Spectral analysis of the Zig-Zag process on the torus

by

Sjoerd Wiarda

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday December 22, 2021 at 13:00 AM.

Student number:	4590252	
Master's program:	Applied Mathematics	
Specialization:	Analysis	
Thesis committee:	dr. ir. J. Bierkens,	TU Delft, supervisor
	prof. dr. ir. M. C. Veraar,	TU Delft
	dr. R. C. Kraaij,	TU Delft

Abstract

In this thesis, we analyse the spectrum of the generator of the one-dimensional Zig-Zag process defined on the torus $\mathbb{T}$. This is a piecewise deterministic Markov process (PDMP) used in Monte Carlo Markov chain methods (MCMC) for sampling from a probability distribution and calculating integrals [PW12], [BFR19], [BVD18]. We show for Lipschitz potentials U and bounded refreshment rates $\lambda_0 \in L^\infty(\mathbb{T})$ that the spectral gap $\kappa = \sup\{\operatorname{Re}\lambda : \lambda \in \sigma(\mathscr{L})\} \setminus \{0\}$ of the associated J-self-adjoint generator $\mathscr{L}$ on $L^2(\mathbb{T}, \nu)$ and $C(\mathbb{T} \times \{+1, -1\})$ is positive. Moreover, we give two lower bounds for κ by making use of one of the Schur complements associated with a block operator that is unitarily equivalent to $\mathscr{L}$. In addition we show that the spectrum of $L^2(\mathbb{T}, \nu)$ and $C(\mathbb{T} \times \{+1, -1\})$ are the same and that the generator defined on both spaces generates a contraction semigroup. Under the assumption of unimodality of the potential U and a zero refreshment rate, we show that a vertical "asymptotic line" exists to which all of the eigenvalues converge. Furthermore, we show that a spectral mapping theorem exists where, due to the spectral line, the spectrum of the semigroup can become uncountable or countable depending on the time parameter of the semigroup $P(t)$ generated by $\mathscr{L}$. Lastly, we show that a discretisation of the spectrum generates a semigroup that converges uniformly on each bounded time interval to the semigroup of the Zig-Zag process and we use these discretisations to numerically analyse the behaviour of general potentials and refreshment rates.

Keywords: spectral gap, J-self-adjoint operator, spectral mapping theorem, Zig-Zag process, piecewise deterministic Markov process, MCMC, spectral theory, Schur complement.

Contents

1
Introduction

In many practical applications, there is a need to create random variables which have a specific distribution. For example, the normal distribution is a very commonly occurring distribution in nature and many phenomena can be approximated by a normal distribution such as human height, human weight, and sums of random variables. If one would want to model such variables they could use functions that are available in most statistical software to create these random variables and then they could perform calculations/simulations. Of course, this is not only possible for normal distributions, but also for many other relatively simple well-known distributions.

However, when it comes to many processes in the world they are not simple enough to be described in terms of the well-known distributions. This becomes especially problematic when the variables that one would like to simulate are high-dimensional and the so-called 'curse of dimensionality' comes in the way. One effective approach to simulating such highly dimensional and complicated distributions can be done using Markov chain Monte Carlo (MCMC) algorithms. These are iterative algorithms that create random variables $X_1, X_2, \dots$ tuned in such a way that they will converge to the desired distribution that one would want for their simulation. More mathematically the function $n \to X_n$ will converge in distribution to the distribution of the desired random variable. Practically this means that there is an amount of iterations N of this algorithm such that the distribution of X_N is as similar to the desired distribution as one would like for their application. Another effective use of MCMC algorithms can be used is for the approximation of certain integrals. Specifically, they can be great at calculating expectations of random variables $\mathbb{E}[f(X)]$ for a wide range of functions. The need for calculating such expectations and creating random variables with specific distributions arises in all sorts of practical applications such as the field of computational physics [Thi07], computational biology [MB20], and computational chemistry [Cep03] that are often interested in calculating such quantities related to various physical systems.

The main difficulty that arises here is what specific MCMC works best for a particular distribution and how many iterations are required before the algorithm reaches this desired distribution sufficiently. There are many different MCMC algorithms which all have their benefits as well as their drawbacks. In recent years several of these MCMC algorithms have shown a lot of promise in terms of their convergence speed and computational costs based on piecewise deterministic Markov processes (PDMP) [BFR19], [BVD18], [PW12]. PDMPs are processes which are partially stochastic (random) and partially deterministic. The specific process that we are interested in is called the Zig-Zag process. This specific PDMP is defined as set of random variables which is defined by two components $\{(X(t), \Theta(t))\, t \geq 0\}$. Where $X(t)_{t\geq 0}$ is related to $\Theta(t)_{t\geq 0}$ by $X(t) = X(0) + \int_0^t \Theta(s)\, ds$ and $\Theta(t)_{t\geq 0}$ is a specific type of a stochastic process that only takes on a finite amount of values. The process $X(t)_{t\geq 0}$ can be defined on different spaces depending on what type of variables one would like to sample. For example, if one would like to sample from the normal distribution then $X(t)_{t\geq 0}$ can be defined on $\mathbb{R}^d$ where it will then approximately get the probability density function defined by $p(x) = \frac{e^{-U(x)}}{\int_{\mathbb{R}^d} e^{-U(y)}\, dy}$ with $U(x) = \frac{||x||^2}{2}$. We refer to U as the potential function and for any continuous probability function $p(x)$ such that $p(x) > 0$ for all x, we can define this potential through $U(x) = -\log p(x)$.

Associated to the Zig-Zag process is a tunable parameter called the refreshment rate $\lambda_0 : \Omega \to [0, \infty)$. This parameter can be adjusted to increase the rate at which the algorithm converges to the desired distribution. However, increasing this parameter too much can have an adverse effect on the convergence rate. It is not clear how to choose this parameter such that this convergence is optimal. To examine the optimal convergence we must define a metric that indicates how close the distribution is. One of them is by examining the convergence of the so-called semigroup of the Zig-Zag process on a class of functions. For example, we could take our class of functions to be the continuous functions on the underlying state space denoted by $C(\Omega)$. For these functions we can define the semigroup by $P_t f := \mathbb{E}[f(X_t)|X_0 = x]$, where the expectation is taken with respect to $p(x)$. As X_t converges in distribution to a variable X with distribution $p(x)$, we will expect that $\mathbb{E}[f(X_t)|X_0 = x]$ will converge in some way to $\mathbb{E}[f(X)]$. For example, we can examine this convergence by looking at how fast $||P_t f - \mathbb{E}[f(X)]||_\infty := \sup_{x\in\Omega} |\mathbb{E}[f(X_t)|X_0 = x] - \mathbb{E}[f(X)]|$ converges to 0 as $t \to \infty$. It turns out that this convergence will happen at an exponential rate (as we shall see in Theorem 4.4.8 for $\Omega = \mathbb{T}$). Specifically we will have that this convergence can be described as $||P_t f - \mathbb{E}[f(X)]||_\infty \leq Me^{-\kappa t}||f - \mathbb{E}[f(X)]||$ where we have that $M, \kappa \geq 0$. In this formulation, the largest possible value of κ is called the convergence rate and it will describe the rate of convergence for the slowest converging function(s). Knowing the largest possible value of κ is important as it can give us certainty for how fast our algorithm is converging for all of the possible functions, so in particular, it would be very useful if we could choose λ_0 such that κ is as large as possible.

There has been previous research concerning the value of κ, specifically it has been shown for $\Omega = \mathbb{R}$ and various assumptions on U that κ is equal to the so-called spectral gap of the generator of the semigroup (Theorem 4.14 [BV21]). Intuitively this spectral gap can be thought of as the largest value of the real components of the eigenvalues of a particular 'matrix'. Determining this value is difficult as unlike with finite-dimensional matrices there is not a straightforward way of calculating all the eigenvalues of the generator. Furthermore, it has been shown that if the potential is sufficiently smooth, meaning $U \in C^2(\mathbb{R})$, if this potential is unimodal, meaning that it has at most 1 minimum, and if the refreshment rate λ_0 is equal to 0 that the spectral gap is positive (Theorem 4.11 [BV21]), however, there is no clear bound on this spectral gap. In the same paper, it was shown that the eigenvalues can be seen as roots of a complex function, however, this gives little information about the actual position of the roots. When it comes to the full description of the spectrum and the norm of the semigroup the only precise values have been found in the case $\Omega = \mathbb{T}$, $\lambda_0 = 0$, and U being constant [MM12].

In this thesis, we will further discuss the properties of these eigenvalues of the Zig-Zag process with $\Omega = \mathbb{T}$. This case can be of interest when trying to sample from a distribution defined on bounded intervals such as $[0,1]$, namely if we have a potential function $U(x) : [0,1] \to \mathbb{R}$ we can create a version of this potential on the torus by extending $U(x) = U(-x)$ for $x \in [-1,0]$ such that U becomes a continuous function on $\mathbb{T}$. This way we can simulate the MCMC process on the torus and use the samples for this process to generate samples and integrals on the bounded interval. We will specifically examine the case where U is a Lipschitz function which is a continuous function such that $|f(x) - f(y)| \leq |x - y|$ for all $x, y \in \mathbb{T}$. Moreover, these functions can be seen as the continuous functions such that the almost everywhere defined derivative U' is almost everywhere bounded, meaning $U' \in L^\infty(\mathbb{T})$. This includes function such as e^{-x^2}, $|x|$, and $x^2 \sin(1/|x|)$.

We begin in Chapter 2 with preliminary theory surrounding the mathematics that is used for analysing the process and the convergence rate. In Chapter 3 we will discuss general spectral properties of the generator associated with the Zig-Zag process by following an analysis using block matrices and Schur complements [Tre08],[KLT04] which allows us to easily construct the inverse of the generator. Moreover, we will show that for Lipschitz potentials as described above that there is always a positive spectral gap with a small assumption on the refreshment rate. Furthermore, we will give two different lower bounds for this spectral gap. In Chapter 4 we will show that similar to [BV21] that the spectral mapping theorem holds for unimodal potentials, moreover, we will prove the existence of an asymptotic line of eigenvalues which gives us that the set of eigenvalues of the semigroup (apart from $\gamma = 0$) is not in general equal to the spectrum of the semigroup. Furthermore, we will discuss spectral properties for both the square-integrable functions $L^2(\mathbb{T}, \nu)$ as well as the set of continuous functions $C(\mathbb{T})$. In Chapter 5 we will show that a certain discretization of the generator of the Zig-Zag process creates a contraction semigroup that converges uniformly to the semigroup of the Zig-Zag process on bounded intervals $[0, t]$, with $t > 0$. Moreover, we will compare the spectrum of the discrete generator with the spectrum of the generator of the Zig-Zag process and make some remarks and observations surrounding certain potentials and refreshment rates. The code used to generate the numerical spectrum can be found in [Wia].

2

Preliminary theory

In order to mathematically describe the Zig-Zag process and its related properties, we need to make use of several mathematical concepts related to the process. First, we will review terminology related to the probability theoretic approach and definition of the Zig-Zag process. Then we will introduce several concepts from functional analysis. After this, we will discuss semigroups which are objects describing the convergence/evolution of the Zig-Zag process. Lastly, we will talk specifically about the Zig-Zag process and related processes.

2.1. Stochastic processes

First, we need to have the definitions in order to properly define the process itself and to get an idea about key properties that the Zig-Zag process has. The Zig-Zag process is a specific case of a large class of random processes over time called Markov processes which are stochastic processes.

Definition 2.1.1 (Stochastic process). *Given a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *and a measurable space* $(Y, \mathscr{Y})$*. A stochastic process on* $[0,\infty)$ *is defined by a set of random variables* $\{X_t, t \in [0,\infty)\}$ *such that for each* $t \in [0,\infty)$ *we have that* X_t *is a random variable defined from the probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *to the measurable space* $(Y, \mathscr{Y})$.

Precise definitions and an extended introduction about probability spaces, measurable spaces, and random variables can be found in most introductory probability books [Fel68].

Stochastic processes can describe an enormous amount of different phenomena. For illustration, a stochastic process can be seen as a set of paths. Namely on an associated probability space $(\Omega, \mathscr{F}, \mathbb{P})$ each realization $\omega \in \Omega$ of the process describes a path $t \to X_t(\omega)$ that this process follows. For a stochastic process, such a realization could describe the price of a stock. Another stochastic process could describe the position of an atomic particle along these paths. Or alternatively, for yet another different process a path could describe the number of customers at a store.

For every process $\{X_t, t \in [0,\infty)\}$ we can describe the distribution $\mathbb{P}(X_t = \cdot)$ of the process at every time t. This distribution can change over time or remain constant. A specific case of a process for which the distribution does not change is a stationary process. A stationary process has the property that for a set of timestamps $t_1, t_2, \ldots t_n \in [0,\infty)$ with $n \in \mathbb{N}$ we have that the cumulative distribution function F_X of $(X_t)_{t\geq 0}$ is time invariant, meaning $F_X(x_{t_1}, x_{t_2}, \ldots, x_{t_n}) = F_X(x_{t_1+\tau}, x_{t_2+\tau}, \ldots, x_{t_n+\tau})$ for all $\tau \in [0,\infty)$. An example of such a process would be Brownian motion. Note that this is something different than having a constant distribution, for example we could create a process with a constant distribution in the following way: If one were to do an infinite amount of independent coin tosses $Y_0, Y_1, Y_2, \ldots$ (with $\mathbb{P}(Y_i = 1)) = \mathbb{P}(Y_i = -1)) = \frac{1}{2}$) and create a stochastic process $\{X_t, t \geq 0\}$ which is equal to the outcome of these tosses such that $X_t = Y_n$ for $t \in [n, n+1)$, then for every $t \geq 0$ we have that the distribution of X_t is the same (So for all $s, t \in [0,\infty)$ we have that $\mathbb{P}(X_t = 1) = \mathbb{P}(X_s = 1) = \mathbb{P}(X_t = -1) = \mathbb{P}(X_s = -1)$). Such a process would have a constant distribution but would not be a stationary process since $F_X(x_{0.2}, x_{0.8}) \neq F_X(x_{0.7}, x_{1.3})$ due to $X_{0.2}, X_{0.8}$ always being equal to each other and $X_{0.7}, X_{1.3}$ possible being unequal to each other.

A lot of processes, in reality, are not stationary, this can be because we have information about the initial state of such processes which we can use to make a better prediction when not a lot of time has elapsed. For example, if we were to create a process that is the average of the coin-tossing process so $Z_t = \frac{1}{\lfloor t \rfloor + 1} \sum_{n=0}^{\lfloor t \rfloor} Y_n$. The distribution of this process is not constant and will change over time, this can specifically be seen by the fact that the range of this variable is constantly changing. However, this specific non-stationary distribution is converging to something, namely, it will converge to the constant 0 in distribution by the law of large numbers. In another example, we could look at the position of an air molecule in a room. If we let this molecule start in the centre of the room we know that after a nanosecond, its positional distribution has not changed much and that the particle is likely still somewhere in the centre of the room. However, assuming that the room is perfectly isolated and that there are no sources of heat in this room, if we were to examine the room after a couple of hours, then its initial position has a negligible influence on its current position and we can safely assume that it is equally likely to be anywhere in the room. With this particular process, we have that the distribution of the position of the particle has converged towards a uniform distribution across the room.

These two examples both have a property that we shall refer to as ergodicity.

Definition 2.1.2 (Ergodicity)**.** *Let $\{X_t, t \geq 0\}$ be a stochastic process defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and measurable space $(Y, \mathscr{Y})$. Define the total variation norm for measures by $||\pi(\cdot)||_{TV} := \sup_{A\in\mathscr{F}} |\pi(A)|$. The stochastic process is said to be ergodic if there is a measure ν on the same probability space such that for all $x \in Y$ we have that*

$$\lim_{t\to\infty} ||P(X_t \in \cdot | X_0 = x) - \nu(\cdot)||_{TV} = 0.$$

Ergodicity has different definitions depending on the context. For our purposes we will refer to this measure ν as the *stationary measure* of $\{X_t, t \geq 0\}$. This definition of ergodicity automatically implies that for every $x \in Y$ such that $X_0 = x$ we have that X_t converges in distribution to this distribution π as $t \to \infty$. This property is very important as it is one of the key reasons why algorithms using these processes such as the Zig-Zag process work in approximating distributions. Specifically, it has been shown that for some assumption on the potential U that the Zig-Zag process is an ergodic process [BRZ19].

When it comes to MCMC algorithms one important property of these stochastic processes is that they are memoryless. This means that if one would have a stochastic process $\{X_t, t \in [0,\infty)\}$where the path of the process is known up to a time $T \in [0,\infty)$ (So the exact value of X_t is known for $t \in [0, T]$) then the only value of the set $\{X_t, t \in [0, T]\}$ that is relevant to predict anything about the future of this process is the value at $t = T$. Using the value of X_T gives us as much information about the future of the process as did the information of the values of this process on the interval $[0, T]$. This property of a stochastic process to be memoryless is called the Markov property. A process with this property is called a Markov process. Specifically, we have that the Zig-Zag process is a Markov process.

Definition 2.1.3 (Markov property)**.** *Let $\{X_t, t \in [0,\infty)\}$ be a stochastic process with corresponding probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and measurable space $(Y, \mathscr{Y})$. Furthermore, define the filtration generated by $\mathscr{F}_T := \{X_t, t \in [0, T)\}$. We say that $\{X_t, t \in [0,\infty)\}$ has the Markov property if for all $S \in \mathscr{Y}$, $s \geq t \geq 0$ we have that*

$$P(X_t \in S|\mathscr{F}_s) = P(X_t \in S|X_s).$$

The Markov property allows us to describe the process in terms of its infinitesimal evolutions, that is we only need to characterize the possible evolution of the process on a tiny interval in order to describe how it will grow over time. This gives rise to the idea to describe the process using semigroups which we will describe in section 2.3. But before we can get there we need to first understand the objects attached to the definition of such a process which are the linear operators as will be described in the following section.

2.2. Operator theory

Describing the Zig-Zag process in terms of a Markov process stochastic process gives a lot of intuition as to how such a process evolves over time, however, it is not the only way in which we can describe it. A different approach to describe this process can be done using operator theory, which allows us to use the theorems and mathematical frameworks from this field in order to get new insights about the process.

When we refer to the convergence of the process we need to be able to express in what sense they are converging and what we mean by this convergence. For this, we need the space in which we work to be a Banach space. The definition of a Banach Space relies on the definition of a Cauchy sequence.

Definition 2.2.1 (Cauchy sequence)**.** *A sequence $(x)_{n\geq 0}^{\infty}$ with values in a normed space X is called Cauchy if for all $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for all $m, n \geq N$ we have that $||x_n - x_m|| < \epsilon$.*

The elements of a Cauchy sequence get infinitesimally close to each other as you compare the elements later in the sequence with each other. When these elements get very close one would expect that they would converge to a common element. This property is called completeness and it is what makes a normed vector space a Banach space.

Definition 2.2.2 (Banach space)**.** *A normed vector space $(X, ||\cdot||)$ is called a Banach space if and only if it is a complete metric space, which means that for all Cauchy sequences $(x)_{n\geq 0}^{\infty} \subseteq X$ we have that there is an $x \in X$ such that $||x - x_n|| \to 0$ as $n \to \infty$.*

There are many different examples of Banach spaces, we list a few:

- All of the real numbers $\mathbb{R}$ with the norm being the absolute value function $|\cdot|$. More general we can take an $n \in \mathbb{N}$ and look at $\mathbb{R}^n$ with the euclidean norm with $x \in \mathbb{R}^n, ||x||_2 := (\sum_{k=1}^n x_i^2)^{\frac{1}{2}}$.
- The space of continuous functions $f : \Omega \to \mathbb{C}$ on $\Omega = \mathbb{R}$ or $\mathbb{T}$ denoted by $C(\Omega)$ together with the uniform norm for $f \in C(\Omega)$ defined by $||f||_\infty := \sup_{x\in\Omega} |f(x)|$.
- The space of k-times differentiable functions on $\Omega = \mathbb{R}$ or $\mathbb{T}$ denoted by $C^k(\Omega)$ with the norm $||f|| = ||f||_\infty + \sum_{i=1}^k ||\partial^k f||_\infty$.
- The space of measurable functions with finite L^p-norm, where $1 \leq p \leq \infty$ on $\Omega = \mathbb{R}$ or $\mathbb{T}$ denoted by $L^p(\Omega)$ which means that for functions $f : \Omega \to \mathbb{C}$ the norm $||f||_{L^p(\Omega)} = (\int_\Omega |f(x)|^p dx)^{\frac{1}{p}}$ is finite. For $p = \infty$ we have $||f||_\infty = \operatorname{ess\,sup}_{x\in\Omega} |f(x)|$.
- The previous example makes use of a standard Lebesgue measure. However, any other measure μ on the measurable space (Ω, Σ) would also work. In later sections we will work with functions $f : \Omega \to \mathbb{C}$ for which the norm $||f|| := (\int_\Omega |f(x)|^p d\mu(x))^{\frac{1}{p}}$ is finite. We denote this space by $L^p(\Omega, \mu)$.

$C(\mathbb{T})$, $C^1(\mathbb{T})$, $L^p(\mathbb{T}, \mu)$ as given in the examples above will be useful for our analysis of the Zig-Zag process. For a function $f : [-\pi, \pi] \to \mathbb{C}$ to be part of $C(\mathbb{T})$ we need to have $f(\pi) = f(-\pi)$ likewise for it to be part of $C^1(\mathbb{T})$ we need that $\partial_x f(\pi) = \partial_x f(-\pi)$. When checking that a function $f : [-\pi, \pi] \to \mathbb{C}$ is differentiable extra care should be taken that it is periodic at the boundaries as well. For example, if we take $g \in C^\infty([-\pi, \pi])$ and define $f = \int_{-\pi}^x g(x)\,dx$ then $f \in C^\infty([-\pi, \pi])$ however $f \notin C^\infty(\mathbb{T})$, moreover it is not even in $f \in C(\mathbb{T})$ unless $\int_{-\pi}^{\pi} g(x)\,dx = 0$ as otherwise we do not have the continuity $f(\pi) = f(-\pi)$.

We also need to have some extra structure/an extension to this $L^p(\mathbb{T}, \mu)$ space by allowing the functions in this space to be differentiable. However, if we were to use the traditional definition of differentiability then it becomes harder to construct a useful Banach Space with the $L^p(\mathbb{T}, \mu)$ norm. Moreover, it removes a lot of functions which we would still like to examine, for example, the function $|x|$ is almost fully differentiable on $\mathbb{T}$ except at the points $x = 0, \pi, -\pi$ and would therefore not be considered a differentiable function. In order to extend this definition of differentiability to include more "almost differentiable" functions, a weaker definition of differentiability is introduced inspired by the integration by parts formula.

Definition 2.2.3 (Weak derivative)**.** *Take $\Omega \subset \mathbb{R}$ or $\Omega = \mathbb{T}$, $k \in \mathbb{N}$ if $f, g^k \in L^1_{loc}(\Omega)$, we say that g^k is the k-th weak derivative of f if for all $\phi \in C_c^\infty(\Omega)$ we have that:*

$$\int_\Omega f(x)\partial_x^k \phi(x)\,dx = (-1)^k \int_\Omega g(x)\phi(x)\,dx.$$

We denote these weak derivatives as $\partial_x^k f := g^k$.

The definition of the weak derivative uses the function space $L^1_{\text{loc}}(\Omega)$. This vector space is defined as the set of measurable functions where each function $f \in L^1_{\text{loc}}(\Omega)$ we have that for each compact set $K \subset \Omega$ that $f|_K \in L^1(K)$ where we mean with $f|_K$ the restriction of f on the set K. The definition of the weak derivative also makes use of $C_c^\infty(\Omega)$ which is the set of functions which are infinitely often differentiable and have finite support. That is for $f \in C_c^\infty(\Omega)$, we have that for each $k \in \mathbb{N}$ that $\partial_x^k f \in C(\Omega)$ and that there is a compact subset $K \subset \Omega$ such that $f(x) = 0$ for $x \in \Omega \setminus K$. In particular for $\Omega = \mathbb{T}$ the requirement in the weak derivative for $f, g^k \in L^1_{\text{loc}}(\mathbb{T})$ is the same as $f, g^k \in L^1(\mathbb{T})$ and $\phi \in C_c^\infty(\mathbb{T})$ is the same as $\phi \in C^\infty(\mathbb{T})$ due to the compactness of $\mathbb{T}$.

The definition of the weak derivative then gives us the ability to construct a Banach space of functions in $L^p(\mathbb{T}, \mu)$ that are weakly differentiable. These spaces are referred to as Sobolev spaces.

Definition 2.2.4 (Sobolev space)**.** *Let (Ω, Σ, μ) be a measure space with $\Omega \subset \mathbb{R}$ or $\Omega = \mathbb{T}$, $k \in \mathbb{N}$, and $1 \le p \le \infty$. The Sobolev space $W^{k,p}(\Omega,\mu)$ consists out of the of functions $f \in L^p(\Omega,\mu)$ such that for all integer $1 \le i \le k$ the weak derivative $\partial_x^k f$ exists and $\partial_x^k f \in L^p(\Omega,\mu)$. Moreover, we can define the Sobolev norm*

$$||f||_{W^{k,p}(\Omega,\mu)} = (||f||^p_{L^p(\Omega,\mu)} + \sum_{i=1}^{k} ||\partial_x^k f||_{L^p(\Omega,\mu)})^{\frac{1}{p}}.$$

Then we define the Sobolev space to be the Banach space $(W^{k,p}(\Omega,\mu), ||\cdot||_{W^{k,p}(\Omega,\mu)})$.

The Sobolev space has an interesting property on a space like $\mathbb{T}$ in that for a function $f \in W^{1,p}(\mathbb{T},\mu)$ with $p \in (1,\infty]$ the weak differentiability and the integrability of the weak derivative implies that function is continuous $f \in C(\mathbb{T})$ (Theorem 6.2 of [HR08]). Moreover, the space that these functions end up in is even a bit more regular than the set of continuous function. Specifically we have that $f \in C^{0,\alpha}(\mathbb{T})$, $\alpha = 1 - \frac{1}{p}$ which is the space of hölder continuous functions. That is the set of functions such that $\sup_{x,y\in\mathbb{T}} \frac{|f(x)-f(y)|}{|x-y|} < \infty$. In particular we have that for $p = \infty$ the set of hölder continuous functions coincides with the set of Lipschitz functions. Moreover, it turns out that the sets are equal $W^{1,\infty}(\mathbb{T}) = C^{0,1}(\mathbb{T})$ (In the sense that there is always an element in one set that is almost everywhere equal to the other). A similar statement holds for $p = 2$ with $W^{1,2}(\mathbb{T}) = \{f \in C(\mathbb{T}) : \partial_x f \in L^2(\mathbb{T})\}$ (Again with equality meaning that the functions of one set are almost everywhere equal to functions in the other set).

We have now introduced the functions that will be relevant for our analysis. Next, we will introduce bounded operators which are linear functions on Banach spaces which are continuous.

Definition 2.2.5 (Bounded operator)**.** *Let X, Y be Banach spaces. An operator $T : X \to Y$ is called bounded if there is a constant $C \ge 0$ such that for all $x \in X$ we have that $||Tx||_Y \le C||x||_X$. The smallest C for which this holds is called the operator norm denoted by $||T|| := \sup_{x\in X} \frac{||Tx||_Y}{||x||_X}$. Sometimes for clarity this is denoted as $||T||_{X\to Y}$ or $||T||_X$ if $X = Y$.*

It is possible to see that bounded operators are continuous in the following way, namely if we have a sequence $(x)_{n\ge0}^\infty \subseteq X$ with limit $x \in X$ such that $||x - x_n||_X \to 0$ then we have that $||Tx - Tx_n||_Y = ||T(x - x_n)||_Y \le ||T||||x - x_n||_X \to 0$. So we have $||Tx - Tx_n|| \to 0$. In fact the bounded operators are exactly the continuous linear functions on Banach spaces. There are many examples of bounded operators such as the following:

- Take $n \in \mathbb{N}$, if we take a matrix $A \in \mathbb{R}^{n\times n}$ and define it on the Banach space $X = (\mathbb{R}^n, ||\cdot||_p)$ then we have that $A : X \to X$ is a bounded operator and the operator norm coincides with the matrix norm $||A||_p$.
- Take $\Omega = \mathbb{R}$ or $\mathbb{T}$ and a multiplier $m \in L^\infty(\Omega)$. We can define the multiplication operator $T_m : L^p(\Omega) \to L^p(\Omega)$ by setting for $f \in L^p(\Omega)$, $Tf := mf$. It can be seen that T_m is bounded $||T_m f||_{L^p(\Omega)} = (\int_\Omega |m(x)f(x)|^p dx)^{\frac{1}{p}} \le (\int_\Omega ||m||^p_{L^\infty(\Omega)} |f(x)|^p dx)^{\frac{1}{p}} = ||m||_{L^\infty(\Omega)} (\int_\Omega |f(x)|^p dx)^{\frac{1}{p}} = ||m||_{L^\infty(\Omega)} ||f||_{L^\infty(\Omega)}$. Where we see that $||T_m||_{L^p(\Omega)} \le ||m||_{L^\infty(\Omega)}$.
- Take $1 \le p \le \infty$, define the operator $T : L^p(\mathbb{T}) \to \mathbb{C}$ by $Tf = \int_\mathbb{T} f dx$. We then have that $|Tf| = |\int_\mathbb{T} f dx| \le \int_\mathbb{T} |f| dx \overset{\text{Hölder's inequality}}{\le} ||f||_{L^p(\mathbb{T})} ||1||_{L^q(\mathbb{T})}$, where q is defined through $1 = \frac{1}{p} + \frac{1}{q}$. We can then see that $||T||_{L^p(\mathbb{T})\to\mathbb{C}} \le ||1||_{L^q(\mathbb{T})}$.
- Take $\Omega = \mathbb{R}$ or $\mathbb{T}$, the derivation operator $D : C^1(\Omega) \to C(\Omega)$ is defined by for $f \in C^1(\Omega)$ by $Tf := \partial_x f$. It can be seen that D is bounded namely $||Df||_{C(\Omega)} = ||\partial_x f||_{C(\Omega)} \le ||\partial f||_{C(\Omega)} + ||f||_{C(\Omega)} = ||f||_{C^1(\Omega)}$ Where we see that $||D||_{C^1(\Omega)\to C(\Omega)} \le 1$.

Another example of bounded operators are semigroups. These are essential to describing the convergence of the Zig-Zag process and will be discussed in the next section.

A more general case of bounded operators are unbounded operators. Terminology can become confusing here as every bounded operator is also an unbounded operator. However, there are unbounded operators which are not bounded. For this reason, 'unbounded operators' are often referred to as 'operators'.

Definition 2.2.6 (Unbounded operator)**.** *Let X, Y be Banach spaces and $Z \subseteq X$ be a linear subspace, furthermore define an operator $\mathscr{L} : Z \subseteq X \to Y$. In this case Z is referred to as the domain of the operator and we denote it by $D(\mathscr{L}) := Z$. An operator with such a domain is called an unbounded operator and is commonly denoted as a pairing $(D(\mathscr{L}), \mathscr{L})$.*

The notation of an unbounded operator and the definition of the domain can often be confusing. For example, we can have that $D(\mathscr{L}) = W^{1,2}(\Omega,\mu)$ and $X = L^2(\Omega,\mu)$ and thus that $\mathscr{L}$ is defined on $W^{1,2}(\Omega,\mu)$. However, the norm that is used on the elements of the domain of $\mathscr{L}$ is not the norm of $W^{1,2}(\Omega,\mu)$, but of $L^2(\Omega,\mu)$. As an example we can look at the operator $\partial_x : L^p(\mathbb{T},\mu) \to L^p(\mathbb{T},\mu)$, this operator can not be defined on the general $L^p(\mathbb{T},\mu)$ space since many functions in this space do not have a well defined (weak) derivative. However, if we define it as $(W^{1,p}(\Omega,\mu),\partial_x)$ the operator becomes well defined. The definition of an unbounded operator is very general and in order to add some structure to the operators such that more can be said about them the notion of densely defined and closed operators is introduced. Because of the general definition of the unbounded operators, we often require some extra structure. This gives the following definitions:

Definition 2.2.7 (Closed operator)**.** *Let X, Y be Banach spaces. The unbounded operator $\mathscr{L}f : D(\mathscr{L}) \subseteq X \to Y$ is called a closed operator if its graph $\{(f,\mathscr{L}f), f \in D(\mathscr{L})\}$ is closed in $X \times Y$, which is equivalent to saying that $D(\mathscr{L})$ is a Banach space when combined with the graph norm defined for $f \in D(\mathscr{L})$ by $||f||_{\mathscr{G}(\mathscr{L})} := ||f||_X + ||\mathscr{L}f||_Y$.*

Definition 2.2.8 (Densely defined operator)**.** *Let X, Y be Banach spaces. The unbounded operator $\mathscr{L}f : D(\mathscr{L}) \subseteq X \to Y$ is said to be densely defined if the domain $D(\mathscr{L})$ is dense in X.*

As a refresher for a set Z to be dense in a Banach space X, we require that for every $x \in X$ there is a sequence $(x_n)_{n\geq 1} \subset Z$ such that $\lim_{n\to\infty} ||x_n - x||_X \to 0$. This way we have that densely defined unbounded operators are almost completely defined on the full set of elements of X.

We illustrate these concepts with some examples.

- All of the bounded operators as defined by Definition 2.2.5 are closed when considering $D(\mathscr{L}) = X$ and they are densely defined operators because $D(\mathscr{L}) = X$.
- Take the Banach spaces $X, Y = C[0,1]$ and $D(\mathscr{L}) = C^1[0,1]$. Then define for $f \in D(\mathscr{L})$ the operator $\mathscr{L}f = \partial_x f$. We then see that $||f||_{\mathscr{G}(\mathscr{L})}$ is equal to the $||f||_{C^1([0,1])}$ norm and thus that $D(\mathscr{L})$ with the graph norm is a Banach space and thus by definition $\mathscr{L}$ is a closed operator. Furthermore, $C^1[0,1]$ is dense in $C[0,1]$ (as can be seen by the Weierstrass approximation theorem). So $(D(\mathscr{L}),\mathscr{L})$ is a closed and densely defined unbounded operator.
- Similar to the example above if we take $X, Y = L^p(\Omega,\mu)$ with $D(\mathscr{L}) = W^{1,p}(\Omega,\mu)$ then the operator $\mathscr{L}f = \partial_x f$ for $f \in D(\mathscr{L})$ is closed. Moreover, it is densely defined.

Densely defined operators are essential to defining the so-called adjoint operators. These adjoints can be seen as a generalization of the transpose of a real-valued matrix. For this thesis, we will only be interested in the definition of the adjoint on Hilbert spaces.

Definition 2.2.9 (Adjoint of an Unbounded operator defined on Hilbert spaces)**.** *Take H_1, H_2 to be Hilbert spaces and take a densely defined unbounded operator $\mathscr{L}f : D(\mathscr{L}) \subset H_1 \to H_2$. We then define the set $D(\mathscr{L}^*) \subset H_2$ to be the set of elements such that for a $g \in D(\mathscr{L}^*)$ there is an $h \in H_1$ such that for every $f \in D(\mathscr{L})$ we have*

$$\langle \mathscr{L}f, g\rangle_{H_2} = \langle f, h\rangle_{H_1}.$$

We then define the adjoint $\mathscr{L}^* : D(\mathscr{L}^*) \subset H_2 \to H_1$ of $\mathscr{L}$ by $\mathscr{L}^* g := h$.

The uniqueness of the elements $h \in H_1$ such that $\mathscr{L}^* g = h$ stems from the so-called Riesz representation theorem (Theorem I.3.4 [Con97]). Moreover, we have that the adjoint is closed and if $\mathscr{L}$ is closed we also have that $(\mathscr{L}^*)^* = \mathscr{L}$. Apart from matrices (for matrices $A \in \mathbb{C}^{n\times n}, n \in \mathbb{N}$ we have that for the elements A_{ij} for A and A^*_{ij} for A^* it holds that $A^*_{ij} = \overline{A}_{ji}$ for all $i, j \in \{1,2,...,n\}$) we can also define for other operators on Hilbert spaces such as ∂_x from the previous examples defined with domain $W^{1,2}(\mathbb{T})$. This operator has the adjoint such that $D(\partial_x^*) = D(\partial_x)$ and $\partial_x^* f = -\partial_x f$. Adjoints are for example useful for analysing the so-called spectrum of an operator which we will introduce later on in this section.

Next, we are introducing the concept of the resolvent and the spectrum of an operator. The resolvent can be seen as the set of values where an operator is very well behaved in the sense that it is boundedly invertible. For this reason, the resolvent is often referred to as the set of regular values.

Definition 2.2.10 (Boundedly invertible operator)**.** *An unbounded operator $\mathscr{L}f : D(\mathscr{L}) \subset X \to Y$ is said to be boundedly invertible if and only if there is a bounded operator $U : Y \to D(\mathscr{L}) \subset X$ such that for $f \in Y$ we have $\mathscr{L}Uf = f$ and for $g \in D(\mathscr{L})$ we have that $U\mathscr{L}g = g$. We denote this bounded inverse by $\mathscr{L}^{-1} := U$.*

Equivalently assuming that $(D(\mathscr{L},)\mathscr{L})$ is a closed operator we have that an unbounded operator is said to be boundedly invertible if and only if Ker($\mathscr{L}$) $= 0$ and Ran($\mathscr{L}$) $= Y$ (Proposition 1.15 of [Con97]).

Definition 2.2.11 (Resolvent set of an unbounded operator)**.** *The resolvent set $\rho(\mathscr{L})$ of an unbounded operator $\mathscr{L}f : D(\mathscr{L}) \subset X \to Y$ is defined by:*

$$\rho(\mathscr{L}) = \{\gamma \in \mathbb{C} : \mathscr{L} - \gamma I \text{ is a boundedly invertible operator}\}.$$

Where we denote I for the identity operator on $D(\mathscr{L})$ (so for $f \in D(\mathscr{L})$ we have $If = f$). We denote $\mathscr{L} - \gamma I$ as $\mathscr{L} - \gamma$. Moreover, we refer to $(\mathscr{L} - \lambda)^{-1}$ as the resolvent of $\mathscr{L}$ at $\gamma \in \rho(\mathscr{L})$ often denoted by $R(\gamma, \mathscr{L})$.

In particular if $\mathscr{L} : D(\mathscr{L}) \subset X \to Y$ is a closed operator we will get that $\gamma \in \rho(\mathscr{L})$ if and only if $\operatorname{Ker} \mathscr{L} - \gamma = 0$ and $\operatorname{Ran} \mathscr{L} - \gamma = D(\mathscr{L})$. We can then see that specifically the eigenvalues of an operator are not part of the resolvent.

Definition 2.2.12 (Eigenvalues and eigenvectors of an unbounded operator)**.** *Let X, Y be Banach spaces and $\mathscr{L}f : D(\mathscr{L}) \subset X \to Y$ be an unbounded operator. We call $f \in D(\mathscr{L})$ an eigenvector if $f \neq 0$ and there is an $\gamma \in \mathbb{C}$ such that $\mathscr{L}f = \gamma f$. γ is called an eigenvalue of the unbounded operator and we denote the set of all eigenvalues by $\sigma_p(\mathscr{L})$. This set is often referred to as the point spectrum.*

Note that an eigenvector must be in X and Y. So if the set of element X, Y have an empty intersection there can not be any eigenvalues/eigenvectors. So we see that $\sigma_p(\mathscr{L}) \subset \mathbb{C} \setminus \rho(\mathscr{L})$. However, unlike with matrices we in general do not have that $\mathbb{C} \setminus \rho \subset \sigma_p(\mathscr{L})$. The full set of all the $\gamma \in \mathbb{C}$ which are not in the resolvent set is called the spectrum.

Definition 2.2.13 (The spectrum of an unbounded operator)**.** *The spectrum $\sigma(\mathscr{L})$ of an unbounded operator $\mathscr{L}f : D(\mathscr{L}) \subset X \to Y$ is defined by:*

$$\sigma(\mathscr{L}) := \mathbb{C} \setminus \rho(\mathscr{L}).$$

So in particular we see $\sigma_p(\mathscr{L}) \subset \sigma(\mathscr{L})$. For general unbounded operators, there are many different parts of the spectrum that are classified for different (technical) reasons such as the continuous, residual, absolutely continuous, essential, discrete, compression, peripheral, singular, and approximate spectrum. Many of these definitions overlap for some points in the spectrum and for some such as the essential spectrum the definition differs between authors. For our purposes, we are mainly interested in the point spectrum/eigenvalues. The reason for this is because of a special class of operators for which the spectrum is almost entirely equal to this spectrum with the exception of $\gamma = 0$. This specific class of bounded operators are the compact operators which act more similar to how matrices work. In fact, all finite-dimensional operators are compact operators.

Definition 2.2.14 (Compact operator)**.** *Let X, Y be Banach spaces. An operator $T : X \to Y$ is called a compact operator if for every bounded sequence $(x_n)_{n \geq 1} \subset X$ we have that the sequence $(Tx_n)_{n \geq 1}$ contains a convergent subsequence.*

The spectrum of a compact operator $\mathscr{L}$ is countable and each non-zero $\gamma \in \sigma(\mathscr{L})$ is an eigenvalue ($\gamma \in \sigma_p(\mathscr{L})$) of finite algebraic multiplicity (Theorem 7.6 of [Con97]), which in particular means that $\dim \operatorname{Ker} \mathscr{L} - \lambda < \infty$. Furthermore, the only possible accumulation point of the spectrum is 0. A couple of examples of compact operators are:

- Finite-dimensional matrices $A \in \mathbb{C}^{N \times N}$, $N \in \mathbb{N}$ defined on any of the finite-dimensional norms. Their eigenvalues correspond to the roots of the polynomial $p(\gamma) := \det(A - \gamma)$. The amount of eigenvalues of a matrix is at most equal to N.
- The resolvents $(\partial_x - \gamma)^{-1}$ of the derivative operator $\partial_x : W^{1,2}(\mathbb{T}) \subset L^2(\mathbb{T}) \to L^2(\mathbb{T})$ is compact, this causes the spectrum of ∂_x to consist purely out of eigenvalues given by $i\mathbb{Z} = \{in : n \in \mathbb{Z}\}$.

The last example $(W^{1,2}(\mathbb{T}), \partial_x)$ is an operator that has a compact resolvent. This means that for all $\gamma \in \rho(\partial_x)$ we have that $(\partial_x - \gamma)^{-1}$ is compact. Interestingly if we have that for one γ the resolvent is compact that it then automatically follows that all of the elements are compact. This is because of the *resolvent identity* which holds for all $\gamma, \beta \in \rho(\mathscr{L})$:

$$(\mathscr{L} - \gamma)^{-1} - (\mathscr{L} - \beta)^{-1} = (\gamma - \beta)(\mathscr{L} - \gamma)^{-1}(\mathscr{L} - \beta)^{-1}.$$

Moreover, if $T : X \to Y$ is a bounded operator and $K : Y \to Z$ is a compact operator then KT is compact as well furthermore the sum of compact operators is again compact, so by rearranging the resolvent identity it can be seen that it is enough to prove that one element of the resolvent gives a compact operator in order to prove that all of the elements of the resolvent are compact.

Another example of compact operators are Hilbert-Schmidt operators which are special operators defined on a Hilbert Schmidt space.

Definition 2.2.15 (Hilbert-Schmidt operator). *For two Hilbert spaces H_1, H_2, an operator $T : H_1 \to H_2$ is Hilbert-Schmidt if for some (equivalently every) orthonormal basis $(h_n)_{n\geq 1}$ of H we have that the Hilbert-Schmidt norm $||T||_{HS} = \sum_{n\geq 1} ||Th_n||_H$ is finite.*

Hilbert-Schmidt operators are compact [Con97] and have the property $||T||_2 \leq ||T||_{HS}$, where $||T||_2$ is the operator norm as in Definition 2.2.5. For every Hilbert-Schmidt operator $T : H_1 \to H_2$ defined on two sigma-finite spaces $H_1 = (X_1, \mathscr{B}_1, \mu_1), H_2 = (X_2, \mathscr{B}_2, \mu_2)$ there is a kernel function $k \in L^2(X_1 \times X_2, \mathscr{B}_1 \times \mathscr{B}_2, \mu_1 \times \mu_2)$ such that for every $f \in H_1$ we have that $Tf(x) = \int_{X_2} k(x, y) f(y)\, d\mu_2(y)\, \mu_2 a.e.$ (Proposition 3.3.5 [Sun16]). Moreover, the adjoint T^* has a similar expression $T^* f(x) = \int_{X_1} \overline{k(y, x)} f(y)\, d\mu_1(y)\, \mu_1 a.e.$. Therefore the associated kernel of T^* is $\overline{k(y, x)}$ Furthermore, we have $||T||_{HS} = ||T^*||_{HS} = ||k||^2_{L^2(X_1\times X_2, \mu_1 \times \mu_2)}$ and $||T||_{HS} = \sum_{n\geq 0} s_n(T)$, where $s_n(T)$ are the singular values of T [Sun16]. These singular values correspond to the eigenvalues of $\sqrt{T^*T}$ counted with multiplicity.

To illustrate this, an example of a Hilbert-Schmidt operator is the Volterra operator $Tf(t) = \int_0^t f(s)ds$ with $f \in L^2([0, 1])$. For this operator, it can be seen that the kernel is defined such that $k(t, s) = 1_{s<t}$. We then calculate the Hilbert-Schmidt norm using the formula for the kernel $||T||^2_{HS} = ||k||^2_{[0,1]\times[0,1]} = 0.5$. Moreover, the eigenvalues of $\sqrt{T^*T}$ are $\gamma_n = \frac{2}{(2n+1)\pi}$, $n \geq 0$, [Thi16]. We can then confirm that $\sum_{n=0}^{\infty} \gamma_n^2 = \sum_{n=0}^{\infty} \frac{4}{(2n+1)^2\pi^2} = 0.5$. Moreover, the norm of the operator as Definition 2.2.5 is $||T||_2^2 = \frac{4}{\pi^2} \approx 0.4053$. The quality of the approximation $||T||_{HS} \leq ||T||_2$ depends on the distribution of the eigenvalues. Namely we have that $||T|| = \gamma_0$ so the difference is $||T||_{HS} - ||T||_2 = \sum_{n=1}^{\infty} \gamma_n^2 = \sum_{n=1}^{\infty} \frac{4}{(2n+1)^2\pi^2}$.

The approximation given by the Hilbert-Schmidt norm will be of use to us when creating lower bounds on the spectral gap of the generator of the Zig-Zag process. The generator is a component of the Semigroups as we will define in the following section.

2.3. Semigroups

One of our main interests is to describe the convergence properties of the Zig-Zag process. Because of the Markov property of the Zig-Zag process we can describe the evolution of the process in terms of a so-called strongly continuous semigroup (or C_0 semigroups) and its generator. These semigroups are sets of bounded operators $\{T(t) : X \to X, t \geq 0\}$ that describe the evolution of a system. To each Markov process, we can associate a semigroup. This object can be used to give a precise definition for the convergence of functions on the stochastic process. They are not unique to just the analysis of stochastic processes like the Zig-Zag process but are used for a wide range of partial differential equations. An extensive description of semigroups can be found in [EN00]. In this section, we will introduce basic notions related to semigroups that we will refer to in later sections. First, we introduce the definition of a strongly continuous semigroup.

