Faculty of Fundamental Problems of Technology
Wrocław University of Science and Technology

Eigenmodes in nearly integrable quantum chains

Jakub Pawłowski
Index number: 250193

Master thesis
under supervision of
prof. dr hab. Marcin Mierzejewski

Wrocław
University
of Science
and Technology

Wrocław 2023

This page is intentionally left blank.

I would like to express my sincere gratitude to prof. dr hab. Marcin Mierzejewski for being my

This page is intentionally left blank.

Abstract

Apparent incompatibility of classical irreversible thermodynamics with

Keywords: *integrals of motion, ETH, integrability breaking, XXZ model*

This page is intentionally left blank.

Contents

This page is intentionally left blank.

1 Introduction

The results concerning spin transport in the long range XXZ have already been published in Mierzejewski et al. [1]. Writing something about computational basis

Write overview of the Krylov chapter, in particulat why we introduce all these things

2

Krylov subspace methods for quantum many-body systems

One of the two purposes of this thesis is to develop and test a set of numerical tools based on the Krylov subspace methods, which is a family of **iterative** methods concerned with projecting high dimensional problems into smaller dimension subspaces and solving them therein. Given a finite dimensional vector space $V \cong \mathbb{C}^m$, a vector $\mathbf{v} \in \mathbb{C}^m$ and a linear operator $A \in \mathbb{C}^{m\times m}$, represented as a matrix, the **k-th Krylov subspace** $\mathcal{K}_k$ is defined as

$$\mathcal{K}_k := \text{span}\{\, \mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \ldots, A^{k-1}\mathbf{v} \,\} \subseteq \mathbb{C}^m \tag{2.1}$$

Maximal dimension of a Krylov subspace is bounded from above by $\text{rank}(A) + 1$[2].

This chapter serves as a pedagogical introduction to the core ideas of these methods, including some of the usually omitted mathematical details. For the initial part of this exposition we follow the excellent textbook of numerical linear algebra by Trefethen and Bau [3], whereas for further applications to quantum many-body physics we rely on the excellent treatments of the topic found in Sandvik [4] and PhD thesis by Crivelli [5].

We start this chapter by quickly sketching the problems with **direct** algorithms such as Exact Diagonalization (ED), and quickly follow with the fundamental iterative algorithm for sparse nonhermitian matrices, the Arnoldi iteration. Its output admits several possible interpretations, however we shall focus on the problem of locating extremal eigenvalues. Afterwards, we restrict our attention to the class of hermitian matrices, to which of course all typical tigh-binding Hamiltonians belong to, and describe the Lanczos algorithm, which allows for efficient calculation of the ground state eigenvalue and eigenvector, and thus the ground state properties of a system. Yet in this work we are mainly interested in infinite temperature calculations, for which in principle sampling of the whole spectrum is required. To this end, in subsequent sections we develop a scheme for time evolution of arbitrary state, called the Krylov propagator[6], and in the last section combine it with the idea of Dynamical Quantum Typicality (DQT), which states that a single pure state can have the same properties as an ensemble density matrix[7–9]. This will produce a numerical algorithm for efficient calculation of time dependent correlation functions without the need for Exact Diagonalization.

2.1 Problems with Exact Diagonalization

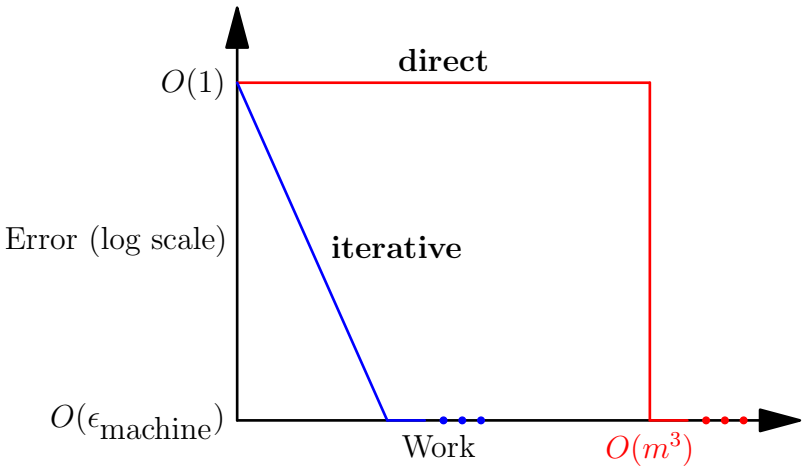

Figure 2.1: Schematic representatation of difference between direct (such as Exact Diagonalization) and iterative (such as Lanczos iteration) algorithms. The advantage of iterative methods comes from the fact that they can be stopped midway, after desired precision is reached. On the other hand, direct algorithms require all $O(m^3)$ operations before any results can be extracted. Figure reproduced from Trefethen and Bau [3].

The most straightforward numerical method for studying discrete quantum many-body systems is without a doubt Exact Diagonalization (ED)[10]. It belongs to the family of the so-called direct algorithms (cf. Fig 2.1) and allows one to obtain numerically exact set of eigenvalues and eigenvectors and subsequently compute any desired properties of the system, be it thermal expectation values, time evolution, Green's functions etc. Unfortunately, the starting point of any ED calculation is the expression of the Hamiltonian as a dense matrix, in the Hilbert space basis of choice. Taking into account the fact that the dimension many-body Hilbert space grows exponentially with the size of the system, the memory cost quickly becomes prohibitive, even when exploiting conservation laws and related symmetries. For example, in the case of a spin chain of length L, with on-site basis dimension being 2, the full dimension of the Hilbert space would be $\mathcal{D} = 2^L$. Taking a modest length of 25 sites, that gives $2^{25} = 33554432 \approx 3.36 \cdot 10^7$ basis states and a memory footprint of Hamiltonian matrix of around 9PB (using double-precision floating point numbers), which is 9000 times more than the typical consumer hard drive capacity of 1TB. Even assuming some kind of distributed memory platform allowing for handling such large matrices, the computational complexity of ED, requiring $O(\mathcal{D}^3)$ operations, is the next major hurdle. Therefore, it is exceedingly difficult to probe the thermodynamic limit physics and ED calculations suffer from finite size effects.

Closer investigation of the Hamiltonian matrix, expressed in computational basis[1] quickly reveals the inefficiency of dense storage. Looking at Figure 2.2, we see that most of the matrix elements are zero. In fact only about $\mu \propto \mathcal{D}$ out of $\mathcal{D}^2$ matrix elements are non-zero. Hence, a numerical scheme leveraging this sparsity is highly desirable. This is exactly what the Krylov subspace algorithms do, by the virtue of requiring only a "black box" computation of

[1] For spin systems, it is the eigenbasis of S^z operator, also called the Ising basis.

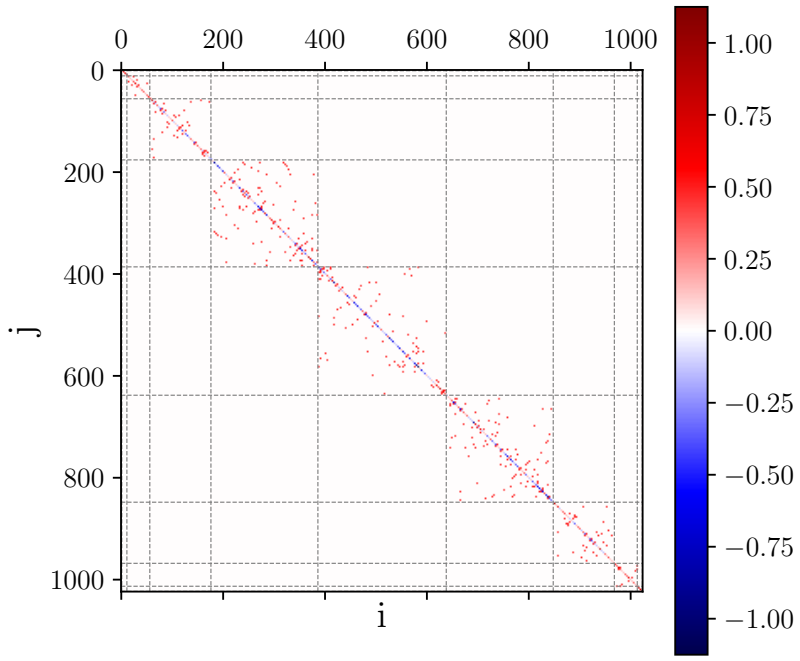

Figure 2.2: Ising basis representation of matrix of the XXZ Hamiltonian $H = J\sum_i \left[\frac{1}{2}\left(S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+\right) + \Delta S_i^z S_{i+1}^z\right]$ on 10 sites, with $J = 1$ and $\Delta = 0.5$. Basis states are sorted according to the magnetization which yields the block structure emphasized by dashed lines. The filling factor μ is approximately 0.005.

matrix-vector product, which can be fairly easily implemented in a way requiring only $O(\mu\mathcal{D})$ operations.

