Analytic Continuation of TBA Equations As A Method of Solving Integrable Models: Application to the Sinh-Gordon Model

BRANDON MICHEAL MORRISON

Analytic Continuation of TBA Equations As A Method of Solving Integrable Models

Application to the Sinh-Gordon Model

Brandon Mícheál Morrison

A Thesis presented for the degree of
Doctor of Philosophy

Centre for Particle Theory
Department of Mathematical Sciences
Durham University
United Kingdom

November 2022

Declaration

The work in this thesis is based on research carried out in the Department of Mathematical Sciences at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification.

Acknowledgements

I would like to thank my family and friends for putting up with me throughout the years. They have always supported me in my pursuits, and without them I would be a lesser man.

I would like to thank Patrick Dorey for agreeing to be my supervisor and for shaping my thinking towards that of a proper mathematician and physicist.

I have always had the good fortune to meet mathematicians and scientists who were always willing to share their knowledge and experience with me wherever I ended up working and living. I am in their debt as well.

When you arise in the morning think of what a privilege it is to be alive, to think, to enjoy, to love ...

— from *Meditations* by Marcus Aurelius

Ní bhíonn saoi gan locht

— from *an old Gaelic Proverb*

Dedicated to

My family

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the primary interests one has when presented with any quantum mechanical model concerns determining its energy spectrum. Given an appropriate Hamiltonian, one is generally interested in solving the system to obtain the spectrum of possible energy values if one was to experimentally measure its total energy.

A similar desire to obtain the energy spectrum exists when one works with a quantum field theory. Unlike quantum mechanics, however, obtaining the energy spectrum of a quantum field theory (QFT) is not nearly as simple. In QFTs, there exist an infinite number of degrees of freedom, with roughly at least one degree of freedom for every point in space. These degrees of freedom are also often coupled together, and this further complicates the theory. The exception being, of course, *free theories*, where one can decouple the dynamics and write down degrees of freedom that evolve independently of all the others. Though given the fact that there are no particle interactions present, these theories cannot describe most real-world phenomena. It would therefore be nice if there were techniques to solve some interacting theories exactly without having to use perturbation theory.

In its most basic incarnation, QFT is the study of renormalization group (RG) flows. For a QFT this roughly means that the continuum conformal field theory in the ultraviolet spectrum, i.e. CFT_{UV}, when perturbed by an operator with scale M initiates a reorganization of the degrees of freedom of the theory which can produce a new CFT_{IR} in the infrared spectrum at large distances ($\gg M^{-1}$). This concerns how the theory evolves as it transverses through the ultraviolet (UV) to the infrared (IR) regimes. General QFTs have a characteristic mass scale ξ, known as the *correlation length*. The parameter r is the ratio $r = \frac{R}{\xi}$, where R is the radius of the cylinder on which the model sits. For the main sinh-Gordon model, the subject of this present study, we do not have a scale invariant theory. In the *ultraviolet limit*, the parameter is sent to zero, i.e. $r \to 0$. We now have a theory governed by conformal field theory

(CFT), which underlies the QFT.

There is also the *infrared limit*, which sends the parameter to infinity, i.e. $r \to \infty$. In this limit, the scattering is governed by the S-matrix. Under certain conditions, the S-matrix takes on a simple factorized form. In particularly special theories, known as *integrable theories*, the S-matrix can be solved for exactly. This feature is key in enabling us to develop a method to solve for the exact spectrum of integrable field theories for both interacting and free particle models.

For Euclidean quantum field theories in (1+1)-dimensions, one may make use of the *thermodynamic Bethe ansatz* (TBA) to analyze finite-size effects of a given model, which allows for the extraction of perturbative and non-perturbative information. To do so, one essentially maps the theory onto a cylindrical surface and studies the spectrum from the resulting Hamiltonian.

Requiring only the exact S-matrix of an integrable QFT, one can derive an integral equation (TBA equation) to compute the ground-state energy of the integrable model. One can then analytically continue the TBA equation along an arbitrary path in the complex plane of the given model, parameterized by r, which implicitly depends upon the coupling-constant. Upon encircling a branch point of the model in the complex plane, we can obtain the next energy level. This corresponds to another sheet on the Riemann surface of the model, and which we will explain in detail later. The TBA equation then solves for a new 'ground-state energy' once again on this new part of the Riemann surface, and we repeat this process to obtain a further excited state. Note that with this method we may not be able to obtain the full spectrum of some integrable models, such as in the case where we have a model described by a disconnected Riemann surface. We analyze this method using the sinh-Gordon model.

In the final part of this thesis, we address the 'negative energy problem' we encountered when thinking about the technique abstractly. If we solve for the finite-size ground state of a free massive boson and attempt to analytically continue it to obtain the excited states, we instead obtain an 'excited energy' that actually has less energy than the ground-state. In discovering why such behaviour occurs, we found a far more elaborate geometric structure to these systems than previously realized. These ideas are further explored later, and will conclude this thesis.

Chapter 2

Riemann Surfaces and Analytic Continuation

This chapter deals with the mathematical theory of Riemann surfaces and analytic continuation, two topics surprisingly not normally taught in most contemporary courses on complex analysis, but of fundamental importance to our method to solve for the spectrum of the sinh-Gordon equation. Since one can find a plethora of good resources dealing with these subjects in full mathematical detail, we seek here only to introduce the motivation and the theory for both topics with the intention to present the material rigorously enough for our purposes. We refer the reader to the sources used in compiling this explanation for a more complete explanation of the details. They are [13], [30], [14], [11], and [4].

2.1 Riemann Surfaces

A first covering of complex analysis introduces one to a theory that rests on the idea that all of the functions, and their derivatives, are single-valued. Values for a single-valued complex function can be easily expressed on the complex plane. One will obtain a single-value for any given value of the variable in the function, which can then be plotted on the complex plane. What happens if we now consider a multi-valued function instead?

The idea of plotting the values of a multi-valued function for a given value of a variable on the complex plane now runs into issues. How can one express an output value for the argument of a complex number as π and 3π in the complex plane clearly given the fact that they are two different values, but would coincide in the complex plane as the latter is a full rotation from the first? More importantly, how do we even make sense of such results? It turns out that we can extend the

theory of analytic functions to include a wide class of multi-valued functions using a geometrical construction known as a Riemann surface.

2.1.1 The Problem

We attempt to illustrate the problem with multi-valued functions we face with an example. The logarithm of a complex variable $\log(z)$ is the function $f(z)$ that satisfies the equation

$$e^{f(z)} = z \tag{2.1.1}$$

If we now set $f(z) = u + iv$, i.e. break up our function into a real and an imaginary part, and make use of the Euler formula,

$$e^{i\phi} = \cos(\phi) + i\sin(\phi), \tag{2.1.2}$$

in (2.1.1) we obtain

$$e^{u+iv} = e^u e^{iv} = e^u(\cos(v) + i\sin(v)) = r(\cos(\theta) + i\sin(\theta)), \tag{2.1.3}$$

where in the last equality we noticed the form was identical to a well-known relation in complex analysis.

This allows us to now equate the real and imaginary parts of (2.1.3) to obtain for u,

$$u = \log(r), \tag{2.1.4}$$

and for v,

$$v = \theta + 2\pi n \qquad (n = 0, \pm 1, \pm 2, ...), \tag{2.1.5}$$

where $u, v \in \mathbb{R}$. Thus we end up with the formula

$$\log(z) = \log(r) + i(\theta + 2\pi n) \qquad (n = 0, \pm 1, \pm 2, ...). \tag{2.1.6}$$

Recall that the *argument* ($arg = \theta + 2\pi n$) of a complex number $z = re^{i\theta}$ is determined up to a constant multiple of 2π, i.e. a single period around a circle. What happens if we were to express this on the complex plane?

Allow us to set $z = 1$. Since this is real, it is implied that $\theta = 0$. Now (2.1.6) becomes

$$\log(1) = \log(1) + i(0 + 2\pi n) \qquad (n = 0, \pm 1, \pm 2, ...). \tag{2.1.7}$$

Taking a look at (2.1.7) we see that the argument

$$\arg = (0 + 2\pi n) \qquad (n = 0, \pm 1, \pm 2, ...) \tag{2.1.8}$$

will have an infinite number of solutions to this equation. It is apparent that for different values of n, we will obtain different values of the function $\log(z)$. These distinct values are, however, now supposed to correspond to equivalent angles on the complex plane. Looking at this another way, we have found that distinct values of $\log(z)$ correspond to the same point on the complex plane. Since we intend to do complex analysis with multi-valued functions, we need to find a way to express multi-valued functions as single-valued functions on the complex plane.

2.1.2 The Solution - Riemann Surfaces

The Idea of the Branch Point and Branch Cut

Let us approach this problem in a particular manner. Say we start with a point $z = z_0$ that lies somewhere on the complex plane $\mathbb{C}$. Let us move this point along an arbitrary path in the complex plane, but ultimately a path that transverses counterclockwise around the origin of the complex plane, returning to the initial point $z = z_0$. As we carry this out, the function $\log(z)$ changes continuously as we move along the path. On completion of a full cycle, the value of $\log(z_0)$ will have increased by $2\pi i$. One may observe this by referencing (2.1.6) to obtain

$$\log(z) = \log(z_0) + 2\pi i. \tag{2.1.9}$$

Now any point z on the complex plane that has the property that the initial function does *not* return to the identical value it held beforehand after one completes any cycle around it we call a *branch point* of the function. For our initial example, we see that $z = 0$ is a branch point for the function $f(z) = \log(z)$. If we also consider the function $f(\frac{1}{z'}) = \log\left(\frac{1}{z'}\right)$ at $z' = 0$, we see that the point at infinity is also a branch point of $\log(z)$. For this particular function, no other branch points exist.

Now allow us to join the two branch points of $\log(z)$ by a curve. The curve's direction and shape can be arbitrary, but simple constructions are the most manageable. Thus we choose to define our curve as a simple line starting at the origin and continuing out toward infinity in an arbitrary direction. When defining the function, let us remove all of the points that lie along this curve from the domain, i.e. we 'cut' the complex plane along this curve. This defines the notion of a *branch cut*. Once again note that branch cuts are entirely arbitrary, although the simpler the construction, the easier it is to work with.

The New Functions

Let us assume that the z-plane is cut along the negative half of the real axis from $x = 0$ to $x = -\infty$. We are able to define the single-valued function

$$f_0(z) = f_0(r, \theta) = \log(r) + i\theta, \tag{2.1.10}$$

where $-\pi < \theta < \pi$ and $r > 0$. We are also able to define the single-valued functions

$$f_1(z) = f_1(r, \theta) = \log(r) + i(\theta + 2\pi), \tag{2.1.11}$$

$$f_{-1}(z) = f_{-1}(r, \theta) = \log(r) + i(\theta - 2\pi), \tag{2.1.12}$$

where $-\pi < \theta < \pi$ and $r > 0$ for both cases. Note that the the range of f_1 and $f_{-1}(z)$ assumes the identical values that the logarithm assumes when $\pi < \theta < 3\pi$ and $-3\pi < \theta < -\pi$ respectively. From these functions, let us now define an infinite series of single-valued functions

$$f_0(z), f_{\pm 1}(z), f_{\pm 2}(z), ...$$

where

$$f_n(z) = f_n(r, \theta) = \log(r) + i(\theta + 2\pi n) \tag{2.1.13}$$

where $-\pi < \theta < \pi$ and $r > 0$. This now causes $f_n(z)$ to assume the same values for $-\pi < \theta < \pi$ that the logarithm takes in the range

$$(2n - 1)\pi < \theta < (2n + 1)\pi \tag{2.1.14}$$

In this construction, we have replaced the multi-valued logarithm function by a series of distinct functions that are *analytic* in the z-plane that was cut along the negative real axis.

The Construction of the Sheets

Recall that if we take a point on the complex plane, move this point along a counterclockwise path around the plane, and end back at the exact point from which we started, we can have a contradiction. The value of a function at the initial point may in fact be different than when we return to that point, i.e. our function has undergone a *non-trivial monodromy*. The construction we have outlined in the previous section takes care of this contradiction. Each of the $f_n(z)$'s are analytic over the entire range on which they are defined, and thus it is no longer possible to encircle the branch point without crossing the branch cut. The f_n's themselves are not defined over the branch cut.

So what to do for the branch cut? There is certainly a discontinuity across the cut, as each $f_n(z)$ has a domain defined from just below to just after the cut itself. We have above the cut

$$f_n(r, \pi - \epsilon) = \log(r) + i[(2n+1)\pi - \epsilon] \qquad \text{for } \epsilon > 0, \tag{2.1.15}$$

and below the cut we have

$$f_n(r, -\pi + \epsilon) = \log(r) + i[(2n-1)\pi + \epsilon] \qquad \text{for } \epsilon > 0. \tag{2.1.16}$$

Therefore the discontinuity $f_n(z)$ across the cut is

$$\lim_{\epsilon \to +0} [f_n(r, \pi - \epsilon) - f_n(r, -\pi + \epsilon)] = 2\pi i. \tag{2.1.17}$$

Note that the value of $f_n(z)$ just above the branch cut is equivalent to the value of $f_{n+1}(z)$ just below the branch cut. This implies that we could superimpose an infinite number of cut complex planes stacked one on top of the other with each plane corresponding to a different n-value, i.e. $n = 0, \pm 1, \pm 2, ...$ Each of these planes is known as a *Riemann sheet*. We can then connect these Riemann sheets via identifying the upper surface of the n^{th}-plane with the lower surface of the $(n+1)$-plane, thereby creating an interconnected surface known as a *Riemann surface*, see Fig. 2.1.

In this construction, all branch points are common for all sheets. Crossing a branch cut corresponds to moving to a different sheet in the Riemann surface. From a sequence of single-valued functions defined in a single complex plane, we have now procured a continuous single-valued function defined on a Riemann surface. Our logarithmic function is now analytic everywhere except at the branch points, which now behave as singularities. We can now perform complex analysis on a multi-valued function on this surface!

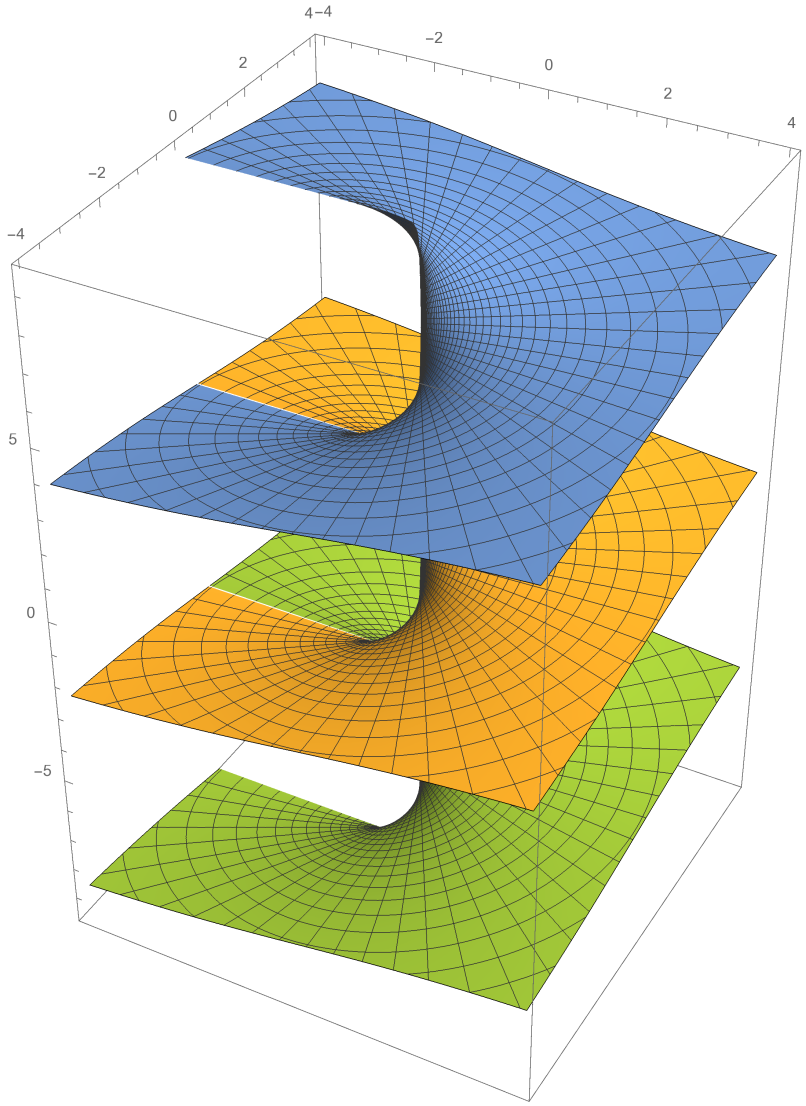


Figure 2.1: Riemann surface for the logarithmic function. Each new color corresponds to a different sheet of the Riemann surface. The line where two sheets (colors) meet is the branch cut. Crossing it takes one onto a new sheet of the surface.

2.2 Analytic Continuation

Analytic functions have many unique properties that make them useful for analysis. Cauchy's integral formula demonstrates that if a function is analytic in a certain region $\mathcal{A}$ of the complex plane and located on a curve delimiting $\mathcal{A}$, then the values of the function within the entire region $\mathcal{A}$ are fully obtained once we determine the values on the boundary curve.

What happens when a function $f(z)$ is only analytic within a smaller subset of a region $\mathcal{D}$? Do there exist certain subsets within $\mathcal{D}$ that have the property that if we

specify the values of $f(z)$ only over these subsets, we can determine $f(z)$ throughout the whole region? This question is of chief importance for the purposes of this thesis, and is addressed by the idea of analytic continuation. To continue, we will first need to explore the idea of uniqueness.

Theorem - Uniqueness:
Let $f_1(z)$ and $f_2(z)$ be two functions of the complex variable z that are both analytic within a region $\mathcal{D}$ of the complex plane. If these two functions are identical in the neighborhood of a point $z \in \mathcal{D}$, on a segment of a curve lying in $\mathcal{D}$, or on a point set with an accumulation point belonging to $\mathcal{D}$, then they are identical throughout D.

Proof: In a region where a function is analytic, it either has only *isolated zeros*, i.e. the zeros do not contain an accumulation point, or the function is identically equal to zero. Since $f_1(z)$ and $f_2(z)$ are surmised to be identical on a point set, we will have a set of zeros for the function

$$f_1(z) - f_2(z) = 0. \tag{2.2.1}$$

Now this set of zeros contains an accumulation point by design, and due to the fact that the function $f_1(z) - f_2(z)$ is analytic throughout $\mathcal{D}$, this function is identically equal to zero in $\mathcal{D}$, i.e.

$$f_1(z) - f_2(z) \equiv 0, \tag{2.2.2}$$

which yields

$$f_1(z) = f_2(z) \tag{2.2.3}$$

Q.E.D.

We have just shown that two *different* analytic functions cannot agree in the neighborhood of an arbitrary point. Thus the behavior of a function in a region where it is analytic is in fact uniquely determine by its behavior within the neighborhood of an arbitrary point.

If we take it that $f(z)$ is analytic in some region $\mathcal{D}$, we can perform an expansion of $f(z)$ in a Taylor series about an arbitrary point $z_0 \in \mathcal{D}$ to obtain

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n, \tag{2.2.4}$$

where a_n are the coefficients of the series with $n = 0, 1, 2,$ etc.

This Taylor series will converge within some disk η_0 defined by

$$r_0 = |z - z_0|, \tag{2.2.5}$$

where r_0 is the *radius of convergence*, which indicates the radius of the largest disk in which $f(z)$ will still converge. Now let us assume that we have knowledge of the behavior of $f(z)$ in the neighborhood of the point $z = z_0$, i.e. we know the coefficients a_n of the Taylor expansion within this neighborhood. This knowledge alone is just enough to determine the behavior of the function in the neighborhood of an arbitrary point $z_0' \in \mathcal{D}$. Note that this new point z_0' may be far from our initial point z_0. Assuming the standard Euclidean topology, a *region* in the complex plane is a subset of $\mathbb{C}^n$ that is open, connected, and non-empty. Thus it is possible to join the points z_0 and z_0' via a continuous path C lying completely within $\mathcal{D}$. We now take another point $z_1 \in C$ such that

$$r_0 > |z_1 - z_0|. \tag{2.2.6}$$

This simply indicates that z_1 lies within the disk η_0. Since the series (2.2.4) converges uniformly for $r_0 > |z - z_0|$ it can be differentiated term by term. In this way one may find all of the derivatives of $f(z)$ from its Taylor expansion at all points within the disk η_0, and specifically at the point z_1. This means that for the series

$$f(z) = \sum_{n=0}^{\infty} a_n^{(1)}(z - z_1)^n \tag{2.2.7}$$

the coefficients $a_n^{(1)}$) are known. Suppose that the sum above converges within a new disk η_1 defined by

$$r_1 = |z - z_1|. \tag{2.2.8}$$

Due to $f(z)$ being analytic on $C \subset \mathcal{D}$, r_1 is always non-zero. Thus if the analytic continuation exists, it is always possible to choose the point z_1, which can be located anywhere within η_0, in such a manner that the disk η_1 lies partially outside of the disk η_0, see Fig. 2.2.

Because of this ability, we are now able to calculate $f(z)$ in the original region η_0, where the domain of $f(z)$ was originally defined, as well as the augmented region consisting of the union η_0 and η_1 This augmented region will have a segment of the curve C that lies outside η_0, but within η_1.

If we repeat this process over and over again, we will cover the curve C by overlapping disks $\eta_0, \eta_1, ...$ which approach the point z_0'. After a number of iterations the point z_0' will finally be covered by one of the disks. This fact will enable us to find the Taylor expansion about this point and thus determine the behavior of the function $f(z)$ about the point $z = z_0'$, the task which we originally set out to accomplish. This technique is known as *analytic continuation*, as we are 'continuing' the domain of analyticity for a given function.

From the uniqueness theorem we proved, we know that analytic continuation is

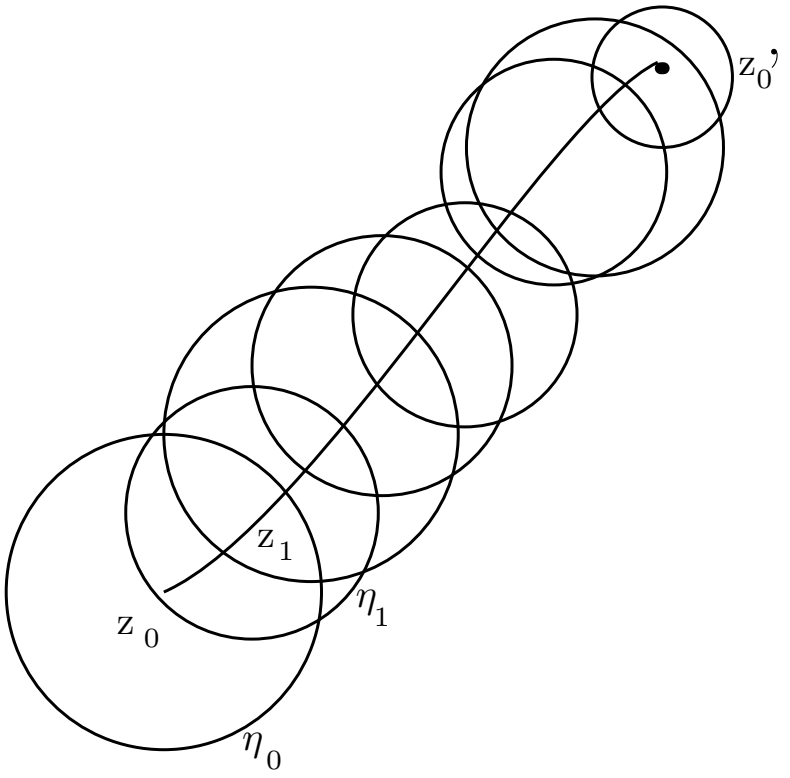


Figure 2.2: Representation of analytic continuation. Analytic continuation allows us to determine the behavior of our function in the neighborhood of an arbitrary point z_0'.

unique provided an analytic continuation of a function exists. Finding it might not always be easy, however. Note that this technique, by consequence of the definitions and assumptions it uses, necessarily determines the location of the singular points of $f(z)$. The radius of convergence of the Taylor expansion of $f(z)$ at a given point is equivalent to the distance from this point to the closest singularity of the function. Analytically continuing a function along such a curve which passes through a singular point of the function causes the radii of the disks, $\eta_0, \eta_1, etc.$, to converge toward zero as we approach the singularity. Thus it is impossible to circumvent the singularity, and the process of analytic continuation must come to a close.

2.3 The Importance Of It All

It is plausible that a reader might wonder why such mathematical techniques are important enough for our project that we include an entire section on them without any reference thus far to physical theories. The answer is rather simple, but of fundamental importance to our method.

When we derive the thermodynamic Bethe ansatz, we will end up with an equation

that is multi-valued. It will yield the ground-state energy of the system outright, but will not yield any more information regarding the excited-states. To determine the excited states, we will begin by computing the TBA on a computer grid. From here we will vary the path of the TBA so that we will move around this grid in appropriately small steps, calculating the TBA at every point we move. We employ the technique of analytic continuation to extend our domain of analyticity for the TBA, allowing us to venture further and further around the grid, and to approach and to circumnavigate branch points for a particular integrable field theory. Since the TBA equation calculates the ground-state energy, encircling the branch point will take us to the next branch of the Riemann surface, and thus the first excited-state. Note that the TBA equation itself may change in the process for reasons we will endeavour to explain later.

Although there is much more to say about this process, we will hold off from explaining all of the details until later. In the meantime, we will focus on building up the necessary theoretical background to understand and employ the technique.

Chapter 3

Conformal Field Theory

As mentioned in the introduction, in the ultraviolet limit a general (1+1)-dimensional QFT is governed by a conformal field theory. The main interest in this section is to introduce and derive the *central charge* which will be of importance in our quest to study the sinh-Gordon model. Therefore this chapter introduces conformal field theory and explains its application and importance to the current problem at hand. No attempt is made at being comprehensive, as the literature on the subject is vast and can be thoroughly studied via consulting the references used in this section. They are [8], [22], [28], [27], and [12].