Definition 2.3.1 (Strongly continuous semigroup). *Let $(X, ||\cdot||)$ be a Banach space. A family of bounded operators $\{T(t) : X \to X, t \geq 0\}$ is called a strongly continuous semigroup if and only if all of the following properties hold:*

1. *$T(0) = I$,*
2. *For all $t, s \geq 0 : T(t)T(s) = T(s + t)$ (Semigroup property),*
3. *For all $f \in X$ we have $||T(t)f - f|| \to 0$ as $t \downarrow 0$.*

There are numerous examples of strongly continuous semigroups:

- The translation semigroup defined for $f \in L^p(\mathbb{R})$ with $1 \leq p < \infty$ by $T(t)f := f(x + t)$.
- The multiplication semigroup defined for $f \in C_0(\mathbb{R}) := \{f : f$ is continuous and $\lim_{x\to\pm\infty} f(x) = 0\}$ by $T(t)f = e^{mt} f$ where $m \in C_0(\mathbb{R})$.
- The heat semigroup defined for $f \in L^2(\mathbb{R}^d)$, with $d \in \mathbb{N}$, by $(K_t * f)(s) := (4\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|s-r|^2}{4t}} f(r)\, dr$.

Due to the semigroup property it can be seen that a strongly continuous semigroup $\{T(t): X \to X, t \geq 0\}$ can be fully defined if for a $\gamma > 0$ the operators $\{T(t) : t \in [0,\gamma]\}$ are known. Specifically, we could take the limit of $\gamma \to 0$ in such a way that the information about the full group is still preserved. We do this by associating to every strongly continuous semigroup an infinitesimal operator $\mathscr{L}$ with a domain $D(\mathscr{L})$ that describes the infinitesimal evolution of the semigroup.

Definition 2.3.2 (Domain and infinitesimal generator)**.** *We define the domain $D(\mathscr{L})$ of the infinitesimal generator of a strongly continuous semigroup $\{T(t): X \to X, t \geq 0\}$ by all the functions $f \in X$ such that*

$$\lim_{t\to 0^+} ||\frac{T(t)f - f}{t}||$$

exists in X.
We then define the infinitesimal generator $\mathscr{L} : D(T) \to X$ of T by

$$\mathscr{L} f := \lim_{t\to 0^+} \frac{T(t)f - f}{t}.$$

We often say that an unbounded generator generates a strongly continuous semigroup if it associated to as an infinitesimal generator to it. Returning to our examples we can associate the following infinitesimal generators:

- For the translation semigroup it turns out that $D(\mathscr{L}) = W^{1,p}(\mathbb{R})$ and for $f \in D(\mathscr{L})$ we have $\mathscr{L} f = f'$.
- Corresponding to the multiplication semigroup we have that $D(\mathscr{L}) = \{f : mf \in C_0(\mathbb{R})\}$, $\mathscr{L} f = mf$ where $m \in C_0(\mathbb{R})$.
- Connected to the heat semigroup there is the corresponding generator $D(\mathscr{L}) = W^{2,2}(\mathbb{R}^d)$ and $\mathscr{L} f = \Delta f$.

For our analysis, the generator will be our main interest. It describes the infinitesimal evolution of our semigroup for $f \in D(\mathscr{L})$. The generator for our process arises because of the Markov property. To illustrate the meaning of the generator further we give an example. If we assume that the generator is a matrix-operator $A \in \mathbb{R}^{n\times n}$ we can define $y(t) = e^{tA}y(0)$, where $e^{tA} = \sum_{k=0}^{\infty} \frac{(At)^k f}{k!}$. This maps every initial value vector $y(0) \in \mathbb{R}^n$ to a vector $y(t) \in \mathbb{R}^n$ such that $y'(t) = Ay(t)$ holds for all $t \geq 0$. This relationship also holds for general generators, namely for all $t \geq 0$ and for all $f \in D(\mathscr{L})$ we have the relationship

$$\frac{d}{dt} T(t)f = \mathscr{L} T(t)f.$$

(Where the limit is to be taken in the corresponding Banach space $\frac{d}{dt}T(t)f := \lim_{h\to 0} \frac{T(t+h)f - T(t)}{h}$). In general $T(t)$ can be seen as a mapping that maps an initial value f to the solution of equation $\frac{d}{dt} f(t) = \mathscr{L} f(t)$.

For generators which are bounded operators $B \in \mathscr{B}(X)$ we have the following connection using functional calculus $T(t)f = e^{tB}f = \sum_{k=0}^{\infty} \frac{(Bt)^k f}{k!}$, which shows a more general connection to the exponential function. Using the boundedness of B we can examine the behaviour of the norm of $T(t)f$, namely we can find an exponential bound $||T(t)f|| = ||\sum_{k=0}^{\infty} \frac{B^k t^k}{k!} f|| \leq \sum_{k=0}^{\infty} \frac{||B||^k t^k}{k!} ||f|| = e^{||B||t}||f||$. Which implies that $||T(t)|| \leq e^{||B||t}$. It turns out that such an exponential bound also exists for general semigroups where B is not bounded.

Proposition 2.3.3 (Exponential bound)**.** *For a semigroup $\{T(t), t \geq 0\}$ we have that there are constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that*

$$||T(t)|| \leq Me^{\omega t}.$$

Proof. A proof of this can be found in [EN00] (Chapter 1 Proposition 5.5). □

This bound is somewhat different from the bound defined by $e^{||B||t}$ apart from that it also holds for semigroups which are defined on an infinite-dimensional Banach space. First, there is a constant M introduced next to the exponential, secondly, we have that the proposition states that $\omega \in \mathbb{R}$, note that we could have equivalently written down that $\omega \in \mathbb{R}_+$ but the main reason for writing the bound like this, is because it can be possible for certain semigroups that $\omega < 0$. However, if we want to find such ω for certain semigroups, we will have to have that $M > 1$. A semigroup that has $M = 1, \omega < 0$ is called a coercive semigroup and a semigroup that has such a bound with $\omega < 0$ only if $M > 1$ is called a hypocoercive semigroup. More explanation surrounding the background of hypocoercive semigroups can be found in section 2 of [Hér17] and section 3 of [Vil06]. The same issue also holds for the Zig-Zag process and other PDMPs [And+19].

As an example we can construct such a hypocoercive estimate in the case of generator being a diagonalize matrix $A \in \mathbb{R}^{n\times n}$ meaning $A = P\Lambda P^{-1}$. With Λ being a diagonal matrix of eigenvalues of A generating a semigroup $T(t)_{t\geq 0}$. Denote $s(A) := -\max_{1\leq i\leq n} \operatorname{Re}\lambda_i$. Take $y(0) \in \mathbb{R}^n$, we then get $||T(t)y(0)|| = ||\sum_{k=0}^{\infty} \frac{A^k t^k}{k!} y(0)|| = ||P\sum_{k=0}^{\infty} \frac{\Lambda^k t^k}{k!} P^{-1} y(0)|| \leq ||P||\cdot||\sum_{k=0}^{\infty} \frac{\Lambda^k t^k}{k!}||\cdot||P^{-1}y(0)|| \leq ||P||\cdot||P^{-1}||\cdot||y(0)||e^{-s(A)t}$. Which implies that $||T(t)|| \leq ||P||\cdot||P^{-1}||e^{-s(A)t}$. This in particular gives us $M = ||P||\cdot||P^{-1}||$ and $\omega = -s(A)$. This can be a much more useful bound than $e^{||B||t}$ since the spectral radius is always smaller than the norm of an operator. And ω can even become negative if we have that $s(A) > 0$. Which would shows that the solutions will decay towards 0 at an exponential rate. This makes $s(A)$ an important quantity and we can generalize its definition to infinite-dimensional Banach spaces.

Definition 2.3.4 (Spectral gap)**.** *Consider an unbounded operator $\mathscr{L} : Y \subseteq X \to X$, we define its spectral gap by:*

$$s(\mathscr{L}) := -\sup\{Re\gamma : \gamma \in \sigma_p(\mathscr{L})\} \setminus \{0\}.$$

Where $\sigma_p(\mathscr{L})$ is the set of eigenvalues/the point spectrum of $\mathscr{L}$ as defined in Definition 2.2.12.

This spectral gap will be of major importance in the further sections concerning convergence properties of the Zig-Zag process. Another specific pairing of the bound in Proposition 2.3.3 is the bound with $M = 1$, $\omega = 0$. The bound would then look like $||T(t)|| \leq 1$ which are referred to as contraction semigroups. For this, the following definition is of use.

Definition 2.3.5 (Dissipative operator)**.** *Let X be Banach spaces with the unbounded operator $\mathscr{L}f : D(\mathscr{L}) \subset X \to X$. We say that $(D(\mathscr{L}),\mathscr{L})$ is dissipative if and only if we have for all $\gamma > 0$ and $f \in Y$ that*

$$||(\lambda - \mathscr{L})f||_X \geq \gamma||f||_X.$$

It holds that if an operator is closed, dissipative and the range $(\mathscr{L} - \gamma)$ is dense in X for some $\gamma > 0$ that such an operator generates a strongly continuous contraction semigroup which is called the Lumer-Phillips Theorem (see Theorem 3.15 of [EN00]). This will be useful to prove that an operator $\mathscr{L}$ generates a strongly continuous semigroup. A convenient way of proving that an operator is dissipative makes use of a duality set. The duality set $J(x)$ for an element $x \in X$ is defined by

$$\mathscr{J}(x) = \{x^* \in X^* : \langle x, x^*\rangle = ||x||_X^2 = ||x^*||_X^2\}. \tag{2.1}$$

The duality set $\mathscr{J}(x)$ is always non-empty as a consequence of the Hahn-Banach Theorem. For example, if X is a Hilbert space we have by the Riesz Representation Theorem that $\mathscr{J}(x) = \{x\}$. We can then prove dissipativity through showing that for every $x \in D(\mathscr{L})$ we have that there is an element from the duality set $x^* \in \mathscr{J}(x)$ such that $\operatorname{Re}\langle x, x^*\rangle \leq 0$. Then Proposition 3.23 from [EN00] gives us that $\mathscr{L}$ is dissipative. They show that this can be seen as $||\gamma x - \mathscr{L}x||_X \geq |\langle \gamma x - \mathscr{L}x, x^*\rangle| \geq \operatorname{Re}\langle \gamma x - \mathscr{L}x, x^*\rangle \geq \gamma||x||_X$. Where the first inequality follows since for every element $x \in X$ we have $||x||_X = \sup_{x^*\in X^*, ||x^*||_{X^*}\leq 1} |\langle x, x^*\rangle|$. The second inequality is a simple property of complex numbers and the third property uses the assumption that $\operatorname{Re}\langle x, x^*\rangle \leq 0$. We will refer back to this duality set multiple times throughout the thesis.

Now that we have the mathematical background of the analysis of the generator associated with the Zig-Zag process we would like to more properly introduce the concept of how *MCMC* algorithms work and how the Zig-Zag process is defined which we will do in the following two sections.

2.4. MCMC algorithms

Let $d \geq 1$ and $\Omega = \mathbb{R}^d$ or $\mathbb{T}^d$. Often in various applications we would want to calculate integrals like the following

$$\mathbb{E}[f(X)] = \int_\Omega f(x)d\mu(x).$$

Where μ is a given probability distribution on Ω. One way in which we could approximate this integral is by creating random variables $X^1, X^2, \ldots X^n$ for some $n \in \mathbb{N}$ which are independent and have as their probability distribution μ. If we then assume that $\operatorname{Var}(f) < \infty$ (or $\int_\Omega (f(x) - \mathbb{E}[f(x)])^2 d\mu(x) < \infty$) we can make use of the law of large numbers. The law of large numbers states that if we have a set of independent variables coming from the same distribution that have finite variance that the average of those variables converges in probability to the expectation of the variables.

$$\frac{1}{n}\sum_{i=1}^{n} f(X^i) \xrightarrow{\mathbb{P}} \int_{\Omega} f(x)\,d\mu(x). \tag{2.2}$$

Note that we have introduced earlier that we are using the Hilbert space $L^2(\mathbb{T},\mu)$. We use this space as these are exactly the functions for which the variance is finite, moreover the use of Hilbert spaces is convenient as there are a lot of convenient properties to them for which many theorems exist. By approaching the approximation of integrals through formula (2.2) it becomes possible to calculate various types of integrals for a large class of functions. However, the convergence of this method can be slow especially for functions and probability measures in higher dimensions. Moreover, when μ is a complicated distribution we are not easily able to create the random variables $X^1, X^2, ...X^n$. Instead, we will create random variables which have approximately these distributions.

One way of doing this is by creating independent Markov processes $(X_t^1)_{t\geq 0}, (X_t^2)_{t\geq 0}, ...(X_t^n)_{t\geq 0}$ which are ergodic and have as their stationary measure μ as described in Definition 2.1.3 and 2.1.2. Then if we simulate these processes correctly they can converge in distribution to the desired target distribution μ such that the approach of formula (2.2) can work.

However, as mentioned before the convergence of the law of large numbers in this way can be very slow. So for our current construct to work we will have to make n very large which would mean that we would have to simulate a lot of processes $(X_t^1)_{t\geq 0}, (X_t^2)_{t\geq 0}, ...(X_t^n)_{t\geq 0}$, creating so many independent processes is in general computationally expensive. However, because of the Markov property, it turns out that it is possible for some Markov processes to make this approximation using only a single one of these stochastic processes $(X_t)_{t\geq 0}$. For these processes, the following convergence in distribution holds by taking a positive real number $b > 0$.

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}}\left(\sum_{k=1}^{N} f(X_{bk}) - \mathbb{E}[f(X)]\right) \xrightarrow{d} N(0,\sigma_{b,f}^2). \tag{2.3}$$

Where we have $\sigma_{b,f}^2 = \tau \mathrm{Var}_\mu(f)$ and $\tau = \sum_{i=0}^{\infty} \mathrm{corr}(X_0, X_{bi})$ and on the right-hand side is a normal distribution with variance $\sigma_{b,f}^2$. A more detailed description of how this convergence is established and what properties for the Markov process allows this can be found in various MCMC literature (for example in section 5 of [RR04]).

This way it can be seen that the convergence rate towards the average in (2.2) is proceeding with rate $\frac{\sigma_{b,f}^2}{\sqrt{N}}$. To get a better view of how fast the process converges explicitly, we could analyse $\sigma_{b,f}^2$ and specifically, we will examine the case when $b \to 0$. Which would change the summation of formula (2.3) into an integral in the time variable. Assuming that $X(0)$ is distributed according to the target measure μ, we then have under certain conditions that the following central limit theorem holds:

$$\lim_{T\to\infty} \frac{1}{\sqrt{T}} \int_0^T f(X_t) - \mathbb{E}[f(X)]\,dt \xrightarrow{d} N(0,\sigma_f^2). \tag{2.4}$$

Here σ_f^2 is called the *asymptotic variance* and we have convergence to a normal distribution with its variance equal to σ_f^2. This is also practical use as it shows that we can make b arbitrarily small and still have that our estimation for the integral $\mathbb{E}[f(X)]$ converges. Moreover, we can use this asymptotic variance to give an estimate of how fast the calculation of the integrals/expectations will be.

A specific Markov process $(X_t)_{t\geq 0}$ for which all of these properties hold is the Zig-Zag process.

2.5. PDMPs and the Zig-Zag process

In recent years, there have been a significant amount of developments and interest towards Piecewise Deterministic Markov Processes (PDMPs) within the MCMC field. Namely, because these processes promise lower computational costs and faster convergence due to a property called non-reversibility [SGS12] [HHS93].

Definition 2.5.1 (Time reversibility). *Assume a process $(X_t)_{t\geq 0}$ has a stationary distribution. This means that X_t has a distribution ν such that if X_0 is distributed according to ν that the distribution of X_t does not change over time (which for the Zig-Zag process by implies that the stationary distribution is equal to the distribution ν in 2.1.2). Moreover, if we then extend the process to the whole number line $\{X_t, t\in\mathbb{R}\}$, then $(X_t)_{t\geq 0}$ is time reversible if the distribution of the process $\{X_t, t\in\mathbb{R}\}$ is the same as $\{X_{-t}, t\in\mathbb{R}\}$.*

An example of a reversible process is the Metropolis-Hastings algorithm which is a well-known MCMC algorithm that is used for similar purposes such as the Zig-Zag algorithm for sampling from complicated high-dimensional distributions. It has been shown that such an algorithm has been outperformed in its asymptotic variance [SGS12] (as described surrounding formula (2.4)) and convergence to equilibrium [HHS93] (as described in Definition 2.1.2) by non-reversible Markov chains.

PDMPs are defined by deterministic dynamics with an underlying stochastic process. Different non-reversible PDMPs have been defined such as the Boomerang Sampler [Bie+20], and more important to our discussion the Bouncy Particle Sampler (BPS) [BVD18] and the Zig-Zag sampler [BFR19]. The Bouncy Particle Sampler has many properties in common with the Zig-Zag sampler, this is because the processes come from the same process which is a variant of the telegraph process [Kac51].

Let $d \in \mathbb{N}$, $\Omega = \mathbb{T}^d$ or $\Omega = \mathbb{R}^d$, and define the *potential function* $U : \Omega \to \mathbb{R}$. Assuming that $\int_\Omega e^{-U(x)}\,dx < \infty$, the Zig-Zag algorithm is used to sample from the *target measure* μ defined for $A \in \mathscr{B}(\Omega)$ by $\mu(A) = \int_A e^{-U(x)}\,dx$. Furthermore, we will define the product measure $\nu(A \times \{\theta\}) = \mu(A)$ for $\theta \in \{1,-1\}$ and $A \in \Omega$. It can be shown that under various assumptions that the Zig-Zag process is ergodic and that it has the stationary distribution with measure ν [BRZ19]. Moreover, this measure can be normalized such that it becomes a probability distribution.

Lastly, in order to define the Zig-Zag process we make use of an inhomogeneous Poisson processes. An inhomogeneous Poisson process with switching rate λ is a stochastic process $(X_t)_{t\geq 0}$ such that for any two timestamps $a, b \in [0,\infty)$ with $b > a$ we have that $X_b - X_a$ is distributed according to Poisson$(\int_a^b \lambda(s)\,ds)$. Where we have that the distribution Poisson(γ) is a discrete distribution with probability mass function $\mathbb{P}(Y = n) = \frac{\gamma^k e^{-\gamma}}{k!}$ with $n \in \mathbb{N} \cup \{0\}$ (where Y is distributed according to the Poisson distribution with parameter γ).

The Zig-Zag process is defined in terms of two processes $(X_t, \Theta_t)_{t\geq 0}$, where $(\Theta_t)_{t\geq 0}$ is the velocity process and $(X_t)_{t\geq 0}$ is the positional process. The positional component is fully defined in terms of the velocity component by $X_t := X_0 + \int_0^t \Theta_s\,ds$. Where we will have that $\Theta_t \in \{+1,-1\}^d$ and $X_t \in \Omega$. The velocity process will be updated coordinate wise based on the inhomogeneous Poisson processes with switching rates $\lambda_i(x,\theta) := (\theta \partial_{x_i} U(x))_+ + \lambda_0(x)$ with $(x)_+ := \max(x,0)$. When an update for the j-th component triggers the corresponding component of the velocity process will flip, meaning we will have that the new velocity will be defined by:

$$(F_j\Theta)_i := \begin{cases} \Theta_i & \text{if } i \neq j, \\ -\Theta_i & \text{if } i = j. \end{cases}$$

So the process can be constructed as follows assuming the initial positions X_0 and Θ_0 are given then for $i \in \mathbb{N}$ we have:

- Define $\xi^i(t) := X^{i-1} + t\Theta^{i-1}$.
- For $j \in \{1,2,...,d\}$, we define τ_j^i to be distributed independently according to

$$\mathbb{P}(\tau_j^i \geq t) = \exp(-\int_0^t \lambda_i(\xi^i(s), \Theta^{i-1})ds).$$

- Then we define $k := \operatorname{argmin}_{j\in\{1,2,...d\}} \tau_j^i$.
- Let $T^i := T^{i-1} + \tau_k^i$.
- Then we define $X^i := \xi^i(T^i)$.
- Finally we have $\Theta^i := F_k\Theta^{i-1}$.

Then the final process can be created for all time $t \geq 0$ by

$$(X(t), \Theta(t)) := (X^i + \Theta^i(t - T^i), \Theta^i), \text{ for } t \in [T^i, T^{i+1}), i \in \mathbb{N} \cup \{0\}.$$

In recent years there has been extensive research done towards the Zig-Zag process because of its promising properties. It has been shown that under certain assumptions that the Zig-Zag process is exponentially ergodic [BRZ19], which means that there is a function $f : \Omega \times \{+1,-1\} \to \mathbb{R}_+$ and a constant $c > 0$ such that $\forall (x,\theta) \in \Omega \times \{+1,-1\}$ we have that for $t \geq 0$

$$||P((X_t, \Theta_t) \in \cdot | (X_0, \Theta_0) = (x,\theta)) - \mu(\cdot)||_{TV} \leq f(x,\theta)e^{-ct}.$$

Which in particular show that the Zig-Zag process is ergodic as in Definition 2.1.2. Furthermore, under certain assumptions on the potential and refreshment rate we have that a central limit theorem holds: If we have examine the case $\Omega = \mathbb{R}$ and if (X_0, Θ_0) is distributed according to the stationary distribution [BD17] then we have the central limit theorem holds as described in the previous section

$$\lim_{T\to\infty} \frac{1}{\sqrt{T}} \int_0^T f(X_t) - \mathbb{E}[f(X)]dt \xrightarrow{d} N(0, \sigma_f).$$

Moreover, for the the asymptotic variance has a concrete formula for specific class of functions g given by

$$\sigma_g^2 = 2\int_{\mathbb{R}} (|U'(x)| + 2\lambda_0)\psi(x)^2 d\mu(x) \tag{2.5}$$

where $\psi = \frac{1}{2}(\phi(x,-1) - \phi(x,1))$ and ϕ is defined through $\mathscr{L}\phi = \mathbb{E}[g] - g$. Furthermore, it has been shown in [LW21] and [And+19] that for $\Omega = \mathbb{R}^d$ that the Zig-Zag process is *hypocoercive*, which means that the Zig-Zag process converges exponentially in $L^2(\mathbb{R}^d, \nu)$, that is there are constants $M > 1, \kappa > 0$ such that

$$||P_t f - \mathbb{E}[f]||_{L^2(\nu)} \leq Me^{-\kappa t}||f - \mathbb{E}[f]||_{L^2(\nu)}. \tag{2.6}$$

holds for all $f \in L^2(\mathbb{R}^d, \nu)$. For this we define the following terms:

$$P_t f := \mathbb{E}[f(X_t, \Theta_t)|(X_0, \Theta_0) = (x, \theta)],$$
$$\mathbb{E}[f] = \int_{\Omega^d} f(x,\theta)d\nu(x,\theta), \text{ and}$$
$$||f||^2_{L^2(\nu)} = \int_{\Omega^d} |f(x,\theta)|^2 d\nu(x,\theta).$$

It turns out that $(P_t)_{t\geq 0}$ is a set of bounded operators and more specifically a strongly continuous semigroup as described in section 2.3. The hypo-coercivity property described in formula (2.6) is a special case of Proposition 2.3.3. Moreover, the hypo-coercivity is necessary to make such an exponential bound with a positive convergence rate (Meaning that the restriction $M > 1$ is required). The previous research by [LW21] and [And+19] towards this property has given bounds on M and κ which are not proven to be sharp or of practical use. Instead, there was an attempt done at analysing the spectral gap of the generator [BV21] as was discussed in section 2.3. Here it was shown that for $d = 1$ and U having one maximum (unimodality) that the κ in the hypo-coercivity equation (2.6) is equal to the spectral gap as defined in Definition 2.3.4 of the generator associated to the semigroup. Therefore in order to bound the convergence rate we need to bound κ.

Although much of the analysis towards the Zig-Zag process has been done on $\mathbb{R}$ it is likely that many properties will be similar and that similar problems/benefits will arise. One such problem is that in formula (2.5) for the asymptotic variance it can be seen that the variance increases as we increase the values of λ_0. This would mean that approximating expectations $\mathbb{E}[f(X)]$ as described in formula (2.4) will take longer as λ_0 is increased. However, this formula only holds if the process is already distributed according to the stationary distribution. The stationary distribution is reached at an exponential rate as is described by the convergence of the semigroup in formula (2.6). However, as we shall see later on for many potentials increasing λ_0 can actually increase the convergence rate of the associated semigroup, meaning that the rate at which the process converges in distribution to this stationary distribution is increased.

In practice, this means that there should be a trade-off in how high this refreshment rate should be: Keeping the refreshment rate λ_0 too small will cause the convergence to the stationary distribution to be smaller such that it will take longer before formula (2.4) can apply. Keeping the refreshment rate λ_0 too large can increase the convergence rate to the stationary distribution significantly, but once this distribution has been reached the central limit theorem of formula (2.4) will have a higher variance meaning that the approximation of $\mathbb{E}[f(X)]$ will take longer. Therefore in practice the refreshment rate λ_0 can be taken as a time-dependent variable that is changed over time as the Zig-Zag process comes closer to the stationary distribution in the sense of (2.6), such that the convergence for (2.4) can be increased. However, in this thesis, we will only look at a constant (over time) refreshment rate λ_0.

We will mainly be concerned with the value of κ in formula (2.6) and we will examine how this value behaves based on properties of U and λ_0. Moreover, we will show how other elements of the spectrum of P_t behave based on the spectrum of the generator $\mathscr{L}$ which we will define in the following chapter.

2.6. Conventions and notations

Throughout the thesis, we make use of some notation of which its definition might be hard to find. Therefore we made the following overview.

$\mathbb{N}$ — The positive integers excluding 0.
$\mathbb{Z}$ — The set of all the integers.
$\mathbb{R}$ — The field of real numbers.
$\mathbb{R}_+$ — The non-negative real numbers $[0,\infty)$.
$\mathbb{C}$ — The field of complex numbers.
$\mathbb{T}$ — The torus $[-\pi,\pi]$ where opposite sides are identified, such that for function f defined on $\mathbb{T}$ we have $f(x) = f(y)$ for $x - y = 2\pi n$ with $n \in \mathbb{Z}$
f^+, f^- — For functions $f : \mathbb{T} \times \{+1,-1\} \to \mathbb{C}$ we denote $f^+(x) := f(x,1)$, $f^-(x) := f(x,-1)$.
$C(\Omega)$ — The set of continuous functions $f : \Omega \to \mathbb{C}$ usually with associated norm $||f||_\infty = \sup_{x\in\Omega} |f(x)|$.
$C^1(\Omega)$ — The set of continuously differentiable functions $f : \Omega \to \mathbb{C}$.
$L^2(\mathbb{T},\mu)$ — The set of functions $f : \mathbb{T} \to \mathbb{C}$ such that $||f||^2_{L^2(\mathbb{T},\mu)} = \int_\mathbb{T} |f(x)|^2 \, d\mu(x) < \infty$, with $d\mu(x) = e^{-U(x)} dx$.
$L^2(\mathbb{T},\nu)$ — The set of functions $f : \mathbb{T} \times \{+1,-1\} \to \mathbb{C}$ such that $\sum_{\theta\in\{+1,-1\}} \int_\mathbb{T} |f(x,\theta)|^2 \, d\mu(x) < \infty$.
$\Omega \setminus \{u\}$ — All the elements of the Hilbert space Ω which are orthogonal to u.
For example, $L^2(\mathbb{T},\mu) \setminus \{e^U\}$ refers to $\{f \in L^2(\mathbb{T},\mu) : \int_\mathbb{T} f(x)e^{U(x)} \, d\mu(x) = 0\}$.
$\mathscr{L}$ — Generator of the Zig-Zag process defined on $L^2(\mathbb{T},\nu)$ with domain $W^{1,2}(\mathbb{T},\nu)$.
$\mathscr{L}_{C^1}$ — Generator of the Zig-Zag process defined on $C(\mathbb{T} \times \{+1,-1\})$ with domain $C^1(\mathbb{T} \times \{+1,-1\})$.
$\mathscr{L}_0, \mathscr{L}_{C^1_0}$ — Generator of the Zig-Zag process defined on the respective domains as above with mean-zero meaning that for all these functions $f : \mathbb{T} \to \mathbb{C}$ we have $\int_\mathbb{T} f(x) d\nu(x) = 0$.
$\kappa, s(\mathscr{L})$ — Interchangeably used for the spectral gap of $\mathscr{L}$.

For the L^2 spaces as defined above we define an associated inner product such that these spaces become Hilbert spaces. For functions $f, g \in L^2(\mathbb{T},\mu)$ we define the inner product $\langle f, g\rangle := \int_\mathbb{T} f(x)\overline{g}(x) \, d\mu(x)$. Then for functions $h, k \in L^2(\mathbb{T},\nu)$ we define the inner product $\langle h, k\rangle = \langle h^+, k^+\rangle + \langle h^-, k^-\rangle$.

3

Spectral analysis of the Zig-Zag process generator with general refreshment rate

In order to examine the convergence rate of the Zig-Zag process we examine the spectrum of the operator. Previous research towards the spectrum of the Zig-Zag process proved many interesting properties and behaviour on the spectrum if $\Omega = \mathbb{R}$ [BV21]. Most interestingly formulas for the eigenvectors and eigenvalues were formulated in the case that the refreshment rate λ_0 is equal to 0 and the potential function U is unimodal. Unfortunately the values of these eigenvalues were described in terms of the roots of a complicated function and it was not possible to create any estimates on the spectral gap κ as defined in Definition 2.3.4 with an associated potential U unless these roots were calculated explicitly using root-finding algorithms. A similar analysis to this previous work is done in Chapter 4 of this thesis.

Analysing the spectrum on $\Omega = \mathbb{R}$ introduces more complications due to extra conditions that have to be put on the potential U (Specifically the rate of growth of the potential as $|x| \to \infty$). These growth conditions are not a problem on the torus due to the boundedness of the domain. Furthermore, for the torus, there has been a full characterisation of the spectrum done in the case that U and λ_0 are constant functions for $\mathbb{T}$ [MM12]. Such an analysis would not be doable on $\mathbb{R}$ due to $e^{-U(x)} \propto 1 \notin L^1(\mathbb{R})$ which is required for U to define a probability distribution. This makes the case of $\mathbb{T}$ easier to analyse as fewer restrictions are applied to the functions for which the analysis is doable.

In order to analyse the case where the refreshment rate λ_0 is non-zero and bounded ($\lambda_0 \in L^\infty(\mathbb{T})$) and the potential U is multi-modal an approach using block operators [Tre08] was proposed. This approach is popular as it allows for complicated operators to be analysed based on more well-behaved operators which together define the full operator. We will introduce the definition of block operators in section 3.2. Moreover, it turns out that for the block operators there is an associated operator family that is called the Schur complement $S_1(\gamma)$ that is defined for $\gamma \in \mathbb{C}$. The Schur complement relates the elements of the point spectrum $\gamma \in \sigma_p(\mathscr{L})$ to kernel of $S_1(\gamma)$. In particular the Schur complement consists out of self-adjoint operators for $\gamma \in \mathbb{R} \setminus \{0\}$. This gives more structure to these operators and allows us to ultimately create a bound that is related to eigenvalues of another operator BB^* that is related to a physical process called the overdamped Langevin equation which is a more known and studied equation.

In this section, we analyse the spectrum of the generator $\mathscr{L}$ of the Zig-Zag process for bounded refreshment rates $\lambda_0 \in L^\infty(\mathbb{T})$ and Lipschitz potentials U. We will calculate the solution to the Poisson equation $\mathscr{L} f = g$, describe bounds on the eigenvalues using block operators and we will create two lower bounds on the spectral gap κ.

3.1. Formulation of the eigenvalue problem

We examine the Zig-Zag process on the one-dimensional torus with a continuous periodic potential $U(x) : \mathbb{T} \to \mathbb{R}$ and a refreshment rate $\lambda_0 : \mathbb{T} \to [0,\infty)$. We will then define numerous assumptions that we will use throughout the following sections.

(A1) **Almost everywhere existence and boundedness of the derivative of the potential** $U' \in L^\infty(\mathbb{T})$.

(A2) **Boundedness of the refreshment rate** $\lambda_0 \in L^\infty(\mathbb{T})$.

(A3) **Smoothness of the potential** $U \in C^1(\mathbb{T})$.

(A4) **Continuity of the refreshment rate** $\lambda_0 \in C(\mathbb{T})$.

(A5) **Bimodality** There is an $x_0 \in [-\pi, \pi]$ such that for $x \in [-\pi, x_0] : U'(x) \leq 0$, and for $x \in [x_0, \pi] : U'(x) \geq 0$.

(A6) **No refreshment** $\lambda_0 = 0$.

Assumption (A1) could have been replaced by $U \in W^{1,\infty}(\mathbb{T})$ or by having U be a Lipschitz function, that is a function such that there is a constant $K > 0$ such that $|f(x) - f(y)| \leq |x - y|$. We will always assume that (A1) and (A2) hold in this chapter. Assumption (A3), (A4),(A5) and (A6) will only be required in Chapter 4.

Next, we will define the generator for the Zig-Zag process. For this, we will introduce some notation. First, we define the square-integrable functions with zero mean with respect to the target measure ν

$$L^2_0(\mathbb{T}, \nu) := \{f \in L^2(\mathbb{T}, \nu) : \int_{\mathbb{T}} f \, d\nu = 0\}.$$

Respectively we denote the Sobolev space with zero mean by

$$W^{1,2}_0(\mathbb{T}, \nu) := L^2_0(\mathbb{T}, \nu) \cap W^{1,2}(\mathbb{T}, \nu).$$

We will use the following notation for the flip operator $F : L^2_0(\mathbb{T}, \nu) \to L^2_0(\mathbb{T}, \nu)$ for $f(x, \theta) \in L^2_0(\mathbb{T}, \nu)$ by writing:

$$(Ff)(x, \theta) := f(x, -\theta). \tag{3.1}$$

We can then define the generator of the Zig-Zag process.

Definition 3.1.1 (Generator of the Zig-Zag process)**.** *The generator for the Zig-Zag process*

$$\mathscr{L}_0 : W^{1,2}_0(\mathbb{T}, \nu) \subset L^2_0(\mathbb{T}, \nu) \to L^2_0(\mathbb{T}, \nu),$$

is an unbounded operator defined for all $f \in D(\mathscr{L}) := W^{1,2}_0(\mathbb{T}, \nu)$ *by*

$$\mathscr{L}_0 f := \theta \partial_x f + \lambda(x, \theta)(Ff - f). \tag{3.2}$$

where $\lambda(\theta, x) = (\theta U(x))_+ + \lambda_0(x)$ *and* $(\theta U(x))_+ = \max(\theta U(x), 0)$

Normally we would define the generator to be defined on the space $\mathscr{L} : W^{1,2}(\mathbb{T}, \nu) \subset L^2(\mathbb{T}, \nu) \to L^2(\mathbb{T}, \nu)$. However, for our analysis, we need to remove the constant functions such that analysing the spectral gap κ becomes more convenient. Moreover, not much information is lost as it can be seen that for all the constant functions we have $\mathscr{L}1 = 0$. So 1 is an eigenvector for the eigenvalue 0 which is not relevant for the definition of the spectral gap.

The generator $\mathscr{L}_0$ is well defined since $U', \lambda_0 \in L^\infty(\mathbb{T})$ and thus $\lambda(x, \theta)(Ff - f) \in L^2(\mathbb{T}, \nu)$ and due to the definition of the Sobolev space we also have that the derivative of the functions exists and that $\theta \partial_x f \in L^2(\mathbb{T}, \nu)$ and thus $\mathscr{L}_0 f \in L^2(\mathbb{T}, \nu)$. Furthermore, by using the relationship $(\lambda(x, 1) - \lambda(x, -1)) = U'(x)$ we have for $f \in D(\mathscr{L})$ that

$$\begin{aligned}
\int_{\mathbb{T}} \mathscr{L} f \, d\nu &= \int_{\mathbb{T}} (\mathscr{L} f)^+ \, d\mu + \int_{\mathbb{T}} (\mathscr{L} f)^+ \, d\mu \\
&= \int_{\mathbb{T}} \partial_x (f^+ - f^-) + (\lambda(x, 1) - \lambda(x, -1))(f^- - f^+) \, d\mu \\
&= \int_{\mathbb{T}} \partial_x (f^+ - f^-) - U'(f^+ - f^-) \, d\mu = \\
&= \int_{\mathbb{T}} (\partial_x (f^+ - f^-) - U'(f^+ - f^-)) e^{-U(x)} \, dx \\
&= \int_{\mathbb{T}} \partial_x (e^{-U(x)} (f^+ - f^-)) \, dx \\
&= e^{-U(\pi)} (f^+ - f^-)(\pi) - e^{-U(-\pi)} (f^+ - f^-)(-\pi) \\
&= 0,
\end{aligned}$$

thus $\int_{\mathbb{T}} \mathscr{L}_0 f \, d\nu = 0$ and thus $\mathscr{L}_0 f \in L^2_0(\mathbb{T},\nu)$. In this derivation we used the product/Leibniz rule for weak derivatives to assert that $\partial_x(e^{-U(x)}(f^+ - f)) = (\partial_x(f^+ - f^-) - U'(f^+ - f^-))e^{-U(x)}$ as $e^{-U(x)} \in W^{1,\infty}(\mathbb{T})$ due to assumption (A1) and using Proposition 4.1.17 from [Web18] the product rule can be applied. Moreover, we are able to evaluate the functions f^+, f^- point-wise due to the Sobolev embedding. This shows why assumption (A1) was chosen, it allows us to make sure that the generator is well-defined.

As described in section 2.3 the generator describes the infinitesimal evolution of the semigroup of the Zig-Zag process. This equation consists out of two parts. $\theta\partial_x f$ describes the change in position over time of the Zig-Zag process without any changes in velocity. If we would have that $\lambda(x,\theta) = 0$ then the generator would describe a semigroup of a particle that moves at a constant unidirectional velocity, so for all $T(t)f(x,\theta) = f(x+t,\theta)$. The $\lambda(x,\theta)(Ff - f)$ part of the generator refers to how the velocity is updated over time. It specifically refers to how the velocity component θ is updated based on an inhomogeneous Poisson process with rate $\lambda(x,\theta)$ as described in section 2.5.

3.2. Block operators

The generator of the Zig-Zag process can be expressed as a block operator. A block operator equivalently describes the mapping using a matrix representation of the operator. It is build by splitting up the original Hilbert space into two other Hilbert spaces using a direct sum of Hilbert spaces such as $L^2(\mathbb{T},\nu) = L^2(\mathbb{T},\mu) \bigoplus L^2(\mathbb{T},\mu)$. Specifically we examine each element $f \in L^2(\mathbb{T},\nu)$ as two elements $f^+, f^- \in L^2(\mathbb{T},\mu)$, where we denote $f^+(x) := f(x,1)$ and $f^-(x) := f(x,-1)$. Moreover, we can define in a similar manner $W^{1,2}(\mathbb{T},\nu) = W^{1,2}(\mathbb{T},\mu) \bigoplus W^{1,2}(\mathbb{T},\mu)$. We can then see that the domain of the operator can be written as $D(\mathscr{L}_0) = \{(f^+, f^-) \in W^{1,2}(\mathbb{T},\mu) \bigoplus W^{1,2}(\mathbb{T},\mu) : \int_{\mathbb{T}} f^+(x)d\mu(x) + \int_{\mathbb{T}} f^-(x)d\mu(x) = 0\}$. We can then describe the generator from equation (3.2) in terms of a 'block operator' in the following way

$$\mathscr{L}_0 \begin{bmatrix} f^+ \\ f^- \end{bmatrix} = \begin{bmatrix} \partial_x - \lambda(x,1) & \lambda(x,1) \\ \lambda(x,-1) & -\partial_x - \lambda(x,-1) \end{bmatrix} \begin{bmatrix} f^+ \\ f^- \end{bmatrix}.$$

This notation can be interpreted using the same vector-matrix multiplication for normal finite-dimensional matrices. This representation of our generator unfortunately does not allow us to separate the domain $D(\mathscr{L}_0)$ into two domains due to the restriction $\int_{\mathbb{T}} f^+(x)d\mu(x) + \int_{\mathbb{T}} f^-(x)d\mu(x) = 0$. and as such the entries in the matrix can not be defined separately. In order for our generator $\mathscr{L}_0$ to be represented as a proper block operator it will have to fulfil the following definition.

Definition 3.2.1 (Unbounded block operator)**.** *Let $\mathscr{H}_1, \mathscr{H}_2$ be Banach spaces. Furthermore, define the unbounded operators $A : D(A) \subset \mathscr{H}_1 \to \mathscr{H}_1$, $B : D(B) \subset \mathscr{H}_2 \to \mathscr{H}_1$, $C : D(C) \subset \mathscr{H}_1 \to \mathscr{H}_2$, and $D : D(D) \subset \mathscr{H}_2 \to \mathscr{H}_2$. We then define the block operator $\mathscr{B}$ on the Banach product spaces $\mathscr{B} : D(\mathscr{B}) \subset \mathscr{H}_1 \bigoplus \mathscr{H}_2 \to \mathscr{H}_1 \bigoplus \mathscr{H}_2$ with $D(\mathscr{B}) := (D(A) \cap D(C)) \bigoplus (D(B) \cap D(D))$ for $(f,g) \in D(\mathscr{A})$ by*

$$\mathscr{B} \begin{bmatrix} f \\ g \end{bmatrix} := \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} Af + Bg \\ Cf + Dg \end{bmatrix}.$$

Where this matrix representation should be interpreted similarly to normal matrix-vector multiplication as can be seen on the right-hand side of the above equation. An operator in this form is called a block operator.

Approaches to analysing operators in terms of block operators for analysing hypocoercivity have been done before [Ber+21]. However, such an approach for the Zig-Zag process and its spectrum have not been done before.

Before we get to analyse the generator as defined in Definition 3.1.1 we first want to rewrite generator $\mathscr{L}_0$ to separate its domain and make the resulting operator more symmetric as a block operator. With a properly separated domain, we can then analyse each entry of the block operator as a separate unbounded operator. Since the generator has a domain that is not separable we will need to transform our generator $\mathscr{L}_0$ into another operator such that the spectrum is the same. This can be done by making use of unitary transformations and we will make use of a simple unitary operator T and create a new operator $\mathscr{A} = T^* \mathscr{L}_0 T$. This new operator $\mathscr{A}$ will thus have the same spectrum as $\mathscr{L}_0$ (so $\sigma(\mathscr{A}) = \sigma(\mathscr{L}_0)$). This new representation might feel a bit arbitrary so we also show how this representation could otherwise be found.

To do this we explicitly write out the eigenvalue problem of the generator of the Zig-Zag process defined in equation (3.2) which can be written into

$$\partial_x f^+ + \lambda(x,1)(f^- - f^+) = \gamma f^+,$$
$$-\partial_x f^- - \lambda(x,-1)(f^- - f^+) = \gamma f^-.$$

The main inspiration for finding this unitary transformation comes from the identities $\lambda(x,1)+\lambda(x,-1) = |U(x)|+2\lambda_0(x)$ and $\lambda(x,1)-\lambda(x,-1) = U'(x)$. We can apply these identities by subtracting and adding the two equations that we have. First, by subtracting these two equations we get:

$$\partial_x(f^+ + f^-) + (\lambda(x,1)+\lambda(x,-1))(f^- - f^+) = \gamma(f^+ - f^-).$$

Using $\lambda(x,1)+\lambda(x,-1) = |U(x)|+2\lambda_0$ we can write this equation as

$$\partial_x(f^+ + f^-) + (|U(x)|+2\lambda_0)(f^- - f^+) = \gamma(f^+ - f^-). \tag{3.3}$$

Likewise by adding up the two equations we get:

$$\partial_x(f^+ - f^-) + (\lambda(x,1)-\lambda(x,-1))(f^- - f^+) = \gamma(f^+ + f^-).$$

Here we can use the identity $\lambda(x,1)-\lambda(x,-1) = U'(x)$ such that we can rewrite the equation into the following form

$$-\partial_x(f^- - f^+) + U'(x)(f^- - f^+) = \gamma(f^+ + f^-). \tag{3.4}$$

We apply the transformation $g = f^- - f^+$ and $h = f^- + f^+$. We can then create a system of equations using (3.3) and (3.4) giving us

$$\begin{aligned} -(|U(x)|+2\lambda_0)g - \partial_x h &= \gamma g, \\ U'g - \partial_x g &= \gamma h. \end{aligned}$$

The interesting thing about the functions g and h is that they are now in a sense 'decoupled'. Namely, the mean-zero restriction is now only required for h, and therefore the function g can be taken independently. To make this all more precise as a unitary transformation we will write this system of equations in terms of an eigenvalue problem to a block operator $\mathscr{A}f = \gamma f$. To define this properly we use notation similar to the definition of $L^2_0(\mathbb{T},\nu)$ and $W^{1,2}_0(\mathbb{T},\nu)$. We define

$$L^2_0(\mathbb{T},\mu) := \{f \in L^2(\mathbb{T},\mu) : \int_{\mathbb{T}} f\,d\mu = 0\}$$

and

$$W^{1,2}_0(\mathbb{T},\mu) := W^{1,2}(\mathbb{T},\mu) \cap L^2_0(\mathbb{T},\mu).$$

We then define the operator $\mathscr{A} : W^{1,2}(\mathbb{T},\mu) \bigoplus W^{1,2}_0(\mathbb{T},\mu) \subset L^2(\mathbb{T},\mu) \bigoplus L^2_0(\mathbb{T},\mu) \to L^2(\mathbb{T},\mu) \bigoplus L^2_0(\mathbb{T},\mu)$ as a block operator

$$\mathscr{A} = \begin{bmatrix} -|U'(x)|-2\lambda_0 & -\partial_x \\ U'-\partial_x & 0 \end{bmatrix}. \tag{3.5}$$

For this, the $\bigoplus$ notation refers to taking direct sums of the underlying Hilbert spaces, such that the resulting product space is again a Hilbert space. The operator $\mathscr{A}$ relates to the Zig-Zag process generator $\mathscr{L}_0$ by $\mathscr{A} = T\mathscr{L}_0T^*$ or $\mathscr{L}_0 = T^*\mathscr{A}T$, where we have the bijective unitary transformation $T : L^2_0(\mathbb{T},\nu) \to L^2(\mathbb{T},\mu) \bigoplus L^2_0(\mathbb{T},\mu)$ such that $T = \frac{1}{\sqrt{2}}\begin{bmatrix} -I & I \\ I & I \end{bmatrix}$. and its adjoint $T^* : L^2(\mathbb{T},\mu) \bigoplus L^2_0(\mathbb{T},\mu) \to L^2_0(\mathbb{T},\nu)$ similarly defined as $T = \frac{1}{\sqrt{2}}\begin{bmatrix} -I & I \\ I & I \end{bmatrix}$. (although the form looks similar we do not have $T^* = T$ as the domains and co-domains are different). Because T is a unitary transformation, we have

$$\sigma(\mathscr{A}) = \sigma(\mathscr{L}_0) \text{ and } \sigma_p(\mathscr{A}) = \sigma_p(\mathscr{L}_0).$$

We will be examining the block operator $\mathscr{A}$ and we will write it in the following form

$$\mathscr{A} = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \tag{3.6}$$

There are several nice properties to this new representation. First of all, we have that this operator fulfils the definition of a block operator with a proper domain. Secondly, whereas the original representation had 4 non-zero entries in the 'block matrix', this new one only has 3. Furthermore, we have that only A depends on λ_0, and under reasonable assumptions, we have that A becomes a bounded operator (and thus relatively well behaved). Moreover, we will have that $B^* = -C$ (this will be proven in lemma 3.2.3). To make this more rigorous we will further define these new operators.