2.2 Calculation of ground state

Our goal in this section is to develop the Lanczos algorithm for ground state search of hermitian matrices, and along the way understand how and why it works.

2.2.1 Arnoldi iteration

The Lanczos algorithm is special case of a more general algorithm, called Arnoldi iteration, designed to transform a general, nonhermitian matrix $A \in \mathbb{C}^{m\times m}$ via a orthogonal [2] similarity transformation to a Hessenberg form $A = QHQ^\dagger$. Such transformation always exist[11]. A square, $m \times m$ matrix H is said to be in **upper Hessenberg form** if $\forall i, j \in \{1, \ldots, n\}$: $i > j + 1 \implies (A)_{i,j} = 0$. It is said to be in **lower Hessenberg form**, if its transpose is in upper Hessenberg form. A Hessenberg matrix differs from a triangular one by one additional super- or subdiagonal. Such form is desirable, because many numerical algorithms in linear algebra experience considerable speedup from leveraging triangular structure of a matrix, and sometimes those benefits carry over to this almost-triangular case. A particularly important strength of the Arnoldi iteration is that it can be interrupted before completion (cf. fig 32.1),

[2] Orthogonal in this context means that $Q^\dagger Q = I_{m\times m}$

thus producing only an approximation of the Hessenberg form in situation where m is so large, that full computations are infeasible (eg. in quantum many-body physics).

Assume now that we are able to only compute the first $n < m$ columns of the equation $AQ = QH$. Let Q_n be the restriction of Q to n columns and let them be denoted by $\mathbf{q_1}, \mathbf{q_2}, \dots \mathbf{q_n} \in \mathbb{C}^m$. Denoting by $\tilde{H}_n$ the $(n+1) \times n$ upper left section of H, which is also a Hessenberg matrix, we can write down the following n-step approximation to the full decomposition

$$AQ_n = Q_{n+1}\tilde{H}_n \tag{2.2}$$

From this equation we can deduce an $n+1$ term recurrence relation for the column $\mathbf{q_{n+1}}$, however it is perhaps best illustrated with a simple example in the first place.

Example 2.1 *Let $A \in \mathbb{C}^{3\times 3}$, $AQ = QH$ be the Hessenberg decomposition and corresponding matrix elements be denoted by lowercase letters. We consider the approximation for $n = 2$, i.e. $AQ_2 = Q_3\tilde{H}_2$. On the right hand side*

$$AQ_2 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \\ q_{31} & q_{32} \end{bmatrix} = \begin{bmatrix} a_{11}q_{11} + a_{12}q_{21} + a_{13}q_{31} & a_{11}q_{12} + a_{12}q_{22} + a_{13}q_{32} \\ a_{21}q_{11} + a_{22}q_{21} + a_{23}q_{31} & a_{21}q_{12} + a_{22}q_{22} + a_{23}q_{32} \\ a_{31}q_{11} + a_{32}q_{21} + a_{33}q_{31} & a_{31}q_{12} + a_{32}q_{22} + a_{33}q_{32} \end{bmatrix}$$
$$= \begin{bmatrix} (A\mathbf{q_1})_1 & (A\mathbf{q_2})_1 \\ (A\mathbf{q_1})_2 & (A\mathbf{q_2})_2 \\ (A\mathbf{q_1})_3 & (A\mathbf{q_2})_3 \end{bmatrix}$$

On the left hand side

$$Q_3H_2 = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ 0 & h_{32} \end{bmatrix} = \begin{bmatrix} q_{11}h_{11} + q_{12}h_{21} & q_{11}h_{12} + q_{12}h_{22} + q_{13}h_{32} \\ q_{21}h_{11} + q_{22}h_{21} & q_{21}h_{12} + q_{22}h_{22} + q_{23}h_{32} \\ q_{31}h_{11} + q_{32}h_{21} & q_{31}h_{12} + q_{32}h_{22} + q_{33}h_{32} \end{bmatrix}$$
$$= \begin{bmatrix} h_{11}(\mathbf{q_1})_1 + h_{21}(\mathbf{q_2})_1 & h_{12}(\mathbf{q_1})_1 + h_{22}(\mathbf{q_2})_1 + h_{32}(\mathbf{q_3})_1 \\ h_{11}(\mathbf{q_1})_2 + h_{21}(\mathbf{q_2})_2 & h_{12}(\mathbf{q_1})_2 + h_{22}(\mathbf{q_2})_2 + h_{32}(\mathbf{q_3})_2 \\ h_{11}(\mathbf{q_1})_3 + h_{21}(\mathbf{q_2})_3 & h_{12}(\mathbf{q_1})_3 + h_{22}(\mathbf{q_2})_3 + h_{32}(\mathbf{q_3})_3 \end{bmatrix}$$

From the above calculation and 2.2 we can read off two identities

$$A\mathbf{q_1} = h_{11}\mathbf{q_1} + h_{21}\mathbf{q_2}$$
$$A\mathbf{q_2} = h_{21}\mathbf{q_1} + h_{22}\mathbf{q_2} + h_{32}\mathbf{q_3}$$

Therefore we get, assuming $\mathbf{q_1}$ is known,

$$\mathbf{q_2} = \frac{A\mathbf{q_1} - h_{11}\mathbf{q_1}}{h_{21}}$$
$$\mathbf{q_3} = \frac{A\mathbf{q_2} - h_{21}\mathbf{q_1} - h_{22}\mathbf{q_2}}{h_{32}}$$

Generalizing the above example, we arrive at the desired $n+1$ term recurrence relation for $\mathbf{q_{n+1}}$

$$\mathbf{q_{n+1}} = \frac{A\mathbf{q_n} - \sum_{m=1}^{n} h_{mn}\mathbf{q_m}}{h_{n+1,n}} \tag{2.3}$$

Algorithm 1 Arnoldi iteration

Input: $\mathbf{v} \in \mathbb{C}^m$, $A \in \mathbb{C}^{m\times m}$, number of steps n

Output: columns of Q_n, matrix elements of H_n

1: $\mathbf{q_1} = \mathbf{v}/\|\mathbf{v}\|$ ▷ components of $\mathbf{v}$ are usually drawn from uniform distribution
2: **for** $i = 1 : n-1$ **do**
3: $\quad \mathbf{q} = A\mathbf{q_i}$
4: $\quad$ **for** $j = 1 : i$ **do**
5: $\quad\quad h_{ji} = \mathrm{cdot}(\mathbf{q_j}, \mathbf{q})$ ▷ cdot is the complex dot product on $\mathbb{C}^m$.
6: $\quad\quad \mathbf{q} = \mathbf{q} - h_{ji}\mathbf{q_j}$ ▷ In exact arithmetic, this enusres orthogonality.
7: $\quad$ **end for**
8: $\quad h_{i+1,i} = \|\mathbf{q}\|$
9: $\quad \mathbf{q_{i+1}} = \mathbf{q}/h_{i+1,i}$
10: **end for**

We can now easily cast the above recurrence into a pseducode algorithm:

Step 9 of the Algorithm 1 may be questionable, as we are dividing by a norm of a vector, which after all can be equal to zero. However, in practical applications of Arnoldi iteration it usually means that our calculations have converged and the iterations may be stopped.

Examining closely the Arnoldi iteration algorithm, we notice that it is essentially the Gram-Schmidt procedure applied to the vectors $\{\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}\}$ and hence the vectors $\{\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_n}\}$ form an orthonormal basis of the Krylov subspace $\mathcal{K}_n$. The orthonormality condition is concisely expressed by the fact that $Q_n^\dagger Q_{n+1}$ is the $n \times (n+1)$ identity matrix. Multiplying the left-hand side of equation (2.2) by $Q_n^\dagger$ we get

$$Q_n^\dagger A Q_n = \underbrace{Q_n^\dagger Q_{n+1}}_{\mathrm{Id}_{n\times(n+1)}} \tilde{H}_n = H_n \in \mathbb{C}^{n\times n} \tag{2.4}$$

where H_n is the Hessenberg matrix $\tilde{H}_n$ with its last row removed.

To understand the meaning of matrix H_n from the point of view of linear algebra, consider the following reasoning. Imagine we are given an endormophism of the space $\mathbb{C}^m$, represented in the standard basis by a matrix A. We would like to restrict it to a endormophism of the Krylov subspace $\mathcal{K}_n$, $n < m$. Of course, as $\mathbf{q} \in \mathcal{K}_n \implies \mathbf{q} \in \mathbb{C}^m$, we can calculate the action of A on a vector from Krylov subspace in a straightforward way. However, the resulting vector $A\mathbf{q}$ in not guaranteed to be an element of $\mathcal{K}_n$. We need to orthogonally project it back to the subspace. Such projection is realized by $Q_n Q_n^\dagger \in \mathbb{C}^{m\times m}$ and hence, with respect to the standard basis on $\mathbb{C}^m$, the desired restriction can be written as $Q_n Q_n^\dagger A$. Transforming it to the basis given by columns of Q_n we get $Q_n^{-1}\left(Q_n Q_n^\dagger A\right) Q_n = Q_n^\dagger A Q_n$. Thus, matrix H_n is the orthogonal projection of A to the subspace $\mathcal{K}_n$, represented in the basis $\{\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_n}\}$.