3.1 Conformal Invariance

Consider the set of differentiable maps $\varphi : \cup \to \cup'$ where $\cup \in M$ and $\cup' \in M'$ are open subsets of manifolds M and M'. We concern ourselves with flat spacetime $g_{\mu\nu} \to \eta_{\mu\nu} = \text{diag}(-1,...,+1,...)$ with line element $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. A *conformal transformation* is an invertable map $x \to x'$ which leaves the metric tensor invariant up to a scale factor $\Lambda(x)$, i.e.

$$\eta'_{\mu\nu} = \Lambda(x)\eta_{\mu\nu}. \tag{3.1.1}$$

Recognizing that the new coordinates must be functions of the original ones, i.e. $x'(x)$, we can express this condition for the flat metric in the following manner:

$$\eta_{\rho\sigma}\frac{\partial x'^\rho}{\partial x^\mu}\frac{\partial x'^\sigma}{\partial x^\nu} = \Lambda(x)\eta_{\mu\nu} \ . \tag{3.1.2}$$

A conformal transformation is *local* if the subset $\cup \neq M$, and *global* if the subset $\cup = M$. Setting the scale factor $\Lambda(x) = 1$ corresponds to the Poincaré group, which consists of the translations together with the Lorentz transformations (rotations and boosts) and truly leaves the metric invariant. We now focus on infinitesimal

coordinate transformations up to first order in a small parameter $\epsilon \ll 1$. We have for our transformation

$$x'^{\rho} = x^{\rho} + \epsilon^{\rho}(x) + \mathcal{O}(\epsilon^2) \ . \tag{3.1.3}$$

Substituting (3.1.3) into (3.1.2), as well as setting $\Lambda(x) = 1$ for a true invariant metric, we now have for the LHS of our equation

$$\eta_{\rho\sigma}\frac{\partial x'^{\rho}}{\partial x^{\mu}}\frac{\partial x'^{\sigma}}{\partial x^{\nu}} = \eta_{\mu\nu} + \left(\frac{\partial \epsilon_{\mu}}{\partial x^{\nu}} + \frac{\partial \epsilon_{\nu}}{\partial x^{\mu}}\right) + \mathcal{O}(\epsilon^2) \ . \tag{3.1.4}$$

Comparing (3.1.4) with (3.1.1), we now see that we must have the following condition for our infinitesimal transformation (3.1.3) if we wish it to remain conformal:

$$\partial_{\mu}\epsilon_{\nu} + \partial_{\nu}\epsilon_{\mu} = Q(x)\eta_{\mu\nu}, \tag{3.1.5}$$

where $Q(x)$ is some function to be determined and we use the notation $\partial_{\mu} \equiv \frac{\partial}{\partial x^{\mu}}$. If we trace both sides of this condition with $\eta^{\mu\nu}$, we are able to solve for the function $Q(x)$, which we substitute back into (3.1.5) to obtain

$$\partial_{\mu}\epsilon_{\nu} + \partial_{\nu}\epsilon_{\mu} = \frac{2}{d}(\partial \cdot \epsilon)\eta_{\mu\nu}, \tag{3.1.6}$$

where d refers to the number of dimensions and $\partial \cdot \epsilon \equiv \partial^{\mu}\epsilon_{\mu}$.

3.2 2-D CFT

For $d \geq 3$, conformal invariance does not really provide one with much more information than would be obtained from simple scale invariance. In $d = 2$ the conformal algebra becomes infinite dimensional, and thus the array of possibilities for 2-D CFTs is greatly restricted. Since our interest lies in (1+1)-dimensional quantum field theory, we have an interest in pursuing our study of CFTs in two-dimensions. We thus set $d = 2$ and take the Euclidean metric, i.e. $\eta_{\mu\nu} = \delta_{\mu\nu}$. This forces (3.1.6) to become

$$\partial_{\mu}\epsilon_{\nu} + \partial_{\nu}\epsilon_{\mu} = (\partial \cdot \epsilon)\eta_{\mu\nu}. \tag{3.2.1}$$

Since we are working in Euclidean (1+1)-dimensional QFT, we have only two unique choices for the metric components of our equation. By setting the components of the metric to $\mu = 0, \nu = 1$, and separately to $\mu = \nu = 1$, we arrive at the holomorphic Cauchy-Riemann conditions from (3.2.1), i.e.

$$\partial_0\epsilon_1 = -\partial_1\epsilon_0, \qquad \partial_0\epsilon_0 = \partial_1\epsilon_1. \tag{3.2.2}$$

If we were to reverse the sign on the right-hand side of both equations in (3.2.2),

it would give the anti-holomorphic version of the Cauchy-Riemann conditions. The appearance of the Cauchy-Riemann conditions is a clear indication that we should switch to complex coordinates before proceeding, and it becomes convenient to write $\epsilon(z) = \epsilon^1 + i\epsilon^2$ and $\overline{\epsilon}(\overline{z}) = \epsilon^1 - i\epsilon^2$. The 2-D conformal transformations thus coincide with the analytic coordinate transformations

$$z \to f(z) \qquad \overline{z} \to \overline{f}(\overline{z}), \tag{3.2.3}$$

which are the solutions to the holomorphic (anti-holomorphic) Cauchy-Riemann equations and possess an infinite-dimensional local algebra. The conformal group for $d = 2$ is thus the set of all analytic maps. It is infinite-dimensional and corresponds to the coefficients of the Laurent series required to specify functions analytic in some neighborhood. The fact that this set is infinite-dimensional is key to why conformal symmetry in 2-D is so special.

For an infinitesimal conformal transformation in 2-D the function $\epsilon(z)$ is required to be holomorphic in some open set, though one may assume that $\epsilon(z)$ is rather a meromorphic function having isolated singularities outside this open set. To consider an infinitesimal conformal transformation, we take $f(z) = z + \epsilon(z)$. Now $f(z)$ is a holomorphic function, which means that our transformation $f(z) = z + \epsilon(z)$ is holomorphic as well. Similar statements hold true for the anti-holomorphic case $\overline{z}$ Thus our line element is modified and the metric tensor transforms as

$$ds^2 = dz\, d\overline{z} \to \left|\frac{\partial f}{\partial z}\right|^2 dz\, d\overline{z}, \tag{3.2.4}$$

where we may identify the scale factor as $\Lambda = \left|\frac{\partial f}{\partial z}\right|^2$.

3.2.1 The Witt and Virasoro Algebra

For infinitesimal conformal transformations in $d = 2$ the function $\epsilon(z)$ must be holomorphic in some open set. Though, as mentioned previously, it is not unreasonable to expect that the infinitesimal change $\epsilon(z)$ is a meromorphic function that contains isolated singularities outside of this set. We proceed by assuming we have a meromorphic function and take the Laurent expansion around the point $z = 0$, i.e.

$$z' = f(z) = z + \epsilon(z) = z + \sum_{n\in\mathbb{Z}} \epsilon_n \left(-z^{n+1}\right), \tag{3.2.5}$$

$$\overline{z}' = \overline{f}(\overline{z}) = \overline{z} + \overline{\epsilon}(\overline{z}) = \overline{z} + \sum_{n\in\mathbb{Z}} \overline{\epsilon}_n \left(-\overline{z}^{n+1}\right), \tag{3.2.6}$$

where the infinitesimal parameters ϵ_n, $\bar{\epsilon}_n$ are constant. Taking a general n^{th}-term in (3.2.5), we now want to determine the generator corresponding to this transformation. An infinitesimal mapping on a spinless, dimensionless field $\phi(z, \overline{z})$ is given by

$$\delta\phi = \epsilon(z)\partial\phi + \overline{\epsilon}(\overline{z})\overline{\partial}\phi. \tag{3.2.7}$$

We offer up a basis for the transformations as

$$z \to z' = z + \epsilon_n(z), \qquad \overline{z} \to \overline{z}' = \overline{z} + \overline{\epsilon}_n(\overline{z}) \qquad (n \in \mathbb{Z}), \tag{3.2.8}$$

where we have

$$\epsilon_n(z) = -z^{n+1}, \qquad \overline{\epsilon}_n(\overline{z}) = -\overline{z}^{n+1} \qquad (n \in \mathbb{Z}). \tag{3.2.9}$$

From the above, we now may define the corresponding infinitesimal conformal generators for a particular n^{th}-term as

$$l_n \equiv -z^{n+1}\partial_z, \qquad \bar{l}_n \equiv -\overline{z}^{n+1}\partial_{\overline{z}} \qquad (n \in \mathbb{Z}), \tag{3.2.10}$$

where the l's and the $\bar{l}$'s satisfy the algebra

$$[l_m, l_n] = (m-n)l_{m+n}, \qquad [\bar{l}_m, \bar{l}_n] = (m-n)\bar{l}_{m+n}, \qquad [l_m, \bar{l}_n] = 0. \tag{3.2.11}$$

Since $n \in \mathbb{Z}$ for (3.2.10), we have an infinite number of independent conformal transformations. The above algebra is called the *Witt algebra*, the Lie algebra of meromorphic vector fields defined on the Riemann sphere ($\mathbb{C} \cup \infty$) that are holomorphic everywhere except at two fixed points. It is important to note that when considering the quantum case, the Witt algebra (3.2.11) will be modified to include a further term which is proportional to a central charge, i.e. a *central extension.*

The representations in QFT are most often *projective representations*, where the overall magnitudes and phases of the state vectors have no physical significance. For a semi-simple compact Lie group, i.e. a group that admits a Lie algebra that is, or is a sum of, Lie algebras that are non-abelian and contain no nonzero proper ideals, a projective representation is a true representation of its universal covering. For non-semi-simple Lie algebras, we introduce a central extension of the group in order to make our projective representation a true representation.

We have been studying the Witt algebra. As we would like to study the true representations of this conformal group, we take and study a central extension. Doing so takes our Witt algebra and turns it into something closely resembling the projective representations of quantum mechanics and QFT, thereby effectively 'quantizing' the Witt algebra.

A central extension $\tilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathbb{C}$ of a Lie algebra $\mathfrak{g}$ by $\mathbb{C}$ is defined by the following commutation relations:

$$[\tilde{x}, \tilde{y}]_{\mathfrak{g}} = [x, y]_{\mathfrak{g}} + c\, p(x, y) \tag{3.2.12}$$

$$[\tilde{x}, c]_{\mathfrak{g}} = 0 \tag{3.2.13}$$

$$[c, c]_{\mathfrak{g}} = 0 \tag{3.2.14}$$

where $\tilde{x}, \tilde{y} \in \tilde{\mathfrak{g}}$, $x, y \in \mathfrak{g}$, $c \in \mathbb{C}$, and the function $p : \mathfrak{g} \times \mathfrak{g}$ is bilinear. Note that c appearing in these equations is the *central charge*, which is an operator that commutes with all of the other symmetry operators in the given algebra. It is the literal center of the symmetry group, i.e. the subgroup which contains elements that commute with every single element of the original group. Its appearance relates to the 'soft' breaking of conformal symmetry upon the introduction of a macroscopic scale into the system.

We denote the elements of the central extension of the Witt algebra by L_n with $n \in \mathbb{Z}$ and apply (3.2.12), (3.2.13), (3.2.14) to the Witt algebra in (3.2.11). The only non-zero term we obtain is

$$[L_m, L_n] = (m - n)L_{m+n} + c\, p(m, n). \tag{3.2.15}$$

We next conduct analysis to determine the form of $p(m, n)$. To preserve the anti-symmetry of the Lie bracket, $p(m, n)$ in (3.2.15) must satisfy the relation $p(m, n) = -p(n, m)$. Additionally we note that by a suitable redefinition of L_n, one may always arrange for $p(1, -1) = 0$ and $p(n, 0) = 0$. With these two facts and suitable redefinitions, one may push through some calculations involving Lie brackets and a carefully selected Jacobi identity to determine that the only non-vanishing central extensions are of the form $p(n, -n)$ for $|n| \geq 2$. The calculation of one more particular Jacobi identity leads to the recursion relation

$$p(n, -n) = \frac{n+1}{n-2}\, p(n-1, -n+1) \tag{3.2.16}$$

$$= \frac{1}{2}\binom{n+1}{3} \tag{3.2.17}$$

$$= \frac{1}{12}(n+1)n(n-1), \tag{3.2.18}$$

where we set $n = 2$ to normalize $p(n - n)$ as $p(2, -2) = \frac{1}{2}$. The normalization is specifically selected so that the central charge has a specific value for the case of the free boson, i.e. $c = 1$. Thus our central extension of the Witt algebra is given by

$$[L_m, L_n] = (m - n)L_{m+n} + \frac{c}{12}\left(m^3 - m\right)\delta_{m+n,0}, \tag{3.2.19}$$

and is known as the *Virasoro algebra*. Note that there does exist a corresponding

formula for $\bar{L}$ and $\bar{c}$.

3.2.2 Global Conformal Transformations

Now diverting our attention to the Witt algebra generated by $\{l_n\}$, we notice that these generators are not defined everywhere on the Euclidean plane $\mathbb{R}^2 \simeq \mathbb{C}$. There exists an ambiguity at $z = 0$, and thus we must now work on the Riemann sphere $S^2 \simeq \mathbb{C} \cup \infty$, which is the conformal compactification of $\mathbb{R}^2$. Not all of the generators are well-defined even on the Riemann sphere S^2. For the point $z = 0$, the generator

$$l_n = -z^{n+1}\partial_z \tag{3.2.20}$$

is non-singular at $z = 0$ when $n \geq -1$. The other troublesome point is $z = \infty$, which is a bonafide part of the Riemann sphere. We perform a change of variables

$$z = -\frac{1}{w} \tag{3.2.21}$$

and take $w \to 0$. We arrive at a generator of the form

$$l_n = -\left(-\frac{1}{w}\right)^{n-1}\partial_w, \tag{3.2.22}$$

which is non-singular at $w = 0$ for $n \leq 1$. Note that we used the notion that $\partial_z = (-w)^2\partial_w$. Thus we find that globally defined conformal transformations on the Riemann sphere are generated via the set of operators $\{l_{-1}, l_0, l_1\}$.

3.2.3 The Conformal Group

Now we move on from defining the operators generating the global conformal transformations to determining the conformal group. The operator l_{-1} generates translations $z \mapsto z + b$. For the operator l_0, we note that $l_0 = -z\partial_z$. This indicates that l_0 generates transformations $z \mapsto az$ with $a \in \mathbb{C}$. We now perform a change of variables such that $z = re^{i\phi}$ to obtain

$$l_0 = -\frac{1}{2}r\partial_r + \frac{i}{2}\partial_\phi, \qquad \bar{l}_0 = -\frac{1}{2}r\partial_r - \frac{i}{2}\partial_\phi. \tag{3.2.23}$$

Note we made use of $r = \sqrt{z\bar{z}}$ and $\phi = \frac{1}{2i}\log\frac{z}{\bar{z}}$ in the transformation $\partial_z = \partial_z r + \partial_z \phi$. We can then form linear combinations from the equations in (3.2.23) to obtain

$$l_0 + \bar{l}_0 = -r\partial_r \tag{3.2.24}$$

and

$$i(l_0 - \bar{l}_0) = -\partial_\phi, \tag{3.2.25}$$

where (3.2.24) is the generator for 2-D *dilations* and (3.2.25) is the generator of *rotations*. We finally turn to l_{+1}, which corresponds to *special conformal transformations*. These are translations for the variable $w = -\frac{1}{z}$. The operators $\{l_{-1}, l_0, l_{+1}\}$ generate transformations of the form

$$f(z) = \frac{az + b}{cz + d}, \tag{3.2.26}$$

and the determinant $ad - bc$ must be non-zero. Equation (3.2.26) is known as a *Möbius transformation*. We can then scale the constants a, b, c, and d to ensure that $ad - bc = 1$. Also note that equation (3.2.26) is invariant under $(a, b, c, d) \longmapsto (-a, -b, -c, -d)$. Taking equation (3.2.26) together with the restrictions given above allows us to infer that the conformal group of the Riemann sphere $S^2 = \mathbb{C} \cup \infty$ is the Möbius group $SL(2, \mathbb{C})/\mathbb{Z}_2$.

3.3 Energy-Momentum Tensor

We next move to a brief introduction to the energy-momentum tensor, as it makes an appearance when we will map our CFT to the cylinder. It its own right, the energy-momentum tensor is of importance in CFT for the simple fact that the algebra of infinitesimal conformal transformations is infinite-dimensional and highly constrains the CFT. This allows us to study a field theory without even knowing the explicit form of its Lagrangian action, as the only information required is the specific behavior under conformal transformations - information encoded by the energy-momentum tensor.

3.3.1 Conformal Invariance and its Implications

Noether's theorem for a field theory states that for every continuous symmetry present, there exists a conserved current j_μ where

$$\partial^\mu j_\mu = 0. \tag{3.3.1}$$

We concern ourselves with theories exhibiting conformal symmetry, i.e. $x^\mu \to x^\mu + \epsilon^\mu(x)$, and thus obtain a conserved current

$$j_\mu = T_{\mu\nu}\epsilon^\nu, \tag{3.3.2}$$

where $T_{\mu\nu}$ is the *energy-momentum tensor* and it is symmetric in its coordinates. This means that we have an infinite number of conserved charges in the theory - a classic sign of an integrable theory. Let us set $\epsilon^\mu =$ constant and make use of the

fact that our current is conserved, equation (3.3.1), to obtain

$$\partial^\mu j_\mu = \partial^\mu \left(T_{\mu\nu}\epsilon^\nu\right) = \left(\partial^\mu T_{\mu\nu}\right)\epsilon^\nu = 0, \tag{3.3.3}$$

where in the third equality we used the normal product rule noting that $\partial\epsilon^\mu = 0$. For (3.3.3) to be true, we conclude that

$$\partial^\mu T_{\mu\nu} = 0. \tag{3.3.4}$$

Now armed with (3.3.4) we seek a relation for more general transformations, where this time we allow ϵ^μ to vary. We obtain

$$\partial^\mu j_\mu = \partial^\mu \left(T_{\mu\nu}\epsilon^\nu\right) + T_{\mu\nu}\left(\partial^\mu\epsilon^\nu\right) \tag{3.3.5}$$

$$= 0 + \frac{1}{2}T_{\mu\nu}\left(\partial^\mu\epsilon^\nu + \partial^\nu\epsilon^\mu\right) \tag{3.3.6}$$

$$= \frac{1}{2}T_{\mu\nu}\eta^{\mu\nu}\left(\partial\cdot\epsilon\right)\frac{2}{d} \tag{3.3.7}$$

$$= \frac{1}{d}T^\mu_\mu\left(\partial\cdot\epsilon\right) = 0. \tag{3.3.8}$$

In (3.3.5) we used (3.3.4) as well as the fact that $T_{\mu\nu}$ is symmetric in the second equality. We obtain the third equality by applying (3.1.6). Since we imposed no condition on arbitary conformal transformations $\epsilon(z)$ when we derived (3.3.8), the result is general and implies that the energy-momentum tensor is traceless, i.e.

$$T^\mu_\mu = 0. \tag{3.3.9}$$

We now restrict our attention to 2-D Euclidean theories and investigate the consequences when $T_{\mu\nu}$ is traceless. We induce a change of coordinates

$$T_{\mu\nu} = \frac{\partial x^\alpha}{\partial x^\mu}\frac{\partial x^\beta}{\partial x^\nu}T_{\alpha\beta} \tag{3.3.10}$$

for the time-coordinate $x^0 = \frac{1}{2}(z+\overline{z})$ and $x^1 = \frac{1}{2i}(z-\overline{z})$ for the spatial. Using (3.3.10) and the change of coordinates above the energy-momentum tensor components in complex coordinates are

$$T_{zz} = \frac{1}{4}\left(T_{00} - 2iT_{10} - T_{11}\right), \tag{3.3.11}$$

$$T_{\overline{z}\overline{z}} = \frac{1}{4}\left(T_{00} + 2iT_{10} - T_{11}\right), \tag{3.3.12}$$

$$T_{z\overline{z}} = T_{\overline{z}z} = \frac{1}{4}\left(T_{00} + T_{11}\right) = \frac{1}{4}T^\mu_\mu = 0, \tag{3.3.13}$$

where for $T_{z\overline{z}}$ we made use of $n_{\mu\nu} = \mathrm{diag}(+1,+1)$ and $T^\mu_\mu = 0$. We also use the fact that $T_{\mu\nu} = T_{\nu\mu}$ in the above. Applying the traceless condition to the third equality

in (3.3.13) yields

$$T_{z\overline{z}} = T_{\overline{z}z} = 0. \tag{3.3.14}$$

Applying the traceless condition to (3.3.11) and (3.3.12) we obtain

$$T_{zz} = \frac{1}{2}\left(T_{00} - iT_{10}\right), \qquad T_{\overline{z}\overline{z}} = \frac{1}{2}\left(T_{00} + iT_{10}\right), \tag{3.3.15}$$

since $\mathrm{tr}\,(T^{\nu}_{\mu}) = T_{00} + T_{11} = 0 \Rightarrow T_{00} = -T_{11}$. Applying the condition for translation invariance (3.3.4) to (3.3.15) we obtain

$$\partial_0 T_{00} + \partial_1 T_{10} = 0, \qquad \partial_0 T_{01} + \partial_1 T_{11} = 0. \tag{3.3.16}$$

For a little more detail, note that here we are applying the condition $g^{\alpha\nu}\partial_\alpha T_{\mu\nu} = 0$ since we lower the contravariant index on the conservation equation (3.3.4). Since we are working with the Euclidean metric, we obtain zero whenever $\alpha \neq \nu$.

Applying the above to the calculation of $\partial_{\overline{z}} T_{zz}$ we obtain

$$\partial_{\overline{z}} T_{zz} = \frac{1}{4}\left(\partial_0 + i\partial_1\right)\left(T_{00} - iT_{10}\right) = \frac{1}{4}\left(\partial_0 T_{00} - i\partial_0 T_{10} + i\partial_1 T_{00} + \partial_1 T_{10}\right) = 0, \tag{3.3.17}$$

where $T_{00} = -T_{11}$ and $T_{10} = T_{01}$ from the traceless condition. A similar consideration also shows that $\partial_z T_{\overline{z}\overline{z}} = 0$. Taken together, we have just shown that the two non-vanishing components of the energy-momentum tensor are a *chiral* and an *anti-chiral* field

$$T_{zz}(z,\overline{z}) \equiv T(z), \qquad T_{\overline{z}\overline{z}}(z,\overline{z}) \equiv \overline{T}(\overline{z}). \tag{3.3.18}$$

3.3.2 Radial Quantization

As we continue to work with Euclidean 2-D CFT, we make the choice that our time coordinate is denoted x^0 and our space coordinate denoted by x^1. Be aware that we may obtain a Lorentzian signature by simply applying a Wick rotation to the Euclidean coordinates, i.e. $x^0 \to ix^0$.

We endeavor to compactify the Euclidean space direction by mapping our infinite plane onto a cylinder. That is, we will map the Euclidean space coordinate x^1 onto a circle of radius $R = 1$ with period $2\pi i$. Note that the value of R can be arbitrary, and that we make our unity selection for simplicity. This mapping ensures that the CFT we obtain will be defined on a cylinder of infinite length, as the time coordinate x^0 extends to infinity in either direction along its length. The new complex coordinates w for this setup are thus defined as

$$w = x^0 + ix^1 \quad \text{and} \quad w \sim w + 2\pi i, \tag{3.3.19}$$

where we included the periodic identification.

A natural question arises immediately - what would motivate one to perform such a mapping, i.e. to place their CFT onto a cylinder? To answer this, let us proceed to map our infinitely-long cylinder onto the complex plane in order to utilize complex analysis for analyzing our CFT. We make the coordinate change of variables

$$z = e^w = e^{x^0+ix^1}. \tag{3.3.20}$$

The CFT's we are concerned about are originally defined on the infinite plane. Due to conformal symmetry, the CFT on the infinite plane is exactly equivalent to the CFT on the cylinder. Points that obey $x^1 \sim x^1 + 2\pi$ are naturally identified since $z = e^{x^0+i(x^1+2\pi)} = e^{x^0+ix^1}$ which is shown by the simple application of the properties of the exponential function. This means that the conjecture that we can treat the spatial coordinate x^1 as periodic is indeed valid, and that mapping the CFT onto a cylinder is indeed reasonable.

The mapping (3.3.20) takes time translations $x^0 \to x^0 + a$ and maps them to complex dilations $z \to e^a z$, while the space translations $x^1 \to x^1 + b$ are mapped to rotations $z \to e^{ib} z$. Note that a and b are constants.

In quantum mechanics the generator of time translations is the Hamiltonian operator H, while the generator of space translations is the momentum operator P. From the above argument that the time translations are mapped to complex dilations, H presently corresponds to the dilation operator. By a similar argument P corresponds to rotations. Using (3.2.24), (3.2.25), and the fact that the central extension for L_0 and $\overline{L}_0$ vanishes, we find that

$$H = L_0 + \overline{L}_0, \qquad P = i(L_0 - \overline{L}_0). \tag{3.3.21}$$

3.4 The Operator-Product Expansion

We proceed to further study the energy-momentum tensor and now will introduce operator formalism for 2-D CFTs.

3.4.1 Conserved Charges

Due to the fact that we have a conserved current $j_\mu = T_{\mu\nu}\epsilon^\nu$ associated with the conformal symmetry, there exists a conserved charge expressed as

$$Q = \int j_0 \, dx^1. \tag{3.4.1}$$

This conserved charge is the generator of symmetry transformations for an operator A, written as

$$\delta A = [Q, A], \tag{3.4.2}$$

where we evaluate (3.4.2) at equal times. In radial quantization this implies that $|z|=$ constant. This changes the integral over space in (3.4.1) into a contour integral, which is counterclockwise by convention. Thus (3.4.1) becomes

$$Q = \frac{1}{2\pi i}\oint_{\mathcal{C}} \left(T(z)\epsilon(z)dz + \overline{T}(\overline{z})\overline{\epsilon}(\overline{z})d\overline{z}\right). \tag{3.4.3}$$

We now can determine the infinitesimal transformation of a field $\phi(z,\overline{z})$ generated by Q. To proceed, we input (3.4.3) and compute $\delta\phi = [Q,\phi]$ to obtain

$$\delta_{\epsilon,\bar{\epsilon}}\phi(w,\overline{w}) = \frac{1}{2\pi i}\oint_{\mathcal{C}}[T(z)\epsilon(z),\phi(w,\overline{w})]dz + \frac{1}{2\pi i}\oint_{\mathcal{C}}[\overline{T}(\overline{z})\overline{\epsilon}(\overline{z}),\phi(w,\overline{w})]dz. \tag{3.4.4}$$

3.4.2 Radial Ordering

In (3.4.4) there is some ambiguity, as we have to determine whether w and $\overline{w}$ lie inside or outside the contour $\mathcal{C}$. In QFT, correlation functions are exclusively defined as a time-ordered product. Since we have changed our coordinate system above from the plane to the cylinder for a CFT, this time-ordering becomes a radial ordering, and we define the radial ordering of two operators as

$$R(A(z)B(w)) \equiv \begin{cases} A(z)B(w) \ \text{ for } |z|>|w| \\ B(w)A(z) \ \text{ for } |w|>|z|. \end{cases} \tag{3.4.5}$$

Due to (3.4.5), we now must solve for the integrals of the commutators in (3.4.4) as

$$\oint [A(z),B(w)]dz = \oint_{|z|>|w|} A(z)B(w)dz - \oint_{|z|<|w|} B(w)A(z)dz \tag{3.4.6}$$

$$= \oint_{\mathcal{C}(w)} R\left(A(z)B(w)\right)dz. \tag{3.4.7}$$

This allows (3.4.4) to become

$$\delta_{\epsilon,\bar{\epsilon}}(w,\overline{w}) = \frac{1}{2\pi i}\oint_{\mathcal{C}} R\left(T(z)\phi(w,\overline{w})\right)dz + \text{anti-chiral term} \tag{3.4.8}$$

Using the fact that the infinitesimal conformal transformation for a primary field transforms as

$$\delta_{\epsilon,\bar{\epsilon}}(w,\overline{w}) = h(\partial_w\epsilon(w))\phi(w,\overline{w}) + \epsilon(w)(\partial_w\phi(w,\overline{w})) + \text{anti-chiral term}, \tag{3.4.9}$$

we can then deduce a relation for the radial ordering of the energy-momentum tensor and a primary field. Now *primary*, *quasi-primary*, and *descendent* fields are defined

by how they transform under an infinitesimal or global conformal transformation. For further detailed information, please consult the references. For a bi-holomorphic field $\phi(w,\overline{w})$ we obtain

$$R\left(T(z)\phi(w,\overline{w})\right) = \frac{h}{(z-w)^2}\phi(w,\overline{w}) + \frac{1}{z-w}\partial_w\phi(w,\overline{w}) + ..., \tag{3.4.10}$$

where the ellipsis represents non-singular terms. We refer to (3.4.10) as an *operator-product expansion* (OPE). The OPE defines the algebraic product structure on the space of quantum fields. Now we will always assume radial order for a product of fields, so $R(A(z)B(w)) \Rightarrow A(z)B(w)$. Using (3.4.10), we can define a *primary field* with conformal dimensions $(h,\overline{h})$ if the OPE between the energy-momentum tensor and $\phi(z,\overline{z})$ appears as

$$T(z)\phi(w,\overline{w}) = \frac{h}{(z-w)^2}\phi(w,\overline{w}) + \frac{1}{z-w}\partial_w\phi(w,\overline{w}) + ..., \tag{3.4.11}$$

$$\overline{T}(\overline{z})\phi(w,\overline{w}) = \frac{\overline{h}}{(\overline{z}-\overline{w})^2}\phi(w,\overline{w}) + \frac{1}{\overline{z}-\overline{w}}\partial_w\phi(w,\overline{w}) + ..., \tag{3.4.12}$$

where again the ellipsis denote non-singular terms. With this definition, we now compute the OPE of the chiral energy-momentum tensor with itself. To do this we first utilize Wick's theorem to change the implicitly written radial ordering to normal ordering, which now means that all fields commute with each other since they are now under normal ordering. Next we will Taylor expand around the point w, noting that to obtain the higher order terms we should Talyor expand term with a single contraction. This yields

$$T(z)T(w) = \frac{1}{\alpha'^2} : \partial X(z)\partial X(z) :: \partial X(z)\partial X(z) : \tag{3.4.13}$$

$$= \frac{2}{\alpha'^2}\left(\frac{-\alpha'}{2}\frac{1}{(z-w)^2}\right)^2 + \frac{4}{\alpha'^2}\left(\frac{-\alpha'}{2}\frac{\partial X(w)\partial X(w)}{(z-w)^2}\right) + ... \tag{3.4.14}$$

$$= \frac{1/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{2\partial^2 X(w)\partial X(w)}{(z-w)} + ... \tag{3.4.15}$$

which yields the final form

$$T(z)T(w) = \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial_w T(w)}{z-w} + ..., \tag{3.4.16}$$

where c is the central charge and $|z|> |w|$.

