proposition 1. *On a tangent space $\mathcal{T}_\psi\mathcal{M} \subset \mathcal{H}$ of a submanifold $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$ we can always find an orthonormal basis $\{|V_\mu\rangle\}$, such that $\boldsymbol{g} \equiv \mathbb{1}$ and the restricted complex structure is represented by the block matrix*

$$\boldsymbol{J}^{\mu}{}_{\nu} \equiv \left(\begin{array}{c|c|c} \begin{array}{cc} & 1 \\ -1 & \\ & \ddots \end{array} & & \\ \hline & \begin{array}{ccccc} c_1 & & & & \\ & -c_1 & & & \\ & & c_2 & & \\ & & & -c_2 & \\ & & & & \ddots \end{array} & \\ \hline & & \begin{array}{cc} 0 & \\ & \ddots \end{array} \end{array}\right) \tag{A.1}$$

with $0 < c_i < 1$. This standard form induces the decomposition of $\mathcal{T}_\psi\mathcal{M}$ into the three orthogonal parts

$$\mathcal{T}_\psi\mathcal{M} = \underbrace{\overline{\overline{\mathcal{T}_\psi\mathcal{M}} \oplus \mathcal{I}_\psi\mathcal{M}}}_{\overline{\mathcal{T}_\psi\mathcal{M}}} \oplus \mathcal{D}_\psi\mathcal{M}\,, \tag{A.2}$$

where $\overline{\overline{\mathcal{T}_\psi\mathcal{M}}}$ is the largest Kähler subspace and $\overline{\mathcal{T}_\psi\mathcal{M}}$ is the largest space on which $\boldsymbol{J}$ and $\boldsymbol{\omega}$ are invertible.

Proof. We focus on a single tangent space $\mathcal{T}_\psi\mathcal{M} \subset \mathcal{H}$ and refer to the Kähler structures on $\mathcal{H}$, rather than the restricted ones on $\mathcal{T}_\psi\mathcal{M}$, as (g, ω, J). To shorten

notation, we define $A := \mathcal{T}_\psi\mathcal{M}$ and B as its orthogonal complement in $\mathcal{H}$ with respect to g, so that $\mathcal{H} = A \oplus B$. We will refer to the restricted Kähler structures on A or B by (g_A, ω_A, J_A) and (g_B, ω_B, J_B), respectively. The relation $J = G\omega = -\Omega g$ implies $g = -\omega J$ or, equivalently[1], $g(v, w) = -\omega(v, Jw)$, and also $g(Jv, Jw) = g(v, w)$ for all $v, w \in \mathcal{H}$. From here, $g(Jv, w) = -g(v, Jw)$ follows and we can derive for $a, a' \in A$

$$g_A(J_A a, a') = g(Ja, a') = -g_A(a, J_A a')\,, \tag{A.3}$$

which implies that J_A is antisymmetric with respect to g_A and thus is diagonalizable, has either vanishing or purely imaginary eigenvalues with the latter appearing in pairs $\pm c_i$. Furthermore, we can always choose an orthonormal basis, such that $g_A = \mathbb{1}$ and J_A is represented by (A.1).
Next, we show that $c_i \in (0, 1]$. We define the orthogonal projectors $\mathbb{P}_A : \mathcal{H} \to A$ and $\mathbb{P}_B : \mathcal{H} \to B$, such that

$$J = \left(\begin{array}{c|c} J_A & J_{AB} \\ \hline J_{BA} & J_B \end{array}\right), \quad \begin{aligned} J_A&\colon\ A \to A,\ a \mapsto \mathbb{P}_A(Ja)\,, \\ J_B&\colon\ B \to B,\ b \mapsto \mathbb{P}_B(Jb)\,, \\ J_{AB}&\colon\ B \to A,\ b \mapsto \mathbb{P}_A(Jb)\,, \\ J_{BA}&\colon\ A \to B,\ a \mapsto \mathbb{P}_B(Ja)\,. \end{aligned} \tag{A.4}$$

We write $J^2 - \mathbb{1} = 0$ in blocks to find

$$\left(\begin{array}{c|c} J_A^2 + J_{AB}J_{BA} - \mathbb{1}_A & J_A J_{AB} + J_{AB} J_B \\ \hline J_B J_{BA} + J_{BA} J_A & J_B^2 + J_{BA} J_{AB} - \mathbb{1}_B \end{array}\right) = 0\,. \tag{A.5}$$

We consider an eigenvector $a \in A$ of J_A with $J_A a = \mathrm{i}ca$ for non-zero c, which implies $J_A^2\, a = -c^2\, a$. We compute

$$\begin{aligned} g(a, a) &= g(J_A a + J_{BA} a, J_A a + J_{BA} a) \\ &= g(J_A a, J_A a) + g(J_{BA} a, J_{BA} a) \\ &\geq g(J_A a, J_A a) = -g(a, J_A^2 a) = c^2 g(a, a) \end{aligned} \tag{A.6}$$

where we used that A and B are orthogonal which eliminates crossing terms. This implies the inequality $c^2 \leq 1$ and thus, we can choose $c_i \in (0, 1]$ as in (A.1). □

Proposition 2. *The Kähler property is equivalent to requiring that $\mathcal{T}_\psi\mathcal{M}$ is not just a real, but also a complex subspace, i.e., for all $|X\rangle \in \mathcal{T}_\psi\mathcal{M}$, we also have $\mathrm{i}|X\rangle \in \mathcal{T}_\psi\mathcal{M}$. Therefore, the multiplication by i commutes with the projector $\mathbb{P}_\psi$, i.e., $\mathbb{P}_\psi \mathrm{i} = \mathrm{i}\mathbb{P}_\psi$ and $\mathbb{P}_\psi$ is complex-linear.*

[1] In the following notation we will treat g and ω as bilinear forms. In other words, $g(v, w) \equiv v^\mu g_{\mu\nu} w^\nu$ for any two vectors $v, w \in \mathcal{H}$

Proof. We want to show that $J_A^2 = -\mathbb{1}_A$ implies that for all $a \in A$, we also have $\mathrm{i}a = Ja \in A$. Therefore, we need to show that $Ja = J_A a$, which is equivalent to $J_{BA} = 0$. For arbitrary $a \in A$, we compute

$$\begin{aligned} g(J_{BA}a, J_{BA}a) &= g(Ja, J_{BA}a) = -g(a, JJ_{BA}a) \\ &= -g(a, J_{AB}J_{BA}a)\,. \end{aligned} \tag{A.7}$$

This expression vanishes if $J_A^2 = -\mathbb{1}_A$, because in that case $J_{AB}J_{BA} = J_A^2 - \mathbb{1}_A = 0$ follows from the first block in (A.5). Since g is non-degenerate, this implies $J_{BA} = 0$. Similarly, we can use the last block in (A.5) to conclude $J_B^2 = -\mathbb{1}_B$, which implies $J_{BA} = 0$. With vanishing J_{AB} and J_{BA}, J is block diagonal and commutes with the projectors. In the language of complex vector spaces, this implies that $\mathbb{P}_\psi \mathrm{i} = \mathrm{i}\mathbb{P}_\psi$. □

Proposition 8. *Given a variational manifold $\mathcal{M}$ we define (according to the Lagrangian action principle) the free projected real time evolution $|\psi(t)\rangle$ as governed by the free Hamiltonian $\hat{H}_0$ and the perturbed projected real time evolution $|\psi_\epsilon(t)\rangle$ as governed by the perturbed Hamiltonian $\hat{H}_\epsilon(t) = \hat{H}_0 + \epsilon\hat{A}(t)$, both with the same initial state $|\psi(0)\rangle$. Then, the propagated perturbation, defined according to* (3.67)*, is given by*

$$\delta x^\mu(t) = -\int_{-\infty}^{t} dt' \, (d\Phi_{t-t'})^\mu{}_\nu \, \mathbf{\Omega}^{\nu\rho} \, \partial_\rho A(t')\big|_{\psi(t')}\,, \tag{A.8}$$

where $d\Phi_t$ is the linearised free evolution flow.

Proof. Let us define the perturbed evolution flow Φ_t^ϵ as the map that sends the coordinates of an initial state $x^\mu(0)$ to the coordinates $x^\mu(t)$ of the state time evolved under the projected perturbed real time evolution. It is governed by

$$\frac{d}{dt}\Phi_t^\epsilon(x) = \mathcal{X}_\epsilon(\Phi_t^\epsilon(x)) = \mathcal{X}_0(\Phi_t^\epsilon(x)) + \epsilon\mathcal{X}_A(\Phi_t^\epsilon(x))\,, \tag{A.9}$$

where $\mathcal{X}_0$ and $\mathcal{X}_A$ are the evolution vector fields associated to the Hamiltonians $\hat{H}_0$ and $\hat{A}$ respectively. We define the free evolution flow Φ_t^0 analogously by just setting $\epsilon = 0$ in the previous expressions.
Let us now define the interaction picture flow $\widetilde{\Phi}_t^\epsilon = \Phi_{-t}^0 \circ \Phi_t^\epsilon$. It has the useful

property that its time evolution only depends on the perturbing vector field:

$$\frac{d}{dt}\widetilde{\Phi}^\epsilon_t(x) = -\mathcal{X}_0(\widetilde{\Phi}^\epsilon_t(x)) + d\Phi^0_{-t}\mathcal{X}_\epsilon(\Phi^\epsilon_t(x)) \tag{A.10}$$

$$= -\mathcal{X}_0(\widetilde{\Phi}^\epsilon_t(x)) + d\Phi^0_{-t}\mathcal{X}_0(\Phi^\epsilon_t(x)) + \epsilon\, d\Phi^0_{-t}\mathcal{X}_A(\Phi^\epsilon_t(x)) \tag{A.11}$$

$$= -\mathcal{X}_0(\widetilde{\Phi}^\epsilon_t(x)) + \frac{d}{dt'}\Big|_{t'=0}\Phi^0_{-t}\Phi^0_{t'}\Phi^\epsilon_t(x) + \epsilon\, d\Phi^0_{-t}\mathcal{X}_A(\Phi^\epsilon_t(x)) \tag{A.12}$$

$$= -\mathcal{X}_0(\widetilde{\Phi}^\epsilon_t(x)) + \frac{d}{dt'}\Big|_{t'=0}\Phi^0_{t'-t}\Phi^\epsilon_t(x) + \epsilon\, d\Phi^0_{-t}\mathcal{X}_A(\Phi^\epsilon_t(x)) \tag{A.13}$$

$$= -\mathcal{X}_0(\widetilde{\Phi}^\epsilon_t(x)) + \mathcal{X}_0(\Phi^0_{-t}\Phi^\epsilon_t(x)) + \epsilon\, d\Phi^0_{-t}\mathcal{X}_A(\Phi^\epsilon_t(x)) \tag{A.14}$$

$$= \epsilon\, d\Phi^0_{-t}\mathcal{X}_A(\Phi^\epsilon_t(x))\,. \tag{A.15}$$

We are interested in the propagated perturbation

$$\delta x^\mu(t) = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\Phi^0_t\widetilde{\Phi}^\epsilon_t(x) = d\Phi^0_t\left(\frac{d}{d\epsilon}\Big|_{\epsilon=0}\widetilde{\Phi}^\epsilon_t(x)\right). \tag{A.16}$$

The quantity $\frac{d}{d\epsilon}\Big|_{\epsilon=0}\widetilde{\Phi}^\epsilon_t(x)$ is for all times a vector of $\mathcal{T}_{\psi(0)}\mathcal{M}$ and its time evolution can be obtained by using (A.15) after having commuted derivatives:

$$\frac{d}{dt}\left[\frac{d}{d\epsilon}\Big|_{\epsilon=0}\widetilde{\Phi}^\epsilon_t(x)\right] = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\left[\frac{d}{dt}\widetilde{\Phi}^\epsilon_t(x)\right] \tag{A.17}$$

$$= \frac{d}{d\epsilon}\Big|_{\epsilon=0}\epsilon\, d\Phi^0_{-t}\mathcal{X}_A(\Phi^\epsilon_t(x)) \tag{A.18}$$

$$= d\Phi^0_{-t}\mathcal{X}_A(\Phi^0_t(x))\,. \tag{A.19}$$

The solution to this equation follows from integrating as

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0}\widetilde{\Phi}^\epsilon_t(x) = \int_{-\infty}^{t} dt'\, d\Phi^0_{-t'}\mathcal{X}_A(\Phi^0_{t'}(x))\,. \tag{A.20}$$

Combining this with (A.16) and the expression (3.9) for the Lagrangian real time evolution vector field $\mathcal{X}_A(\Phi^0_{t'}(x))$ leads to the result (A.8). □

A.2 Kähler manifolds

Kähler manifolds play a central role in this manuscript. For completeness, in this appendix we will discuss their definition and properties. More information can

be found in [81]. A Kähler manifold is a manifold $\mathcal{M}$ equipped with a metric $\boldsymbol{g}$ and a symplectic form $\boldsymbol{\omega}$ that satisfy several properties. These include some local properties, that is that $\boldsymbol{J}^{\mu}{}_{\nu} = -\boldsymbol{G}^{\mu\sigma}\boldsymbol{\omega}_{\sigma\nu}$ verifies $\boldsymbol{J}^2 = -\mathbb{1}$ at all points, and also some non-local properties (closedness of $\boldsymbol{\omega}$ and vanishing Nijenhuis tensor). The precise definition is as follows.