We define the following unbounded operators $(A, D(A)), (B, D(B)), (C, D(C))$, and $(D, D(D))$. Their definition follows from the relationship $\mathscr{A} = T\mathscr{L}_0 T^*$. First, we define the multiplier operator A which is the only operator dependent on the refreshment rate λ_0 by

$$\begin{aligned}
&A : D(A) \subset L^2(\mathbb{T},\mu) \to L^2(\mathbb{T},\mu),\\
&D(A) = W^{1,2}(\mathbb{T},\mu),\\
&Af(x) := -W(x)f(x) := -(|U'(x)| + 2\lambda_0(x))f.
\end{aligned}$$

Here we also defined the non-negative function $W(x) := |U'(x)| + 2\lambda_0(x)$ which will turn out to be an important function for the estimates/behaviour of the spectrum. This map is well defined because $\lambda_0, U' \in L^\infty(\mathbb{T})$ are bounded on $\mathbb{T}$ by assumption (A2) and (A1). So we have $Wf \in L^2(\mathbb{T},\mu)$. Next, we have a simple derivative map

$$\begin{aligned}
&B : D(B) \subset L^2_0(\mathbb{T},\mu) \to L^2(\mathbb{T},\mu),\\
&D(B) = W^{1,2}_0(\mathbb{T},\mu),\\
&Bf := -\partial_x f.
\end{aligned} \tag{3.7}$$

This map is also well defined by the definition of the Sobolev space. Finally we have another derivative map

$$\begin{aligned}
&C : D(C) \subset L^2(\mathbb{T},\mu) \to L^2_0(\mathbb{T},\mu),\\
&D(C) = W^{1,2}(\mathbb{T},\mu),\\
&Cf := U'f - \partial_x f.
\end{aligned} \tag{3.8}$$

This map is also well defined since $U'f \in L^2(\mu,\mathbb{T})$ due to the boundedness of U'. Furthermore, we have for $f \in D(C)$

$$\begin{aligned}
\int_{\mathbb{T}} U'f - \partial_x f \, d\mu &= \int_{\mathbb{T}} (U'f - \partial_x f)e^{-U(x)} dx\\
&= -\int_{\mathbb{T}} \partial_x(e^{-U(x)} f) dx\\
&= e^{-U(-\pi)} f(-\pi) - e^{-U(\pi)} f(\pi)\\
&= 0.
\end{aligned}$$

Here we similarly to proving the well-definedness of the generator $\mathscr{L}_0$ made use of the product rule since $e^{-U(x)} \in W^{1,\infty}(\mathbb{T})$ combined with Proposition 4.1.17 from [Web18]. This implies that $Cf \in L^2_0(\mu,\mathbb{T})$ and thus this makes C a well-defined operator. Finally there is also an implicit map for the bottom right entry of our block operator which is a nul operator which we will define for completeness and in order to apply certain theorems

$$\begin{aligned}
&D : D(D) \subset L^2_0(\mathbb{T},\mu) \to L^2_0(\mathbb{T},\mu),\\
&D(D) = W^{1,2}(\mathbb{T},\mu),\\
&Df := 0,
\end{aligned}$$

which is trivially well defined.

There are still some issues with these 'natural' definitions of these operators. For example, the operator A and D are not closed operators. (For this their domain would need to be $L^2(\mathbb{T},\mu)$ and $L^2_0(\mathbb{T},\mu)$ respectively). This will create some problems for defining some derived operators later on. Luckily both of these operators are closable and because $D(A) = D(C)$ and $D(B) = D(D)$ extending these operators does not alter the operator $\mathscr{A}$ by definition of the domain $D(\mathscr{A})$ as a block operator.

Lemma 3.2.2. *$(A, D(A))$ and $(D, D(D))$ have closed extensions given by $\overline{A} : L^2(\mathbb{T},\mu) \subset L^2(\mathbb{T},\mu) \to L^2(\mathbb{T},\mu)$, $\overline{A}f = -Wf$ and $\overline{D} : L^2_0(\mathbb{T},\mu) \subset L^2_0(\mathbb{T},\mu) \to L^2_0(\mathbb{T},\mu)$, $\overline{D}f = 0$, moreover $\overline{D}$ and $\overline{A}$ are bounded self-adjoint operators.*

Proof. This can be shown since for all $f \in L^2(\mathbb{T},\mu)$ we have $||Af||_{L^2(\mathbb{T},\mu)} = ||(|U'|+2\lambda_0)f||_{L^2(\mathbb{T},\mu)} \leq ||(|U'|+2\lambda_0)||_{L^\infty(\mathbb{T})}||f||_{L^2(\mathbb{T},\mu)}$, where $||(|U'|+2\lambda_0)||_{L^\infty(\mathbb{T})}$ is finite due to assumption (A2) and (A1) so A is bounded and thus well defined and we clearly have for all $f \in D(A)$ that $\overline{A}f = Af$ and thus $\overline{A}$ is an extension to A. The function W is a real-valued multiplier therefore the operator A is symmetric and because it is bounded it is also self-adjoint. The exact same argument can be made to prove the properties for D. □

The spectra of $\overline{A}$ and $\overline{D}$ are respectively given by $\sigma(\overline{A}) = \text{ess im}(-W)$ (The essential range of $-W$) and $\sigma(\overline{D}) = \{0\}$. Moreover, because we have

$$\mathscr{A} = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = \begin{bmatrix} \overline{A} & B \\ C & \overline{D} \end{bmatrix} \tag{3.9}$$

since $D(\mathscr{A}) = D(A)\cap D(C)\bigoplus D(B)\cap D(D) = D(\overline{A})\cap D(C)\bigoplus D(B)\cap D(\overline{D})$ we will often use $\overline{A}$ and A interchangeable and will often refer to $\overline{D}$ with D. Fortunately all of the operators are densely defined because the sets of $W^{1,2}(\mathbb{T},\mu), W_0^{1,2}(\mathbb{T},\mu)$ are respectively dense in $L^2(\mathbb{T},\mu), L_0^2(\mathbb{T},\mu)$. This allows us to define the adjoints of the operators. Next, we prove the assertion about the symmetry $B^* = -C$ that was mentioned before.

Lemma 3.2.3. *The unbounded operators* $(B, D(B)), (C, D(C))$ *are closed and* $B^* = -C$.

Proof. First, we prove the closedness of B: Take a sequence $g_n \in D(B)$, $g_n \to g \in L_0^2(\mathbb{T},\mu)$ with $Bg_n = -\partial_x g_n \to f \in L^2(\mathbb{T},\mu)$. We then have that g_n is Cauchy in $W_0^{1,2}(\mathbb{T},\mu)$ which is a Banach space and thus its limit has the property $-\partial_x g_n \to -\partial_x g$ in $L^2(\mathbb{T},\mu)$. Thus we have $g \in D(B)$ and by uniqueness of the limit in $L^2(\mathbb{T},\mu)$ we have $Bg = f$.

Next, we prove ($B^* = -C$): Because B is densely defined we know $(B^*, D(B^*))$ exists and is closed. Suppose $g \in D(B^*)$ and take an $f \in C_c^\infty(\mathbb{T})$. Furthermore, define $k(x) := e^{U(x)}f(x) - \int_\mathbb{T} e^{U(\xi)} f\, d\mu(\xi)$ such that $k \in D(B)$ and take the $h \in L_0^2(\mathbb{T},\mu)$ such that $(Bk|g) = (k|h)$ (which exists by definition of the adjoint). We then have

$$\begin{aligned}
\int_\mathbb{T} g\overline{\partial_x(f)}\, dx &= \int_\mathbb{T} g\overline{e^U\partial_x(f)}\, d\mu \\
&= \int_\mathbb{T} g\overline{(\partial_x(e^U f) - U'e^U f)}\, d\mu \\
&= \int_\mathbb{T} g\overline{\partial_x(e^U f)}\, d\mu - \int_\mathbb{T} \overline{gU'e^U f}\, d\mu \\
&= -\overline{(Bk|g)} - \int_\mathbb{T} g\overline{U'e^U f}\, d\mu \\
&\overset{\text{Adjoint}}{=} -\int_\mathbb{T} h\overline{k}\, d\mu - \int_\mathbb{T} g\overline{U'e^U f}\, d\mu \\
&= -\int_\mathbb{T} h\overline{f} + g\overline{U'f}\, dx + \int_\mathbb{T} h\, d\mu \overline{\int_\mathbb{T} e^{U(\xi)} f\, d\mu(\xi)} \\
&= -\int_\mathbb{T} (h + U'g)\overline{f}\, dx.
\end{aligned}$$

Where in the last equality we use that $\int_\mathbb{T} h\, d\mu = 0$ because h has mean zero. Since the equality holds for arbitrary $f \in C_c^\infty(\mathbb{T})$ we have by the definition of the weak derivative that $\partial_x g = h + U'g$, and because $h, U'g \in L^2(\mathbb{T},\nu)$, we have that $\partial_x g \in L^2(\nu)$. Thus we have $g \in D(C)$. Furthermore, by rearranging terms we see that $B^* g = h = \partial_x(g) - U'g = -Cg$. So we have $B^* \subset -C$.

Next, we assume that $g \in D(C)$, and we take an $f \in D(B)\cap C_c^\infty(\mathbb{T})$. For these functions we have the following identity

$$
\begin{aligned}
(Bf|g) &= \int_{\mathbb{T}} Bf\overline{g}\,d\mu \\
&= -\int_{\mathbb{T}} \partial_x(f)\overline{g}\,d\mu \\
&= -\int_{\mathbb{T}} \partial_x(f)\overline{g}e^{-U}\,dx \\
&= \int_{\mathbb{T}} f\partial_x(\overline{g}e^{-U})\,dx \\
&= \int_{\mathbb{T}} f(\partial_x\overline{g} - U'g)\,d\mu \\
&= -(f|Cg).
\end{aligned}
$$

For an arbitrary $f \in D(B)$ we have by denseness of $C_c^\infty(\mathbb{T}) \cap \{f : \int f\,d\mu = 0\}$ in $W_0^2(\mathbb{T},\mu)$ that there is a sequence $f_n \in D(B) \cap C_c^\infty(\mathbb{T})$ such that $||f_n - f||_{W^{1,2}(\mathbb{T},\mu)} \le \frac{1}{2\min\{||Cg||_{L^2(\mathbb{T},\mu)}, ||g||_{L^2(\mathbb{T},\mu)}\}}\epsilon$. We then have the following

$$
\begin{aligned}
|(Bf|g) - -(f|Cg)| &\le |(Bf|g) - (Bf_n|g) + (Bf_n|g) - (f_n|Cg) + (f_n|Cg) - (f|Cg)| \\
&\le |(Bf|g) - (Bf_n|g)| + |(f|Cg) - (f_n|Cg)| \\
&= |(B(f - f_n)|g)| + |(f - f_n|Cg)| \\
&\overset{\text{Hölder}}{\le} ||B(f - f_n)||_{L^2(\mathbb{T},\mu)}||g||_{L^2(\mathbb{T},\mu)} + ||f - f_n||_{L^2(\mathbb{T},\mu)}||Cg||_{L^2(\mathbb{T},\mu)} \\
&\le \epsilon.
\end{aligned}
$$

So we have $(Bf|g) = -(f|Cg)$ for all $f \in D(B)$ and we can see that $f \to (Bf|g)$ is continuous because of the following, for a sequence $f_n \to f \in D(B)$ we have

$$
\begin{aligned}
|(Bf|g) - (Bf_n|g)| &= |(f|Cg) - (f_n|Cg)| \\
&= |(f - f_n|Cg)| \\
&\overset{Hölder}{\le} ||f - f_n||_{L^2(\nu,\mathbb{T})}||Cg||_{L^2(\mathbb{T},\mu)} \to 0.
\end{aligned}
$$

So we have $g \in D(B^*)$ and thus we have $B^* = -C$ and it follows that C is closed as well. □

The closedness of B and C gives us the closedness of $\mathscr{A}$ as can be seen by the following lemma.

Lemma 3.2.4. *$(\mathscr{A}, D(\mathscr{A}))$ is a closed operator.*

Proof. Because B and C are closed operators this follows from Theorem 2.2.7 i of [Tre08]. This theorem can be applied because $\mathscr{A}$ is off-diagonally dominant of order 0. This can be seen because the operator D is trivially B-bounded of order 0 (see Definition 2.1.2 of [Tre08]). And operator A is C-bounded because it is a bounded operator on the elements of $D(A)$ (and thus also on $D(C)$). It is therefore C-bounded of order 0. Which makes $\mathscr{A}$ is off-diagonally dominant of order 0. □

Now that we have $B^* = -C$ we get that $\mathscr{A}$ has the following form

$$
\mathscr{A} = \begin{bmatrix} A & B \\ -B^* & D \end{bmatrix}.
$$

An operator of this form is a J-symmetric block operator. The name refers to an the operator
$J : L^2(\mathbb{T},\nu) \bigoplus L_0^2(\mathbb{T},\nu) \to L^2(\mathbb{T},\nu) \bigoplus L_0^2(\mathbb{T},\nu)$ defined by

$$
J := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{3.10}
$$

with the property $J^2 = I$ and such that $\mathscr{A}J$ and $J\mathscr{A}$ become symmetric operators. Even stronger than this we have that $\mathscr{A}$ is J-self-adjoint as follows from the following corollary, where we are able to prove the adjoint of $\mathscr{A}$ by knowing the adjoints of B and C and having A and D be bounded operators.

Lemma 3.2.5. *The adjoint of* $(\mathscr{A}, W^{1,2}(\mathbb{T},\nu) \bigoplus W_0^{1,2}(\mathbb{T},\nu))$ *is given by*

$$\mathscr{A}^* = \mathscr{A}_{sym} - \mathscr{A}_{asym} = \begin{bmatrix} A & -B \\ B^* & 0 \end{bmatrix},$$

where

$$\mathscr{A}_{sym} := \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \mathscr{A}_{asym} := \begin{bmatrix} 0 & B \\ -B^* & 0 \end{bmatrix}.$$

Which are the symmetric and asymmetric parts of the operator $\mathscr{A}$ *such that we also have* $\mathscr{A} = \mathscr{A}_{sym} + \mathscr{A}_{asym}$. *Moreover, we have the domain* $D(\mathscr{A}^*) = W^{1,2}(\mathbb{T},\nu) \bigoplus W_0^{1,2}(\mathbb{T},\nu)$. *We also have that* $\mathscr{A}$ *is J-self-adjoint meaning* $\mathscr{A}^* = JAJ$ *with J defined in formula* (3.10).

Proof. Because of the boundedness of $\mathscr{A}_{sym}$ we have that $\mathscr{A} = \mathscr{A}_{sym} + \mathscr{A}_{asym}$. It can then be seen that $\mathscr{A}_{sym}^* = \mathscr{A}_{sym}$ and using Proposition 2.6.3 of [Tre08] we have that

$$\mathscr{A}_{asym}^* = \begin{bmatrix} 0 & B \\ -B^* & 0 \end{bmatrix}^* = \begin{bmatrix} 0 & -B^{**} \\ B^* & 0 \end{bmatrix} = \begin{bmatrix} 0 & -B \\ B^* & 0 \end{bmatrix} = -\mathscr{A}_{asym}$$

where we have that $B^{**} = B$ due to the closedness of B from lemma 3.2.3. We then see again using the boundedness of $\mathscr{A}_{sym}$ that $\mathscr{A}^* = (\mathscr{A}_{sym} + \mathscr{A}_{asym})^* = \mathscr{A}_{sym} - \mathscr{A}_{asym}$. □

Corollary 3.2.6. *We have* $\mathscr{L}_0^* = F\mathscr{L}_0 F$, *where* $F : L_0^2(\mathbb{T},\nu) \to L_0^2(\mathbb{T},\nu)$ *is the flip operator as defined in formula* (3.1).

Proof. This can be seen by noting that $\mathscr{L}_0^* = T^*\mathscr{A}^*T = T^*J^*\mathscr{A}JT = T^*J^*T\mathscr{L}_0T^*JT$ and observing that $F = -T^*JT$. □

This gives us that $\mathscr{A}$ is J-self-adjoint and allows us to express the adjoint of $\mathscr{A}$ as $\mathscr{A}^* = J\mathscr{A}J$. J-self-adjoint operators are operators which are self-adjoint in Krein spaces, which are indefinite inner product spaces that have a decomposition $\mathscr{H}_1, \mathscr{H}_2$ with inner products $(\cdot|\cdot)_{\mathscr{H}_1}$ and $-(\cdot|\cdot)_{\mathscr{H}_2}$ such that the operator becomes self-adjoint on $\mathscr{H}_1 \bigoplus \mathscr{H}_2$ (The direct sum of the two Hilbert spaces) this creates an indefinite inner product space. The spectrum of self-adjoint (block) operators have been well studied with many interesting properties about their behaviour such a spectral mapping theorems and variational principles [Tre08], [Sch12], [Kna17]. The J-self-adjoint operators are often studied under the assumption that they are diagonally dominant and sometimes when they are upper dominant [LS17]. This means that the operator A, D, A, B are more influential operators in the block operator in the sense that they relatively bound the other operators (see Definition 2.2.1 from [Tre08]). Often when it comes to off-diagonally dominant J-self-adjoint operators there is an assumption that the spectrum is real-valued [LLT02], [KLT04]. When the operator is diagonally dominant there even exist a Gershgorin Theorem similar to the one defined for matrices [Gir+20]. However, with off-diagonally dominant matrices, such theorems do not apply. However, we can use similar objects as defined in these papers and see if other statements apply. Before we get to that, we will first note some useful properties for B and B^* that we will use later on.

First, we prove that the operator B can be redefined to become a boundedly invertible operator $\tilde{B}$ if we remove a function e^U from the co-domain. Similarly, we can turn C into a boundedly invertible operator $\tilde{C}$ if we remove e^U from its domain.

Lemma 3.2.7. *B as defined in formula* (3.7) *is injective and C as defined in formula* (3.8) *is surjective. Moreover, if B is defined as a mapping* $\tilde{B} : W_0^{1,2}(\mathbb{T},\mu) \subset L_0^2(\mathbb{T},\mu) \to L^2(\mathbb{T},\mu) \setminus \{e^U\}$ *and* $\tilde{C} : W^{1,2}(\mathbb{T},\mu) \setminus \{e^U\} \subset L^2(\mathbb{T},\mu) \setminus \{e^U\} \to L_0^2(\mathbb{T},\mu)$ *(where with the notation* $\Omega \setminus \{e^U\}$ *we refer to the Banach space* $\Omega \setminus \{e^U\} := \{f \in \Omega : \int_{\mathbb{T}} f(x)dx = 0\}$ *with the norm inherited from* Ω*) then we again have* $\tilde{B}^* = -\tilde{C}$ *and we have that B and C are boundedly invertible with the bounded inverse of* $\tilde{B}$ *given by*

$$\tilde{B}^{-1}f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{y} f(\xi)\,d\xi\, e^{-U(y)}dy - \int_{-\pi}^{x} f(\xi)\,d\xi,$$

and the bounded inverse of $\tilde{C}$ *given by*

$$\tilde{C}^{-1}f(x) = \frac{e^{U(x)}}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{y} f(\xi)\,e^{-U(\xi)}d\xi\,dy - e^{U(x)}\int_{-\pi}^{x} f(\xi)\,e^{-U(\xi)}d\xi.$$

Moreover, we have that $(\tilde{B}^{-1})^* = -\tilde{C}^{-1}$.

Proof. First, we prove the injectivity of B and we prove the surjectivity of C at the end of the proof of this lemma. If we have $f, g \in D(B)$ such that $B(f-g)=0$, which means that we have $-\partial_x(f-g)=0$, we then have that $f-g=K$, where K is a constant. And because of the definition of $D(B)$ we have that $\int_{\mathbb{T}} f - g \, d\mu = 0$ And thus we have that $K=0$ and thus $f=g$, so B is injective.

We first note that the mapping $\tilde{B}$ and $\tilde{C}$ are well defined, because $\tilde{C}$ is a restriction of C and because $(Bf|e^U) = \int_{\pi}^{\pi} -\partial_x f \, dx = 0$ due to the periodicity of $f \in W^{1,2}(\mathbb{T},\mu)$. Because $W^{1,2}(\mathbb{T},\mu) \setminus \{e^U\}$ is still a Hilbert space, the proof surrounding the closedness of $\tilde{B}$ and $\tilde{B}^* = -\tilde{C}$ goes exactly the same like lemma 3.2.3. We then prove that the mapping $\tilde{B}$ is boundedly invertible. Since B is injective then we certainly have that $\tilde{B}$ is injective as the domain is the same and the co-domain is smaller. For the surjectivity of $\tilde{B}$ if we have an $f \in L^2(\mathbb{T},\mu) \setminus \{e^U\}$ we can define the inverse of $\tilde{B}$ as denoted above. We need to prove that $\tilde{B}f \in W_0^{1,2}(\mathbb{T},\mu)$ for all $f \in L^2(\mathbb{T},\mu) \setminus \{e^U\}$. Take an $f \in L^2(\mathbb{T},\mu) \setminus \{e^U\}$ we then have the point-wise estimate

$$|\int_{-\pi}^{x} f \, d\xi| \leq ||f||_{L^2(\mathbb{T},\mu)} ||e^U||_{L^2(\mathbb{T},\mu)}$$

and thus we have

$$||\int_{-\pi}^{\cdot} f \, d\xi||_{L^2(\mathbb{T},\mu)} \leq ||f||_{L^2(\mathbb{T},\mu)} ||e^{U(\cdot)}||_{L^2(\mathbb{T},\mu)} ||1||_{L^2(\mathbb{T},\mu)}.$$

Furthermore, for the double integral term we have

$$|\int_{-\pi}^{\pi}\int_{-\pi}^{y} f \, d\xi \, e^{-U(y)} \, dy| \leq \int_{-\pi}^{\pi} ||f||_{L^2(\mathbb{T},\mu)} ||e^{U(\cdot)}||_{L^2(\mathbb{T},\mu)} e^{-U(y)} \, dy = ||f||_{L^2(\mathbb{T},\mu)} ||e^{U(\cdot)}||_{L^2(\mathbb{T},\mu)} ||1||^2_{L^2(\mathbb{T},\mu)}$$

and thus we have

$$||\int_{-\pi}^{\pi}\int_{-\pi}^{y} f \, d\xi \, e^{-U(y)} \, dy||_{L^2(\mathbb{T},\mu)} \leq ||f||_{L^2(\mathbb{T},\mu)} ||e^{U(\cdot)}||_{L^2(\mathbb{T},\mu)} ||1||^3_{L^2(\mathbb{T},\mu)}.$$

So we see that $\tilde{B}^{-1}f \in L^2(\mathbb{T},\mu)$. We also have that $\tilde{B}^{-1}f$ is weakly differentiable and $\partial_x \tilde{B}^{-1} f = -f \in L^2(\nu)$. Specifically we have that $\tilde{B}^{-1}f(\pi) = \tilde{B}^{-1}f(-\pi)$ because $\int_{-\pi}^{\pi} f \, d\xi = 0$ due to the orthogonality of f on e^U. So we have $\tilde{B}^{-1}f \in W^{1,2}(\mathbb{T},\mu)$. Finally we have

$$\begin{aligned}(\tilde{B}^{-1}f|1) &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi}\int_{-\pi}^{\eta} f \, d\xi e^{-U(\eta)} \, d\eta - \int_{-\pi}^{\pi}\int_{-\pi}^{x} f \, d\xi \, e^{-U(x)} \, dx \\ &= \int_{-\pi}^{\pi}\int_{-\pi}^{\eta} f \, d\xi \, e^{-U(\eta)} \, d\eta - \int_{-\pi}^{\pi}\int_{-\pi}^{x} f \, d\xi e^{-U(x)} \, dx = 0.\end{aligned}$$

So we get $\tilde{B}^{-1}f \in W_0^{1,2}(\mathbb{T},\mu)$. So we see that $\tilde{B}$ is surjective combined with the injectivity of B we see that B is boundedly invertible. Then using Theorem 5.30 from [Kat95] and lemma 3.2.3 we get that $(B^{-1})^* = -C^{-1}$.

The surjectivity of C can be seen by taking an $f \in L_0^2(\mathbb{T},\mu)$, we then see that have if we use the definition of $\tilde{C}^{-1}f$ that $\tilde{C}^{-1}f \in W^{1,2}(\mathbb{T},\mu)$ by also noting that the function is continuous at $\tilde{C}^{-1}f(\pi) = \tilde{C}^{-1}f(-\pi)$ because of a similar argument shown below, moreover by simple calculations we see that we have $C\tilde{C}^{-1}f = f$ and we have that $Cg = f$ so C is surjective.

□

It turns out that the inverses of B and C are a special type of bounded operators, namely they are Hilbert Schmidt operators.

Lemma 3.2.8. $(\tilde{B}^*)^{-1} = \tilde{C}^{-1} : L_0^2(\mathbb{T},\mu) \to W^{1,2}(\mathbb{T},\mu) \setminus \{e^U\}$ *is a Hilbert-Schmidt operator given by*

$$(\tilde{B}^*)^{-1} f(x) = \int_{-\pi}^{\pi} k(x,\xi) f(\xi) \, d\mu(\xi)$$

where the kernel function k is given by

$$k(x,\xi) = \frac{e^{U(x)}}{2\pi}(\pi - \xi) - e^{U(x)} 1_{x>\xi}.$$

Furthermore, the Hilbert-Schmidt norm of $(\tilde{B}^*)^{-1}$ *is given by*

$$\begin{aligned}||(\tilde{B}^*)^{-1}||_{HS}^2 &= ||k||^2_{L^2(\mathbb{T}\times\mathbb{T},\mu\times\mu)}\\ &= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(\frac{1}{2\pi}(\pi-\xi)-1_{x>\xi})^2 e^{-U(\xi)+U(x)}\,dx\,d\xi.\end{aligned}$$

Similarly, we have that $\tilde{B}^{-1} = (\tilde{C}^*)^{-1} : L^2(\mathbb{T},\mu)\setminus\{e^U\} \to W_0^{1,2}(\mathbb{T},\mu)$ *is a Hilbert-Schmidt operator given by:*

$$\tilde{B}^{-1}f = \int_{-\pi}^{\pi} k(\xi,x) f(\xi)\,d\mu(\xi)$$

and we again have

$$||\tilde{B}^{-1}||_{HS}^2 = ||k||^2_{L^2(\mathbb{T}\times\mathbb{T},\mu\times\mu)}.$$

Proof. We can use the definition of $(\tilde{B}^*)^{-1}$ from lemma 3.2.7 where we had the following formula

$$\begin{aligned}(\tilde{B}^*)^{-1}f(x) &= \frac{e^{U(x)}}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{y} f(\xi)\,e^{-U(\xi)}\,d\xi\,dy - e^{U(x)}\int_{-\pi}^{x} f(\xi)\,e^{-U(\xi)}\,d\xi\\ &= \frac{e^{U(x)}}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} 1_{y>\xi} f(\xi)\,e^{-U(\xi)}\,d\xi\,dy - e^{U(x)}\int_{-\pi}^{\pi} 1_{x>\xi} f(\xi)\,e^{-U(\xi)}\,d\xi\\ &\overset{\text{Fubini}}{=} \frac{e^{U(x)}}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} 1_{y>\xi} f(\xi)\,e^{-U(\xi)}\,dy\,d\xi - e^{U(x)}\int_{-\pi}^{\pi} 1_{x>\xi} f(\xi)\,e^{-U(\xi)}\,d\xi\\ &= \int_{-\pi}^{\pi}\frac{e^{U(x)-U(\xi)}}{2\pi} f(\xi)\int_{\xi}^{\pi} 1\,dy\,d\xi - \int_{-\pi}^{\pi} 1_{x>\xi} f(\xi)\,e^{U(x)-U(\xi)}\,d\xi\\ &= \int_{-\pi}^{\pi}\frac{e^{U(x)}}{2\pi} f(\xi)(\pi-\xi)\,d\mu(\xi) - \int_{-\pi}^{\pi} 1_{x>\xi} f(\xi)\,e^{U(x)}\,d\mu(\xi)\\ &= \int_{-\pi}^{\pi} k(x,\xi) f(\xi)\,d\mu(\xi)\end{aligned}$$

Since the kernel consists of a bounded function we have that $||k||^2_{L^2(\mathbb{T}\times\mathbb{T},\mu\times\mu)} < \infty$. This makes $(\tilde{B}^*)^{-1}$ a Hilbert Schmidt operator according to Proposition 3.3.5 from [Sun16]. Moreover, it gives us that $||(\tilde{B}^*)^{-1}||_{HS} = ||k(t,s)||_{L^2(\mathbb{T}\times\mathbb{T})}$. Similarly, we have for Hilbert-Schmidt operators with kernel $k(x,y)$ that their adjoints are again Hilbert-Schmidt operators with as kernel $\overline{k(y,x)}$. This gives the statement about $\tilde{B}^{-1}$. □

This will gives us in particular that $(\tilde{B}^*)^{-1}$ is compact, but more importantly it gives us an estimate for the norm $||\tilde{B}^*||_{L^2(\mathbb{T},\mu)} \leq ||k||_{L^2(\mathbb{T}\times\mathbb{T})}$. Using lemma 3.2.7 we can prove that $0 \in \rho(\mathscr{L}_0)$. Meaning that we can find a function $f \in L_0(\mathbb{T},\nu)$ such that there is a unique solution u to the Poisson equation $\mathscr{L}_0 u = f$, where u is in $W_0^{1,2}(\mathbb{T},\nu)$. Moreover, we get a specific formula for u.

Proposition 3.2.9. *The generator of the Zig-Zag process* $\mathscr{L}_0 : W_0^{1,2}(\mathbb{T},\nu) \subset L_0^2(\mathbb{T},\nu) \to L_0^2(\mathbb{T},\nu)$ *is boundedly invertible if and only if* $||W||_\infty > 0$. *Take* $(h^+,h^-) \in L_0^2(\mathbb{T},\nu)$, *we then have*

$$\begin{bmatrix}\tilde{f}^+\\ \tilde{f}^-\end{bmatrix} = \mathscr{L}_0^{-1}\begin{bmatrix}h^+\\ h^-\end{bmatrix}$$

where we have that

$$\tilde{f}^+ = -Ke^{U(x)} + e^{U(x)}\int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy + \tilde{B}^{-1}((-h^+ + h^-)(x) + W(x)e^{U(x)}(K - \int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy)),$$

$$\tilde{f}^- = Ke^{U(x)} - e^{U(x)}\int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy + \tilde{B}^{-1}((-h^+ + h^-)(x) + W(x)e^{U(x)}(K - \int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy)),$$

with

$$K = -\frac{\int_{-\pi}^{\pi}(-h^+ + h^-)(x) - W(x)e^{U(x)}\int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy\,dx}{\int_{-\pi}^{\pi} W(x)e^{U(x)}\,dx}.$$

If $||W||_\infty = 0$ *and by taking a constant* $Q \in \mathbb{C}\setminus\{0\}$, *then we have the eigenvector*

$$\mathscr{L}_0\begin{bmatrix}Q\\ -Q\end{bmatrix} = \mathscr{A}\begin{bmatrix}Q\\ 0\end{bmatrix} = 0.$$

Proof. Take $(g^+, g^-) \in L^2(\mathbb{T},\mu) \bigoplus L_0^2(\mathbb{T},\mu)$. We then have the following equations

$$\mathscr{A}\begin{bmatrix} f^+ \\ f^- \end{bmatrix} = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} f^+ \\ f^- \end{bmatrix} = \begin{bmatrix} Af^+ + Bf^- \\ Cf^+ \end{bmatrix} = \begin{bmatrix} g^+ \\ g^- \end{bmatrix}.$$

We can make use of the surjectivity of C and knowledge of its kernel from lemma 3.2.7 to see that

$$f^+(x) = Ke^{U(x)} - e^{U(x)}\int_{-\pi}^{x} g^-(y)e^{-U(y)}\,dy$$

for some $K \in \mathbb{C}$. Moreover, this lemma gives us that $f^+ \in W^{1,2}(\mathbb{T},\mu)$. Combining this with the equation $Af^+ + Bf^- = g^+$ gives us that

$$Bf^-(x) = g^+(x) + KW(x)e^{U(x)} - W(x)e^{U(x)}\int_{-\pi}^{x} g^-(y)e^{-U(y)}\,dy.$$

We know from lemma 3.2.7 that for the operator B an inverse exists if and only if the right-hand side is in $L^2(\mathbb{T},\mu) \setminus \{e^U\}$ which means that

$$\int_{-\pi}^{\pi} g^+(x) + KW(x)e^{U(x)} - W(x)e^{U(x)}\int_{-\pi}^{x} g^-(y)e^{-U(y)}\,dy\,dx = 0.$$

Which means that we need $||W||_\infty > 0$ such that we can define

$$K = -\frac{\int_{-\pi}^{\pi} g^+(x) - W(x)e^{U(x)}\int_{-\pi}^{x} g^-(y)e^{-U(y)}\,dy\,dx}{\int_{-\pi}^{\pi} W(x)e^{U(x)}\,dx}.$$

This then gives us that

$$f^-(x) = \tilde{B}^{-1}(g^+(x) + W(x)e^{U(x)}(K - \int_{-\pi}^{x} g^-(y)e^{-U(y)}\,dy)).$$

We then have from lemma 3.2.7 that $(f^+, f^-) \in W^{1,2}(\mathbb{T},\mu) \bigoplus W_0^{1,2}(\mathbb{T},\mu)$.

By using the relationship $\mathscr{A} = T\mathscr{L}_0T^*$ we see that $\mathscr{L}_0^{-1} = T^*\mathscr{A}^{-1}T$. Which means that if we have $(h^+, h^-) \in W_0^{1,2}(\mathbb{T},\nu)$ then we set $g^+ = -h^+ + h^-$ and $g^- = h^+ + h^-$ such that

$$\begin{aligned} f^+(x) &= Ke^{U(x)} - e^{U(x)}\int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy, \\ f^-(x) &= B^{-1}((-h^+ + h^-)(x) + W(x)e^{U(x)}(K - \int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy)). \end{aligned}$$

Then we have that

$$\mathscr{L}_0^{-1}\begin{bmatrix} h^+ \\ h^- \end{bmatrix} = \begin{bmatrix} -f^+ + f^- \\ f^+ + f^- \end{bmatrix} =: \begin{bmatrix} \tilde{f}^+ \\ \tilde{f}^- \end{bmatrix}$$

which means that

$$\begin{aligned} \tilde{f}^+ &= -Ke^{U(x)} + e^{U(x)}\int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy + \tilde{B}^{-1}((-h^+ + h^-)(x) + W(x)e^{U(x)}(K - \int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy)), \\ \tilde{f}^- &= Ke^{U(x)} - e^{U(x)}\int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy + \tilde{B}^{-1}((-h^+ + h^-)(x) + W(x)e^{U(x)}(K - \int_{-\pi}^{x}(h^+ + h^-)(y)e^{-U(y)}\,dy)). \end{aligned}$$

The statement about the eigenvector if $||W||_\infty = 0$ is simple as in this case the operator would become $\mathscr{L}_0 f = \theta\partial_x f$ for which we clearly have that with a $Q \in \mathbb{C} \setminus \{0\}$ that $f(x,\theta) = \theta Q$ has the property $\mathscr{L}_0 f = 0$ and $\int_{\mathbb{T}} f\,d\nu = 0$. □

Now that we have introduced $\mathscr{L}_0$ as an equivalent block operator $\mathscr{A}$ and proved some properties about the operators making up the full operator we can start to examine some bounds on the spectrum which we will do in the following section.

3.3. The numerical range and the Schur complements

In general when trying to examine the spectrum of an operator there is not a single method that works every time. One popular approach is to use the numerical range of an operator. The numerical range of an operator is defined by $\mathscr{W}(\mathscr{L}_0) := \{\frac{\langle \mathscr{L}_0 f, f\rangle}{\langle f,f\rangle} : f \in D(\mathscr{L}_0) \setminus \{0\}\}$ [GR97]. In particular we can take $f \neq 0$ to be an eigenvector corresponding to an eigenvalue γ since $\mathscr{L}_0 f = \gamma f$ implies $\langle \mathscr{L}_0 f, f\rangle = \gamma \langle f, f\rangle$ which implies $\frac{\langle \mathscr{L}_0 f, f\rangle}{\langle f,f\rangle} = \gamma$ and we can see that that $\sigma_p(\mathscr{L}_0) \subset \mathscr{W}(\mathscr{L}_0)$. Unfortunately, the numerical range does not give a lot of information about the operator $\mathscr{A}$ or $\mathscr{L}_0$.

Proposition 3.3.1. *Assume that* $\mu(\{x \in \mathbb{T} : W(x) = 0\}) = 0$ *and* $\mu(\{x \in \mathbb{T} : W(x) = -||W||_\infty\}) = 0$. *Then the numerical range* $\mathscr{W}(\mathscr{L}_0)$ *of* $\mathscr{L}_0$ *is given by*

$$\mathscr{W}(\mathscr{L}_0) = \{\gamma \in \mathbb{C} : \mathrm{Re}\,\gamma \in (-||W||_\infty, 0)\} \cup \{0\}.$$

Proof. Assume that $f \in D(\mathscr{A})$ we then have

$$\begin{aligned}
\langle \mathscr{A} f, f\rangle &= \langle \begin{bmatrix} A & B \\ -B^* & 0 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \rangle \\
&= \langle A f_1, f_1\rangle + \langle B f_2, f_1\rangle - \langle B^* f_1, f_2\rangle \\
&= \langle A f_1, f_1\rangle + \langle B f_2, f_1\rangle - \overline{\langle B f_2, f_1\rangle} \\
&= \langle A f_1 | f_1\rangle + 2i\,\mathrm{Im}\langle B f_2 | f_1\rangle.
\end{aligned}$$

This representation shows a decomposition of the real and imaginary part of elements of the numerical range. We see that the real part of $\mathscr{A}$ is equal to $\langle A f_1, f_1\rangle$. By taking arbitrary f_1 we can see that the numerical range can not take values in $\{\gamma \in \mathbb{C} : \mathrm{Re}\,\gamma \notin (-||W||_\infty, 0]\}$.

We will make use of the convexity of the numerical range (Toeplitz-Hausdorff Theorem [Gus70]) and identify certain points in the spectrum to conclude that the convex hull of these points is in the numerical range.

By choosing $f_1 = 0$ and let $f_2 \in W_0^{1,2}(\mathbb{T},\mu)$ such that $||f_2||_{L^2(\mathbb{T},\mu)} = 1$. This implies that $0 \in \mathscr{W}(\mathscr{L}_0)$. Next, by choosing $f_2 = 0$ and letting f_1 be in $W^{1,2}(\mathbb{T},\mu)$ such that $||f_1||_{L^2(\mathbb{T},\mu)} = 1$. It can then seen that the numerical range of A is in the numerical range of $\mathscr{L}_0$, so we have $\mathscr{W}(A) \subset \mathscr{W}(\mathscr{L}_0)$. Then using the convexity of the numerical range we have in particular that all of the values on the real axis $(-||W||_\infty, 0] \subset \mathscr{W}(\mathscr{L}_0)$ are in the numerical range of $\mathscr{L}_0$.

Next, by taking $f_2 = f_1 = e^{i\gamma x + U(x)}$ with $\gamma \in \mathbb{Z} \setminus \{0\}$ we can see that $\int_\mathbb{T} e^{i\gamma x + U(x)}\, d\mu(x) = \int_\mathbb{T} e^{i\gamma x}\, dx = 0$ and thus we have $f_2 \in W_0^{1,2}(\mathbb{T},\mu)$. Moreover, we have that $B^* f_1 = i\gamma B^* f_1$ thus we get that $\mathrm{Im}\langle \mathscr{A} f, f\rangle = 2\,\mathrm{Im}\langle f_1, B^* f_1\rangle = -2\gamma\langle f_1, f_1\rangle = -2\gamma \int_\mathbb{T} e^{U(x)}\, dx$. Moreover, we have that $\langle f, f\rangle = 2\langle f_1, f_1\rangle = 2\int_\mathbb{T} e^{U(x)}\, dx$. So we have $\frac{\langle \mathscr{L}_0 f, f\rangle}{\langle f,f\rangle} = i\gamma$. Due to the convexity of the numerical range [Gus70] we then get that $\{\gamma \in \mathbb{C} : \mathrm{Re}\,\gamma \in (-||W||_\infty, 0)\} \subset \mathscr{W}(\mathscr{L}_0)$. The only values to still check are $\{\gamma \in \mathbb{C} \setminus \{0\} : \mathrm{Re}\,\gamma = 0\}$. This part of the spectrum can only be included in the numerical range if we have that $\langle A f_1, f_1\rangle = 0$ which implies that $f_1 = 0$, but then we have that $\langle \mathscr{A} f, f\rangle = 0$. So we have that $\{\gamma \in \mathbb{C} \setminus \{0\} : \mathrm{Re}\,\gamma = 0\} \subset \mathbb{C} \setminus \mathscr{W}(\mathscr{L}_0)$. □

From the numerical range we can see that the spectrum of the generator has to lie in a vertical strip with the boundary lines of this strip going through the points $(0,0)$ and $(-||W||_\infty, 0)$. The assumptions on the two sets in Proposition 3.3.1 being equal to zero only has an influence on whether the boundaries of this strip are included or not. Either way, the description of the spectrum using the numerical range does not give any information about the magnitude of the spectral gap.

As an improvement for accurately describing the spectrum of a block operator the quadratic numerical range was introduced [Tre08]. For a block operator $\mathscr{A}$ with domain $D(\mathscr{A}) = D_1 \bigoplus D_2$ the quadratic numerical range is defined by examining the eigenvalues of $2x2$ complex-valued matrices of the following form

$$\mathscr{A}_{f,g} := \begin{bmatrix} \langle A f, f\rangle & \langle B g, f\rangle \\ \langle C f, g\rangle & \langle D g, g\rangle \end{bmatrix}.$$

The quadratic numerical range $\mathscr{W}^2(\mathscr{A})$ of $\mathscr{A}$ is then defined by

$$\mathscr{W}^2(\mathscr{A}) := \bigcup_{\substack{f \in D_1, g \in D_2, \\ ||f|| = ||g|| = 1}} \sigma_p(\mathscr{A}_{f,g}). \tag{3.11}$$

The intuition behind the quadratic numerical range is similar to that of the numerical range and improves it by examining the numerical range of the two separate equations of the eigenvalue problem. It can then be shown that $\sigma_p(\mathscr{A}) \subset \mathscr{W}^2(\mathscr{A}) \subset \mathscr{W}(\mathscr{A})$ (Theorem 2.5.3, Theorem 2.5.9 [Tre08]).

Most of the theorems that we could find that would be applicable to our block operator assume that $\mathscr{A}$ either is self-adjoint, or that $\mathscr{W}^2(\mathscr{A}) \subset \mathbb{R}$, or that $\mathscr{A}$ is diagonally dominant which does not apply to our case. An overview of many of these theorems can be found in section 2 of [Tre08] (specifically section 2.6 is about J-self-adjoint block operators). The quadratic numerical range is not necessarily convex anymore, in fact in our case it is not convex. Later we will come back to discussing the quadratic numerical range. The reason for this is because we will first examine the numerical range of the Schur complements $S_1(\gamma), S_2(\gamma)$ in Proposition 3.4.3. And as we will later show surrounding formula (3.13) it turns out that the numerical range of the Schur complement is included in the quadratic numerical range of $\mathscr{A}$, so we have $\mathscr{W}(S_1(\gamma)) \subset \mathscr{W}^2(\mathscr{A})$.

In order to analyse the spectrum, we will make use of an operator family called the Schur complements. It is not just that the numerical range of these operators has the spectrum of $\mathscr{L}_0$ included in it, but there it will turn out that the kernel functions $\{f \in D(S_1(\gamma)) : S_1(\gamma)f = 0\}$ correspond one to one to the eigenvectors of $\mathscr{L}_0$.

Definition 3.3.2 (Schur complements)**.** *The Schur complements S_1 and S_2 of a block operator $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ are defined by*

$$S_1(\gamma) := A - \gamma - B(D-\gamma)^{-1}C \qquad \text{for } \gamma \in \rho(D),$$
$$S_2(\gamma) := D - \gamma - C(A-\gamma)^{-1}B \qquad \text{for } \gamma \in \rho(A).$$

The Banach spaces that these complements are defined on are inherited from the Banach spaces that the underlying operators A, B, C, and D are defined on.

In general the domains $D(S_1(\gamma))$ and $D(S_2(\gamma))$ vary for different $\gamma \in \mathbb{C}$. Note that we are using the definition of the D, A with their closed domains as in lemma 3.2.2 such that $S_1(\gamma)$ is defined for $\gamma \in \mathbb{C} \setminus \{0\}$ and $S_2(\gamma)$ is defined for $\gamma \in \mathbb{C} \setminus \{\text{essrange}(-W(x))\}$. This makes $S_1(\gamma)$ a more useful candidate as it allows us to look at almost all of the spectrum (apart from $\gamma = 0$, but this value has already been analysed with Proposition 3.2.9). Moreover, we have a clear definition for $\rho(D)$ as this operator is a null operator ($D = 0$) we have that $\rho(D) = \mathbb{C} \setminus \{0\}$. On the other hand the spectrum of A is the essential range of $-W$. This would not allow us to analyse some of the real-valued spectral values. For our purposes $S_1(\gamma) = -\frac{1}{\gamma}BB^* + (A - \gamma)$ is therefore the more useful operator family of the two which we will refer to when talking about the Schur complement. By definition of the concatenation of multiple unbounded operators the domain of $S_1(\gamma)$ is given by $D(S_1(\gamma)) = \{f \in D(C) : Cf \in D(B)\} = \{f \in D(B^*) : B^*f \in D(B)\} = D(BB^*)$. Since Cf is a map into mean-zero functions this domain can be written as $D(S_1(\gamma)) = \{f \in W^{1,2}(\mathbb{T}, \mu) : \partial_x f - U'f \in W^{1,2}(\mathbb{T}, \mu)\}$. Under assumption (A3) we have that this becomes equal to $W^{2,2}(\mathbb{T}, \mu)$ due to the product rule being applicable as $U' \in W^{1,\infty}(\mathbb{T})$ and using Proposition 4.1.17 from [Web18]. This domain is useful as it allows us to turn $S_1(\gamma)$ into a closed operator for $\gamma \in \mathbb{C} \setminus \{0\}$.

Now that we have properly defined the domain of the Schur complement we can link the eigenvalues and eigenvectors of the Schur complement to the eigenvectors and eigenvalues of the generator $\mathscr{L}_0$.

Lemma 3.3.3. *For the the Schur complement of $\mathscr{A}$ defined for $\gamma \in \mathbb{C} \setminus \{0\}$ by $S_1(\gamma) := -\frac{1}{\gamma}BB^* + (A - \gamma)$ we have $\sigma_p(\mathscr{L}_0) = \sigma_p(\mathscr{A}) = \sigma_p(S_1(\gamma)) := \{\gamma \in \mathbb{C} \setminus \{0\} : 0 \in \sigma_p(S(\gamma))\}$.*

Proof. Take $\gamma \in \sigma_p(\mathscr{A}) \setminus \{0\}$ with eigenvector $(f, g) \in D(\mathscr{A})$ such that $\mathscr{A}\begin{bmatrix} f \\ g \end{bmatrix} = \gamma \begin{bmatrix} f \\ g \end{bmatrix}$. This implies that $Af + Bg = \gamma f$ and $Cf = \gamma g$, which in particular gives us that $Cf \in W_0^{1,2}(\mathbb{T}, \mu)$ such that $f \in D(S_1(\gamma))$. Which allows us to conclude that $Af + \frac{1}{\gamma}BCf = \gamma f$ which implies that $Af - \gamma f - \frac{1}{\gamma}BB^*f = 0$ and thus that $S_1(\gamma)f = 0$. If $f = 0$ then we would have $Bg = 0$ and because B is injective we would have $g = 0$ which would contradict (f, g) being an eigenvector for $\mathscr{A}$. Thus we have that f is an eigenvector for $S_1(\gamma)$ and thus $\gamma \in \sigma_p(S_1(\gamma))$.