H_n is once again a square matrix, so we can talk about its eigenvalues $\{\theta_i\}_{i=1}^n$ in the usual fashion. These numbers are called the *Arnoldi eigenvalues estimates at step n*, or the *Ritz values with respect to* $\mathcal{K}_n$. Given the interpretation above, we may suspect that they would be related to the eigenvalues of the original matrix A. Indeed, as we shall see in a moment, some of the Ritz values are extremally good approximations of some of the original eigenvalues.

2.2.2 Polynomial approximation and eigenvalues

By carrying out the Arnoldi iterations for succesive steps, and at each step n (or at just some of the steps) calculating the eigenvalues of the Hessenberg matrix H_n, we are left with sequences of Ritz values. Some of them often converge rapidly to, what we reasonably assume, eigenvalues of the original matrix A. However in practive, the maximal accessible n is much smaller than m, so we cannot expect to find all eigenvalues. As it turns out, Arnoldi iteration typically finds extremal eigenvalues, which fortunately are those that we are interested in.

Before we will understand the details, let us introduce different, seemingly unrelated problem of *polynomial approximation*. We can take any $\mathbf{q} \in \mathcal{K}_n$ and using the defining basis of Krylov subspace $\mathcal{K}_n$(Definition 2.1), expand is as

$$\begin{aligned}\mathbf{q} &= a_0\mathbf{v} + a_1 A\mathbf{v} + a_2 A^2\mathbf{v} + \ldots + a_{n-1}A^{n-1}\mathbf{v} \\ &= \left(a_0\mathbb{1} + a_1 A + a_2 A^2 + \ldots + a_{n-1}A^{n-1}\right)\mathbf{v}\end{aligned}$$

Utilizing the special structure of vectors from $\mathcal{K}_n$, we can define a polynomial $p(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_{n-1}z^{n-1}$, and concisely write our vector as $\mathbf{q} = p(A)\mathbf{v}$. As the vector $\mathbf{q}$ was arbitrary, we have estabilished an isomorphism between the n-th Krylov subspace and the space of complex polynomials of maximal degree $n-1$. We are now ready to state the problem:

Arnoldi Approximation Problem

Given a matrix $A \in \mathbb{C}^{m\times m}$ and a vector $\mathbf{v} \in \mathbb{C}^m$, find
$p \in P^n := \{\, a_0 + a_1 z + \ldots + a_{n-1}z^{n-1} + z^n \mid a_0, a_1, \ldots a_{n-1} \in \mathbb{C}\,\}$
such that $\|p(A)\mathbf{v}\|_2$ is minimized.

Remarkably, the Arnoldi approximation is the exact solution to this problem. This fact is interesting enough that we state it here as a theorem and, following Trefethen and Bau [3], provide a complete proof.

Theorem 2.1 *If $dim(\mathcal{K}_n) = n$, i.e. matrix having columns $\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}$ is of rank n, then the Arnoldi Approximation Problem has a unique solution $p^n \in P^n$, given by the characteristic polynomial of the matrix H_n, defined by (2.4).*

Proof. We start with an observation, that given a polynomial $p \in P^n$, the vector $p(A)\mathbf{v}$ can be written as $p(A)\mathbf{v} = A^n\mathbf{v} - Q_n\mathbf{r}$ for some $\mathbf{r} \in \mathbb{C}^n$. To see that, note that $(A^n\mathbf{v} - p(A)\mathbf{v}) \in \mathcal{K}_n$ and columns of Q_n form an orthonormal basis of $\mathcal{K}_n$. Now, we can recast our problem into a slighlty different language, namely finding a vector in $\mathcal{K}_n$ that is the closest in the sense of L_2 norm to $A^n\mathbf{v}$. In short:

$$\mathbf{r}^* = \min_{\mathbf{r}\in\mathbb{C}^n} \|A^n\mathbf{v} - Q_n\mathbf{r}\|$$

To achieve that, we need to have $p(A)\mathbf{v} \perp \mathcal{K}_n$, that is $p(A)\mathbf{v}$ must be orthogonal to all basis vectors spanning $\mathcal{K}_n$. This is consisely expressed as $Q_n^\dagger p(A)\mathbf{v} = \mathbf{0} \in \mathbb{C}^n$.

Now, we know that the Hessenberg factorization $A = QHQ^\dagger$ exists, and is approximated by n steps of the Arnoldi iteration. Thus, the matrices Q and H can have the following block

structure:

$$Q = \begin{bmatrix} Q_n & V \end{bmatrix}, \quad H = \begin{bmatrix} H_n & 0_{n\times(m-n)} \\ Y & 0_{(m-n)\times(m-n)} \end{bmatrix} \tag{2.5}$$

where $V \in \mathbb{C}^{m\times(m-n)}$ is a matrix with orthonormal columns, which are also orthogonal to columns of Q_n, and matrix $Y \in \mathbb{C}^{(m-n)\times n}$ has only the upper-right entry different from zero (the one from the last row of $\tilde{H}_n$). Using the Hessenberg factorization we can write our condition as $Q_n^\dagger Q p(H) Q^\dagger \mathbf{v} = \mathbf{0}$, and because equation (2.5) introduces partitions into *conformable* blocks, we can use the rules of block-matrix algebra to simiplify it further[12].

First, let us investigare closely the structure of $p(H)$. We observe that

$$\begin{aligned}
H^2 &= \begin{bmatrix} H_n & 0_{n\times(m-n)} \\ Y & 0_{(m-n)\times(m-n)} \end{bmatrix}^2 = \begin{bmatrix} H_n^2 & 0_{n\times(m-n)} \\ YH_n & 0_{(m-n)\times(m-n)} \end{bmatrix} \\
H^3 &= \begin{bmatrix} H_n & 0_{n\times(m-n)} \\ Y & 0_{(m-n)\times(m-n)} \end{bmatrix}^3 = \begin{bmatrix} H_n^3 & 0_{n\times(m-n)} \\ YH_n^2 & 0_{(m-n)\times(m-n)} \end{bmatrix} \\
&\dots \\
H^n &= \begin{bmatrix} H_n & 0_{n\times(m-n)} \\ Y & 0_{(m-n)\times(m-n)} \end{bmatrix}^n = \begin{bmatrix} H_n^n & 0_{n\times(m-n)} \\ YH_n^{n-1} & 0_{(m-n)\times(m-n)} \end{bmatrix}
\end{aligned}$$

Thus $p(H)$ can be written as

$$\begin{aligned}
p(H) &= a_0\mathbb{1} + a_1 H + a_2 H^2 + \ldots + a_{n-1}H^{n-1} + H^n \\
&= \begin{bmatrix} a_0\mathbb{1} + a_1 H_n + a_2 H_n^2 + \ldots + a_{n-1}H_n^{n-1} + H_n^n & 0_{n\times(m-n)} \\ a_0\mathbb{1} + a_1 Y + a_2 Y H_n + \ldots + a_{n-1} Y H_n^{n-2} + Y H_n^{n-1} & 0_{(m-n)\times(m-n)} \end{bmatrix} \\
&= \begin{bmatrix} p(H_n) & 0 \\ \tilde{Y} & 0 \end{bmatrix}
\end{aligned}$$

We have now all the pieces to simplify the orthogonality condition:

$$\begin{aligned}
\mathbf{0} &= Q_n^\dagger Q p(H) Q^\dagger \mathbf{v} \\
&= \begin{bmatrix} Q_n^\dagger \end{bmatrix} \begin{bmatrix} Q_n & V \end{bmatrix} \begin{bmatrix} p(H_n) & 0_{n\times(m-n)} \\ \tilde{Y} & 0_{(m-n)\times(m-n)} \end{bmatrix} \begin{bmatrix} Q_n^\dagger \\ U^\dagger \end{bmatrix} \mathbf{v} \\
&= \begin{bmatrix} \mathbb{1}_{n\times n} & 0_{n\times(m-n)} \end{bmatrix} \begin{bmatrix} p(H_n) & 0_{n\times(m-n)} \\ \tilde{Y} & 0_{(m-n)\times(m-n)} \end{bmatrix} \begin{bmatrix} Q_n^\dagger \\ U^\dagger \end{bmatrix} \mathbf{v} \\
&= \begin{bmatrix} p(H_n) & 0_{n\times(m-n)} \end{bmatrix} \begin{bmatrix} Q_n^\dagger \\ U^\dagger \end{bmatrix} \mathbf{v} \\
&= p(H_n) Q_n^\dagger \mathbf{v}
\end{aligned}$$

As a final step, notice that by construction the first row of $Q_n^\dagger$ is $\mathbf{v}/\|\mathbf{v}\|$, and all the remaining rows are orthogonal to $\mathbf{v}$, therefore only the first column of H_n, or the first n elements of the first column of H are required to be 0. By Caylel-Hamilton theorem, this is guaranteed if we take $p = p^n$, where p^n is the characteristic polynomial of H_n. For the uniqueness part, suppose that there exists another polynomial, say q^n such that $q^n \perp \mathcal{K}_n$. But then $p^n - q^n$ is

a nonzero polynomial of degree $n-1$ (because $p^n,\ q^n$ are monic) such that $(p^n - q^n)(A)\mathbf{v} = \mathbf{0}$, and hence vectors $\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}$ are linearly dependent, which violates assumption that $\dim(\mathcal{K}_n) = n$.