Definition 4 (Kähler manifold)**.** *A real manifold $\mathcal{M}$ is called Kähler if each tangent space is equipped with a positive definite metric $\boldsymbol{g}$ and a compatible symplectic form $\boldsymbol{\omega}$ as in Definition 1, such that $\boldsymbol{J}^{\mu}{}_{\nu} = -\boldsymbol{G}^{\mu\sigma}\boldsymbol{\omega}_{\sigma\nu}$ with $\boldsymbol{J}^2 = -\mathbb{1}$, and the following conditions are satisfied:*

- ***Symplectic form $\boldsymbol{\omega}$ is closed ($d\boldsymbol{\omega} = 0$) with***

$$\begin{aligned}(d\boldsymbol{\omega})_{\mu\nu\sigma} = \tfrac{1}{6}\,(&\partial_\mu\boldsymbol{\omega}_{\nu\sigma} + \partial_\nu\boldsymbol{\omega}_{\sigma\mu} + \partial_\sigma\boldsymbol{\omega}_{\mu\nu} \\ &-\partial_\mu\boldsymbol{\omega}_{\sigma\nu} - \partial_\nu\boldsymbol{\omega}_{\mu\sigma} - \partial_\sigma\boldsymbol{\omega}_{\nu\mu})\ .\end{aligned} \tag{A.21}$$

- ***Nijenhuis tensor $N_{\boldsymbol{J}}$ vanishes ($N_{\boldsymbol{J}} = 0$) with***

$$\begin{aligned}(N_{\boldsymbol{J}})^{\mu}_{\nu\sigma} = \boldsymbol{J}^{\lambda}{}_{\sigma}\,\partial_\lambda\boldsymbol{J}^{\mu}{}_{\nu} - \boldsymbol{J}^{\lambda}{}_{\nu}\,\partial_\lambda\boldsymbol{J}^{\mu}{}_{\sigma} \\ + \boldsymbol{J}^{\mu}{}_{\lambda}(\partial_\nu\boldsymbol{J}^{\lambda}{}_{\sigma} - \partial_\sigma\boldsymbol{J}^{\lambda}{}_{\nu})\,.\end{aligned} \tag{A.22}$$

In essence, a Kähler manifold is simulteneously a Riemannian, a symplectic and a complex manifold, such that the respective structures in every tangent space are compatible in the sense of Definition 1.

For the purpose of the methods presented in this manuscript, it is of interest only whether the restricted Kähler structures on $\mathcal{M}$ satisfy the compatibility conditions from Definition 1. If they do, the manifold is known as an almost-Hermitian manifold. We do not use the additional properties of $\boldsymbol{\omega}$ being closed or $N_{\boldsymbol{J}}$ vanishing.

However, as shown in the following proposition, if an almost-Hermitian manifold $\mathcal{M}$ is also a submanifold of a Kähler manifold, the additional non-local conditions are automatically satisfied and $\mathcal{M}$ is itself a Kähler manifold. In the context of this manuscript we always deal with manifolds $\mathcal{M} \subset \mathcal{P}(\mathcal{H})$, where projective Hilbert space $\mathcal{P}(\mathcal{H})$ is known to be a Kähler manifold [181]. For this reason, for all the manifolds we encounter, the local compatibility conditions from Definition 1 are sufficient conditions for the manifold to be Kähler and we will therefore refer to manifolds that satisfy them as Kähler.

Proposition 15. *Given a Kähler manifold $\tilde{\mathcal{M}}$ with compatible Kähler structures $(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{\omega}}, \tilde{\boldsymbol{J}})$, a sub manifold $\mathcal{M} \subset \tilde{\mathcal{M}}$ equipped with the restricted Kähler structures $(\boldsymbol{g}, \boldsymbol{\omega}, \boldsymbol{J} = -\boldsymbol{G}\boldsymbol{\omega})$ is itself a Kähler manifold provided that $\boldsymbol{J}^2 = -\mathbb{1}$.*

Proof. $\mathcal{M}$ satisfies all local Kähler conditions. We therefore only need to show that $\boldsymbol{\omega}$ is closed and $N_{\boldsymbol{J}} = 0$. We consider local coordinates $\tilde{x}^{\tilde{\mu}}$ on $\tilde{\mathcal{M}}$, such that $x^{\tilde{\mu}} \equiv (x^{\mu}, x'^{\mu'})$ where changes in x^{μ} preserve the submanifold $\mathcal{M}$, while changes in $x'^{\mu'}$ are orthogonal to it. We can further choose x^{μ} and $x'^{\mu'}$ locally, such that the matrix representations of the Kähler structures $(\boldsymbol{\omega}, \boldsymbol{g}, \boldsymbol{J})$ with respect to the decomposition $\tilde{\mu} \equiv (\mu, \mu')$ are

$$\tilde{\boldsymbol{g}} \equiv \left(\begin{array}{c|c} \boldsymbol{g} & 0 \\ \hline 0 & \boldsymbol{g}' \end{array}\right), \; \tilde{\boldsymbol{\omega}} \equiv \left(\begin{array}{c|c} \boldsymbol{\omega} & 0 \\ \hline 0 & \boldsymbol{\omega}' \end{array}\right), \; \tilde{\boldsymbol{J}} \equiv \left(\begin{array}{c|c} \boldsymbol{J} & 0 \\ \hline 0 & \boldsymbol{J}' \end{array}\right), \tag{A.23}$$

which is a consequence of $\boldsymbol{J}^2 = -\mathbb{1}$, as proven in proposition 1. Thus, this implies that $\tilde{\boldsymbol{J}} = \boldsymbol{J} \oplus \boldsymbol{J}'$ with respect to this decomposition $\mathcal{T}_{\psi}\tilde{\mathcal{M}} = \mathcal{T}_{\psi}\mathcal{M} \oplus (\mathcal{T}_{\psi}\mathcal{M})^{\perp}$.

- Symplectic form is closed. In the above basis, $(d\boldsymbol{\omega})_{\mu\nu\sigma}$ corresponds to a sub block of the array $(d\tilde{\boldsymbol{\omega}})_{\tilde{\mu}\tilde{\nu}\tilde{\sigma}}$. Consequently, $d\tilde{\boldsymbol{\omega}} = 0$ implies $d\boldsymbol{\omega} = 0$.

- We restrict $\tilde{N}_{\tilde{\boldsymbol{J}}}$ on $\tilde{\mathcal{M}}$ to $\mathcal{M}$ to find

$$\begin{aligned} (\tilde{N}_{\tilde{\boldsymbol{J}}})^{\mu}_{\nu\sigma} = {} & \boldsymbol{J}^{\tilde{\lambda}}{}_{\sigma}\, \partial_{\tilde{\lambda}} \boldsymbol{J}^{\mu}{}_{\nu} - \boldsymbol{J}^{\tilde{\lambda}}{}_{\nu}\, \partial_{\tilde{\lambda}} \boldsymbol{J}^{\mu}{}_{\sigma} \\ & + \boldsymbol{J}^{\mu}{}_{\tilde{\lambda}} (\partial_{\nu} \boldsymbol{J}^{\tilde{\lambda}}{}_{\sigma} - \partial_{\sigma} \boldsymbol{J}^{\tilde{\lambda}}{}_{\nu}) \end{aligned} \tag{A.24}$$

  which is not obviously equal to $(N_{\boldsymbol{J}})^{\mu}{}_{\nu\sigma}$ due to the contraction over $\tilde{\lambda}$, which takes the full manifold into account. However, our previous considerations showed that $\tilde{\boldsymbol{J}} = \boldsymbol{J} \oplus \boldsymbol{J}'$. This implies that $\boldsymbol{J}^{\tilde{\lambda}}{}_{\mu} = \boldsymbol{J}^{\lambda}{}_{\mu}$, which proves the equality. Consequently $N_{\tilde{\boldsymbol{J}}} = 0$ implies $N_{\boldsymbol{J}} = 0$.

We therefore conclude that any submanifold $\mathcal{M}$ of a Kähler manifold $\tilde{\mathcal{M}}$ with $\boldsymbol{J}^2 = -\mathbb{1}$ everywhere is again a Kähler manifold. Note that this implies in particular that $\mathcal{M}$ is also a complex and a symplectic manifold. □

A.3 Variational methods with generalised group-theoretic coherent states

The main application of a family of states $|\psi(x)\rangle$ such as the one defined in (5.2) (where we indicate with x collectively all the parameters defining the state) is to use it as the *ansatz* for a variational calculation. In this appendix we show that all the relevant quantities one needs to compute for such application can be brought to linear combinations of terms of the form

$$\langle \mu | \mathcal{U}(g)\, \hat{Z}_{i_1} \cdots \hat{Z}_{i_d} | \mu \rangle \; . \tag{A.25}$$

To do this we use the result of section 5.3 that the adjoint action of $\mathcal{V}(M)$ on any polynomial of operators $\hat{Z}_i$ gives rise to a linear combination of products of group operations and algebra operators.

Given a Hamiltonian $\hat{H}$ defined on $\mathcal{H}$, an *ansatz* $|\psi(x)\rangle$ may be used both to approximate the ground state of $\hat{H}$ and to simulate the real time dynamics of the system. This can be done according to different variational principles, as discussed in Chapters 2 and 3 and reference [69] and illustrated for Gaussian states in [52, 122, 182]. To do so it is necessary to be able to compute the following quantities:

$$\langle\psi(x)|\hat{H}|\psi(x)\rangle\,,\ \langle V_\mu(x)|\hat{H}|\psi(x)\rangle\,,\ \langle V_\mu(x)|V_\nu(x)\rangle\,, \tag{A.26}$$

where $|V_\mu(x_0)\rangle = \frac{\partial}{\partial x^\mu}|\psi(x)\rangle\,|_{x=x_0}$ is a so-called *tangent vector* of the variational manifold.

Here, we have assumed that the group $\mathcal{G}$ and its representation have been chosen so that $\hat{H}$ can be expressed as a polynomial in the operators $\hat{Z}_i$. For what concerns the computation of the tangent vectors, it can be shown (see Section 4.2.2 and reference [69]) that the derivatives of $\mathcal{U}(g)$ with respect to a suitable parametrization of the group can be written as linear combinations of terms of the form $\mathcal{U}(g)\,\hat{Z}_i$. Similarly, it holds that

$$\frac{\partial}{\partial M^{ab}}\mathcal{V}(M) = \frac{\mathrm{i}}{2}\,\mathcal{V}(M)\hat{H}_a\hat{H}_b\,. \tag{A.27}$$

Consequently, we have that for generalised group-theoretic coherent states tangent vectors have the form

$$\begin{aligned}|V_\mu(x)\rangle =& C_1^i\,\mathcal{U}(g_1)\hat{Z}_i\mathcal{V}(M)\mathcal{U}(g_2)\,|\mu\rangle \\ &+ C_2^i\,\mathcal{U}(g_1)\mathcal{V}(M)\mathcal{U}(g_2)\hat{Z}_i\,|\mu\rangle \\ &+ C_3^{ab}\,\mathcal{U}(g_1)\mathcal{V}(M)\hat{H}_a\hat{H}_b\mathcal{U}(g_2)\,|\mu\rangle\ .\end{aligned} \tag{A.28}$$

With this in mind, one sees immediately that the quantities (A.26) are made up of terms where one has to evaluate repeatedly the adjoint action of $\mathcal{U}(g)$ or $\mathcal{V}(M)$ on products of operators $\hat{Z}_i$ and then compute the expectation value of the result on $|\mu\rangle$. Using the results (5.15) and (5.16) these give rise to linear combinations of further products of operators $\hat{Z}_i$ and potentially of group transformations $\mathcal{U}(g)$. Using then (5.13) as explained in Section 5.3 to commute all the group transformations to the left, they can thus be all brought to linear combinations of terms of the form (A.25).

A.4 Calculations for the Bose-Hubbard model

A.4.1 Review of Bogoliubov theory

In Chapter 6, we generalise the well-known Bogoliubov theory of the Bose-Hubbard model by extending the underlying variational family from coherent to Gaussian coherent states. To allow for a fair comparison, we review here the three steps involved in traditional Bogoliubov theory.

For later computations, it is useful to write the Bose-Hubbard Hamiltonian (6.1) in momentum space

$$\hat{H} = \sum_k \varepsilon_k \hat{b}_k^\dagger \hat{b}_k + \frac{U}{2N} \sum_{k,p,q} \hat{b}_{k+q}^\dagger \hat{b}_{p-q}^\dagger \hat{b}_k \hat{b}_p \tag{A.29}$$

where we defined $\hat{b}_k = \frac{1}{\sqrt{N}} \sum_n e^{-\mathrm{i}k\cdot x_n} \hat{b}_n$ on the reciprocal lattice and introduced the non-interacting dispersion relation

$$\varepsilon_k = -2 \sum_{d=1}^{\dim} \cos \frac{2\pi k_d}{N_d} - \mu \,, \tag{A.30}$$

where N_d refers to the number of lattice sites in the d-th direction, such that $N = \prod_d N_d$.

Step 1 (coherent variation). Bogoliubov theory approximates the ground state within the class of translationally invariant coherent states, *i.e.*, the states

$$|\beta_0\rangle = \mathcal{D}(\beta_0)\,|0\rangle \quad \text{with} \quad \mathcal{D}(\beta_0) = e^{\beta_0 \hat{b}_0^\dagger - \beta_0^* \hat{b}_0} \,, \tag{A.31}$$

satisfying $\langle \beta_0 |\, b_k \,| \beta_0 \rangle = \beta_0 \delta_{0,k}$. Within this class, the average energy value is minimised for $|\beta_0|$ equal to

$$\beta_0^{\mathrm{c}} := \sqrt{-\varepsilon_0 N / U} \tag{A.32}$$

leading to the expectation value

$$E_{|\beta_0^{\mathrm{c}}\rangle} = \langle \beta_0^{\mathrm{c}} |\, \hat{H} \,| \beta_0^{\mathrm{c}} \rangle = \varepsilon_0 \,|\beta_0^{\mathrm{c}}|^2 + \frac{U}{2N} \,|\beta_0^{\mathrm{c}}|^4 = -\frac{\epsilon_0^2 \, N}{2U} \,, \tag{A.33}$$

provided that $\varepsilon_0 < 0$. There is a larger set of solutions given by $\beta_0 = \beta_0^{\mathrm{c}} e^{\mathrm{i}\varphi}$ associated to the spontaneously broken U(1) symmetry generated by $\hat{N} = \sum_k \hat{b}_k^\dagger \hat{b}_k$.