Now assume that we have $\gamma \in \sigma_p(S_1(\gamma))$ and an eigenvector $f \in D(S_1(\gamma))$ such that $S_1(\gamma)f = 0$. This means that we have $Af + \frac{1}{\gamma}BCf = \gamma f$. We then define $g := \frac{1}{\gamma}Cf$ which shows us that $g \in W^{1,2}(\mathbb{T}, \mu)$ by definition of $D(S_1(\gamma))$. So we have $Cf = \lambda g$ and $Af + Bg = \gamma f$, so $\mathscr{A}\begin{bmatrix} f \\ g \end{bmatrix} = \gamma \begin{bmatrix} f \\ g \end{bmatrix}$ which implies that $\gamma \in \sigma_p(\mathscr{L}_0)$. □

Note that the eigenvectors of the Schur complement are not the same as the eigenvectors of $\mathscr{L}_0$ but it can be seen that for an eigenvector f such that $\mathscr{L}_0 f = \gamma f$ that we then have that $S_1(\gamma)(f^+ - f^-) = 0$.

Lemma 3.3.4. *For $\gamma \in \mathbb{C}\setminus\{0\}$, $S_1(\gamma) : D(S_1(\gamma)) = D(BB^*) \subset L^2(\mathbb{T},\mu) \to L^2(\mathbb{T},\mu)$ is a closed operator.*

Proof. Take a sequence $(f)_{n\geq 1} \subset D(S_1(\gamma)) = D(BB^*)$ such that $f_n \xrightarrow{L^2(\mathbb{T},\mu)} f$ and $S_1(\gamma)f_n \xrightarrow{L^2(\mathbb{T},\mu)} g \in L^2(\mathbb{T},\mu)$ as $n\to\infty$. We then have that $(A-\gamma)f_n \xrightarrow{L^2(\mathbb{T},\mu)} (A-\gamma)f$ by the boundedness of A. The convergence of $S_1(\gamma)f_n$ gives us that

$$BB^* f_n \xrightarrow{L^2(\mathbb{T},\mu)} -\gamma(g-(A-\gamma)f)$$

and in particular that $(BB^* f_n)_{n\geq 1}$ is a Cauchy sequence. Then for each $\epsilon > 0$ take an N such that for $n,m \geq N$ we have that $||BB^*(f_n - f_m)||_{L^2(\mathbb{T},\mu)} \leq \epsilon$ and $||(f_n - f_m)||_{L^2(\mathbb{T},\mu)} \leq \epsilon$ we then have

$$\begin{aligned} ||B^*(f_n - f_m)||^2_{L^2(\mathbb{T},\mu)} &= |\langle B^*(f_n - f_m), B^*(f_n - f_m)\rangle| \\ &= |\langle BB^*(f_n - f_m), (f_n - f_m)\rangle| \\ &\leq ||BB^*(f_n - f_m)||_{L^2(\mathbb{T},\mu)} ||(f_n - f_m)||_{L^2(\mathbb{T},\mu)} \leq \epsilon^2. \end{aligned}$$

This gives us that $(B^* f_n)_{n\geq 1}$ is a Cauchy sequence and thus that $B^* f_n \xrightarrow{L^2(\mathbb{T},\mu)} h$ for some $h \in L^2(\mathbb{T},\mu)$. Because U' is bounded we have that $U' f_n \xrightarrow{L^2(\mathbb{T},\mu)} U' f$. So we have that ∂f_n is Cauchy by the relation $\partial_x f_n = U' f_n - B^* f_n$. Because $W^{1,2}(\mathbb{T},\mu)$ is a Banach space this means that $\partial_x f_n \to \partial_x f$ and thus we have that $f \in W^{1,2}(\mathbb{T},\mu)$ and $\partial_x f = U' f - h$. So we have that $B^* f = h$.

Since $(BB^* f_n)_{n\geq 1}$ is Cauchy, we have that Bh_n is Cauchy in $L^2(\mathbb{T},\mu)$ with $BB^* f = Bh_n$ and therefore $(h_n)_{n\geq 1}$ is Cauchy in the Sobolev space $W^{1,2}(\mathbb{T},\mu)$. Similarly, this gives us that $Bh_n \xrightarrow{L^2(\mathbb{T},\mu)} Bh = -\gamma(g-(A-\gamma)f)$. So we have that $f \in D(S_1(\gamma))$. Moreover, we see that $-\frac{1}{\gamma}BB^* f + (A-\gamma)f = g$, so $S_1(\gamma)f = g$. So $(S_1(\gamma), D(BB^*))$ is closed. □

3.4. The Schur complement and bounds on the eigenvalues

Now we can make use of the Schur complement and treat its point spectrum $\sigma_p(S_1)$ as being the same as the spectrum of the generator of the Zig-Zag process minus $\gamma = 0$. By multiplying the Schur complement with γ we get the equation $\gamma S(\gamma) = -BB^* + \gamma A - \gamma^2 I$. We then see that if we want to find the eigenvalues of $S_1(\gamma)$, we need to find the eigenvalues of a quadratic eigenvalue problem. Analyses towards such problems have been done for example in the case where the operators are matrices [TM01]. Even in the finite-dimensional case, the analysis of quadratic eigenvalue problems has significant differences from the analysis of the conventional linear eigenvalue problems. For example, there can be an infinite amount of eigenvalues and an eigenvector can correspond to multiple eigenvalues at the same time. Analyses have been done in a general case towards operators of this form [EL04], here these operators have been associated with oscillations of a non-homogeneous string inside a viscous substance. However, they describe a more general case with fewer assumptions on B and A and were only able to describe the eigenvalues in the range $(0,\infty)$.

To examine the locations of the eigenvalues we will first examine the numerical range of the Schur complement in order to prove bounds on the eigenvalues of the generator. Specifically we have that all of the real components of the eigenvalues are bounded on an interval and thus that all the eigenvalues lie on a vertical strip. Furthermore, this strip is a smaller strip than the one we found in Proposition 3.3.1.

Lemma 3.4.1. *For all $\gamma \in \sigma_p(\mathscr{L}_0)$ we have that $\mathrm{Re}\,\gamma \in [-||W||_\infty, 0] = [-|||U'(x)| + 2\lambda_0(x)||_\infty, 0]$. Moreover, if $\gamma \in \mathbb{C}\setminus\mathbb{R}$ we also have that $\mathrm{Re}\,\gamma \in [-\frac{1}{2}||W||_\infty, -\operatorname{ess\,inf}_x \frac{W}{2}(x)] = [-||(\frac{|U'(x)|}{2} + \lambda_0(x))||_\infty, -\operatorname{ess\,inf}_x(\frac{|U'(x)|}{2} + \lambda_0(x))]$.*

Proof. We have from lemma 3.3.3 that $\sigma_p(\mathscr{L})_0 = \sigma_p(S_1)$. We take $a + bi = \gamma \in \sigma_p(S_1)$ and a corresponding eigenvector $f \in W^{2,2}(\mathbb{T},\mu)$ such that $S_1(\gamma)f = 0$. We first prove the second bound so assume that $b \neq 0$, we then have that

$$\begin{aligned} S_1(\gamma)f = 0 \quad &\to \quad BB^* f = \gamma(A-\gamma)f, \\ &\to \quad \langle BB^* f, f\rangle = \langle \gamma(A-\gamma)f, f\rangle, \\ &\to \quad ||B^* f||^2_{L^2(\mathbb{T},\mu)} = \langle \gamma(A-\gamma)f, f\rangle, \\ &\to \quad 0 = \mathrm{Im}\langle \gamma(A-\gamma)f, f\rangle \\ &= -\int_{\mathbb{T}} \mathrm{Im}(\gamma(W+\gamma))|f|^2 d\mu \\ &= -\int_{\mathbb{T}} \mathrm{Im}((a+bi)W + (a^2+2abi-b^2))|f|^2 d\mu \\ &= -\int_{\mathbb{T}} b(W+2a)|f|^2 d\mu. \end{aligned}$$

We have that this implies that $W+2a$ cannot be an almost everywhere positive or negative function as otherwise we would have $\int_{\mathbb{T}}(W+2a)|f|^2 d\mu \neq 0$ which would lead to a contradiction. This would imply that we would need both $\operatorname{ess\,sup}_x W + 2a \geq 0$ and $\operatorname{ess\,inf} W + 2a \leq 0$. Which implies that $a \in [-\operatorname{ess\,sup}_x \frac{W}{2}(x), -\operatorname{ess\,inf} \frac{W}{2}(x)]$. This gives us the bound for all the non-real eigenvalues.

In order to prove the bound for all of $\sigma_p(\mathscr{L}_0)$ we only have to look at the real eigenvalues as the statement for the non-real eigenvalues already implies this bound. So now assume that $\gamma \in \mathbb{R}$.

$$\begin{aligned} S(\gamma)f = 0 \quad &\to \quad BB^* f = \gamma(A-\gamma)f, \\ &\to \quad \langle BB^* f, f\rangle = \langle \gamma(A-\gamma)f, f\rangle, \\ &\to \quad ||B^* f||^2_{L^2(\mathbb{T},\mu)} = \langle \gamma(A-\gamma)f, f\rangle, \\ &\to \quad 0 \leq \langle \gamma(A-\gamma)f, f\rangle \\ &\quad\quad = \langle Af, \gamma f\rangle - \langle \gamma f, \gamma f\rangle. \end{aligned}$$

Because we have $\gamma < 0$ this implies that

$$\begin{aligned} \to \quad & \gamma ||f||^2_{L^2(\mathbb{T},\mu)} \geq \langle Af, f\rangle, \\ & = -\int_{\mathbb{T}} W|f|^2 d\nu, \\ \to \quad & \gamma \geq -\frac{\langle \sqrt{W}f, \sqrt{W}f\rangle}{\langle f, f\rangle} \\ & \geq -||\sqrt{W}||^2_{B(L^2(\mathbb{T},\mu))} = -||W||_{L^\infty} = -\operatorname{ess\,sup}_x W(x). \end{aligned}$$

So we have $\gamma \in [-\operatorname{ess\,sup}_x W(x), 0]$ □

We see that the imaginary eigenvalues are bound by the essential supremum and essential infimum of $\frac{W(x)}{2}$. To visualize this bound we can use an approximation fo the generator as described in Chapter $\mathscr{L}_N$ to visualize the bounds described above for two potentials.

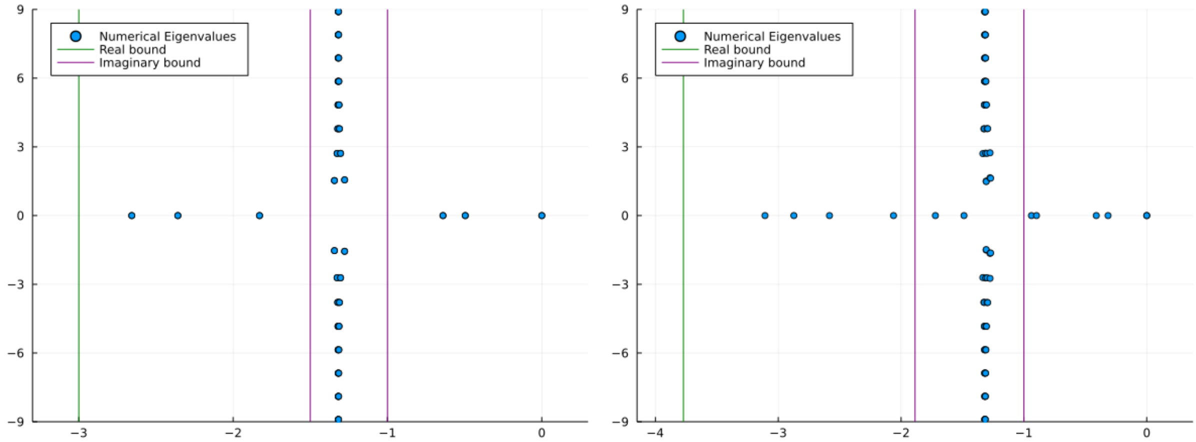

Figure 3.1: Numerical eigenvalues of the operator $\mathscr{L}_N$ with $N = 1000$ which is used to approximate the eigenvalues of $\mathscr{L}$ as described in Chapter 5. The left image corresponds to the $U(x) = 2(\frac{1-\cos(x)}{2})$ and refreshment rate $\lambda_0 = 1$. The right image has the potential $U(x) = 2(\frac{1-\cos(x)}{2})^4$ and the refreshment rate $\lambda_0 = 1$. The purple lines describe the bound on the non-real eigenvalues and the green line represents the bound on the real-valued eigenvalues described in lemma 3.4.1.

Lemma 3.4.1 allows us to get complete control over the distance of the non-real eigenvalues to the line $i\mathbb{R}$ for a given potential $U \in W^{1,\infty}(\mathbb{T})$. In an extreme case we could choose λ_0 such that all of the imaginary values get stuck on a single line by choosing $\lambda_0 = c - \frac{|U'(x)|}{2}$ with $c > \operatorname{ess\,sup}_x \frac{|U'(x)|}{2}$, we would have that $W = 2c$ and thus all the eigenvalues with a non-zero imaginary component of $\mathscr{L}_0$ will have the same value and be on the vertical line with the real component being equal to -c. Moreover, for potentials that have a discontinuous derivative, we have that automatically all of the non-real eigenvalues are placed away from $i\mathbb{R}$ without having to add a refreshment rate.

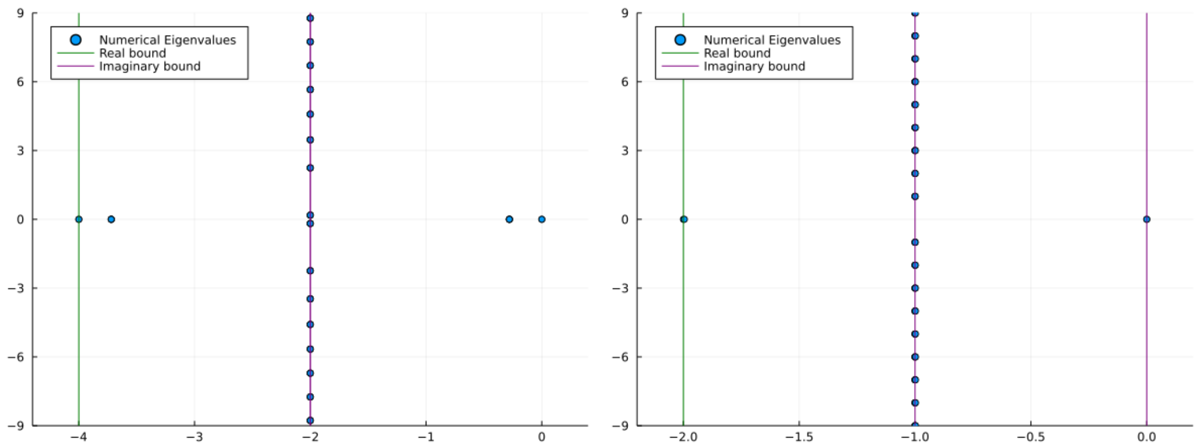

Figure 3.2: Numerical eigenvalues of the operator $\mathscr{L}_N$ with $N = 1000$ which is used to approximate the eigenvalues of $\mathscr{L}_0$ as described in Chapter 5. The left image corresponds to the $U(x) = \frac{1-\cos(x)}{2}$ and refreshment rate $\lambda_0(x) = 2 - \frac{|\sin(x)|}{4}$. By choosing this refreshment rate we get that $W = 4$ and thus according to lemma 3.4.1 we have that the non-real eigenvalues appear on a line with its real component being equal to 2. The right image has the potential $U(x) = 2|x|$ and the refreshment rate $\lambda_0 = 0$. Because we have that $W(x) = |U'(x)| = 2|\operatorname{sign}(x)| = 2$ we have again that all of the non-real eigenvalues are on a line with real components equal to 1. The purple lines describe the bound on the non-real eigenvalues and the green line represents the bound on the real-valued eigenvalues described in lemma 3.4.1.

Remarkably the potential $U(x) = \sigma|x|$, with $\sigma \in \mathbb{R}$ seems to have its spectral gap increase $\kappa \to \infty$ as $\sigma \to \infty$ which turns out to be true as we shall see in Chapter 4. In general, it seems to be the case that if $\lambda_0 = 0$ that a scaling of the form $\sigma \to \sigma|U'|$ gives us that the spectral gap increases indefinitely by using the discretization of Chapter 5, however, this is not necessarily of practical relevance as we want to obtain the optimal λ_0 for a given potential that we want to estimate.

Using lemma 3.4.1 and Proposition 3.2.9 we can now get a bound on all of the eigenvalues. Moreover, this gives us a bound on the spectral gap.

Theorem 3.4.2. *Take $U \in W^{1,\infty}(\mathbb{T})$ and $\lambda_0 \in L^\infty(\mathbb{T})$. If we have* $\operatorname{ess\,inf}_{x\in\mathbb{T}} W(x) = \operatorname{ess\,inf}_{x\in\mathbb{T}} |U'(x)| + 2\lambda_0(x) > 0$ *then we have that $\kappa > 0$ (the spectral gap is positive). Moreover, a lower bound on κ is given by*

$$\kappa \geq \min\{\operatorname{ess\,inf} \frac{1}{2} W(x), ||\mathscr{L}_0^{-1}||^{-1}_{L^2(\mathbb{T},\nu)}\}.$$

Proof. We have by lemma 3.4.1 that the real component of eigenvalues on $\mathbb{C}\setminus\{\mathbb{R}\}$ are bounded by $\inf_{x\in\mathbb{T}} W(x) > 0$. Next, we have by lemma 3.2.9 that $0 \in \rho(\mathscr{L}_0)$. We can then make use of the Neumann series of the resolvent around 0 to conclude that $(\frac{1}{||\mathscr{L}_0^{-1}||}, 0] \subset \rho(\mathscr{L}_0)$ which gives us a bound on the real-valued eigenvalues as well. Taking a minimum of the bound for the real-valued eigenvalues and the eigenvalues with an imaginary component gives the bound for all of the eigenvalues. Moreover, Proposition 4.3.2 gives us that $\sigma(\mathscr{L}_0) = \sigma_p(\mathscr{L}_0)$ so these bounds on the eigenvalues are also bounds on the full spectrum of $\mathscr{L}_0$ and thus also the spectral gap. □

The lower bound on the spectral gap as it is presented is difficult to calculate due to the norm of the resolvent $||\mathscr{L}_0^{-1}||^{-1}_{L^2(\mathbb{T},\nu)}$ being unknown. In a similar way to lemma 3.2.8 it can be seen that $\mathscr{L}_0^{-1}$ is also a Hilbert-Schmidt operator. Then one way of making a bound on this resolvent norm is by making use of the Hilbert-Schmidt norm of the operator. We then would have $||\mathscr{L}_0^{-1}||_{L^2(\mathbb{T},\nu)} \leq ||\mathscr{L}_0^{-1}||_{HS}$ which would imply that another lower bound on the spectral gap is given by

$$\kappa \geq \min\{\operatorname{ess\,inf} \frac{1}{2} W(x), ||\mathscr{L}_0^{-1}||^{-1}_{HS}\}.$$

Whereas $\kappa = \sup\{\operatorname{Re}\gamma : \gamma \in \sigma_p(\mathscr{L}_0)\}$ the Hilbert-Schmidt norm has the property $\sqrt{\sum_{\lambda\in\sigma_p(\mathscr{L}_0)} \frac{1}{|\lambda|^2}} \leq ||\mathscr{L}_0^{-1}||_{HS}$ so such an estimate can never be equal to the spectral gap, however it could be of a similar order when compared to scaling of W, U or λ_0.

Lemma 3.4.1 gives us a bound on all of the eigenvalues of $\mathscr{L}_0$, it says that all of the eigenvalues are in a vertical strip with the width of this strip being equal to $||W||_\infty$. Furthermore, it gives us an upper bound on the real component of the eigenvalues with a non-zero imaginary component. Using this we can say that the spectral gap will be determined by a real eigenvalue of $S_1(\gamma)$ or it will at least be equal to $-\operatorname{ess\,inf}_x \frac{W}{2}(x)$ which is controllable by increasing λ_0 on the locations where $|U'|$ has a small value. So in order to investigate the spectral gap, we should examine the real eigenvalues of the Schur complement.

The bounds on the eigenvalues and the difference between the eigenvalues in $\mathbb{R}$ compared to those in $\mathbb{C}\setminus\mathbb{R}$ can be better examined by looking at the following proposition.

Proposition 3.4.3. *For all $\gamma \in \sigma_p(S_1)$ with a corresponding eigenvector f such that $S_1(\gamma)f = 0$ we have*

$$\gamma = -\frac{||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)}}{2||f||^2_{L^2(\mathbb{T},\mu)}} \pm \sqrt{\frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}} - \frac{||B^*f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}}.$$

Specifically this shows that if $\gamma \in \mathbb{R}$ that we then have

$$\frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}} \geq \frac{||B^*f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}.$$

Furthermore, if $\gamma \in \mathbb{C}\setminus\mathbb{R}$ then we have

$$\frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}} < \frac{||B^*f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}},$$

$$\operatorname{Re}\gamma = -\frac{||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)}}{2||f||^2_{L^2(\mathbb{T},\mu)}} = \frac{\langle Af, f\rangle}{2||f||^2_{L^2(\mathbb{T},\mu)}} = -\frac{\langle(\frac{U'}{2} + \lambda_0)f, f\rangle}{||f||^2_{L^2(\mathbb{T},\mu)}},$$

and

$$|\gamma|^2 = \frac{||B^*f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}.$$

Proof. Take $a, b \in \mathbb{R}$ and define $\gamma = a + bi \in \sigma_p(S_1)$ and a corresponding eigenvector f such that $S_1(\gamma)f = 0$, we then have

$$\begin{aligned} S(\gamma)f = 0 \quad &\to \quad BB^* f = \gamma(A-\gamma)f \\ &\to \quad \langle BB^* f, f\rangle = \langle \gamma(A-\gamma)f, f\rangle \\ &\to \quad ||B^* f||^2_{L^2(\mathbb{T},\mu)} = \langle \gamma(A-\gamma)f, f\rangle. \end{aligned}$$

By using the equality $\langle Af, f\rangle = -||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)}$ this final equality then implies

$$||B^* f||^2_{L^2(\mathbb{T},\mu)} + \gamma||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)} + \gamma^2||f||^2_{L^2(\mathbb{T},\mu)} = 0.$$

This is a quadratic formula with its solution being

$$\gamma = -\frac{||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)}}{2||f||^2_{L^2(\mathbb{T},\mu)}} \pm \sqrt{\frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}} - \frac{||B^* f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}}.$$

If we assume that $\gamma \in \mathbb{R}$ we can see that the square root has to be real and therefore we have that

$$\frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}} \geq \frac{||B^* f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}.$$

Otherwise if we assume that $\gamma \in \mathbb{C}\setminus\mathbb{R}$ we see that this square root has to become imaginary meaning that

$$\frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}} < \frac{||B^* f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}.$$

We then also see that for these eigenvalues with non-zero imaginary component that the real component of the eigenvalue has to be given by

$$\operatorname{Re}\gamma = \frac{\langle Af, f\rangle}{2||f||^2_{L^2(\mathbb{T},\mu)}} = -\frac{||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)}}{2||f||^2_{L^2(\mathbb{T},\mu)}}. \tag{3.12}$$

And the imaginary component is given by

$$\operatorname{Im}\gamma = \pm\sqrt{\frac{||B^* f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}} - \frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}}}.$$

Furthermore, we then see that the magnitude of the eigenvalues with non-zero imaginary component are given by

$$|\gamma|^2 = (\operatorname{Re}\gamma)^2 + (\operatorname{Im}\gamma)^2 = \frac{||B^* f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}.$$

□

Proposition 3.4.3 gives some insight in how the eigenvalues of an eigenvector f behave with respect to their corresponding eigenfunctions. Setting $||f||_{L^2(\mathbb{T},\mu)} = 1$ we see that whether an eigenvalue is real or not has to do with the relationship between $||B^* f||^2_{L^2(\mathbb{T},\mu)}$ and $||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}$. Which directly relates to the comparison of the operator A and B^* of the block operator $\mathscr{A}$. Moreover, we get an interesting relationship

$$\sum_{\gamma\in\sigma_p(\mathscr{L}_0), f\in W_0^{1,2}(\mathbb{T},\mu): \mathscr{L}_0 f = \gamma f} \frac{||f^+ - f^-||^2_{L^2(\mathbb{T},\nu)}}{||B^*(f^+ - f^-)||^2_{L^2(\mathbb{T},\mu)}} = \sum_{\gamma\in\sigma_p(\mathscr{L}_0)} \frac{1}{|\gamma|^2} \leq ||\mathscr{L}_0^{-1}||^2_{HS}.$$

This shows that the rate at which the ordered sequence $(\frac{||B^* f||}{||f||})_{n\geq 1}$ grows when having the corresponding list of the eigenvectors $(f_n)_{n\geq 1}$ should be at a pace high enough for the sequence to converge.

Returning to the quadratic numerical range $W^2(\mathscr{A})$ we see that similar forms for an estimation of the spectrum comes up. We do this by examining the following functionals for $x, y \in W^{1,2}(\mathbb{T},\mu) \bigoplus W_0^{1,2}(\mathbb{T},\mu)$ defined by

$$\lambda_\pm \begin{pmatrix} x \\ y \end{pmatrix} := \frac{1}{2}\left(\frac{\langle Ax, x\rangle}{||x||^2_{L^2(\mathbb{T},\mu)}} + \frac{\langle Dy, y\rangle}{||y||^2_{L^2(\mathbb{T},\mu)}} \pm \sqrt{\left(\frac{\langle Ax, x\rangle}{||x||^2_{L^2(\mathbb{T},\mu)}} - \frac{\langle Dy, y\rangle}{||y||^2_{L^2(\mathbb{T},\mu)}}\right)^2 + 4\frac{\langle By, x\rangle\langle Cx, y\rangle}{||x||^2_{L^2(\mathbb{T},\mu)}||y||^2_{L^2(\mathbb{T},\mu)}}}\right) \tag{3.13}$$

which are exactly the elements of $\sigma_p(\mathscr{A}_{f,g})$ as defined in formula (3.11) (they are the solutions of the determinants of the 2x2 matrices defined by the quadratic numerical range). For our case they reduce to

$$\lambda_\pm \begin{pmatrix} x \\ y \end{pmatrix} := \frac{1}{2}\left(\frac{\langle Ax, x\rangle}{||x||^2_{L^2(\mathbb{T},\mu)}} \pm \sqrt{\left(\frac{\langle Ax, x\rangle}{||x||^2_{L^2(\mathbb{T},\mu)}}\right)^2 - 4\frac{|\langle By, x\rangle|^2}{||x||^2_{L^2(\mathbb{T},\mu)}||y||^2_{L^2(\mathbb{T},\mu)}}}\right)$$

We can then see if we take $y = B^* x$ that the above expression reduces to that of the expression of the eigenvalues of Proposition 3.4.3 such that $\lambda_\pm \begin{pmatrix} x \\ B^* x \end{pmatrix} = S_1(\gamma)x$. Which gives us that $W(S_1(\gamma)) \subset W^2(\mathscr{A})$.

By choosing y to orthogonal to $B^* x$ we get that $\frac{|\langle By, x\rangle|^2}{||x||^2_{L^2(\mathbb{T},\mu)}||y||^2_{L^2(\mathbb{T},\mu)}} = 0$ moreover we can choose $y = B^* x$ which gives $\frac{|\langle By, x\rangle|^2}{||x||^2_{L^2(\mathbb{T},\mu)}||y||^2_{L^2(\mathbb{T},\mu)}} = \frac{|\langle B^* x, B^* x\rangle|^2}{||x||^2_{L^2(\mathbb{T},\mu)}||B^* x||^2_{L^2(\mathbb{T},\mu)}} = \frac{\langle B^* x, B^* x\rangle}{||x||^2_{L^2(\mathbb{T},\mu)}}$. We can then see this as the numerical range of the linear operator $y \to \langle y, B^* x\rangle$ and so we can make use of the Toeliptz-Hausdorff Theorem [Gus70] to see that all of the values between 0 and $\frac{||B^* x||_{L^2(\mathbb{T},\mu)}}{||x||_{L^2(\mathbb{T},\mu)}}$ can be attained for every $x \in W^{1,2}(\mathbb{T},\mu)$. This in particular implies that by taking an x such that $\langle Ax, x\rangle$ is non-zero that all the small values around 0 on the real axis are part of the quadratic numerical range and thus that the quadratic numerical range will not allow us to say something about the real part of the spectral gap.

However, this is where the usefulness of the Schur complement comes in, namely for the Schur complement we can prove that the elements close to 0 on the real number line are not part of the numerical range of the Schur complement. So we really have that the inclusion of the numerical range of the Schur complement into the quadratic numerical is a proper one.

We therefore return to the analysis of the Schur complement. In Proposition 3.4.3 we see that there are two roots to every eigenfunction. This is often included in the analysis of the quadratic numerical range, for this the following sets are defined:

$$\Lambda_\pm = \{\lambda_\pm \begin{pmatrix} x \\ y \end{pmatrix} : (x, y) \in D(\mathscr{A}) \text{ and } \left(\frac{\langle Ax, x\rangle}{||x||^2_{L^2(\mathbb{T},\mu)}}\right)^2 - 4\frac{|\langle By, x\rangle|^2}{||x||^2_{L^2(\mathbb{T},\mu)}||y||^2_{L^2(\mathbb{T},\mu)}} \geq 0\}.$$

Again the big difference being the $\frac{|\langle By, x\rangle|^2}{||x||^2_{L^2(\mathbb{T},\mu)}||y||^2_{L^2(\mathbb{T},\mu)}}$ term. These sections as in this definition are usually only about eigenvalues which are real-valued. Moreover, most observations that use these functionals are for operators that have real quadratic numerical range (and thus a real spectrum).

For J-self-adjoint operators these two branches have a particular meaning if we have an eigenvalue $\gamma \in \mathbb{C}\setminus\mathbb{R}$ with an eigenvector f such that $S_1(\gamma)f = 0$ we can then see that $S_1(\overline{\gamma})\overline{f} = 0$. So we see that $\overline{\gamma} \in \sigma_p(S_1)$ is an eigenvalue with its eigenvector given by $\overline{f}$. However, we do not have such a connection between the eigenvectors when it comes to real eigenvalues. Nevertheless, when it comes to numerical approximations these paired eigenvalues do seem to behave similarly when scaling the potential and the refreshment rate. The lack of connection for real-valued eigenvalues, causes the eigenvalues to be categorized into three groups and allows us to give an expression of the spectral gap. We define the set of eigenvectors of $S_1(\gamma)$ by $M = \{f \in D(BB^*) : ||f||_{L^2(\mathbb{T},\mu)} = 1, S_1(\gamma)f = 0\}$. Furthermore, define $\gamma_\pm(f) = \frac{1}{2}||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)} \pm \sqrt{\frac{1}{4}||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)} - ||B^* f||^2_{L^2(\mathbb{T},\mu)}}$

then the spectral gap is given by

$$\kappa = \inf_{f\in M}\begin{cases}\frac{1}{2}||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)} & \text{if } \frac{1}{4}||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)} < ||B^*f||^2_{L^2(\mathbb{T},\mu)},\\ \gamma_+(f) & \text{if } S_1(\gamma_+(f))f = 0,\\ \gamma_-(f) & \text{otherwise.}\end{cases}$$

One of the issues with this formula for κ is that we do not know much about the functions f in M. If we were to take the infimum over all of the functions of $D(BB^*)$ we would have that the spectral gap is equal to 0 due to the function $e^{U(x)} \in D(BB^*)$. This function corresponds to a function which has no 'curvature' with respect to the derivative operator B^* as $B^*e^{U(x)} = 0$. It can also be seen as an eigenfunction to $S_1(0)$ if we define it as the limit of the Schur complement viewed as $\gamma S_1(\gamma) = -BB^* + \gamma A - \gamma^2$. We will come back to this function later.

Another issue is that we have strong conditionals in the expression of κ, specifically the distinction between $S_1(\gamma_+(f))f = 0$ (and implicitly $S_1(\gamma_-(f))f = 0$). To solve this we can ignore this distinction and create a lower bound to $\tilde{\kappa} \le \kappa$ defined by

$$\tilde{\kappa} := \inf_{f\in M}\frac{1}{2}||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)} - 1_{\{\frac{1}{4}||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)} \ge ||B^*f||^2_{L^2(\mathbb{T},\mu)}\}}\sqrt{\frac{1}{4}||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)} - ||B^*f||^2_{L^2(\mathbb{T},\mu)}}.$$

We can then see that for the eigenfunctions with imaginary component the spectral gap is bounded below by $\inf_{f\in D(BB^*)}\frac{1}{2}||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)}$. So to find a lower bound we should only examine the real eigenvalues and find a lower bound for the expression $\frac{1}{2}||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)} - \sqrt{\frac{1}{4}||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)} - ||B^*f||^2_{L^2(\mathbb{T},\mu)}}$. Although the condition $\frac{||\sqrt{W}f||^4_{L^2(\mathbb{T},\mu)}}{4||f||^4_{L^2(\mathbb{T},\mu)}} \ge \frac{||B^*f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}}$ restricts the function space, we still have that the function space is too large due to the kernel function of B^* which is $e^{U(x)}$. If we could only examine the orthogonal complement of $e^U(x)$ then because B^* becomes invertible and we would get the following "bound"

$$\overline{\kappa} := \begin{cases}\min\{\operatorname{ess\,inf}\frac{1}{2}W(x),\ \frac{1}{2}||\sqrt{W}||^2_\infty - \sqrt{\frac{1}{4}||\sqrt{W}||^4_\infty - ||(B^*)^{-1}||^{-1}}\} & \text{if} \frac{1}{4}||\sqrt{W}||^4_\infty > ||(B^*)^{-1}||^{-1},\\ \operatorname{ess\,inf}\frac{1}{2}W(x) & \text{otherwise.}\end{cases}$$

It turns out that we can fix this problem however the bound will become slightly different in how it is written. If we can prove that the eigenvalues corresponding to real eigenvalues hold to a bound similar to this bound, then we are done since this bound already holds for all of the eigenvalues in $\mathbb{C}\setminus\mathbb{R}$ due to lemma 3.4.1. To further investigate the real eigenvalues we will examine $S_1(\gamma)$ when $\gamma \in \mathbb{R}$. We will first prove that the spectrum of this operator is real. We do this by making use of the Friedrichs Extension Theorem. The Friedrichs extension defines a unique self-adjoint operator from a quadratic form. We can construct every $S_1(\gamma)$ as such a quadratic form, however a more convenient approach is by examining the quadratic form $q(f,g) = (B^*f, B^*g)$ associated to the operator $BB^*f = \partial_x^2 f - U'\partial_x f - U''f$.

Lemma 3.4.4. *For $\gamma \in \mathbb{C}\setminus\{0\}$ we have that $(S_1(\gamma))^* = S_1(\overline{\gamma})$. In particular this means for $\gamma \in \mathbb{R}\setminus\{0\}$ that $S_1(\gamma)$ is self-adjoint.*

Proof. We first examine the operator $(BB^*, D(BB^*))$. We have for $f \in D(BB^*)$ that $\langle BB^*f|f\rangle = \langle B^*f|B^*f\rangle \ge 0$ meaning that this operator is positive. So by the Friedrichs Extension Theorem we have that there is a unique positive self-adjoint extension $(BB^*)^\times$ such that $BB^* \subset (BB^*)^\times$.

The associated linear form of $(BB^*)^\times$ is $q : D(B^*) \times D(B^*) \to \mathbb{C}$ defined by $\forall f,g \in D(B^*)$ by $q(f,g) = \langle B^*f|B^*g\rangle$. It holds that the domain of $(BB^*)^\times$ is a subset of the domain of the completion of q defined by $D((BB^*)^\times) = \{f \in D(\overline{q}) : \exists g \in L^2(\mathbb{T},\mu) \text{ with } (g|v) = q(f,v)\forall v \in D(\overline{q})\}$. For q to be closed we need that the Hilbert space associated to the norm $||f||_q^2 := q(f|f) + (f|f)_{L^2(\mathbb{T},\mu)}$ defines a complete space on $D(B^*)$. This holds automatically because B^* is a closed operator and this $||f||_q^2$ norm is the same as the graph norm of B^* which is thus closed. So we have that $D(\overline{q}) = D(B^*)$.

We have then have that $D((BB^*)^\times) = \{f \in W^{1,2}(\mathbb{T},\mu) : \exists g \in L^2(\mathbb{T},\mu) \text{ with } (g|v) = (B^*f|B^*v)\ \forall v \in W^{1,2}(\mathbb{T},\mu)\}$. This shows that $B^*f \in D(B^{**}) = D(B) = W_0^{1,2}(\mathbb{T},\mu)$. By definition this means that $f \in D(BB^*)$. This shows $D((BB^*)^\times) \subset D(BB^*)$. But because we had $BB^* \subset (BB^*)^\times$ we have that $BB^* = (BB^*)^\times$.

This gives us that $(BB^*)^* = BB^*$. Furthermore, because in general it holds that if K is a bounded operator that we have $(BB^* + K)^* = BB^* + K^*$ we then see that

$$\begin{aligned}(S_1(\gamma))^* &= (-\frac{1}{\gamma}(BB^*) + (A-\gamma))^* \\ &= -\frac{1}{\overline{\gamma}}((BB^*))^* + (A-\gamma)^*, \\ &= -\frac{1}{\overline{\gamma}}(BB^*) + (A-\overline{\gamma}), \\ &= (S_1(\overline{\gamma})).\end{aligned}$$

□

The relationship $S_1(\gamma)^* = S_1(\overline{\gamma})$ makes $S_1(\gamma)$ a self-adjoint pencil which have some interesting properties [Mar88]. However, for our proof, we will not make use of these properties.

Next, we want to show that $S_1(\gamma)$ has a compact resolvent. This can be confusing terminology because it could imply that only $S_1(\gamma)^{-1}$ is compact for all $\gamma \in \mathbb{C} \setminus \{0\}$ but equivalently we can show that a wider range of these operators are compact as shown in the following lemma.

Lemma 3.4.5. *For $\gamma \in \mathbb{C} \setminus \{0\}$ we have that $S_1(\gamma)$ has compact resolvent, that is for all $\eta \in \mathbb{C}$ such that $(\eta - S_1(\gamma))^{-1}$ exists and is bounded we have that $(\eta - S_1(\gamma))^{-1}$ is a compact operator*

Proof. We have that BB^* is a self-adjoint operator as we have proven in lemma 3.4.4. Furthermore, it is bounded from below $0 \le \langle BB^* f, f\rangle$. This means that $||(BB^* - \eta)^{-1}||$ exists for $\eta < 0$. We also have the estimate $||(BB^* - \eta)^{-1}||^{-1} \ge d(\eta, \sigma(BB^*))$. Because the spectrum of BB^* is contained in $[0,\infty)$ we have that we can make this lower bound as large as possible by choosing a η further away from 0. In particular we can choose η such that $||\frac{-1}{\gamma}(BB^* - \eta)^{-1}||^{-1} \ge ||(A-\gamma)||$. We can then make use of Theorem IV.3.1 of [Kat95] (in particular remark 3.2) where we use $T = \frac{-1}{\gamma}BB^*$ and $S = S_1(\gamma)$ to conclude that $(S_1(\gamma) - \eta)^{-1}$ exists and is bounded. We then also have that $(S_1(\gamma) - \eta)^{-1}$ is a bounded operator from $(L^2(\mathbb{T},\nu), ||\cdot||_{L^2(\mathbb{T},\nu)}) \to (D(BB^*), ||\cdot||_{\mathscr{G}(S_1(\gamma))})$. Next, we have that the embedding $i_{Graph} : (D(BB^*), ||\cdot||_{\mathscr{G}(S_1(\gamma))}) \to (D(BB^*), ||\cdot||_{W^{1,2}(\mathbb{T},\mu)})$ is continuous. This is because we have for $f \in D(BB^*)$ that

$$\begin{aligned}||f||^2_{W^{1,2}(\mathbb{T},\mu)} &\le 2||\partial_x f||^2_{L^2(\mathbb{T},\mu)} + 2||f||^2_{L^2(\mathbb{T},\mu)} \\ &\le 4||B^* f||^2_{L^2(\mathbb{T},\mu)} + 4||U' f||^2_{L^2(\mathbb{T},\mu)} + 2||f||^2_{L^2(\mathbb{T},\mu)} \\ &\le 4(f, BB^* f) + K||f||^2_{L^2(\mathbb{T},\mu)} \\ &\le 4||BB^* f||_{L^2(\mathbb{T},\mu)}||f||_{L^2(\mathbb{T},\mu)} + K||f||^2_{L^2(\mathbb{T},\mu)} \\ &\le 4|\gamma|||S_1(\gamma) f||_{L^2(\mathbb{T},\mu)}||f||_{L^2(\mathbb{T},\mu)} + 4|\gamma|||(A-\gamma)||_{L^2(\mathbb{T},\mu)}||f||^2_{L^2(\mathbb{T},\mu)} + K||f||^2_{L^2(\mathbb{T},\mu)} \\ &\le C||S_1(\gamma) f||_{L^2(\mathbb{T},\mu)}||f||_{L^2(\mathbb{T},\mu)} + C||f||^2_{L^2(\mathbb{T},\mu)}. \\ &\le C||f||_{\mathscr{G}(S_1(\gamma))}||f||_{W^{1,2}(\mathbb{T},\mu)}.\end{aligned}$$

Dividing both sides by $||f||_{W^{1,2}(\mathbb{T},\mu)}$ gives us that the embedding is continuous. Finally we also have the compact Sobolev embedding $i_S : (D(BB^*), ||\cdot||_{W^{1,2}(\mathbb{T},\mu)}) \to (D(BB^*), ||\cdot||_{L^2(\mathbb{T},\mu)})$. We get that

$$(S_1(\gamma) - \eta)^{-1} = i_S \circ i_{Graph} \circ (S_1(\gamma) - \eta)^{-1}$$

where the operator $(S_1(\gamma) - \eta)^{-1}$ on the left-hand side is seen as a mapping $(L^2(\mathbb{T},\nu), ||\cdot||_{L^2(\mathbb{T},\nu)}) \to (D(BB^*), ||\cdot||_{L^2(\mathbb{T},\nu)})$ and on the right-hand side it is seen as a mapping on $(L^2(\mathbb{T},\nu), ||\cdot||_{L^2(\mathbb{T},\nu)}) \to (D(BB^*), ||\cdot||_{\mathscr{G}(S_1(\gamma))})$.

Due to all of the operators being continuous and i_S being a compact mapping, we get that $S_1(\gamma) - \eta)^{-1}$ is compact and by the resolvent identity, all $(\eta - S_1(\gamma))^{-1}$ are compact for all $\eta \in \mathbb{C}$ when this inverse exists and is bounded. □

Theorem III.6.29 [Kat95] then give us that the spectrum of $S_1(\gamma)$ exists entirely out of eigenvalues of finite multiplicity. That is if $(S_1(\gamma) - \eta)^{-1}$ does not exists or cannot be bounded then we have that there is an $f \in D(BB^*)$ such that $(S_1(\gamma) - \eta) f = 0$.

Spectral theorems about self-adjoint operators with compact resolvent then gives us that are an infinite amount of eigenvalues $\eta \in \mathbb{C}$ with the property $S_1(\gamma)f = \eta f$. Since $S_1(\gamma)$ is bounded below for $\mathrm{Re}\gamma < 0$ since for $f \in D(BB^*)$ we have $\langle S(\gamma)f, f\rangle \geq -||\sqrt{W}||^2_{L^2(\mathbb{T},\mu)} - \gamma$ due to $-\frac{1}{\gamma}\langle B^* f, B^* f\rangle \geq 0$ and $A-\gamma$ being a bounded symmetric operator. This gives us in particular that these eigenvalues can be ordered $\eta_1 \leq \eta_2....$

As an interesting side-note for a given potential and a constant refreshment rate λ_0 we have that as we increase λ_0 and if we denote by $S_1^{\lambda_0}(\gamma)f$ the Schur complement corresponding to potential U with refreshment rate λ_0 that we then have.

$$S_1^{\lambda_0}(\gamma) = \frac{-1}{\gamma}BB^* - |U'| - 2\lambda_0 - \gamma = S_1^0(\gamma)f - 2\lambda_0.$$

Since $S_1^0(\gamma)$ has an infinite amount of eigenvalues we see that for an infinite amount of $\lambda_0 \in \mathbb{R}$ that there are $f \in D(BB^*)$ such that we also have that $S_1^{\lambda_0}(\gamma)f = 0$. This gives an indication that if we scale λ_0 that there will be more real eigenvalues corresponding to $\mathscr{L}_0$. We see this happening when using the numerical approximations of $\mathscr{L}_N$ but we will not make use of this in this thesis.

Now that we have for $\gamma \in \mathbb{R}\setminus\{0\}$ that $S_1(\gamma)$ is self-adjoint and that the operator has a compact resolvent, we have that $S_1(\gamma)$ only has a real spectrum with real eigenvalues. Because $S_1(\gamma)$ is also bounded below as mentioned above we can then make use of variational principles, specifically, we can make use of the max-min theorem (Theorem XIII.2 [RS78] and noting that $\sigma_{ess} = \emptyset$ due to the compact resolvent). This theorem is useful for finding the eigenvalues of a self-adjoint operator. Given a self-adjoint operator with eigenvalues $\beta_1 \leq \beta_2 \leq \beta_3 \leq ...$ and its spectrum being equal to its point spectrum, it states that the eigenvalues of the self-adjoint operator $S_1(\gamma)$ are given by:

$$\beta_n = \sup_{f_1,f_2,...f_{n-1}\in D(S_1(\gamma))} \inf_{f\in \mathrm{span}(\{f_1,f_2,...f_{n-1}\})^\perp \cap D(S_1(\gamma)), ||f||_{L^2(\mathbb{T},\mu)}=1} \langle S_1(\gamma)f, f\rangle.$$

The orthogonal complement is defined with respect to our standard inner product $\langle\rangle$ on $L^2(\mathbb{T},\mu)$. For example, we have $\beta_1 = \inf_{f\in D(S_1(\gamma)), ||f||_{L^2(\mathbb{T},\mu)}=1}\langle S_1(\gamma)f, f\rangle$ and by the lower bound which was described above we can see that for small γ that the smallest eigenvalue of $S_1(\gamma)$ is negative. We can use this max-min theorem to prove the following bound on the real eigenvalues of $\mathscr{L}_0$.