■

This theorem allows us to interpret the Arnoldi eigenvalues estimates $\{\theta_i\}$ as the roots of the optimal polynomial. Following the above proof, it is relatively easy to see that they are scale invariant, i.e. if $A \to \alpha A$ for some $\alpha \in \mathbb{C}$, then $\{\theta_i\}_{i=1}^n \to \{\alpha\theta_i\}_{i=1}^n$ and invariant under unitary transformations, i.e. if $A \to UAU^\dagger$ and $\mathbf{v} \to U\mathbf{v}$ for some unitary U, then the Arnoldi estimates are unchanged. Furthermore, owing to the properties of monic polynomials, they are also translationally invariant, namely if $A \to A + \alpha\mathbb{1}$ for some $\alpha \in \mathbb{C}$, then $\{\theta_i\}_{i=1}^n \to \{\theta_i + \alpha\}_{i=1}^n$.

In the end we see that the direct purpose of Arnoldi iteration is to solve a polynomial approximation problem and not to find eigenvalues. However, those two problems have enough in common, that the Arnoldi iteration produces some correct eigenvalues as a 'by-product'. We can reason along the following lines. If our task is to find a polynomial $p \in P^n$ minimizing $\|p(A)\|$, it may be a good idea to select a polynomial that has roots close to the eigenvalues of A. In an extreme situation, when there exists a diagonalization of A and it posses only $n \ll m$ distinct eigenvalues, the minimal polynomial[3] will coincide with the characteristic polynomial computed via Arnoldi iteration after n steps and the Arnoldi eigenvalue approximations will be exact, provided we start from $\mathbf{v}$ having nonzero overlap with all eigenvectors of A. In most practical situations however, the agreement is only approximate, namely Arnoldi eigenvalues are close to real eigenvalues, and computed polynomial is such that $\|p(A)\|$ is small.

There is more to this story than we have told here, particularly a nice geometric interpreation of Algorithm 1 via *Arnoldi lemniscates*, which illustrates why extremal eigenvalues are found first, however we shall not concern ourselves with those matters any further. Interested readers are once again referred to Trefethen and Bau [3], whereas we turn our attention to the case of utmost interest in quantum mechanics, namely Arnoldi iteration for hermitian matrices.

2.2.3 Restriction to hermitian matrices: Lanczos iteration

After the mathematical detour of previous section, armed with deeper understanding of Krylov subspace and Arnoldi iteration, we are now going to investigate the algorithms that are of direct relevance to condensed matter physics, starting with Lanczos iteration. From this point onwards, we shall switch to the favored by physicsts Dirac bra-ket notation, namely $|v\rangle \equiv \mathbf{v}$ and $\langle v| = \mathbf{v}^\dagger$. Moreover, we assume the matrix A to be hermitian, as in most use cases it will be the Hamiltonian of our system.

It immediately follows from (2.4) that, given A is hermitian, the Hessenberg matrix H_n will also be hermitian. But a matrix that is both Hessenberg and hermitian, must of course be tridiagonal! Indeed, to see this directly, let us write the equation for matrix elements $(H_n)_{ij}$ of H_n:

$$(H_n)_{ij} = \sum_{r,s} (Q_n^\dagger)_{ir} (A)_{rs} (Q_n)_{sj} = \langle q_i | A | q_j \rangle \tag{2.6}$$

[3] A minimal polynomial of matrix A is a polynomial p of the smallest degree such that $p(A) = 0$. It always divides the characteristic polynomial.

where $|q_i\rangle, |q_j\rangle$ are respectively i-th and j-th columns of matrix Q_n. From the recurrence relation (2.3) we know that $A|q_j\rangle \in \text{span}\{|q_1\rangle, \ldots, |q_{j+1}\rangle\}$, and that it is orthogonal to all $|q_i\rangle$ with $i > j+1$. Therefore $(H_n)_{ij} = 0$ for $i > j+1$. Simiarly, by taking the hermitian conjugate of equation (2.6), we get

$$(H_n)_{ij} = \langle q_j|A^\dagger|q_i\rangle \triangleq \langle q_j|A|q_i\rangle \tag{2.7}$$

where $\triangleq$ follows from assumed hermiticity of A. Repeating the above reasoning we quickly obtain that $(H_n)_{ij} = 0$ also for $j > i+1$ and hence the matrix is tridiagonal. In literature the diagonal is usually denoted by $\alpha_i \equiv (H_n)_{ii}$, whereas the sub- and superdiagonal are denoted by $\beta_i \equiv (H_n)_{i,i+1} = (H_n)_{i+1,i}$. The relation for $|q_{n+1}\rangle$ becomes a 3-step recurrence:

$$|q_{n+1}\rangle = \frac{A|q_n\rangle - \beta_{n-1}|q_{n-1}\rangle - \alpha_n|q_n\rangle}{\beta_n} \tag{2.8}$$

This has a tremendous impact on the practical applications of the algorithm, as both computational and memory costs decrease significantly. We are now ready to state the simplified version of the Algorithm 1. Another important observation is that α_n's are diagonal elements

Algorithm 2 Lanczos iteration

Input: $|v\rangle \in \mathbb{C}^m$, $A \in \mathbb{C}^{m\times m}$ such that $A^\dagger = A$, number of steps n

Output: columns of Q_n, tridiagonal matrix H_n

$\beta_0 = 0$
$|q_0\rangle = \mathbf{0} \in \mathbb{C}^m$
$|q_1\rangle = |v\rangle / \||v\rangle\|$
for $i = 1 : n-1$ **do**
    $|q\rangle = A|q_i\rangle$
    $\alpha_i = \langle q_i|q\rangle$
    $|q\rangle = |q\rangle - \beta_{i-1}|q_{i-1}\rangle - \alpha_i|q_i\rangle$
    $\beta_i = \||q\rangle\|$
    $|q_{i+1}\rangle = |q\rangle / \beta_i$
end for

of a hermitian matrix, and β_n's are norms of vector $|q\rangle$ in subsequent interations, both of which are real. Therefore, even if our Hamiltonian is complex, the numbers α_n and β_n can be stored as vectors of real floating point numbers, decreasing memory requirements even further. This is the first algorithm implemented for the purpose of this thesis, using the Armadillo linear algebra library[13] and Intel MKL.

Matrix Q_n is of dimension $m \times n$, so keeping it in full in the memory can still be costly. Fortunately, at each step of the Lanczos iteration no more than three vectors are necessary $(|q\rangle, |q_i\rangle, |q_{i-1}\rangle)$ so the storage of full matrix Q_n is reduntant. Extremal eigenvalues are then obtained by explicit diagonalization of the constructed matrix $H_n = V_n D_n (V_n)^\dagger$, which can be done efficiently using specialized routines for tridiagonal matrices. However, this approach has it drawbacks when we are interested also in the ground state eigenvector, which will be the case in further applications.

It turns out that the Lanczos iteration can approximate not only eigenvalues, but also corresponding eigenvectors. They are the eigenvectors of the tridiagonal matrix H_n, transformed back to the original Hilbert space. Given the full Hessenberg decomposition we would have

$$A = QHQ^\dagger = Q(VDV^\dagger)Q^\dagger = (QV)D(QV)^\dagger \tag{2.9}$$

Restriction to n-step iteration produces an approximation $A \approx (Q_nV_n)D_n(Q_nV_n)^\dagger$. The simplest form of Lanczos iteration presented in Algorithm 2 is sufficient to obtain only the ground state eigenvector with machine precision, because eigenvectors of excited states are plagued by loss of orthogonality stemming from the nature of floating point numbers. We shall have a brief look at this problem at the end of this section.