Step 2 (mean field Hamiltonian). We can use $|\beta_0^{\mathrm{c}}\rangle$ to define the mean field Hamiltonian

$$[\hat{H}]_{|\beta_0^{\mathrm{c}}\rangle} = E_{|\beta_0^{\mathrm{c}}\rangle} + \frac{1}{2} \sum_k \Big(U_k^{\mathrm{c}} (\delta \hat{b}_k^{\mathrm{c}})^\dagger \delta \hat{b}_k^{\mathrm{c}} + V_k^{\mathrm{c}} \delta \hat{b}_k^{\mathrm{c}} \delta \hat{b}_{-k}^{\mathrm{c}} + \mathrm{H.c.} \Big) \,, \tag{A.34}$$

where $\delta\hat{b}_k^{\rm c} = \hat{b}_k - \delta_{k,0}\beta_0^{\rm c}$, $U_k^{\rm c} = \varepsilon_k + \frac{2U}{N}\,|\beta_0^{\rm c}|^2 = \varepsilon_k - 2\varepsilon_0$ and $V_k^{\rm c} = \frac{U}{N}\,(\beta_0^{\rm c})^2 = -\varepsilon_0$. Here, we define $[\hat{H}]_{|\psi\rangle}$ to be the quadratic Hamiltonian resulting from the quadratic truncation of $\hat{H}$ written as normal ordered polynomial in creation and annihilation operators $\delta\hat{b}_k^\dagger$ and $\delta\hat{b}_k$ associated to the Gaussian state $|\psi\rangle$, *i.e.*, in our case, $\delta\hat{b}_k\,|\beta_0^{\rm c}\rangle = 0$.

Step 3 (squeezed ground state). The mean field Hamiltonian $[\hat{H}]_{|\beta_0^{\rm c}\rangle}$ is quadratic, implying that we can diagonalize it by applying the Bogoliubov transformation

$$\mathcal{S}(\lambda) = \exp\left(\tfrac{1}{2}\textstyle\sum_k(\lambda_k\,\hat{b}_k^\dagger\hat{b}_{-k}^\dagger - \lambda_k^*\,\hat{b}_k\hat{b}_{-k})\right)\,. \tag{A.35}$$

We perform the transformation by expressing $\hat{b}_k$ in terms of new creation and annihilation operators

$$\begin{aligned}\delta\hat{B}_k &= \mathcal{D}(\beta)\mathcal{S}(\lambda)\,\hat{b}_k\,\mathcal{S}^\dagger(\lambda)\mathcal{D}^\dagger(\beta) &\text{(A.36)}\\ &= u_k(\hat{b}_k - \beta_k) - v_k(\hat{b}_{-k}^\dagger - \beta_k)\,, &\text{(A.37)}\end{aligned}$$

with $u_k = \cosh\lambda_k$ and $v_k = \sinh\lambda_k$. The diagonalization is accomplished by $\beta = \beta_0^{\rm c}$ and $\lambda_k^{\rm c}$ given by

$$\tanh 2\lambda_k^{\rm c} = -\frac{V_k^{\rm c}}{U_k^{\rm c}}\,, \tag{A.38}$$

such that the Hamiltonian takes the form

$$[\hat{H}]_{|\beta_0^{\rm c}\rangle} = E_{|\beta_0^{\rm c}\rangle} - \Delta^{\rm c} + \sum_k \mathcal{E}_k^{\rm c}\,(\delta\hat{B}_k^{\rm c})^\dagger\delta\hat{B}_k^{\rm c}\,, \tag{A.39}$$

where the energy shift $\Delta^{\rm c}$ and the excitations $\mathcal{E}_k^{\rm c}$ are

$$\begin{aligned}\Delta^{\rm c} &= \frac{1}{2}\sum_k \frac{\varepsilon_0^2}{\varepsilon_k - 2\varepsilon_0 + \sqrt{(2\varepsilon_0-\varepsilon_k)^2 - \varepsilon_0^2}}\,, &\text{(A.40)}\\ \mathcal{E}_k^{\rm c} &= \sqrt{(U_k^{\rm c})^2 - (V_k^{\rm c})^2} = \sqrt{(\varepsilon_k - 2\varepsilon_0)^2 - \varepsilon_0^2}\,. &\text{(A.41)}\end{aligned}$$

The formal ground state of this Hamiltonian is given by a state $|\beta_0^{\rm c}, \lambda^{\rm c}\rangle = \mathcal{D}(\beta_0^{\rm c})\mathcal{S}(\lambda^{\rm c})\,|0\rangle$, such that $\delta\hat{B}_k^{\rm c}\,|\beta_0^{\rm c}, \lambda^{\rm c}\rangle = 0$. The Bogoliubov spectrum is gapless, *i.e.*, we have $\mathcal{E}_0^{\rm c} = 0$.

The full spectrum of the mean field Hamiltonian $[\hat{H}]_{|\beta_0^{\rm c}\rangle}$ is constructed from the 1-particle Bogoliubov spectrum $\mathcal{E}_k^{\rm c}$, whose excitations are non-interacting. Depending on the shape of the function $\mathcal{E}_k^{\rm c}$, *i.e.*, if there exist k, q with $\mathcal{E}_k^{\rm c} > \mathcal{E}_{k-q}^{\rm c} + \mathcal{E}_q^{\rm c}$, it is possible to find a superposition of excitations with total momentum k whose

total energy is less than $\mathcal{E}_k$. To get the lower bound on the excitation continuum, we need to compute

$$\mathcal{E}_k^{\min} = \min_{q_i} \sum_i \mathcal{E}^{\mathrm{c}}_{q_i} \quad \text{with} \quad k = \sum_i q_i \,. \tag{A.42}$$

For the system in one dimension, it is sufficient to compute the slope for $k = 0$, namely $\mathcal{E}'_0$, to find the explicit form of the lower bound to be given by

$$\mathcal{E}_k^{\min} = \min\left(\mathcal{E}_k, \sqrt{4+2\mu}\,\frac{2\pi k}{N}\right) . \tag{A.43}$$

We can derive the condition on μ, such that there is some range of k, for which $\mathcal{E}_k^{\min}$ lies underneath the Bogoliubov spectrum $\mathcal{E}^{\mathrm{c}}_k$. This condition is given by $\partial_k^2 \mathcal{E}^{\mathrm{c}}_k(\mu) = 0$, *i.e.*, the second derivative of $\mathcal{E}^{\mathrm{c}}_k$ must vanish. Its solution is

$$\frac{2\pi k}{N} = \cos^{-1}\left(\frac{8+3\mu-\sqrt{5\mu^2+32\mu+48}}{4}\right) , \tag{A.44}$$

which only exists for $\mu \leq 4$. In higher dimensions, we can consider the slice $k = (k_x, 0, \ldots, 0)$, which leads to an effective rescaling of $\mu \to \mu + 2(\dim -1)$. In this case, we therefore have the condition $\mu < 6 - 2\dim$ to have part of the continuum spectrum to lie underneath the 1-particle dispersion relation $\mathcal{E}_k$.

Let us make the following three important remarks. First, the Bogoliubov energy

$$E_{\text{Bogoliubov}} = E_{|\beta_0^{\mathrm{c}}\rangle} - \Delta^{\mathrm{c}} \tag{A.45}$$

is not variational, *i.e.*, it is not the expectation value of the state $|\beta_0^{\mathrm{c}}, \lambda^{\mathrm{c}}\rangle$ with respect to the full Hamiltonian, but rather the minimal energy of the mean field Hamiltonian $[\hat{H}]_{|\beta_0^{\mathrm{c}}\rangle}$. Only the energy $E_{|\beta_0^{\mathrm{c}}\rangle}$ is variational, *i.e.*, it minimises the energy expectation value within the class of coherent states.

Second, the state $|\beta_0^{\mathrm{c}}, \lambda^{\mathrm{c}}\rangle$ is actually ill defined in the zero mode, due to $\lambda_k^{\mathrm{c}} \to \infty$ for $k \to 0$. Put differently, the minimal energy $E_{\text{Bogoliubov}}$ is only reached in the limit of an infinitely squeezed state, whose energy with respect to the full Hamiltonian actually diverges.

Third, we could have computed the Bogoliubov dispersion relation without defining the mean field Hamiltonian $[\hat{H}]_{|\beta_0^{\mathrm{c}}\rangle}$, but rather just by studying the real time flow of the full Hamiltonian projected on the manifold of coherent states and linearised around the stationary coherent state $|\beta_0^{\mathrm{c}}\rangle$

To see this last point more in detail, consider the manifold of coherent states

$$|\beta\rangle = \mathcal{D}(\beta)\,|0\rangle \quad \text{with} \quad \mathcal{D}(\beta) = e^{\sum_k \left(\beta_k \hat{b}_k^\dagger - \beta_k^* \hat{b}_k\right)} . \tag{A.46}$$

Here, β is a vector written in the momentum basis, *i.e.*, its components are labeled by k. The tangent plane at the state $|\beta\rangle$ is spanned by vectors of the form $\mathcal{D}(\beta)\hat{b}_k^\dagger|0\rangle$. Therefore, the projected real time evolution can be computed from the quantity

$$h_k(\beta) = \langle 0|\hat{b}_k\mathcal{D}^\dagger(\beta)\hat{H}\mathcal{D}(\beta)|0\rangle \ , \tag{A.47}$$

which for the Bose Hubbard model evaluates to

$$h_k(\beta) = \varepsilon_k\beta_k + \frac{U}{N}\sum_{k_1,k_2}\beta^*_{k_1+k_2-k}\beta_{k_1}\beta_{k_2} \, . \tag{A.48}$$

Expressed in real components, the resulting evolution is

$$\begin{pmatrix} \mathrm{Re}\dot{\beta} \\ \mathrm{Im}\dot{\beta} \end{pmatrix} = \begin{pmatrix} \mathrm{Im}\, h(\beta) \\ -\mathrm{Re}\, h(\beta) \end{pmatrix} = -\mathrm{i}\frac{1}{\sqrt{2}}T^{-1}\begin{pmatrix} h(\beta) \\ -h^*(\beta) \end{pmatrix} , \tag{A.49}$$

where we introduced the transformation matrix

$$T = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbb{1} & \mathrm{i}\mathbb{1} \\ \mathbb{1} & -\mathrm{i}\mathbb{1} \end{pmatrix} . \tag{A.50}$$

The symbols β and h denote the column vectors that group the values of β_k and h_k for all values of k.

The linearisation around $|\beta_0^{\mathrm{c}}\rangle$ is then given by

$$\boldsymbol{K} = \left(\tfrac{\partial}{\partial \mathrm{Re}\beta}, \ \tfrac{\partial}{\partial \mathrm{Im}\beta} \right)\begin{pmatrix} \mathrm{Re}\dot{\beta} \\ \mathrm{Im}\dot{\beta} \end{pmatrix} \tag{A.51}$$

$$= \sqrt{2}\left(\tfrac{\partial}{\partial \beta}, \ \tfrac{\partial}{\partial \beta^*} \right)\begin{pmatrix} \mathrm{Re}\dot{\beta} \\ \mathrm{Im}\dot{\beta} \end{pmatrix} T$$

$$= -\mathrm{i}\, T^{-1}\left(\tfrac{\partial}{\partial \beta}, \ \tfrac{\partial}{\partial \beta^*} \right)\begin{pmatrix} h(\beta) \\ -h^*(\beta) \end{pmatrix} T \, , \tag{A.52}$$

where we expressed the derivatives with respect to $\mathrm{Re}\beta$ and $\mathrm{Im}\beta$ in terms of derivatives with respect to β and β^*, taken as independent variables. All derivatives are evaluated at $\beta_k = \delta_{k,0}\beta_0^{\mathrm{c}}$. The matrix $\mathrm{i}\boldsymbol{K}$, whose eigenvalues $\pm\omega$ represent the TDVP estimate of the 1-particle excitation energies of the model, is then, up to similarity transformations, equal to

$$\boldsymbol{K} = \left[\left(\tfrac{\partial}{\partial \beta}, \ \tfrac{\partial}{\partial \beta^*} \right)\begin{pmatrix} h(\beta) \\ -h^*(\beta) \end{pmatrix} \right]_{\beta_k = \delta_{k,0}\beta_0^{\mathrm{c}}} . \tag{A.53}$$

This matrix decomposes into blocks of the form

$$\boldsymbol{K}_k = \begin{pmatrix} U_k^{\mathrm{c}} & V_k^{\mathrm{c}} \\ -V_k^{\mathrm{c}} & -U_k^{\mathrm{c}} \end{pmatrix} \tag{A.54}$$

with $U_k^{\mathrm{c}} = \varepsilon_k - 2\varepsilon_0$ and $V_k^{\mathrm{c}} = -\varepsilon_0$ as before. The eigenvalues $\pm\omega_k$ are given by

$$\omega_k = \sqrt{(U_k^{\mathrm{c}})^2 - (V_k^{\mathrm{c}})^2} = \mathcal{E}_k^{\mathrm{c}}\,, \tag{A.55}$$

which is in full agreement with Bogoliubov theory (A.41).