3.5 CFT Mapped Onto the Cylinder

To conclude this chapter and tie in CFT with our research goals, we map the complex plane onto a single strip with a width L via

$$z \to w = \frac{2\pi}{L} \log z \tag{3.5.1}$$

with $\{z \in \mathbb{C} | 0 \leq \operatorname{Im} z < L\}$. The Hamiltonian is given via

$$H_{cyl} = \frac{2\pi}{L} \left(L_0 + \overline{L}_0 - \frac{c}{12} \right) \tag{3.5.2}$$

For simplicity we identify the length with a single period $L = 2\pi$ and proceed by calculating the energy-momentum tensor on a cylinder T_{cyl}. Using the OPE between the energy-momentum tensor and itself allows us to determine the infinitesimal transformation expression

$$\delta_\epsilon T(z) = \frac{1}{2\pi i} \oint_{\mathcal{C}(z)} dw\, \epsilon(w) T(w) T(z) \tag{3.5.3}$$

$$= \frac{1}{2\pi i} \oint_{\mathcal{C}(z)} dw\, \epsilon(w) \left(\frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial_w T(w)}{z-w} + ... \right) \tag{3.5.4}$$

$$= \frac{c}{12} \partial_z^3 \epsilon(z) + 2T(z) \partial_z \epsilon(z) + \epsilon(z) \partial_z T(z), \tag{3.5.5}$$

which can be integrated to give

$$T(z) = (df)^2 T(f(z)) + \frac{c}{12} S(f(z), w). \tag{3.5.6}$$

where under $z \to f(z)$ the *Schwartzian derivative* $S(f, z)$ is given by

$$S(z, w) = \frac{\partial_z f\, \partial_z^3 f - \frac{3}{2} (\partial_z^2 f)^2}{(\partial_z f)^2}. \tag{3.5.7}$$

By comparing (3.4.16) with (3.4.11), we see that for non-vanishing central charges, $T(z)$ is not a primary field. Under conformal transformations $z = e^w$, the energy-momentum tensor behaves as

$$T(w) = \left(\frac{dz}{dw} \right)^2 T(z) + \frac{c}{12} S(z, w), \tag{3.5.8}$$

where $S(z, w)$ becomes

$$S(z, w) = \frac{\partial_w z\, \partial_w^3 z - \frac{3}{2} (\partial_w^2 z)^2}{(\partial_w z)^2}. \tag{3.5.9}$$

For $z = e^w$ we obtain $S(z, w) = -\frac{1}{2}$ and $\frac{dz}{dw} = z$, which yields

$$T_{cyl}(w) = z^2 T(z) - \frac{c}{24}, \tag{3.5.10}$$

If we substitute the mode expansion $T(z) = \sum L_n z^{-n-2}$ into (3.5.8), where $L_n = \frac{1}{2\pi i} \oint dz\, z^{n+1}\, T(z)$ we obtain the following expression

$$T_{cyl}(w) = \sum_{n\in\mathbb{Z}} L_n z^{-n} - \frac{c}{24} = \sum_{n\in\mathbb{Z}} \left(L_n - \frac{c}{24}\delta_{n0} \right) e^{-nw}. \tag{3.5.11}$$

From this we can obtain the translation generator $(L_0)_{cyl}$ in terms of the dilation generator L_0 on the plane:

$$(L_0)_{cyl} = L_0 - \frac{c}{24}. \tag{3.5.12}$$

In a QFT one can customarily enact two equivalent actions, either shift the energy of the vacuum by a constant or change the normalization of a functional integral. In a CFT, scale and rotational invariance of the $SL(2,\mathcal{C})$ invariant vacuum on the plane fixes L_0 and $\overline{L}_0$ to have an eigenvalue with value zero on the vacuum, which fixes the zero of the energy.

A similar argument and solution for the anti-holomorphic term $\overline{L}_0$ allows us to input such arguments into (3.3.21) to obtain the result

$$H_{cyl} = (L_0)_{cyl} + (\overline{L}_0)_{cyl} \tag{3.5.13}$$

$$= L_0 + \overline{L}_0 - \frac{c}{12}. \tag{3.5.14}$$

Our result (3.5.14) will be of special importance when we later analytically continue the TBA equation and wish to ensure that our results are correct.

Chapter 4

S-Matrices

The 2-particle exact S-matrix is the only requisite quantity for the computation of the ground-state energy for a given model via the TBA equation. In the last chapter we described CFT, which in turn describes short distance scales. Since they are scale invariant, CFTs have an infinite correlation length. When the distances become larger, provided our theory is massive, we now are working with a finite correlation length. Thus instead of a CFT, we now have a theory governed by the S-matrix. Due to its central importance in this method, we devote this chapter to introducing the S-matrix and how it is constrained by integrability. This section chiefly follows the reviews by Bombardelli [10], Dorey [16], and Eden and company [21].

4.1 Asymptotic States and the S-Matrix

It is possible in quantum theory to define the time evolution via a unitary operator $U(t, t_0)$, which conserves probability and produces the time-evolved state $|\Phi(t)\rangle$ by acting on an initial state $|\Phi(t_0)\rangle$ as shown below:

$$|\Phi(t)\rangle = U(t, t_0)\,|\Phi(t_0)\rangle\,. \tag{4.1.1}$$

It turns out that we only need to know the values of t_0 and t in the operator $U(t, t_0)$ at its asymptotics, i.e. when $t_0 \to -\infty$ and $t \to +\infty$ in order to analyze a scattering process. Values of t at any point other than its asymptotics are unnecessary. If we assume that interactions between any of the particles only occur in a narrow region of spacetime, then these particles can be considered free at distances far way from the interaction region. Thus we are motivated to formally define these quantum states of free excitations.

To do so, we begin by introducing the *asymptotic states* given by

$$|p_1, p_2, ..., p_n\rangle^{in/out}_{a_1 a_2 ... a_n}, \tag{4.1.2}$$

where n is the number of particles in the state, p_i are their momenta, and a_i are the particle flavours. These asymptotic states describe collections of wave packets with approximate positions at given times, i.e. n free particles at time $t \to -\infty$ for the *in states* and at time $t \to +\infty$ for the *out states*. The order of momenta is taken to be $p_1 > p_2 > ... > p_n$, and any intermediate state can equivalently be expanded in the *in* or *out* bases. The S-matrix is then defined as the linear operator that maps the final asymptotic states into initial asymptotic states or vice versa depending on the convention adopted, the chosen convention being related to an inversion of this linear operator. The S-matrix is given via the relation

$$|...\rangle_{in} = S\,|...\rangle_{out} \tag{4.1.3}$$

where S is the time-evolution operator from $t \to -\infty$ to $t \to +\infty$, and based on (4.1.1), can be represented as a unitary matrix:

$$S = \lim_{\substack{t_0 \to -\infty \\ t \to +\infty}} U(t, t_0). \tag{4.1.4}$$

In the interaction picture, where both operators and states depend on time, we have a Hamiltonian given by

$$H = H_0 + H_{int}, \tag{4.1.5}$$

where H_0 is the Hamiltonian of the free system and H_{int} represents the interacting part (we're in interaction picture). The S-matrix can then be expressed as follows:

$$S = \mathscr{T} \exp\left[-i \int_{-\infty}^{+\infty} dt\, H_{int}(t)\right], \tag{4.1.6}$$

where $\mathscr{T}$ refers to the time-ordering operator, which is applied to the series expansion of the argument in the exponential in (4.1.6).

4.1.1 Assumptions and General Properties

We will now discuss some general assumptions, which are motivated by physical properties fulfilled by usual QFTs. As mentioned before, interactions among particles are assumed to only occur at *short range*. We also assume that the *principle of asymptotic completeness* is valid. This says that asymptotic states form a complete basis for the initial and final states. Therefore, as mentioned before, any *incoming* state can be expanded in the basis of *outgoing* states and vice versa through the

linear time evolution operator S given in equation (4.1.3).

We also assume *probability conservation*. We let $|\phi\rangle$ represent any state. Probability conservation implies

$$1 = \sum_m |\langle m| S |\phi\rangle|^2, \tag{4.1.7}$$

where we expand $|\phi\rangle$ in an orthogonal basis of states $|\phi\rangle = \sum_n a_n |n\rangle$. This means that $|m\rangle$ and $|n\rangle$ are orthogonal, complete basis vectors generating the Hilbert space of the asymptotic states. Using (4.1.7) we have,

$$\begin{aligned} 1 = \sum_m |\langle m| S |\phi\rangle|^2 &= \sum_m \langle\phi| S^\dagger |m\rangle \langle m| S |\phi\rangle \\ &= \langle\phi| S^\dagger S |\phi\rangle = \sum_{n,m} a_n^* a_m \langle n| S^\dagger S |m\rangle , \end{aligned} \tag{4.1.8}$$

where in the third equality the *resolution of the identity* was utilized, i.e. $1 = \sum_m |m\rangle \langle m|$. Therefore, as a direct consequence for probability conservation (4.1.7), we find that S must be *unitary*, i.e. $S^\dagger S = 1$.

Since we are concerned with relativistic theories, we must ensure *Lorentz invariance* is satisfied. Naturally this will have consequences for the possible forms of the S-matrix. Given a generic Lorentz transformation $L |m\rangle = |m'\rangle$, where the prime denotes the transformed coordinates, any requirement of invariance under such a transformation at the level of the S-matrix is equivalent to requiring that

$$\langle m'| S |n'\rangle = \langle m| S |n\rangle . \tag{4.1.9}$$

Now let us concern ourselves with the 2-to-2 particle scattering process, where the incoming particles have momenta p_1, p_2 and the outgoing particles have momenta p_3, p_4. In a relativistic (1+1)-dimensional theory, the energies and momenta of the scattered particles are represented by *Mandelstam variables*, a set of relativistic invariants given by

$$s = (p_1 + p_2)^2, \quad t = (p_1 - p_3)^2, \quad u = (p_1 - p_4)^2, \tag{4.1.10}$$

where $p_i = (p_i^{(0)} = E_i, p_i^{(1)})$. Due to conservation of momenta we have $p_1 + p_2 = p_3 + p_4$. We also require the *on-shell* condition $p_i^2 = m_i^2$ to hold, as we are working with real (not virtual) particles. Taking these together yields a condition, namely $s + t + u = \sum_{i=1}^4 m_i^2$. This relation implies the scattering amplitude is determined solely by two Mandelstam variables, since they are not independent of each other. In fact, in $d = 2$ only one Mandelstam variable is independent.

In two-dimensions one can parametrize energy and momentum in terms of the *rapidity* θ, i.e. the hyperbolic angle which differentiates two frames of reference in relative motion, via $p_i = m_i \sinh(\theta_i)$ and $E_i = m_i \cosh(\theta_i)$. One is then able to

re-formulate the Mandelstam variables in terms of the rapidity as

$$s = m_1^2 + m_2^2 + 2m_1 m_2 \cosh(\theta_{12}), \tag{4.1.11}$$

$$t = m_1^2 + m_3^2 - 2m_1 m_3 \cosh(\theta_{13}), \tag{4.1.12}$$

$$u = m_1^2 + m_4^2 - 2m_1 m_4 \cosh(\theta_{14}), \tag{4.1.13}$$

where $\theta_{ij} = \theta_i - \theta_j$. Therefore, as a consequence of Lorentz invariance, the scattering phases depend on the difference of the rapidities alone.

We now assume *macrocausality* holds, which implies that the outgoing particles propagate only after the incoming particles have scattered when viewed on large time scales. We note that this property can be violated on microscopic time scales, but won't address this further. Our final assumption is that the S-matrices are *analytic*. This simply requires that the S-matrices are analytic functions in the θ-plane which have a minimal number of singularities dictated by specific physical processes of the model. Armed with the consequences of all of these assumptions, we are ready to proceed to consider the form of an S-matrix in an integrable theory.

4.2 Consequences of Integrability

In QFT, the common perception of a system being 'integrable' lies in a system containing an infinite number of independent, conserved charges Q_s that are in involution, i.e. they commute. This implies that all of the charges can be diagonalized simultaneously as

$$Q_s \left|p\right\rangle_a = q_s^{(a)}(p) \left|p\right\rangle_a . \tag{4.2.1}$$

If these charges can be expressed as integrals of local densities, i.e the charges are *local*, then their action on asymptotic states is approximately additive. This means that the action of each charge will act separately on each particle in the state:

$$Q_s \left|p_1, ..., p_n\right\rangle_{a_1,...,a_n} = (q_s^{(a_1)}(p_1) + ... + q_s^{(a_n)}(p_n)) \left|p_1, ..., p_n\right\rangle_{a_1,...,a_n} . \tag{4.2.2}$$

In short, this is what it generally means for a QFT to be integrable, and this has profound consquences on the form of the S-matrix.

In relativistic scattering theories, relatively few symmetries are possible. The Coleman-Mandula theorem, satisfying the assumptions made previously for our scattering theory, demonstrates that in $d > 2$ dimensions the S-matrix is trivial ($S = 1$) given the existence of at least one conserved charge transforming as a tensor of second or higher rank. In $d = 2$ dimensions, the S-matrix does not trivialize. To

find it will require a bit of work. Despite this fact, integrability still significantly limits the possible form of the S-matrix and implies the following properties:

- conservation of particle number,
- conservation of momentum,
- factorizability.

4.2.1 Conservation of Particle Number and Momentum

Recall we assumed that the charges Q_s were conserved as part of our general notion of integrability. Then an initial eigenstate of Q_s will time-evolve into a superposition of states that all share the same eigenvalue,

$$\sum_{i\in in}^{n} q_s^{(a_i)}(p_i) = \sum_{j\in out}^{m} q_s^{(b_j)}(p'_j) \ . \tag{4.2.3}$$

Given we have an infinite set of mutually commuting charges Q_s, we must have an infinite set of constraints given by equation (4.2.3) that imply $n = m$, $p_i = p'_j$, and thus $q_s^{(a_i)} = q_s^{(b_i)} \ \forall \ i$. The consequences from this simple argument are two-fold. The number of particles before the scattering event is equal to the number after, i.e. we have no particle production. We also have that the initial and final sets of momenta are conserved as well! Thus, integrability forces the scattering of 2-particles to be *elastic*.

4.2.2 Factorizability

Next we wish to show that integrability forces the S-matrix to be *factorizable*, i.e. one can factorize an n-particle scattering into a product of two-particle events

$$S_n(p_1, ..., p_n) = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} S_2(p_i, p_j). \tag{4.2.4}$$

To show factorizability we proceed by considering an n-particle configuration space $\mathbb{R}^n$. Let us define a permutation σ of ordered coordinates as $x_{\sigma_1} < x_{\sigma_2} < ... < x_{\sigma_n}$ and a permutation of momenta as $p_{\sigma_1} > p_{\sigma_2} > ... > p_{\sigma_n}$. We further assume that the particle interaction range $\mathcal{D}$ is short. From this setup we are able to then consider that we have $n!$ disconnected domains where the particles are well-separated from each other, i.e. $|x_{\sigma_{i+1}} - x_{\sigma_i}| \gg \mathcal{D}$. Thus we may consider the particles essentially free. Due to the conservation of particle number and momentum, the wavefunction

that describes the particles in any one domain is a superposition of a finite number of n-particle plane waves, i.e.

$$\psi_\sigma(x_1,...,x_n) = \sum_{\sigma'} c(\sigma,\sigma') e^{i\left(p_{\sigma'_1} x_{\sigma_1} + ... + p_{\sigma'_n} x_{\sigma_n}\right)}. \tag{4.2.5}$$

Note that σ and σ' are permutations of $p_1,...,p_n$, which are permissible by the conservation of particle number and momentum. The particle numbers are fixed, and changes in the momenta of the particles can only occur via an interaction - scattering.

Recall that we speculated that there exists an asymptotic region, where the particles freely move, for any permutation of the particles in the configuration space. We can then treat the scattering process as if it were a plane wave propagating from one asymptotic region to the other via traveling through the boundary interaction regions. The propagation path then can always be selected in a manner to ensure that it travels through *only* interaction regions where two particles alone are close enough to interact. How is this so?

The key is to realise that we're working with a local, causal quantum field theory and consider the effect of the conserved charges on localised wavepackets. We begin with a single-particle state, which has a position space wavefunction given by

$$\psi(x) \propto \int_{-\infty}^{\infty} dp e^{-a^2(p-p_1)^2} e^{ip(x-x_1)}. \tag{4.2.6}$$

Such a wavefunction describes a particle with a spatial momentum which is approximately p_1 and a position which is approximately x_1. Now let us act on the wavefunction with an operator that yields a momentum-dependent phase factor $e^{-i\phi(p)}$. Now our wavefunction becomes

$$\psi(x) \propto \int_{-\infty}^{\infty} dp e^{-a^2(p-p_1)^2} e^{ip(x-x_1)} e^{-i\phi(p)}. \tag{4.2.7}$$

Most of the integral comes from $p \sim p_1$, and $\phi(p)$ can be expanded in powers of $(p-p_1)$ to find $\tilde{p}_1$ and $\tilde{x}_1$. Thus the revised values of the momentum and position are

$$\tilde{p}_1 = p_1 \quad , \quad \tilde{x}_1 = x_1 + \phi'(p_1). \tag{4.2.8}$$

For a multiparticle state, a product of one-particle wavefunctions will be a good approximation when the particles are well separated. On such a state $|p_a p_b ...\rangle$, the action is to shift the position of a particle a by $\phi'(p_a)$, shift the position of particle b by $\phi'(p_b)$, and so on and so fourth.

Now let us assume conservation of operators P_s which act on one-particle states and well-separated multiparticle states as $(P_1)^s$, with P_1 being the spatial part of the two-momentum operator. Now if we act with $e^{-i\alpha P_s}$, the phase factor becomes

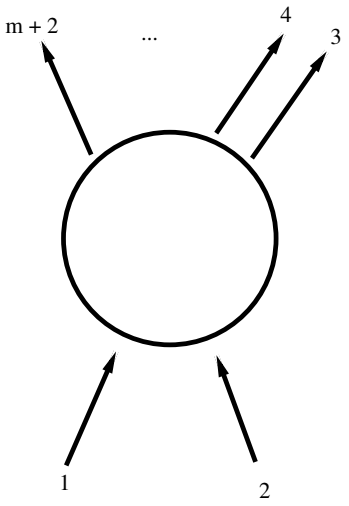


Figure 4.1: A $2 \to m$ process.

$\phi_s(p) = \alpha p^s$, so a particle with momentum p_a will have its position shifted by an amount $s\alpha p_a^{s-1}$. The case where $s = 1$, which is just the momentum itself, simply translates every particle by the same amount α. For $s > 1$, however, particles with different momentum are moved by different amounts. This very fact is critical to our following argument.

Let us initially consider a $2 \to m$ scattering process, which is shown and labelled in Fig. 4.1. For the amplitude to be non-vanishing, the time t_{12} when the first two particles collide must precede the time t_{23} when the trajectory of particle 2, which is the slower incoming particle, intersects the trajectory of particle 3, which is the fastest outgoing particle. Thus we have

$$t_{12} \leq t_{23}. \tag{4.2.9}$$

Note that nothing will occur until the wavepackets of particles 1 and 2 overlap. After this scattering event, it suffices to follow the path of the rightmost particle until all particles have separated in order to establish the inequality above. Note that this can be violated on microscopic timescales, but not on macroscopic ones, i.e. this property is termed *macrocausality*.

If there is a conserved higher-spin charge P_s in the model, it must commute with the S-matrix, and we have

$$\langle final| S |initial\rangle = \langle final| e^{i\alpha P_s} S e^{-i\alpha P_s} |initial\rangle . \tag{4.2.10}$$

Thus $e^{i\alpha P_s}$ can be used to rearrange the initial and final configurations without changing the amplitude. If any of the outgoing rapidities are different from θ_1 or θ_2, then shifting the configurations around in this way will give a pattern of

trajectories for which $t_{12} \geq t_{23}$. By macrocausality, the amplitude for this pattern must vanish, and then by the insensitivity of the amplitudes to shifts induced by $e^{i\alpha P_s}$, all of the other amplitudes, including the one initially under consideration, must also vanish. Hence the only possibility for the two incoming particles when scattering are two outgoing particles with the same pair of rapidities as before the interaction. Therefore scattering events with more than two particles must *factorize* into a product of two-particle scattering amplitudes. For further details, as well as alternative arguments, please see [16] or [10]. We must now try to figure out how to determine the structure of these all-important two-particle S-matrices.

4.3 The Analytic S-Matrix

Given the incoming rapidities θ_1, θ_2 and the outgoing rapidities θ_3, θ_4 for a two-particle elastic scattering event, the S-matrix is given via

$$|\theta_1, \theta_2\rangle_{i,j}^{in} = S_{ij}^{kl}(\theta_1 - \theta_2)\, |\theta_1, \theta_2\rangle_{k,l}^{out} \tag{4.3.1}$$

where $\theta_1 > \theta_2$. Note that due to elastic scattering, $\theta_3 = \theta_2$ and $\theta_4 = \theta_1$. Using Mandelstam variables where $u = 0$ and $t(\theta_{12}) = s(i\pi - \theta_{12})$, the S-matrix now depends on only one Mandelstam variable, i.e. the Mandelstam variable s. Now we must determine the matrix elements. We first will focus on the constraints given by *discrete symmetries* that are usually obeyed by physical quantum field theories. In the following we will refer to Fig. 4.2 of the two-particle elastic S-matrix element $S_{ij}^{kl}(\theta_1 - \theta_2)$. If the theory in question is invariant under *parity*, i.e. invariant under the reflection of spatial coordinates, our particles i and j are swapped with particles k and l to obtain

$$S_{ij}^{kl}(\theta) = S_{ji}^{lk}(\theta). \tag{4.3.2}$$

Parity implies that we could read Fig. 4.2 from left to right or equivalently from right to left. Either choice will not change the result.

If our theory is invariant under *time reversal*, as its name implies, our theory will remain the same if we run the entire process in reverse. This is equivalent to reading Fig. 4.2 from bottom to top or equivalently from top to bottom, and then exchanging particles i with l and j with k. It is expressed as

$$S_{ij}^{kl}(\theta) = S_{lk}^{ji}(\theta). \tag{4.3.3}$$

If our theory is invariant under *charge conjugation*, i.e. where a particle is replaced

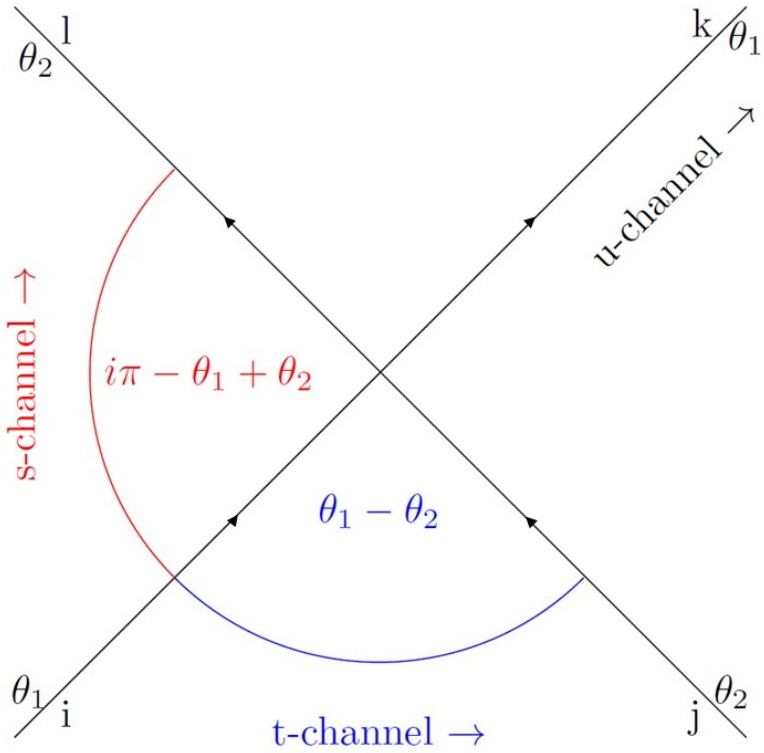


Figure 4.2: Representation of the S-matrix.

by an anti-particle, the S-matrix must remain invariant and is given by

$$S_{\bar{i}\bar{j}}^{\bar{k}\bar{l}}(\theta) = S_{ij}^{kl}(\theta), \tag{4.3.4}$$

where a particle i when charge-conjugated is denoted via barred indicies such as $\bar{i}$.