The ground state vector $|\psi_0\rangle$ can be then read of as the first column Q_nV_n. However, there is a problem. To conserve memory, we have not constructred the whole matrix Q_n explicitly, but only three of its columns at a given time and hence do not have access to the matrix product Q_nV_n. We need a second pass of the Lanczos iteration, with a single line added for iteretive calculation of the first column. It can be summarized by the following piece of pseducode:

Algorithm 3 Second pass of Lanczos iteration, for calculating ground state eigenvector

Input: $|\psi_0\rangle = \mathbf{0} \in \mathbb{C}^m$, matrix V_n from Alg. 2, rest of input data from Alg. 2

Output: columns of Q_n, tridiagonal matrix H_n

$\beta_0 = 0$
$|q_0\rangle = \mathbf{0} \in \mathbb{C}^m$
$|q_1\rangle = |v\rangle / \||v\rangle\|$
for $i = 1 : n-1$ **do**
    $|\psi_0\rangle = |\psi_0\rangle + (V_n)_{i,1} |q_i\rangle$ ▷ this is the only difference from Alg. 2
    $|q\rangle = A |q_i\rangle$
    $\alpha_i = \langle q_i|q\rangle$
    $|q\rangle = |q\rangle - \beta_{i-1} |q_{i-1}\rangle - \alpha_i |q_i\rangle$
    $\beta_i = \||q\rangle\|$
    $|q_{i+1}\rangle = |q\rangle / \beta_i$
end for

To finish this section, let us discuss quickly the convergence properties of Lanczos iteration. We have one free parameter, namely the number of iterations n. If we had carried out the full Arnoldi iteration, as described in 1, the orthogonality of subsequent columns of matrix Q_n would be guaranteed by the explicit Gram-Schmidt procedure and we in principle could continue it up to $n = m$ obtaining the full Hessenberg decomposition. However, restricting ourselves to a three step recurrence in Lanczos iteration we rely on mathematical identities to force the orthogonality of $|q_i\rangle$ with all previous vectors. Those are valid in exact arithmetic, but can quickly break down when using floating point numebers, as it is done in practice. Therefore, the iteration is unstable and should be stopped as soon as desired accuracy is reached. Taking E_n^1 to be the lowest eigenvalues of H_n, the convergence criterion can be defined as $|E_{n+1}^1 - E_n^1|/|E_n^1| < \varepsilon$ for some small ε, e.g. 10^{-14}. As long as it is nondegenerate, the convergence usually happens quite quickly for both lowest eigenvalue and corresponding

eigenvector. To reliably obtain higher eigenstates one needs to perform reorthogonalization, but it requires keeping the matrix Q_n in memory which can be very costly.

This instability can manifest itself in an interesting way, namely in form of additional eigenvalues, called Lanczos 'ghosts'. They are spurious copies of already found eigenvalues, which start apperating after too many iteration steps (cf. fig. ...) plot with Lanczos ghosts They are difficult to understand rigorously, however Trefethen and Bau [3] offer a nice heuristic explanatation. Reaching convergence of some Arnoldi approximated eigenvalue to the true eigenvalue of A causes annihilation of corresponding eigenvector component in the vector $|q\rangle$. However, rounding errors originating from floating point arithmetic cause $|q\rangle$ to again develop component in the direction of the eigenvector and thus, after certain number of iterations, appearance of another Arnoldi approximated eigenvalue is necessary to annihilate it. This can go on and on producing more and more Lanczos 'ghosts'. Fortunately, for all further applications in this thesis we will only require the lowest eigenstate, so we can avoid most of the instability problems by stopping the iteration sufficiently quick.

2.3 Time evolution via the Krylov propagator

Even though the original puprose of Lanczos iteration was to approximate boundaries of the spectrum of matrices and solve systems of linear equation[14,15], it can also be employed to evaluate functions of hermitian matrices. Given any unitary (or orthogonal) decomposition of matrix $A = QHQ^\dagger$, and an analytic function f[4], we have $f(A) = Qf(H)Q^\dagger$. And indeed, Lanczos iteration gives us such decomposition, albeit an approximate one $A \approx Q_n H_n Q_n^\dagger$. Because of this approximate character and convergence probles discussed at the end of previous section, global approximation of $f(A)$ via Lanczos iteration is a hopeless endevour. However, the goal of this section is to calculate time evolution, which boils down to evaluation how the time-evolution operator $\exp\left(-i\hat{H}t\right)$ acts on some state $|v\rangle$ in the Hilbert space for a given Hamiltonian $\hat{H}$. It is thus enough to restrict our attention the problem of calculating $f(A)|v\rangle$, for which the Lanczos iteration turns out to be an excellent tool.

Let us now assume that the matrix A is our Hamiltonian $\hat{H}$, acting on Hilbert space $\mathcal{H} \cong \mathbb{C}^m$ and $|v\rangle \in \mathcal{H}$ is some fixed state. Using the approximate factorization derived from (2.9), we get

$$f(\hat{H})\,|v\rangle \approx (Q_n V_n) f(D_n)(Q_n V_n)^\dagger\,|v\rangle \tag{2.10}$$

where D_n is a real diagonal matrix, so $f(D_n)$ is easy to compute. Previously, we have started the construction of Krylov subspace basis from some random vector. Now, let us change this slighlty and start from the vector $|v\rangle$ instead, which we assume to be normalized. Then, the first column of Q_n will be $|v\rangle$, whereas all subsequent columns will be orthogonal (up to some numerical errors), yielding $Q_n^\dagger\,|v\rangle = |e_1\rangle$, which is the first vector in canonical basis of $\mathbb{C}^n$, i.e. $|e_1\rangle = [1, 0, 0, \ldots, 0]$. Equation (2.10) then simplifies to

$$f(\hat{H})\,|v\rangle \approx Q_n(V_n f(D) V_n^\dagger)Q_n^\dagger\,|v\rangle = Q_n(V_n f(D) V_n^\dagger)\,|e_1\rangle = Q_n f(H_n)\,|e_1\rangle \tag{2.11}$$

[4] A function f is real (complex) analytic if and only if its Taylor series about x_0 converges in some neighborhood of x_0 pointwise to the function, for every x_0 in the domain.

Moreover, $f(H_n)\,|e_1\rangle$ is just the first column of $f(H_n)$, because $(f(H_n)\,|e_1\rangle)_i = \sum_j (f(H_n))_{i,j}\,\delta_{1,j} = (f(H_n))_{i,1}$. So we do not need to compute the full matrix, but only a single vector of the form

$$(f(H_n))_{i,1} = \sum_j (V_n)_{i,j}\, f(D_n)_j \left(V_n^\dagger\right)_{j,1} = \sum_j (V_n)_{i,j}\, (V_n^*)_{1,j}\, f(D_n)_j \tag{2.12}$$

where the diagonal matrix $f(D_n)$ is treated as a vector. Assuming that $f(D_n)$ is already computed, it boils down to a single dot product for each element of the column.

We are now ready to apply this procedure to the problem of interest, namely the time-evolution of a pure state $|\psi(t)\rangle$ in the interval $(t, \Delta t)$. It is described by the Schrödineger equation $i\partial_t\,|\psi(t)\rangle = \hat{H}\,|\psi(t)\rangle$, having a formal solution

$$|\psi(t+\Delta t)\rangle = \exp\left(-i\hat{H}\Delta T\right)|\psi(t)\rangle\,, \tag{2.13}$$

under the assuption that $\hat{H}$ does not depend explicitly on time. Exact exponentiation of matrix amounts to exact diagonalization which, as mentioned at the beginning of this chapter, is a difficult task. However, by letting $f(\hat{H}) = \exp\left(-i\hat{H}\Delta t\right)$ and $|v\rangle = |\psi(t)\rangle$ in equation (2.11) we get an approximation of the action of the time evolution operator on a state, called the **Krylov based propagator**[16], which was originally proposed in the 1980's by Park and Light [6] and has been since used with great success to investigate many different systems[17–20]. The final equation reads

$$|\psi(t+\Delta t)\rangle = \exp\left(-i\hat{H}\Delta T\right)|\psi(t)\rangle \approx Q_n \exp\left(-iH_n\Delta t\right) Q_n^\dagger\,|\psi(t)\rangle \tag{2.14}$$

$$= Q_n \exp\left(-iH_n\Delta t\right)|e_1\rangle = \sum_{j=1}^{n} \left(f(H_n)\right)_{j,1}|q_j\rangle \tag{2.15}$$

where $|q_j\rangle$ are columns of the unitary (orthogonal) matrix Q_n. Because of this orthogonality, the approximation error is bouned from above by the last coefficient $(f(H_n))_{n,1}$, i.e.

$$\left\| |\psi(t+\Delta t)\rangle_{\text{exact}} - |\psi(t+\Delta t)\rangle_{\text{Krylov}} \right\| \leq \left| (f(H_n))_{n,1} \right| \tag{2.16}$$

assuming that $\||\psi(t)\rangle\| = 1$. There also exists an estimate of the sufficient dimension of Krylov subspace[21]

$$n \lessapprox 1.5\rho_{\hat{H}}\Delta t > 10 \tag{2.17}$$

where $\rho_{\hat{H}}$ is the spectral radius of Hamiltonian and can be calculated as the difference between highest and lowest eigenvalue. It is now very simple to cast these expressions into a concrete algorithm

Algorithm 4 Krylov propagator

Input: input data from Algorithm 2, with $|v\rangle = |\psi(t)\rangle$, time step Δt

Output: propagated state $|\psi(t+\Delta t)\rangle$

1: Run Alg. 2, obtaining eigenvalues D_i
2: Calculate $(f(H_n))_{i,1}$ using eq. (2.12)
3: Run modified Alg. 3, with $(f(H_n))_{i,1}$ instead of $(V_n)_{i,1}$

We finish this section with a remark, that real-time propagator is not the only function of the Hamiltonian which can be calculated using Lanczos techniques. Fairly often encounterd

are also the imaginary-time propagator $\exp(-\beta H)$ and pure state Green's function of some obervable Q: $\langle\psi_{\epsilon_0}| Q\left(\omega + i\eta + \epsilon_0 + H\right)^{-1} Q\,|\psi_{\epsilon_0}\rangle$[22]. However, we will not need them for the purpose of this thesis.