Proposition 3.4.6. *Define*

$$a := \begin{cases} \max\{-\frac{\int_{\mathbb{T}} W(x)e^{U(x)}\,dx}{\int_{\mathbb{T}} e^{U(x)}\,dx}, -\frac{1}{2}||W||_\infty + \sqrt{\frac{1}{4}||W||^2_\infty - ||(\tilde{B}^*)^{-1}||^{-2}_{L^2(\mathbb{T},\mu)}}\} & \text{if } \frac{1}{2}||W||_\infty \geq ||(\tilde{B}^*)^{-1}||^{-1}_{L^2(\mathbb{T},\mu)} \\ -\frac{\int_{\mathbb{T}} W(x)e^{U(x)}\,dx}{\int_{\mathbb{T}} e^{U(x)}\,dx} & \text{otherwise.} \end{cases}$$

We then have that $\forall \gamma \in (a, 0]$ *that* $\gamma \in \mathbb{C}\setminus\sigma_p(\mathscr{L}_0) = \rho(\mathscr{L}_0)$

Proof. We will often make use of the identity $\langle S_1(\gamma)f, f\rangle = \frac{-1}{\gamma}||B^* f||^2_{L^2(\mathbb{T},\mu)} - ||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)} - \gamma||f||^2_{L^2(\mathbb{T},\mu)}$. Take $\gamma \in (a, 0]$ and take the function $x \to Ke^{U(x)} \in D(BB^*)$ such that $\frac{1}{K} := ||e^{U(x)}||_{L^2(\mathbb{T},\mu)}$. We can then use the max-min theorem to prove that the first eigenvalue of $S_1(\gamma)$ has the following negative upper bound

$$\begin{aligned}\beta_1 &= \inf_{f\in D(BB^*), ||f||_{L^2(\mathbb{T},\mu)}=1}\langle S_1(\gamma)f, f\rangle \leq \langle S_1(\gamma)Ke^U, Ke^U\rangle \\ &= -\frac{1}{\gamma}||B^* Ke^U||^2_{L^2(\mathbb{T},\mu)} - ||\sqrt{W}Ke^U||^2_{L^2(\mathbb{T},\mu)} - \gamma \\ &= -||\sqrt{W}Ke^U||^2_{L^2(\mathbb{T},\mu)} - \gamma.\end{aligned}$$

So we see that $\beta_1 < 0$ if $-||\sqrt{W}Ke^U||^2_{L^2(\mathbb{T},\mu)} - \gamma < 0$ which means that $-||\sqrt{W}Ke^U||^2_{L^2(\mathbb{T},\mu)} = -\frac{\int_{\mathbb{T}} W(x)e^{U(x)}\,dx}{\int_{\mathbb{T}} e^{U(x)}\,dx} < \gamma$

which holds as $a < \gamma$. Next, we examine a positive lower bound on the second eigenvalue

$$\begin{aligned}
\beta_2 &= \sup_{g\in D(BB^*)} \inf_{f\in D(BB^*): f\perp g, ||f||_{L^2(\mathbb{T},\mu)}=1} \langle S_1(\gamma)f, f\rangle \geq \inf_{f\in D(BB^*): f\perp e^U, ||f||_{L^2(\mathbb{T},\mu)}=1} \langle S_1(\gamma)f, f\rangle \\
&= \inf_{f\in D(BB^*): f\perp e^U, ||f||_{L^2(\mathbb{T},\mu)}=1} -\frac{1}{\gamma}||B^* f||^2_{L^2(\mathbb{T},\mu)} - ||\sqrt{W}f||^2_{L^2(\mathbb{T},\mu)} - \gamma \\
&\geq -||\sqrt{W}||^2_\infty - \gamma - \frac{1}{\gamma} \inf_{f\in D(BB^*): f\perp e^U} \frac{||B^* f||^2_{L^2(\mathbb{T},\mu)}}{||f||^2_{L^2(\mathbb{T},\mu)}} \\
&\geq -||\sqrt{W}||^2_\infty - \gamma - \frac{1}{\gamma} \inf_{g\in L^2(\mathbb{T},\mu)} \frac{||g||^2_{L^2(\mathbb{T},\mu)}}{||(\tilde{B}^*)^{-1}g||^2_{L^2(\mathbb{T},\mu)}} \\
&= -||\sqrt{W}||^2_\infty - \gamma - \frac{1}{\gamma} \frac{1}{\sup_{g\in L^2(\mathbb{T},\mu)} \frac{||(\tilde{B}^*)^{-1}g||^2_{L^2(\mathbb{T},\mu)}}{||g||^2_{L^2(\mathbb{T},\mu)}}} \\
&= -||\sqrt{W}||^2_\infty - \gamma - \frac{1}{\gamma}||(\tilde{B}^*)^{-1}||^{-2}_{L^2(\mathbb{T},\mu)}.
\end{aligned}$$

So we see that $\beta_2 > 0$ if $-||\sqrt{W}||^2_\infty - \gamma - \frac{1}{\gamma}||(\tilde{B}^*)^{-1}||^{-2}_{L^2(\mathbb{T},\mu)} > 0$ which means that $-||(\tilde{B}^*)^{-1}||^{-2}_{L^2(\mathbb{T},\mu)} - \gamma||\sqrt{W}||^2_\infty - \gamma^2 < 0$. This is a parabola and using the quadratic formula we get its roots $\alpha_\pm = -\frac{1}{2}||\sqrt{W}||^2_\infty \pm \sqrt{\frac{1}{4}||\sqrt{W}||^4_\infty - ||(\tilde{B}^*)^{-1}||^{-2}_{L^2(\mathbb{T},\mu)}}$. Since this is a parabola that goes to $-\infty$ if $\gamma \to \pm\infty$ we see then specifically if $\gamma > \alpha_+$ that then $\beta_2 > 0$. Furthermore, because we also have $\beta_1 < 0$ we see that 0 is not an eigenvalue for $S_1(\gamma)$ so we see by lemma 3.3.3 that $\gamma \notin \sigma_p(\mathscr{L}_0)$. Moreover, Proposition 4.3.2 gives us that $\sigma(\mathscr{L}_0) = \sigma_p(\mathscr{L}_0)$ so because values in the interval $(a, 0]$ are in $\sigma_p(\mathscr{L}_0)$ we also have that they are not in $\sigma(\mathscr{L}_0)$ due to Proposition 4.3.2 and thus this interval is in $\rho(\mathscr{L}_0)$. □

What we have done is that we moved the max-min theorem towards the operator BB^* and then created a bound using the first and second eigenvalue of this operator to bound the first real eigenvalue of $\mathscr{L}_0$. In a similar way, we can repeatedly use the same argument as above to create bounds on an infinite amount of eigenvalues of $S_1(\gamma)$ in terms of the eigenvalues of BB^*. We see specifically that if $\frac{1}{2}||W||^2_\infty < ||(\tilde{B}^*)^{-1}||^{-1}_{L^2(\mathbb{T},\mu)}$ that there is at most 1 eigenvalue that is real since $\beta_2 > 0$ for all $\gamma \in (-\infty, 0]$. By repeatedly using the same argument we can create similar statements about the maximal amount of eigenvalues that can be real depending on the eigenvalues of BB^* but this is not relevant for finding the spectral gap so we omit this.

The operator BB^* is associated with the overdamped Langevin diffusion operator [PP17]. Specifically it is the operator associated to the Kolmogorov forward equation of the stochastic differential equation

$$dX_t = \nabla U(X_t)dt + \sqrt{2}dW_t.$$

This equation describes a particle that receives forces over time as described by ∇U while also receiving random noise from dW_t. This process has as its stationary distribution the function e^U which is the inverse of e^{-U}. Note that due to $\Omega = \mathbb{T}$ that $-U$ is also a valid potential.

Another interesting observation is that as long as $\frac{1}{2}||W||_\infty < ||(\tilde{B}^*)^{-1}||^{-1}_{L^2(\mathbb{T},\mu)}$ then we have that the bound on the spectral gap can not decrease as we scale the refreshment rate. This can also be seen by the fact that we do not have any real eigenvalues in this case and that the bound that we have on the imaginary eigenvalues is determined by an infimum involving the refreshment rate, so increasing this the refreshment rate then increases the bound on the spectral gap. However, this is not any assurance that the spectral gap actually increases, only that the lower bound that we calculated increases.

For many examples we seem to have $\frac{\int_\mathbb{T} W(x)e^{U(x)}\,dx}{\int_\mathbb{T} e^{U(x)}\,dx} \geq \frac{1}{2}||W||_\infty$ which would often reduce the first case of the above expression. However, in general this does not hold. This can be seen by for example taking for every $\epsilon > 0$, $\lambda_0 = \frac{1}{\epsilon}\cdot 1_{|x|\leq\epsilon}$ with $U = 0$. We then have

$$\frac{\int_\mathbb{T} W(x)e^{U(x)}\,dx}{\int_\mathbb{T} e^{U(x)}\,dx} = \frac{1}{2\pi}\int_\mathbb{T} 2\lambda_0(x)\,dx = \frac{4}{2\pi}.$$

Whereas for the other side of the expression would have

$$\frac{1}{2}||W||_\infty = \frac{1}{\epsilon}.$$

Moreover, by taking $\lambda_0 = 0$ we also have that this equality does not hold in general as we can see with $U(x) = \cos(x)^{20}$. Then we have that $\frac{\int_{\mathbb{T}} W(x)e^{U(x)}\,dx}{\int_{\mathbb{T}} e^{U(x)}\,dx} \approx 0.868$ and $\frac{1}{2}||W||_\infty \approx 1.374$. In general this term can therefore not be ignored without any extra justification.

We can then use Proposition 3.4.6 to prove a new bound on the spectral gap.

Theorem 3.4.7. *Assume $U \in W^{1,\infty}(\mathbb{T})$ and $\lambda_0 \in L^\infty(\mathbb{T})$ and define*

$$\overline{\kappa} := \begin{cases} \min\{\frac{1}{2}||W||_\infty - \sqrt{\frac{1}{2}||W||_\infty^2 - ||(\tilde{B}^*)^{-1}||^{-2}_{L^2(\mathbb{T},\mu)}}, \frac{1}{2}\operatorname{ess\,inf}_{x\in\mathbb{T}} W(x)\} & \text{if } \frac{1}{2}||W||_\infty^2 \geq ||(\tilde{B}^*)^{-1}||^{-1}_{L^2(\mathbb{T},\mu)} \\ \frac{1}{2}\operatorname{ess\,inf}_{x\in\mathbb{T}} W(x) & \text{otherwise.} \end{cases}$$

Then we have $\overline{\kappa} \leq \kappa$, where κ is the spectral gap of $\mathscr{L}_0$. Specifically this means that if $\operatorname{ess\,inf}_{x\in\mathbb{T}} W(x) > 0$ then $\kappa > 0$.

Proof. The bound on the imaginary eigenvalues follow from lemma 3.4.1 and the bounds for the real eigenvalues follow from Proposition 3.4.6. Moreover, note that we can ignore the term $\frac{\int_{\mathbb{T}} W(x)e^{U(x)}\,dx}{\int_{\mathbb{T}} e^{U(x)}\,dx}$ since $\int_{\mathbb{T}} W(x)e^{U(x)}\,dx \geq \operatorname{ess\,inf}_{y\in\mathbb{T}} W(y)\int_{\mathbb{T}} e^{U(x)}\,dx$ and thus we have

$$\frac{\int_{\mathbb{T}} W(x)e^{U(x)}\,dx}{\int_{\mathbb{T}} e^{U(x)}\,dx} \geq \operatorname{ess\,inf}_{y\in\mathbb{T}} W(y)\frac{\int_{\mathbb{T}} e^{U(x)}\,dx}{\int_{\mathbb{T}} e^{U(x)}\,dx} = \operatorname{ess\,inf}_{y\in\mathbb{T}} W(y) \geq \frac{1}{2}\operatorname{ess\,inf}_{y\in\mathbb{T}} W(y).$$

Moreover, since $\sigma_p(\mathscr{L}_0) = \sigma(\mathscr{L}_0)$ by Proposition 4.3.2 we have that the spectrum of $\mathscr{L}_0$ is bounded and thus also the spectral gap. □

We first show an example of this by using $U' = 0$ and $\forall x \in \mathbb{T}$ we define $\lambda_0(x) = \lambda \in \mathbb{R}_+$ such that $\forall x \in \mathbb{T}$ we have $W(x) = 2\lambda$. This gives us that $\mu \propto 1$. We then have that this gives us that the approximation on the spectral of Theorem 3.4.7 gap is given by $\kappa(\lambda) = \lambda - 1_{\lambda>1}\sqrt{\lambda^2-1}\}$. Where we calculated $||(\tilde{B}^*)^{-1}||^{-1}_{L^2(\mathbb{T},\mu)}$ by noting that this operator becomes the Volterra operator for which we know that the exact norm on this domain is equal to 1 [Thi16].

Now we will calculate all of the eigenvalues of this potential and refreshment rate and we will see how well this approximation is. We have to solve $S(\gamma)f = 0$, with $f \in W^{2,2}(\mathbb{T})$ we see that this implies $-\partial_x^2 f - 2\gamma\lambda - \gamma^2 = 0$ which has the functions $f(x) = e^{\pm i\sqrt{\gamma(2\lambda+\gamma)}x}$ as its solution. In order to make these functions continuous we must have $f(-\pi) = f(\pi)$ which implies that $\sqrt{\gamma(2\lambda+\gamma)} = n$ and $\gamma(2\lambda+\gamma) = n^2$ with $n \in \mathbb{Z}$. Solving this gives us $\gamma = -\lambda \pm \sqrt{\lambda^2 - n^2}$. Because we ignore $\gamma \in 0$ this gives us that the spectral gap is given by $\kappa(\lambda) = \lambda - 1_{\lambda>1}\sqrt{\lambda^2-1}\}$ which is the exact same formula that the appromation of Theorem 3.4.7 gave us.

Because $\lambda - \sqrt{\lambda^2-1} = \frac{1}{\lambda+\sqrt{\lambda^2-1}}$ is a decreasing function for $\lambda > 1$ we see that the optimal spectral gap is reached when $\lambda = 1$, with $\kappa(1) = 1$. Moreover, we see that the spectral gap is reached exactly at the moment that an eigenvalue becomes real. This is behaviour that we observed for many other potentials using the approximation of Chapter 4 that as soon an eigenvalue becomes real we have very quickly that the spectral gap becomes worse.

In general for other potentials calculating $||(\tilde{B}^*)^{-1}||^{-1}_{L^2(\mathbb{T},\mu)}$ can be a hard thing to do exactly, however due to lemma 3.2.8 we can create an approximation $||(\tilde{B}^*)^{-1}||_{L^2(\mathbb{T},\mu)} \leq ||(\tilde{B}^*)^{-1}||_{HS}$. The difference between these two terms is that $||(\tilde{B}^*)^{-1}||_{L^2(\mathbb{T},\mu)} = \frac{1}{\lambda_1(BB^*)}$ whereas for the other norm we have $||(\tilde{B}^*)^{-1}||^2_{HS} = \sum_{i\neq 0}\frac{1}{\lambda_i(BB^*)^2}$ where $\lambda_i(BB^*)$ refers to the ith eigenvalue of BB^* counted with multiplicity, here we skip λ_0 as this eigenvalue has been removed as we examine the domain $D(BB^*) \setminus \{e^U\}$. The downside by making such an approximation is that the period for which we can assure that there are no real eigenvalues for $\mathscr{L}_0$ becomes smaller as we scale λ_0 and the bound on the spectral gap becomes worse at a faster rate.

This bound on the eigenvalues that are real is very similar to Theorem 2.11.13 of [Tre08]. However, there are several differences in the result and the assumptions. To interpret the result we should flip the operator $-\mathscr{L}_0$ such that the spectrum is increasing $\lambda_1 \leq \lambda_2 \leq \ldots$. Then the results of Theorem 2.11.13 assume that the operator is diagonally dominant through some assumptions on the domains of the entries of the block operator. So effectively they assume that the operator A and D are more influential than the B and C operators are. Moreover, they assume that the quadratic range (and thus also the spectrum) is real-valued and that there are numbers between $\overline{\Lambda_+(\mathscr{A})}$ and $\overline{\Lambda_+(\mathscr{A})}$. Moreover, if an actual lower bound on the eigenvalues and thus the spectral gap is to be found there is an additional assumption that B is bounded strongly by the minimal spectral value of A.

Under the assumption that U is a $C^\infty(\mathbb{T})$ Morse function (meaning a function with critical points that are non-degenerate and have distinct values) and that $\inf_{x\in\mathbb{T}} \lambda_0(x) > 0$ it has been proven that the eigenvalues of the generator become arbitrarily small at an exponential rate for a large enough scaling of the form $h \to hW$ (Theorem 1 [GN20]). This would seem to indicate that the max-min principle of the Schur complement $S_1(\gamma)$ gives an equality for the spectral gap for small enough h. However, we do not have this in general and the assumption on the refreshment rate seems to be required. The main example is $U(x) = |x|$, as we will show in Chapter 4 for which the spectral gap is always caused by an eigenvalue that is in $\mathbb{C} \setminus \mathbb{R}$ no matter how much this potential is scaled. Therefore it does not have the property that the eigenvalues become real at a large enough scaling. Therefore only under the assumption that the scaling is large enough and the infimum of the refreshment rate is non-zero can we could say that the spectral gap of $\mathscr{L}_0$ can be exactly determined by the first few eigenvalues of $S_1(\gamma)$.

3.5. The Schur complement and the semiclassical Witten Laplacian

It turns out that the Schur complement has a connection with the semiclassical Witten Laplacian [Mic19], which is an operator which is of interest to theoretical physicists. In this section, we will make the link between these two operators. The semiclassical Witten Laplacian has the following form

$$\Delta_{h,U} = -h^2\partial_x^2 + |U'|^2 - hU''.$$

The Schur complement can be written in the following manner for an eigenvector $f \in D(BB^*)$ corresponding to the eigenvalue γ we have

$$\begin{aligned}\gamma S(\gamma)f &= -BB^*f + \gamma(A-\gamma)f \\ &= \partial_x(\partial_x(f) - U'f) - \gamma(W+\gamma)f \\ &= \partial_x^2 f - U'\partial_x f - U''f - \gamma(W+\gamma)f = 0.\end{aligned}$$

This can be rewritten into the following form

$$\partial_x^2 f + p(x)\partial_x f + q(x)f = 0.$$

Where we have the following functions

$$p(x) = -U' \text{ and } q(x) = -U'' - \gamma(W+\gamma).$$

We then make the substitution $h = e^{-\frac{U}{2}}f$, which will give us:

$$-\partial_x^2 h + V(x)h = 0, \tag{3.14}$$

where we have that V is defined as

$$V(x) = -q(x) + \frac{1}{2}\partial_x p(x) + \frac{1}{4}p^2(x).$$

And this can be explicitly written as

$$V(x) = \gamma(W+\gamma) + \frac{|U'|^2}{4} + \frac{U''}{2}.$$

Here we see terms that are similar to the Witten Laplacian with the exception of $\gamma(W+\gamma)$. We then compare the 2nd order differential equation with the semi-classical Witten Laplacian which is defined for $a, r \in \mathbb{R}$ by

$$\Delta_{h,aU} = -h^2\partial_x^2 + |a|^2|U'|^2 - ahU''.$$

In order to get this definition to be similar to our 2nd order differential equation of formula (3.14) we must have

$$\frac{|a|^2}{h^2}|U'|^2 - \frac{a}{h}U'' = \frac{1}{4}|U'|^2 + \frac{1}{2}U''.$$

Which leads to the condition $\frac{a}{r} = -\frac{1}{2}$. So we have for all $r \in \mathbb{R}$ that $S_1(\gamma)$ can be equivalently written as

$$\Delta_{r,-\frac{r}{2}U} + \gamma(W+\gamma) = -\partial_x^2 + (\frac{1}{4}|U'|^2 + \frac{1}{2}U'') - \gamma(W+\gamma) = 0.$$

Note that due to the domain being a torus we have that $-\frac{rU}{2}$ defines a valid potential similar to how we did this for BB^*. In fact, the relationship is stronger to that case as all that we have done is here calculate the following equality

$$E^*\gamma S_1(\gamma)E = -\Delta_{r,-\frac{r}{2}U} - \gamma(W+\gamma).$$

where $E : L^2(\mathbb{T},\mu) \to L^2(\mathbb{T})$ is the unitary transformation given by $Ef = e^{-\frac{U}{2}}$. This unitary operator allows us to better analyse the asymptotic behaviour as the norm/inner-product $L^2(\mathbb{T})$ are not changed when scaling U. Moreover, we have from this the identity

$$EBB^*E^* = \Delta_{r,-\frac{r}{2}U}.$$

As we have discussed the operator BB^* is the Kolmogorov forward equation for the overdamped Langevin diffusion and we see once again that the BB^* plays an important role in the analysis of the spectrum of $\mathscr{L}_0$. Because the operator BB^* is a more 'known' operator than $\mathscr{L}_0$ this allows us to make more statements about $\mathscr{L}_0$ given BB^* such as in Theorem 3.4.7. A similar connection has been made in [GN20] where the connection was made via the anti-symmetric part of $\mathscr{L}_0$. Examining lemma 3.2.5 shows that this connection was to be expected as the anti-symmetric part of $\mathscr{L}_0$ is directly related to BB^* through the Schur complement defined for this part. It seems as if the analysis towards the Witten-Laplacian is, therefore, crucial for the further approximation/understanding of the eigenvalues of the Schur complement.

4

Spectral analysis of the Zig-Zag process generator with $\lambda_0 = 0$

When analysing the bound defined for $f \in X$ by $||T(t)f|| \leq Me^{-\omega t}||f||$ as defined in 2.3.3, different norms and spaces can be used for the underlying Banach Space. This change in the topology/Banach space could have an effect on the value of κ. Moreover, changing the underlying Banach space could change the positions of the eigenvalues and the spectral gap. Previously this analysis has been done for $X = L^2(\mathbb{R}, \nu)$ [BV21], however in practice it can also be useful to know the convergence for continuous functions with a supremum norm as this is usually the set of functions which are used for the analysis of the Zig-Zag process using Stochastic approaches such as examining a Feller process (more about this type of approach can be found in, for example, [BD17]).

Moreover, we would like to know how the spectrum of $\mathscr{L}$ translates to the semigroup that it might generate. In this section we will prove such a spectral theorem for the generator of the Zig-Zag process with the Banach space $L^2(\mathbb{T}, \nu)$ as well as $C(\mathbb{T} \times \{+1, -1\})$. We will show a similar analysis to [BV21] towards the case where $\lambda_0 = 0$ and the potential is unimodal and we will find similar results. Furthermore, we will see that the spectral mapping theorem gives some complications for the torus and we will show how we could solve these complications. We will first introduce the generator on the continuous function space.

4.1. Spectral analysis on $C(\mathbb{T} \times \{+1, -1\})$

The generator $\mathscr{L}$ can be defined similarly for the space of continuous functions $C(\mathbb{T} \times \{+1, -1\})$ as we were able to do it for $L^2(\mathbb{T}, \nu)$. This is the space of continuous functions in the sense that for $f(x,\theta) \in C(\mathbb{T} \times \{+1, -1\})$ we have that $f^+ = f(x, 1)$ and $f^- = f(x, -1)$ are continuous. This vector space then becomes a Banach space if we associate the norm $||f||_{C(\mathbb{T}\times\{+1,-1\})} = \sup_{x\in\mathbb{T},\theta\in\{+1,-1\}} |f(x,\theta)|$ to it. Then as for the domain for this new generator we take the continuously differentiable functions $C^1(\mathbb{T} \times \{+1, -1\})$. That is the functions $f \in C(\mathbb{T} \times \{+1, -1\})$ such that $f^+, f^-(x)$ have derivatives $\partial_x f^+(x), \partial_x f^-(x)$ which are fully defined and continuous on $\mathbb{T}$. In order to make it more clear to which operator we are referring we denote $\mathscr{L}_{C^1} : C^1(\mathbb{T} \times \{+1, -1\}) \subset C(\mathbb{T} \times \{+1, -1\}) \to C(\mathbb{T} \times \{+1, -1\})$ when we are referring to the generator on $C^1(\mathbb{T} \times \{+1, -1\})$. So for $f \in C^1(\mathbb{T} \times \{+1, -1\})$ we define

$$(\mathscr{L}_{C^1})f(x,\theta) := \theta\partial_x f + \lambda(x,\theta)(Ff - f).$$

Moreover, we denote the space $C_0^1(\mathbb{T} \times \{+1, -1\})$ to be the mean-zero continuous functions, that is

$$C_0(\mathbb{T} \times \{+1, -1\}) := \{f \in C(\mathbb{T} \times \{+1, -1\} : \int_{\mathbb{T}} f(x)d\nu = 0\}$$

and in a similar way we can define $C_0^1(\mathbb{T}\times\{+1,-1\})$ on which we will define the generator $\mathscr{L}_{C_0^1} : C_0^1(\mathbb{T}\times\{+1,-1\}) \subset C_0(\mathbb{T} \times \{+1, -1\}) \to C_0(\mathbb{T} \times \{+1, -1\})$ with the same formula for $f \in C_0^1(\mathbb{T} \times \{+1, -1\})$ by

$$(\mathscr{L}_{C_0^1})f(x,\theta) := \theta\partial_x f + \lambda(x,\theta)(Ff - f).$$

We have that the codomain is the same Banach space as the Banach space the operator is defined on similar to how we did this for $\mathscr{L}_0$. For this, we will need that $\mathscr{L}_{C^1} f \in C(\mathbb{T} \times \{+1, -1\})$ for all $f \in C(\mathbb{T} \times \{+1, -1\})$. This gives us that $\theta\partial_x f + \lambda(x,\theta)(Ff - f) \in C(\mathbb{T}\times\{+1,-1\})$. Which implies that $\partial_x(f^+ - f^-) + U'(f^+ - f^-) \in C(\mathbb{T}\times\{+1,-1\})$. This causes us to make the assumption that $U' \in C(\mathbb{T})$ or $U \in C^1(\mathbb{T})$ which is assumption (A3). Similarly, this will give us that $\lambda_0 \in C(\mathbb{T})$ which is assumption (A4).

Under these assumption we have that the mapping $\mathscr{L}_{C^1} : C^1(\mathbb{T}\times\{+1,-1\}) \subset C(\mathbb{T}\times\{+1,-1\}) \to C(\mathbb{T}\times\{+1,-1\})$ is well defined. Moreover, it can be seen by a similar argument for $\mathscr{L}_0$ that $\mathscr{L}_{C^1_0}$ is also well-defined. Furthermore, we have that these operators are closed.

Lemma 4.1.1. *Assume (A3) and (A4) we then have that* $(\mathscr{L}_{C^1}, C^1(\mathbb{T}\times\{+1,-1\}))$, $(\mathscr{L}_{C^1_0}, C^1(\mathbb{T}\times\{+1,-1\}))$ *are closed operators.*

Proof. Take a sequence $f_n \in C^1(\mathbb{T}\times\{+1,-1\})$ such that $f_n \xrightarrow{C(\mathbb{T}\times\{+1,-1\})} f$ and $\mathscr{L}_{C^1} f_n \xrightarrow{C(\mathbb{T}\times\{+1,-1\})} g$. We then have that for every $\epsilon > 0$ that there is an N such that for $n, m \geq N$

$$||\partial_x f_n - \partial_x f_m||_{C(\mathbb{T}\times\{+1,-1\})} \leq ||\mathscr{L}_{C^1} f_n - \mathscr{L}_{C^1} f_m||_{C(\mathbb{T}\times\{+1,-1\})} + 2||\lambda(x,\theta)||_{C(\mathbb{T},\{+1,-1\})}||f_n - f_m||_{C(\mathbb{T}\times\{+1,-1\})} < \epsilon.$$

Where we used that $(f_n)_{n\geq 1}$ and $(\mathscr{L}_{C^1})_{n\geq 1}$ are Cauchy. So we have that $(\partial_x f_n)_{n\geq 1}$ is Cauchy and we have for the limit that $f \in C^1(\mathbb{T}\times\{+1,-1\}$ holds. It then follows that

$$||\mathscr{L}_{C^1} f_n - \mathscr{L}_{C^1} f||_{C(\mathbb{T}\times\{+1,-1\})} \leq ||\partial_x f_n - \partial_x f||_{C(\mathbb{T}\times\{+1,-1\})} + 2||\lambda(x,\theta)||_{C(\mathbb{T}\times\{+1,-1\})}||f_n - f||_{C(\mathbb{T}\times\{+1,-1\})}.$$

Where the right-hand side goes to 0 as $n \to \infty$ so we see that $g = \mathscr{L}_{C^1} f$ by the uniqueness of the limit so we have that $\mathscr{L}_{C^1}$ is closed. Similarly, by following the exact same proof and replacing $C^1(\mathbb{T}\times\{+1,-1\})$ by $C^1_0(\mathbb{T}\times\{+1,-1\})$ we see that $\mathscr{L}_{C^1}$ is closed as well. □

For the next proposition (that is Proposition 4.1.3) we will first need to proof that $\tilde{B}$ is bijective and that C is surjective with kernel function e^U when these operators are defined on the continuous/differentiable functions as we did this in lemma 3.2.7. For this, we will denote $\tilde{B}_{C^1_0}$ to denote $\tilde{B}$ as defined in lemma 3.2.7 but then on a domain of differentiable functions, furthermore we use $C_{C^1_0}$ to denote the differentiable version of the operator C defined in formula (3.8).

Lemma 4.1.2. *Assume (A3) and (A4) then the unbounded operator* $\tilde{B}_{C^1_0} : C^1_0(\mathbb{T}\times\{+1,-1\}) \subset C_0(\mathbb{T}\times\{+1,-1\}) \to C(\mathbb{T}\times\{+1,-1\}) \setminus \{e^U\}$ *is bijective and the unbounded operator* $C_{C^1_0} : C^1(\mathbb{T}\times\{+1,-1\}) \subset C(\mathbb{T}\times\{+1,-1\}) \to C_0(\mathbb{T}\times\{+1,-1\})$ *is surjective with* $\mathrm{Ker}(C_{C^1_0}) = \{e^U\}$.

Proof. The injectivity of $\tilde{B}_{C^1_0}$ has the same argument as in lemma 3.2.7. As for the surjectivity $\tilde{B}_{C^1_0}$ take an $f \in C(\mathbb{T}\times\{+1,-1\}) \setminus \{e^U\}$. We can see that

$$\tilde{B}^{-1}_{C^1_0} f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{y} f(\xi)\, d\xi\, e^{-U(y)} dy - \int_{-\pi}^{x} f(\xi)\, d\xi$$

is in $C(\mathbb{T}\times\{+1,-1\})$ because $\int_{-\pi}^{\pi} f(\xi)\, d\xi = 0$ due to the orthogonality on e^U in the domain. Moreover, for $x \in (-\pi,\pi)$ we have clearly $\partial_x \tilde{B}^{-1}_{C^1_0} f = -f$ by the fundamental theorem of calculus. For the continuity points $x = \pi, -\pi$ we clearly have that $\lim_{h\to 0^-} \frac{(\tilde{B}^{-1}_{C^1_0} f)(x+h) - (\tilde{B}^{-1}_{C^1_0} f)(x)}{h} = -f(\pi)$. Moreover, we have that $\lim_{h\to 0^+} \frac{(\tilde{B}^{-1}_{C^1_0} f)(x+h) - (\tilde{B}^{-1}_{C^1_0} f)(x)}{h} = -\frac{1}{h}\int_{-\pi}^{-\pi+h} f(\xi) d\xi = -f(-\pi)$. And due to the continuity of f we have that $f(\pi) = f(-\pi)$ and thus we have that $(\partial_x \tilde{B}^{-1}_{C^1_0} f)(\pi) = (\partial_x \tilde{B}^{-1}_{C^1_0} f)(-\pi)$ so we have that $\tilde{B}^{-1}_{C^1_0} f \in C^1(\mathbb{T}\times\{+1,-1\})$. Moreover, we have $(\tilde{B}^{-1}_{C^1_0} f | e^U) = 0$ for the same reason as in lemma 3.2.7. So we have that $\tilde{B}^{-1}_{C^1_0} f \in C^1_0(\mathbb{T},\mu) \setminus \{e^U\}$ and thus $\tilde{B}_{C^1_0}$ is bijective.

Next, we examine the kernel of $C_{C^1_0}$. Assume that there is an $f \in C^1(\mathbb{T}\times\{+1,-1\})$ and we have $C_{C^1_0} f = U'f - \partial_x f = 0$. This implies that $\partial_x(e^{-U} f) = 0$, thus we have $f = Ke^U$ for some $K \in \mathbb{C}$. So we see that the kernel of $C_{C^1_0}$ is spanned by $\{e^U\}$. Lastly, we examine the surjectivity of $C_{C^1_0}$ we take again an element $f \in C^1_0(\mathbb{T}\times\{+1,-1\}$ we then define

$$g(x) := -e^{U(x)}\int_{-\pi}^{x} f(\xi)\, e^{-U(\xi)} d\xi.$$

this function is continuous since $\int_{-\pi}^{x} f(\pi)\, e^{-U(\xi)} d\xi = 0$ due to $(f|1) = 0$. Moreover, for every $x \in (-\pi,\pi)$ we have that $g(x)$ is differentiable and continuous as we have that $\partial_x g(x) = U'f(x) - f(x)$. Furthermore, for the boundary points $x=-\pi,\pi$ we have that $\lim_{h\to 0^-}\frac{(g(x+h)-g(x)}{h}=U'(\pi)f(\pi)-f(\pi)$. Moreover, we have that $\lim_{h\to 0^+}\frac{g(x+h)-g(x)}{h}=-\frac{h}{\int}{}^{-\pi h}_{-\pi}f(\xi)d\xi=U'(-\pi)f(-\pi)-f(-\pi)$. And since $f,U'\in C(\mathbb{T}\times\{+1,-1\}$ we see that $\partial_x g(\pi)=\partial_x g(-\pi)$ so we see that $g\in C^1(\mathbb{T}\times\{+1,-1\}$ so $C_{C_0^1}$ is surjective. □

Proposition 4.1.3. *Assume (A3) and (A4) then the generator* $\mathscr{L}_{C_0^1}: C_0^1(\mathbb{T}\times\{+1,-1\})\subset C_0(\mathbb{T}\times\{+1,-1\})\to C_0(\mathbb{T}\times\{+1,-1\})$ *is boundedly invertible if and only if* $||W||_\infty>0$. *The formula for this inverse is exactly the same as for* $\mathscr{L}$ *given in Proposition 3.2.9.*

Proof. The exact same proof of Proposition 3.2.9 can be followed by replacing $L^2(\mathbb{T},\nu),W^{1,2}(\mathbb{T},\nu)$ by $C_0(\mathbb{T}\times\{+1,-1\}),C_0^1(\mathbb{T}\times\{+1,-1\})$ and by replacing $\tilde{B},C$ by $\tilde{B}_{C_0^1},C_{C_0^1}$. □

We now want to work towards proving that $\mathscr{L}_{C^1}$ generates a strongly continuous semigroup on $C(\mathbb{T}\times\{+1,-1\}$. We do this by proving that this generator is dissipative which is more discussed in the preliminary theory with Definition 2.3.5. To prove dissipativity for the non-Hilbertian Banach space $C^1(\mathbb{T}\times\{+1,-1\})$ its duality set is less obvious than with a Hilbert space. Luckily Proposition 3.23 of [EN00] only requires us to find a single element from the duality set to prove dissipativity. Moreover, it follows that the dissipativity property holds for all of the elements of the duality set if the operator $\mathscr{L}_{C^1}$ generates a contraction semigroup which we will prove later in lemma 4.1.5.

Lemma 4.1.4. *Assume (A3) and (A4) then* $(\mathscr{L}_{C^1},C^1(\mathbb{T}\times\{+1,-1\})),(\mathscr{L}_{C_0^1},C_0^1(\mathbb{T}\times\{+1,-1\}))$ *are dissipative.*

Proof. For every $f\in C^1(\mathbb{T}\times\{+1,-1\})$ or $C_0^1(\mathbb{T}\times\{+1,-1\})$, we take (x_0,θ_0) such that f is maximal and thus $|f(x_0,\theta_0)|=||f||_\infty$. We then define the functional $f':C(\mathbb{T}\times\{+1,-1\})\to\mathbb{C}$ by $\langle g,f'\rangle:=g(x_0,\theta_0)\overline{f(x_0,\theta_0)}$. Which has the property $\langle f,f'\rangle=||f||_\infty^2$ and thus we have that f' is in the duality set of f. We then have

$$\begin{aligned}\mathrm{Re}\langle\mathscr{L}f,f'\rangle&=(\theta_0\partial_x f(x_0,\theta_0)+\lambda(x_0,\theta_0)(f(x_0,-\theta_0)-f(x_0,\theta_0)))\overline{f(x_0,\theta_0)}\\&=\mathrm{Re}(\lambda(x_0,\theta_0)(f(x_0,-\theta_0)-f(x_0,\theta_0)))\overline{f(x_0,\theta_0)}\\&=\lambda(x_0,\theta_0)(\mathrm{Re}\,f(x_0,-\theta_0)\overline{f(x_0,\theta_0)}-|f(x_0,\theta_0)|^2)\le 0.\end{aligned}$$

In the second equality we used that $\partial_x f(x_0,\theta_0)=0$ since f is maximal in this point. Then by Proposition 3.23 [EN00] we have that $(\mathscr{L}_{C^1},C^1(\mathbb{T}\times\{+1,-1\}))$ and also $(\mathscr{L}_{C_0^1},C_0^1(\mathbb{T}\times\{+1,-1\}))$ are dissipative. □

Now that we have that $(\mathscr{L},C^1(\mathbb{T}\times\{+1,-1\}))$ is dissipative we can make use of Lumer-Phillips theorem in order to prove that the operator $\mathscr{L}$ generates a contraction semigroup. To completely prove this we also make use of Proposition 4.1.3.

Lemma 4.1.5. *Assume (A3), (A4) and* $||W||_\infty>0$ *then* $(\mathscr{L}_{C^1},C^1(\mathbb{T}\times\{+1,-1\}))$ *and* $(\mathscr{L}_{C_0^1},C_0^1(\mathbb{T}\times\{+1,-1\}))$ *generate strongly continuous contraction semigroups* $P(t):C(\mathbb{T}\times\{+1,-1\})\to C(\mathbb{T}\times\{+1,-1\})$ *and* $P^0(t):C_0(\mathbb{T}\times\{+1,-1\})\to C_0(\mathbb{T}\times\{+1,-1\})$.

Proof. Since $0\in\rho(\mathscr{L}_{C_0^1})$ by lemma 3.2.9 we have by the series expansion of the resolvent $(\mathscr{L}_{C_0^1}-\lambda)^{-1}$ that $[0,\frac{1}{||\mathscr{L}^{-1}||})\subset\rho(\mathscr{L}_{C_0^1})$. Moreover, this gives us that for $\gamma\in(0,\frac{1}{||\mathscr{L}^{-1}_{C_0^1}||})$ we have that $||(\gamma-\mathscr{L}_{C_0^1})^{-1}||\le\frac{1}{\gamma}$ due to the dissipativity from lemma 4.1.4. Specifically taking $\gamma=\frac{1}{||\mathscr{L}^{-1}||}-\epsilon$ for a small enough epsilon allows us to repeat the argument about the serious expansion of the resolvent to conclude that $[0,\frac{2}{||\mathscr{L}^{-1}||}-2\epsilon)\subset\rho(\mathscr{L})$. Again repeating this argument we eventually get that $[0,\infty)\subset\rho(\mathscr{L})$ and we have that $||(\gamma-\mathscr{L}_{C_0^1})^{-1}||\le\frac{1}{\gamma}$. This allows us to use Theorem II.3.5 from [EN00] to conclude that $(\mathscr{L},C_0^1(\mathbb{T}\times\{+1,-1\}))$ generates a strongly continuous contraction semigroup $P^0(t)$ on $C_0(\mathbb{T},\{+1,-1\})$.

Now if we take $f=1$ we have that $(\gamma-\mathscr{L}_{C^1})f=\gamma f$ so we have that $(\gamma-\mathscr{L}_{C^1})$ is surjective since for any element $f\in\mathscr{L}_{C^1}$ we can write this element as $f=g_1+g_2$ with $g_1\in C_0^1(\mathbb{T}\times\{+1,-1\})$ and $g_2\in\mathrm{span}\{1\}$ then we have that there is a $h_1\in C_0^1(\mathbb{T}\times\{+1,-1\})$ such that $(\gamma-\mathscr{L}_{C^1})h_1=(\gamma-\mathscr{L}_{C_0^1})h_1=g_1$ by the surjectivity of $\gamma-\mathscr{L}_{C_0^1}$ on $C_0^1(\mathbb{T}\times\{+1,-1\})$ and as we have seen there is a $h_2\in\mathrm{span}\{1\}$ such that $(\gamma-\mathscr{L}_{C^1})g_2=h_2$ (namely $h_2=\gamma g_2$) so $\mathscr{L}_{C^1}$ is surjective and we can use Lumer-Phillips (Theorem II.3.5 from [EN00]) to conclude that $\mathscr{L}_{C^1}$ generates a strongly continuous contraction semigroup $P(t):C(\mathbb{T}\times\{+1,-1\})\to C(\mathbb{T}\times\{+1,-1\})$. □

4.2. Spectral analysis on $L^2(\mathbb{T}, \nu)$

Now we want to repeat the analysis that we just did for the operator $\mathscr{L}_{C^1}$ on the space $C(\mathbb{T} \times \{+1, -1\})$ to the operator $\mathscr{L}$ on the space $L^2(\mathbb{T}, \nu)$.

Lemma 4.2.1. *$(\mathscr{L}, W^{1,2}(\mathbb{T}, \nu))$, $(\mathscr{L}_0, W_0^{1,2}(\mathbb{T}, \nu))$, and $(\mathscr{A}, W^{1,2}(\mathbb{T}, \mu) \bigoplus W_0^{1,2}(\mathbb{T}, \mu))$ are dissipative. Furthermore, we have for $f \in D(\mathscr{L})$*

$$\mathrm{Re}\langle \mathscr{L} f, f\rangle = -\frac{1}{2}\langle W(Tf)^+, (Tf)^+\rangle_{L^2(\mathbb{T},\mu)} = -\frac{1}{2}\int_{\mathbb{T}} (|U'| + 2\lambda_0)|f^+ - f^-|^2 d\mu = -\frac{1}{2}\int_{\mathbb{T}} W|f^+ - f^-|^2 d\mu$$

and for $f \in D(\mathscr{A})$

$$\mathrm{Re}\langle \mathscr{A} f, f\rangle = -\frac{1}{2}\langle W f^+, f^+\rangle_{L^2(\mathbb{T},\mu)} = -\frac{1}{2}\int_{\mathbb{T}} (|U'| + 2\lambda_0)|f^+|^2 d\mu = -\frac{1}{2}\int_{\mathbb{T}} W|f^+|^2 d\mu.$$

Proof. For this, we make use of lemma 3.2.5. Take an $f \in D(\mathscr{L}) = W^{1,2}(\mathbb{T}, \nu)$ we then have

$$\begin{aligned}\mathrm{Re}\langle \mathscr{L} f, f\rangle_{L^2(\mathbb{T},\nu)} &= \mathrm{Re}\langle \mathscr{A} Tf, Tf\rangle_{L^2(\mathbb{T},\nu)} \\ &= \langle \mathscr{A}_{sym} Tf, Tf\rangle_{L^2(\mathbb{T},\nu)} \\ &= \frac{1}{2}\langle A(Tf)^+, (Tf)^+\rangle_{L^2(\mathbb{T},\mu)} \\ &= -\frac{1}{2}\langle W(Tf)^+, (Tf)^+\rangle_{L^2(\mathbb{T},\mu)}\end{aligned}$$

This directly shows that $\mathscr{L}$ is dissipative and also that $\mathscr{A}$ is dissipative because T is a unitary transformation, more precisely take a $g \in D(\mathscr{A})$ this gives us for all $g \in D(\mathscr{A})$ that we have $\langle \mathscr{A} g, g\rangle = \langle T\mathscr{L}T^* g, g\rangle = \langle \mathscr{L}T^* g, T^* g\rangle = -\frac{1}{2}\langle W g^+, g^+\rangle_{L^2(\mathbb{T},\mu)}$ so $\mathscr{A}$ is dissipative. □

Interestingly $\mathscr{A}$ is not dissipative on the space $C(\mathbb{T} \times \{+1, -1\})$. However, we will not need to make use of this.

Lemma 4.2.2. *We have that $(\mathscr{L}, W^{1,2}(\mathbb{T}, \nu))$ and $(\mathscr{L}_0, W_0^{1,2}(\mathbb{T}, \nu))$ generate strongly continuous contraction semigroups $P(t) : L^2(\mathbb{T}, \nu) \to L^2(\mathbb{T}, \nu)$ and $P^0(t) : L_0^2(\mathbb{T}, \nu) \to L_0^2(\mathbb{T}, \nu)$.*

Proof. Using Theorem II.3.17 from [EN00] and lemma 4.2.1 we have that $\mathscr{L}, \mathscr{L}_0$ generate strongly continuous contraction semigroup. □

4.3. Combined analysis for $C(\mathbb{T} \times \{+1, -1\})$ and $L^2(\mathbb{T}, \nu)$

Now that we have proven that there exist contraction semigroups which are generated by $\mathscr{L}$ and $\mathscr{L}_{C^1}$ we will prove that all of these generators have a compact resolvent. That is for every $\gamma \in \rho(\mathscr{L})$ we have that the operator $(\gamma - \mathscr{L})^{-1}$ is a compact operator. The proof for this can be done for all of the 4 operators at once so for this proof we will denote all of them by $\mathscr{L}$.

Proposition 4.3.1. *We have that the operators $(\mathscr{L}, D(\mathscr{L}))$ with $D(\mathscr{L}) = W^{1,2}(\mathbb{T}, \nu)$, $W_0^{1,2}(\mathbb{T}, \nu)$ have a compact resolvent. Moreover, assuming (A3) and (A4) we have that for $D(\mathscr{L}) = C^1(\mathbb{T} \times \{+1, -1\})$, and $C_0^1(\mathbb{T}, \{+1, -1\})$ the resolvent is compact as well.*

Proof. Using lemma 4.2.2 and 4.1.5 and by using Theorem II.3.15 from [EN00] we get that there is a $\gamma > 0$ such that $(\gamma - \mathscr{L})^{-1}$ exists and is bounded. Because of the relationship $\mathscr{L} = T^{-1}\mathscr{A}T$, where T is a bounded bijective operator we have that $(\gamma - \mathscr{A})^{-1}$ also exists and is bounded and they relate through $(\gamma - \mathscr{L})^{-1} = T^{-1}(\gamma - \mathscr{A})T$. Denote by X the set of elements of $D(\mathscr{A})$ and by $||\cdot||_X$ the norm of the respective Banach space that the operator is defined on. Furthermore, define by $||\cdot||_{D(\mathscr{A})}$ the Sobolev norm $||\cdot||_{W^{1,2}(\mathbb{T},\nu)}$ if $D(\mathscr{A}) = W^{1,2}(\mathbb{T}, \nu)$ or $D(\mathscr{A}) = W_0^{1,2}(\mathbb{T}, \nu)$ and the $C^1(\mathbb{T} \times \{+1, -1\})$ norm $||\cdot||_{C^1(\mathbb{T}\times\{+1,-1\})}$ if $D(\mathscr{A}) = C^1(\mathbb{T} \times \{+1, -1\})$ or if $D(\mathscr{A}) = C_0^1(\mathbb{T} \times \{+1, -1\})$. We will prove that the operator

$$\tilde{R}(\gamma, \mathscr{A}) : (X, ||\cdot||_X) \to (D(\mathscr{A}), ||\cdot||_{\mathscr{G}(\mathscr{A})})$$

defined the same as the resolvent operator such that $\forall f \in X$, $\tilde{R}(\mathscr{A}, \gamma) f = R(\gamma, \mathscr{A}) f$ but on a codomain with a different norm, is continuous. Moreover, we will prove that the inclusion

$$i_{Graph} : (D(\mathscr{A}), ||\cdot||_{\mathscr{G}(\mathscr{A})}) \to (D(\mathscr{A}), ||\cdot||_{D(\mathscr{A})})$$

is continuous. We can then make use of the compact operator

$$i_C : (D(\mathscr{A}), ||\cdot||_{D(\mathscr{A})}) \to (D(\mathscr{A}), ||\cdot||_X)$$

(Compact because of the Sobolev embedding for $D(\mathscr{A}) = W^{1,2}(\mathbb{T},\nu)$ or $W_0^{1,2}(\mathbb{T},\nu)$ and Arzelà–Ascoli Theorem for $D(\mathscr{A}) = C^1(\mathbb{T}\times\{+1,-1\})$ or $C_0^1(\mathbb{T}\times\{+1,-1\})$) to then prove that the composition which is equal to the resolvent map

$$R(\gamma,\mathscr{A}) = i_C \circ i_{Graph} \circ \tilde{R}(\mathscr{A},\gamma) : (X, ||\cdot||_X) \to (D(\mathscr{A}), ||\cdot||_X)$$

becomes a compact operator.