Now to continue with our defintion of analyticity, we must recall the definitions of the Mandelstam variables (4.1.10). We see that s, t, and u are the square of the center-of-mass energies in the channels defined by their respective processes, as can be seen in Fig 4.2. These are $i + j \to k + l$ for the *s-channel*, $i + \bar{l} \to k + \bar{j}$ for the $t - channel$, and finally $i + \bar{k} \to l + \bar{j}$ for the *u-channel*. Since only one of the Mandelstam variables is independent, it is standard to focus on s, the square of the forward-channel momentum. For physically possible processes θ_{12} is *real*. By consequence s is also real and satisfies $s \geq (m_i + m_j)^2$. We then study the analytic continuation of S(s) to the s-plane.

Such a study of analytic continuation necessarily requires that we introduce a Riemann surface $\mathcal{R}$. Denoting the orthogonal projection as $\rho : \mathcal{R} \to \mathbb{C}$ and allowing for $\alpha \in \mathbb{C}$, then the number of elements contained within the set $\rho^{-1}(\alpha)$ is equivalent to the number of unique results that may be obtained from analytically continuing the function along paths with an endpoint α. We place a branch cut

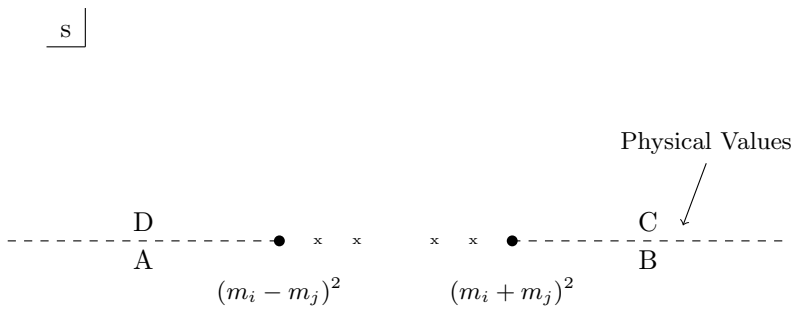


Figure 4.3: Representation of the first sheet of the full Riemann surface (*the physical sheet*). The dashed lines indicate the branch cuts along the sheet. Note that the four small x's in the diagram to correspond to possible poles for bound states. These do not occur in the sinh-Gordon model.

along the real axis from $s = (m_i + m_j)^2$ to $+\infty$. Due to *crossing symmetry*, which for the diagonal case is

$$S_{ij}(\theta) = S_{i\bar{j}}(i\pi - \theta), \tag{4.3.5}$$

there exists another branch cut that starts from $s = (m_i - m_j)^2$ to $-\infty$. These are the only two branch cuts when our matrix is factorized, and this setup imparts a particular set of properties on the function S.

First off, S itself is now a single-valued, meromorphic function on the complex plane. Physical values for s are located just above the right-hand branch cut. Additionally S is now *hermitian-analytic*, i.e.

$$S_{ij}^{kl}(s^*) = (S_{kl}^{ij})^*(s). \tag{4.3.6}$$

If we add to this property time reversal symmetry, we obtain the stronger condition of *real-analycity*, i.e.

$$S_{ij}^{kl}(s^*) = (S_{ij}^{kl})^*(s). \tag{4.3.7}$$

If s is real, then so is the function $S(s)$. We then have the situation $(m_i - m_j)^2 \leq s \leq (m_i + m_j)^2$. Note that this, in general, actually means that *real* S-matrices do *not* describe actual physical processes. We illustrate this condition in Fig. 4.3.

The next step is to consider *unitarity*. Whenever s^+ is a physical value for s, i.e. $s^+ = s + i0$ and $s \geq (m_i + m_j)^2$, unitary demands $S(s)$ satsifies the condition $S(s^+)S^\dagger(s^+) = 1$. This condition lends itself to a matrix equation with a sum over a complete set of asymptotic states nestled in between S and $S^\dagger$. Note that as s^+ grows it becomes energetically permissible for states that contain further particles to participate in the sum. This would normally force one to consider further S-matrix

elements, such as $2 \to 3$, etc. For integrable theories, matrix elements greater than $2 \to 2$ all vanish and *physical unitarity* becomes

$$S_{ij}^{kl}(s^+)(S_{lk}^{mn})^*(s^+) = \delta_i^n \delta_j^m. \tag{4.3.8}$$

If we further take into consideration real analyticity, (4.3.7), then (4.3.8) becomes

$$S_{ij}^{kl}(s^+)(S_{lk}^{mn})(s^-) = \delta_i^n \delta_j^m, \tag{4.3.9}$$

where $s^- = s - i0$ which lies just below the right-hand branch cut. This equation means that

$$S(s^+)S_\gamma(s^-) = 1, \tag{4.3.10}$$

where S_γ is the S-matrix analytically continued below the cut and around the first branch point $(m_i + m_j)^2$.

At this point, it would be helpful to know a bit more about the branch points. Unitary implies

$$S(s^+)S_\gamma(s^+) = 1, \tag{4.3.11}$$

Note that (4.3.11) implies that the branch cut is of square root type. We now focus on S_γ where we continue twice around the branch point to find

$$S_\gamma(s^-) = (S(s^+))^{-1} = (S_\gamma(s^+))^{-1} = S(s^+). \tag{4.3.12}$$

Note that we again made use of real analyticity (4.3.7) in the final equivalence above. Note that the first and the last equalities hold due to (4.3.11) and the middle equality is due to the definition of analytic continuation. Thus if we continue twice around this branch point, we find that we end up with the identical S-matrix from which we started. A similar calculation can be made for the second branch point, which yields a similar result. Therefore the branch cuts are of square-root type. Note that this does not imply that $\rho^{-1}(\alpha)$ has only two elements for all complex points α. From the outset, we have no idea what will occur if we encircle one branch point and then proceed to encircle the other.

In order to make things simpler to work with, knowing what is in store ahead, we will switch from the variable s to the difference of the rapidities by inversion of (4.3.12). This change forces the physical sheet to now map to the strip $0 \leq \mathrm{Im}(\theta_{12}) \leq \pi$. For the next sheet, we have $\pi \leq \mathrm{Im}(\theta_{12}) \leq 2\pi$. This process continues with periodicity $2\pi i$. Thus the branch cuts of the s-plane open up in such a manner that causes all of the Riemann sheets to be mapped into strips $n\pi \leq \mathrm{Im}(\theta_{12}) \leq (n+1)\pi$ with $S(\theta)$ analytic at the images $in\pi$ of the branch points. Thus $S(\theta)$ is a meromorphic function of θ and is *real* on the imaginary axis of θ.

Analyzing this another way, if we have the diagonal property, we also assume that each particle type is distinguished by the charges, then

$$S_{ij}^{kl} = 0, \tag{4.3.13}$$

whenever $i \neq k$ or $j \neq l$. We can then simply denote this as S_{ij}. Note that for an integrable theory this assumption is not always true, but it is true in our case as the sinh-Gordon model has only a single type of particle. In the θ-picture unitarity implies

$$S_{ij}(\theta)S_{ij}(-\theta) = 1. \tag{4.3.14}$$

To derive an important identity for our purposes, we will first make the definition that $\theta' = \theta + 2\pi i$. This is only to ensure that our final form complies with the convention shown in the literature. Note that because we are dealing with the diagonal case, we are able to treat these S-matrices as mere numbers mathematically. We begin with the crossing symmetry equation, i.e. (4.3.5), to obtain

$$\begin{aligned} S_{ij}(\theta') &= S_{i\bar{j}}(i\pi - \theta') \qquad \text{(Crossing)} && (4.3.15)\\ &= \frac{1}{S_{i\bar{j}}(-i\pi + \theta')} \qquad \text{(Unitarity)} && (4.3.16)\\ &= \frac{1}{S_{ij}(2\pi i - \theta')} \qquad \text{(Crossing)} && (4.3.17)\\ &= S_{ij}(\theta' - 2\pi i) \qquad \text{(Unitarity)} && (4.3.18) \end{aligned}$$

Now applying $\theta' = \theta + 2\pi i$ yields the identity

$$S_{ij}(\theta + 2\pi i) = S_{ij}(\theta), \tag{4.3.19}$$

which demonstrates that our Riemann surface $\mathcal{R}$ of the analytic continuation of S is simply a cylinder! Equation (4.3.19), with respect to s, is a *double cover*, i.e. whenever we cross two branch cuts we return to the initial sheet from whence we came.

At this point, we have covered the theories that govern how our QFT will behave in the infrared limit and in the uv-limit respectively. In light of our goal, we next will introduce the TBA.

Chapter 5

The Thermodynamic Bethe Ansatz

This chapter introduces and derives the asymptotic Bethe ansatz and the TBA. It culminates with the full and general derivation for the TBA equations. Along the way, we make specific selections to obtain the fermonic case, which we will eventually apply to the sinh-Gordon model. There are a plethora of resources that cover various Bethe ansatz techniques, including the TBA. This section was written to provide a unique and intuitive introduction to the TBA and its derivation. This section makes use of references [34], [24], [3], [7], [12], [26], and [25].

5.1 Background

In the original presentation of the Bethe ansatz, Hans Bethe obtained the energy eigenstates of the one-dimensional version of the ferromagnetic Heisenberg model of interacting, localized spins in a solid. This model is defined via a Hamiltonian, and to solve it amounts to finding all of the eigenstates and associated energies. Thus, the Bethe ansatz solves the system via diagonalizing the Hamiltonian. In principle one must prove that one indeed obtains eigenstates for the problem in question. On the other hand, the TBA follows from the fact that the scattering is purely elastic, and no other additional assumptions, nor input, are involved in its derivation. The TBA is thus unique compared to the other incarnations of the Bethe ansatz, and in many ways more fundamental.

At its roots, the TBA is a procedure that allows one to take the problem of calculating the complete infinite-volume thermodynamics of a purely elastic scattering theory (PEST) and turn it into a problem of finding the solution to a set of coupled non-

linear integral equations for the single-particle excitation energies and the rapidity distributions of the particles for the theory in question.

5.2 Statistical and Relativistic Quantum Theory Equivalence

5.2.1 Transfer Matrix and the Partition Function

We will begin our derivation of the TBA equation by solving for the partition function, which is part of classical statistical theory. One might wonder how a classical statistical system with short-range interactions is equivalent to a relativistic quantum theory. The key is the *transfer matrix* method. We then must introduce the transfer matrix, and we do so via the Ising model in 1-dimension.

The Ising model consists of a regular line of spins σ_i which assume unit values ± 1. We begin by replacing the topology of the straight, open chain by a closed chain where we identify the n^{th} spin as a neighbor of the first, i.e. $\sigma_{n+1} \equiv \sigma_1$. This eliminates the difficulties of the endpoint effects in 1-D and preserves the thermodynamic properties of the infinite open chain. This prescription allows us to write the Ising Hamiltonian as

$$H_n(\sigma_i) = -J \sum_i^n \sigma_i \sigma_{i+1} - \mu B \sum_i^n \sigma_i. \tag{5.2.1}$$

Working with the simpler case of an absent external magnetic field ($B = 0$) and setting $J = -1$ (antiferromagnetic case), we note that the interaction energy between the spin on lattice sites i and $i + 1$ is assumed to be

$$H_0(\sigma_i, \sigma_{i+1}) = \sum_i^n \sigma_i \sigma_{i+1}, \tag{5.2.2}$$

which implies that the total energy of a configuration of spins is given by

$$H = \sum_{i=1}^n H_0(\sigma_i, \sigma_{i+1}). \tag{5.2.3}$$

Given our goal of introducing the concept of the tranfer matrix, we do not need to concern ourselves with the exact form of the interaction energy $H_0(\sigma_i, \sigma_{i+1})$. For the case of a classical discrete system in the canonical ensemble, the partition function is defined as

$$Z = \sum_i^n e^{-\beta H_i}. \tag{5.2.4}$$

The partition function of our system with a general interaction energy becomes

$$Z = \sum_{\{\sigma_i\}} e^{-\beta \sum_{i=1}^{N} H_0(\sigma_i, \sigma_{i+1})} \tag{5.2.5}$$

$$= \sum_{\{\sigma_i\}} \prod_{i=1}^{N} e^{-\beta H_0(\sigma_i, \sigma_{i+1})} \tag{5.2.6}$$

$$= \prod_{i=1}^{N} \sum_{\substack{\sigma_i = \pm 1, \\ \sigma_{i+1} = \pm 1}} e^{-\beta H_0(\sigma_i, \sigma_{i+1})}. \tag{5.2.7}$$

The partition function in the final form above motivates the introduction of a 2 x 2 matrix with elements

$$T = (T_{ij}) = \left(e^{-\beta H_0(\sigma_i, \sigma_{i+1})}\right) = \begin{pmatrix} e^{-\beta H_0(+1,+1)} & e^{-\beta H_0(+1,-1)} \\ e^{-\beta H_0(-1,+1)} & e^{-\beta H_0(-1,-1)} \end{pmatrix}. \tag{5.2.8}$$

The matrix T is known as the *transfer matrix*. Note that the interaction Hamiltonian $H_0(\sigma_i, \sigma_{i+1})$ is symmetrical in its arguments, which implies that the transfer matrix is a real, symmetric matrix with positive matrix elements.

The transfer matrix elements represent all of the possible Boltzmann weights of two coupled spins. If we were to take the product of two transfer matrices, we would obtain a new matrix given by $T * T = T^2$. This new matrix T^2 represents all of the possible Boltzmann weights of a short chain of three spins, where the inner spin has been summed over. This implies that T^2 depends only on the first and the third spins. If we were to take further self-products of the transfer matrix, we would continue to create new matrices where each new resulting matrix has their inner spins summed over. Thus for each resulting matrix, it will only depend on the first and the last spins. Therefore we produce the expression (5.2.7) via successive self-products of the transfer matrix, terminating once we reached the final lattice site n. Once we have reached length n, we must take into account the periodicity of our chain. Having identified $\sigma_{n+1} \equiv \sigma_1$, we are compelled to take the trace of our matrix to obtain equation (5.2.7) expressed in terms of the transfer matrix as

$$Z_n = \mathrm{Tr}\left(T^n\right). \tag{5.2.9}$$

For our purposes it would be more convienient to work with the 2-D case, and so we seek to generalize our findings for the case of the 2-D Ising model. After having introduced everything for the 1-D case, the extension to 2-D is done quite readily. Instead of a closed spin chain in 1-D, we now operate upon a 2-D regular square lattice of spins where we have i-rows and j-columns. The Hamiltonian of the system

is only slightly modified from the 1-D case, and now reads

$$H = -J \sum_{<ij>} \sigma_i \sigma_j - \mu B \sum_{i}^{N} \sigma_i. \tag{5.2.10}$$

In 1-D we identified the n^{th} spin as a nearest neighbor of the first. For the 2-D lattice, we must do so for both the m^{th} spin for the rows and the n^{th} spin for the columns

$$\sigma_{ij} \equiv \sigma_{i+m,j} \text{ and } \sigma_{ij} \equiv \sigma_{i,j+n}. \tag{5.2.11}$$

Note that this identification now gives our lattice the shape of a torus, compared to the circular chain for the 1-D case. This shape will become important later on.

One now introduces a complex vector space V of row spin configurations $|\mu_i\rangle$, where we introduce a bra-ket notation analogous with quantum mechanics and denote μ_i to be the configuration of spins on the i^{th} row

$$\mu_i = \{\sigma_{i1}, ..., \sigma_{in}\} \tag{5.2.12}$$

where there are 2^n such configurations.

We now define the transfer matrix T by its matrix elements $\langle\mu| T |\mu'\rangle$, and from this we can write the partition function in terms of the transfer matrix via

$$Z = \sum_{\mu_1 ... \mu_m} \langle\mu_1| T |\mu_2\rangle \langle\mu_2| T |\mu_3\rangle ... \langle\mu_m| T |\mu_1\rangle . \tag{5.2.13}$$

Summing (5.2.13) via the resolution of the identity, we arrive at the result

$$Z_m = \sum_{\mu_1} \langle\mu_1| T |\mu_1\rangle = \mathrm{Tr}\,(T^m) . \tag{5.2.14}$$

Comparing (5.2.14) with (5.2.9) shows that both the 1-D and the 2-D cases lead to identical expressions for the partition function, so we don't need to concern ourselves with how the dimension of the problem might affect the form of the partition function. Only the structure of the lattice on which the problem lives (1-D on a circular chain and 2-D on a torus) has changed.

5.2.2 Why Are Lattices So Important?

Attempts to directly work with QFTs on a continous spacetime manifold often present problems. Arguably the fundamental problem in QFT remains the divergence of Feynmann integrals and how to handle them. Thus it is apparent that for most QFTs, one cannot properly define them directly. Since it is well understood that

a statistical lattice model can be viewed as the regularized version of a continuous QFT, the workaround to this issue is to define the lattice version of a particular QFT first. Based on one's knowledge or thoughts of what the form of the continous theory should be, constraints may be imposed on the lattice model in order that when one sends the lattice spacing to zero (*scaling limit*) to obtain the continous QFT, one recovers the desired behavior of the continuous theory.

5.2.3 Towards a Quantum Field Theory

Considering our ultimate problem deals with a quantum field theory, we would like to make the connection from the statistical model to the field theoretical one.

One may define a Hamiltonian operator $\hat{H}$ as

$$T \equiv e^{-a\hat{H}}, \tag{5.2.15}$$

where a is the lattice spacing and T is the transfer matrix. Inserting the above into (5.2.14) immediately yields the partition function of the form

$$Z = \text{Tr } (T^m) = \text{Tr } \left(e^{-am\hat{H}}\right), \tag{5.2.16}$$

where we see that Z now depends on the number of lattice points m.

Recall that to obtain (5.2.16), we acted on the 2^n-dimensional space of row spin configurations. The lattice version of our QFT associated with the 2-D Ising model (5.2.10) has a Hilbert space of states isomorphic to the space of row spin configurations V. In this setup, the space direction is fixed and oriented along the rows while the time-direction is also fixed and oriented along the columns of the lattice. Alternatively we could have also decided to start with a transfer matrix T' that acted on the 2^m-dimensional space of column spin configurations. We would then have had the opposite orientation for the space and time directions in our lattice QFT, and have obtained a new Hamiltonian $\hat{H}'$. This would have given us a partition function

$$Z = \text{Tr } (T^n) = \text{Tr } \left(e^{-an\hat{H}'}\right) \tag{5.2.17}$$

where we now have n lattice sites and a is the lattice spacing. Note that the expressions for the two partition functions, (5.2.16) and (5.2.17), are equivalent. Like the dimension of our problem, it also does not matter which direction we orient the space and time directions on our lattice.

5.3 The Geometric Setup

We now begin our derivation of the TBA by considering a relativisitic field theory in toroidal geometry in 1+1-dimensions, with the intention that at some point in the future we will end up with a cylindrical surface as the limit of a torus in order to analyze the finite size effects we are after. Recall previously that our 2-D lattice formed a torus when we identified the next nearest neighbors of the rows and columns of our regular square lattice, lending support to the idea that the torus is the proper surface to consider for a 2-D integrable model.

We first construct a flat torus from two orthogonal circles C_1 and C_2 with circumferences R and L respectively. We will hereafter refer to the two circles making up the torus by their circumferences. We then superimpose an artificial cartesian coordinate system upon our torus, with the x-axis running along the contour of R and the y-direction running parallel with respect to L. From such a setup, there are two topologically distinct ways one can develop the Hamiltonian approach to this problem.

Approach 1: One may choose the field theory states to lie on circle R. We denote the corresponding space of states as $\mathcal{H}_R$, as in Fig. 5.1a. By consequence of our choice to place the space of states along R, governed by the x-axis, the y-axis now must play the role of the time-direction. That is, time now must lie along L. The states are time-evolved by the Hamiltonian

$$H_R = P_y^{(R)} = \frac{1}{2\pi}\int_R T_{yy}\,dx, \tag{5.3.1}$$

where $T_{\mu\nu}$ is the stress tensor of the theory. The space momentum

$$P_R = P_x^{(R)} = \frac{1}{2\pi}\int T_{xy}\,dx, \tag{5.3.2}$$

is quantized and its eigenvalues are given by $\frac{2\pi n}{R}$, where n is an integer in space $\mathcal{H}_R$.

Approach 2: One may instead choose contour L as the quantization surface. We then denote the corresponding space of states as $\mathcal{H}_L$, as in Fig. 5.1b. Based on this alternative choice for the field theory states, time is now constrained to point along R. The states are then chosen to time-evolve in the $-x$-direction, which preserves frame orientation, and are given by the Hamiltonian

$$H_L = -P_x^{(L)} = \frac{1}{2\pi}\int_L T_{xx}\,dy. \tag{5.3.3}$$

Thus for this setup, the space momentum is given by

$$P_L = P_y^{(L)} = \frac{1}{2\pi}\int T_{xy}\,dy, \tag{5.3.4}$$

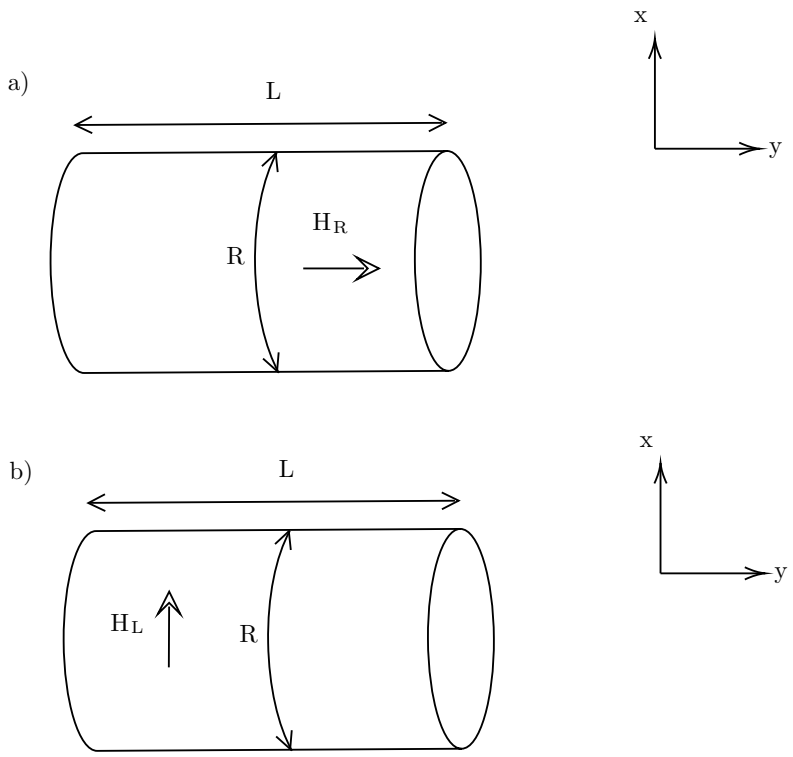


Figure 5.1: Representation of the two topologically distinct ways in which one could select space and time directions upon a torus. The arrow and Hamiltonian labels H_R and H_L denote the direction of time and the operators generating the time evolution respectively.

which has eigenvalues $\frac{2\pi n}{L}$, where once again n is an integer in space $\mathcal{H}_L$.

Now recall the form of the partition function (5.2.16) for our lattice construction. It depends on the lattice spacing a and the number of row lattice points m. The combination am in that equation was the circumference 'R' for that torus of lattice points. To go from a discrete (lattice) to a continuous theory we take the scaling limit, which implies

$$\mathbf{Z} = \mathrm{Tr}\, e^{-am\mathbf{H_L}} \longrightarrow Z \equiv \mathrm{Tr}\, e^{-RH_L}, \tag{5.3.5}$$

where here we used boldface to distinguish the discrete theory from the continuous theory (without boldface).

Recall that in the lattice theory, we were free to choose our orientations for space and time. Either choice gave us an equivalent partition function. In our current construction of the torus, where we need make no mention of the lattice theory, we arrive at the identical configuration as before. The new partition function is given

by

$$Z = \mathrm{Tr}\, e^{-LH_R} = \mathrm{Tr}\, e^{-RH_L}, \tag{5.3.6}$$

where we see that Z for our torus has an equivalent form to Z derived from lattice considerations.

For the purposes of continuing our derivation, we will take *approach 1* and will now consider the thermodynamic limit $L \to \infty$ with $L \gg R$. In the limit $L \longrightarrow \infty$ the torus regains a cylindrical shape with circumference R. We hope that the space of states detected by H_L can be described sufficiently well via Asymptotic Bethe Ansatz (AsBA) states in order to capture the partition function $Z(R, L)$. Note here we use the abbreviation AsBA since the abbreviation ABA is generally recognized as referring to the Algebraic Bethe ansatz, which does not come into play in our work. We will introduce the AsBA in the next subsection.

At the critical point the largest eigenvalues of T coalesce. When $L \to \infty$, the partition function $Z(R, L)$ of the field theory is dominated by the ground-state of H_R with (ground-state) energy $E(R)$. The partition function behaves in this limit as

$$Z(R, L) \sim e^{-E(R)L}. \tag{5.3.7}$$

Note that R remains finite in the thermodynamic limit. Since $L \longrightarrow \infty$, the ground-state energy $E(R)$ should be exact. If after the ground-state alone, we now determine that we do not ever need to consider a lattice theory in deriving the TBA. This is due to the fact that we will arrive at the equivalent partition function simply by constructing a torus from the start. Note that in (5.3.7), the space of states was taken to be $\mathcal{H}_R$, that is along the R-direction of the torus. Keeping the same limit $L \longrightarrow \infty$, we also could have taken the other point of view and chose the quantization surface along L. The partition function from this choice reads

$$Z(R, L) = \mathrm{Tr}\, \left[e^{-RH_L}\right] \tag{5.3.8}$$

where the trace is taken over the Hilbert space $\mathcal{H}_L$.

We now assume that we have control over the Hilbert space of states $\mathcal{H}_L$, which we certainly do not have for $\mathcal{H}_R$ when R is finite. In this limit, L is consequently very large. We can therefore assume that states in $\mathcal{H}_L$ are able to be described as N particles which freely propagate except when their respective paths approach each other with close proximity, where they scatter via the (exact) S-matrix.

This description of particles in $\mathcal{H}_L$ is reminiscent of particles in asymptotic states, which behave as free particles before and after any interaction between two particles. Since the particles are well-separated for all times other than when they scatter, we

can likely speak of a wavefunction (assuming particle conservation). This suggests that the space of states seen by the Hamiltonian H_L can be accurately described by the Asymptotic Bethe Ansatz (AsBA). Note that this is initially an assumption we have, but it turns out to be correct. The AsBA is detailed in the following section.

5.4 The Bethe Wavefunction and the Asyptotic Bethe Ansatz

As discussed in the section of this thesis on S-matrices, in an integrable theory the scattering has the following special features:

- The number of particles is conserved in any scattering event.
- The momenta are not mutable, but only redistributed between particles.
- The S-matrix for multi-particle scattering factorizes, i.e. the S-matrix for any number of particles can be written as a product of $2 \to 2$ S-matrices.

One cannot normally introduce the familiar wavefunction from single-particle quantum mechanics in QFT due to the production of real and virtual particles. There is an exception due to the features an integrable system presents. Primarily due to particle conservation, we are able to introduce a wavefunction in at least some regions of the configuration space. From the periodicity of the wavefunction, one can then derive quantization conditions which determine the spectrum of the theory.

Let us begin by envisioning a model where particles live on a line of length L. This is appropriate as asymptotic scattering has a particle coming in from infinity, scattering, and then going out towards infinity. Hence this is just a line. Since we are solving for a theory on a cylinder, we may identify the endpoints of this line. This leaves us with a circle. Our theory is integrable, so the number of particles is conserved. We can therefore speak of a wavefunction as we did in single-particle quantum mechanics. If there are n-particles on a circle, the wavefunction Ψ consequently must be periodic. We now endeavor to construct an appropriate wavefunction.