2.4 Correlation functions and Quantum Typicality

We are now ready to introduce **time dependent correlation functions** and a method of calculating them using already developed Krylov subspace machinery. This function is defined for a pair of operators A, B as

$$\tilde{C}_{AB}(t) \equiv \mathrm{Re}\langle A(t)B\rangle = \mathrm{Re}\,\mathrm{Tr}\left(\hat{\rho}A(t)B\right) \tag{2.18}$$

where $\hat{\rho} = e^{-\beta H}/\mathcal{Z}$, $\mathcal{Z} = \mathrm{Tr}\left(e^{-\beta H}\right)$ is the canonical ensemble at temperature $T = 1/\beta$, and $A(t) = e^{i\hat{H}t}Ae^{-i\hat{H}t}$ is understood via the Heisenberg picture. In this thesis we are interested only in the infinite temperature properties, so $\hat{\rho} \to \frac{1}{\mathcal{D}}\mathbb{1}$, where $\mathcal{D}$ is the dimension of Hilbert space, however let us keep the temperature finite for a while, so we are able to see the results of subsequent developments in full. Correlation functions allow us to probe the complex dynamics of interacting many-body systems and within Linear Response Theory are directly related to the transport properties[23], so they have been an object of intense study[24–30]. We can use a complete set of eigensetates $|n\rangle$ of the Hamiltonian $\hat{H}$ in order to transform the correlation function into the so called spectral representation

$$\begin{aligned}
\langle A(t)B\rangle &= \frac{1}{\mathcal{Z}}\mathrm{Tr}\left(A(t)Be^{-\beta H}\right) = \frac{1}{\mathcal{Z}}\mathrm{Tr}\left(e^{i\hat{H}t}Ae^{-i\hat{H}t}Be^{-\beta H}\right) \\
&= \frac{1}{\mathcal{Z}}\sum_k\sum_{m,n}\langle k|\,|n\rangle\langle n|\,e^{i\hat{H}t}Ae^{-i\hat{H}t}\,|m\rangle\langle m|\,Be^{-\beta H}\,|k\rangle \\
&= \frac{1}{\mathcal{Z}}\sum_k\sum_{m,n}\delta_{k,n}e^{i(\varepsilon_n-\varepsilon_m)t}e^{-\beta\varepsilon_k}\,\langle n|A|m\rangle\;\langle m|B|k\rangle \\
&= \frac{1}{\mathcal{Z}}\sum_{m,n}e^{i(\varepsilon_n-\varepsilon_m)t}e^{-\beta\varepsilon_n}\,\langle m|A|n\rangle^*\;\langle m|B|n\rangle
\end{aligned} \tag{2.19}$$

which is useful for the calculations using Exact Diagonalization. Unfortunately, we have already estabilished that ED calculations suffer greatly from the exponential growth of the Hilbert space. We would like to have a more efficient method for calculating $C_{AB}(t)$, capable of accessing larger systems sizes, far beyond the reach of ED. The question is, how the Krylov subspace methods developed so far can help us? After all, we only know how to find the ground state and calculate time evolution of any state, whereas calculating a correlation function requires taking the trace over full ensemble of states. This were the concept of **(Dynamical) Quantum Typicality** ((D)QT) comes into play, which broadly speaking postulates that a set of states with a common feature e.g. the same energy, should give a narrow distribution of some other feature e.g. expected value of some observable[31]. For the pedagogical purposes, we shall now briefly review two approaches to (D)QT, first in Sec. 2.4.1 following the article by Popescu et al. [9] focusing on a conceptual point of view, and second in Sec. 2.4.2 following Bartsch and Gemmer [31] and Steinigeweg et al. [32], giving us a concrete numerical tool for evaluating correlation functions, complete with rigorous error analysis.

2.4.1 General canonical principle

Roughly speaking, Quantum Typicality is an attempt to replace the fundamental postulate of statistical mechanics[33], the equal *a priori* probability postulate, by a principle that is fundamentally different, referring not to statistical ensembles or time averages, but to individual states. Another key characteristic of this new postulate, dubbed **general canonical principle**[9] or **canonical typicality**[8] is the existence of a rigorous mathematical proof, unlike in the case of equal *a prior* probability postulate. Let us now consider an isolated quantum system, called the universe U, partitioned into two components, the system S and the much larger environment E. In language of Hilbert spaces, this decomposition is $\mathcal{H}_U = \mathcal{H}_S \otimes \mathcal{H}_E$ such that $\mathcal{D}_s = \dim(\mathcal{H}_S) \ll d_E = \dim(\mathcal{H}_E)$. We can also impose some global constraint R for the universe, represented as restriction of the allowed states to some subspace $\mathcal{H}_R \subseteq \mathcal{H}_S \otimes \mathcal{H}_E$. We take the restricted universe to be in a maximally entangled state

$$\rho_R = \frac{1}{\mathcal{D}_R}\mathbb{1}_R \tag{2.20}$$

capturing our lack of knowledge about the system and being consistent with our intuition from statistical mechanics, about assigning *a priori* equal probability to each pure state. Now, we define a canonical state of our system S, as the density matrix obtaiend from ρ_R by tracing out the degrees of freedom of the environment E

$$\rho_S^{\mathrm{C}} = \mathrm{Tr}_E(\rho_R) \tag{2.21}$$

The crucial insight of canonical typicality is that we can take the universe to be in some pure state $\rho_R = |\psi\rangle\langle\psi|$ and the state of the system

$$\rho_S(\psi) = \mathrm{Tr}_R(|\psi\rangle\langle\psi|) \tag{2.22}$$

will be very close to the canonical state ρ_S^{C}. Moreover, this 'closeness' can be quanitfied very precisely, using a mathematical result from the asymptotic theory of finite dimensional normed spaces, called the Levy's lemma[34], which tells us about properties of typical points on on high-dimensional hypherspheres. Because of normalization, pure quantum states can be represented as point of a hyphersphere, hence the lemma is applicable. Let us now introduce some concepts necessary for the precise statement of the result. A precise notion of distance between two objects requires a metric and in our case a suitable metric will be induced by a norm on the vetor space of operators. There are two norms relevant for this problem, the **trace norm**

$$\|\rho\|_1 = \mathrm{Tr}\,|\rho| = \mathrm{Tr}\left(\sqrt{\rho^\dagger\rho}\right) \tag{2.23}$$

and the **Hilbert-Schmidt norm**

$$\|\rho\|_2 = \sqrt{\mathrm{Tr}\,(\rho^\dagger\rho)}\,. \tag{2.24}$$

The trace norm is used directly in the preciste statements of the general canonical principle, because $\|\rho_1 - \rho_2\|_1$ quantifies how hard is to tell apart ρ_1 and ρ_2 using measurements. Indeed, it can be shown that $\|\rho\|_1 = \sup_{\|A\|\leq 1}\mathrm{Tr}(\rho A)$, where $\|\cdot\|$ is the operator norm. The Hilbert-Schmidt norm is used during the proof, as it is a bit easier to manipulate and can be easily

related to the trace norm using Jensen's inequality[35] for convex functions applied to $\phi(x) = x^2$. Taking $\{\lambda_i\}_{i=1}^{\mathcal{D}}$ to be the eigenvalues of ρ we have

$$\begin{aligned}\|\rho\|_1^2 &= \left(\sum_{i=1}^{\mathcal{D}} |\lambda_i|\right)^2 = \mathcal{D}^2 \left(\sum_{i=1}^{\mathcal{D}} \frac{1}{\mathcal{D}} |\lambda_i|\right)^2 = \mathcal{D}^2 \phi\left(\sum_{i=1}^{\mathcal{D}} \frac{1}{\mathcal{D}} |\lambda_i|\right) \\ &\leq \mathcal{D}^2 \sum_{i=1}^{\mathcal{D}} \frac{1}{\mathcal{D}} \phi(\lambda_i) = \mathcal{D} \sum_{i=1}^{\mathcal{D}} |\lambda_i|^2 = \mathcal{D} \|\rho\|_2^2 \end{aligned} \tag{2.25}$$