A.4.2 Computation of Gaussian ground state and equations of motion

Here, we review the underlying analytical and semi-analytical methods associated to Section 6.2, that enabled us to compute the best Gaussian state, *i.e.*, the Gaussian state $|\psi_{\mathrm{g}}\rangle$ with the lowest energy expectation value with respect to $\hat{H}$. We can restrict ourselves to searching for the ground state in the translationally invariant submanifold, which is parametrized by β_0 and $\lambda_k := \lambda_{0,k}$. Due to the U(1) invariance, it is always possible to find a ground state in which both these parameters are real. It turns out to be very convenient to parametrize the state in terms of the Bogoliubov parameters β_k, $u_k = \cosh\lambda_k$ and $v_k = \sinh\lambda_k$ such that

$$\begin{aligned} \delta\hat{B}_k &= \mathcal{U}(\beta,\lambda)\,\hat{b}_k\,\mathcal{U}^\dagger(\beta,\lambda) \\ &= u_k(\hat{b}_k - \beta_k) - v_k(\hat{b}^\dagger_{-k} - \beta_k) \end{aligned} \tag{A.56}$$

will annihilate the Gaussian state $|\beta,\lambda\rangle = \mathcal{U}(\beta,\lambda)\,|0\rangle$.

The stationary point $|\psi_{\mathrm{g}}\rangle$ is characterized by vanishing $\mathbb{P}_{|\psi_{\mathrm{g}}\rangle}(-\hat{H})\,|\psi_{\mathrm{g}}\rangle$, which translates into the conditions

$$\langle 0|\hat{b}_0\mathcal{U}^\dagger(\beta,\lambda)H|\psi(\beta,\lambda)\rangle = 0\,, \tag{A.57}$$

$$\langle 0|\hat{b}_k\hat{b}_{-k}\mathcal{U}^\dagger(\beta,\lambda)H|\psi(\beta,\lambda)\rangle = 0\,. \tag{A.58}$$

Rewriting these conditions in terms of our parameters (β_0,λ_k) gives

$$\varepsilon_0 + \frac{U}{N}(\beta_0^2 + A + 2B) = 0\,, \tag{A.59}$$

$$\left[\varepsilon_k + \frac{2U}{N}(\beta_0^2 + B)\right] u_k v_k + \frac{U}{2N}(\beta_0^2 + A)(u_k^2 + v_k^2) = 0\,, \tag{A.60}$$

where we defined $A = \sum_k u_k v_k$ and $B = \sum_k v_k^2$ and ε_k is the dispersion relation in (A.30). Equations (A.59-A.60) are solved by

$$\beta_0^{\mathrm{g}2} = -\frac{N\varepsilon_0}{U} - A - 2B\,, \tag{A.61}$$

$$u_k^{\mathrm{g}} = \frac{1}{\sqrt{2}}\sqrt{(1-T_k^2)^{-\frac{1}{2}}+1}\,, \tag{A.62}$$

$$v_k^{\mathrm{g}} = \frac{1}{\sqrt{2}}\,\mathrm{sign}\,T_k\,\sqrt{(1-T_k^2)^{-\frac{1}{2}}-1}\,, \tag{A.63}$$

where we introduced the convenient parameter

$$T_k = -\left(1+\frac{2U}{N}\frac{B}{\varepsilon_0}\right)\left(2-\frac{\varepsilon_k}{\varepsilon_0}+\frac{2U}{N}\frac{A+B}{\varepsilon_0}\right)^{-1}. \tag{A.64}$$

This expression for the solution depends on the final values of the quantities A and B which have to be obtained from the coupled equations

$$\begin{aligned} A &= \frac{1}{2}\sum_k \frac{2BU+N\varepsilon_0}{\sqrt{(2U(A+2B)+N(3\varepsilon_0-\varepsilon_k))\,(2AU+N(\varepsilon_0-\varepsilon_k))}}\,,\\ B &= \frac{1}{2}\sum_k \frac{N(\varepsilon_k-2\varepsilon_0)-2(A+B)U}{\sqrt{(2U(A+2B)+N(3\varepsilon_0-\varepsilon_k))\,(2AU+N(\varepsilon_0-\varepsilon_k))}} - \frac{N}{2}\,, \end{aligned} \tag{A.65}$$

which can be solved numerically efficiently independently of the dimensionality of the system and with a linear dependence on the system size.

We can similarly express the energy E and particle density n of a Gaussian state $|\psi\rangle$ in terms of β_0, u_k, v_k, $A = \sum_k u_k v_k$ and $B = \sum_k v_k^2$ as

$$\begin{aligned} E &= \sum_k \varepsilon_k v_k^2 - \frac{N\varepsilon_0^2}{2U} - (A+2B)\varepsilon_0 - \frac{U}{N}(2A+B)B\,,\\ n &= \frac{\langle\hat{N}\rangle}{N} = -\frac{\varepsilon_0}{U} - \frac{A+B}{N}\,. \end{aligned} \tag{A.66}$$

In particular, we can use β_0^{g}, u_k^{g} and v_k^{g} to compute $E_{|\psi_{\mathrm{g}}\rangle}$.

Next, we present a comprehensive derivation of the linearisation matrix $\boldsymbol{K}^{\mu}{}_{\nu}$. It is here convenient to consider a slightly different parametrization of the manifold than the one used in Section 6.3.1, namely $|\psi(\tilde{x})\rangle = \mathcal{U}(x_{\mathrm{g}})\mathcal{U}(\tilde{x})\,|0\rangle$. Note that this is an example of the parametrisation discussed in Section 4.2.2 The physical quantities we are interested in, *i.e.*, the spectrum of $\boldsymbol{K}$, are independent of the choice of parametrization.

Using the parametrization $|\psi(\tilde{x})\rangle$, we can write the projected Schrödinger equation as

$$\frac{d}{dt}|\psi(\tilde{x})\rangle = \frac{d\tilde{x}^\mu}{dt}\partial_\mu|\psi(\tilde{x})\rangle = \mathbb{P}_{|\psi(\tilde{x})\rangle}(-\mathrm{i}\hat{H})|\psi(\tilde{x})\rangle \ . \tag{A.67}$$

The tangent vectors $|v_\mu(\tilde{x})\rangle = \partial_\mu|\psi(\tilde{x})\rangle$ define the metric

$$\boldsymbol{g}_{\mu\nu}(\tilde{x}) = \mathrm{Re}\,\langle v_\mu(\tilde{x})|v_\nu(\tilde{x})\rangle \ , \tag{A.68}$$

whose inverse $\boldsymbol{G}^{\mu\nu}$ is characterized by the property $\boldsymbol{G}^{\mu\rho}(\tilde{x})\boldsymbol{g}_{\rho\nu}(\tilde{x}) = \delta^\mu{}_\nu$. Note that we we use Einstein's sum convention where we sum over repeated indices. We point out here that we are employing a real formalism in which we use the real parametrization (6.5) and we consider the tangent plane as a real vector space, *i.e.*, we consider the vectors $|W\rangle$ and $|W'\rangle = \mathrm{i}|W\rangle$ to be linearly independent (and orthogonal with respect to the real inner product $\mathrm{Re}\,\langle W|W'\rangle$).

Using $\boldsymbol{G}^{\mu\nu}$, we can solve for the time evolution of the coordinates

$$\frac{d\tilde{x}^\mu}{dt} = \tilde{\mathcal{X}}^\mu(\tilde{x}) = \boldsymbol{G}^{\mu\nu}(\tilde{x})\,\mathrm{Re}\,\langle v_\nu(\tilde{x})|\,(-\mathrm{i}\hat{H})|\psi(\tilde{x})\rangle \ . \tag{A.69}$$

The choice of parametrization is convenient because it implies $\boldsymbol{G}^{\mu\nu}(0) = \delta_{\mu\nu}$ as $|v_\mu(0)\rangle = \mathcal{U}(x_\mathrm{g})|W_\mu\rangle$, where

$$\{|W_\mu\rangle\} = \left\{\hat{b}_k^\dagger|0\rangle\,, \hat{b}_{k-q}^\dagger\hat{b}_q^\dagger|0\rangle\,, \mathrm{i}\hat{b}_k^\dagger|0\rangle\,, \mathrm{i}\hat{b}_{k-q}^\dagger\hat{b}_q^\dagger|0\rangle\right\} \tag{A.70}$$

is a set of orthonormal vectors. For a generic $\tilde{x} \neq 0$, the vectors $|v_\mu(\tilde{x})\rangle$ will not be orthonormal, but still form a basis of the tangent space $\mathcal{T}_{|\psi(\tilde{x})\rangle}\mathcal{M} = \mathrm{span}\,\{\mathcal{U}(x_\mathrm{g})\mathcal{U}(\tilde{x})|W_\mu\rangle\}$. We can expand them as

$$|v_\mu(\tilde{x})\rangle = M^\nu{}_\mu(\tilde{x})\,\mathcal{U}(x_\mathrm{g})\mathcal{U}(\tilde{x})|W_\nu\rangle \ , \tag{A.71}$$

with respect to the orthonormal basis $\mathcal{U}(x_\mathrm{g})\mathcal{U}(\tilde{x})|W_\nu\rangle$. The expansion coefficients are encoded in a matrix $M^\nu{}_\mu(\tilde{x})$. For our purposes it suffices to see that $M^\nu{}_\mu(0) = \delta^\nu{}_\mu$. Indeed, exploiting this and the fact that the approximate ground state is also the stationary point of real time evolution, *i.e.*,

$$\mathcal{X}^\mu(0) = \mathrm{Re}\,\langle W_\mu|\mathcal{U}^\dagger(x_\mathrm{g})(-\mathrm{i}\hat{H})|\psi(0)\rangle = 0, \tag{A.72}$$

we can linearise the equation of motion (A.69) around $\tilde{x} = 0$ finding

$$\begin{aligned}\boldsymbol{K}^\mu{}_\nu &= \frac{\partial}{\partial\tilde{x}^\mu}\bigg|_{\tilde{x}=0}\tilde{\mathcal{X}}^\mu(\tilde{x}) \\ &= \frac{\partial}{\partial\tilde{x}^\mu}\bigg|_{\tilde{x}=0}\mathrm{Re}\,\langle W_\mu|\,\mathcal{U}^\dagger(\tilde{x})\mathcal{U}^\dagger(x_\mathrm{g})(-\mathrm{i}\hat{H})\,|\psi(\tilde{x})\rangle \ .\end{aligned} \tag{A.73}$$

This is equivalent to linearising the equations of motion for x^μ around x^μ_{g} as described in (6.12).

Evaluating the matrix K from (A.73) reduces to calculating expectation values using Wick's theorem and taking derivatives. It leads to the form

$$\boldsymbol{K} = -\mathrm{i}T^{-1}ST \tag{A.74}$$

where the matrix T is defined in (A.50), taking the subdivision into blocks to refer to the split between real and imaginary parameters in (6.5).

The matrix S is block diagonal with each block S_k referring to a fixed total momentum. Each block can be written as the sum of a diagonal matrix and a rank 5 matrix, that is $S_k = E + CR$, with

$$E = \begin{pmatrix} E_k & 0 & 0 & 0 \\ 0 & \Delta_{q,\tilde{q}} & 0 & 0 \\ 0 & 0 & -E_k & 0 \\ 0 & 0 & 0 & -\Delta_{q,\tilde{q}} \end{pmatrix}, \; R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & a_{k,\tilde{q}} & 0 & a_{k,\tilde{q}} \\ 0 & b_{k,\tilde{q}} & 0 & c_{k,\tilde{q}} \\ 0 & c_{k,\tilde{q}} & 0 & b_{k,\tilde{q}} \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & G_k & \frac{2U}{N}\beta_0(u_k+v_k) & \frac{U}{N}\beta_0 v_k & \frac{U}{N}\beta_0 u_k \\ F_k & \bar{F}_k & \frac{2U}{N}a_{k,q} & \frac{U}{2N}b_{k,q} & \frac{U}{2N}c_{k,q} \\ -G_k & 0 & -\frac{2U}{N}\beta_0(u_k+v_k) & -\frac{U}{N}\beta_0 u_k & -\frac{U}{N}\beta_0 v_k \\ -\bar{F}_k & -F_k & -\frac{2U}{N}a_{k,q} & -\frac{U}{2N}c_{k,q} & -\frac{U}{2N}b_{k,q} \end{pmatrix}$$

with newly introduced parameters

$$\begin{aligned} E_k &= (\varepsilon_k + \frac{2U}{N}(\beta_0^2 + B))(u_k^2 + v_k^2) + \frac{2U}{N}(\beta_0^2 + A)u_k v_k, \\ \Delta_{q,\tilde{q}} &= (\delta_{q,\tilde{q}} + \delta_{q,-\tilde{q}})(E_{\frac{k}{2}+\tilde{q}} + E_{\frac{k}{2}-\tilde{q}}), \\ G_k &= 2(\varepsilon_k + \frac{2U}{N}(\beta_0^2 + B))u_k v_k + \frac{U}{N}(\beta_0^2 + A)(u_k^2 + v_k^2), \\ F_k &= \frac{U}{N}\beta_0\left[2a_{k,q}(u_k + v_k) + b_{k,q}v_k + c_{k,q}u_k\right], \\ \bar{F}_k &= \frac{U}{N}\beta_0\left[2a_{k,q}(u_k + v_k) + b_{k,q}u_k + c_{k,q}v_k\right], \\ a_{k,q} &= u_{\frac{k}{2}+q}v_{\frac{k}{2}-q} + u_{\frac{k}{2}-q}v_{\frac{k}{2}+\tilde{q}}, \\ b_{k,q} &= 2v_{\frac{k}{2}+q}v_{\frac{k}{2}-q}, \\ c_{k,q} &= 2u_{\frac{k}{2}+q}u_{\frac{k}{2}-q}, \end{aligned}$$

where u_k and v_k have to be evaluated at the solutions corresponding to the ground state approximation defined in Appendix A.4.2, *i.e.*, at β_0^{g}, u_k^{g} and v_k^{g} from (A.61-A.63), and ε_k is the dispersion relation in (A.30).