Let us define the first particle's momentum as p_1 and send this particle around the circle, returning the particle to its original starting point at the end of this loop. See Figure 5.2 for reference. If this particle was the only particle in the theory, or if the theory was non-interacting, the wavefunction would acquire a phase factor $\Psi = e^{ip_1 L}$. We call such a region where the particle does not interact a free region, and denote it by $\{i_1, ..., i_p\}$ if $x_{i_1} < ... < x_{i_N}$. From the periodicity of the wavefunction,

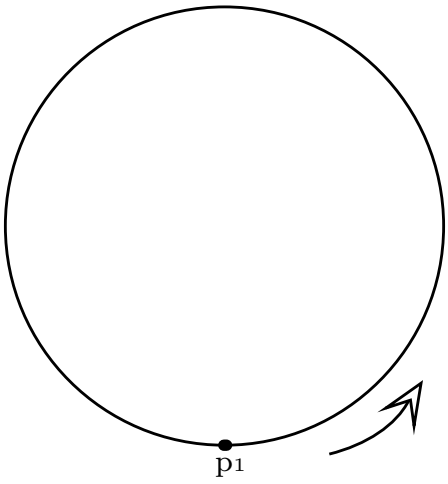


Figure 5.2: Representation of asymptotic scattering on a periodic surface. Note here that we only have one particle with momentum p_1 on the circle. Thus here we do not have a single scattering event take place. We only obtain a periodic factor for the wavefunction once the particle completes one full rotation around the circle.

$$e^{ip_1 x} = e^{ip_1(x+L)} \longrightarrow e^{ip_1 L} = 1, \tag{5.4.1}$$

and the momentum would be quantized according to

$$p_1 = \frac{2\pi k}{L},\ k \epsilon \mathbb{Z}. \tag{5.4.2}$$

Now allowing for more than one particle, we must take into account the interaction of the first particle with all of the other particles along the circle as we transverse once around the circle. The *interaction range* is defined as the inverse mass scale of the theory in question. If L is large compared to the interaction range between the particles on the circle, i.e. $|x_i - x_j| >> \xi$ where ξ is the correlation length, then the particles are usually well-separated from each other.

In transition regions where two or more particles are close to each other, relativistic effects become paramount and we cannot use the wavefunction any longer. The scattering theory provides conditions to match wave functions in the adjacent free regions, so we are not at a total loss and may continue. Given the purely elastic case, every transition, i.e. $\{i_1, ..., i_p, i_{p+1}, ..., i_N\} \longrightarrow \{i_1, ..., i_{p+1}, i_p, ..., i_N\}$, results in the multiplication of the wavefunction by the S-matrix. For real rapidities θ, the scattering amplitude is a unimodular number with real phase $\delta(\theta)$ given by

$$S(\theta) = e^{i\delta(\theta)}. \tag{5.4.3}$$

Due to particle number conservation, when we take the particle around the circle, the first particle will scatter through all the other particles. Please see Figure 5.3 for reference. For each scattering, the wavefunction is multiplied by the S-matrix, which manifests as a simple phase factor $S(\theta_1 - \theta_2) = e^{i\alpha(\theta_1 - \theta_2)}$.

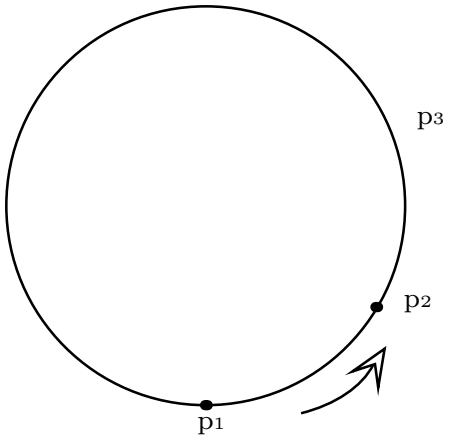


Figure 5.3: Representation of asymptotic scattering on a periodic surface with multiple particles, each having their own momentum. We take it that p_1 is the only particle that moves, while p_2 and p_3 remain stationary. As p_1 moves around the circle, it will come into contact with p_2 and p_3 separately. Each interaction is a two-particle scattering event, and is recorded by multiplying our wavefunction by the S-matrix for each scattering event.

We are left with a product of S-matrix phases which combine with the phase (*Bethe wavefunction*) we acquired from the initial propagation of the free particle, $e^{ip_1 L}$, around the circle. Once again we require the wavefunction to be periodic, which yields a condition

$$e^{ip_1 L} S(\theta_1 - \theta_2) S(\theta_1 - \theta_3) ... S(\theta_1 - \theta_n) = 1. \tag{5.4.4}$$

Now we generalize (5.4.4) for any particle. Our equation now reads

$$e^{ip_i L} \prod_{j \neq i}^{N} S(\theta_i - \theta_j) = 1, \tag{5.4.5}$$

where (5.4.5) are known as the *asymptotic Bethe ansatz equations* for the model. This is a key component in the derivation of the TBA equation.

5.5 Putting It All Together - The TBA Derivation

The TBA equations emerge in the limit $L >> R_c$, where one can dismiss off-shell effects. Note that R_c refers to a finite correlation length. In order to avoid complicating matters, we will consider a simple scattering theory in which we have only one neutral particle of mass m and only one pair of scattering amplitudes $S(\theta_1 - \theta_2)$.

The *rapidities* θ_1 and θ_2 of the particles parameterize their on-shell energies and momenta

$$e_i(\theta) = m\cosh\theta_i, \qquad p_i(\theta) = m\sinh\theta_i. \tag{5.5.1}$$

and the amplitude $S(\theta)$ satisfies both unitarity and crossing symmetry respectively, i.e.

$$S(\theta)S(-\theta) = 1, \qquad S(\theta) = S(i\pi - \theta). \tag{5.5.2}$$

Now using the Euler formula ($1 = e^{2\pi i}$), (5.4.3), and the definition of the on-shell momenta (5.5.1) in (5.4.5), we are able to take the logarithm of both sides to immediately yield

$$mL\sinh\theta_i + \sum_{j\neq i}^{N}\delta(\theta_i - \theta_j) = 2\pi n_i. \tag{5.5.3}$$

with N integer numbers n_i. Equation (5.5.3) is a system of transcendental equations that picks out a proper set of rapidities $(\theta_1, ..., \theta_N)$ in the free regions $\{i_1, ..., i_p\}$ of $\mathcal{H}_L$. The energy and momentum of the state $(\theta_1, ..., \theta_N)$ is given by

$$H_L(\theta) = \sum_{i=1}^{N} m\cosh\theta_i, \qquad P_L(\theta) = \sum_{i=1}^{N} m\sinh\theta_i. \tag{5.5.4}$$

Substituting the scattering amplitude (5.4.3) in the unitarity relation (5.5.2) yields

$$\delta(\theta) + \delta(-\theta) = 2\pi n \tag{5.5.5}$$

for an integer n. So if we used (5.5.5) and the momentum (5.5.4) in equation (5.5.3), we would immediately get the condition

$$e^{-P_L L} = 1. \tag{5.5.6}$$

This is a welcome confirmation that our assumptions were correct, as one would indeed expect that the wavefunction would be periodic for a periodic circle L.

Now as per our assumptions for the problem, we are dealing with a simple scattering theory consisting of identical particles. This implies that the Bethe wavefunction must either be symmetrized or antisymmeterized depending on the particle statistics.

Returning to our unitarity condition (5.5.2) once again, we see that for identical particles

$$S^2(0) = 1, \tag{5.5.7}$$

which implies via $S(0) = \pm 1$ that two different cases are possible.

CASE I: $S(0) = -1$

- *Bosons* - Let us assume the identical particles are bosons and allow for two particles to have identical rapidities. If we exchange these two particles, $S(0)$ indicates that we multiply the wavefunction by -1. This implies they are antisymmetric in the coordinates, which is *not* a feature of Bose statistics. In the bosonic case, such states for the rapidities must be excluded and each value of the rapidity can be occupied by *at most* one particle. This implies all of the integers n_i are unique. This is akin to the behavior of fermions, and thus this situation in the Bethe ansatz equations is referred to as '*fermionic*'.

- *Fermions* - Let us assume the identical particles are now fermions with two particles having identical rapidities. Exchange is allowed in this case, as Fermi statistics do allow for identical rapidities. Fermions therefore can occupy each rapidity value in any number, which is reminiscent of the behavior of bosons. Hence, this case of the Bethe ansatz equations is referred to as '*bosonic*'.

CASE II: $S(0) = 1$

- *Bosons* - Identical particles are bosons with identical rapidies. Everything is compatible with Bose statistics and the particles behave like bosons properly. Thus any value of the rapidity can be occupied any number of times. The Bethe ansatz equations are '*bosonic*'.

- *Fermions* - Identical particles are fermions with identical rapidies. Upon exchange the wavefunction will be multiplied by 1 (from S(0)). This is incompatible with Fermi statistics and the particles behave like proper fermions (all rapidities must differ). The Bethe ansatz equations are '*fermionic*'.

It turns out that all of the currently known interacting cases are of the "fermionic" type, and therefore the set of $\{n_i\}$'s in (5.5.3) are strictly increasing. Note that in the sinh-Gordon model, which we will cover in a later chapter, the particles are bosons. Therefore the sinh-Gordon model is an example of case I, and the particles are bosons.

Returning to our set of transcedental equations (5.5.3), this situation is receptive to analysis in the thermodynamic limit. As $L \longrightarrow \infty$, the number of particles in the thermodynamic states increases. The spectrum of rapidities given by (5.5.3) condenses and the distance between adjacent levels behaves as $\theta_i - \theta_{i+1} \sim \frac{1}{mL}$.

Since the rapidities condense in the limit $L \longrightarrow \infty$, it now becomes fruitful to introduce a continuous *rapidity density of particles* $\rho_1(\theta)$ and proceed to use this new variable in our analysis of the Bethe equations. If we take a tiny rapidity interval $\Delta\theta$ containing n particles, we are able to define the rapidity density as

$$\rho_1(\theta) = \frac{n}{\Delta\theta}. \tag{5.5.8}$$

Note that the smooth function θ is independent of the choice of our interval $\Delta\theta$ while $\frac{1}{mL} \ll \Delta\theta \ll 1$. Also note that the phase sum, i.e. $\sum_{j\neq i} \delta(\theta_i - \theta_j)$, in (5.5.3) is nearly constant when one varies from one θ_i the next θ_{i+1}. We can therefore estimate this phase sum as an integral, and (5.5.3) now has the form

$$mL \sinh\theta_i + \int \delta(\theta_i - \theta')\rho_1(\theta')d\theta' = 2\pi n_i. \tag{5.5.9}$$

Equation (5.5.9) can be thought of as the equation for rapidity levels, where rapidity levels are solutions to this equation for all integers n_i. We reiterate that solutions are for all integer numbers n_i on the right-hand side of the equation, and do *not* restrict this to mean only the n_i's that correspond to the actual state of the system. This situation is akin to that of a system of free particles, where the quantization condition for a single-particle determines the set of allowed levels and we speak of free and occupied levels. In contrast to the free state, the set of levels is now arranged self-consistently with the particle distribution.

We now introduce yet another new variable, the *level density* $\rho(\theta)$, which will take the place of the set of integers n_i. Our equation (5.5.9) now becomes

$$mL \sinh\theta_i + \int \phi(\theta_i - \theta')\rho_1(\theta')d\theta' = 2\pi\rho(\theta), \tag{5.5.10}$$

where

$$\phi(\theta) = \frac{\partial\delta(\theta)}{\partial\theta} \tag{5.5.11}$$

The energy of the system is also modified from (5.5.4) to

$$H_L(\theta) = \int m \cosh\theta\rho_1(\theta)d\theta. \tag{5.5.12}$$

We pause here to mention that with this equation and (5.5.10), we now have an equation for $\rho(\theta)$ in terms of $\rho_1(\theta)$. This type of equation is known as *iterative*, in that its solution depends on the equation itself. The final form of the TBA will

retain this feature.

For every consistent pair of densities ρ and ρ_1 in the limit $L \longrightarrow \infty$ there exist a large number of corresponding quantum states. We now partition the rapidity axis into small intervals $\Delta\theta_\alpha \ll 1$. If the condition $\Delta\theta_\alpha \gg \frac{1}{mL}$ is also satisfied, then for every interval $\Delta\theta$ there exists a large number $N_\alpha \sim \rho(\theta_\alpha)\Delta\theta_\alpha$ of levels in each interval with roughly $n \sim p_1(\theta_\alpha)\Delta\theta_\alpha$ particles distributed between them. Note that the averaged densities are not sensitive to a *local* redistribution of particles amongst them. We are interested in the total number of rearrangements in the interval $\Delta(\theta)$. The following analysis has two possibilities depending whether our case is 'fermionic' or 'bosonic'. In the upcoming derivation of the TBA equation we follow the work in [34] without further explanation or derivation.

CASE A: '*Fermionic*'
The number of distributions in the interval $\Delta\theta$ is

$$\mathscr{N}(\rho, \rho_1) = \frac{(N_\alpha)!}{(n_\alpha)!\,(N_\alpha - n_\alpha)!}. \tag{5.5.13}$$

CASE B: '*Bosonic*'
The number of distributions in the interval $\Delta\theta$ is

$$\mathscr{N}(\rho, \rho_1) = \frac{(N_\alpha + n_\alpha - 1)!}{(n_\alpha)!\,(N_\alpha - 1)!}, \tag{5.5.14}$$

as given in [34].

As $L \longrightarrow \infty$, the *number of states* $\mathscr{N}(\rho, \rho_1)$ corresponding to given consistent densities ρ and ρ_1 is estimated via the entropy ($\mathcal{S} = \log \mathscr{N}(\rho, \rho_1)$). While the condition $\frac{1}{mL} \ll \Delta\theta \ll 1$ is satisfied, $\mathscr{N}(\rho, \rho_1)$ is given by the product of either (5.5.13) or (5.5.14) over the intervals $\Delta\theta_\alpha$ depending on the system configuration, i.e. 'fermionic' or 'bosonic'. The entropy is given by

CASE A:

$$\mathcal{S} = \int [\rho \log(\rho) - \rho_1 \log(\rho_1) - (\rho - \rho_1)\log(\rho - \rho_1)]d\theta \tag{5.5.15}$$

CASE B:

$$\mathcal{S} = \int [(\rho + \rho_1)\log(\rho + \rho_1) - \rho \log(\rho) - \rho_1 \log(\rho_1)]d\theta \tag{5.5.16}$$

With the entropy now accounted for, the summation of states found in (5.3.8) now amounts to finding the minimum of the free energy in ρ and ρ_1, which are macroscopic characteristics. The minimum of the free energy is given by

$$-RLf(\rho, \rho_1) = -RH_{\mathcal{H}_L}(\rho_1) + \mathscr{N}(\rho, \rho_1), \tag{5.5.17}$$

which is constrained by (5.5.9). We now introduce the '*pseudoenergy*' $\epsilon(\theta)$, given by

CASE A:

$$\frac{\rho_1}{\rho} = \frac{e^{-\epsilon}}{1+e^{-\epsilon}}; \qquad e^{-\epsilon} = \frac{\rho_1}{\rho-\rho_1}; \quad , \tag{5.5.18}$$

CASE B:

$$\frac{\rho_1}{\rho} = \frac{e^{-\epsilon}}{1-e^{-\epsilon}}; \qquad e^{-\epsilon} = \frac{\rho_1}{\rho+\rho_1}; \quad . \tag{5.5.19}$$

Using this notation, the extremum condition has the form

CASE A:

$$-Rm\cosh\theta + \epsilon(\theta) + \frac{1}{2\pi}\int \phi(\theta_i-\theta')\log\left(1+e^{\epsilon(\theta))}\right)d\theta' = 0, \tag{5.5.20}$$

CASE B:

$$-Rm\cosh\theta + \epsilon(\theta) - \frac{1}{2\pi}\int \phi(\theta_i-\theta')\log\left(1-e^{\epsilon(\theta))}\right)d\theta' = 0, \tag{5.5.21}$$

and we note that ρ and ρ_1 enter the expression for the extremal free energy f in the ratio $\frac{\rho_1}{\rho}$ only for

$$Rf(R) = \mp m\frac{1}{2\pi}\int \cosh(\theta)\log\Big(1\pm e^{-\epsilon(\theta)}\Big)d\theta \tag{5.5.22}$$

where the upper sign refers to the 'fermionic' case and the lower sign refers to the 'bosonic' case. As mentioned previously, the sinh-Gordon model is of 'fermionic'-type and the particles in the model are bosons. In order to proceed with the derivation of the appropriate form of the TBA for our purposes, we choose CASE I with bosonic particles and all instances of CASE A.

In the general case there are various types of particles $A_a, a = 1, 2, ..., N$ each with their own mass m_a. The PEST is described by a symmetric $N\mathrm{x}N$ matrix of pair transition amplitudes given by $S_{ab}(\theta)$. Each $S_{ab}(\theta)$ satisfies unitarity, i.e. $S_{ab}(\theta)S_{ab}(-\theta) = 1$. If charged particles are present, then the crossing symmetry relation in (5.5.2) may need a bit of an adjusting. One is concerned with N level densities $\rho^a(\theta)$ and N particle densities $\rho_1^{(a)}$, which modify our integral equation for the rapidity levels from (5.5.9) to

$$\frac{m_a L}{2\pi}\cosh(\theta) + \phi_{ab} * \rho_1^{(b)} = \rho^{(a)}, \tag{5.5.23}$$

where * indicates that we take the convolution of the terms

$$\phi * \rho_1 = \frac{1}{2\pi}\int \phi(\theta-\theta')\rho_1(\theta')d\theta', \tag{5.5.24}$$

and ϕ_{ab} is the *symmetric matrix kernel* given by

$$\phi_{ab}(\theta) = -i\frac{d}{d\theta}\log(S_{ab}(\theta)). \tag{5.5.25}$$

The 'fermionic' type pseudo-energies are given by

$$\rho_1^{(a)}(\theta) = \frac{e^{-\epsilon_a(\theta)}}{1+e^{-\epsilon_a(\theta)}}\rho^{(a)}(\theta). \tag{5.5.26}$$

Note that for the 'bosonic'-type, the addition in the denominator would change sign.

The extremum equation also transforms into a system of nonlinear integral equations for N pseudo-energies $\epsilon_a(\theta)$. With foreknowledge of our final form, we wish to simplify some of the notation. In this effort, we introduce a new term L given by

$$L_a(\theta) = \pm\log\Big(1 \pm e^{-\epsilon_a(\theta)}\Big), \tag{5.5.27}$$

where the plus sign is for the 'fermionic' type and the negative sign is for the 'bosonic' type. We choose the 'fermionic' type. Now we combine (5.5.20), (5.5.23), and (5.5.27) to arrive at the final form of our *TBA equations* which read

$$-m_a R\cosh\theta + \epsilon_a + \sum_{b=1}^{N}\phi_{ab} * L_b = 0. \tag{5.5.28}$$

Note that the TBA equations are analytic in a strip, which we refer to as the *analyticity strip*. It is model dependent and comes from the kernel ϕ in (5.5.28). These equations allow us to calculate the ground-state energy of the model, which is given by

$$E(R) = Rf(R) = -\frac{1}{2\pi}\sum_{a=1}^{N} m_a \int L_a(\theta)\cosh\theta\, d\theta, \tag{5.5.29}$$

where f is the bulk free energy. As we mentioned previously in this section, it is only the ground-state energy that we are able to obtain in this manner. Using the central charge from CFT, we can now show that the results are valid for the model in question.

5.5.1 The TBA and the UV-Limit

Taking the definition of the parameter r as

$$r \equiv mR, \tag{5.5.30}$$

we now take the uv-limit, i.e. $r \to 0$, expecting to recover the underlying CFT of the given theory. Previously we found that a CFT which lies on a cylinder with

circumference R has a Hamiltonian given by

$$H_{cyl} = \frac{2\pi}{R}\left(L_0 + \overline{L}_0 - \frac{c}{12}\right). \tag{5.5.31}$$

Since the vacuum is invariant with respect to global conformal transformations, this implies that $L_0 |0\rangle = 0$ and $\overline{L}_0 |0\rangle = 0$. Therefore in the the uv-limit the vacuum expectation value is equal to zero, and our original selection for the energy and momentum of the state (5.5.4) is in accordance with the naturally fixed vacuum energy of CFT's. Thus for a fixed value of R, namely $R = R_0$, we have the limiting behavior

$$\lim_{m\to 0} E(m, R) = E_{CFT}(R) \tag{5.5.32}$$

$$= -\frac{c\pi}{6R}, \tag{5.5.33}$$

where $E(m, R)$ is the *ground-state energy* and c the *central charge* of the CFT.

Notice that the TBA equation solely depends upon the product $r = mR$. Its solution $\epsilon(\theta)$ thus depends on $r = mR$ alone as well. To find the solution for the central charge of the CFT, we now take the equation for the limiting behavior of the ground state energy (5.5.33) and solve for c, from which we define the *effective central charge* as

$$c(r) \equiv -\frac{6R}{\pi} E(m, R). \tag{5.5.34}$$

Notice that the effective central charge also *only* depends upon the value r, which refers to the system size in units of correlation length. For a fixed value of R, namely $R = R_0$, we use (5.5.33) and (5.5.34) to obtain

$$\lim_{r\to 0} c(r) = \lim_{m\to 0} -\frac{6R_0}{\pi} E(m, R_0) \tag{5.5.35}$$

$$= -\frac{6R_0}{\pi}\left(\lim_{m\to 0} E(m, R_0)\right) \tag{5.5.36}$$

$$= -\frac{6R_0}{\pi}\left(-\frac{c\pi}{6R_0}\right) \tag{5.5.37}$$

$$= c. \tag{5.5.38}$$

From the above derivation $\lim_{r\to 0} c(r) = c$ we may gauge if equations (5.5.28) and (5.5.29) yield viable solutions.

As we previously mentioned, the TBA will yield the correct ground-state energy for the model in question. This is as well as the TBA will do on its own. To obtain the higher energy levels, one must resort to another method - analytic continuation of the TBA.

Chapter 6

Analytic Continuation and the Excited States

As should be clear by now, the TBA equation only yields information on the ground state energy of the model in question. It can provide no other direct information regarding the excited states of any applicable theory, which is an issue if one wants to find the exact energy spectrum of an integrable model.

Bender and Wu first used the concept of analytic continuation to investigate the spectrum of the anharmonic oscillator in [6]. Analytic continuation was further employed by Dorey and Tateo in [18] to obtain the excited states of the Yang-Lee model. Here we introduce the general method to find the excited states of a relativistic integrable quantum field theory via analytic continuation. We also make use of references [16] and [25].

6.1 General Idea

For a quantum field theory Hamiltonian, we have the Hamiltonian of the underlying CFT plus a perturbation,

$$H^{QFT} = H^{CFT} + V, \tag{6.1.1}$$

where V is the perturbation. A toy example for a QFT Hamiltonian is

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{6.1.2}$$

which is of the form above where the first term on the RHS models the Hamiltonian of the underlying CFT and the second term on the RHS models the perturbation.

If we solve this Hamiltonian as an eigenvalue problem, we obtain for the energy

$$E_{\pm} = \pm\sqrt{1+\lambda^2}, \tag{6.1.3}$$

where $\lambda \in \mathbb{R}$.

Let us now examine the plot of equation (6.1.3). Referring to Fig. 6.1, we clearly see two distinct and disconnected solutions, which is expected given our solution set.

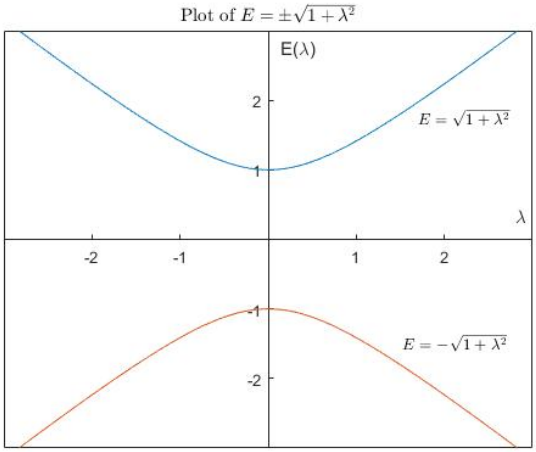


Figure 6.1: Image showing disconnected energy levels of square-root solution to an eigenproblem.

Now we will make a slight modification to equation (6.1.3) by taking an analytic continuation in the complex plane of the parameter λ, i.e. $\lambda \to i\lambda$, which yields

$$E_{\pm} = \sqrt{1-\lambda^2}. \tag{6.1.4}$$

Let us plot this continued solution, as we did before for the disconnected case. We now notice in Fig. 6.2 that instead of two disconnected energy levels, we have a single energy level. Thus we discovered that by an analytic continuation, we are able to reach the other eigenvalue of the problem with a smooth continuous path that would have been unreachable with $\lambda \in \mathcal{R}$. This is, of course, if the analytic continuation exists in the first place, as we have previously stated. Thus with an analytic continuation, see Fig 6.3, we can begin with the ground state and end up with the excited state (other eigenvalue of the Hamiltonian) by taking a smooth continuous path around the branch point.

The takeaway message from this realisation is that in eigenvalue problems, continuing some parameter around a closed contour returns us to the same problem, but not

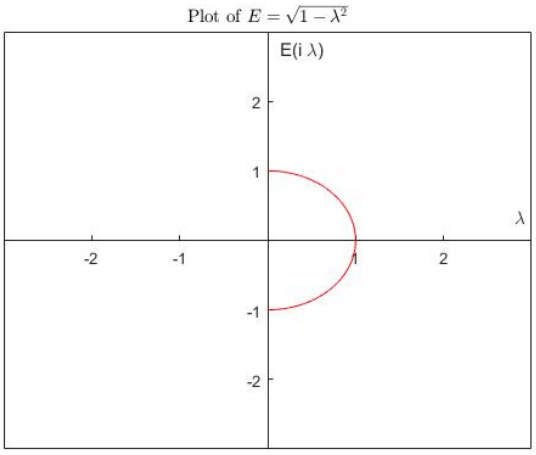


Figure 6.2: Image showing disconnected energy levels of square-root solution to an eigenproblem that become connected after an analytic continuation.

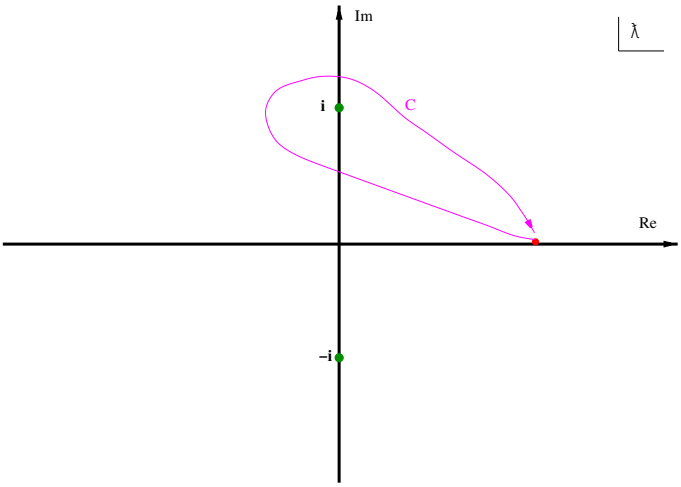


Figure 6.3: Demonstration of analytic continuation in eigenvalue problems. Continuing around the path C swaps the eigenvalue from E_0 to E_1.

necessarily the same eigenvalue. Obviously the problem did not change as we still are dealing with an eigenvalue problem and the same Hamiltonian. So if our analytically continued eigenvalue is different, then it must be another eigenvalue of the original problem.