We first examine the boundedness of $\tilde{R}(\mathscr{A},\gamma)$. We have for all $f \in X$ that

$$\begin{aligned} ||\tilde{R}(\mathscr{A},\gamma)f||_{\mathscr{G}(\mathscr{A})} &= ||(\gamma-\mathscr{A})^{-1}f||_{\mathscr{G}(\mathscr{A})} \\ &= ||(\gamma-\mathscr{A})^{-1}f||_X + ||\mathscr{A}(\gamma-\mathscr{A})^{-1}f||_X \\ &= ||(\gamma-\mathscr{A})^{-1}f||_X + ||(\gamma-\mathscr{A}-\gamma)(\gamma-\mathscr{A})^{-1}f||_X \\ &\leq (\gamma+1)||(\gamma-\mathscr{A})^{-1}||_X||f||_X + ||(\gamma-\mathscr{A})(\gamma-\mathscr{A})^{-1}f||_X \\ &= K||f||_X \end{aligned}$$

where $K = (\gamma+1)||(\gamma-\mathscr{A})^{-1}||_X + 1$. So we see that $\tilde{R}(\mathscr{A},\gamma)$ is continuous.

Next, we prove that i_{Graph} is a continuous embedding (a continuous identity operator). Which means that for every $f \in D(\mathscr{A})$ we will have to bound

$$||f||_{D(\mathscr{A})} = ||f||_X + ||\partial_x f||_X$$

by

$$||f||_{\mathscr{G}(\mathscr{A})} = ||f||_X + ||\mathscr{A}f||_X$$

First, note that we can trivially bound $||f||_X$ and that $||\partial_x f||_X \leq ||(\partial_x f)^+||_{\tilde{X}} + ||(\partial_x f)^-||_{\tilde{X}}$. Where $||\cdot||_{\tilde{X}}$ is the norm of $L^2(\mathbb{T},\mu)$ for $X := L^2(\mathbb{T},\nu)$ or $L_0^2(\mathbb{T},\nu)$ and $||\cdot||_{\tilde{X}}$ is the norm of $C^1(\mathbb{T})$ for $X = C^1(\mathbb{T},\{+1,-1\})$ or $X = C_0^1(\mathbb{T},\{+1,-1\})$. We also have that $||\mathscr{A}f||_X$ is equivalent in norm to $||A(f)^+ + B(f)^-||_{\tilde{X}} + ||C(f)^+||_{\tilde{X}}$.

Due to $C = U' - \partial_x$, we have that

$$||(\partial_x f)^+||_{\tilde{X}} \leq ||C(f)^+||_{\tilde{X}} + ||U'f^+||_{\tilde{X}} \leq K(||\mathscr{A}f||_X + ||f^+||_{\tilde{X}}) = K||f||_{\mathscr{G}(\mathscr{A})}.$$

and we have that

$$||(\partial_x f)^-||_{\tilde{X}} = ||B(f)^-||_{\tilde{X}} \leq ||B(f)^- + A(f)^+||_{\tilde{X}} + ||Af^-||_{\tilde{X}} \leq \tilde{K}(||\mathscr{A}f||_X + ||f^+||_{\tilde{X}}) = \tilde{K}||f||_{\mathscr{G}(\mathscr{A})}.$$

Combining these terms gives us that there is a constant $M > 0$ such that

$$||f||_{D(A)} \leq M||f||_{\mathscr{G}(\mathscr{A})}.$$

This implies that i_{Graph} is continuous. Lastly, we have that the embedding $i_c : W^{1,2}(\mathbb{T},\nu) \to L^2(\mathbb{T},\nu)$ is compact due to the Sobolev Embedding Theorem and $i_c : C^1(\mathbb{T}) \to C(\mathbb{T})$ is compact due to Arzelà-Ascoli Theorem. Finally we can write $(\gamma-\mathscr{L})^{-1} = i_c \circ i_e \circ \tilde{R}(\mathscr{L},\gamma)$, where on the left-hand we mean the mapping as $(\gamma-\mathscr{L})^{-1} : X \to D(\mathscr{L})$ and on the right-hand the mapping as $\tilde{R}(\mathscr{L},\gamma) : X \to \mathscr{G}(\mathscr{L})$. So we see that $(\gamma-\mathscr{L})^{-1} : X \to D(\mathscr{L})$ can be written as a concatenation of bounded and compact operators and thus it is also compact. □

Proposition 4.3.2. *We have that the spectrum of* $(\mathscr{L}, D(\mathscr{L}))$*, with* $D(\mathscr{L}) = W^{1,2}(\mathbb{T},\nu)$ *or* $W_0^{1,2}(\mathbb{T},\nu)$ *consists out of isolated eigenvalues (which implies* $\sigma(\mathscr{L}) = \sigma_p(\mathscr{L})$*) with finite multiplicity (the dimension of the eigenvectors corresponding to an eigenvalue is finite), moreover there is no accumulation point of the eigenvalues (other than* ∞*).*

Moreover, if we have assumption (A3), (A4) then the spectrum of the operators with domain $C^1(\mathbb{T},\{+1,-1\})$*, or* $C_0^1(\mathbb{T},\{+1,-1\})$ *also consist only out of isolated eigenvalues and the same statements about them holds. Furthermore, under these assumptions we also have that*

$$\sigma((\mathscr{L}_0, W_0^{1,2}(\mathbb{T},\nu)) = \sigma((\mathscr{L}_{C_0^1}, C_0^1(\mathbb{T},\{+1,-1\})) = \sigma((\mathscr{L}, W^{1,2}(\mathbb{T},\nu)) \setminus \{0\} = \sigma((\mathscr{L}_{C^1}, C^1(\mathbb{T}\times\{+1,-1\})) \setminus \{0\}$$

Proof. Using Theorem 6.29 from [Kat95] we get that the spectrum consists out of isolated eigenvalues with finite multiplicity and no accumulation point other than ∞. Moreover, by using lemma 3.3.3. We see that all the eigenvectors of $(\mathscr{L}_0, W_0^{1,2}(\mathbb{T},\nu))$ are in $C_0^1(\mathbb{T},\{+1,-1\})$, since the assumption $U \in C^1(\mathbb{T})$ gives $D(BB^*) = W_0^{2,2}(\mathbb{T})$ and thus these eigenvectors are in $C_0^1(\mathbb{T},\{+1,-1\})$ and thus belong to $\sigma((\mathscr{L}_{C_0^1}, C_0^1(\mathbb{T}\times\{+1,-1\}))$ as well. The reverse inclusion is trivial as every element from $C^1(\mathbb{T},\{+1,-1\})$ is also in $W^{1,2}(\mathbb{T},\nu)$ and thus we have.

$$\sigma(\mathscr{L}_0) = \sigma(\mathscr{L}_{C_0^1}).$$

Lastly, since we have that $(\mathscr{L}, W^{1,2}(\mathbb{T},\nu)$ is an extension of $(\mathscr{L}, W_0^{1,2}(\mathbb{T},\nu))$. So we have $\sigma_p((\mathscr{L}, W_0^{1,2}(\mathbb{T},\nu))) \subset \sigma_p((\mathscr{L}, W^{1,2}(\mathbb{T},\nu))$. Moreover, if we have an $\gamma \in \sigma_p((\mathscr{L}, W^{1,2}(\mathbb{T},\nu))$ with eigenvector $f \in W^{1,2}(\mathbb{T},\nu))$, then we have $\lambda(f,1) = (\mathscr{L},1) = 0$ so we get that $f \in W_0^{1,2}(\mathbb{T},\nu)$, so we have $\sigma_p((\mathscr{L}, W^{1,2}(\mathbb{T},\nu)))\setminus\{0\} \subset \sigma_p((\mathscr{L}, W^{1,2}(\mathbb{T},\nu))$. Since 0 is not an eigenvalue for $(\mathscr{L}, W_0^{1,2}(\mathbb{T},\nu))$ we get that

$$(\mathscr{L}, W_0^{1,2}(\mathbb{T},\nu)) = (\mathscr{L}, W^{1,2}(\mathbb{T},\nu))\setminus\{0\}.$$

The exact same argument gives us that

$$(\mathscr{L}, C_0^1(\mathbb{T}\times\{+1,-1\})) = (\mathscr{L}, C^1(\mathbb{T}\times\{+1,-1\}))\setminus\{0\}.$$

□

To get a full description of the point spectrum we would need to solve the differential equation $\mathscr{L}f = \gamma f$ explicitly which turns out to be difficult to do as we have to solve a complicated differential system explicitly. We will now give an explicit description of the eigenvalues and eigenvectors of the generator under the assumption that we have no refreshment rate, so $\lambda_0 = 0$ as in assumption (A6) and that the potential $U(x)$ has at most one maximum, and consequentially one minimum as in assumption (A5). For these specific assumptions, it becomes doable to write out the eigenvectors and eigenvalues in a more elegant manner. To describe these eigenvectors we define the following functions

$$\phi(\gamma,z,x) := \int_z^x U'(\xi)e^{-2\gamma(\xi-z)-(U(\xi)-U(z))}\,d\xi. \tag{4.1}$$

Moreover, we define the corresponding set of roots

$$Z(\gamma) = (e^{-2\gamma\pi-U(\pi)+U(x_0)}-1)^2 - \phi(\gamma,x_0,\pi)\phi(-\gamma,x_0,-\pi) = 0. \tag{4.2}$$

We then define the set of roots of this function

$$\Sigma := \{\gamma \in \mathbb{C} : Z(\gamma) = 0\}$$

These definitions are then used in the following theorem:

Theorem 4.3.3. *Suppose assumptions (A5) and (A6) are satisfied, we then have that for the unbounded operator* $(\mathscr{L}, W^{1,2}(\mathbb{T},\nu))$ *that* $\sigma(\mathscr{L}) = \sigma_p(\mathscr{L}) = \Sigma$. *Furthermore, we have for* $\gamma \in \sigma_p(\mathscr{L})$ *that the corresponding eigenvector* $f \in D(\mathscr{L})$ *is given by*

$$f^+(x) = \begin{cases} c^+e^{\gamma(x-x_0)} & \text{if } -\pi \le x \le x_0, \\ e^{\gamma(x-x_0)+U(x)-U(x_0)}(c^+ - c^-\phi(\gamma,x_0,x)) & \text{if } x_0 \le x \le \pi, \end{cases}$$

$$f^-(x) = \begin{cases} e^{-\gamma(x-x_0)+U(x)-U(x_0)}(c^- - c^+\phi(-\gamma,x_0,x)) & \text{if } -\pi \le x \le x_0, \\ c^-e^{-\gamma(x-x_0)} & \text{if } x_0 \le x \le \pi. \end{cases}$$

Moreover, for the resolvent $(\mathscr{L}-\gamma)^{-1}h = f$ *with* $h \in L^2(\mathbb{T},\mu)$ *we have*

$$f^-(x) = e^{-\gamma(x-x_0)+U(x)-U(x_0)}(c^- + \int_{x_0}^x e^{\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^-(\xi) - U'(\xi)e^{\gamma(\xi-x_0)}(c^+ - \int_{x_0}^{\xi} e^{-\gamma(\eta-x_0)}h^+(\eta)\,d\eta))\,d\xi), -\pi \le x \le x_0,$$

and

$$f^+(x) = e^{\gamma(x-x_0)+U(x)-U(x_0)}(c^+ - \int_{x_0}^x e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^+(\xi) + U'(\xi)e^{-\gamma(\xi-x_0)}(c^- + \int_{x_0}^{\xi} e^{\gamma(\eta-x_0)}h^-(\eta)\,d\eta))\,d\xi), x_0 \le x \le \pi.$$

Then the coefficients c^+, c^- are given by

$$\begin{bmatrix} c^+ \\ c^- \end{bmatrix} = \frac{1}{Z(\gamma)} A(\gamma)K(\gamma) \begin{bmatrix} h^+ \\ h^- \end{bmatrix}$$

Where we have that $Z(\gamma)$ is given by formula (4.2)*, and we have that*

$$A(\gamma) := \begin{bmatrix} e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} & e^{U(\pi)-U(x_0)}\phi(\gamma, x_0, \pi) \\ e^{U(\pi)-U(x_0)}\phi(-\gamma, x_0, -\pi) & e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} \end{bmatrix}$$

and

$$K(\gamma)h := \begin{bmatrix} -e^{-2\gamma\pi}\int_{x_0}^{-\pi} e^{-\gamma(\xi-x_0)} h^+(\xi)\,d\xi + e^{U(\pi)-U(x_0)}\int_{x_0}^{\pi} e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^+(\xi) + U'(\xi)e^{-\gamma(\xi-x_0)}\int_{x_0}^{\xi} e^{\gamma(\eta-x_0)} h^-(\eta)\,d\eta)\,d\xi \\ e^{-2\gamma\pi}\int_{x_0}^{\pi} e^{\gamma(\xi-x_0)} h^-(\xi)\,d\xi + e^{U(\pi)-U(x_0)}\int_{x_0}^{-\pi} e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^-(\xi) - U'(\xi)e^{-\gamma(\xi-x_0)}\int_{x_0}^{\xi} e^{\gamma(\eta-x_0)} h^+(\eta)\,d\eta)\,d\xi \end{bmatrix}.$$

Moreover, if we additionally assume (A3) then we have that all of the above also holds for $(\mathscr{L}_{C^1}, C^1(\mathbb{T}\times\{+1,-1\}))$. *With the resolvent equation having $h \in C(\mathbb{T}\times\{+1,-1\})$ in its definition.*

Proof. Using Proposition 4.3.2 we see that $\sigma(\mathscr{L}) = \sigma_p(\mathscr{L})$ so we will examine the eigenvalue problem $\mathscr{L}f = \gamma f$ with $\gamma \in \sigma_p(\mathscr{L})$ and $\gamma \in D(\mathscr{L})$, which can be written as the following system

$$\begin{aligned} \partial_x f^+ + \lambda(x,1)(f^- - f^+) &= \gamma f^+, \\ -\partial_x f^- + \lambda(x,-1)(f^+ - f^-) &= \gamma f^-. \end{aligned}$$

Since f is a periodic function we will define it on the interval $[-\pi, x_0]$ and $[x_0, \pi]$ which will give us a definition for f on all of $\mathbb{T}$. Since U has one maximum at $x = \pi$ and a minimum at $x = x_0$, we get that $U'(x) \geq 0$ and thus $\lambda(x,-1) = 0$ for $x_0 \leq x \leq \pi$ this gives us $f^- = c^- e^{-\gamma(x-x_0)}$ for this interval. Likewise if $U'(x) \leq 0$ then we have $\lambda(x,1) = 0$ for $-\pi \leq x \leq x_0$ and this gives us $f^+ = c^+ e^{\gamma(x-x_0)}$ for $-\pi \leq x \leq x_0$. We can then use variations of parameters and require continuity at x_0 to get the following form with constants $c^+, c^- \in \mathbb{C}$.

$$f^+(x) = \begin{cases} c^+ e^{\gamma(x-x_0)} & \text{if } -\pi \leq x \leq x_0, \\ e^{\gamma(x-x_0)+U(x)-U(x_0)}(c^+ - c^-\phi(\gamma, x_0, x)) & \text{if } x_0 \leq x \leq \pi, \end{cases}$$

$$f^-(x) = \begin{cases} e^{-\gamma(x-x_0)+U(x)-U(x_0)}(c^- - c^+\phi(-\gamma, x_0, x)) & \text{if } -\pi \leq x \leq x_0, \\ c^- e^{-\gamma(x-x_0)} & \text{if } x_0 \leq x \leq \pi, \end{cases}$$

where we used ϕ to refer to the function defined in formula (4.1). We then investigate the continuity at $x = \pi, -\pi$. First, $f^+(\pi) = f^+(-\pi)$ implies

$$\begin{aligned} & e^{\gamma(\pi-x_0)+U(\pi)-U(x_0)}(c^+ - c^-\phi(\gamma, x_0, \pi)) = c^+ e^{\gamma(-\pi-x_0)} \\ \to \quad & c^+(e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi}) - c^- e^{U(\pi)-U(x_0)}\phi(\gamma, x_0, \pi) = 0, \end{aligned}$$

and $f^-(\pi) = f^-(-\pi)$ implies

$$\begin{aligned} & e^{-\gamma(-\pi-x_0)+U(\pi)-U(x_0)}(c^- - c^+\phi(-\gamma, x_0, -\pi)) = c^- e^{-\gamma(\pi-x_0)} \\ \to \quad & c^-(e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi}) - c^+ e^{U(\pi)-U(x_0)}\phi(-\gamma, x_0, -\pi) = 0, \end{aligned}$$

Combining these two linear requirements for c^+ and c^-, gives us the following system of equations

$$\begin{bmatrix} e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} & -e^{U(\pi)-U(x_0)}\phi(\gamma, x_0, \pi) \\ -e^{U(\pi)-U(x_0)}\phi(-\gamma, x_0, -\pi) & e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} \end{bmatrix} \begin{bmatrix} c^+ \\ c^- \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This system has a non-trivial solution if and only if the determinant is zero, meaning that the following needs to hold

$$Z(\gamma) = (e^{-2\gamma\pi-U(\pi)+U(x_0)} - 1)^2 - \phi(\gamma, x_0, \pi)\phi(-\gamma, x_0, -\pi) = 0.$$

Which proves that $\sigma(\mathscr{L}) \subset \Sigma$.

Next, we prove that $\Sigma \subset \sigma(\mathscr{L})$ which is the same as proving that $\mathbb{C}\setminus\sigma(\mathscr{L}) = \rho(\mathscr{L}) \subset \mathbb{C}\setminus\Sigma$ So taking a $\gamma \in \rho(\mathscr{L})$ gives us that for all $h \in L^2(\mathbb{T},\nu)$ there exists an $f \in W^{1,2}(\mathbb{T},\nu)$ (or $h \in C(\mathbb{T},\{+1,-1\})$ and $f \in C^1(\mathbb{T}\times\{+1,-1\})$ for $D(\mathscr{L}) = C^1(\mathbb{T}\times\{+1,-1\})$)) such that $(\gamma-\mathscr{L})^{-1} f = h$. We then have the following system

$$\gamma f^+ - \partial_x f^+ - \lambda(x,1)(f^- - f^+) = h^+,$$
$$\gamma f^- + \partial_x f^- - \lambda(x,-1)(f^+ - f^-) = h^-.$$

In a similar way to how we solved the system for the eigenvectors we can solve this system by examining the part of $U' \leq 0$ and $U' \geq 0$ separately. Where we will again introduce constants $c^-, c^+ \in \mathbb{C}$. This gives us the vector

$$f^+(x) = e^{\gamma(x-x_0)}(c^+ - \int_{x_0}^{x} e^{-\gamma(\xi-x_0)} h^+(\xi)\,d\xi),\ -\pi \leq x \leq x_0,$$

and

$$f^-(x) = e^{-\gamma(x-x_0)}(c^- + \int_{x_0}^{x} e^{\gamma(\xi-x_0)} h^-(\xi)\,d\xi),\ x_0 \leq x \leq \pi,$$

and by variations of constants together with requiring the continuity at $x = x_0$ we get

$$f^-(x) = e^{-\gamma(x-x_0)+U(x)-U(x_0)}(c^- + \int_{x_0}^{x} e^{\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^-(\xi) - U'(\xi)e^{\gamma(\xi-x_0)}(c^+ - \int_{x_0}^{\xi} e^{-\gamma(\eta-x_0)} h^+(\eta)\,d\eta))\,d\xi), -\pi \leq x \leq x_0,$$

and

$$f^+(x) = e^{\gamma(x-x_0)+U(x)-U(x_0)}(c^+ - \int_{x_0}^{x} e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^+(\xi) + U'(\xi)e^{-\gamma(\xi-x_0)}(c^- + \int_{x_0}^{\xi} e^{\gamma(\eta-x_0)} h^-(\eta)\,d\eta))\,d\xi), x_0 \leq x \leq \pi,$$

again we will require $f^+(\pi) = f^+(-\pi)$ and $f^-(\pi) = f^-(-\pi)$. This will give us the following system:

$$\begin{bmatrix} e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} & -e^{U(\pi)-U(x_0)}\phi(\gamma,x_0,\pi) \\ -e^{U(\pi)-U(x_0)}\phi(-\gamma,x_0,-\pi) & e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} \end{bmatrix} \begin{bmatrix} c^+ \\ c^- \end{bmatrix} = K(\gamma)h, \tag{4.3}$$

where

$$K(\gamma)h = \begin{bmatrix} -e^{-2\gamma\pi}\int_{x_0}^{-\pi} e^{-\gamma(\xi-x_0)} h^+(\xi)\,d\xi + e^{U(\pi)-U(x_0)}\int_{x_0}^{\pi} e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^+(\xi) + U'(\xi)e^{-\gamma(\xi-x_0)}\int_{x_0}^{\xi} e^{\gamma(\eta-x_0)} h^-(\eta)\,d\eta)\,d\xi \\ e^{-2\gamma\pi}\int_{x_0}^{\pi} e^{\gamma(\xi-x_0)} h^-(\xi)\,d\xi + e^{U(\pi)-U(x_0)}\int_{x_0}^{-\pi} e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^-(\xi) - U'(\xi)e^{-\gamma(\xi-x_0)}\int_{x_0}^{\xi} e^{\gamma(\eta-x_0)} h^+(\eta)\,d\eta)\,d\xi. \end{bmatrix}$$

Take $a, b \in \mathbb{C}$, by choosing $h^+ = -ae^{\gamma(\xi-x_0)}1_{[-\pi,x_0]}\frac{1}{x_0+\pi}e^{2\gamma\pi}$ and $h^- = be^{-\gamma(\xi-x_0)}1_{[x_0,\pi]}\frac{1}{\pi-x_0}e^{2\gamma\pi}$ we get $K(\gamma)h = (a,b)$. So we see that $K : L^2(\mathbb{T},\nu) \to \mathbb{C}^2$ is surjective. So if we want our resolvent to be defined on all of $L^2(\mathbb{T},\nu)$ (or $C(\mathbb{T}\times\{+1,-1\})$) we will need that the matrix on the left-hand side of formula (4.3) is non-singular in order to span all of $\mathbb{C}^2$ as well. This means that $Z(\gamma) \neq 0$ so that means $\gamma \notin \Sigma$. Furthermore, it can be seen that $(\mathscr{L}-\gamma)^{-1}h$ is in $D(\mathscr{L})$ by explicit checking that that f and its derivative are in $L^2(\mathbb{T},\mu)$ for $D(\mathscr{L}) = W^{1,2}(\mathbb{T},\nu)$ which we will also do in lemma 4.4.2, and thus we have $\Sigma \subset \sigma_p(\mathscr{L})$, which we can combine with the first part of this proof to conclude that $\sigma_p(\mathscr{L}) = \Sigma$.

Lastly, we have thus have that matrix on the left-hand side of formula (4.3) can be inverted with determinant determinant $Z(\gamma)$ so we get the formulas for the coefficients

$$\begin{bmatrix} c^+ \\ c^- \end{bmatrix} = \frac{1}{Z(\gamma)} A(\gamma)K(\gamma)\begin{bmatrix} h^+ \\ h^- \end{bmatrix},$$

where we have that

$$A(\gamma) = \begin{bmatrix} e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} & e^{U(\pi)-U(x_0)}\phi(\gamma,x_0,\pi) \\ e^{U(\pi)-U(x_0)}\phi(-\gamma,x_0,-\pi) & e^{U(\pi)-U(x_0)} - e^{-2\gamma\pi} \end{bmatrix}.$$

□

Theorem 4.3.3 allows us to calculate the eigenvalues of the spectrum using algorithms for finding roots on the spectrum. Specifically, we make use of a root-finding algorithm called qz-40 [DSZ02] which uses contour integrals to calculate winding numbers of the harmonic function $Z(\gamma)$.

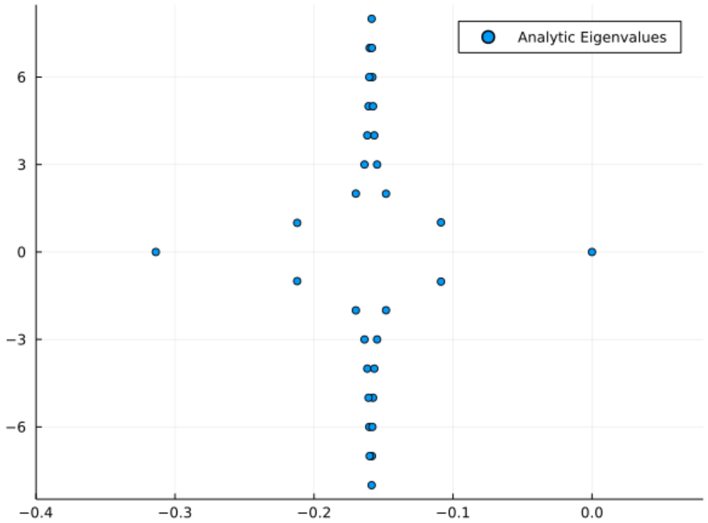

Figure 4.1: Eigenvalues of $\mathscr{L}$ on the complex plane with $U'(x) = \frac{1}{2}\sin(x)$, $\lambda_0(x) = 0$. This figure has been generated by calculating contour integrals using qz-40 [DSZ02] and finding the zeroes of $Z(\gamma)$ as defined in formula (4.2) which correspond to the eigenvalues of $\mathscr{L}$ due to theorem 4.3.3.

We can for example investigate $U(x) = |x|\sigma$ with $\sigma \in \mathbb{R}_+$. Calculating formula (4.1) gives us that the only eigenvalues are $\{\frac{-\sigma}{2} + in : n \in \mathbb{Z}\}$ and $\gamma = \sigma, 0$. This gives an indication that a piecewise discontinues/zig-zagging potential have an ever increasing spectral gap when $\sigma \to \infty$. This seems to confirm the observations in figure 3.2.

4.4. Spectral mapping theorem

Next, we would like to associate the spectrum of $\sigma(\mathscr{L})$ with the spectrum of $P(t)$ for $t > 0$. The theorems associated to this mapping are usually referred to as spectral mapping theorems. We have for the case $\Omega = \mathbb{R}$ [BV21] that the following spectral mapping theorem holds

$$\{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\} = \sigma(P(t)) \setminus \{0\}.$$

Usually zero for $P(t)$ is omitted as no eigenvalue of $\mathscr{L}$ could correspond to this eigenvalue.

The spectral mapping theorem for the torus follows a similar proof to the one of [BV21], however, it will be a bit more involved and we shall see that the full spectral mapping theorem as stated above will in general not hold for the case $\Omega = \mathbb{T}$. Theorems that quickly give the spectral mapping theorem such as Corollary 3.12 from [EN00] can in general not be applied as they require that $P(t)_{t\geq 0}$ is eventually differentiable, eventually compact, or eventually norm continuous. For example, for the case $U'(x) = 0$ and $\lambda_0(x) = \beta \in \mathbb{R}_+$ or for the case $U(x) = |x|$ we have seen that there are an infinite amount of eigenvalues in the strip $\{z \in \mathbb{C} : 2\beta \leq \operatorname{Re} z \leq 0\}$ and then using Theorem II.4.18 from [EN00] and the overview underneath Theorem II.4.29 from [EN00] we get that we do not have the eventually differentiability, eventually compactness or eventually continuous properties. This does not allow us to apply the standard spectral mapping theorem.

We are also given an extra complication as could be seen in figure 4.1. Namely, we have that the real component of all of the eigenvalues seems to converge to a vertical asymptotic line. Such a spectral line is not unique to our operator, similarly, the simple operator ∂_x with domain $W^{1,2}(\mathbb{T})$ also has a spectrum of the form $i\mathbb{Z}$. The difference is that for the generator $\mathscr{L}$ this line is in general not perfectly on $\operatorname{Re}\gamma = 0$ as the eigenvalues have a real component spread along the real axis. Furthermore, what also seems to happen is that the eigenvalues have an imaginary component that is close to an integer (usually not equal to an integer as with the case $U' = 0$). Luckily now that we have proven 4.3.3 we can prove the convergence of the eigenvalues to this asymptotic line.

Proposition 4.4.1. *Assuming (A5), (A6), and if we have a sequence $(\gamma_n)_{n\geq 1} \subseteq \sigma_p(\mathscr{L}) = \sigma(\mathscr{L})$ with the property that $|\gamma_n| \to \infty$ as $n \to \infty$, then it follows that*

$$\operatorname{Re}\gamma_n \to L := -\frac{U(\pi)-U(0)}{2\pi} = \frac{-1}{2\pi}\int_{-\pi}^{\pi}\frac{|W(x)|}{2}dx.$$

Moreover, we have that $\lim_{n\to\infty}\inf_{k\in\mathbb{Z}}|\operatorname{Im}\gamma_n - k| = 0$.

If additionally we assume that (A3) and $||\partial_x^2(e^{-U(\xi)})||_{L^1([-\pi,\pi])} < \infty$ we have the inequality

$$|(e^{-2\gamma_n\pi - U(\pi)+U(x_0)}-1)^2| \leq \frac{1}{4|\gamma_n|^2}||\partial_x^2(e^{-U(\xi)})||_{L^1([x_0,\pi])}||\partial_x^2(e^{-U(\xi)})||_{L^1([-\pi,x_0])}e^{4||W||_\infty\pi}$$

Proof. Because of Proposition 4.3.2 we do not have any accumulation points in the spectrum and since the real part of the spectrum is bounded by lemma 3.4.1 we must have that $|\operatorname{Im}\gamma_n| \to \infty$, we then have by the Riemann-Lebesgue lemma that $|\phi(\pm\gamma_n, a, b)| \to 0$. Then we can use Theorem 4.3.3 which states that $\forall \gamma_n$ it holds that

$$Z(\gamma_n) = (e^{-2\gamma_n\pi - U(\pi)+U(x_0)}-1)^2 - \phi(\gamma_n, x_0, \pi)\phi(-\gamma_n, x_0, -\pi) = 0.$$

Since $|\phi(\gamma_n, x_0, \pi)\phi(-\gamma_n, x_0, -\pi)| \to 0$ we must have that $(e^{-2\gamma_n\pi-U(\pi)+U(x_0)}-1)^2 \to 0$ which implies that $\operatorname{Re}\gamma_n \to -\frac{U(\pi)-U(0)}{2\pi}$ and $\lim_{n\to\infty}\inf_{k\in\mathbb{Z}}|\operatorname{Im}\gamma_n - k| = 0$ if the later were not to be true then we could find that there is an n large enough such that $|\arg(e^{2\pi\gamma_n})| > \epsilon$ for some $\epsilon > 0$. Which would mean that there would be some $\alpha > 0$ such that $\liminf_{n\to\infty}|e^{-2\pi\operatorname{Im}\gamma_n}-1| > \alpha$. But then we would have

$$\begin{aligned}\alpha < \liminf_{n\to\infty}|e^{-2\pi\operatorname{Im}\gamma_n}-1| &= \liminf_{n\to\infty}|e^{-2\pi\operatorname{Im}\gamma_n - 2\pi\operatorname{Re}\gamma_n - U(\pi)+U(x_0)+(-2\pi\operatorname{Re}\gamma_n - U(\pi)+U(x_0))}-1| \\ &= \liminf_{n\to\infty}|e^{(-2\pi\gamma_n - U(\pi)+U(x_0))+(-2\pi\operatorname{Re}\gamma_n-U(\pi)+U(x_0))}-1| \to 0\end{aligned}$$

which would be a contradiction so $\lim_{n\to\infty}\inf_{k\in\mathbb{Z}}|\operatorname{Im}\gamma_n - k| = 0$.

For the last statement we assume that $U \in C^1(\mathbb{T})$ we then have the following bound

$$\begin{aligned}|\phi(\pm\gamma, x_0, \pm\pi)| &= |\int_{x_0}^{\pm\pi}\partial_x(e^{-U(x)+U(x_0)})e^{\mp 2\gamma(x-x_0)}dx| \\ &= \frac{1}{2|\gamma|}|\partial_x(e^{-U(x)+U(x_0)})e^{\mp 2\gamma x}|_{x=x_0}^{x=\pm\pi} - \int_{x_0}^{\pm\pi}\partial_x^2(e^{-U(x)+U(x_0)})e^{\mp 2\gamma(x-x_0)}dx| \\ &= \frac{1}{2|\gamma|}|\int_{x_0}^{\pm\pi}\partial_x^2(e^{-U(x)+U(x_0)})e^{\mp 2\gamma(x-x_0)}dx| \\ &\leq \frac{1}{2|\gamma|}||\partial_x^2(e^{-U(\xi)})||_{L^1([x_0,\pm\pi])}e^{2||W||_\infty(\pi\pm x_0)}.\end{aligned}$$

We used that $|\partial_x(e^{-U(x)+U(x_0)})e^{\mp 2\gamma x}|_{x=x_0}^{x=\pm\pi} = 0$ because there is a maximum at x_0 and $\pm\pi$. Moreover, we used lemma 3.4.1 to bound $-\gamma$ with $||W||_\infty$. □

The inequality acquired for $U \in C^1(\mathbb{T})$ can be used to classify points on whether they are included in the spectrum. It can therefore serve as a bound on the eigenvalues in $\mathbb{C}\setminus\mathbb{R}$ for eigenvalues that are far away from the origin.

Now that we have an accurate description of the spectral line we can make the following theorem which states that the resolvent $(\gamma - \mathscr{L})^{-1}$ can be uniformly bounded for (almost) every vertical line in the complex plane. The only case where this does not hold is for the asymptotic spectral line. The proof of this is required for proving the spectral mapping theorem.

Lemma 4.4.2. *Assume (A5) and (A6), then for all $\alpha \in \mathbb{R}$ such that $\alpha \neq L = \frac{-1}{2\pi}\int_{-\pi}^{\pi}\frac{|W(x)|}{2}dx = \frac{U(\pi)-U(x_0)}{2\pi}$ we have that there is a constant $C(\alpha)$ such that*

$$\limsup_{\beta\in\mathbb{R}, |\beta|\to\infty}||(\alpha + i\beta - \mathscr{L})^{-1}|| \leq C(\alpha).$$

Moreover, if we assume (A3) and (A4) then we have that the above statements also hold for $\mathscr{L}_{C^1}$.

Proof. Take $\alpha + i\beta = \gamma \in \rho(\mathscr{L})$ and take $(h^+, h^-) \in W^{1,2}(\mathbb{T}, \nu)$ (or in $C^1(\mathbb{T}, \{+1,-1\})$ for $\mathscr{L}_{C^1}$), then define $\begin{bmatrix} f^+ \\ f^- \end{bmatrix} = (\mathscr{L} - \gamma)^{-1} \begin{bmatrix} h^+ \\ h^- \end{bmatrix}$. From Theorem 4.3.3 we get a formula for these function. We then bound these functions by looking separately at the intervals $[-\pi, x_0]$ and $[x_0, \pi]$. First, we bound the part of $f^+(x)$ defined on $x \in [-\pi, x_0]$. Using Theorem 4.3.3 we see that this the function has the following form

$$f^+(x) = e^{\gamma(x-x_0)}(c^+ - \int_{x_0}^{x} e^{-\gamma(\xi - x_0)} h^+(\xi)\, d\xi), -\pi \le x \le x_0.$$

We can then make the following point-wise estimates by using the Cauchy-Schwarz inequality

$$|e^{\gamma(x-x_0)} \int_{x_0}^{x} e^{-\gamma(\xi-x_0)} h^+(\xi)\, d\xi| \le e^{\alpha(x-x_0)} \int_{x_0}^{-\pi} e^{-\alpha(\xi-x_0)} |h^+(\xi)|\, d\xi \le e^{\alpha(x-x_0)} ||e^{-\alpha(\cdot - x_0)} e^{U(\cdot)}||_{L^2(\mathbb{T},\mu)} ||h^+||_{L^2(\mathbb{T},\mu)}$$

which then implies

$$||e^{\gamma(x-x_0)} \int_{x_0}^{x} e^{-\gamma(\xi-x_0)} h^+(\xi)\, d\xi||_{L^2(\mathbb{T},\mu)} \le ||e^{\alpha(x-x_0)}||_{L^2(\mathbb{T},\mu)} ||e^{-\alpha(\xi-x_0)} e^{U}||_{L^2(\mathbb{T},\mu)} ||h^+||_{L^2(\mathbb{T},\mu)}.$$

Moreover, we have $||e^{\gamma(x-x_0)} c^+||_{L^2(\mathbb{T},\mu)} \le ||e^{\alpha(x-x_0)}||_{L^2(\mathbb{T},\mu)} |c^+|$. So we have that there is a constant function $C_1(\alpha)$ which only depends on the real part of the eigenvalue α such that

$$||f^+(x) 1_{x \le x_0}||_{L^2(\mathbb{T},\mu)} \le C_1(\alpha)(||h^+||_{L^2(\mathbb{T},\mu)} + |c^+|)$$

Next, we bound $f^+(x)$ for $x \in [x_0, \pi]$ for these values of x we have the following representation

$$f^+(x) = e^{\gamma(x-x_0)+U(x)-U(x_0)}(c^+ - \int_{x_0}^{x} e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^+(\xi) + U'(\xi) e^{-\gamma(\xi-x_0)}(c^- + \int_{x_0}^{\xi} e^{\gamma(\eta-x_0)} h^-(\eta)\, d\eta))\, d\xi), x_0 \le x \le \pi.$$

In a similar way we bounded $f(x)1_{x \le x_0}$ we can estimate the inner-parts of the integrals using Cauchy Schwarz and no difficulties are encountered due to every function being bounded and defined on bounded domains or in $L^2(\mathbb{T}, \mu)$. First, we can bound

$$||c^- + \int_{x_0}^{\xi} e^{\gamma(\eta-x_0)} h^-(\eta)\, d\eta||_{L^2(\mathbb{T},\mu)} \le C_2(\alpha)(|c^-| + ||h^-||_{L^2(\mathbb{T},\mu)})$$

again we can then estimate

$$||e^{-\gamma(\xi-x_0)-U(\xi)+U(x_0)}(h^+(\xi) + U'(\xi) e^{-\gamma(\xi-x_0)} C_2(\alpha)(|c^-| + ||h^-||_{L^2(\mathbb{T},\mu)})||_{L^2(\mathbb{T},\mu)} \le C_3(\alpha)(|c^-| + ||h^-||_{L^2(\mathbb{T},\mu)} + ||h^+||_{L^2(\mathbb{T},\mu)})$$

lastly we can estimate

$$||e^{\gamma(x-x_0)+U(x)-U(x_0)}(c^+ - \int_{x_0}^{x} C_3(\alpha)(|c^-| + ||h^-||_{L^2(\mathbb{T},\mu)} + ||h^+||_{L^2(\mathbb{T},\mu)})\, d\xi||_{L^2(\mathbb{T},\mu)} \le C_4(\alpha)(|c^+| + |c^-| + ||h^-||_{L^2(\mathbb{T},\mu)} + ||h^+||_{L^2(\mathbb{T},\mu)})$$

Combining the estimates for $f^+(x)_{x \le x_0}$ and $f^+(x)_{x \ge x_0}$ we can bound $f^+(x)$ by a constant $C_5(\alpha)(|c^+| + |c^-| + ||h||_{L^2(\mathbb{T},\nu)})$. Similarly, we can perform the same bounds but then on the Banach space $C(\mathbb{T} \times \{+1,-1\})$ to get a similar bound (with a different constant $C_5(\alpha)$).

We have for the coefficients c^+, c^-

$$\begin{bmatrix} c^+ \\ c^- \end{bmatrix} = \frac{1}{Z(\gamma)} A(\gamma) K(\gamma) \begin{bmatrix} h^+ \\ h^- \end{bmatrix}.$$

By repeating the same argument we can see that K is bounded by the exact same reasons for why $f^+(x)$ and $f^-(x)$ are bounded so we have that there is a constant $C_6(\alpha)$ such that $||K(h)|| \le C_6(\alpha)||h||$. Furthermore, by the Riemann-Lebesgue lemma we have that $\phi(\gamma, x_0, \pi) \to 0$ and $\phi(\gamma, x_0, \pi) \to 0$ as $|\beta| \to \infty$ so we see that there is a constant $C_7(\alpha)$ such that $||A(\gamma)|| \le C_7(\alpha)$. Finally we have that for $Z(\gamma)$ which is defined as

$$Z(\gamma) = (e^{-2\gamma\pi - U(\pi) + U(x_0)} - 1)^2 - \phi(\gamma, x_0, \pi)\phi(-\gamma, x_0, -\pi) = 0.$$

that $\phi(\gamma, x_0, \pi)\phi(-\gamma, x_0, -\pi) \to 0$ due to the Riemann Lebesgue. Moreover, if we have that $\alpha \ne \frac{-1}{2\pi}\int_{-\pi}^{\pi} \frac{|W(x)|}{2}\, dx = \frac{-1}{2\pi} \frac{|U'(x)|}{2}\, dx = \frac{U(\pi)-U(x_0)}{2\pi}$ that the term $|(e^{-2\gamma\pi - U(\pi)+U(x_0)} - 1)^2|$ is bounded from below by $|(e^{-2\alpha\pi - U(\pi)+U(x_0)} - 1)^2|$ which is non-zero. So we are done as the coefficients are bounded functions of h independent of β.

□

Now we can finally get to the spectral mapping theorem where we then make use of Theorem 4.5 from [Gea78]. This theorem states that if $\gamma \in \rho(\mathscr{L}) \setminus \{0\}$ that we then have for the following set:

$$\Gamma_{\gamma,t} := \{\eta \in \mathbb{C} : e^{\eta t} = \gamma\} \tag{4.4}$$

that for all $\eta \in \Gamma_{\gamma,t}$ we have that $\eta \notin \sigma(\mathscr{L})$ and we have $\sup_{\eta\in\Gamma_{\gamma,t}} ||(\eta - \mathscr{L})^{-1}|| < \infty$. This is why Proposition 4.4.2 was required as the condition $\sup_{\eta\in\Gamma_{\gamma,t}} ||(\eta - \mathscr{L})^{-1}|| < \infty$ can now be checked for all $\gamma \in \rho(\mathscr{L}) \setminus \{0\}$.

Proposition 4.4.3. *Assume (A5) and (A6), then for all* $t > 0$ *we have*

$$\{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\} \subset \sigma(P(t)) \setminus \{0\} \subset \{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\} \cup O_L$$

where

$$O_L := \{e^{\gamma t} : \gamma \in \mathbb{C}, \operatorname{Re}\gamma = L\}$$
$$L := \frac{-1}{2\pi}\int_{-\pi}^{\pi} \frac{|W(x)|}{2}\,dx = \frac{-1}{2\pi}\int_{-\pi}^{\pi} \frac{|U'(x)|}{2}\,dx.$$

furthermore if assumption (A3) holds then the above inclusions also hold for $\mathscr{L}_{C^1}$.

Proof. By Theorem IV.3.7 from [EN00] we have the identity $\sigma_p(P(t)) \setminus \{0\} = \{e^{\gamma t} : \gamma \in \sigma_p(\mathscr{L})\}$ and since $\{e^{\gamma t} : \gamma \in \sigma_p(\mathscr{L})\} = \{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\}$ by Proposition 4.3.2 we have $\{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\} \subset \sigma(P(t)) \setminus \{0\}$.

Now assume that $\gamma \in \sigma(P(t))$. We then have by Theorem 4.5 from [Gea78] that there is a $\xi \in \Gamma_{\gamma,t}$ with $\Gamma_{\gamma,t}$ defined in equation (4.4) such that $\xi \in \sigma(\mathscr{L})$ or $\sup_{\eta\in\Gamma_{\gamma,t}} ||(\eta - \mathscr{L})^{-1}|| = \infty$. In the first case we would have $\gamma \in \{e^{\eta t} : \eta \in \sigma(\mathscr{L})\}$. In the second case we have to have that $\sup_{n\in\mathbb{Z}} ||(\mathscr{L} - (\frac{\ln|\gamma| + i(\arg(\gamma)+2\pi n)}{t})^{-1}|| = \infty$ Now if $\ln|\gamma| \neq \frac{-1}{2\pi}\int_{-\pi}^{\pi} \frac{|W(x)|}{2}\,dx$ we can use lemma 4.4.2 to see that this would cause a contradiction so only the first case can occur or we have $\operatorname{Re}\xi = a$ and thus we have $\gamma \in \{e^{\eta t} : \eta \in \sigma(\mathscr{L})\} \cup O_L$. □

The complication that occurs now in order to prove that equality of sets is surrounding the eigenvalues $\gamma \in \sigma(P(t)) \setminus \{0\}$ such that $\ln|\gamma| = L$. When there are only a finite amount of eigenvalues for the operator $\mathscr{L}$ then using the same argument as for $a \neq L$ we can show that $\{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\} = \sigma(P(t)) \setminus \{0\}$. However, when it comes to the values of O_L we do not trivially have an inclusion in $\sigma(P(t)) \setminus \{0\}$ in general if the operator $\mathscr{L}$ has an infinite amount of eigenvalues. Moreover, it seems as if in all the cases that we examined that there are an infinite amount of eigenvalues for example we had this for the case $U' = 0, \lambda_0 = c$ and $U = |x|$. Furthermore, we have that there are an infinite amount of eigenvalues for all other cases that we approximated using the scheme in Chapter 5. To prove the existence of an infinite amount of eigenvalues an approach surrounding operator stencils [Mar88] could be employed as was done by [ET18] or perhaps a relationship with Kneser's oscillation criterion as was done by [EL04].

Moreover, the eigenvalues seem to distribute themselves as $L + in$ as n becomes larger with $n \in \mathbb{Z}$. Although we can prove that if there are an infinite amount of eigenvalues that they then must be near these locations, we cannot prove that they all will be near all the points $L + in$ for a large enough n. This then gives us the following conjecture.

Conjecture 4.4.4. *There are an infinite amount of eigenvalues in* $\mathscr{L}$ *and each of these eigenvalues gets arbitrary close with its imaginary component to n for some* $n \in \mathbb{Z}$. *That is for every* $\epsilon > 0$ *there is* $N \in \mathbb{N}$ *such that for every* $n \geq N$ *there is a corresponding* $\gamma_n \in \sigma_p(\mathscr{L})$ *such that we have* $|\gamma_n - L - in| \leq \epsilon$. *With L as defined in 4.4.3.*

A consequence of this conjecture would be that we can more realistically characterize what happens to the bounds on the resolvents as seen in the following lemma.

Lemma 4.4.5. *Assume (A5), (A6), and conjecture 4.4.4 then if* $\alpha = L = \frac{-1}{2\pi}\int_{-\pi}^{\pi} \frac{|W(x)|}{2}\,dx$ *and if we have a set* $S \subset \mathbb{R}$ *then define* $M := \liminf_{\beta\in S, |\beta|\geq N, N\to\infty} \inf_{n\in\mathbb{Z}} |\beta - n|$, *with* $M = \infty$ *if S is a bounded setbounded. We then have that*

$$\limsup_{\beta\in S, |\beta|\to\infty} ||(\alpha + i\beta - \mathscr{L})^{-1}|| = \infty$$

if and only if $M = 0$. *Moreover, if we assume (A3) and (A4) then we have that the above statements also hold for* $\mathscr{L}_{C^1}$.

Proof. Take $\gamma = \alpha + \beta i \in \mathbb{C}$ with $\alpha = \frac{-1}{2\pi}\int_{-\pi}^{\pi}\frac{|W(x)|}{2}dx$ and assume that $M > 0$. First, due to the value of α we then have that $|(e^{-2\gamma\pi - U(\pi)+U(x_0)} - 1)^2| = |(e^{-i2\beta\pi} - 1)^2|$, moreover if we have that $M > 0$ then we have that $\liminf_{\beta\in S,|\beta|\to\infty}|(e^{-i2\beta\pi}-1)^2| = |(e^{-i2M\pi}-1)^2|$. So we have $\limsup_{\beta\in S,|\beta|\to\infty}|\frac{1}{Z(\gamma)}| < \frac{1}{|(e^{-i2M\pi}-1)^2|}$ so then we have again that c^+ and c^- are bounded coefficients as in the proof of lemma 4.4.2 and we again are done.