Given the simple structure of the blocks S_k, it is easy to diagonalize them numerically. Their eigenvalues are the zeros of the function $f(\omega) := \det[1 + R(E-\omega)^{-1}C]$. Evaluation only scales linearly with the system size N and moreover, we can characterize analytically some properties of the spectrum in the thermodynamic limit.

More specifically, the function $f(\omega)$ presents a series of poles, given by the diagonal elements of E. Its zeros (*i.e.*, the eigenvalues of the system) are positioned one in between each pair of subsequent poles. One subset of the poles, that is the diagonal elements of $\Delta_{q,\tilde{q}}$, for $N \to \infty$ come closer together, creating in the thermodynamic limit a continuous line. The zeros that are in between such poles will therefore also come together to a continuum that represents the continuum in the spectrum of S_k. The boundaries of this continuum can thus be inferred by computing the values of the minimal and maximal diagonal elements of $\Delta_{q,\tilde{q}}$. In particular, we can identify the minimum, given by $2E_k$, with the Higgs excitation mode. In order to give an expression for the Higgs gap at zero momentum, we need to evaluate $2E_0$.

We want to do this at constant filling n, which is equivalent to imposing $\varepsilon_0 = -U\left(n + \frac{A+B}{N}\right)$, due to (A.66). Substituting this condition into equations (A.65), we find equations for A and B at fixed n. These equations admit constant solutions in the limit $U \to 0$. Inserting these solutions in the expression for E_0, we find the asymptotics of the Higgs gap at constant density for $U \to 0$

$$2E_0 \sim \alpha(N,n)\, U \quad \text{as} \quad U \to 0\,. \tag{A.75}$$

The function α has a complicated analytical expression that admits the large N asymptotics

$$\alpha(N,n) \sim 2\sqrt[3]{2}n^{\frac{2}{3}}N^{-\frac{1}{3}} \quad \text{as} \quad N \to \infty\,. \tag{A.76}$$

Finally, notice that $\boldsymbol{K}$ has eigenvectors $e^{\mu}(\omega)$ appearing in complex conjugate pairs satisfying

$$\begin{aligned} \boldsymbol{K}^{\mu}{}_{\nu} e^{\nu}(\omega) &= +\mathrm{i}\omega e^{\mu}(\omega)\,, \\ \boldsymbol{K}^{\mu}{}_{\nu} e^{*\nu}(\omega) &= -\mathrm{i}\omega e^{*\mu}(\omega)\,. \end{aligned} \tag{A.77}$$

We assume to normalise them such that $\delta(\omega) = \pm 1$, where

$$\delta(\omega) = [\mathrm{Im}\, e^{\mu}(\omega)]\Omega_{\mu\nu}[\mathrm{Re}\, e^{\nu}(\omega)]\,. \tag{A.78}$$

A.5 Calculations for the Sherrington-Kirkpatrick model

A.5.1 Variational Method

In this appendix we will present the algorithm used to find the variational ground states for the generalised atomic coherent states of the form

$$|\Psi\rangle = \mathcal{U}(y)\mathcal{V}(M)\,|\phi(x)\rangle\ , \tag{A.79}$$

defined in Chapter 7.

The GCS variational ground state is described by the parameters $\zeta_0 = (y_0,\, x_0,\, M_0)$ which minimise the energy $E(\zeta) = \langle\Psi(\zeta)|\,\hat{H}\,|\Psi(\zeta)\rangle$, where $\hat{H}$ is the Hamiltonian of the system. For optimisation of the parameters we used the natural gradient descent (natural GD) algorithm (also known as imaginary time evolution, see Section 3.5). Like for standard gradient descent (GD), natural GD starts at some initial state and iteratively updates the parameters, with an update based on the local structure, until a minimum in the energy is reached. However, while the parameters in GD are updated in the direction $\mathcal{X}$ of the energy gradient $\mathcal{X} = -\nabla E$, the direction for natural gradient descent is defined by

$$g\mathcal{X} = -\nabla E \tag{A.80}$$

and encodes additional information on the curvature in terms of the local metric $g_{\mu\nu} = 2\mathrm{Re}\,\langle V_\mu|V_\nu\rangle$, where the tangential vectors $|V_\mu\rangle$ are specified below. Natural GD in general leads to enhanced convergence compared to GD, however, like GD can get stuck in local, non-optimal minima of the energy. In order to avoid this, we employ an adiabatic updating procedure. Thereby, we start at $g = 0$, where the system is exactly described by a product state and apply the natural GD algorithm to a large number of random initial CS (usually 10.000) and use the state with minimal energy as the variational ground state for both CS and GCS at $g = 0$. Then, iteratively for increasing transverse field values $g > 0$, we use $\zeta_0(g) + \eta$, that is the variational ground state parameters of the point g with some small perturbation η, as the starting point of the natural GD algorithm to find the optimal parameters $\zeta_0(g + \delta g)$ for the point $g + \delta g$.

In order to perform the optimisation procedure described above, we will need to compute quantities of the form

$$\langle\Psi|\,\hat{H}\,|\Psi\rangle\,,\ \langle\Psi|\,\hat{H}\,|V_\mu\rangle\,,\ \langle V_\mu|V_\nu\rangle\ , \tag{A.81}$$

corresponding to the energy of the state, the derivative of the energy with respect to the variational parameters and the local structure of the variational manifold,

respectively. The $|V_\mu\rangle$ are the so called tangential vectors, describing the change in the state $|\Psi\rangle$ upon infinitesimal change in the variational parameters

$$|V_\mu\rangle = \mathbb{Q}_\Psi \frac{\partial}{\partial \zeta^\mu} |\Psi\rangle \ , \tag{A.82}$$

where we take the derivative with respect to the μ-th variational parameter. The projection $\mathbb{Q}_\Psi |\phi\rangle = |\phi\rangle - \langle\Psi|\phi\rangle |\Psi\rangle$ removes all directions which lead only to a change in phase or amplitude of the state $|\Psi\rangle$, *i.e.*, which would not change the physical state.

There are three different kinds of tangential vectors

$$\begin{aligned} |\mathcal{X}_n^i\rangle &= \mathbb{Q}_\Psi \mathcal{U}(y)\mathcal{V}(M)\mathcal{U}(x)(i\hat{\sigma}_i^n)\,|\uparrow\rangle^{\otimes N} \ , \\ |\mathcal{M}_{nm}\rangle &= \mathbb{Q}_\Psi \mathcal{U}(y)\mathcal{V}(M)\left(\frac{-i}{4}\hat{\sigma}_z^n\sigma_z^m\right)|\phi(x)\rangle \ , \\ |\mathcal{Y}_m^j\rangle &= \mathbb{Q}_\Psi \mathcal{U}(y)(i\hat{(\sigma)}_j^m)\mathcal{V}(M)\,|\phi(x)\rangle \ . \end{aligned} \tag{A.83}$$

corresponding to the three kinds of variational parameters x_n^i, M_{nm} and y_n^i, respectively.

To evaluate the quantities (A.81) let us first observe that the adjoint action of the rotation unitaries $\mathcal{U}(x)$, defined below equation (7.2), on a product of Pauli operators simply results in independently rotating each Pauli operator according to

$$\mathcal{U}(x)^\dagger \hat{\sigma}_i^n \hat{\sigma}_j^m \ldots \mathcal{U}(x) = \left(\sum_{i'} R_n^{ii'}\hat{\sigma}_{i'}^n\right)\left(\sum_{j'} R_m^{jj'}\hat{\sigma}_{j'}^m\right)\ldots \ , \tag{A.84}$$

with the orthogonal matrices $\boldsymbol{R}_n(x) = \boldsymbol{R}_n(x_n)$ depending only on the parameters for the n-th spin x_n^i, where $i = x, y, z$. Hence, for any of the quantities in equation (A.81), we can take care of the action of $\mathcal{U}(y)$ simply by rotating the Pauli operators that appear in $\hat{H}$.

Let us now consider product operators $\hat{O} = \bigotimes_{n=1}^N \hat{O}_n$, where $\hat{O}_n$ acts only on the n-th spin, since any operator is a linear combination of such product operators. Another direct consequence of equation (A.84) is that for two CS $|\phi\rangle$ and $|\chi\rangle$ the quantity $\langle\phi|\,\hat{O}\,|\chi\rangle$, where $\hat{O}$ is an arbitrary product operator, factorizes into a product of N single-spin terms

$$\langle\chi|\,\hat{O}\,|\psi\rangle = \prod_n \langle\chi_n|\,\hat{O}_n\,|\psi_n\rangle \tag{A.85}$$

and is thus efficiently calculable.

Finally, in order to compute expectation values for GCS, the key observation is

$$\mathcal{V}(M)^\dagger \hat{\sigma}_\alpha^n \mathcal{V}(M) = \hat{\sigma}_\alpha^n \exp(\frac{\alpha i}{4} \sum_m M_{nm} \hat{\sigma}_z^m) \tag{A.86}$$
$$\equiv \hat{\mathcal{O}}_n^\alpha(M)\,,$$

for $\alpha \in \{+, -, 0\}$ and $\hat{\sigma}_{\alpha=0} \equiv \hat{\sigma}_z$. Note that in what follows Greek subscripts α, β, ... will refer to $+$, $-$, 0. The relation (A.86) is a direct consequence of the commutation relations $[\hat{\sigma}_\alpha^n, \hat{\sigma}_z^m] = \alpha \delta_{nm} \hat{\sigma}_\alpha^n$.

Notice that the operator $\hat{\mathcal{O}}_n^\alpha(M)$ is a product operator. Moreover, for multiple Pauli operators we can insert identities $\mathbb{1} = \mathcal{V}\mathcal{V}^\dagger$, such that

$$\mathcal{V}^\dagger \hat{\sigma}_\alpha^n \hat{\sigma}_\beta^m \ldots \mathcal{V} = \mathcal{V}^\dagger \hat{\sigma}_\alpha^n \mathcal{V}\mathcal{V}^\dagger \hat{\sigma}_\beta^m \mathcal{V}\mathcal{V}^\dagger \ldots \mathcal{V} \tag{A.87}$$
$$= \hat{\mathcal{O}}_n^\alpha \hat{\mathcal{O}}_m^\beta \ldots \tag{A.88}$$

is again a product operator.

Thus, using the special relation (A.86), as well as the explicit form of the GCS (7.3) and the tangential vectors (A.83), one immediately finds that the quantities (A.81) are simply sums of expectation values of product operators with respect to the CS part $|\phi(x)\rangle$ of the GCS and can thus be computed efficiently.

Let us point out that the procedure described above for the computation of expectation values $\langle \hat{\sigma}_i^n \hat{\sigma}_j^m \ldots \rangle_\Psi$ with respect to a GCS $|\Psi\rangle$ scales polynomially in the system size, but at the expense of scaling exponentially in the number of Pauli operators. However, for the present application one will have to compute expectation values of products of at most 4 Pauli operators, so this scaling does not pose a problem.

A.5.2 Rényi-2 entropy

In this appendix we will present a method that can be used to efficiently estimate numerically the second Rényi entropy of entanglement for states of the form

$$|\Psi\rangle = \mathcal{U}(y)\mathcal{V}(M)\,|\phi(x)\rangle\,, \tag{A.89}$$

that is GCS as defined in (7.3). Let us consider a system of N spins and a partition of the spins into two sets A and A^c constituted of L and $N-L$ spins, respectively. We are interested in computing the Rényi-2 entropy $S_2 = -\log_2(q_A)$ of the reduced state $\rho_A = \mathrm{tr}_{A^c} |\Psi\rangle\langle\Psi|$, where q_A is the purity $q_A = \mathrm{tr}(\rho_A^2)$. Notice that the local unitaries contained in $\mathcal{U}(y)$ do not modify this quantity in any way, so in what follows we will assume them to be all equal to the identity.

We will show that the quantity q_A can be rewritten in terms of a sampling problem of a set of L classical spin-1 variables, taking values -1, 0 and $+1$. That is, we have

$$q_A = \sum_{j_1,\ldots,j_L=-1,0,+1} P_1(j_1)\cdots P_L(j_L)F(j_1,\ldots,j_L)\,, \tag{A.90}$$

for a certain function F and certain probability distributions P_n. Thus, one can estimate q_A by sampling configurations of the classical spins $\{j_n\}$ according to the product probability distribution $P_1\cdots P_L$ and computing the expectation value A.90 as the mean value of F. To achieve an error ϵ on q_A it is sufficient to sample $\sim 1/\epsilon^2$ configurations, rather than compute all the exponentially many terms in the sum (A.90). Notice, that the entropy S_2 is invariant if one exchanges the sets A and A^c, so we can always choose A to be the smallest of the two.

To rewrite q_A let us consider an ancillary system also made up of N spins and prepared to be in a copy of the state $|\Psi\rangle$. We will denote quantities relative to this ancillary system with primes. We then have

$$q_A = \langle \Psi,\, \Psi|\mathcal{S}_{AA'}|\Psi,\, \Psi\rangle\ , \tag{A.91}$$

where $\mathcal{S}_{AA'}$ is the swap operator acting between the spins in A and the corresponding ancillas in A'.

Note that the terms in $\mathcal{V}(M)$ that only connect spins within A or within A^c do not contribute to (A.91). We can therefore replace $\mathcal{V}(M)$ with

$$\tilde{\mathcal{V}}(M) = \exp(-\frac{i}{4}\sum_{n\in A^c}\hat{\sigma}_z^n\sum_{m\in A}M_{nm}\hat{\sigma}_z^m)\,. \tag{A.92}$$

We will also assume that $|\phi(x)\rangle = \bigotimes_{n=1}^N |\phi_n\rangle$ with $|\phi_n\rangle = c_n^0|\uparrow\rangle + c_n^1|\downarrow\rangle$.