In finite dimensions, this notion is clearly correct. In quantum mechanics and QFT, we are dealing with Hilbert spaces and Fock spaces, which are infinite dimensional. Thus one might question the validity of this idea when applying it to a QFT. Though for perturbed conformal field theories where the truncated conformal space approach

yields good finite-dimensional approximations to the finite-dimensional problem indicates that this property should hold. Note that we may not be able to obtain all eigenvalues this way, as we may find ourselves in situations where the Riemann surface has disconnected sectors. Later in Chapter 9 we will further discuss the subtleties associated with continuing the boundary conditions.

6.2 Analytical Outline

Now that we have the general idea behind our method of obtaining excited states, we shall illustrate what happens analytically via the same simple eigenvalue problem,

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{6.2.1}$$

where $\lambda \in \mathbb{C}$,

Solving the eigenvalue equation for (6.2.1) yields

$$E_{\pm}(\lambda) = \pm\sqrt{1+\lambda^2}. \tag{6.2.2}$$

The branch points for the above energy equation are located at $\pm i$, so we will choose to continue around the branch point $+i$ for this example. To continue in λ we make the definition

$$\lambda = i\left(1 - re^{i\theta}\right), \tag{6.2.3}$$

where r will be the radius of our circular path around the branch point and θ is the angle between r and the imaginary axis. This can be seen clearly in Fig. 6.4. Substituting our λ definition into (6.2.2) yields

$$E = \pm\sqrt{2re^{i\theta}} = \pm\sqrt{2r}e^{i\frac{\theta}{2}}, \tag{6.2.4}$$

where we make the approximation that $r \ll 1 \implies 1+\lambda^2 = 1-i\left(1-re^{i\theta}\right)^2 \approx 2re^{i\theta}$. Using (6.2.5), we will start with the negative solution as our ground state energy. Obviously we have that $\theta = 0$ for the initial value before we performed any analytic continuation. Thus for our ground state energy we have

$$E_- = -\sqrt{2r}e^{i\frac{\theta}{2}} = -\sqrt{2r}e^{0} = -\sqrt{2r}. \tag{6.2.5}$$

This result shouldn't be surprising since we started with E_- and didn't actually continue yet ($\theta = 0$), thus we should stay where we started. Now we endeavour to analytically continue in a circular path and end at the point $\theta = 2\pi$ to obtain

$$E_\gamma = -\sqrt{2r}e^{i\frac{\theta}{2}} = -\sqrt{2r}e^{i\pi} = \sqrt{2r} = E_+, \tag{6.2.6}$$

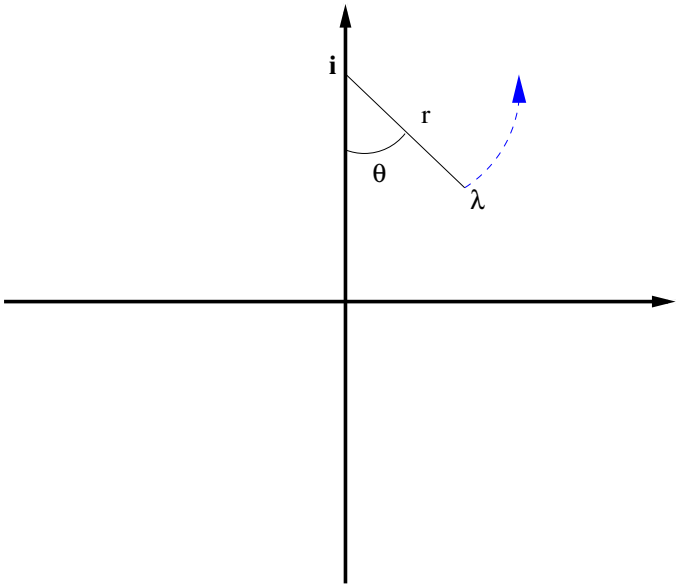


Figure 6.4: Figure showing analytic continuation in the λ-plane around the branch point i. Note that we keep r fixed and small, and we continue θ from 0 to 2π. The blue broken line is the continuation path.

where E_γ is the continued energy. Thus we clearly see that analytically continuing in a full circular path where the path $\gamma \in [0, 2\pi]$ will pick out the other eigenvalue in this simple eigenvalue problem when starting our continuation from the ground state. This illustrates the idea behind our method for a toy example.

6.3 The TBA and Analytic Continuation

6.3.1 Singularities of Integrals

As mentioned in the last chapter, the TBA equations are a set of integral equations. As part of our aim, we fully intend to analytically continue the TBA equations throughout the Riemann surface for $c(r)$ determined by the S-matrix of a given model. Note that we initially determine $c(r)$ for real r. A natural question to ask is whether a function defined by an integral can have a singularity and in which circumstances may it arise. We now proceed toward a general discussion of this behavior, and only later apply this knowledge to the TBA equations.

We begin by defining $g(z, w)$, which we allow to be an analytic function of two complex variables. We also allow for a finite contour $\mathcal{C}$ to exist within the complex

w-plane. We proceed to define a function $f(z)$ via

$$f(z) = \int_{\mathcal{C}} g(z, w)dw. \tag{6.3.1}$$

We assume that the singularities of the integrand $g(z, w)$ are known and that their locations in the w-plane are given by

$$w = w_r(z), \tag{6.3.2}$$

where $r = 1, 2, ...$ We now make one further presumption which is that for points z within a neighborhood of a point z_0, there exists a neighborhood of the contour $\mathcal{C}$ in the w-plane that does not contain the singularities w_r. Accepting these assumptions, our definition in (6.3.1) implies $f(z)$ is analytic at the point z_0. We now focus on how a singularity of $f(z)$ may originate when the function is analytically continued away from z_0.

We start by moving z away from z_0. As we do, the singularities w_r will move around in the w-plane. The function $f(z)$ will remain analytic until one of the singularities reaches the contour $\mathcal{C}$, at which point we are left with an undefined integral. At first, this appears to be a problem as it seems like analytic continuation may only be carried out in any direction until a singularity reaches the contour. Fortunately this is *not* the case, and we can often continue the function further.

We now allow our contour $\mathcal{C}$ to have endpoints A and B, which are initially fixed. We also assume we have another contour $\mathcal{C}'$ which shares the identical endpoints. If the region bounded by the two curves is free of singularities w_r then we have that

$$f(z) = \int_{\mathcal{C}'} g(z, w)dw \tag{6.3.3}$$

by consequence of Cauchy's theorem. Thus the strategy is that as w_r approaches the contour $\mathcal{C}$, we switch to a different contour $\mathcal{C}'$. The singularity will avoid this contour, at least for a little bit, and we will end up with the analytic continuation of the original function given in (6.3.1). This process can then be repeated as necessary.

Now it was previously mentioned that we may 'often' continue the function further, and we outlined a strategy on how this is accomplished. This naturally causes us to wonder in which cases we are prevented from such continuation. It turns out that there are three such cases:

1. *Endpoint Singularities* - Occurs if one of the singularities w_r reaches one of the endpoints of the contour. Since the endpoints must remain fixed for our contour switching strategy to work, we are not able to vary the contour in such a way as to avoid the singularity. The corresponding point, say z_1, may thus be a singularity of the function $f(z)$.

2. *Pinch Singularities* - Occurs when we have two (or more) singularities that approach the contour from opposing directions and coincide. The contour will thus inevitably become trapped between these two singularities with no way for the contour to escape this fate. Note that one of the singularities could be fixed, while the other approaches it. The important aspect is that the contour is 'pinched' between the singularities. The corresponding point, say z_2, may then be a singularity of the function $f(z)$.

3. *Infinite Deformations* - Occurs when the contour $\mathcal{C}$ is being deformed to avoid a singularity w_r. During the deformation, if w_r moves off toward infinity and drags $\mathcal{C}$ with it, the integral may diverge since $\mathcal{C}$ is no longer finite. Note that this case can be reduced to a special case of the pinch singularity by making a change of variables.

In our initial discussion, we assumed that our contour $\mathcal{C}$ had fixed endpoints. There is nothing fundamentally different if the endpoints were functions of z. Furthermore the contour may be closed and have no endpoints in the first place, in which case *only* pinch singularities will be encountered. We may further drop the restriction that $\mathcal{C}$ is finite, as coordinate transformations demonstrate that the point at infinity is akin to any other point.

6.3.2 Analysis of the Singularities

The singularities we encountered in the last section are usually branch points of the function $f(z)$, to which one may attach a branch cut. In some simple cases one may determine the discontinuity of $f(z)$ across a given cut. This is especially true when the singularities w_r of the integrand $g(z, w)$ are poles. We now endeavour to illustrate how one handles the discontinuity of $f(z)$ across a given branch cut by using the Residue theorem to recover (6.3.1), which is integrated over the original contour $\mathcal{C}$, from (6.3.3), which is integrated over $\mathcal{C}'$. Note that the branch points in our problem arise from the pinch singularity and not the endpoint singularity, as we are about to further describe.

We begin with the assumption that endpoint A is fixed. We now identify $z = z_1$ as an endpoint singularity that is produced by w_1. This is the statement that

$$w_1(z_1) = A. \tag{6.3.4}$$

If $w_1(z)$ is analytic at z_1, then it is equal to its Taylor series in an open interval, and we obtain

$$w(z) = A + (z - z_1)\frac{dw_1}{dz}\Big|_{z=z_1} \tag{6.3.5}$$

where we also assumed that $w_1(z)$ has a non-zero derivative at the point z_1.

If we choose a new point $z = z_1 + \epsilon$ in close proximity to z_1, then the point $w_1(z)$ is in close proximity to point A. This modifies our Taylor series to

$$w(z_1 + \epsilon) = A + \epsilon \frac{dw_1}{dz}|_{z=z_1}. \tag{6.3.6}$$

Now if we choose our path z to form a complete circle around the singularity z_1, i.e. $z = z_1 + |\epsilon|e^{i\theta}$ with θ varying from 0 to 2π, the path for w_1 will describe a similar circle of radius $\left|\epsilon \frac{dw_1}{dz}\right|$ around A. It is imperative that the contour is distorted, as it must avoid the singularity w_1 as it traces out its circular path. This situation is described in Figure 6.5. The endpoint singularity illustration is a good way to understand how branch points arise.

Our initial situation in the w-plane, before the circle is traced out, is given by Fig. 6.5*a*. In Fig. 6.5*b* we see how the contour is distorted to avoid having the singularity w_1 run into the contour. If we apply Cauchy's theorem to this distorted contour, the situation is now equivalent to that shown in Fig. 6.5*c*. As we see, we now have the original straight-line contour plus a new circular contour around w_1.

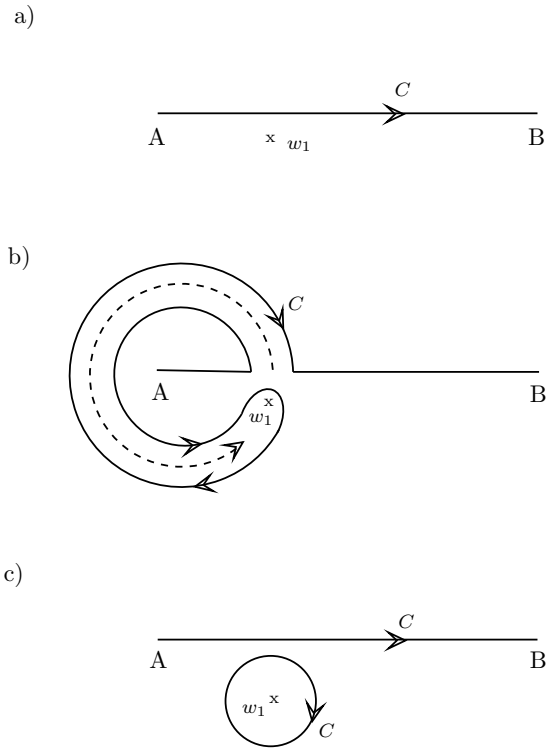


Figure 6.5: Representation of the movement of a singularity across the contour. This is a singularity $w_1(z)$ of the integrand of (6.3.3) when the variable z encircles an endpoint singularity z_1.

So returning to our original function (6.3.1), we can mathematically express the

above description as follows:

$$f_\gamma(w_1(z)) = \int_{\mathcal{C}'} g(z_1, w_1)\, dw \tag{6.3.7}$$

$$= \int_{\mathcal{C}} g(z_1, w_1)\, dw + \oint_{w_1(z)} g(z_1, w_1)\, dw \tag{6.3.8}$$

$$= \int_{\mathcal{C}} g(z_1, w_1)\, dw + 2\pi i \operatorname{Res}_f(g_{z_1}; w_1(z_1)). \tag{6.3.9}$$

where f_γ denotes the function f analytically continued along the path γ and Res refers to the *residue*, i.e. the complex number proportional to the contour integral of a meromorphic function along a path enclosing one of its singularities. The general situation can be expressed via

$$f_\gamma(w_1(z)) = \int_{\mathcal{C}} g(z_1, w_1)\, dw + 2\pi i \cdot \sum_{r=1}^{n} \operatorname{Res}_f(g_{z_r}; w_r(z_r)) \cdot \operatorname{Ind}_\gamma(w_r(z_r)). \tag{6.3.10}$$

This general situation now contains the *index*, given by Ind in the equation. The index simply counts the number of times the path completes a 'full circle' around the singularity in question, and therefore must be an integer. Note that 'full circle' refers to any path, no matter how jagged or wobbly, that eventually completes a full counterclockwise path (θ from 0 to 2π) around a given singularity. Since we pick up the same discontinuity z_1 each time, we know that the $z = z_1$ singularity of $f(z)$ is logarithmic. Now we apply this reasoning to the TBA equations. Before we do, we would like to demonstrate this theory in action.

6.3.3 Free Fermion Ising Model

As an example of how this actually works in practice, we will use the simple but illuminating example of the free fermion Ising model. We will endeavour to study the off-critical Ising model in the thermal direction. We utilize the reference by Klassen and Meltzer in [23]. There is a single mass m and the correlation length of the model is given by $\xi = \frac{1}{m}$.

The ground-state energy on the circle in closed form is

$$E_0(m, R) = E_{bulk} - \frac{\pi}{6R} c^{Ising}(mR) \tag{6.3.11}$$

The central charge of the Ising field theory is given by

$$c^{Ising}(r) = \frac{1}{2} - \frac{3r^2}{2\pi^2}\left[\log\frac{1}{r} + \frac{1}{2} + \ln\pi - \gamma_E\right] + \frac{6}{\pi}\sum_{k=1}^{\infty}\left(\sqrt{r^2 + (2k-1)^2\pi^2} - (2k-1)\pi - \frac{r^2}{2(2k-1)\pi}\right), \tag{6.3.12}$$

where $\gamma_E \approx 0.57721566$ is the Euler-Mascheroni constant and $E_{bulk} \sim R^2 \log R$. For $c(0) = \frac{1}{2}$, note that we have

$$E_0(R \to 0) = \frac{2\pi}{R}\left(h + \bar{h} - \frac{c}{12}\right), \tag{6.3.13}$$

which is correct for the ground state where h and $\bar{h}$ are the conformal dimensions and $h + \bar{h} = 0$ in the $c_{1/2}$ CFT, which is found as $R \to 0$.

Now referring to (6.3.12), we start on the real axis and analytically continue in the complex plane around the branch points for the Ising field theory, the branch points are located at $(2k-1)\pi i$, and they are of square root form. Thus if we take a path around k_1, k_2, etc., this action flips the signs of some of the square roots in the infinite sum of square roots in (6.3.12) from positive to negative.

Thus the basic idea is that the ground state energy is given via

$$E_0(M, R) = E_{bulk}(M, R) - \frac{\pi}{6R}c_0(r). \tag{6.3.14}$$

When some of those square roots in (6.3.12) flip from positive to negative, the energy in (6.3.14) is actually increased. Thus the exited state is of the form

$$E_1(M, R) = E_{bulk}(M, R) - \frac{\pi}{6R}c_1(r). \tag{6.3.15}$$

Note that the relationship between E_0 and c_0 and between E_1 and c_1 are equal.

Thus upon our return to the real axis, from which we left to analytically continue around the branch point, we obtain

$$E_{k_1,k_2,\ldots} = E_0(m, R) + \frac{2}{R}\sum_{i=1}^{n}\sqrt{r^2 + (2k-1)^2\pi^2}. \tag{6.3.16}$$

Note that (6.3.16) is an excited state with higher energy than the ground state. Thus this model illustrates the idea of how analytic continuation in the complex plane can yield information about the excited state of an integrable model. We now focus on the analytically continuing the TBA and the nuances involved.

6.3.4 The Analytic Continuation of the TBA Equations

Recall that at the end of the chapter on conformal field theory we discussed the CFT on a cylinder. This was the exact same shape we ended up with when we first employed the transfer matrix method for a QFT defined on a 2-D lattice (recall that the torus with $L \to \infty$ was a cylinder of circumference R). The ultimate aim was to find the analytic continuation of the effective central charge $c(r)$ from CFT. By locating a branch point of the TBA equations and continuing c(r) along a path that

encircles the singularity, we should be able to obtain $c(r)'$ - a new effective central charge which corresponds to the first excited state. To obtain its energy, we can use (5.5.34) to calculate the second excited state. The process will then be repeated starting from the first excited state.

We begin by noting that the effective central charge is actually defined via an integral equation, which we can see via substituting (5.5.29) into (5.5.34) to obtain

$$c(r) = \frac{3}{\pi^2}\int_{-\infty}^{\infty} r\cosh\theta' \log\left(1+e^{-\epsilon(\theta')}\right)d\theta'. \tag{6.3.17}$$

This equation for $c(r)$ depends upon the solution ϵ of the TBA equation, so we must inevitably analytically continue the TBA equations. Noting that $r = mR = \gamma(t)$, we see that r is the path which we set, and by which we ultimately analytically continue the TBA.

To see how to do this, we proceed by making the substitutions $z \to mR$ and $w \to \theta$ in the previous section. We now see that we can apply our former reasoning to our current problem. We select a particular path r and solve the TBA equation for $\epsilon(\theta)$, using this TBA solution in (6.3.17) to obtain the effective central charge.

Unlike in the previous section, however, we now are forced to solve for the TBA equation at every step along the path. There is simply no guarantee that if we have a TBA solution ϵ and continue it along a closed path, we will end up with the same solution once we return to the initial starting point. The TBA is *not* static and ϵ could have undergone non-trivial monodromies.

The Continued TBA

In the previous chapter, the TBA equations were given by

$$-m_a R\cosh\theta + \epsilon_a + \sum_{b=1}^{N}\phi_{ab} * L_b = 0. \tag{6.3.18}$$

To simplify our discussion, let us make the substitution $mR \to r$ and assume we only have one type of particle, which is true for the sinh-Gordon model. We now have a TBA equation which reads

$$\epsilon(\theta) = r\cosh\theta - \phi * L(\theta). \tag{6.3.19}$$

Now we endeavour to solve (6.3.19) for some r. We wish to study how (6.3.19) changes as r varies along a path in the complex plane. Recall that $r = mR = m(\lambda)R$ where λ is the *coupling constant*. Thus in reality we will be implicitly continuing the coupling constant in the complex plane. We define a general path γ which admits

as its argument a set of points t, i.e. $\gamma(t)$. With an idea of what is to come, we will refer to these points as 'steps' along the path γ.

Now we want to analytically continue the TBA equation in r. Thus we set $r = \gamma(t)$ and continue the TBA equation, solving it at every step along the path. We will end up with a solution to the TBA, i.e. the function $\epsilon(\theta)$, which we assume to be *entire* on the complex θ-plane (holomorphic on the complex plane).

From our definition of L in the previous chapter, we now choose our variant of (5.5.27) with respect our problem to obtain

$$L(\theta) = \log\left(1 + e^{-\epsilon(\theta)}\right). \tag{6.3.20}$$

Now (6.3.20) has two conditions which cause singularities to appear:

1. $e^{\epsilon(\theta)} = -1$,

Given the symmetry of the TBA, i.e. $\theta \to -\theta$, singularities will always come in pairs $\{\theta_0, -\theta_0\}$. As we analytically continue our TBA via a path r of our choosing in the complex plane, these singularities in general avoid the real axis and present us with no difficulty.

For some paths the pair of singularities can actually approach the real axis as r is varied. For such a case the convolution

$$(\phi * L(\theta))_{\mathbb{R}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(\theta - \theta')\log\left(1 + e^{-\epsilon(\theta')}\right) d\theta' \tag{6.3.21}$$

is threatened. Note that the limits of integration indicate that the convolution contour is computed over the real axis. Refer to Figure 6.6*a*. We can avoid this issue to an extent by simply deforming the contour. As the singularities approach the real line, we can deform the contour so that they do not run into the real line. As long as the singularities do not cross the real line, deforming the contour allows us to analytically continue the same TBA equation. We simply now solve the TBA via a modified convolution contour. This modifies the TBA slightly to

$$-r\cosh(\theta) + \epsilon + (\phi * L)_{\mathcal{C}} = 0, \tag{6.3.22}$$

where $(\phi * L(\theta))_{\mathcal{C}}$ denotes the convolution along the deformed contour given by

$$(\phi * L(\theta))_{\mathcal{C}} = \frac{1}{2\pi}\int_{\mathcal{C}} \phi(\theta - \theta')\log\left(1 + e^{-\epsilon(\theta')}\right) d\theta', \tag{6.3.23}$$

where the region enclosed by $\mathcal{C}$ and the real axis is free from zeros of $L(\theta)$. As the enclosed region has no zeros, we may use Cauchy's theorem to find

$$(\phi * L(\theta))_{\mathcal{C}} = (\phi * L(\theta))_{\mathbb{R}}. \tag{6.3.24}$$

Thus we see that the modified convolution contour and the original convolution contour are equivalent. Distorting the contour does *not* change our problem, and the results obtained for (6.3.19) and (6.3.22) are identical and both yield the correct analytic continuation of the TBA equation. We may thus stick with (6.3.19), i.e. the unmodified contour. Thus we have shown deforming the contour is completely valid and we can continue to use the same TBA equation we started with initially. As long as the singularities do not cross the real axis, we can simply deform the contour and continue on with our work.

Now if our chosen path causes the singularities to actually cross the real axis, as

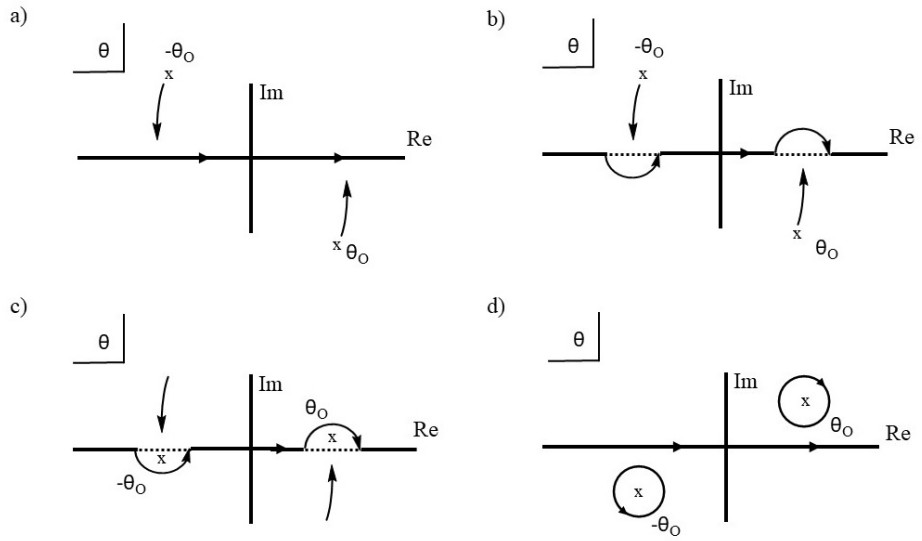


Figure 6.6: Representation of the analytic continuation of the TBA.

in Figure 6.6*b* and *c*, we now must use (6.3.22), as the modified contour is now no longer equivalent to (6.3.19). To recover the original convolution contour, we must pick up the residues explicitly. To see how this is done, recall that

$$\phi = -i\frac{d}{d\theta}\log(S). \tag{6.3.25}$$

We begin with the modified convolution contour

$$(\phi * L)_{\mathcal{C}} = \frac{1}{2\pi}\int_{\mathcal{C}} \phi(\theta-\theta')L(\theta')d\theta'. \tag{6.3.26}$$

Noting that for large θ, the $\cosh(\theta)$ term of the TBA dominates. Thus asymptotically $\epsilon = r\cosh(\theta)$. Focusing on the case that Re $r > 0$, ϵ will grow larger and so $L = \log\left(1+e^{-\epsilon}\right) \to 0$. Now when we integrate (6.3.26) by parts, we obtain

$$(\phi * L)_{\mathcal{C}} = \frac{1}{2\pi i}\int_{\mathcal{C}} \log(S(\theta-\theta'))\frac{\epsilon'(\theta')}{1+e^{\epsilon(\theta')}}d\theta' \tag{6.3.27}$$

$$= \frac{1}{2\pi i}\int_{w(\theta)} \log(S(\theta-\theta'))\frac{\epsilon'(\theta')}{1+e^{\epsilon(\theta')}}d\theta' \tag{6.3.28}$$

where ϵ' is assumed to be an entire function, i.e. it contains no singularities at any point in $\mathbb{C}$, and $\int_{w(\theta)}$ refers to the singularity integral which we solve via the Residue Theorem to obtain

$$(\phi * L)_{\mathcal{C}} = \frac{1}{2\pi i}\int_{\mathbb{R}} \log(S(\theta-\theta'))\frac{\epsilon'(\theta')}{1+e^{\epsilon(\theta')}}d\theta' + \text{Res}(f;-\theta_0)) - \text{Res}(f;\theta_0) \tag{6.3.29}$$

where f is the integrand

$$f : \theta' \mapsto \left(\frac{\log(S(\theta-\theta'))\epsilon'(\theta')}{1+e^{\epsilon(\theta')}}\right). \tag{6.3.30}$$