Hence, $\|\rho\|_1 \leq \sqrt{\mathcal{D}} \|\rho\|_2$. We shall meet the Hilbert-Schmidt norm again in the next section, when discussing the algorihm searching for local integrals of motion. The precise theorem estabilishing the typicality is as follows

Theorem 2.2 *Let V be a function assigning to each subset of $\mathcal{H}_U$ its volume (in the sense of a suitable Haar measure[34]). Then, the following inequality holds*

$$\frac{V\left[\{|\psi\rangle \in \mathcal{H}_R \mid \frac{1}{2}\|\rho_S(\psi) - \rho_S^{\mathrm{C}}\|_1 \geq \eta\}\right]}{V\left[\{|\psi\rangle \in \mathrm{H}_R\}\right]} \leq \eta' \tag{2.26}$$

where

$$\eta = \epsilon + \frac{1}{2}\sqrt{\frac{\mathcal{D}_S}{\mathcal{D}_E^{\mathrm{eff}}}}$$
$$\eta' = 4e^{-\frac{2}{9\pi^3}\mathcal{D}_R\epsilon^2}$$

and the effective dimension of environment subspace is $\mathcal{D}_E^{\mathrm{eff}} = \frac{1}{\mathrm{Tr}\left(\rho_E^2\right)} \geq \frac{\mathcal{D}_R}{\mathcal{D}_S}$, where $\rho_E = \mathrm{Tr}_S(\rho_R)$.

Mathematically inclined readers are referred to Popescu et al. [36] for the full proof of this theorem, but for us it is important what this theorem means, namely that all but exponentially rare pure states of the universe are on the level of the system indistinguishable from the canonical state ρ_S^{C}. For our purposes in this thesis, we are interested in the case were the constraint R is that the total energy in the universe is close to some fixed value E. Assuming that the systems is weakly coupled with the environment, it becomes a standard exercise in statistical mechanics to show that the canonical state is the Gibbs canonical ensemble

$$\rho_S^{\mathrm{C}} \propto \exp\left(-\frac{H_S}{k_B T}\right) \tag{2.27}$$

where H_s is the Hamiltonian of the system and T is the temperature set by the energy E.

2.4.2 Dynamical Quantum Typicality

Another approach to Quantum Typicality is not concern directly with quantum states, but with expectation values of quantum observables instead. It was shown that for states drawn from a particular distribution in Hilbert space, the expectation values of a generic observable Q are very similar[37]. This result was further extended in the case of a unitarly invariant probability distribution, that is normalized states of the form

$$|\psi\rangle = \sum_{i=1}^{\mathcal{D}} c_i |i\rangle \tag{2.28}$$

where $\mathrm{Re}\, c_i$ and $\mathrm{Im}\, c_i$ are drawn from multidimensional Gaussian distribuion with zero mean, and $\{|i\rangle\}_{i=1}^{\mathcal{D}}$ is an arbitrary basis. Technicaly, because $\sum_{i=1}^{\mathcal{D}} |c_i|^2 = 1$, the coefficients are not indepentent, thus the full distribution is not necessarily Gaussian. However, Central Limit Theorem ensures that for $\mathcal{D}$ suitably large the distribution is indeed close Gaussian, with the standard deviation equal $1/\sqrt{2\mathcal{D}}$[38]. Using the **Hilbert space average method**, analytical expressions for both the average HA and variance HV of $\langle\psi|Q|\psi\rangle$ were derived[31]. They are as follows

$$\mathrm{HA}\left[\langle\psi|Q|\psi\rangle\right] = \frac{\mathrm{Tr}\,(Q)}{\mathcal{D}} \tag{2.29}$$

$$\mathrm{HV}\left[\langle\psi|Q|\psi\rangle\right] = \frac{1}{\mathcal{D}+1}\left(\frac{\mathrm{Tr}\,(Q^2)}{\mathcal{D}} - \left(\frac{\mathrm{Tr}\,(Q)}{\mathcal{D}}\right)^2\right) \tag{2.30}$$

Proof of the above equations is not difficult conceputally, however requires evaluation of rather cumbersome integrals over high-dimensional hyphersphers so we shall refrain from spelling it out in full. Interested reader can find all the details in a book by Gemmer et al. [38]. Let us now take $Q(t) = \hat{\rho}A(t)B$ and and define a quanitity

$$\alpha = \mathcal{D}\,\langle\psi|\sqrt{\hat{\rho}}A(t)B\sqrt{\hat{\rho}}|\psi\rangle \tag{2.31}$$

Note that because density matrix $\hat{\rho}$ is positive semi-definite and Hermitian, the square root $\sqrt{\hat{\rho}}$ exists and is well defined. The next step is to plug α into equations (2.29) and (2.30)

$$\mathrm{HA}\left[\alpha\right] = \mathrm{Tr}\left(\sqrt{\hat{\rho}}A(t)B\sqrt{\hat{\rho}}\right) = \mathrm{Tr}\left(\hat{\rho}A(t)B\right) = \frac{\mathrm{Tr}\left(e^{-\beta H}A(t)B\right)}{\mathcal{Z}} \tag{2.32}$$

$$\begin{aligned}\mathrm{HV}\left[\alpha\right] &= \frac{\mathcal{D}^2}{\mathcal{D}+1}\left(\frac{\mathrm{Tr}\left(\left(\sqrt{\hat{\rho}}A(t)B\sqrt{\hat{\rho}}\right)^2\right)}{\mathcal{D}} - \left(\frac{\mathrm{Tr}\left(\sqrt{\hat{\rho}}A(t)B\sqrt{\hat{\rho}}\right)}{\mathcal{D}}\right)^2\right) \\ &\leq \frac{\mathcal{D}}{\mathcal{D}+1}\,\mathrm{Tr}\left(\hat{\rho}A(t)B\hat{\rho}A(t)B\right) < \mathrm{Tr}\left(\hat{\rho}A(t)B\hat{\rho}A(t)B\right)\end{aligned} \tag{2.33}$$

Looking at eq. (2.32) we immediately see the desired way of calculating the correlation function.

$$\tilde{C}_{AB}(t) = \mathrm{Re}\,\frac{\mathrm{Tr}\left(e^{-\beta H}A(t)B\right)}{\mathcal{Z}} = \mathrm{Re}\,\frac{\mathcal{D}}{\mathcal{Z}}\,\langle\psi|e^{-\frac{\beta H}{2}}e^{iHt}Ae^{-iHt}Be^{\frac{-\beta H}{2}}|\psi\rangle + \mathrm{Re}\,\epsilon \tag{2.34}$$

where ϵ is the error we made by using just one random state $|\psi\rangle$. Let us massage this expression a bit more by introducing two auxiliary states $|\psi_\beta(t)\rangle = e^{-iHt}e^{-\frac{-\beta H}{2}}|\psi\rangle$ and $|\phi_\beta(t)\rangle = e^{-iHt}Be^{-\frac{-\beta H}{2}}|\psi\rangle$. We can also calculate the partition function using the random state $|\psi\rangle$ as

$$\mathcal{Z} = \mathrm{Tr}\left(e^{-\beta H}\right) = \mathcal{D}\,\langle\psi|e^{-\beta H}|\psi\rangle = \mathcal{D}\,\langle\psi_\beta(0)|\psi_\beta(0)\rangle \tag{2.35}$$

Combining (2.34) and (2.35), we arrive at the final expression, as seen in literature[24,29,30]

$$\tilde{C}_{AB}(t) = \mathrm{Re}\,\frac{\langle\psi_\beta(t)|A|\phi_\beta(t)\rangle}{\langle\psi_\beta(0)|\psi_\beta(0)\rangle} + \mathrm{Re}\,\epsilon \tag{2.36}$$

We have sucessfully shifted time evolution and the action of density matrix to state vectors instead of operators, hence we may apply the Krylov time propagator, studied in previous section, to calculate both real and imaginary time evolution. Apart from that, the only other

numerical calculations are sparse matrix-vector multiplication[5] and inner products of vectors, which are much less demanding than full exact diagonalization.