We can then write

$$q_A = \langle\phi(x),\phi(x)|\,\tilde{\mathcal{V}}^\dagger\tilde{\mathcal{V}}'^\dagger\mathcal{S}_{AA'}\tilde{\mathcal{V}}\tilde{\mathcal{V}}'\mathcal{S}_{AA'}\,|\phi(x),\phi(x)\rangle\ , \tag{A.93}$$

where we exploited the fact that $\mathcal{S}_{AA'}\,|\phi(x),\phi(x)\rangle = |\phi(x),\phi(x)\rangle$ to add an extra swap operator. We then act with the swap operators on $\tilde{\mathcal{V}}\tilde{\mathcal{V}}'$ (exchanging system and ancilla operators in A) to obtain

$$\hat{Q} \equiv \tilde{\mathcal{V}}^\dagger\tilde{\mathcal{V}}'^\dagger\mathcal{S}_{AA'}\tilde{\mathcal{V}}\tilde{\mathcal{V}}'\mathcal{S}_{AA'} \tag{A.94}$$

$$= \exp(\frac{i}{4}\sum_{n\in A^c}(\hat{\sigma}_z^n - \hat{\sigma}_z^{n\,\prime})\sum_{m\in A}M_{nm}(\hat{\sigma}_z^m - \hat{\sigma}_z^{m\,\prime})) \tag{A.95}$$

$$= \prod_{n\in A^c}\exp(\frac{i}{4}(\hat{\sigma}_z^n - \hat{\sigma}_z^{n\,\prime})\sum_{m\in A}M_{nm}(\hat{\sigma}_z^m - \hat{\sigma}_z^{m\,\prime})) \tag{A.96}$$

$$\equiv \prod_{n\in A^c}\hat{Q}_n\,. \tag{A.97}$$

Each operator $\hat{Q}_n$ has support on the single spin $n \in A^c$ and on all of A. We can therefore first take the expectation value of each $\hat{Q}_n$ on the term $|\phi_n, \phi_n\rangle$ corresponding to the spin n (of the system and of the ancilla) within the product state $|\phi(x), \phi(x)\rangle$. We easily find

$$\langle \phi_n, \phi_n | \, \hat{Q}_n \, | \phi_n, \phi_n \rangle = 1 - 4 p_n \sin^2 \left[\frac{1}{4} \sum_{m \in A} M_{nm} (\hat{\sigma}_z^m - \hat{\sigma}_z^{m\prime}) \right] , \tag{A.98}$$

where we set $p_n = |c_n^0|^2 |c_n^1|^2$.

We then proceed to take the expectation value of $\prod_{n \in A^c} \langle \phi_n, \phi_n | \, \hat{Q}_n \, | \phi_n, \phi_n \rangle$ on the remaining part of the state $|\phi(x), \phi(x)\rangle$ corresponding to the subsystem A which leads to the expression (A.90), once we define

$$P_m(0) = |c_m^0|^4 + |c_m^1|^4 \tag{A.99}$$

$$P_m(+1) = P_m(-1) = |c_m^0|^2 |c_m^1|^2 , \tag{A.100}$$

which gives rise to a well-defined probability distribution. The function F turns out to be

$$F(j_1, \ldots, j_L) = \prod_{n \in A^c} \left[1 - 4 p_n \sin^2 \left(\frac{1}{2} \sum_{m \in A} M_{nm} j_m \right) \right] , \tag{A.101}$$

which can be evaluated efficiently for any configuration of js. Note how the classical spin-1 variables j_m emerge as the possible eigenvalues of the operators $(\hat{\sigma}_z^m - \hat{\sigma}_z^{m\prime})/2$.

The form (A.90) of the purity can also be used to prove that the ensemble of random weighted graph states discussed in Section 7.3 around equation (7.8) must have a volume law scaling of the entanglement.

For this, let us use the fact that S_2 is invariant under exchange of A and A^c to rewrite (A.90) as

$$\begin{aligned} S_2(L) &= - \log_2 \left[\sum_{\{j\}} P(j_{L+1}) \cdots P(j_N) F(j_{L+1}, \ldots, j_N) \right] \\ &= - \log_2 \left[\sum_{\{j\}} P(j_{L+1}) \cdots P(j_N) \prod_{n=1}^{L} f(X_n) \right] , \end{aligned} \tag{A.102}$$

where $f(x) = 1 - \sin^2 x$ and X_n are random variables defined by

$$X_n = \frac{1}{2} \sum_{m=L+1}^{N} M_{nm} j_m \, . \tag{A.103}$$

Notice, that in the case of weighted graph states where we fix $|\phi(x)\rangle = |+\cdots+\rangle$ the probability distributions P_n are all the same for each j_n and are given by $P(0) = 1/2$, $P(\pm 1) = 1/4$.

The variables X_n are the sum of a large number of independently distributed random numbers. By the Central Limit Theorem we can therefore assume that, in the limit of large N (and fixed L), the variables X_n are distributed according to normal distributions with mean and variance given by

$$\langle X_n \rangle = 0 \tag{A.104}$$

$$\left\langle X_n^2 \right\rangle = \frac{1}{8} \sum_{m=L+1}^{N} M_{nm}^2 \,, \tag{A.105}$$

where by $\langle \ \cdot \ \rangle$ we denote averaging over the variables j. Note that each entry M_{nm} of the matrix M is an independent identically normally distributed variable. We can therefore assume that in the large N limit the sum $\sum_m M_{nm}^2$ will approximate the variance of M_{nm}. More precisely

$$\sum_{m=L+1}^{N} M_{nm}^2 \approx (N-L)\,\overline{M_{nm}^2} = (N-L)\,\frac{1}{N} \,. \tag{A.106}$$

It follows that $\langle X_n^2 \rangle \to 1/8$ for $N \to \infty$.

We can also assume that the variables X_n are independently distributed. Indeed, their correlator is given by

$$\langle X_n X_m \rangle = \frac{1}{8} \sum_{l=L+1}^{N} M_{nl} M_{ml} \,. \tag{A.107}$$

For $n \neq m$ this correlator has vanishing average with respect to the disorder of M. Its variance is a function of $\overline{M^2}$ and can be seen to decay as $1/N$.

From all these considerations we can conclude that equation (A.102) will ultimately reduce to

$$S_2(L) = -\log_2 \left\langle \prod_{n=1}^{L} f(X_n) \right\rangle \tag{A.108}$$

$$= -\log_2 \prod_{n=1}^{L} \langle f(X_n) \rangle \tag{A.109}$$

$$= -\sum_{n=1}^{L} \log_2 \langle f(X_n) \rangle \tag{A.110}$$

$$= -\sum_{n=1}^{L} \log_2 \frac{1 + e^{-2\langle X_n^2 \rangle}}{2} \,. \tag{A.111}$$

The factorisation in (A.109) is valid only up to corrections containing the correlator $\langle X_n X_m \rangle$, however we have seen that this correlator will go to zero at least as $1/N$ in the limit $N \to \infty$. In step (A.111) we have simply computed the average of $f(x)$ over a normally distributed variable with zero mean and variance $\langle X_n^2 \rangle$.

In the preceding paragraphs we have seen how, in the large N limit, the variances $\langle X_n^2 \rangle$ actually neither depends on n nor on the specific realization of M, but rather all tend to $1/8$. We can therefore arrive at the result

$$\overline{S_2(L)} = L \left(-\log_2 \frac{1 + e^{-\frac{1}{4}}}{2} \right) = C\, L \tag{A.112}$$

which shows that the volume law entanglement entropy scaling holds in the limit of large N and fixed L. The value of the constant C that we have derived analytically here coincides numerically with the one that can be extracted from the functional fits discussed in Section 7.3.

A similar behaviour can be expected also in the case of the QSK ground states. There, however, the product state $|\phi(x)\rangle$ has some structure which will make the distributions $P_n(j)$ depend on n. It follows, that the variances $\langle X_n^2 \rangle$ will also depend on n in a non-trivial way. However, they will still be of order 1 in the large N limit and therefore the expression (A.111) remains extensive in L.

Acknowledgements

Achieving the successes as a researcher, which are collected in this thesis, and getting through the more difficult and frustrating moments, which are equally part of a scientist's life, have both been possible thanks to the help and support of many people, to whom I owe a word of thanks.

Before anyone else, I would like to thank my supervisor Ignacio Cirac. The amount of time and energy he dedicates to providing high quality advice and supervision to many students and young researchers is admirable. I feel lucky and honoured to have been one of them. His unbounded knowledge of our field was an irreplaceable resource for completing this thesis. His approach to science, research and leadership, as well as the great respect that he tirelessly shows to any person or idea, will with no doubt remain a model for me throughout my career.

During my time as a PhD student I have also had the pleasure of collaborating with several other outstanding scientists who have greatly contributed to my understanding of physics. Firstly Eugene Demler, with whom one can always have endless and inspiring conversations, that have often given me completely new perspectives about many Condensed Matter Physics problems. I would like to thank him also for the warmest hospitality and kindness shown to me on my visits to Cambridge and Zurich. Then Lucas Hackl, who provided me with indispensable guidance in the first years of my PhD, both in what concerns science and in what concerns life in general, and from whom it is always possible to learn something new on Mathematical Physics. Finally Tao Shi who, with energy and humour, has constantly provided me with inspiration and motivation to discover more about Variational Methods. Together with them, I would also like to thank all my other collaborators and coauthors: Claudius Hubig, Jutho Haegeman, Paul Schindler and Zongping Gong.

Every member of the MPQ theory group, past and present, would deserve to be thanked individually: for creating a rich and stimulating environment, for having solutions to most problems a physicist will encounter, for all the fun conversations over lunch or coffee, for all the good times outside of work hours. Here I would like to give a special mention to my B1.42 officemates, Johannes Knörzer, Christoph Sünderhauf and Arthur Christianen, who have always made the atmosphere both

productive and fun and have taught me a lot: from Gaussian states to German language, from cooking to literature. A grateful thank you also to Andrea Kluth, Regina Jasny, Elena Wiggert and Sonya Gzyl for the support in administrative tasks and much more.

A heartfelt *grazie* goes to my family: my parents, brothers and grandparents have contributed to making me who I am and given me the confidence and qualities necessary to achieve all that I have achieved, from long before I even knew what a Hamiltonian is.

All my friends from back in Italy continue to have a special role in my life. It is amazing to know that, after all this time and scattered as we are, there will always be friendly faces to go back to and to count on. Thank you Fabbri, Tommy, Mario, Anna, Ginny, Ste and all the others from Unimi for all the time together, be it on a beach in Silvi, eating *parmigiana di melanzane* on a special occasion or just going out for a beer. Thank you Leo and Sasha for all the inspiring, amusing, adventurous and relaxing journeys that we have been on together. Thank you Bica for being there to share experiences and emotional support.

The greatest thank you goes to Marija for her love: *ačiū, kad tu esi.*

References

[1]S. Lloyd, "Universal Quantum Simulators", Science **273**, 1073–1078 (1996).

[2]M. Hilbert and P. López, "The World's Technological Capacity to Store, Communicate, and Compute Information", Science **332**, 60–65 (2011).

[3]S. Lloyd, "Computational Capacity of the Universe", Physical Review Letters **88**, 237901 (2002).

[4]R. P. Feynman, "Simulating physics with computers", International Journal of Theoretical Physics **21**, 467–488 (1982).

[5]J. I. Cirac and P. Zoller, "Goals and opportunities in quantum simulation", Nature Physics **8**, 264–266 (2012).

[6]A. Acín, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, and F. K. Wilhelm, "The quantum technologies roadmap: a European community view", New Journal of Physics **20**, 080201 (2018).

[7]J. Kempe, A. Kitaev, and O. Regev, "The Complexity of the Local Hamiltonian Problem", SIAM Journal on Computing **35**, 1070–1097 (2006).

[8]E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, "Quantum Computation by Adiabatic Evolution", arXiv:quant-ph/0001106, Report No. MIT-CTP-2936 (2000).

[9]E. Farhi, J. Goldstone, and S. Gutmann, "A Quantum Approximate Optimization Algorithm", arXiv:1411.4028, Report No. MIT-CTP-4610 (2014).

[10]A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor", Nature Communications **5**, 4213 (2014).

[11]N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, A. Kandala, A. Mezzacapo, P. Müller, W. Riess, G. Salis, J. Smolin, I. Tavernelli, and K. Temme, "Quantum optimization using variational algorithms on near-term quantum devices", Quantum Science and Technology **3**, 030503 (2018).

[12]J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, "The theory of variational hybrid quantum-classical algorithms", New Journal of Physics **18**, 023023 (2016).

[13]J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, "Architectures for Quantum Simulation Showing a Quantum Speedup", Physical Review X **8**, 021010 (2018).

[14]S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, "Characterizing quantum supremacy in near-term devices", Nature Physics **14**, 595–600 (2018).

[15]F. Arute, K. Arya, R. Babbush, et al., "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505–510 (2019).

[16]I. Georgescu, S. Ashhab, and F. Nori, "Quantum simulation", Reviews of Modern Physics **86**, 153–185 (2014).

[17]F. Franchini, *An Introduction to Integrable Techniques for One-Dimensional Quantum Systems*, Lecture Notes in Physics (Springer International Publishing, Cham, 2017).

[18]M. Schottenloher and U. H. M. Seminar, *A Mathematical Introduction to Conformal Field Theory*, A Mathematical Introduction to Conformal Field Theory: Based on a Series of Lectures Given at the Mathematisches Institut Der Universität Hamburg (Springer, 1997).