We now need to determine the residues in (6.3.29). In (6.3.30) both the numerator and denominator are holomorphic in a neighborhood of $\pm\theta_0$, and there exists a non-zero derivative of the denominator of $\pm\theta$. Thus we may compute the residue $Res(f,\theta_0)$ via the relation

$$Res(f,\theta_0) = \frac{h(\theta_0)}{g'(\theta_0)}, \tag{6.3.31}$$

where $h(\theta_0)$ is the numerator of f evaluated at θ_0 and $g'(\theta_0)$ is the derivative of the denominator of f evaluated at θ_0. For f we see that we have a singularity when the denominator is equal to zero. We thus have $e^{\epsilon(\theta_0)} = -1$ from before, the knowledge that the TBA is symmetric ($\theta = -\theta$) so $e^{\epsilon(-\theta_0)} = e^{\epsilon(\theta_0)}$. Using this to calculate the residue at θ_0 yields

$$Res(f;\theta_0) = \frac{\log(S(\theta-\theta_0))\epsilon'(\theta_0)}{\epsilon'(\theta_0)e^{\epsilon(\theta_0)}} \tag{6.3.32}$$

$$= \frac{\log(S(\theta-\theta_0))}{-1} \tag{6.3.33}$$

$$= -\log(S(\theta-\theta_0)). \tag{6.3.34}$$

Now calculating the residue at $-\theta_0$ yields

$$Res(f;-\theta_0) = \frac{\log(S(\theta-(-\theta_0)))\epsilon'(-\theta_0)}{\epsilon'(-\theta_0)e^{\epsilon(-\theta_0)}} \tag{6.3.35}$$

$$= \frac{\log(S(\theta+\theta_0))}{-1} \tag{6.3.36}$$

$$= -\log(S(\theta+\theta_0)). \tag{6.3.37}$$

Now plugging (6.3.34) and (6.3.37) into (6.3.29) yields

$$(\phi * L)_{\mathcal{C}} = (\phi * L)_{\mathbb{R}} + (-\log(S(\theta-\theta_0))) - (-\log(S(\theta+\theta_0))) \tag{6.3.38}$$

$$= (\phi * L)_{\mathbb{R}} + \log\left(\frac{S(\theta-\theta_0)}{S(\theta+\theta_0)}\right). \tag{6.3.39}$$

Plugging in (6.3.39) into the TBA equation with a modified convolution contour, i.e. (6.3.22), yields

$$-r\cosh(\theta)+\epsilon(\theta)+(\phi*L)_R(\theta)+\log\left(\frac{S(\theta-\theta_0)}{S(\theta+\theta_0)}\right)=0. \tag{6.3.40}$$

Note that we solve for the terms in (6.3.40) iteratively at fixed r until $\epsilon(\theta)$ and $e^{-\epsilon(\theta_0)}$ converge. Note also that here θ_0 must satisfy $e^{-\epsilon(\theta_0)}=-1$, which is obtained from $e^{\epsilon(\theta_0)}=-1$ by dividing through by $e^{\epsilon(\theta_0)}$. In Chapter 8, we solve (6.3.40) iteratively at fixed r until $\epsilon(\theta)$ and $Y(\theta_0)$ converge. For another pair of singularities crossing the real axis, we start with (6.3.40) and perform the exact same steps found in Figure 6.6 for the new pair of singularities to obtain

$$-r\cosh(\theta)+\epsilon(\theta)+(\phi*L)_R(\theta)+\log\left(\frac{S(\theta-\theta_0)}{S(\theta+\theta_0)}\right)+\log\left(\frac{S(\theta-\theta_1)}{S(\theta+\theta_1)}\right)=0, \tag{6.3.41}$$

where θ_0 refers to the first pair of singularities, and θ_1 the second pair. This process repeats itself for all subsequent pairs of singularities to yield the general result

$$-r\cosh(\theta)+\epsilon(\theta)+(\phi*L)_R(\theta)+\sum_i\log\left(\frac{S(\theta-\theta_i)}{S(\theta+\theta_i)}\right)=0. \tag{6.3.42}$$

6.3.5 The Analytic Continuation of c(r)

Recall that ultimately we want the central charge $c(r)$. Its analytic continuation is obtained by integrating

$$c(r)=\frac{3}{\pi^2}\int_{-\infty}^{\infty}r\cosh\theta'\log\left(1+e^{-\epsilon(\theta')}\right)d\theta'. \tag{6.3.43}$$

along the distorted contour. Note that this equation is exactly the same as (6.3.17). Performing a residue calculation quite similar to the ones we have done in detail for the TBA, and we obtain

$$\begin{aligned} c(r)&=\frac{3}{\pi^2}\int_{\mathcal{C}}r\cosh\theta' L(\theta')d\theta' && (6.3.44)\\ &=\frac{12r}{\pi}i\sinh(\theta_0)+\frac{3}{\pi^2}\int_{\mathbb{R}}r\cosh(\theta')L(\theta')d\theta'. && (6.3.45)\end{aligned}$$

6.3.6 Analytic Continuation of c(r) Summary

The entire process to obtain the analytic continuation of the central charge $c(r)$ via the analytic continuation of the TBA is as follows:

1. First locate a branch point of the analytically continued effective central charge $c(r)$. These are of the pinch-type, as introduced previously. Two singularities

w_r will converge toward the same point, moving toward each other from opposite sides of the contour. This will essentially 'pinch' the contour and create a branch point.

2. Choose a path r in $\mathbb{C}$ that admits an initial $r > 0$, on which one will continue $c(r)$ around the branch point and back to the initial value of the path.

3. One must solve the TBA at every step along this chosen path, modifying it as necessary.

4. Once one transverses the path and arrives back at the same initial starting point, one will have obtained the first excited state of $c(r)$.

5. Further excited states are obtained by taking the previous excited state to be the ground state and repeating the process.

This chapter was essentially the heart of the technique - how to use the TBA equation to obtain the effective central charge for excited states. To achieve this analytic continuation, we are actually required to extend the function ϵ, i.e. the solution of the TBA, to the entire complex plane. This is model-dependent and will be explained in the next section in detail when we apply this technique to the sinh-Gordon model.

Chapter 7

Sinh-Gordon Model

We now are ready to apply the formalism we developed to a physical model, and in our case we chose the sinh-Gordon model (ShG). The sinh-Gordon model is an interacting, integrable quantum field theory. The particles involved are bosons. It was originally quantized exactly via the quantum inverse scattering method. It represents a model of suitably perturbed non-compact sigma models of considerable interest in string theory [32]. This section primarily makes use of the paper by Alexei Zamolodchikov found in [32], and also includes some insight from [25].

7.1 Derivation and Definitions

We begin by considering the real scalar relativistic field ϕ in 1+1-dimensional Minkowski spacetime that satisfies the equation of motion

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} + \sinh b\phi = 0. \tag{7.1.1}$$

This equation is known as the sinh-Gordon equation, which is defined by the Euclidean action

$$S_{ShG} = \int \left[\frac{1}{4\pi}(\partial_\nu\phi)^2 + 2\mu\cosh(2b\phi)\right] d^2x \tag{7.1.2}$$

where μ is a dimensional coupling constant of dimension $\mu \sim [mass]^{2+2b^2}$, b is the dimensionless ShG parameter. Note that for this section we are referring to the previous work of Zamolodchikov in [32], which we take as input in our study without further derivation or explanation.

We restrict our study to real, positive b-values, i.e. $0 < b < \infty$. Since $\cosh(-\theta) = \cosh(\theta)$, the negative range for the parameter is already covered by studying to the positive range. Furthermore at $b = 0$, the result turns out to be a bit singular. The

$\cosh(\theta)$ term vanishes, and one is left with a quadratic equation that can be solved via the Euler-Lagrange equation for fields. The result is merely the wave equation.

Returning to our choice of positive b-values for our domain, we choose to define and use different parameters for the model. This choice is made based on the knowledge of what will come when analyzing the model. We define

$$p = \frac{b^2}{1+b^2} \qquad \text{and} \qquad a = 1 - 2p \tag{7.1.3}$$

and will henceforth use them in place of b.

The perturbing operators $e^{\pm 2b\phi}$ in the action given in (7.1.2) have negative dimension $\Delta = -b^2$. We must fix a normalization of these operators to make the coupling μ a strict sense. The operators in our normalization are implied to be normal ordered with respect to the unperturbed (massless) vacuum such that in the unperturbed theory we have

$$\langle e^{2b\psi(x_1)} ... e^{2b\psi(x_n)} e^{-2b\psi(y_1)} ... e^{-2b\psi(y_n)} \rangle_{\mu=0} = \frac{\prod_{i,j} |x_i - y_j|^{4b^2}}{\prod_{i>j} (|x_i - x_j||y_i - y_j|)^{4b^2}}. \tag{7.1.4}$$

This is a known integrable model, and can be solved for many important characteristics. The spectrum has only one neutral particle subject to a factorized scattering. This yields a rather simple two-particle S-matrix of the form

$$S(\theta) = \frac{\sinh(\theta) - i\pi p}{\sinh(\theta) + i\pi p}. \tag{7.1.5}$$

With our normalization (7.1.4) the mass of this particle is related to the scale parameter μ as

$$\pi\mu \frac{\Gamma(b^2)}{\Gamma(1-b^2)} = [mZ(p)]^{2+2b^2} \tag{7.1.6}$$

where

$$Z(p) = \frac{1}{8\sqrt{\pi}} p^p (1-p)^{1-p} \Gamma\left(\frac{1-p}{2}\right) \Gamma\left(\frac{p}{2}\right). \tag{7.1.7}$$

Note that the scattering theory happens to be invariant under the weak-strong coupling duality transformation $b \to 1/b$, which from (7.1.3) means $p \to 1-p$. This ensures that the physical content of the model is invariant up to the overall mass scale. Since (7.1.7) is also invariant under $p \to 1-p$, the mass scale also remains invariant if the coupling constant μ is simultaneously substituted by the 'dual' coupling constant $\tilde{\mu}$ which is related to μ by

$$\left(\pi\mu \frac{\Gamma(b^2)}{\Gamma(1-b^2)}\right)^{\frac{1}{b}} = \left(\pi\tilde{\mu} \frac{\Gamma\left(\frac{1}{b^2}\right)}{\Gamma\left(1-\frac{1}{b^2}\right)}\right)^{b}. \tag{7.1.8}$$

From this we see that the sinh-Gordon model is invariant under the duality $b \to \frac{1}{b}$ and

$\mu \to \tilde{\mu}$. Due to this symmetry it is sufficient to only consider the region $0 < b^2 \leq 1$ or $0 < p \leq \frac{1}{2}$. Note that a is simply reflected $a = -a$ under $p \to 1 - p$ and therefore can be taken to be non-negative, i.e. $0 \leq a < 1$.

7.1.1 TBA Equation Application - The X- and Y-Functions

As mentioned in the last chapter, we must extend ϵ to the whole complex plane. As this is model-dependent, in order to understand how to implement the analytic continuation of the TBA equation, we now show how this is accomplished for the sinh-Gordon model.

Noting the structure of the TBA equation,

$$mR\cosh(\theta) = \epsilon(\theta) + \phi * \log\left(1 + e^{-\epsilon(\theta)}\right), \tag{7.1.9}$$

we are able to deduce that $\epsilon(\theta) = \epsilon(-\theta)$, i.e. the function $\epsilon(\theta)$ is even. It is also analytic in the strip $|\mathrm{Im}\,\theta| < \frac{\pi}{2} - \frac{\pi a}{2}$. This strip has the asymptotic $\epsilon(\theta) \sim \frac{mRe^{\theta}}{2}$ when $\mathrm{Re}(\theta) \to \infty$. Thus we may introduce and use the so-called Y-function,

$$Y(\theta) = e^{-\epsilon(\theta)}, \tag{7.1.10}$$

which is analytic and non-zero in this strip. At $\mathrm{Re}\,(\theta) \to \infty$ (7.1.10) behaves as

$$Y(\theta) = e^{\left(\frac{-mRe^{\theta}}{2}\right)}. \tag{7.1.11}$$

The symmetry $Y(\theta) = Y(-\theta)$ relates the asymptotic at $\mathrm{Re}(\theta) \to -\infty$ to (7.1.11).

Now in the original TBA equation, we have an integral representation that is valid for $|\mathrm{Im}(\theta)| < \pi p$. On the boundary of this strip, however, we encounter a singularity in $\phi(\theta - \theta')$, so we cannot use equation (7.1.9). We need a new strip from which we can analytically continue ϵ to all $\mathbb{C}$, following the work of Zamolodchikov in [32]. We now define the X-function as

$$X(\theta) = \exp\left(-\frac{mR}{2\sin(\pi p)}\cosh(\theta) + \int \frac{\log(1 + Y(\theta'))}{\cosh(\theta - \theta')}\frac{d\theta'}{2\pi}\right), \tag{7.1.12}$$

which is analytic and non-zero in the strip $|\mathrm{Im}(\theta)| < \frac{\pi}{2}$ and when $\mathrm{Re}(\theta) \to \infty$ in the strip, we have

$$X(\theta) \sim \exp\left(-\frac{mR}{4\sin(\pi p)}\exp(\theta)\right). \tag{7.1.13}$$

So the question becomes how do we make use of this X-function? By multiplying two instances of (7.1.12) together, albeit with modified arguments that are quite

similar for each instance of $X(\theta)$, we are able to obtain

$$X\left(\theta+\frac{ia\pi}{2}\right)X\left(\theta-\frac{ia\pi}{2}\right)=e^{-r\cosh(\theta)+\frac{1}{2\pi}\int_{-\infty}^{\infty}L(\theta')\phi(\theta-\theta')d\theta'} \tag{7.1.14}$$

$$=e^{-r\cosh(\theta)+(\phi*L)_{\mathbb{R}}(\theta)}, \tag{7.1.15}$$

where we made use of the identities

$$\cosh\left(\theta+\frac{ia\pi}{2}\right)+\cosh\left(\theta-\frac{ia\pi}{2}\right)=2\cosh(\theta)\cosh\left(\frac{ia\pi}{2}\right), \tag{7.1.16}$$

$$\cosh\left(\theta+\frac{ia\pi}{2}\right)\cosh\left(\theta-\frac{ia\pi}{2}\right)=\frac{\cosh(2\theta)+\cos(2\pi p)}{2}, \tag{7.1.17}$$

and

$$\sin(\pi p)=\cos\left(\frac{a\pi}{2}\right) \tag{7.1.18}$$

in what amounts to essentially a glorified algebra problem. Comparing (7.1.15) with the TBA equation yields

$$X\left(\theta+\frac{ia\pi}{2}\right)X\left(\theta-\frac{ia\pi}{2}\right)=e^{-\epsilon} \tag{7.1.19}$$

$$=Y(\theta). \tag{7.1.20}$$

With (7.1.20) we now have a relation between $X(\theta)$ and $Y(\theta)$, to which henceforth we refer to as the *XY-system*. On the real axis $Y(\theta)$ is real and positive, and thus we expect a strip $|\mathrm{Im}(\theta)|<\delta$ with some finite $\delta>0$ where $1+Y(\theta)\neq 0$. Thus the analyticity condition for $X(\theta)$ can be extended to the strip $|\mathrm{Im}(\theta)|<\frac{\pi}{2}+\delta$. We now proceed to compute the function

$$X\left(\theta+\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}\right). \tag{7.1.21}$$

In (7.1.12), recall that $L=\log(1+e^{-\epsilon})$. The logarithm remains well-defined within the strip $|\mathrm{Im}(\theta)|<\delta$. If we start with a path located within the strip $|\mathrm{Im}(\theta)|<\frac{\pi}{2}$ and analytically continue $X(\theta)$, we may obtain the function $X\left(\theta+\frac{i\pi}{2}\right)$. As we approach the singularity at the point $\theta+\frac{i\pi}{2}$, we must deform the contour of integration for $X(\theta)$ found in (7.1.12) so that θ does not collide with it. As long as the contour remains within the strip, the logarithm remains well-defined and this deformed contour remains equivalent to (7.1.12) via Cauchy's theorem. Note that $X(\theta)$ is also defined on the boundary as well, and thus represents the true analytic continuation of $X(\theta)$ to the boundary of the strip.

Using the fact that $\cosh\left(\theta+\frac{i\pi}{2}\right)=-\cosh\left(\theta-\frac{i\pi}{2}\right)$ and Cauchy's theorem we obtain

$$X\left(\theta+\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}\right)=e^{\int_{\mathcal{C}}\frac{L(\theta')}{\cosh\left(\theta+\frac{i\pi}{2}-\theta'\right)}\frac{d\theta'}{2\pi}}, \tag{7.1.22}$$

where $\mathcal{C}$ is a circular contour that runs clockwise around θ. Taking the residue of (7.1.22) gives

$$X\left(\theta+\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}\right)=1+Y(\theta). \tag{7.1.23}$$

Now using (7.1.20) and (7.1.23) we quickly obtain

$$X\left(\theta+\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}\right)=1+X\left(\theta+\frac{ia\pi}{2}\right)X\left(\theta-\frac{ia\pi}{2}\right) \tag{7.1.24}$$

This equation is henceforth known as the *X-system*. With (7.1.24) we now can extend the original analyticity strip beyond $|\mathrm{Im}(\theta)|<\frac{\pi}{2}$. In fact it is (7.1.24) that we can use to extend the original analyticity strip to the whole complex plane, so that $X(\theta)$ is an entire function of θ. It follows from (7.1.24) that the asymptotic (7.1.13) holds in the larger strip $|\mathrm{Im}(\theta)|<\pi$.

7.1.2 Extended Representation of $X(\theta)$ and the $Y(\theta)$-System

We mentioned that we can use (7.1.24) to extend the original analyticity of the strip to the whole complex plane. To accomplish this, we will have to derive a representation of the $X(\theta)$-system outside this strip. We begin by recalling (7.1.24):

$$X\left(\theta+\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}\right)=1+X\left(\theta+\frac{ia\pi}{2}\right)X\left(\theta-\frac{ia\pi}{2}\right). \tag{7.1.25}$$

For a representation of $X(\theta)$ above the initial analyticity strip (an '*upper shift*'), we make use of the shift $\theta\to\theta-\frac{i\pi}{2}$. We now have

$$X\left(\theta\right)X\left(\theta-i\pi\right)=1+X\left(\theta+(a-1)\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}(a+1)\right). \tag{7.1.26}$$

We now simply solve for the term $X(\theta)$ to obtain our *upper shift*, i.e.

$$X\left(\theta\right)=\frac{1+X\left(\theta+(a-1)\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}(a+1)\right)}{X\left(\theta-i\pi\right)} \tag{7.1.27}$$

To obtain the *lower shift* we instead use the shift $\theta\to\theta+\frac{i\pi}{2}$ in (7.1.24) to obtain

$$X\left(\theta\right)=\frac{1+X\left(\theta+(a+1)\frac{i\pi}{2}\right)X\left(\theta-\frac{i\pi}{2}(1-a)\right)}{X\left(\theta+i\pi\right)}. \tag{7.1.28}$$

This representation of $X(\theta)$ outside the initial analyticity strip is analytic in the new strip $|\mathrm{Im}(\theta)|<\frac{\pi}{2}+\pi p$, where $a=1-2p$. Recalling what we learned from Chapter 1 and using the Identity theorem, we are assured that (7.1.27) is the correct analytic continuation of (7.1.24). We can further use (7.1.27) to extend $X(\theta)$ to the strip

$|\mathrm{Im}(\theta)|< \frac{\pi}{2}$ simply by successively adding new strips of width πp.

Note that $Y(\theta)$ is also an entire function of θ. Given relations (7.1.20) and (7.1.23) which relate $X(\theta)$-system and the original $Y(\theta)$-system, we can derive the $Y(\theta)$-system equation for the sinh-Gordon model, which are quite similar to the functional relations appearing in the TBA study of integrable perturbed rational conformal field theories. We simply plug in (7.1.20) and (7.1.23) into (7.1.24) to yield

$$Y\left(\theta+\frac{i\pi}{2}\right)Y\left(\theta-\frac{i\pi}{2}\right)=\left(1+Y\left(\theta+\frac{ia\pi}{2}\right)\right)\left(1+Y\left(\theta-\frac{ia\pi}{2}\right)\right). \qquad (7.1.29)$$

This equation is known as the *Y-system*. Note that usually the existence of a Y-system implies a periodicity of the Y-functions in θ with some imaginary period related to the scale dimension of the perturbing operator, see [32] and [33]. In contrast, the Y-system for the sinh-Gordon, and even the X-system for that matter, do *not* imply a periodic structure of $Y(\theta)$ in θ.

All of this works together by allowing us to first find the TBA pseudoenergy $\epsilon(\theta)$ on the contour. We will then use it to find the *X-function* on the initial strip ($|\mathrm{Im}(\theta)|< \frac{\pi}{2}$). We then will use the *X-system* to extend the *X-function* to all of $\mathbb{C}$. We will then be in a position to reconstruct the *Y-function* using the $XY-system$, as after all the *Y-function* is directly defined from the TBA equation and needed for our results.

Having discussed the sinh-Gordon model specifically, we now turn to a discussion of its numerical study and our results.

Chapter 8

Computational Solution and Numerical Results for the Sinh-Gordon Model

In this chapter we address the numerical solution of the TBA equations and the computational procedure for its analytic continuation. To accomplish this task, we must first address the formulation of an iterative solution for the TBA and then set about finding branch points for the effective central charge c in the sinh-Gordon model. We then make use of these to obtain the excited state energies of the model. We also discuss the many regularities found in the model as $p \to 0$.

8.1 Numerical Solution of the Ground State TBA

Recall that the solution of the TBA equation is iterative. To solve it numerically, we make use of the iteration

$$\epsilon_{n+1} = (1-\lambda)\epsilon_n + \lambda F[\epsilon_n], \tag{8.1.1}$$

where

$$F[\epsilon_n] = r\cosh(\theta) - \frac{1}{2\pi}\int_{\mathcal{R}} \phi(\theta-\theta')\log\left(1+e^{-\epsilon(\theta')}\right)d\theta', \tag{8.1.2}$$

and where $0 < \lambda \leq 1$. Note that (8.1.1) corresponds to iterating the TBA equation with a damping factor λ added. We fix this constant in the iteration to maximize efficiency and to help with convergence. The iteration is more stable when λ is smaller, but the program runs far slower. We will have more to say about the damping in Section 8.5.4 to come. We then proceed to solve for a series of different

values of r, which contains the coupling constant g in the form $r = m(g)R$. We set $\epsilon^0 \equiv 0$ in the TBA equation, which enables us to obtain a solution for the TBA equation for this initial value of r. For the next value r, we utilize the solution ϵ obtained from solving the TBA for the initial value of r as the 'seed' for the next square of iterations. Using these parameters, we now have a subsequent solution for the TBA equation. For subsequent values of r, we utilize the solution of the TBA for the previous r-value, and use them as seeds to obtain a subsequent solution to the TBA equation. Analytic continuation is automatically implemented if the step size is small, and thus we don't need to implement a special procedure to ensure analytic continuation is carried out.

Note that for the naive iteration $\lambda = 1$, we have simply $F[\epsilon_n]$ on the RHS of (8.1.1), which is just the TBA equation. Now if we chose any value of λ, notice that when substituting that value in (8.1.1), we will obtain the full TBA equation on the RHS as well. Note also that for any fixed point of (8.1.1), it is also a solution to the TBA as well.

8.2 Branch Points

8.2.1 Off the Real Axis

Now that we know how to numerically solve the TBA equation and carry out analytic continuation, we now need to be able to find branch points for the model to obtain the excited states by crossing the branch point and travelling to another sheet of the Riemann surface. In Dorey and Tateo's paper [18], branch points of the central charge $c(r)$ were found in the half-plane of positive real values of r. Their work was for the Yang-Lee model. For the sinh-Gordon model, branch points were noticeably absent in this region. The situation where Re $r < 0$ is more difficult to treat, we encounter our first complication with this model.

For some practicalities, let us now consider $\pm\mathrm{Re}(\theta) \gg 1$. The driving term in the TBA, $\pm r\cosh(\theta)$, and the TBA itself acquire a kink form, i.e.

$$\pm\, r\cosh(\theta) \sim \frac{1}{2} r e^{\pm\theta} \qquad \text{and} \qquad \epsilon(\theta) = \frac{1}{2} r e^{\pm\theta} - \phi * L(\theta). \tag{8.2.1}$$

For real r, $\epsilon(\theta)$ has an accumulation of points as $|\mathrm{Re}(\theta)| \to \infty$ at which

$$Y(\theta) = e^{-\epsilon(\theta)} = -1 \tag{8.2.2}$$

all with $|\mathrm{Im}(\theta)| = \frac{\pi}{2}$. This follows from (7.1.11), and can be seen in figure 8.1(b) below in the accumulations of concentric circles along the lines $|\mathrm{Im}(\theta)| = \pm\pi/2$, indicating

zeros of the function

$$Z(\theta) \equiv 1 + Y(\theta)\,. \tag{8.2.3}$$

When r becomes complex with a positive argument, these points are shifted in the imaginary direction in the complex θ plane, downwards in the right half plane and upwards in the left half plane, as can be seen in figure 8.4. More precisely, if $r = |r|e^{i\gamma}$ with $\gamma > 0$ the asymptotic shift in the imaginary part is by $-\gamma$ to the right, and $+\gamma$ to the left. As a result, to solve the TBA equations when $\arg(r) > \frac{\pi}{2}$ one must also distort the integration contour away from the real axis at $\pm\infty$, since otherwise points at which $Y(\theta) = -1$ would hit the integration contour.

Let us begin by distorting the contour such that at the left of the imaginary axis it has a small positive imaginary value, which we call σ. Note that the distorted contour still crosses the imaginary axis at the origin. To the right of the imaginary axis, the distorted contour has a small negative imaginary value, which we identify as $-\sigma$. Note that in our actual programs σ is referred to as the 'contour height,' and these terms will be used interchangeably from now on. Note that σ is constrained to be less than the size of boundary of the analyticity strip.

Now for an example, we will use the contour $\mathcal{C}$ parameterized via

$$t \mapsto t + i\sigma \tanh(-t), \quad -\infty < t < \infty. \tag{8.2.4}$$

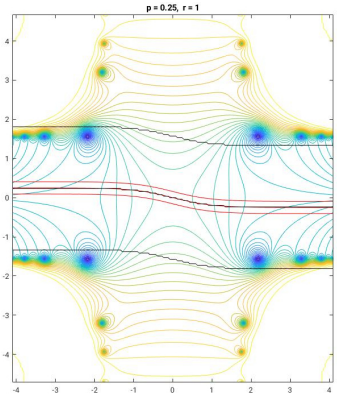
Using this parameterization we can avoid the zeros that approach the real line due to equation (7.1.11).

8.2.2 Continuation to the Whole Complex Plane

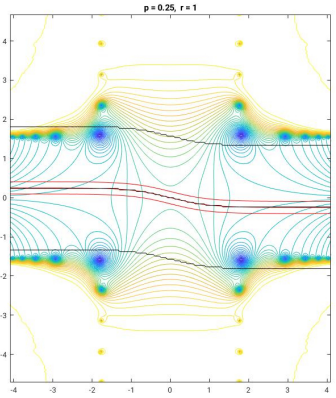
Recall back in Chapter 7 that we had to extend ϵ to the entire complex plane, and thereby came up with the $X-$ and $Y-$ system equations to analytically continue the function $Y(\theta)$ to the whole complex plane. We need to be able to numerically implement the $X-$ and $Y-$ system equations in our TBA programs.

We begin by fixing a value for p. The program will begin by solving the TBA equation along the distorted contour $\mathcal{C}$, yielding the solution ϵ along $\mathcal{C}$. We can now compute $X(\theta)$ via (7.1.12) in a strip of width $\frac{\pi}{2}$ above and below $\mathcal{C}$. Note that the integration is to be taken along $\mathcal{C}$. We then must use the X-system equation (7.1.24) to extend $X(\theta)$ to the entire plane, of course in theory. In reality we must confine ourselves to a subset of interest which our computer can handle.