Now assume that $M = 0$. First, note that we have in general the inequality $||(\gamma - \mathscr{L})^{-1}|| \geq d(\gamma, \sigma(\mathscr{L}))$ for elements in the resolvent. Now this implies that

$$\limsup_{\beta\in S,|\beta|\to\infty} ||(\alpha + i\beta - \mathscr{L})^{-1}|| \geq \limsup_{\beta\in S,|\beta|\to\infty} \frac{1}{d(\alpha + i\beta, \sigma(\mathscr{L}))}.$$

Furthermore, we have

$$\begin{aligned}\liminf_{\beta\in S,|\beta|\to\infty} d(\alpha + i\beta, \sigma(\mathscr{L})) &= \liminf_{\beta\in S,|\beta|\to\infty}\inf_{\eta\in\sigma(\mathscr{L})} |\alpha + i\beta - \eta| \\ &\leq \liminf_{\beta\in S,|\beta|\to\infty}\inf_{\eta\in\sigma(\mathscr{L})}\inf_{n\in\mathbb{Z}} |\eta - a - in| + |i\beta - in|.\end{aligned}$$

Now due to conjecture 4.4.4 and due to lemma 4.4.1 we have that there is an N big enough such that for $|\gamma| > N$ we have that $|\gamma - a - in| \leq \epsilon$. Giving us $\inf_{\eta\in\sigma(\mathscr{L})}\inf_{n\in\mathbb{Z}}|\eta - a - in| = 0$ for all $\beta \in S$. Combining this with the assumption surrounding $M = 0$ gives us that

$$\liminf_{\beta\in S,|\beta|\to\infty}\inf_{\eta\in\sigma(\mathscr{L})}\inf_{n\in\mathbb{Z}} |\gamma - a - in| + |i\beta - in| = 0.$$

So we have

$$\limsup_{\beta\in S,|\beta|\to\infty} \frac{1}{d(\alpha + i\beta, \sigma(\mathscr{L}))} = \infty$$

which then implies

$$\limsup_{\beta\in S,|\beta|\to\infty} ||(\alpha + i(\beta + \eta) - \mathscr{L})^{-1}|| = \infty$$

which gives the result.

□

To prove the conjecture we would need a statement such as Weyl's law but then for eigenvalues on the complex plane. Weyl's law states that for eigenvalues of the Laplace-Beltrami operator $\Delta = \partial_x^2$ acting on functions that vanish at the boundary of a bounded domain $\Omega \subset \mathbb{R}$ that if we define the eigenvalue counting function $N(\lambda) = \#\{\gamma \in \sigma_p(\Delta) : |\gamma| \leq |\lambda|\}$ that we then have that $\lim_{|\lambda|\to\infty}\frac{2\pi N(\lambda)}{\lambda^{\frac{1}{2}}\,\mathrm{vol}(\Omega)} = 1$. The proof of this lemma makes use of the self-adjointness of the Laplace-Beltrami operator which would not (directly) apply to our operator. Similarly, if we now take $N(\lambda)$ to be the counting function for $\mathscr{L}$ then there should be a statement that would show that $\lim_{|\lambda|\to\infty}\frac{N(\lambda)}{2|\lambda|} \to 1$. However, this would still not be enough as it could also be that the eigenvalues do not fully satisfy the condition of conjecture 4.4.4. For example, the set of eigenvalues $\{L + i2n, n \in \mathbb{Z}\} \cup \{L + \sqrt{L^2 - (2n)^2}, n \in \mathbb{Z}\}$ would also satisfy this Weyl's law but it would not confirm conjecture 4.4.4. So ultimately quite a strong statement about the asymptotic distribution of the eigenvalues needs to be proven which was not something we were able to do.

Note that the spectral mapping theorem is somewhat different for $\Omega = \mathbb{T}$ compared to $\Omega = \mathbb{R}$ for the latter the problem that we encounter was not there as the only real-valued accumulation point of $\sigma_p(\mathscr{L})$ seems to be at $\mathrm{Re}\,\gamma \to -\infty$ which would correspond to the point 0 for $\sigma(P(t))$ (effectively we also have that $L = \infty$ for $\Omega = \mathbb{R}$). This point is already omitted by some spectral theories such as $IV.3.12$ from [EN00] whereas for the torus the accumulation point on the generator is $\mathrm{Re}\,\gamma = L$.

However, assuming the conjecture it is possible to perfectly classify which points of O_L are included in the spectrum of the semigroup $P(t)$. For example, if we take $t = 2\pi$ then we get that for the set $S = \{L + \frac{i(arg(\gamma)+2\pi k)}{t} : k \in \mathbb{Z}\}$

$$\liminf_{n\in\mathbb{Z},|n|\to\infty} |a + nb - n| = \liminf_{n\in\mathbb{Z},|n|\to\infty} |a| = a$$

$$\liminf_{k\in\mathbb{Z},|k|\geq N,N\to\infty} \inf_{n\in\mathbb{Z}} |a+kb-n| = \liminf_{k\in\mathbb{Z},|k|\geq N,N\to\infty} \inf_{n\in\mathbb{Z}} |a+k-n| \tag{4.5}$$

Which is 0 if and only if a is an integer. Which implies that $\arg(\gamma)$ is an integer multiple of 2π. Now using Theorem IV.3.7 [Gea78] we then see that only in that case these points of O_L are included (Which would correspond to the point $\gamma = 1$). The points with $\arg(\gamma) = 0$ form an accumulation point in $P(t)_{t\geq 0}$ as can be seen in figure 4.2. More generally we would have

Proposition 4.4.6. *Assume (A5), (A6), and that conjecture 4.4.4 holds and take* $t > 0$ *such that* $\frac{2\pi}{t} \in \mathbb{R}\setminus\mathbb{Q}$ *we then have*

$$\sigma(P(t))\setminus\{0\} = \{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\} \cup O_L$$

where O_L *is defined in Proposition 4.4.3.*

Otherwise if $\frac{2\pi}{t} = \frac{p}{q} \in \mathbb{Q}$ *and* p *and* q *are integers that are coprime (this can always be done for any rational number) then we have*

$$\sigma(P(t))\setminus\{0\} = \{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\} \cup O_p.$$

where we define following set

$$O_p := \{e^{Lt+i\frac{k}{p}2\pi}, k \in \{0,\ldots p-1\}\}.$$

Moreover, if we assume (A3) and (A4) then we have that the above statements also hold for $\mathscr{L}_{C^1}$.

Proof. Assume that $\frac{2\pi}{t} \in \mathbb{R}\setminus\mathbb{Q}$. Then assume that we have a $\gamma \in O_L \setminus \{e^{\eta t}, \eta \in \sigma(\mathscr{L})\} = \{e^{\eta t}, \eta \in \mathbb{C}, \mathrm{Re}\,\eta = L, \eta \notin \sigma(\mathscr{L})\}$. To prove the inclusion $\gamma \in \sigma(P(t))\setminus\{0\}$ we must thus have that

$$\sup_{k\in\mathbb{Z}} ||(\mathscr{L} - (L + \frac{i(\arg(\gamma)+2\pi k)}{t})^{-1}|| = \infty$$

by Theorem 4.5 from [Gea78] which would happen by lemma 4.4.5 if

$$\liminf_{k\in\mathbb{Z},|k|\geq N,N\to\infty} \inf_{n\in\mathbb{Z}} |a+kb-n| = 0. \tag{4.6}$$

With $b = \frac{2\pi}{t}$ and $a = \frac{\arg(\gamma)}{t}$. Now if b is irrational we can take kb to be a positive number furthermore choose $n = \lfloor kb \rfloor$ we then have that $kb - n \in [0,1]$. Now we can view this as having $kb \in \mathbb{R}\setminus\mathbb{Z}$ and we have that multiplying an irrational number with an integer gives us a dense set, that is $\{|kb| \mod 1 : k \in \mathbb{Z}\}$ is dense in $\mathbb{R}\setminus\mathbb{Z}$ so we can get arbitrarily close to a (mod 1) with $kb - n$. Furthermore, we can always choose k large enough such that this holds for a given N, so the lim inf does not pose a problem so we have that the equality always holds which implies that $\gamma \in \sigma(P(t))\setminus\{0\}$.

Now assume that $\frac{2\pi}{t} = \frac{p}{q} \in \mathbb{Q}$ with p, q coprime and assume that $\gamma \in O_p \setminus \{e^{\gamma t} : \gamma \in \sigma(\mathscr{L})\}$. We then have that $\gamma \in \sigma(P(t))\setminus\{0\}$ if and only if formula (4.6) holds with $b = \frac{p}{q}$ and $a = \frac{\arg(\gamma)p}{2\pi q}$, moreover this gives that $\arg(\gamma) = 2\pi\frac{k}{p}$ such that we have that $a = \frac{k}{q}$. So we have that $\gamma \in \sigma(P(t))\setminus\{0\}$ if and only if

$$\liminf_{z\in\mathbb{Z},|z|\geq N,N\to\infty} \inf_{n\in\mathbb{Z}} |\frac{k}{q} + z\frac{p}{q} - n| = 0.$$

We can then multiply everything by q to get

$$\liminf_{z\in\mathbb{Z},|z|\geq N,N\to\infty} \inf_{n\in\mathbb{Z}} |k + zp - nq| = 0.$$

But this equation only has a solution if k is an integer due to the Chinese remainder theorem. Moreover, due to the modulo equality given by the Chinese remainder theorem, we can get arbitrarily high values of z so the lim inf does not pose a problem so we have $\gamma \in \sigma(P(t))\setminus\{0\}$.

□

Specifically we then see that the only accumulation points for $\frac{2\pi}{t} = \frac{p}{q}$ are the points of O_p and for irrational $\frac{2\pi}{t}$ the full circle of O_L are accumulation points.

We can then plot these properties for rational $\frac{2\pi}{t}$ as seen in figure 4.2. Note that this is done numerically so we could never really calculate t correctly as an irrational number, however, if we only plot the first few eigenvalues then the accumulation points should show up correctly as can be seen below. If we take a value of t which is further away from such a rational number (only a small amount of significant numbers would be required) then we see that the accumulation points do not appear as already the first few eigenvalues with small imaginary component diverge.

Next, a final useful part of having an infinite amount of eigenvalues is that we can prove a stronger upper bound on the convergence rate of $P^0(t)$.

Proposition 4.4.7. *For all $t \geq 0$ we have $e^{-\kappa t} \leq ||P^0(t)||$, in particular this means that $e^{-||W||_\infty t} \leq ||P^0(t)||$, moreover if we have an infinite amount of eigenvalues we also have $e^{-Lt} \leq ||P^0(t)||$ with L as defined in lemma 4.4.1.*

Proof. For the first statement this is due to Proposition IV.2.2 from [EN00]. The second statement uses a different upper bound on the spectral gap given by lemma 4.4.1. □

Lastly, we now have due to 4.4.3 that we prove the hypo-coercive estimate using the spectral gap κ as the convergence rate.

Theorem 4.4.8. *Assume (A5) and (A6) then we have that there is an $M > 0$ such that for all $t > 0$ and all $f \in L^2(\mathbb{T}, \nu)$ that*

$$||P(t)f - \int_{\mathbb{T}} f\, d\nu||_{L^2(\mathbb{T},\nu)} \leq Me^{-\kappa t}||f - \int_{\mathbb{T}} f\, d\nu||_{L^2(\mathbb{T},\nu)}.$$

Moreover, if we assume (A3) then we also have for all $g \in C(\mathbb{T} \times \{+1, -1\})$ that

$$||P_{C^1}(t)g - \int_{\mathbb{T}} g\, d\nu||_{C(\mathbb{T}\times\{+1,-1\})} \leq Me^{-\kappa t}||g - \int_{\mathbb{T}} g\, d\nu||_{C(\mathbb{T}\times\{+1,-1\})}$$

Proof. Take $h = f - \int_{\mathbb{T}} f$ we then have that $h \in L^2_0(\mathbb{T}, \nu)$. Moreover, we have that we can redo 4.4.3 but then for $\mathscr{L}_0$. We then have using Proposition IV.2.2 from [EN00] that there is a constant $M > 0$ such that $||P^0(t)h|| \leq Me^{-\kappa t}$ holds for all $f \in L^2(\mathbb{T}, \nu)$. The same argument holds for $\mathscr{L}_{C^1_0}$. □

Proposition IV.2.2 from [EN00] also gives us that this is the most optimal constant ω for which an inequality of the form $||P(t)f - \int_{\mathbb{T}} f\, d\nu|| \leq Me^{-\kappa t}||f - \int_{\mathbb{T}} f\, d\nu||_{L^2(\mathbb{T},\nu)}$ can hold.

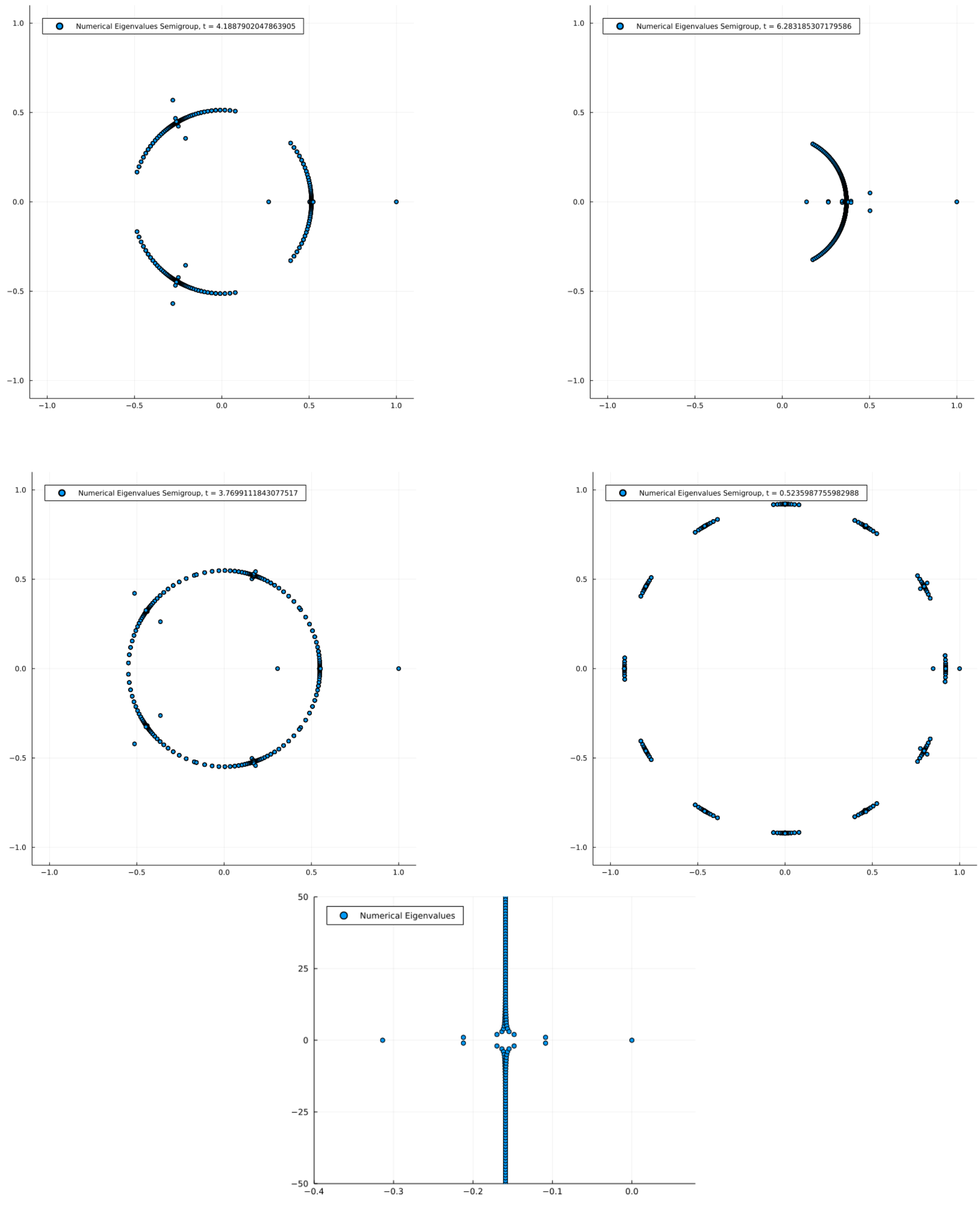

Figure 4.2: Approximation of the point spectrum of $P(t)$ with $U(x) = \frac{1}{2}(1-\cos(x))$, $\lambda_0 = 0$ by calculating the spectrum of $\mathscr{L}_N$ from Chapter 5 with $N = 6200$ and then plotting the set $\{e^{\gamma t} : \gamma \in \mathscr{L}_N, |\operatorname{Im}\gamma| < 100\}$. By taking such a large N and by only taking the eigenvalues with a small imaginary component we have that the spectrum converges. In the top-left image we have $\frac{2\pi}{t} = \frac{3}{2}$, the top-right image we have $\frac{2\pi}{t} = 1$, in the middle-left image we have $\frac{2\pi}{t} = \frac{5}{3}$ and in the middle-right image we have $\frac{2\pi}{t} = 12$. Lastly, at the bottom we have the spectrum of $\mathscr{L}_N$. We can see that the spectrum has accumulation points on the set O_p as described in Proposition 4.4.6.

5

Numerical approximation

In order to analyse the spectrum of the generator $\mathscr{L}$ for a given potential $U(x)$ and refreshment rate λ_0 there is a lot of behaviour that could happen to the eigenvalues. It is often hard to see what behaviour is common between a class of potentials. Furthermore, it is difficult to verify how sharp created bounds are or to properly test whether the propositions/theorems hold. For these reasons, it would be useful to calculate the eigenvalues of the generators. To do this we can make use of a finite difference numerical approximation $\mathscr{L}_N$ on a discretized subspace $\mathbb{T}_N$ of $\mathbb{T}$ such that $\mathbb{T}_N := (x_i)_{i=1}^N$, with $x_i := 2\pi\frac{i-1}{N} - \pi$. The main benefit of this discretized matrix $\mathscr{L}_N$ is that we can calculate its eigenvalues using standard algorithms for finite-dimensional matrices. We could then make this finite-difference approximation converge to the original generator and examine if the spectrum converges.

Numerical analysis on the convergence of eigenvalues of such a finite difference approximation $\mathscr{L}_N$ to the actual generator $\mathscr{L}$ has been done before, particularly in the field of Sturm-Liouville problem. However, to establish the connection between the eigenvalues and eigenvectors between the two operators it is often required to explicitly know the formula for the eigenvalues and eigenfunctions of both the finite difference operator as well as the full operator [PHA81], [And89] and [Con99]. Moreover, the arguments often rely on operators to be self-adjoint/normal such that the resolvent relationship for $\gamma \in \rho(\mathscr{L})$, $||(\mathscr{L}-\gamma)^{-1}||_X = d(\gamma, \sigma(\mathscr{L}))$ can be used. Unfortunately, we do not have such a relationship and therefore the proof behind this can become significantly harder.

For this reason, we prove a weaker result in that a constructed generator $\mathscr{L}_N^1$ has its semigroup $T_N(t)$ converge to the semigroup $T(t)$ as described in Proposition 5.1.2. his can be done by looking at the difference between the operator $T(t)$ and $T_N(t)$ with respect to the elements from the Banach space of $T(t)$. To make this comparison we introduce operators such that a function on the discretized space $\mathbb{T}_N$ can be compared with a function in $L^2(\mathbb{T}, \nu)$ for this we will introduce a projection $P_N : X \to X_N$ and a lift operator $E_N : X_N \to X$. Such that for large enough $N > 0$ we have that for $x \in X$

$$||E_N T_N(t) P_N x - T(t)x||_{L^2(\mathbb{T},\nu)} \to 0.$$

on bounded intervals of t. Moreover, we introduce a different discretization $\mathscr{L}_N$ whose spectrum seems to converge if N is an even number at a much faster rate than $\mathscr{L}_N^1$ however due to this operator not being dissipative we cannot establish a similar convergence result as for $\mathscr{L}_N^1$. Finally, we will talk about some generalizations of theorems that seem to hold true for general potentials. The code used to generate the numerical spectrum can be found in [Wia].

5.1. Discretization of the generator

We will now define and construct all of the spaces and operators required to make a statement about convergence. For the discrete torus $\mathbb{T}_N$ we will also define the external nodes $x_{N+1} := x_1$ and $x_0 := x_N$ and we define the difference between nodes by $\Delta x_N := \frac{2\pi}{N}$. We define $C(\mathbb{T}_N \times \{+1,-1\})$ to be the functions $f : \mathbb{T}_N \times \{+1,-1\} \to \mathbb{C}$ with associated norm $||f||_{C(\mathbb{T}_N\times\{+1,-1\})} = \sup_{x\in\mathbb{T}_N,\theta\in\{+1,-1\}} |f(x,\theta)|$. We define the projection operator P_N : $C(\mathbb{T} \times \{+1,-1\}) \to C(\mathbb{T}_N \times \{+1,-1\})$ by

$$(P_N f)(x_i, \theta) = f(x_i, \theta).$$

Furthermore, we define the lifting operator $E_N : C(\mathbb{T}_N \times \{+1,-1\}) \to C(\mathbb{T} \times \{+1,-1\})$ by

$$(E_N f)(x,\theta) := f(x_i,\theta) + \frac{f(x_{i+1},\theta) - f(x_i,\theta)}{\Delta x_N}(x - x_i) \begin{cases} x \in [x_i, x_{i+1}]\, for\ i \in \{1,2..N-1\}, \\ x \in [x_N, \pi]\, for\ i = N. \end{cases}$$

Which linearly interpolates the values of $f \in C(\mathbb{T}_N \times \{+1,-1\})$ to create a function in $C(\mathbb{T} \times \{+1,-1\})$. Finally we denote the identity operator on $C(\mathbb{T}_N \times \{+1,-1\})$ by I_N.

We then define for $N \in \mathbb{N} \setminus \{0\}$ the discretization of the generator of the Zig-Zag process $\mathscr{L}_N^1 : C(\mathbb{T}_N \times \{+1,-1\}) \to C(\mathbb{T}_N \times \{+1,-1\})$ by

$$(\mathscr{L}_N^1 f)(x,\theta) := \frac{f(x_{i+\theta},\theta) - f(x_i,\theta)}{\Delta x_N} + \lambda(x_i,\theta)(f(x_i,-\theta) - f(x_i,\theta)).$$

Lemma 5.1.1. $\mathscr{L}_N^1 : C(\mathbb{T}_N \times \{+1,-1\}) \to C(\mathbb{T}_N \times \{+1,-1\})$ *is dissipative and generates a contraction semigroup* $T_N(t) : C(\mathbb{T}_N \times \{+1,-1\}) \to C(\mathbb{T}_N \times \{+1,-1\})$.

Proof. Similar to the proof of lemma 4.1.4 we first prove the dissipativity using Proposition 3.23 from [EN00]. We take a function $f \in C(\mathbb{T}_N \times \{+1,-1\})$ and take a (x_i^*, θ^*) such that $||f||_{C(\mathbb{T}_N \times \{+1,-1\})} = |f(x_i^*,\theta^*)|$. If we then take the function $\delta_{x_i^*,\theta^*}(x_j)\overline{f(x_i^*,\theta^*)} := \overline{f(x_i^*,\theta^*)1_{x_j = x_i^*}}$ for all $x_j \in \mathbb{T}_N$ we then have that $\langle f, \delta_{x_i^*,\theta^*}\rangle = |f(x_i^*,\theta^*)|^2 = ||f||^2_{C(\mathbb{T}_N \times \{+1,-1\})}$ and thus that $\delta_{x_i^*,\theta^*}(x_i)\overline{f(x_i^*,\theta^*)}$ is in the duality set $\mathscr{J}(f)$ as defined in formula (2.1). We then have

$$\begin{aligned}
&\mathrm{Re}\langle \mathscr{L}_N^1 f, \delta_{x_i^*,\theta^*}\overline{f(x_i^*,\theta^*)}\rangle \\
&= \mathrm{Re}\,\frac{f(x_{i+\theta^*}^*,\theta^*) - f(x_i^*,\theta^*)}{\Delta x_N}\overline{f(x_i^*,\theta^*)} + \lambda(x_i^*,\theta^*)(f(x_i^*,-\theta^*) - f(x_i^*,\theta^*))\overline{f(x_i^*,\theta^*)} \\
&= \mathrm{Re}\,\frac{f(x_{i+\theta^*}^*,\theta^*)\overline{f(x_i^*,\theta^*)} - |f(x_i^*,\theta^*)|^2}{\Delta x_N} + \lambda(x_i^*,\theta^*)(f(x_i^*,-\theta^*)\overline{f(x_i^*,\theta^*)} - |f(x_i^*,\theta^*)|^2) \\
&\le 0
\end{aligned}$$

so $\mathscr{L}_N^1$ is dissipative. We then have that $\mathscr{L}_N^1 - \gamma I_N$ is injective for $\gamma > 0$ and since this operator can be written as a square matrix we also have that it surjective. Applying Lumer-Phillips Theorem (Theorem 3.15 of [EN00]) will then give us that $\mathscr{L}_N^1$ generates a contraction semigroup. □

In order to compare the convergence of the discretization of the generator to the generator itself, we could examine the convergence in terms of their resolvents and generated semigroups. In order to do this, we must be able to compare the function spaces on which $T_N(t)$ and $T(t)$ are defined. One such result is given by a formulation of Trotter-Kao (Theorem 2.1 [IK98]) which will give us the following result:

Proposition 5.1.2. *For the Zig-zag process with generator* $\mathscr{L}_N^1$ *we have for all* $f \in C(\mathbb{T} \times \{+1,-1\})$ *and for all* $\gamma > 0$ *that*

$$\lim_{N\to\infty} ||E_n(\gamma I_N - \mathscr{L}_N^1)^{-1} P_N f - (\gamma I - \mathscr{L}_{C^1})^{-1} f||_{C(\mathbb{T}\times\{+1,-1\})} = 0$$

and for $t \ge 0$

$$\lim_{N\to\infty} ||E_N T_N(t) P_N f - T(t) f||_{C(\mathbb{T}\times\{+1,-1\})} = 0 \tag{5.1}$$

uniformly on bounded intervals of t.

Proof. In order to prove this we would like to apply Proposition 3.1 from [IK98], which requires us to verify the following assumptions:

1. There are constants M_1, M_2 independent of N, such that for all $N \in \mathbb{N}$ we have $||P_N|| \le M_1$ and $||E_N|| \le M_2$.

2. We have for all $f \in C(\mathbb{T} \times \{+1,-1\})$ that $\lim_{N\to\infty} ||E_N P_N f - f||_{C(\mathbb{T}\times\{+1,-1\})} = 0$.

3. For all $N \in \mathbb{N}$ we have $P_N E_N = I_N$.

4. The semigroups $T(t), T_N(t)$ generated by $\mathscr{L}_{C^1}, \mathscr{L}_N$ share constants $M \geq 0, \omega \in \mathbb{R}$ such that $||T_N(t)||, ||T(t)|| \leq Me^{\omega t}$.

5. For every $f \in C^1(\mathbb{T} \times \{+1,-1\})$ we have that $\lim_{N\to\infty} ||E_N \mathscr{L}_N^1 P_N f - \mathscr{L}_{C_1} f||_{C(\mathbb{T}\times\{+1,-1\})} = 0$.

We will verify all of these points to conclude that Proposition 5.1.2 holds.

(Assumption 1) We have for $f \in C(\mathbb{T} \times \{+1,-1\})$ that $||P_N f||_{C(\mathbb{T}_N\times\{+1,-1\})} = \sup_{x\in\mathbb{T}_\mathbb{N},\theta\in\{+1,-1\}} |f(x,\theta)|$ $\leq \sup_{x\in\mathbb{T},\theta\in\{+1,-1\}} |f(x,\theta)| = ||f||_{C(\mathbb{T}\times\{+1,-1\})}$. So P_N is uniformly bounded with $M_1 = 1$. Because of the linear interpolation between points we also have that $||E_N f||_{C(\mathbb{T}\times\{+1,-1\})} \leq ||f||_{C(\mathbb{T}_N\times\{+1,-1\})}$. So E_N is uniformly bounded with $M_2 = 1$. Therefore the first assumption holds with $M_1 = M_2 = 1$.

(Assumption 2) By Heine-Cantor we have that the continuous functions on $\mathbb{T}$ are uniformly continuous. Therefore for an $f \in C(\mathbb{T} \times \{+1,-1\})$ we have that for all $\epsilon > 0$ there is a $\delta > 0$ and an $N \in \mathbb{N}$ such that for all $x \in \mathbb{T}$ with the closest points on the grid $x_i, x_{i+1} \in \mathbb{T}_N$, we have that $|x - x_i|, |x - x_{i+1}| \leq \Delta x_N < \delta$ implies that $|f(x_{i+1},\theta) - f(x,\theta)|, |f(x_i,\theta) - f(x,\theta)| < \epsilon$. Furthermore, we can then make the following point-wise estimate

$$
\begin{aligned}
&|E_N P_N f(x,\theta) - f(x,\theta)| \\
&= |f(x_i,\theta) - f(x,\theta) + \frac{f(x_{i+1},\theta) - f(x_i,\theta)}{\Delta x_N}(x - x_i)| \begin{cases} x \in [x_i, x_{i+1}] \text{ for } i \in \{1,2,..N-1\} \\ x \in [x_N, \pi] \text{ for } i = N \end{cases} \\
&= |(1 - \frac{(x - x_i)}{\Delta x_N})(f(x_i,\theta) - f(x,\theta)) + \frac{(x - x_i)}{\Delta x_N}(f(x_{i+1},\theta) - f(x,\theta))| \begin{cases} x \in [x_i, x_{i+1}] \text{ for } i \in \{1,2,..N-1\} \\ x \in [x_N, \pi] \text{ for } i = N \end{cases} \\
&< (1 - \frac{(x - x_i)}{\Delta x_N})\epsilon + \frac{(x - x_i)}{\Delta x_N}\epsilon \begin{cases} x \in [x_i, x_{i+1}] \text{ for } i \in \{1,2,..N-1\} \\ x \in [x_N, \pi] \text{ for } i = N \end{cases} \\
&= \epsilon.
\end{aligned}
$$

Thus we have $\lim_{N\to\infty} ||E_N P_N f - f||_{C(\mathbb{T}\times\{+1,-1\})} = 0$.

(Assumption 3) This assumption holds trivially by the definition of the projection P_N on the discrete torus and the linear interpolation E_N.

(Assumption 4) This follows from lemma 5.1.1 and lemma 4.1.5 which then implies that the common constants are $M = 1$ and $\omega = 0$.

(Assumption 5) From the definition of $\mathscr{L}_{C^1}, \mathscr{L}_N^1$ we see that for $f \in C^1(\mathbb{T} \times \{+1,-1\})$ there are two parts of the generator $\mathscr{L}_{C^1}$ to approximate. First, we approximate $\theta\partial_x f$ with $E_N \theta \partial_x^N P_N f$, where $(\partial_x^N P_N f)(x_i,\theta) := \frac{f(x_{i+\theta},\theta) - f(x_i,\theta)}{\Delta x_N}$. Since $\partial_x f$ is continuous, it is also uniformly continuous. Therefore we have for all $\epsilon > 0$ that there is a $\delta > 0$ and an $N \in \mathbb{N}$ such that for all $\xi \in \mathbb{T}$ with closest points $|\xi - x_{i-1}|, |\xi - x_i|, |\xi - x_{i+1}| \leq \Delta x_N < \delta_N$ implies that we have that $|\partial_x f(\xi,\theta) - \partial_x f(x_i,\theta)|, |\partial_x f(\xi,\theta) - \partial_x f(x_{i+1},\theta)| \leq \epsilon$. Furthermore, we have by the mean value theorem that there are $\xi_{i,1} \in [x_i, x_{i+1}]$ and $\xi_{i,-1} \in [x_{i-1}, x_i]$ such that $\partial_x^N f(x_i,\theta) = \frac{f(x_{i+\theta},\theta) - f(x_i,\theta)}{\Delta x_N} = \theta\partial f(\xi_{i,\theta},\theta)$. We then have

$$
\begin{aligned}
&|E_N \partial_x^N P_N f(x,\theta) - \theta\partial_x f(x,\theta)| \\
&= |(1 - \frac{(x - x_i)}{\Delta x_N})(\partial_x^N f(x_i,\theta) - \theta\partial_x f(x,\theta)) + \frac{(x - x_i)}{\Delta x_N}(\partial_x^N f(x_{i+1},\theta) - \theta\partial_x f(x,\theta))| \begin{cases} x \in [x_i, x_{i+1}] \text{ for } i \in \{1,2,..N-1\} \\ x \in [x_N, \pi] \text{ for } i = N \end{cases} \\
&= |(1 - \frac{(x - x_i)}{\Delta x_N})(\theta\partial_x f(\xi_{i,\theta},\theta) - \theta\partial_x f(x,\theta)) + \frac{(x - x_i)}{\Delta x_N}(\theta\partial_x f(\xi_{i,\theta},\theta) - \theta\partial_x f(x,\theta))| \begin{cases} x \in [x_i, x_{i+1}] \text{ for } i \in \{1,2,..N-1\} \\ x \in [x_N, \pi] \text{ for } i = N \end{cases} \\
&\leq |(1 - \frac{(x - x_i)}{\Delta x_N})\epsilon + \frac{(x - x_i)}{\Delta x_N}\epsilon \begin{cases} x \in [x_i, x_{i+1}] \text{ for } i \in \{1,2,..N-1\} \\ x \in [x_N, \pi] \text{ for } i = N \end{cases} \\
&= \epsilon.
\end{aligned}
$$

For the second part, we note that we have

$$\begin{aligned} E_N(\mathscr{L}_N^1 - \partial_x^N)P_N f &= E_N \lambda(x_i,\theta)(f(x_i,-\theta) - f(x_i,\theta)) \\ &= E_N P_N \lambda(x,\theta)(f(x,-\theta) - f(x,\theta)). \end{aligned}$$

Because $\lambda(x,\theta)(f(x,-\theta)-f(x,\theta))$ is continuous we can apply the proof of assumption 3 to see that $\lim_{N\to\infty} ||E_N(\mathscr{L}_N^1 - \partial_x^N)P_N f - (\mathscr{L}_{C_1} - \theta\partial_x)f||_{C(\mathbb{T}\times\{+1,-1\})} = 0$. So by combining both parts of the generator we see that $\lim_{N\to\infty} ||E_N \mathscr{L}_N^1 P_N f - \mathscr{L}_{C^1} f||_{C(\mathbb{T}\times\{+1,-1\})} \to 0$ which proves assumption 5. □

We can then calculate the eigenvalues of $\mathscr{L}_N^1$ and see if they converge to previous examples. Unfortunately, for the convergence of the eigenvalues of $\mathscr{L}_N^1$ to the eigenvalues of $\mathscr{L}_{C^1}$ it is often required to have a high value of N even for simple potentials as can be seen in figure 5.1.

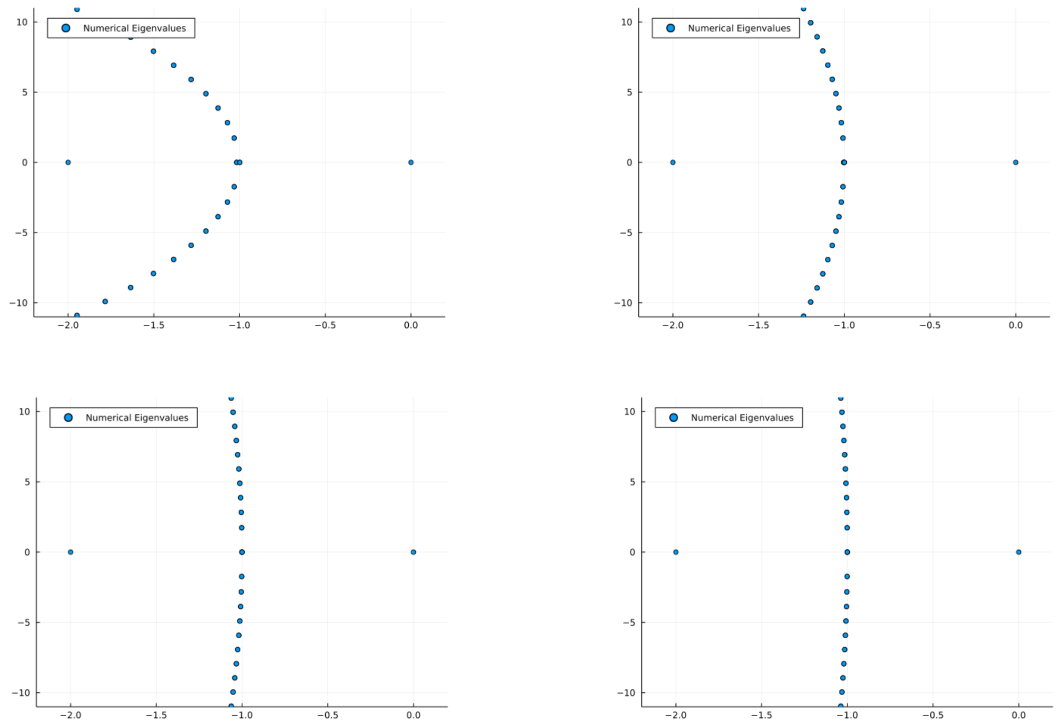

Figure 5.1: Convergence of the eigenvalues of $\mathscr{L}_N^1$ towards the eigenvalues of $\mathscr{L}_{C^1}$ for the potential $U' = 0$ and refreshment rate $\lambda_0 = 1$. The spectrum of $\mathscr{L}_{C^1}$ is equal to $\{1 \pm \sqrt{1-n^2} : n \in \mathbb{Z}\}$. In each image we calculate the eigenvalues of the matrix of $\mathscr{L}_N^1$ with in the top-left image we have $N = 400$, top-right image we have $N = 1600$, bottom-left we have $N = 6000$, and in the bottom-right we have $N = 10000$. It can be seen that the convergence does not complete even for $N = 10000$, even when only examining the first eigenvalues.

Moreover, even when calculating the spectrum for $N = 10000$ we do not get visible convergence to the spectrum especially for eigenvalues with a larger imaginary component. The main use of the numerical approximation is to get insight into the behaviour of the spectrum. However, with $\mathscr{L}_N^1$ this process can take a lot of time to converge or even become impractical to calculate on a normal computer, for example calculating the eigenvalues of the matrix corresponding to $N = 10000$ took half an hour on a modern computer. For this reason, in order to speed up the convergence, another discretized operator is introduced

$$(\mathscr{L}_N f)(x,\theta) := \theta \frac{f(x_{i+1},\theta) - f(x_{i-1},\theta)}{2\Delta x_N} + \lambda(x_i,\theta)(f(x_i,-\theta) - f(x_i,\theta)).$$

In this discretization we use the central difference method to approximate the derivative part $\theta\partial_x f$ of the generator $\mathscr{L}$. We suspect that the convergence of the spectrum of the generator $\mathscr{L}_N^1$ is slow due to the first-order approximation of the derivative as this is the part of the matrix that causes the operator $\mathscr{L}_{C^1}$ to be able to have an infinite amount of eigenvalues and causes the operator to be not bounded. However, by defining the approximation in this manner we have that for large enough N, that $\mathscr{L}_N$ is not dissipative.

Proposition 5.1.3. *For $U : \mathbb{T} \to \mathbb{R}$, $\lambda_0 : \mathbb{T} \to [0,\infty)$ such that $|U'|$ and λ_0 are bounded we have that there is a large enough $N \in \mathbb{T}$ such that $(\mathscr{L}_N f)$ is not dissipative.*

Proof. For any constant $D \in (0,1)$, there is an N such that we have $\frac{D}{2\Delta x_N} > ||\lambda(\cdot,1)||_\infty$. We then construct a function $f \in C(\mathbb{T}_N, \times\{+1,-1\})$ such that $-D = f(x_1,1) < f(x_2,1) = 1$, and $f(x_i,\theta) = 0$ otherwise. Then for $f^* \in \mathscr{J}(f)$ we have that because $C(\mathbb{T}_N \times \{+1,-1\})$ is a finite-dimensional vector space, that the dual space is also finite dimensional. Moreover, we that $f^* \in C^*(\mathbb{T}_N \times \{+1,-1\}) = L^1(\mathbb{T}_N \times \{+1,-1\})$ is of the form $\langle f, f^*\rangle = f^*(x_1,1)f(x_1,1) + f^*(x_2,1)f(x_2,1)$ and that $||f^*||_{L^1(\mathbb{T}_N\times\{+1,-1\})} = \sum_{x\in\mathbb{T}_N,\theta\in\{+1,-1\}} |f(x,\theta)|$. Furthermore, we have that $||f||_{C(\mathbb{T}_N,\times\{+1,-1\})} = |f(x_2,1)| = 1$ and thus because $f^* \in \mathscr{J}(f)$ we must have that $||f^*|| = 1$. Suppose that $|f^*(x_1,1)| > 0$. We would then have that $1 = |\langle f, f^*\rangle| \le |f^*(x_1,1)||f(x_1,1)| + |f^*(x_2,1)||f(x_2,1)| < |f^*(x_1,1)| + |f^*(x_2,1)| = 1$. Which would contradict that $\langle f, f^*\rangle = 1$. So we have that $f^*(x_1,1) = 0$. But then in order to satisfy $\langle f, f^*\rangle = 1$, we must have $f^*(x_2,1) = 1$ and $f^*(x,\theta) = 0$ otherwise. So we have that

$$\langle \mathscr{L}_N f, f^*\rangle = \frac{f(x_3,1) - f(x_1,1)}{2\Delta x_N} f(x_2,1) + \lambda(x_2,1)(f(x_2,-1) - f(x_2,1)) = \frac{D}{2\Delta x_N} - \lambda(x_2,1) > 0.$$

So $\mathscr{L}_N$ is not dissipative. □

The corresponding matrix of this generator has the following form

$$\tilde{\mathscr{L}}_N f = \begin{bmatrix} J & K \\ L & M \end{bmatrix}$$

where we have matrices $J, K, L, M \in \mathbb{R}^{NxN}$ such that J, M are almost tridiagonal matrices with the exception of 2 non-zero entries J_{1N}, J_{N1} and M_{1N}, M_{N1} respectively due to the periodicity of the functions, furthermore if we subtract the diagonal part of these matrices, then they are anti-symmetric $(J - J_{diag})^T = -(J - J_{diag}), (M - M_{diag})^T = -(M - M_{diag})$. Moreover, K, L are diagonal matrices which come from the $\lambda(x_i,\theta)(f(x_i,-\theta)$ term such that for each $i \in \{1,..,N\}$ we have that K_{ii} and L_{ii} never positive simultaneously if $\lambda_0 = 0$ due to $\lambda(x_i,\theta)$ only being positive for one θ at the time.

The discretized generator $\mathscr{L}_N$ can be used to analyse the spectrum of the original generator $\mathscr{L}$ and to verify analytical results. There is some strange behaviour of the spectrum of the discrete generator that we have not fully explained. For this reason, we cannot be certain that the spectrum properly converges to the spectrum of the generator $\mathscr{L}$. Nevertheless, the spectrum seems to be approximated correctly for some situations which we are able to calculate exactly.

The first issue with the convergence of the spectrum has to do with extra eigenvalues which appear when the size of the discretization N is odd. Specifically, it seems as if for odd N that the eigenvalues of $\mathscr{L}_N$ are doubled and interwoven with the normal amount of eigenvalues. An example where this is clearly visible is for the case $U' = 0$ and $\lambda = 1$. The spectrum then gives the following images.

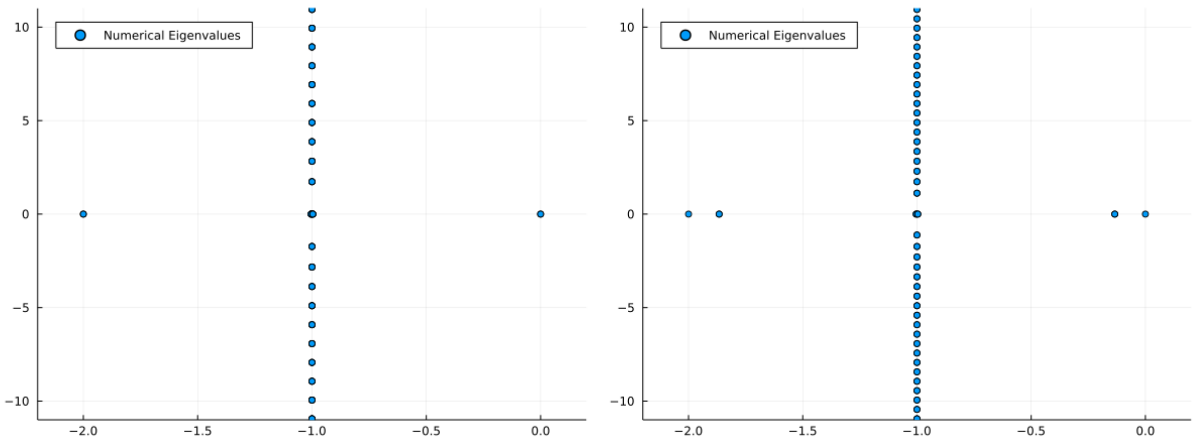

Figure 5.2: Numerical eigenvalues of $\mathscr{L}_N$ on the complex plane with $U'(x) = 0$, $\lambda_0(x) = 1$. In the left figure the corresponding generator has $N = 1024$ and in the right figure $N = 1025$. Since the spectrum is given by $-1 \pm \sqrt{1-n^2}$ with $n \in \mathbb{Z}$ we can see that for $N = 1025$ a double amount of eigenvalues are generated.

It turns out that for $N = 1025$ the eigenvalues seem to follow the formula $-1 \pm \sqrt{1 - (\frac{n}{2})^2}$ with $n \in \mathbb{Z}$. We can do the same thing but now for a unimodal potential with $U' = \sin(x)/2$ and $\lambda_0 = 0$.

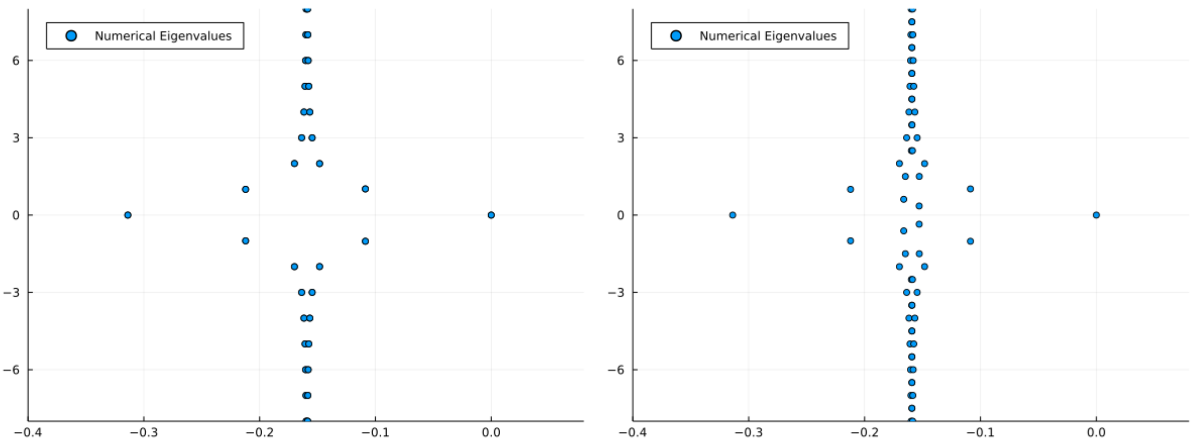

Figure 5.3: Numerical eigenvalues of $\mathscr{L}_N$ on the complex plane with $U'(x) = \frac{1}{2}\sin(x)$, $\lambda_0 = 1$. In the left figure the corresponding generator has $N = 1024$ and in the right figure $N = 1025$. Since the correct spectrum is displayed in figure 4.1 we can see that for $N = 1025$ a lot of extra incorrect eigenvalues are generated.