[19]F. Assaad and H. Evertz, "World-line and Determinantal Quantum Monte Carlo Methods for Spins, Phonons and Electrons", in *Computational Many-Particle Physics*, edited by H. Fehske, R. Schneider, and A. Weiße (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008), pp. 277–356.

[20]A. W. Sandvik and J. Kurkijärvi, "Quantum Monte Carlo simulation method for spin systems", Physical Review B **43**, 5950–5961 (1991).

[21]M. Troyer and U.-J. Wiese, "Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations", Physical Review Letters **94**, 170201 (2005).

[22]W. Ritz, "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik.", Journal für die reine und angewandte Mathematik **1909**, 1 –61.

[23]G. D. Purvis and R. J. Bartlett, "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples", The Journal of Chemical Physics **76**, 1910–1918 (1982).

[24]R. O. Jones and O. Gunnarsson, "The density functional formalism, its applications and prospects", Reviews of Modern Physics **61**, 689–746 (1989).

[25]P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", Physical Review **136**, B864–B871 (1964).

[26]S. Sorella, "Wave function optimization in the variational Monte Carlo method", Physical Review B **71**, 241103 (2005).

[27]C. Weber, "Variational Study of Strongy Correlated Electronic Models", PhD Thesis (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, Dec. 2007).

[28]G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks", Science **355**, 602–606 (2017).

[29]S. Anders, H. J. Briegel, and W. Dür, "A variational method based on weighted graph states", New Journal of Physics **9**, 361–361 (2007).

[30]U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states", Annals of Physics, January 2011 Special Issue **326**, 96–192 (2011).

[31]F. Verstraete, V. Murg, and J. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems", Advances in Physics **57**, 143–224 (2008).

[32]F. Verstraete and J. I. Cirac, "Matrix product states represent ground states faithfully", Physical Review B **73**, 094423 (2006).

[33]N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, "Entropy Scaling and Simulability by Matrix Product States", Physical Review Letters **100**, 030504 (2008).

[34]J. Eisert, M. Cramer, and M. B. Plenio, "Colloquium: Area laws for the entanglement entropy", Reviews of Modern Physics **82**, 277–306 (2010).

[35]M. B. Hastings, "An area law for one-dimensional quantum systems", Journal of Statistical Mechanics: Theory and Experiment **2007**, 08024–08024 (2007).

[36]S. R. White, "Density matrix formulation for quantum renormalization groups", Physical Review Letters **69**, 2863–2866 (1992).

[37]P. Calabrese and J. Cardy, "Evolution of entanglement entropy in one-dimensional systems", Journal of Statistical Mechanics: Theory and Experiment **2005**, 04010 (2005).

[38]F. Verstraete and J. I. Cirac, "Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions", arXiv:cond-mat/0407066 (2004).

[39]G. Vidal, "Class of Quantum Many-Body States That Can Be Efficiently Simulated", Physical Review Letters **101**, 110501 (2008).

[40]A. Tilloy and J. I. Cirac, "Continuous Tensor Network States for Quantum Fields", Physical Review X **9**, 021040 (2019).

[41]A. Tilloy, "Relativistic continuous matrix product states for quantum fields without cutoff", Physical Review D **104**, 096007 (2021).

[42]G. C. Wick, "The Evaluation of the Collision Matrix", Physical Review **80**, 268–272 (1950).

[43]B. Windt, A. Jahn, J. Eisert, and L. Hackl, "Local optimization on pure Gaussian state manifolds", SciPost Physics **10**, 066 (2021).

[44]C. V. Kraus and J. I. Cirac, "Generalized Hartree–Fock theory for interacting fermions in lattices: numerical methods", New Journal of Physics **12**, 113004 (2010).

[45]J. Bardeen, L. N. Cooper, and J. R. Schrieffer, "Microscopic Theory of Superconductivity", Physical Review **106**, 162–164 (1957).

[46]Z. Gong and T. Guaita, "Topology of Quantum Gaussian States and Operations", arXiv:2106.05044 (2021).

[47]N. N. Bogoliubov, "On the Theory of Superfluidity", Journal of Physics (USSR) **11**, 23–32 (1947).

[48]E. P. Gross, "Structure of a quantized vortex in boson systems", Il Nuovo Cimento **20**, 454–477 (1961).

[49]L. P. Pitaevskii, "Vortex Lines in an Imperfect Bose Gas", JETP **13**, 451 (1961).

[50]C. David Sherrill and H. F. Schaefer, "The Configuration Interaction Method: Advances in Highly Correlated Approaches", in *Advances in Quantum Chemistry*, Vol. 34, edited by P.-O. Löwdin, J. R. Sabin, M. C. Zerner, and E. Brändas (Academic Press, Jan. 1999), pp. 143–269.

[51]M. Werther, S. L. Choudhury, and F. Großmann, "Coherent state based solutions of the time-dependent Schrödinger equation: hierarchy of approximations to the variational principle", International Reviews in Physical Chemistry **40**, 81–125 (2021).

[52]T. Shi, E. Demler, and J. I. Cirac, "Variational study of fermionic and bosonic systems with non-Gaussian states: Theory and applications", Annals of Physics **390**, 245–302 (2018).

[53]Y. Ashida, T. Shi, M. C. Bañuls, J. I. Cirac, and E. Demler, "Solving Quantum Impurity Problems in and out of Equilibrium with the Variational Approach", Physical Review Letters **121**, 026805 (2018).

[54]Y. Wang, I. Esterlis, T. Shi, J. I. Cirac, and E. Demler, "Zero-temperature phases of the two-dimensional Hubbard-Holstein model: A non-Gaussian exact diagonalization study", Physical Review Research **2**, 043258 (2020).

[55]C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information", Reviews of Modern Physics **84**, 621–669 (2012).

[56]A. Ferraro, S. Olivares, and M. G. A. Paris, *Gaussian states in continuous variable quantum information* (Bibliopolis, Napoli, 2005).

[57]X.-B. Wang, T. Hiroshima, A. Tomita, and M. Hayashi, "Quantum information with Gaussian states", Physics Reports **448**, 1–111 (2007).

[58]S. Bravyi, "Lagrangian representation for fermionic linear optics", arXiv:quant-ph/0404180 (2004).

[59]L. Hackl and E. Bianchi, "Bosonic and fermionic Gaussian states from Kähler structures", SciPost Physics Core **4**, 025 (2021).

[60]W.-M. Zhang, D. H. Feng, and R. Gilmore, "Coherent states: theory and some applications", Rev. Mod. Phys. **62**, 867–927 (1990).

[61]J. P. Gazeau, *Coherent States in Quantum Physics*, 1st ed. (John Wiley & Sons, Ltd, 2009).

[62]R. Gilmore, "Geometry of symmetrized states", Annals of Physics **74**, 391–463 (1972).

[63]A. M. Perelomov, "Coherent states for arbitrary Lie group", Communications in Mathematical Physics **26**, 222–236 (1972).

[64]F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, "Atomic coherent states in quantum optics", Phys. Rev. A **6**, 2211–2237 (1972).

[65]C. D. Batista and G. Ortiz, "Algebraic approach to interacting quantum systems", Advances in Physics **53**, 1–82 (2004).

[66]S. M. Davidson and A. Polkovnikov, "SU(3) semiclassical representation of quantum dynamics of interacting spins", Physical Review Letters **114**, 045701 (2015).

[67]E. Bianchi, L. Hackl, and M. Kieburg, "Page curve for fermionic Gaussian states", Physical Review B **103**, L241118 (2021).

[68]M. Kitagawa and M. Ueda, "Squeezed spin states", Physical Review A **47**, 5138–5143 (1993).

[69]L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, and J. I. Cirac, "Geometry of variational methods: dynamics of closed quantum systems", SciPost Physics **9**, 48 (2020).

[70]Y. Ashida, T. Shi, M. C. Bañuls, J. I. Cirac, and E. Demler, "Variational principle for quantum impurity systems in and out of equilibrium: Application to Kondo problems", Physical Review B **98**, 024103 (2018).

[71]Y. Ashida, T. Shi, R. Schmidt, H. Sadeghpour, J. I. Cirac, and E. Demler, "Quantum Rydberg central spin model", Physical review letters **123**, 183001 (2019).

[72]Y. Ashida, T. Shi, R. Schmidt, H. Sadeghpour, J. I. Cirac, and E. Demler, "Efficient variational approach to dynamics of a spatially extended bosonic Kondo model", Physical Review A **100**, 043618 (2019).

[73]P. Sala, T. Shi, S. Kühn, M. C. Bañuls, E. Demler, and J. I. Cirac, "Variational study of U(1) and SU(2) lattice gauge theories with Gaussian states in 1+ 1 dimensions", Physical Review D **98**, 034505 (2018).

[74]G. Fubini, "Sulle metriche definite da una forma hermitiana: nota", Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti **63**, 502–513 (1904).

[75]E. Study, "Kürzeste Wege im komplexen Gebiet", Mathematische Annalen **60**, 321–378 (1905).

[76]J. Haegeman, M. Mariën, T. J. Osborne, and F. Verstraete, "Geometry of matrix product states: metric, parallel transport, and curvature", Journal of Mathematical Physics **55**, 021902 (2014).

[77]R. J. Glauber, "Coherent and incoherent states of the radiation field", Physical Review **131**, 2766 (1963).

[78]D. Perez-García, F. Verstraete, M. M. Wolf, and J. I. Cirac, "Matrix product state representations", Quantum Information and Computation **7**, 401–430 (2007).

[79]G. Vidal, "Entanglement renormalization", Physical review letters **99**, 220405 (2007).

[80]L. Hackl and R. H. Jonsson, "Minimal energy cost of entanglement extraction", Quantum **3**, 165 (2019).

[81]D. Huybrechts, *Complex geometry: an introduction* (Springer Science & Business Media, 2005).

[82]P. Kramer and M. Saraceno, *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics*, Springer-Verlag, Lecture Notes in Physics 140 (Springer-Verlag, Berlin Heidelberg New York, 1981).

[83]A. McLachlan, "A variational solution of the time-dependent Schrödinger equation", Molecular Physics **8**, 39–44 (1964).

[84]P. A. Dirac, "Note on exchange phenomena in the Thomas atom", in Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 26 (Cambridge University Press, 1930), pp. 376–385.

[85]J Frenkel, "Wave mechanics, advanced general theory", Bull. Amer. Math. Soc **41**, 776 (1935).

[86]T. W. Kibble, "Geometrization of quantum mechanics", Communications in Mathematical Physics **65**, 189–201 (1979).

[87]E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations* (Springer Science & Business Media, 2006).

[88]C. Lubich, *From quantum to classical molecular dynamics: reduced models and numerical analysis* (European Mathematical Society, 2008).

[89]R. Martinazzo and I. Burghardt, "Local-in-time error in variational quantum dynamics", Physical Review Letters **124**, 150601 (2020).

[90]J. Broeckhove, L. Lathouwers, E. Kesteloot, and P. Van Leuven, "On the equivalence of time-dependent variational principles", Chemical Physics Letters **149**, 547–550 (1988).

[91]J. Williamson, "On the algebraic problem concerning the normal forms of linear dynamical systems", American journal of mathematics **58**, 141–163 (1936).

[92]J. Haegeman, T. J. Osborne, and F. Verstraete, "Post-matrix product state methods: To tangent space and beyond", Physical Review B **88**, 075133 (2013).

[93]S. T. Smith, "Optimization techniques on Riemannian manifolds", Fields institute communications **3**, 113–135 (1994).

[94]R. Mahony, "Optimization algorithms on homogeneous spaces", PhD thesis (PhD thesis, Australian National University, Canberra, 1994).

[95]A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints", SIAM journal on Matrix Analysis and Applications **20**, 303–353 (1998).

[96]P.-A. Absil, R Mahony, and R. Sepulchre, "Optimization on manifolds: methods and applications", in *Recent advances in optimization and its applications in engineering* (Springer, 2010), pp. 125–144.

[97]K. Huper and J. Trumpf, "Newton-like methods for numerical optimization on manifolds", in Conference record of the thirty-eighth asilomar conference on signals, systems and computers, 2004. Vol. 1 (IEEE, 2004), pp. 136–139.

[98]W. Ring and B. Wirth, "Optimization methods on Riemannian manifolds and their application to shape space", SIAM Journal on Optimization **22**, 596–627 (2012).

[99]C. Qi, K. A. Gallivan, and P.-A. Absil, "Riemannian BFGS algorithm with applications", in *Recent advances in optimization and its applications in engineering* (Springer, 2010), pp. 183–192.

[100]W. Huang, K. A. Gallivan, and P.-A. Absil, "A Broyden class of quasi-newton methods for riemannian optimization", SIAM Journal on Optimization **25**, 1660–1685 (2015).

[101]T. Guaita, L. Hackl, T. Shi, E. Demler, and J. I. Cirac, "Generalization of group-theoretic coherent states for variational calculations", Physical Review Research **3**, 023090 (2021).

[102]R. Gilmore, "On the properties of coherent states", Revista Mexicana de Física **23**, 143–187 (1974).

[103]A. Perelomov, *Generalized coherent states and their applications* (Springer Science & Business Media, 2012).

[104]H. Georgi, *Lie algebras in particle physics: from isospin to unified theories* (CRC Press, 2018).

[105]A. W. Knapp, *Lie groups beyond an introduction* (Birkhäuser, 1996).

[106]M. A. de Gosson, *Symplectic Geometry and Quantum Mechanics* (Birkhäuser Basel, 2006).

[107]A. Fetter and J. Walecka, *Quantum Theory of Many-particle Systems*, International series in pure and applied physics (McGraw-Hill, 1971).