The final thing left is to estimate the error ϵ, in order to show that this approach actually makes sense. From eq. (2.33) it is clear that the Hilbert Space Average of ϵ is zero, as

$$\mathrm{HA}(\epsilon) = \mathrm{HA}\left(\alpha - \mathrm{Tr}\left(\hat{\rho}A(t)B\right)\right) = 0 \tag{2.37}$$

Equation (2.33), for Hilber Space Variance, allows us to estimate the standard deviation

$$\begin{aligned}
(\sigma(\epsilon))^2 &= \mathrm{HV}\left[\epsilon\right] = \mathrm{HV}\left[\alpha\right] < \mathrm{Tr}\left(\hat{\rho}A(t)B\hat{\rho}A(t)B\right) \\
&= \sum_{m,n} \frac{e^{-\beta\varepsilon_m}}{\mathcal{Z}} \langle m|A(t)B|n\rangle \frac{e^{-\beta\varepsilon_n}}{\mathcal{Z}} \langle n|A(t)B|m\rangle \\
&< \sum_{m,n} \frac{e^{-\beta\varepsilon_m}}{\mathcal{Z}} \langle m|A(t)B|n\rangle \frac{e^{-\beta\varepsilon_0}}{\mathcal{Z}} \langle n|A(t)B|m\rangle \\
&= \frac{1}{\mathrm{Tr}\left(e^{-\beta(H-\varepsilon_0)}\right)} \sum_{m} \frac{e^{-\beta\varepsilon_m}}{\mathcal{Z}} \langle m|A(t)BA(t)B|m\rangle \\
&= \frac{1}{\mathrm{Tr}\left(e^{-\beta(H-\varepsilon_0)}\right)} \mathrm{Tr}\left(\hat{\rho}A(t)BA(t)B\right)
\end{aligned} \tag{2.38}$$

Where the red inequality follows from the fact that Boltzmann factor is a strictly decreasing function and assuming that the spectrum of H is ordered in increasing fashion. Defining the effective dimension $\mathcal{D}_{\mathrm{eff}} \equiv \mathrm{Tr}\left(e^{-\beta(H-\varepsilon_0)}\right)$, we finally obtain the upper bound on standard deviation of error as

$$\sigma\left(\mathrm{Re}\,\epsilon\right) < \sqrt{\frac{\mathrm{Re}\langle A(t)BA(t)B\rangle}{\mathcal{D}_{\mathrm{eff}}}} \tag{2.39}$$

From this bound we see that at infinite temperature, the error is exponentially supressed in system size and thus for suitably large systems even a single pure state $|\psi\rangle$ is enough to obtain a very good approximation of the correlation function. It becomes progressively worse with lower temperature, however as the mean error is 0 we can always average over a few random states. Fortunately, in this thesis we are only interested in the case $\beta \to 0$, so usually a single run will be enough.

As an example application, we shall look at spin current in XXZ model. Sum rule, normalized corr etc.

[5] Matrices representing local observables will also be sparse.

3

Spin transport in long range XXZ model

The results concerning spin transport in the long range XXZ have already been published in Mierzejewski et al. [1].

4

Relaxation eigenmodes in long range XXZ model

The results concerning spin transport in the long range XXZ have already been published in Mierzejewski et al. [1]. Section about Krylov approach to lioms moved here.

5 Summary

AAA

Bibliography

[1] M. Mierzejewski, J. Wronowicz, J. Pawłowski, and J. Herbrych, "Quasiballistic transport in the long-range anisotropic Heisenberg model", Phys. Rev. B **107**, 045134 (2023).

[2] V. Simoncini and L. ELDÉN, "Krylov subspaces", The Princeton Companion to Applied Mathematics, 113–114 (2015).

[3] L. N. Trefethen and D. I. I. I. Bau, *Numerical Linear Algebra* (Philadelphia, 1997).

[4] A. W. Sandvik, "Computational studies of quantum spin systems", AIP Conf. Proc. **1297**, 135–338 (2010).

[5] D. Crivelli, "Particle and energy transport in strongly driven one-dimensional quantum systems", PhD thesis (Uniwersytet Śląski w Katowicach, 2016).

[6] T. J. Park and J. C. Light, "Unitary quantum time evolution by iterative Lanczos reduction", J. Chem. Phys. **85**, 5870–5876 (1986).

[7] J. Gemmer and G. Mahler, "Distribution of local entropy in the Hilbert space of bipartite quantum systems: origin of Jaynes' principle", Eur. Phys. J. B **31**, 249–257 (2003).

[8] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, "Canonical Typicality", Phys. Rev. Lett. **96**, 050403 (2006).

[9] S. Popescu, A. J. Short, and A. Winter, "Entanglement and the foundations of statistical mechanics", Nat. Phys. **2**, 754–758 (2006).

[10] A. Weisse and H. Fehske, "Exact Diagonalization Techniques", in *Computational many-particle physics*, edited by H. Fehske, R. Schneider, and A. Weisse, Lecture Notes in Physics (Springer, Berlin, Heidelberg, 2008), pp. 529–544.

[11] S. R. Garcia and R. A. Horn, *A Second Course in Linear Algebra*, Google-Books-ID: nXOuDgAAQBAJ (Cambridge University Press, May 2017).

[12] H. Eves, *Elementary Matrix Theory* (Dover Publications, New York, Apr. 1980).

[13] C. Sanderson and R. Curtin, "Armadillo: a template-based C++ library for linear algebra", The Journal of Open Source Software **1**, 26 (2016).

[14] V. Simoncini, "Computational Methods for Linear Matrix Equations", SIAM Rev. **58**, 377–441 (2016).

[15] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain", Science **49**, 64 (1994).

[16] C. Moler and C. Van Loan, "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later", SIAM Rev. **45**, 3–49 (2003).

[17] P. Schmitteckert, "Nonequilibrium electron transport using the density matrix renormalization group method", Phys. Rev. B **70**, 121302 (2004).

[18] D. Stanek, C. Raas, and G. S. Uhrig, "Dynamics and decoherence in the central spin model in the low-field limit", Phys. Rev. B **88**, 155305 (2013).

[19] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, "Time-evolving a matrix product state with long-ranged interactions", Phys. Rev. B **91**, 165112 (2015).

[20] P. E. Dargel, A. Wöllert, A. Honecker, I. P. McCulloch, U. Schollwöck, and T. Pruschke, "Lanczos algorithm with matrix product states for dynamical correlation functions", Phys. Rev. B **85**, 205119 (2012).

[21] N. Mohankumar and S. M. Auerbach, "On time-step bounds in unitary quantum evolution using the lanczos method", Comput. Phys. Commun. **175**, 473–481 (2006).

[22] E. Dagotto, "Correlated electrons in high-temperature superconductors", Rev. Mod. Phys. **66**, 763–840 (1994).

[23] G. D. Mahan, *Many-particle physics* (Springer US, Boston, MA, 2000).

[24] R. Steinigeweg, J. Gemmer, and W. Brenig, "Spin-current autocorrelations from single pure-state propagation", Phys. Rev. Lett. **112**, 120601 (2014).

[25] J. Sirker, R. G. Pereira, and I. Affleck, "Diffusion and Ballistic Transport in One-Dimensional Quantum Systems", Phys. Rev. Lett. **103**, 216602 (2009).

[26] R. Steinigeweg and J. Gemmer, "Density dynamics in translationally invariant spin-12 chains at high temperatures: a current-autocorrelation approach to finite time and length scales", Phys. Rev. B **80**, 184402 (2009).

[27] C. Karrasch, J. Hauschild, S. Langer, and F. Heidrich-Meisner, "Drude weight of the spin-$\frac{1}{2}$ XXZ chain: Density matrix renormalization group versus exact diagonalization", Phys. Rev. B **87**, 245128 (2013).

[28] C. Karrasch, J. H. Bardarson, and J. E. Moore, "Finite-Temperature Dynamical Density Matrix Renormalization Group and the Drude Weight of Spin-$1/2$ Chains", Phys. Rev. Lett. **108**, 227206 (2012).

[29] R. Steinigeweg, J. Gemmer, and W. Brenig, "Spin and energy currents in integrable and nonintegrable spin-1/2 chains: A typicality approach to real-time autocorrelations", Phys. Rev. B **91**, 104404 (2015).

[30] J. Richter and R. Steinigeweg, "Combining dynamical quantum typicality and numerical linked cluster expansions", Phys. Rev. B **99**, 094419 (2019).

[31] C. Bartsch and J. Gemmer, "Dynamical typicality of quantum expectation values", Phys. Rev. Lett. **102**, 8–11 (2009).

[32] R. Steinigeweg, A.Khodja, H. Niemeyer, C. Gogolin, and J. Gemmer, "Pushing the limits of the eigenstate thermalization hypothesis towards mesoscopic quantum systems", Phys. Rev. Lett. **112**, 1–5 (2014).

[33] K. Huang, *Statistical Mechanics* (Wiley, 1987).

[34] G. S. Vitali D. Milman, *Asymptotic theory of finite dimensional normed spaces* (Springer Berlin Heidelberg, 1986).

[35] J. L. W. V. Jensen, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes", Acta Math. **30**, 175–193 (1906).

[36] S. Popescu, A. J. Short, and A. Winter, *The foundations of statistical mechanics from entanglement: Individual states vs. averages*, arXiv:quant-ph/0511225, Nov. 2006.

[37] P. Reimann, "Typicality for generalized microcanonical ensembles", Phys. Rev. Lett. **99**, 160404 (2007).

[38] J. Gemmer, M. Michel, and G. Mahler, "Quantum thermodynamics", Lect. Notes. Phys. **657**, 10.1007/b98082 (2009).

A

Hilbert subspaces with fixed momentum