This behaviour seems to extend to all potentials of which we can calculate the spectrum exactly. We suspect that the reason why this behaviour happens has something to do with the lack of the dissipative property of $\mathscr{L}_N$ and that the process that is generated by $\mathscr{L}_N$ is not reducible for specific N. An approach as in [DM13] could be used to prove more about the connection between the spectrum of this generator $\mathscr{L}_N$ and the generator $\mathscr{L}$.

5.2. Observations and conjectures

In this section, we will use the discretization $\mathscr{L}_N$ to make figures of the eigenvalues of the generator and we propose some conjectures state some observations drawn from this.

5.2.1. Spectral asymptotic line

In all of the figures of the spectrum of the discrete approximation. The eigenvalues seem to have a vertical asymptotic line to which all of the eigenvalues converge. For $\lambda_0 = 0$ and unimodal U we are able to prove the location of this asymptotic line in Proposition 4.4.1. Proving this formula for general λ_0 and general U was not possible in the same way as the proof required formula (4.2) to be known which required solving the system $\mathscr{L}f = \gamma f$, which for $\lambda_0 = 0$ could be possible to do for general potentials by keeping track of all of the continuity points and solving the system in-between each critical point, however it would not provide much more insight as to why this asymptotic line exists. However, it is possible to examine how this formula should look like and to give some interpretation to it. Using the numerical approximation and trying many different examples the formula that seemed to always predict the asymptotic line is the following

Conjecture 5.2.1. *If we have a sequence $(\gamma_n)_{n\geq 1} \subseteq \sigma_p(\mathscr{L})$ with the property that $\forall M > 0$, $\exists N \in \mathbb{N}$, such that $\forall n \geq N$ we have $|\gamma_n| > M$, then it follows that* $\operatorname{Re}\gamma_n \to -\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{W(x)}{2}\,dx$.

It can be seen for unimodal U with $\lambda_0 = 0$ that the formula for the asymptotic line holds since $\frac{U(\pi)-U(0)}{2\pi} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|U'(x)|}{2}\,dx = \int_{-\pi}^{\pi} \frac{|W(x)|}{2}\,dx$. This general formula came up when puzzling around with the formula for $\gamma S_1(\gamma)f = -BB^*f - \gamma Wf - \gamma^2 f = 0$. This formula can be seen as a quadratic formula with two roots with as a critical point $-\langle Wf, f\rangle$, specifically if we look at the numerical range we see that the real component of eigenvalues in $\mathbb{C}\setminus\mathbb{R}$ is given by $-\langle Wf, f\rangle$. Clearly, W is an important function for the real component of the spectral line and since we have the formula for U unimodal and $\lambda_0 = 0$ we constructed this general formula. The formula can be seen as the average of the function $\frac{W}{2}$ on the torus. For refreshment rates equal to zero it can also be seen as the difference between all of the consecutive maximums and maxima of the potential U. The distance between the maximum and minimum showing up is not something new, similar behaviour has also been seen for the escape time between two wells such as in Theorem 1 of [Mon16].

We can then examine potentials that have a higher amount of maximums than a unimodal distribution and we can examine a potential with non-constant refreshment rates as displayed in figure 5.4.

Another interesting thing that seems to occur for potentials of the form $U(x) = -\cos(kx)$ is that depending on the value of k we have that every k-th eigenvalue on the imaginary axis is a bit more spread out than the

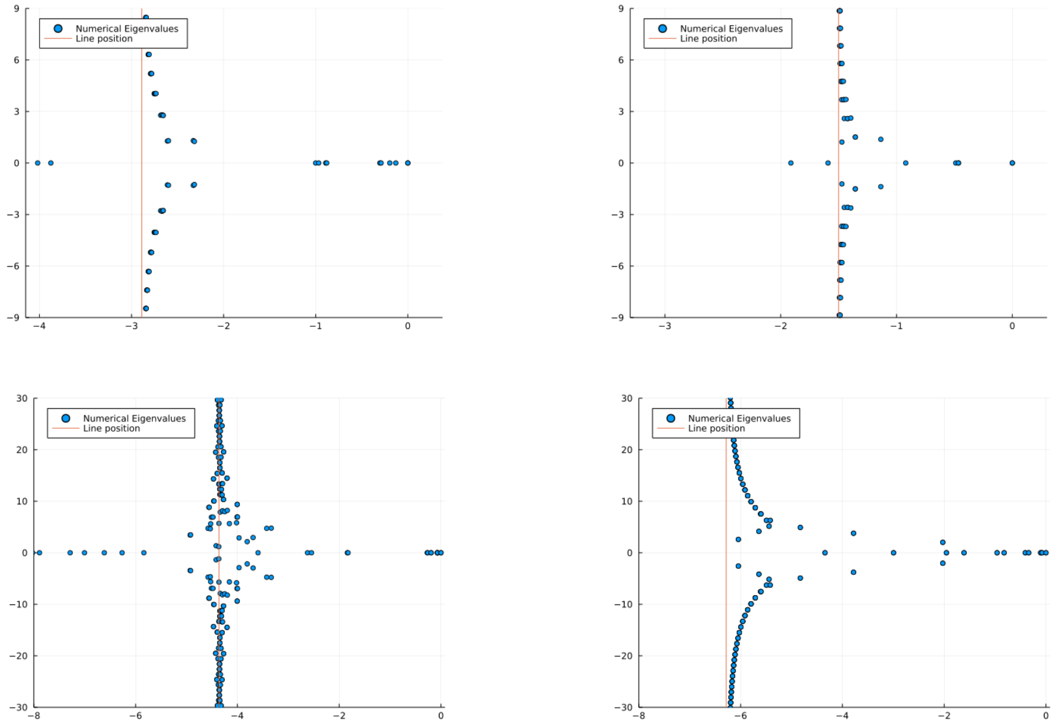

Figure 5.4: The spectrum of $\mathscr{L}_N$ for different potentials and refreshment rates together with a vertical line coming from conjecture 5.2.1. It can be seen that the spectral line (the orange line) coincides with the asymptotic line of the spectrum. In the top-left image we have $U(x) = 2(\frac{(1-\cos(x))}{2})^4$ and $\lambda(x) = 1 + |x|$. In the top-right image we have $U(x) = |x|$ and $\lambda(x) = 1 + \cos(x)$. In the bottom-left image we have $U(x) = 2(\frac{(1-\cos(5x))}{2})$ and $\lambda(x) = 1.2 + |x|$. In the bottom-right image we have $U(x) = 0$ and $\lambda(x) = 4|x|$. For all discretization we used $N = 2000$.

other eigenvalues as can be seen in the bottom-left picture of figure 5.4. This seems to indicate that these amplitudes may be related to the frequencies of the underlying potentials which is something we will also see in the following section.

5.2.2. Eigenvalues of the high-temperature limit

In this section we will examine the spectrum of high-temperature limit, meaning that we will look at the behaviour of the eigenvalues for a potential which is close to constant potential. To examine this more thoroughly we will denote the generator of Zig-Zag process in the following way, by taking $\sigma \in \mathbb{C}$

$$\mathscr{L}_\sigma f = \theta \partial_x f + \sigma((\theta U'(x))_+ + \lambda_0)(Ff - f).$$

We want to know what will happen when σ is close to zero. For example, if we were to look at the operator $\mathscr{L}_0 = \theta \partial_x f$ we would expect that our spectrum would come close to $\{in : n \in \mathbb{Z}\} \setminus \{0\}$ as this is the spectrum of the generator corresponding to the derivative operator on the torus of which we know that the eigenvectors are sine and cosine waves with integer frequencies.

We can say something about this point around $\sigma = 0$ by examining $\mathscr{L}_\sigma$ as an operator valued function such that $\mathscr{L}_\sigma = T + \sigma T'$. We could then use perturbation techniques from [Kat95] to predict the movement of the eigenvalues.

For this, we have that our operator family $(\mathscr{L}_\sigma)_{\sigma \in \mathbb{C}}$ with the domain $W^{1,2}(\mathbb{T}, \mu)$ is a holomorphic family of operators of type A as described in Definition VII.2.1 of [Kat95]. Since we have already proven that $\mathscr{L}_\sigma$ is a closed operator for $\sigma > 0$ due to lemma 3.2.4 we also have that $\mathscr{L}_\sigma$ is a closed operator for $\sigma \in \mathbb{C}$ due to the difference between the operators being a bounded operator. Moreover, we have that the domains and Hilbert spaces are equivalent to each other due to e^{-U} being a function that attains a minimum and a maximum on a bounded domain and we can switch between the Banach spaces using the unitary transformation $e^{\frac{-U}{2}} : L^2(\mathbb{T}, \mu) \to L^2(\mathbb{T}))$. Thus by Theorem VII.2.6 of [Kat95] we have that $(\mathscr{L}_\sigma)_{\sigma \in \mathbb{C}}$ is a holomorphic family of type A.

However, we run into a complication, namely, we see that the eigenvalues split/branch into two different eigenvalues when we perform a scaling of σ. This is because the operator $\mathscr{L}_{\sigma=0} = \theta \partial_x$ has two eigenfunctions for each eigenvalue (namely $f^+ = e^{cx}, f^- = 0$, and $f^+ = 0, f^- = 0$). This is due to the lack of a connection between the functions f^+, f^- for this specific operator. But when we start introducing σ both of the functions connect through the term $Ff - f$ and we see that this causes the spectrum to branch/split. This makes it difficult to apply many perturbation theorems such as from [Kat95]. Namely, it is possible to state that a weighted mean of the eigenvalues has an analytic expansion, however, due to the branching, it is no longer possible to state if the expansion has any algebraic singularities.

For example, we could have an operator of the following form

$$T(\sigma) = \begin{bmatrix} 0 & 1 \\ \sigma & 0 \end{bmatrix}.$$

We then have that the eigenvalues are given by $\lambda_\pm = \pm\sqrt{\sigma}$. So these eigenvalues are analytic and they have an algebraic singularity at $\sigma = 0$ so it is not possible to conclude that an analytic expansion occurs at $\sigma = 0$. It can also be seen that for this operator that we have for $\sigma = 0$ that the eigenvalues split similar to our situation.

Usually it is easy to prove that these singularities do not occur if the underlying operator is self-adjoint, we can then have an example such as the following

$$T(\sigma) = \begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix}.$$

We then have that the eigenvalues are given by $\lambda_\pm = \pm\sigma$ and we see that the eigenvalues do not have an algebraic singularity and a simple analytic expansion at $\sigma = 0$. However, we do not have a self-adjoint operator.

However, our operator is holomorphic as a parameter of σ and our operator is regular in the sense that it is J-self-adjoint. And although we were not able to prove if the eigenvalues are analytic or not, we can try to create an analytic expansion of the eigenvalues and see using the discretization $\mathscr{L}_N$ if such an expansion converges.

Proposition 5.2.2. *Assuming (A3), (A5), (A6), and assume that the eigenvalues of $\mathscr{L}_\sigma$ admit an analytic expansion without any algebraic singularities for the operator family $\sigma \to \mathscr{L}_\sigma$. Then for each eigenvalue $\gamma_j \in \sigma_p(\mathscr{L}_\sigma)$ there is an analytic function $\gamma_j(\sigma)$ and an $\epsilon_j > 0$ such that for $\sigma \in (0, \epsilon_j)$ we have that $\gamma_j(\sigma) \in \sigma_p(\mathscr{L}_\sigma)$. We then have the following properties*

$$\gamma_j(\sigma) \to in \quad \textit{for some} \quad n \in \mathbb{Z}, \textit{ and}$$

$$\frac{1}{h}(\gamma_j(\sigma) - in) \to \gamma_j^{(1)}.$$

For this limit $\gamma_j^{(1)}$ it holds that

$$\gamma_j^{(1)} = -\frac{U(\pi) - U(x_0)}{2\pi} \pm \frac{1}{2\pi}\sqrt{(\int_{x_0}^{\pi} U'(x)e^{-2in(x-x_0)}\,dx)(\int_{x_0}^{-\pi} U'(x)e^{2in(x-x_0)}\,dx)},$$

Furthermore, if U is symmetric and $x_0 = 0$, we get

$$\frac{2}{h^2}(\gamma_j(\sigma) - h\gamma_j^{(1)} - in) \to \gamma_j^{(2)}$$

where we have

$$\gamma_j^{(2)} = \frac{1}{2\pi}(2\gamma_j^{(1)}\pi + U(\pi) - U(0))^2 \pm \frac{1}{\pi}\int_0^{\pi}((U(x) - U(0)) + 2\gamma_j^{(1)}x)U'(x)e^{-2inx}x\,dx.$$

Proof. Since we have an analytic expansion without singularities for our eigenvalues we can write them as an analytic expansion such that for a small enough ϵ this expansion converges. We then have $\gamma_j(\sigma) = \gamma_j^{(0)} + \sigma\gamma_j^{(1)} + \frac{\sigma^2}{2}\gamma_j^{(2)} + o(\sigma^2)$. Where $\frac{o(\sigma^2)}{\sigma^2} \to 0$ as $\sigma \to 0$ Using Thm 4.3.3 we can see that these functions satisfy:

$$(e^{-2\gamma_j(\sigma)\pi - \sigma(U(\pi) - U(x_0))} - 1)^2 = \phi(\gamma_j(\sigma), x_0, \pi)\phi(-\gamma_j(\sigma), x_0, -\pi). \tag{5.2}$$

By taking $\sigma = 0$ we see that

$$(e^{-2\gamma_j^{(0)}\pi} - 1)^2 = \phi(0, x_0, \pi)\phi(0, x_0, -\pi)$$

Now $\phi(0,x_0,\pi)=\phi(0,x_0,-\pi)=0$ when $\sigma=0$. So we get $(e^{-2\gamma_j^0\pi}-1)^2=0$ Which implies that $\gamma_j^0=in$ for an $n\in\mathbb{Z}$
Next, we divide both sides of (5.2) by σ^2 and we take $\sigma\to 0$. We then get that the left-hand becomes equal to

$$\lim_{\sigma\to 0}\frac{(e^{-2(in+\sigma\gamma_j^{(1)}+o(h))\pi-\sigma(U(\pi)-U(x_0))}-1)^2}{\sigma^2}=\lim_{\sigma\to 0}(\frac{(e^{\sigma(-2\gamma_j^{(1)}\pi-U(\pi)+U(x_0))}-1)}{\sigma})^2=(2\gamma_j^{(1)}\pi+U(\pi)-U(x_0))^2$$

The right-hand will become equal to

$$\begin{aligned}
&\lim_{\sigma\to 0}\frac{1}{\sigma^2}(\int_{x_0}^{\pi}\sigma U'(x)e^{-\sigma(U(x)-U(x_0))-2(in+\gamma_j^{(1)}\sigma+o(\sigma))(x-x_0)}\,dx)(\int_{x_0}^{-\pi}\sigma U'(x)e^{-\sigma(U(x)-U(x_0))+2(in+\sigma\gamma_j^{(1)}+o(\sigma))(x-x_0)}\,dx)\\
&=\lim_{\sigma\to 0}(\int_{x_0}^{\pi}U'(x)e^{-\sigma(U(x)-U(x_0))-2(in+\gamma_j^{(1)}+o(\sigma^2))(x-x_0)}\,dx)(\int_{x_0}^{-\pi}U'(x)e^{-\sigma(U(x)-U(x_0))+2(in+\sigma\gamma_j^{(1)}+o(\sigma))(x-x_0)}\,dx)\\
&=(\int_{x_0}^{\pi}U'(x)e^{-2in(x-x_0)}\,dx)(\int_{x_0}^{-\pi}U'(x)e^{2in(x-x_0)}\,dx).
\end{aligned}$$

So we have:

$$(2\gamma_j^{(1)}\pi+U(\pi)-U(x_0))^2=(\int_{x_0}^{\pi}U'(x)e^{-2in(x-x_0)}\,dx)(\int_{x_0}^{-\pi}U'(x)e^{2in(x-x_0)}\,dx)$$

Solving for $\gamma_j^{(1)}$ will give us:

$$\gamma_j^{(1)}=-\frac{U(\pi)-U(x_0)}{2\pi}\pm\frac{1}{2\pi}\sqrt{(\int_{x_0}^{\pi}U'(x)e^{-2in(x-x_0)}\,dx)(\int_{x_0}^{-\pi}U'(x)e^{2in(x-x_0)}\,dx)},$$

Finally we will examine the value of $\gamma_j^{(2)}$. For simplicity of the calculations we will assume that U is symmetric around $x=0$. We then have that equation (5.2) becomes

$$e^{-2\gamma_j(\sigma)\pi-\sigma(U(\pi)-U(0))}-1=\pm\phi(\gamma_j(\sigma),0,\pi). \tag{5.3}$$

We then take the derivative in σ of the left-hand twice which becomes

$$\begin{aligned}
\partial_\sigma^2(e^{-2\gamma_j(\sigma)\pi-\sigma(U(\pi)-U(0))}-1)&=\partial_\sigma(e^{-2\gamma_j(\sigma)\pi-\sigma(U(\pi)-U(0))}(-2\gamma_j^{(1)}(\sigma)\pi-(U(\pi)-U(0)))\\
&=e^{-2\gamma_j(\sigma)\pi-\sigma(U(\pi)-U(0))}((-2\gamma_j^{(1)}(\sigma)\pi-U(\pi)+U(0))^2-2\gamma_j^{(2)}(\sigma)\pi).
\end{aligned}$$

Evaluating this at $\sigma=0$ gives us that the left side of equation (5.2) is equal to $(-2\gamma_j^{(1)}\pi-U(\pi)+U(0))^2-2\gamma_j^{(2)}\pi$.
Now for the right-hand we will first calculate the first derivative in σ

$$\begin{aligned}
\partial_\sigma\int_0^{\pi}\sigma U'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx&=\int_0^{\pi}U'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx\\
-\int_0^{\pi}\sigma U'(x)(U(x)-U(0))e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx&-\int_0^{\pi}2\sigma\gamma_j^{(1)}(\sigma)xU'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx.
\end{aligned}$$

Now we need to take a second derivative, we will do this for all of the three integrals and then evaluate them at $\sigma=0$.
For the first integral we have:

$$\begin{aligned}
&\partial_\sigma\int_0^{\pi}U'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&=-\int_0^{\pi}U'(x)(U(x)-U(0))e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}-2\int_0^{\pi}\gamma_j^{(1)}(\sigma)xU'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&=-\int_0^{\pi}U'(x)(U(x)-U(0))e^{-2inx}\,dx-\int_0^{\pi}2\gamma_j^{(1)}xU'(x)e^{-2inx}dx.
\end{aligned}$$

For the second integral we have:

$$\begin{aligned}
&-\partial_\sigma\int_0^{\pi}\sigma U'(x)(U(x)-U(0))e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&=-\int_0^{\pi}U'(x)(U(x)-U(0))e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}+\int_0^{\pi}\sigma U'(x)(U(x)-U(0))^2e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&+2\int_0^{\pi}\gamma_j(\sigma)\sigma U'(x)(U(x)-U(0))xe^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&=-\int_0^{\pi}U'(x)(U(x)-U(0))e^{-2inx}\,dx.
\end{aligned}$$

For the third integral we have:

$$
\begin{aligned}
&-\partial_\sigma \int_0^\pi 2\sigma\gamma_j^{(1)}(\sigma)xU'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&=-\int_0^\pi 2\gamma_j^{(1)}(\sigma)xU'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}-\int_0^\pi 2\sigma\gamma_j^{(2)}(\sigma)xU'(x)e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&+\int_0^\pi 2\sigma\gamma_j^{(1)}(\sigma)xU'(x)(U(x)-U(0))e^{-\sigma(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&+\int_0^\pi 2\sigma(\gamma_j^{(1)}(\sigma)x)^2U'(x)e^{-\sigma(2\gamma_j^{(1)}(\sigma)x)^2(U(x)-U(0))-2\gamma_j(\sigma)x}\,dx|_{\sigma=0}\\
&=-\int_0^\pi 2\gamma_j^{(1)}xU'(x)e^{-2inx}\,dx.
\end{aligned}
$$

Now equating the right-hand with the left-hand will give us:

$$(2\gamma_j^{(1)}\pi+U(\pi)-U(0))^2-2\gamma_j^{(2)}\pi=\pm\int_0^\pi(2(U(x)-U(0))+4\gamma_j^{(1)}x)U'(x)e^{-2inx}x\,dx.$$

Furthermore if then solve for $\gamma_j^{(2)}$ we get:

$$\gamma_j^{(2)}=\frac{1}{2\pi}(2\gamma_j^{(1)}\pi+U(\pi)-U(0))^2\pm\frac{1}{\pi}\int_0^\pi((U(x)-U(0))+2\gamma_j^{(1)}x)U'(x)e^{-2inx}x\,dx.$$

□

One thing that should be noted is that using a similar strategy and by taking more derivatives of $Z(\gamma(\sigma))$ it should be possible to get even more terms of the Taylor expansion. However, this becomes very difficult to do as the derivatives get higher due to the many times that the product rule will be used. For this reason, we also assumed that U was symmetric for the second derivative, but this is not an inherent requirement and was simply chosen for the convenience of the computation.

We can then experiment with using a non-zero value for the refreshment rate λ_0. To see how λ_0 influences the spectrum for small values of λ_0. It turns out that as long as σ and λ_0 are small at the same time, that the paths of the eigenvalues are linearly dependent on λ_0. Furthermore, we already have a description of the spectrum of $\mathscr{L}_{0,\lambda_0}$ [MM12], for small λ_0 this causes all but one of the eigenvalues to lie on a vertical line at $\gamma=-\lambda_0$. This leads us to the following conjecture.

Conjecture 5.2.3. *Assume the same assumptions as in Proposition 5.2.2 except for λ_0. Moreover assume that $\lambda_0(x)$ is a constant function independent of $x\in\mathbb{T}$. We then have that for the eigenvalue expansions $\gamma_j(\sigma)$ the following holds*

$$\gamma_j(\sigma)\to in \quad \textit{for some} \quad n\in\mathbb{Z},\ \textit{and}$$

$$\frac{1}{h}(\gamma_j(\sigma)-in)\to\gamma_j^{(1)}.$$

For this limit $\gamma_j^{(1)}$, it holds for $n\neq 0$ that

$$\gamma_j^{(1)}=-\frac{U(\pi)-U(x_0)}{2\pi}-\lambda_0\pm\frac{1}{2\pi}\sqrt{(\int_{x_0}^{\pi}U'(x)e^{-2in(x-x_0)}\,dx)(\int_{x_0}^{-\pi}U'(x)e^{2in(x-x_0)}\,dx)},$$

and for $n=0$ we have

$$\gamma_j^{(1)}=-\sigma\frac{U(\pi)-U(0)}{\pi}-2\lambda_0,$$

or

$$\gamma_j^{(1)}=0.$$

It was not possible to find a formula for general λ_0, this one likely exists by introducing in a similar way the function W. For now, this conjecture gives us that all of the spectrum moves linearly fast based on λ_0 if σ is close to 0.

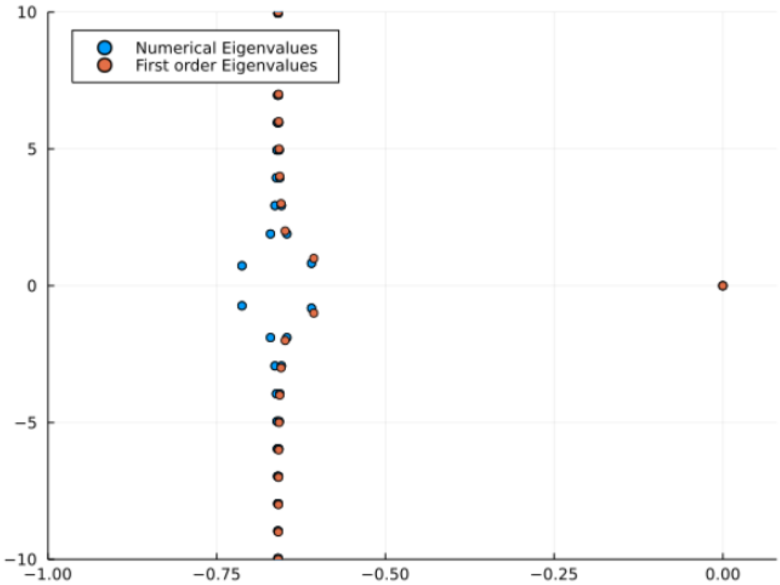

Figure 5.5: The spectrum of $\mathscr{L}_N$ and the approximated eigenvalues with $N = 900$, potential $U(x) = -\frac{\cos(x)}{2}$, and $\lambda_0 = 0.5$. The eigenvalues of $\mathscr{L}_N$ are displayed by the blue dots. The approximated eigenvalues from conjecture 5.2.3 are displayed by the orange dots. Note that we only displayed the positive roots for these approximated eigenvalues.

6
Discussion

In this thesis, we have studied the spectral gap κ of the generator of the Zig-Zag process $\mathscr{L}$ on $\mathbb{T}$ for both square-integrable functions $L^2(\mathbb{T},\nu)$ and continuous functions $C(\mathbb{T} \times \{+1,-1\})$. We did this by analysing the Schur complement $S_1(\gamma)$,$\gamma \in \mathbb{C} \setminus \{0\}$ of a block operator $\mathscr{A}$ that is unitarily equivalent to the generator of the Zig-Zag process. We proved that the generator has a point spectrum $\sigma(\mathscr{L}) = \sigma_p(\mathscr{L})$. We then made bounds on the numerical range of the Schur complement in order to bound the eigenvalues $\sigma_p(\mathscr{L}) \cap \mathbb{C} \setminus \mathbb{R}$. Moreover, we proved that $\gamma = 0$ is an element of the resolvent set and explicitly calculated the corresponding inverse for the generator. This inverse allowed us to also make a bound on the eigenvalues $\sigma_p(\mathscr{L}) \cap \mathbb{R}$ and consequently a bound on the spectral gap. After this, we noticed that the Schur complement has a point spectrum and that the Schur complement becomes a self-adjoint operator for real parameters ($S_1(\gamma)^* = S_1(\gamma)$ for $\gamma \in \mathbb{R} \setminus \{0\}$). Using these two facts we could apply the min-max theorem and noticed that we could carry over this min-max theorem to an operator BB^* corresponding to the Kolmogorov backward equation of the overdamped Langevin diffusion. Using this, we were able to create a second bound on the spectral gap using the second eigenvalue of BB^* which we could then bound by using the Hilbert-Schmidt norm of $\tilde{B}^*$. After this, we proved that the spectrum of $\mathscr{L}$ is the same regardless of whether Banach space consists out of the square-integrable functions or continuous functions. Moreover, we proved for unimodal potentials that the spectrum could be defined by the roots of a complex holomorphic function $Z(\gamma)$. We used this function to prove the existence of an asymptotic spectral line L to which all the eigenvalues converge. This line caused some complications for the spectral mapping theorem which were not there for the Zig-Zag process defined on the real numbers $\Omega = \mathbb{R}$. We then proved a spectral mapping theorem and noticed that this could result in the spectrum of the associated semigroup of $\mathscr{L}$ having an uncountable amount of eigenvalues. Moreover this allowed us to prove that the best possible value of κ for the hypocoercive estimate $||P(t)f - \int_{\mathbb{T}} f d\nu||_X \leq Me^{-\kappa t}$ is equal to the spectral gap for both $X = L^2(\mathbb{T},\nu)$ and $X = C(\mathbb{T} \times \{+1,-1\})$. Lastly, we created a discretization of the generator $\mathscr{L}_N^1$ and proved that this generator generated a strongly continuous contraction semigroup that converged uniformly on bounded time intervals for all the elements of $C(\mathbb{T} \times \{+1,-1\})$. We noticed that the eigenvalues of this discretization seemed to converge correctly to the spectrum of $\mathscr{L}$ for some examples but that the convergence was very slow. To improve on this we introduced another generator $\mathscr{L}_N$ that converged at a significantly higher rate and used this to numerically analyse the spectrum of the generator for some specific potentials and refreshment rates. Using this discretized generator we conjectured a generalization of the spectral asymptotic line for multimodal potentials and general refreshment rates. Lastly, we made use of a perturbation of the generator by considering a family of generator $\mathscr{L}_\sigma$ such that $\mathscr{L}_{\sigma=0} = \theta\partial_x$ to state some formula's for how the eigenvalues behave for small σ.

The analysis of the spectrum has a lot of different aspects to it which also gives a lot of directions for further research. We will list several points and observations that could help for further research.

Firstly, the results are likely to extend to other PDMPs such as the bouncy particle sampler similar to the work of [GN20]. This is mainly because many results in the past have been easy to transfer between the PDMPs and because these operators are similar in their definition. We did not do this generalization in this thesis as the focus was on finding many properties surrounding the generator of the Zig-Zag process rather than surrounding general PDMPs.

The spectrum of the generator defined on the Banach spaces $L^p(\mathbb{T},\nu)$ with $1 \le p \le \infty$ is likely to be the same as we can likely prove similarly that the eigenvectors of the generators have to sufficiently smooth. This could be useful for analysing the properties of the eigenvectors/eigenvalues as different norms for the Banach space can give different estimations on the eigenvalues/eigenvectors.

The analysis could be extended to the position space $\mathbb{T}^d$ with $d \ge 2$ however there can be some complications with explicitly calculating the inverse of $\mathscr{L}$ and $\tilde{B}$. Moreover applying Sobolev embeddings can give rise to more complications in higher dimensions as different embeddings apply. Nevertheless analysing the spectrum further with the one-dimensional case using methods we used such as the Schur complement can still be useful to prove certain properties that then might generalize but require a different proof for the higher-dimensional case.

We have not compared how the bounds that we created behave for certain limits of the potential (high-temperature/low-temperature limit). Moreover, it could be possible to compare the two given bounds numerically by creating new bounds on them using the Hilbert-Schmidt norm of the inverse of the generator $\mathscr{L}^{-1}$ and of $\tilde{B}$. We can then compare how these bounds behave by comparing the bounds with the actual spectral gap with the use of $\mathscr{L}_N$. Perhaps the bounds work better than each other depending on the type of potential or refreshment rate.

By using the unitary transformation $e^{\frac{-U}{2}}$ it is possible to rewrite the operator family $\mathscr{L}_\sigma = T + \sigma T'$ as we did it in section 5.2.2 in such a way that T becomes a skew adjoint operator and that $T' = \begin{bmatrix} W & \frac{U'}{2} \\ \frac{-U'}{2} & 0 \end{bmatrix}$. Which could then be written as an operator with only the entry W which would create another self-adjoint operator and an operator with the entries $\frac{U'}{2}$ which would create a skew-adjoint operator with only 2 entries. Since this operator then has many symmetrical properties it could be more doable to prove a stronger result regarding the analytic expansion of the eigenvalues.

The discretization $\mathscr{L}_N$ can also be used to analyse the eigenvectors of $\mathscr{L}_N$, we have observed that these eigenvectors behave more like sine and cosine functions as the imaginary component of the corresponding eigenvalue goes to infinity. There is likely a stronger connection with the operator $\theta\partial_x$ as the imaginary component becomes larger. For this reason, perturbation techniques could also help prove more properties about the asymptotic line as regardless of how large σ is in the scaling of the operator family $\mathscr{L}_\sigma$ we always seem to have for large eigenvalues that the spectrum looks similar to this operator. Moreover, such perturbation techniques could also give us that for large enough imaginary components that the resolvent of the operator is well behaved or that the resolvent has a large radius of convergence for large enough eigenvalues.

There can also be weaker assumptions made on λ_0 and U. Many assumptions were made in order to make the generator map $L^2_0(\mathbb{T},\nu)$ and $C(\mathbb{T} \times \{+1,-1\})$ into themselves. For example, for $L^2_0(\mathbb{T},\nu)$ we required U to be Lipschitz to perform the product rule for the well-definedness of $\mathscr{L}$, but we could still create a well-defined generator if we simply change the domain and co-domain to less 'perfect' domains. Although by taking such a generalization we no longer have that Theorem 3.4.7 holds due to the unboundedness of the derivative of U (although Theorem 3.4.2 could still hold).

Similar to using a numerical tool such as $\mathscr{L}_N$ for analysing the spectra we could also make use of pseudospectra which in particular can give more information about the growth of the resolvent. Perhaps this could create a stronger image of what is happening as the imaginary component of the eigenvalues goes to infinity and would allow us to prove a more general spectral mapping theorem.

We have that for $\sigma = 0$ that the operator family $\mathscr{L}_\sigma$ has a multiplicity of 2 for each eigenvalue after which these eigenvalues branch when σ becomes positive. This seems to indicate that for most eigenvalues we will have that the multiplicity of the eigenvalue is 1. Moreover, for small perturbations it could be that it is always equal to 1 which seems to be the case for most numerical analyses that we did. This could be useful as it could also give us a conjecture of the form of the spectral projections as they were given in [BV21] for rank one projections.

Due to the resolvent being Hilbert-Schmidt we have a lot of regularity that is added to our operator that we have not examined properly. Such as that the difference between resolvents is a trace operator due to the resolvent identity. Furthermore, we could likely use properties like this to say more about the expansion rate of the eigenvalues away from the origin since the squared norm of the inverse of the eigenvalues needs to be finite. Moreover, there could be general properties corresponding to the general class of J-self-adjoint operators that have a Hilbert-Schmidt resolvent that we have not yet examined.

Due to the invertibility of the generator only requiring that $||W||_\infty > 0$, the assumption of Theorem 3.4.2 and 3.4.7 that $\operatorname{essinf}_{x\in\mathbb{T}} W(x) > 0$ can likely be improved. This main assumptions is made for the eigenvalues in $\mathbb{C} \setminus \mathbb{R}$. In general, it seems to be that the spectral gap is always positive based on the numerical analysis. So the analysis on whether the gap is always positive should require us to examine these non-real eigenvalues. This analysis could become easier as we have more structure about these eigenvalues in the sense that we know what their real and imaginary parts are equal to in terms of the eigenvectors themselves (Proposition 3.4.3). We suspect that there could be more analysis done by also using the second Schur complement $S_2(\gamma)$ as this operator family is only defined for the non-real eigenvalues.

In the literature surrounding the quadratic numerical range, there is often analysis done towards the positive and negative eigenvectors of the quadratic range. These correspond to which root we take of the eigenvalues similar to how we do this in Proposition 3.4.3. In [Tre08] it turns out that these eigenvectors have some distinct behaviour and this could allow us to perhaps couple these eigenvectors together in such a way that we get more regularity in how they can behave for real eigenvalues.

Lastly, the field of quadratic eigenvalue problems is often concerned with oscillators that are receiving strong drag forces such as in [TM01]. Here there is an interpretation behind the eigenvalues first being imaginary and then becoming real as the drag forces increase. Namely, as these correspond to waves that are underdamped, overdamped, and critically damped. We suspect that such a connection is happening here with the eigenvalues and eigenvectors of $\mathscr{L}_N$ and that perhaps similar analysis techniques could be applied.

Bibliography

[And89] Alan L Andrew. "Correction of finite difference eigenvalues of periodic Sturm-Liouville problems". In: *The ANZIAM Journal* 30.4 (1989), pp. 460–469.

[And+19] Christophe Andrieu et al. *Hypocoercivity of Piecewise Deterministic Markov Process-Monte Carlo.* 2019. arXiv: 1808.08592 [stat.CO].

[Ber+21] E. Bernard et al. *Hypocoercivity with Schur complements.* 2021. arXiv: 2003.00726 [math.AP].

[BD17] Joris Bierkens and Andrew Duncan. "Limit theorems for the zig-zag process". In: *Advances in Applied Probability* 49.3 (Sept. 2017), pp. 791–825. ISSN: 1475-6064. URL: http://dx.doi.org/10.1017/apr.2017.22.

[BFR19] Joris Bierkens, Paul Fearnhead, and Gareth Roberts. "The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data". In: *The Annals of Statistics* 47.3 (June 2019). ISSN: 0090-5364. URL: http://dx.doi.org/10.1214/18-AOS1715.

[BRZ19] Joris Bierkens, Gareth O. Roberts, and Pierre-André Zitt. "Ergodicity of the zigzag process". In: *The Annals of Applied Probability* 29.4 (Aug. 2019). ISSN: 1050-5164. URL: http://dx.doi.org/10.1214/18-AAP1453.

[BV21] Joris Bierkens and Sjoerd M. Verduyn Lunel. *Spectral analysis of the zigzag process.* June 2021. URL: http://arxiv.org/abs/1905.01691.

[Bie+20] Joris Bierkens et al. "The Boomerang Sampler". In: *Proceedings of the 37th International Conference on Machine Learning.* Ed. by Hal Daumé III and Aarti Singh. Vol. 119. Proceedings of Machine Learning Research. PMLR, 13–18 Jul 2020, pp. 908–918. URL: https://proceedings.mlr.press/v119/bierkens20a.html.

[BVD18] Alexandre Bouchard-Côté, Sebastian J. Vollmer, and Arnaud Doucet. "The Bouncy Particle Sampler: A Nonreversible Rejection-Free Markov Chain Monte Carlo Method". In: *Journal of the American Statistical Association* 113.522 (2018), pp. 855–867. DOI: 10.1080/01621459.2017.1294075.

[Cep03] David Ceperley. *Metropolis Methods for Quantum Monte Carlo Simulations.* June 2003. DOI: 10.1063/1.1632120.

[Con99] D.J. Condon. "Corrected finite difference eigenvalues of periodic Sturm–Liouville problems". en. In: *Applied Numerical Mathematics* 30.4 (July 1999), pp. 393–401. ISSN: 01689274. DOI: 10.1016/S0168-9274(98)00093-2.

[Con97] John B. Conway. *A course in functional analysis.* en. 2nd ed. Graduate texts in mathematics 96. New York: Springer, 1997. ISBN: 978-0-387-97245-9.

[DSZ02] Michael Dellnitz, Oliver Schütze, and Qinghua Zheng. "Locating all the zeros of an analytic function in one complex variable". In: *Journal of Computational and Applied Mathematics* 138 (Jan. 2002), pp. 325–333. DOI: 10.1016/S0377-0427(01)00371-5.

[DM13] Persi Diaconis and Laurent Miclo. "On the spectral analysis of second-order Markov chains". In: *Annales de la Faculté des Sciences de Toulouse. Mathématiques.* 22.3 (2013). 36 pages, pp. 573–621. URL: https://hal.archives-ouvertes.fr/hal-00719047.

[EN00] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations.* Springer-Verlag, 2000. ISBN: 0-387-98463-1.

[ET18] Christian Engström and Axel Torshage. *Accumulation of Complex Eigenvalues of a Class of Analytic Operator Functions.* 2018. arXiv: 1709.01462 [math.FA].

[EL04] David Eschwé and Matthias Langer. "Variational Principles for Eigenvalues of Self-adjoint Operator Functions". In: *Integral Equations and Operator Theory* 49.3 (July 2004), pp. 287–321. ISSN: 1420-8989. URL: https://doi.org/10.1007/s00020-002-1209-5.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. New York: John Wiley & Sons Inc., 1968, pp. xviii+509.

[Gea78] Larry Gearhart. "Spectral theory for contraction semigroups on Hilbert space". en. In: *Transactions of the American Mathematical Society* 236.0 (1978), pp. 385–394. ISSN: 0002-9947, 1088-6850. DOI: 10.1090/S0002-9947-1978-0461206-1.

[Gir+20] Juan Giribet et al. "Spectral enclosures for a class of block operator matrices". In: *Journal of Functional Analysis* 278.10 (June 2020), p. 108455. ISSN: 0022-1236. URL: http://dx.doi.org/10.1016/j.jfa.2019.108455.

[GN20] Arnaud Guillin and Boris Nectoux. "Low lying eigenvalues and convergence to the equilibrium of some Piecewise Deterministic Markov Processes generators in the small temperature regime". working paper or preprint. Jan. 2020. URL: https://hal.archives-ouvertes.fr/hal-02436593.

[Gus70] Karl E. Gustafson. *The Toeplitz-Hausdorff theorem for linear operators.* 1970. DOI: https://doi.org/10.1090/S0002-9939-1970-0262849-9.

[GR97] Karl E. Gustafson and Duggirala K. M Rao. *Numerical Range: The Field of Values of Linear Operators and Matrices.* English. ISBN: 9781461384984 OCLC: 1159051455. 1997.

[HR08] Emmanuel Hebey and Frédéric Robert. "Sobolev spaces on manifolds". en. In: *Handbook of Global Analysis.* Elsevier, 2008, pp. 375–415. ISBN: 978-0-444-52833-9. DOI: 10.1016/B978-044452833-9.50008-5.

[Hér17] Frédéric Hérau. "Introduction to hypocoercive methods and applications for simple linear inhomogeneous kinetic models". In: *Lectures on the Analysis of Nonlinear Partial Differential Equations.* Vol. 5. Morning Side Lectures in Mathematics Series MLM 5. 2017, pp. 119–147. URL: https://hal.archives-ouvertes.fr/hal-01616979.

[HHS93] Chii-Ruey Hwang, Shu-Yin Hwang-Ma, and Shuenn-Jyi Sheu. "Accelerating Gaussian Diffusions". In: *Annals of Applied Probability* 3 (1993), pp. 897–913.

[IK98] Kazufumi Ito and Franz Kappel. "The Trotter-Kato theorem and approximation of PDEs". en. In: *Mathematics of Computation* 67.221 (1998), pp. 21–44. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/S0025-5718-98-00915-6.

[Kac51] Mark Kac. *On Some Connections between Probability Theory and Differential and Integral Equations.* 1951.

[Kat95] Tosio Kato. *Perturbation Theory for Linear Operators, Second Edition (Classics in Mathematics) by.* Springer-Verlag Berlin Heidelberg New York, 1995. ISBN: 3-540-58661-X.

[Kna17] Anthony W. Knapp. *Advanced Real Analysis: Digital Second Edition, Corrected version.* en. East Setauket, New York: Anthony W. Knapp, 2017. ISBN: 978-1-4297-9991-1. DOI: 10.3792/euclid/9781429799911.

[KLT04] Margarita Kraus, Matthias Langer, and Christiane Tretter. "Variational principles and eigenvalue estimates for unbounded block operator matrices and applications". en. In: *Journal of Computational and Applied Mathematics* 171.1-2 (Oct. 2004), pp. 311–334. ISSN: 03770427. DOI: 10.1016/j.cam.2004.01.024.

[LLT02] Heinz Langer, Matthias Langer, and Christiane Tretter. "Variational Principles for Eigenvalues of Block Operator Matrices". In: *Indiana University Mathematics Journal* 51.6 (2002), pp. 1427–1459. ISSN: 00222518, 19435258. URL: http://www.jstor.org/stable/24902845.

[LS17] Matthias Langer and Michael Strauss. "Spectral properties of unbounded J-self-adjoint block operator matrices". en. In: *Journal of Spectral Theory* 7.1 (2017), pp. 137–190. ISSN: 1664-039X. DOI: 10.4171/JST/158.

[LW21] Jianfeng Lu and Lihan Wang. *On explicit L^2-convergence rate estimate for piecewise deterministic Markov processes in MCMC algorithms.* 2021. arXiv: 2007.14927 [math.PR].

[Mar88] A. S. Markus. *Selfadjoint pencils.* English. Vol. 71. American Mathematical Society, Sept. 1988. ISBN: 978-0821890820.

[Mic19] Laurent Michel. "About small eigenvalues of the Witten Laplacian". In: *Pure and Applied Analysis* 1.2 (Apr. 2019), pp. 149–206. ISSN: 2578-5893. URL: http://dx.doi.org/10.2140/paa.2019.1.149.

[Mon16] Pierre Monmarché. *Piecewise deterministic simulated annealing*. 2016. arXiv: 1410.1656 [math.PR].

[MM12] Pierre Monmarché and Laurent Miclo. *Étude spectrale minutieuse de processus moins indécis que les autres*. 2012. arXiv: 1209.3588 [math.PR].

[MB20] Nicola F. Müller and Remco R. Bouckaert. "Adaptive Metropolis-coupled MCMC for BEAST 2". en. In: *PeerJ* 8 (Sept. 2020), e9473. ISSN: 2167-8359. DOI: 10.7717/peerj.9473.

[PHA81] John W Paine, Frank R de Hoog, and Robert S Anderssen. "On the correction of finite difference eigenvalue approximations for Sturm-Liouville problems". In: *Computing* 26.2 (1981), pp. 123–139.

[PW12] E. A. J. F. Peters and G. de With. "Rejection-free Monte Carlo sampling for general potentials". In: *Physical Review E* 85.2 (Feb. 2012). ISSN: 1550-2376. URL: http://dx.doi.org/10.1103/PhysRevE.85.026703.

[PP17] Yves Pomeau and Jarosław Piasecki. "The Langevin equation". en. In: *Comptes Rendus Physique* 18.9-10 (Nov. 2017), pp. 570–582. ISSN: 16310705. DOI: 10.1016/j.crhy.2017.10.001.

[RS78] Michael Reed and Barry Simon. *Methods of modern mathematical physics. IV.* English. OCLC: 936547877. 1978. ISBN: 978-0-08-057045-7.

[RR04] Gareth O. Roberts and Jeffrey S. Rosenthal. "General state space Markov chains and MCMC algorithms". In: *Probability Surveys 1* (2004), pp. 20–71. DOI: 10.1214/154957804100000024.

[Sch12] Konrad Schmüdgen. *Unbounded Self-adjoint Operators on Hilbert Space*. Vol. 265. Graduate Texts in Mathematics. Dordrecht: Springer Netherlands, 2012. ISBN: 978-94-007-4752-4 978-94-007-4753-1. DOI: 10.1007/978-94-007-4753-1.

[SGS12] Yi Sun, Faustino Gomez, and Juergen Schmidhuber. *Improving the Asymptotic Performance of Markov Chain Monte-Carlo by Inserting Vortices*. 2012. arXiv: 1209.6048 [stat.ME].

[Sun16] V. S. Sunder. *Operators on Hilbert Space*. Singapore, 2016. DOI: 10.1007/978-981-10-1816-9.

[Thi16] John Thickstun. *The Volterra operator*. 2016. URL: https://johnthickstun.com/docs/volterra.pdf.

[Thi07] Jos Thijssen. "The Monte Carlo method". In: *Computational Physics*. 2nd ed. Cambridge University Press, 2007, pp. 295–337. DOI: 10.1017/CBO9781139171397.012.

[TM01] Françoise Tisseur and Karl Meerbergen. "The Quadratic Eigenvalue Problem". In: *SIAM Review* 43 (Aug. 2001). DOI: 10.1137/S0036144500381988.

[Tre08] Christiane Tretter. *Spectral theory of block operator matrices and applications*. Imperial College Press, 2008. ISBN: 978-1-86094-768-1.

[Vil06] C. Villani. "Hypocoercivity". In: (Sept. 2006). URL: http://arxiv.org/abs/math/0609050.

[Web18] Joa Weber. *Introduction to Sobolev spaces*. Dec. 2018. URL: http://www.math.stonybrook.edu/~joa/PUBLICATIONS/SOBOLEV.pdf.

[Wia] Sjoerd Wiarda. *Numerical computation of the spectrum of the Zig-Zag process on the torus*. URL: https://github.com/SjoerdWiarda/ZigZagSpectrum.