[108]M. Foss-Feig, K. R. A. Hazzard, J. J. Bollinger, and A. M. Rey, "Nonequilibrium dynamics of arbitrary-range Ising models with decoherence: An exact analytic solution", Phys. Rev. A **87**, 042101 (2013).

[109]M. Ringel and V. Gritsev, "Dynamical symmetry approach to path integrals of quantum spin systems", Physical Review A **88**, 062105 (2013).

[110]M. Lakshmanan, "The fascinating world of the Landau–Lifshitz–Gilbert equation: an overview", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **369**, 1280–1300 (2011).

[111]M. Mathur and H. S. Mani, "SU(N) coherent states", Journal of Mathematical Physics **43**, 5351–5364 (2002).

[112]V. Galitski, "Quantum-to-classical correspondence and Hubbard-Stratonovich dynamical systems: a Lie-algebraic approach", Physical Review A **84**, 012118 (2011).

[113]Lang, I.J. and Firsov, Y.A., "Kinetic Theory of Semiconductors with Low Mobility", Journal of Experimental and Theoretical Physics **16**, 1301 (1963).

[114]Y. E. Shchadilova, F. Grusdt, A. N. Rubtsov, and E. Demler, "Polaronic mass renormalization of impurities in Bose-Einstein condensates: correlated Gaussian-wave-function approach", Phys. Rev. A **93**, 043606 (2016).

[115]J. Kondo, "Resistance Minimum in Dilute Magnetic Alloys", Progress of Theoretical Physics **32**, 37–49 (1964).

[116]S. Florens, L. Fritz, and M. Vojta, "Kondo Effect in Bosonic Spin Liquids", Physical Review Letters **96**, 036601 (2006).

[117]L. D. Landau and S. I. Pekar, "Effective Mass of a Polaron", Journal of Experimental and Theoretical Physics **18**, 419 (1948).

[118]H. Fröhlich, "Interaction of electrons with lattice vibrations", Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **215**, 291–298 (1952).

[119]D. M. Kennes, E. Y. Wilner, D. R. Reichman, and A. J. Millis, "Transient superconductivity from electronic squeezing of optically pumped phonons", Nature Physics **13**, 479–483 (2017).

[120]O. Gunnarsson, *Alkali-doped Fullerides: Narrow-band Solids with Unusual Properties* (World Scientific Publishing, Singapore, 2004).

[121]T. D. Lee, F. E. Low, and D. Pines, "The Motion of Slow Electrons in a Polar Crystal", Physical Review **90**, 297–302 (1953).

[122]T. Guaita, L. Hackl, T. Shi, C. Hubig, E. Demler, and J. I. Cirac, "Gaussian time-dependent variational principle for the Bose-Hubbard model", Physical Review B **100**, 094529 (2019).

[123]I. Bloch, J. Dalibard, and W. Zwerger, "Many-body physics with ultracold gases", Rev. Mod. Phys. **80**, 885–964 (2008).

[124]D. Pekker, B. Wunsch, T. Kitagawa, E. Manousakis, A. S. Sørensen, and E. Demler, "Signatures of the superfluid to Mott insulator transition in equilibrium and in dynamical ramps", Phys. Rev. B **86**, 144527 (2012).

[125]S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, UK, 1999).

[126]S. D. Huber, B. Theiler, E. Altman, and G. Blatter, "Amplitude mode in the quantum phase model", Phys. Rev. Lett. **100**, 050404 (2008).

[127]S. D. Huber, E. Altman, H. P. Büchler, and G. Blatter, "Dynamical properties of ultracold bosons in an optical lattice", Phys. Rev. B **75**, 085106 (2007).

[128]U. Bissbort, M. Buchhold, and W. Hofstetter, "Quasi-Particle Theory for the Higgs Amplitude Mode", arXiv:1401.4466 (2014).

[129]K. Sengupta and N. Dupuis, "Mott-insulator–to–superfluid transition in the Bose-Hubbard model: A strong-coupling approach", Phys. Rev. A **71**, 033629 (2005).

[130]M. R. C. Fitzpatrick and M. P. Kennett, "Contour-time approach to the Bose–Hubbard model in the strong coupling regime: Studying two-point spatio-temporal correlations at the Hartree–Fock–Bogoliubov level", Nuclear Physics B **930**, 1 –44 (2018).

[131]M. Knap, E. Arrigoni, and W. von der Linden, "Variational cluster approach for strongly correlated lattice bosons in the superfluid phase", Phys. Rev. B **83**, 134507 (2011).

[132]C. Menotti and N. Trivedi, "Spectral weight redistribution in strongly correlated bosons in optical lattices", Phys. Rev. B **77**, 235120 (2008).

[133]M. Bijlsma and H. T. C. Stoof, "Variational approach to the dilute Bose gas", Phys. Rev. A **55**, 498–512 (1997).

[134]M. Endres, T. Fukuhara, D. Pekker, M. Cheneau, P. Schauß, C. Gross, E. Demler, S. Kuhr, and I. Bloch, "The 'Higgs' amplitude mode at the two-dimensional superfluid/Mott insulator transition", Nature **487**, 454 (2012).

[135]U. Bissbort, S. Götze, Y. Li, J. Heinze, J. S. Krauser, M. Weinberg, C. Becker, K. Sengstock, and W. Hofstetter, "Detecting the Amplitude Mode of Strongly Interacting Lattice Bosons by Bragg Scattering", Phys. Rev. Lett. **106**, 205303 (2011).

[136]C. Schori, T. Stöferle, H. Moritz, M. Köhl, and T. Esslinger, "Excitations of a superfluid in a three-dimensional optical lattice", Phys. Rev. Lett. **93**, 240402 (2004).

[137]S. T. Beliaev, "Energy-Spectrum of a Non-ideal Bose Gas", JETP **7**, 299 (1958).

[138]C. Kollath, A. Iucci, T. Giamarchi, W. Hofstetter, and U. Schollwöck, "Spectroscopy of Ultracold Atoms by Periodic Lattice Modulations", Phys. Rev. Lett. **97**, 050402 (2006).

[139]A. González-Tudela and J. I. Cirac, "Markovian and non-Markovian dynamics of quantum emitters coupled to two-dimensional structured reservoirs", Phys. Rev. A **96**, 043811 (2017).

[140]P. M. Schindler, T. Guaita, T. Shi, E. Demler, and J. I. Cirac, "A Variational Ansatz for the Ground State of the Quantum Sherrington-Kirkpatrick Model", arXiv:2204.02923 (2022).

[141]K. Binder and A. P. Young, "Spin glasses: experimental facts, theoretical concepts, and open questions", Rev. Mod. Phys. **58**, 801–976 (1986).

[142]H. Nishimori, *Statistical physics of spin glasses and information processing: an introduction* (Oxford University Press, Oxford; New York, 2001).

[143]M. Baity Jesi, "An introduction to spin glasses: history, simulations and phase transition", in *Spin glasses: criticality and energy landscapes* (Springer International Publishing, Cham, 2016), pp. 3–42.

[144]M. Mezard and A. Montanari, *Information, physics, and computation* (Oxford University Press, Oxford, 2009).

[145]S. Sachdev, "Spin glasses enter the quantum regime", Physics World **7**, 25 (1994).

[146]G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, "Theory of Quantum Annealing of an Ising Spin Glass", Science **295**, 2427–2430 (2002).

[147]T. Albash and D. A. Lidar, "Adiabatic quantum computation", Rev. Mod. Phys. **90**, 015002 (2018).

[148]C. L. Baldwin and C. R. Laumann, "Quantum algorithm for energy matching in hard optimization problems", Phys. Rev. B **97**, 224201 (2018).

[149]V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, "The quantum adiabatic algorithm applied to random optimization problems: the quantum spin glass perspective", Physics Reports **523**, 127–205 (2013).

[150]E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, "The Quantum Approximate Optimization Algorithm and the Sherrington-Kirkpatrick Model at Infinite size", arXiv:1910.08187 (2019).

[151]D. Sherrington and S. Kirkpatrick, "Solvable Model of a Spin-Glass", Phys. Rev. Lett. **35**, 1792–1796 (1975).

[152]H. Ishii and T. Yamamoto, "Effect of a transverse field on the spin glass freezing in the Sherrington-Kirkpatrick model", Journal of Physics C: Solid State Physics **18**, 6225–6237 (1985).

[153]K. Usadel, "Spin glass transition in an Ising spin system with transverse field", Solid State Communications **58**, 629–630 (1986).

[154]D. Thirumalai, Q. Li, and T. R. Kirkpatrick, "Infinite-range Ising spin glass in a transverse field", Journal of Physics A: Mathematical and General **22**, 3339–3349 (1989).

[155]T. Yamamoto and H. Ishii, "A perturbation expansion for the Sherrington-Kirkpatrick model with a transverse field", Journal of Physics C: Solid State Physics **20**, 6053–6060 (1987).

[156]K. Takahashi, "Quantum fluctuations in the transverse Ising spin glass model: a field theory of random quantum spin systems", Phys. Rev. B **76**, 184422 (2007).

[157]J. V. Alvarez and F. Ritort, "Quantum Monte Carlo study of the infinite-range Ising spin glass in a transverse field", Journal of Physics A: Mathematical and General **29**, 7355–7366 (1996).

[158]P.-Y. Lai and Y. Y Goldschmidt, "Monte Carlo studies of the Ising spin-glass in a transverse field", Europhysics Letters (EPL) **13**, 289–294 (1990).

[159]P. Ray, B. K. Chakrabarti, and A. Chakrabarti, "Sherrington-Kirkpatrick model in a transverse field: absence of replica symmetry breaking due to quantum fluctuations", Phys. Rev. B **39**, 11828–11832 (1989).

[160]B. K. Chakrabarti, A. Dutta, and P. Sen, *Quantum Ising phases and transitions in transverse Ising models* (Springer, Berlin, Heidelberg, 1996).

[161]A. Das and B. K. Chakrabarti, "Reaching the ground state of a quantum spin glass using a zero-temperature quantum Monte Carlo method", Phys. Rev. E **78**, 061121 (2008).

[162]S. Mukherjee, A. Rajak, and B. K. Chakrabarti, "Classical-to-quantum crossover in the critical behavior of the transverse-field Sherrington-Kirkpatrick spin glass model", Phys. Rev. E **92**, 042107 (2015).

[163]Y. W. Koh, "Effects of low-lying excitations on ground-state energy and energy gap of the Sherrington-Kirkpatrick model in a transverse field", Phys. Rev. B **93**, 134202 (2016).

[164]S. Mukherjee, S. Nag, and A. Garg, "Many-body localization-delocalization transition in the quantum Sherrington-Kirkpatrick model", Phys. Rev. B **97**, 144202 (2018).

[165]S. Mukherjee, A. Rajak, and B. K. Chakrabarti, "Possible ergodic-nonergodic regions in the quantum Sherrington-Kirkpatrick spin glass model and quantum annealing", Phys. Rev. E **97**, 022146 (2018).

[166]A. P. Young, "Stability of the quantum Sherrington-Kirkpatrick spin glass model", Physical Review E **96**, 032112 (2017).

[167]S. Pappalardi, A. Polkovnikov, and A. Silva, "Quantum echo dynamics in the Sherrington-Kirkpatrick model", SciPost Phys. **9**, 21 (2020).

[168]G. Parisi, "A sequence of approximated solutions to the S-K model for spin glasses", Journal of Physics A: Mathematical and General **13**, L115–L121 (1980).

[169]D. Panchenko, "The Sherrington-Kirkpatrick model: an overview", Journal of Statistical Physics **149**, 362–383 (2012).

[170]L. Hartmann, J. Calsamiglia, W. Dür, and H. J. Briegel, "Weighted graph states and applications to spin chains, lattices and gases", Journal of Physics B: Atomic, Molecular and Optical Physics **40**, S1–S44 (2007).

[171]S. Anders, M. B. Plenio, W. Dür, F. Verstraete, and H.-J. Briegel, "Ground-State Approximation for Strongly Interacting Spin Systems in Arbitrary Spatial Dimension", Physical Review Letters **97**, 107206 (2006).

[172]M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms", Nature Reviews Physics **3**, 625–644 (2021).

[173]P. M. Schindler, "A Variational Method for the Quantum Sherrington Kirkpatrick Model", PhD thesis (Ludwig–Maximilians–Universität, Munich, Oct. 2021).

[174]C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators", Journal of research of the National Bureau of Standards **45**, 255–282 (1950).

[175]P. Sen, P. Ray, and B. K. Chakrabarti, "Quantum critical behavior of the infinite-range transverse Ising spin glass : an exact numerical diagonalization study", arXiv:cond-mat/9705297 (1997).

[176]H. Leschke, C. Manai, R. Ruder, and S. Warzel, "Existence of replica-symmetry breaking in quantum glasses", Phys. Rev. Lett. **127**, 207204 (2021).

[177]A. Pal and D. A. Huse, "Many-body localization phase transition", Phys. Rev. B **82**, 174411 (2010).

[178]M. Koashi and A. Winter, "Monogamy of quantum entanglement and other correlations", Physical Review A **69**, 022309 (2004).

[179]B. M. Terhal, "Is entanglement monogamous?", IBM Journal of Research and Development **48**, 71–78 (2004).

[180]C. Manai and S. Warzel, "The de Almeida–Thouless line in hierarchical quantum spin glasses", Journal of Statistical Physics **186**, 14 (2021).

[181]A. Ashtekar and A. Magnon, "Quantum fields in curved space-times", Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **346**, 375–394 (1975).

[182]C. V. Kraus and J. I. Cirac, "Generalized Hartree–Fock theory for interacting fermions in lattices: numerical methods", New Journal of Physics **12**, 113004 (2010).