Once we have $X(\theta)$ on the entire plane, we can use (7.1.20) to obtain $Y(\theta)$. To iterate we choose $p = 0.25$ for convenience, as $a = 2p$. If we want to obtain the value of $X(\theta)$ at a certain grid point, setting $p = 0.25$ will ensure we will only ever need



(a) Plot of the functions $\frac{|X(\theta)|}{1+|X(\theta)|}$



(b) Plot of the function $\frac{|Z(\theta)|}{1+|Z(\theta)|}$

Figure 8.1: Plots of the functions $\frac{|X(\theta)|}{1+|X(\theta)|}$ and $\frac{|Z(\theta)|}{1+|Z(\theta)|}$. Note that the middle black line is the original contour line, the red lines is the contour line shifted via the *contourshift* variable in our program, and the outer black lines are the contour line shifted by the *Yshift* variable in our program, i.e. the value of $\frac{\pi}{2}$ which defines the original analyticity strip.

the value of $X(\theta)$ at other grid points, i.e. we can compute $X(\theta)$ anywhere using values of $X(\theta)$ at other points.

We note that we are after the values when $Z(\theta) = 0$, since these zeros indicate the points where $Y(\theta) = -1$. These points are the singularities of $\log(1 + Y(\theta))$, which matter for TBA continuation. The function $Z(\theta)$ is simply a convenient function to plot to understand the behavior of the Y-system. Now that we can obtain $X(\theta)$ on the entire plane, we choose to plot the functions

$$\theta \mapsto \frac{|X(\theta)|}{1 + |X(\theta)|} \quad \text{and} \quad \theta \mapsto \frac{|Z(\theta)|}{1 + |Z(\theta)|} \tag{8.2.5}$$

using the distorted contour in Figure 8.1. These functions yield a visual snapshot of the $X(\theta)$ and $Z(\theta)$ functions over the entire complex plane. Note that $\theta = \alpha + i\beta$ in the complex plane, where our continuation takes place. Thus for our TBA continuation we have $(\theta - \theta_0) = \alpha + i\beta$. This implies that $|\theta| = \sqrt{\alpha^2 + \beta^2}$, which shares the same form as the equation of a circle. Thus our singularities located at $Y(\theta) = -1$ appear as circular rings, as also can be seen via Figure 8.1.

Thus in Figure 8.1, the circular rings correspond to local minima. As stated before in Chapter 6, if none of these zeros of $Z(\theta)$ cross the real axis, the distorted and the non-distorted contour are equivalent. Thus it should not matter whether we use the distorted or non-distorted contour. These two figures demonstrate this fact quite convincingly, thereby adding further numerical confirmation that the functions $X(\theta)$ and $Z(\theta)$ are indeed behaving analytically, as we expected.

We see that this both plots in Fig. 8.1 match those found in [25], lending evidence toward the fact that our programs are correctly computing the grid and functions for the sinh-Gordon model. In fact, it has been noticed that the symmetry is easily broken in these programs if a programming error has occurred. Thus the symmetry of the programs has also been a useful guide to our programs' ability to produce correct results.

8.2.3 Finding the Branch Points

Using the modified contour (8.2.4) and a modified TBA routine to scan the contour, we are able to locate branch points. In these scan programs, we scan over a grid by plotting the real and imaginary parts of the central charge $c(r)$. If the values do not converge, another variable records these values. This is what is plotted on Fig. 8.2, and gives us a very clear idea of the approximate locations of the branch points. Thus once again our analytical understanding of what should happen at a branch point is being numerically confirmed. In Fig. 8.2 the locations at where the program breaks down are indeed the square-root branch points we spoke about in Chapter 6, and they will be the key to obtaining the excited states. Note that the exact position of the branch point is found from an analysis of the motion of the zeros of Z in the complex plane.

To demonstrate how to obtain a more accurate approximation of a branch point location, we will make use of the 1-particle TBA program, see Appendix C. We first perform a wide scan as shown in Fig. 8.2. Once we obtain an idea of the branch point locations from the wide scan plot, we can then modify the parameters to reduce the scan window to isolate a single branch point at a given value of p. This is what we have done in Fig. 8.3 for the 1-singularity program with $p = 0.25$.

Our scan programs actually produce a MATLAB data file with parameters as its title string, which can be imported into our TBA analytic continuation programs. One can then visualize the path in relation to the branch point, and this is the source of the smaller plots of the branch point to the right of the main TBA analytic continuation plot which will be seen later.

8.2.4 Obtaining an excited state

After we have located a branch point using our scan program, we then must select a path by which we would analytically continue the TBA equation.

We begin by taking the parameters we noted in the scan program and input them into our analytic continuation program. We also take the appropriate data file produced

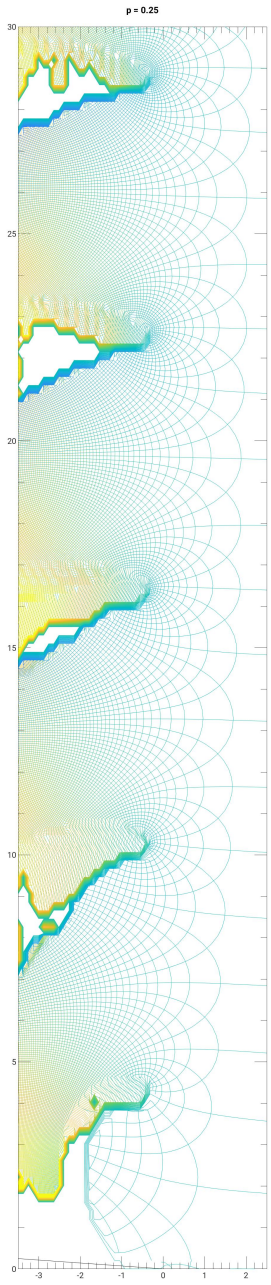


Figure 8.2: Wide frame capture of grid scan for the 0-singularity program with $p = 0.25$. Note that each endpoint of the protrusions appearing to the left of the imaginary axis, i.e. the branch points, are spaced approximately evenly along the imaginary axis. The function plotted here are the real and imaginary parts of c, the central charge.

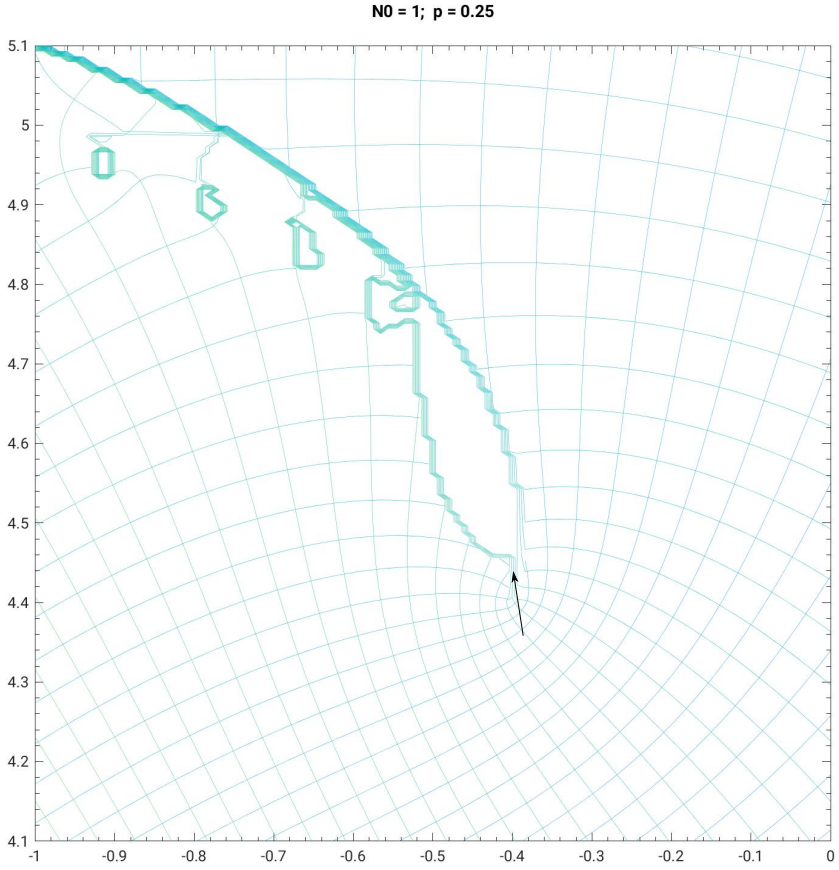


Figure 8.3: Plot of a branch point in the first $2\pi i$ segment for the 1-singularity TBA with $p = 0.25$. Note that to obtain a better idea of the location of the branch point, one can increase the grid fineness and zoom in on the image produced. This image resolution was chosen due to the ease of identifying the approximate location of the branch point with respect to the rest of the grid. The arrow identifies where the branch point is approximately located, which is the point on the grid at which the finest part of the grid (the smallest 'squares') meet with the broken down region of the grid (all the non-square-like, jagged lines). The approximate branch point here is (-0.4, 4.45).

by the scan program and park it into a separate 'datafiles' folder within the folder containing the programs. This allows our analytic continuation program to import the particular branch point scan result into the analytic continuation program under the heading 'continuation path', see Fig. 8.4. This allows us to visualize the branch point in the analytic continuation program as the continuation is carried out, as we mentioned previously. Our program actually plots the path taken by the analytic

continuation program on the scan program so one can check that one really does encircle the branch point in question with the chosen path they supply. Please refer to Figs. 8.4 - 8.6 for the following explanation.

Based on knowledge of the region gained from running the scan programs, I begin my chosen path with the r-value

$$r = -0.200 + 3.5i. \tag{8.2.6}$$

For this initial value of r, the central charge is calculated to be

$$c_1(r) = 0.23848 + 1.15319i. \tag{8.2.7}$$

Note that this r-value simply corresponds to the regular point (-0.2, 3.5) on the complex plane, where in this ordered pair (x, y) the usual x-value denotes the *real* value and the y-value denotes the value on the *imaginary* axis. For simplicity we will roughly take a modified rectangular path. At every point along the way, our programs will calculate the r-value, the central charge, and other parameters. These are directly output to the command window in MATLAB. We will start from the point (-0.2,3.5), as seen in Fig. 8.4, and endeavour to take an anti-clockwise path. Note that the Z-plot (left-hand larger plot in Fig. 8.4) in this figure shows precisely why the distorted contour is needed.

The grid for the model will break down at the branch point, which can easily be seen in the scan image. We know from Chapter 6 that at the location where the singularities of the model pinch the contour, thereby trapping it between them, a branch point is created. Once the pair of singularities cross the real axis after going round a branch point, we were forced to calculate the residues of the singularities which led to a logarithmic term in the TBA that carried the S-matrix as its argument. This was how we obtained the excited state TBA equation.

To do this numerically, starting from the ground state and moving to the first excited state, we must employ the use of both the 0-singularity program and the 1-singularity program. The 0-singularity program calculates the analytic continuation of the ground state TBA throughout the complex plane. The areas in the complex grid where the grid appears to break down are the branch points, and thus we cannot move our paths through these areas and endeavor to obtain a correct analytic continuation by only using the 0-singularity program.

As we move toward the branch point, we must match the branch points for the exact same values on both the 0-singularity program and the 1-singularity program. Once we do this, we are able to take a path from the 0-singularity program and stop it right when we enter the region of the grid that breaks down, see Fig. 8.5. We then

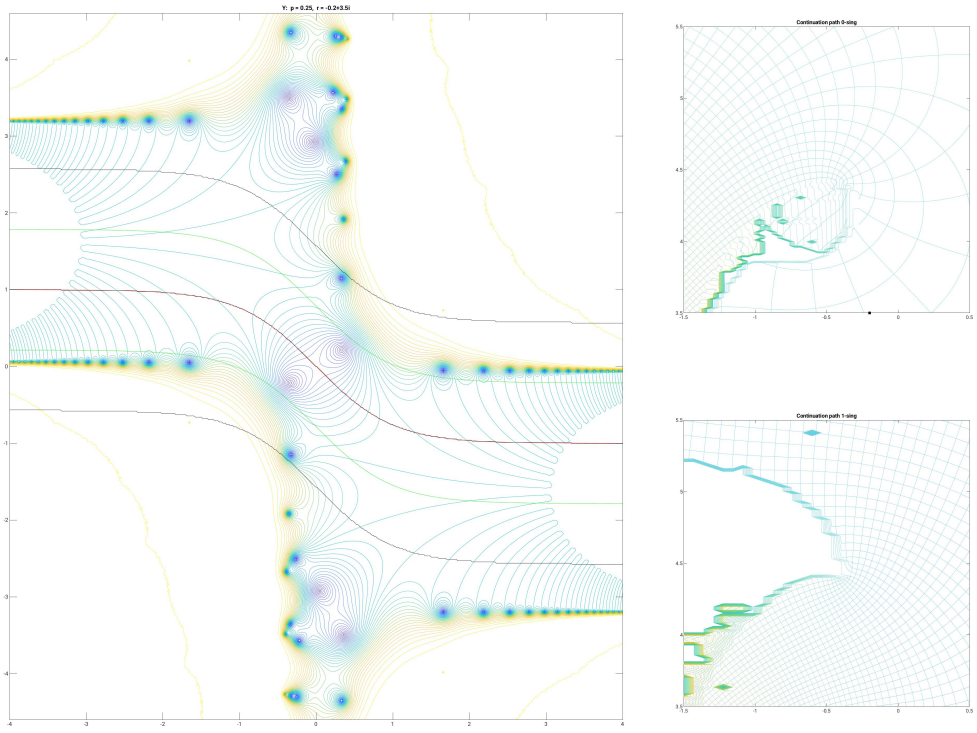


Figure 8.4: 0-singularity program run at the point (-0.2, 3.5). It is from this point that we start our path. Note that even though we are plotting Z for convenience, recall that $Z = 1 + Y(\theta)$ and thus we are still plotting Y when we plot Z.

pick up the continuation on the 1-singularity program exactly where we left off from the 0-singularity program, see Fig. 8.6, returning to the original starting value from which we began our path. In doing so, our analytic continuation of the TBA remains continuous.

Notice immediately that the location of the singularities of Z are different in Figs. 8.4 and 8.6, despite the fact that the starting and endpoint locations are identical for both paths - we really have moved onto a new sheet of the Riemann surface. Thus we have also verified that we are indeed able to numerically navigate the Riemann surface structure of our model based upon our analytical understanding. The key here is to line up the branch point protrusions obtained by the scan programs for the starting case and the excited case. Doing so will allow us to maintain the necessary smooth continuous analytical path between the Riemann sheets that our continuation requires.

At the completion of our path, the final value of r is once again

$$r = -0.2 + 3.5i, \tag{8.2.8}$$

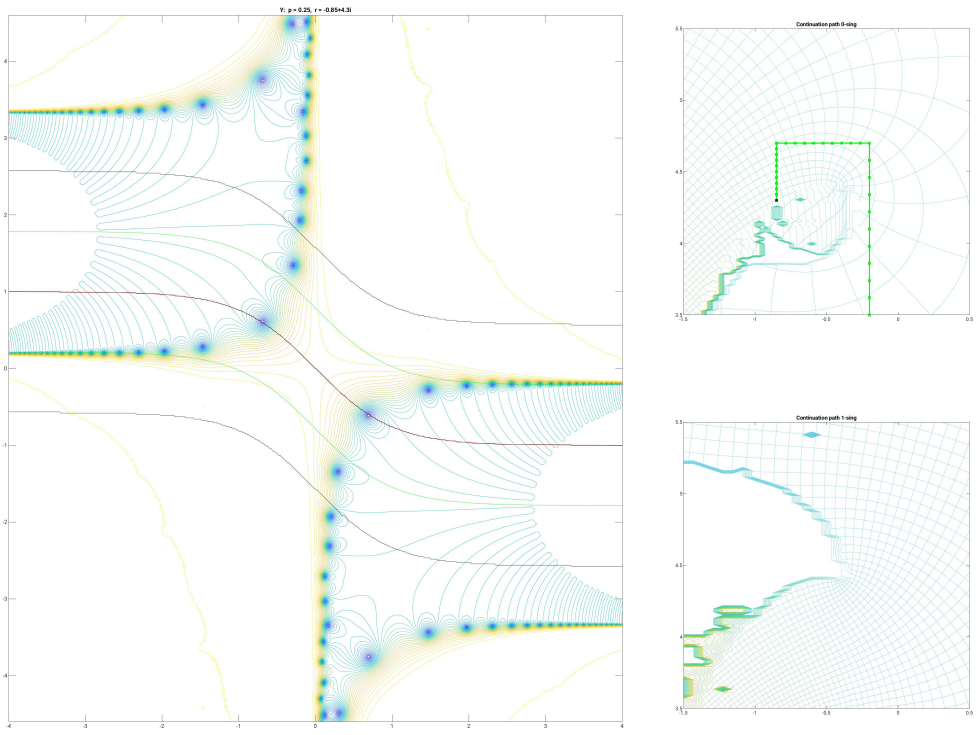


Figure 8.5: 0-singularity program run at the region where we are crossing the branch point. Notice our path, which began at (-0.2,3.5) and moves anti-clockwise. As we cross the branch point on the 0-singularity program, we must stop and resume our continuation on the 1-singularity program, as crossing the branch point brought us to another Riemann surface and gave us the excited state TBA. The point where we halted our path is (-0.85, 4.3). Hence this is why our path as seen in the Continuation Path window ends in the region where the grid is breaking down. Note that the vertical lines with black points are attached to a pair of singularities. This just helps users track the pattern of movement for the singularities.

as we expected. However, we are on a new sheet of the Riemann surface and thus expect a new effective central charge which is required to ultimately calculate the excited state energy of the model. Numerically our new effective central charge is now

$$c_1(r) = -12.87531 - 6.06842i. \tag{8.2.9}$$

The different value for $c(r)$ is another indicator that we are indeed on another branch of the Riemann surface. Note that sometimes the sign(s) may flip for $c(r)$, as in this example. For a different path, the sign of $c(r)$ may not flip. The important point is that $c(r)$ is different after transversing the branch point and returning to the point

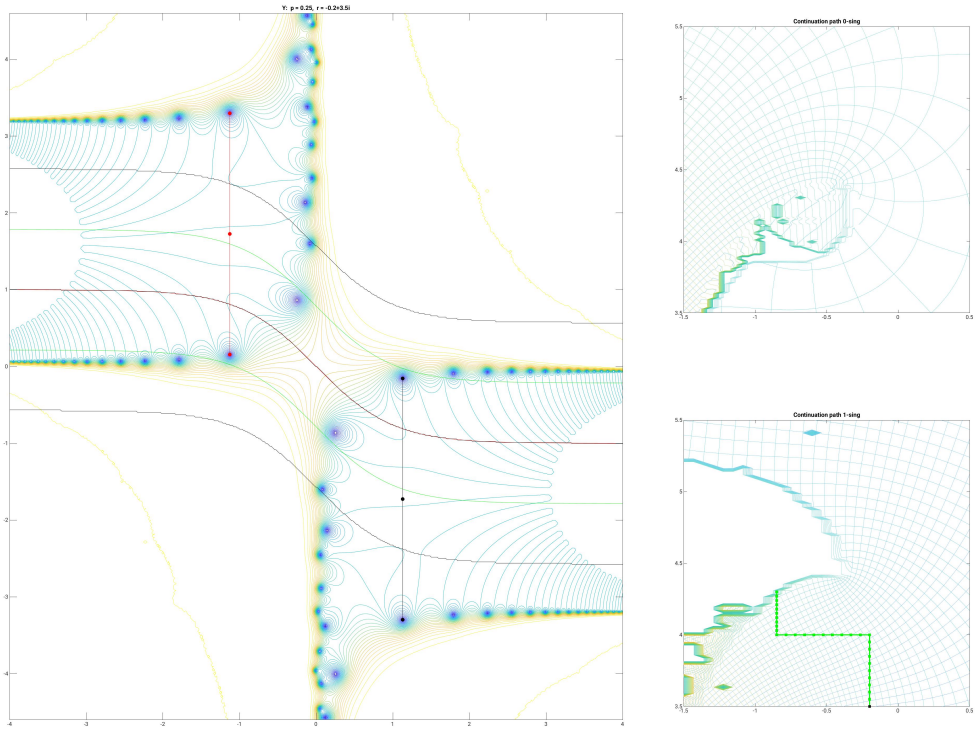


Figure 8.6: The 1-singularity program run after returning to the point (-0.2, 3.5). Notice that we resumed our continuation at the point (-0.7, 4.4), exactly the same point we ended the 0-singularity program on. This allows for continuity. Note that the scan image of the grid, seen in the Continuation Path window, has two branch points quite close to each other. When selecting a path one must take care to avoid crossing other branch points and to ensure that their continuation moves on good parts of the grid. Choosing a clockwise path for this example would have had our continuation in this figure move though the white space between the two branch points above, where the grid has already broken down. This would not have been a good continuation to take.

of origin.

Now all we have left to do is to calculate the energy at this point, once at the start of our journey and once at the end. We make use of equation (5.5.34), rearranging it to obtain

$$E(m,R) = -\frac{c(r)\pi}{6R}. \tag{8.2.10}$$

The full equation for the ground state energy includes (8.2.10), and is given by

$$E_0(m,R) = E_{bulk}(m,R) - \frac{c_0(r)\pi}{6R}, \tag{8.2.11}$$

where E_{bulk} is the bulk energy term and $c_0(r)$ is the ground state central charge.

When we transverse a branch point, as we have done above, we end up with a new value for the central charge. We saw above we started with the ground state central charge, c_0, and ended up with c_1, which we identify as the first excited state central charge. Our new energy equation, i.e. the first excited state energy, is now given by

$$E_1(m, R) = E_{bulk}(m, R) - \frac{c_1(r)\pi}{6R}. \tag{8.2.12}$$

Note that the relation between the values E_1 and c_1 in (8.2.12) is the same as that between E_0 and c_0 in (8.2.11).

Recall that we can only obtain the ground state energy using the TBA equation, i.e. the ground state eigenvalue is obtained. Since the branch points have already been determined to be of square-root type, encircling the branch point takes us to the first-excited state TBA, which picks out the other eigenvalue, which is the energy of the first excited state.

To continue to the next excited state, one now solves the excited state TBA, i.e. starts with the 1-singularity program and matches a branch point on the 1-singularity program with a branch point on the 2-singularity program. One then repeats the same exercise to obtain the next higher excited state. Repeating this process allows one to obtain the entire energy spectrum.

8.2.5 The Art of Numerical Analytic Continuation

Given that the process of analytic continuation is so central to our project and this thesis, this section will cover a few of the nuances regarding the art of selecting a proper analytic continuation.

As mentioned above, it is absolutely critical that we must have two programs to successfully analytically continue to another Riemann surface. We need a program from which we begin our continuation, i.e. a program from which we calculate the 'ground state', and another program that is one singularity greater to which we continue our path. In the above case, we utilized the actual ground state 0-singularity program and the first excited state 1-singularity program. We will stick with this set for the remainder of our discussion.

In selecting a good continuation path, care does indeed need to be taken. Riemann surface theory was invented to do complex analysis on multi-valued functions. Analytic continuation requires a continuous path, and therefore any path we take from the 0-singularity program to the 1-singularity program, around a branch point of course, must be continuous as well. How do we ensure this to be true?

This is where the 'art' part comes into play. As explained above, we begin by first scanning the grid in the same sector for branch points. Our task here becomes to match up a branch point from the program we started on with a branch point on the program we wish to continue onto. We do this by matching the exact locations numerically found via the scan programs. Once we have done this, we must devise a suitable path.

Since a suitable path must continue around a branch point and be continuous, we must think carefully about our scans. We must take note of where the matching branch points occur and where they start to break down on each scan program.

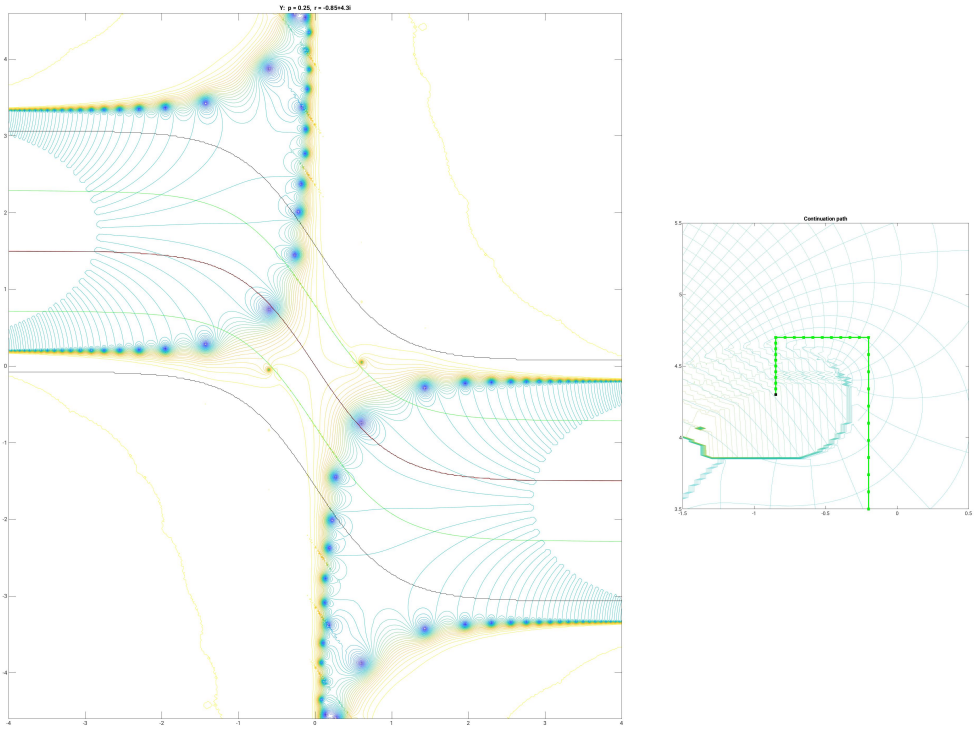


Figure 8.7: Example of poor path selection for the 0-singularity program. Note that this is the same branch point and path taken as found in Fig. 8.5. In this scan the contour height was set at 1.5 instead of 1.0, which caused the breakdown for the branch point on the 0-singularity program to occur at a larger imaginary value than the other. Hence we could not use the same path as we did previously, as we enter the broken down region sooner in this case. Furthermore it does not allow for a good match with the branch point on the 1-singularity program.

Taking a look at Figure 8.7. This figure is exactly the same branch point and we have taken the same exact path as the one found in Figure 8.5 earlier. Matching this branch point to the one in Figure 8.6 reveals that we have a problem. Since

this is the exact same path as used in the previous and correct example, we know it will produce identical results to Figure 8.6. Going back to Figure 8.7, we see that this very same path has already crossed into the region of the scan where the grid has broken down. Running this program also causes the singularities in the larger left-side window of these plots to jump randomly once we cross into this broken down region rather than move smoothly and in an apparently continuous pattern. This is a second indication that we no longer have a good continuous path.

A good path should end right before we enter the grid breakdown by the branch point on the first program and continue from that same point on the second program. The breakdown portions of the branch points on the grids where our path transverses should overlap between the two programs, i.e. the path should not move within a region where the grid breaks down in either program. Note that this means that a clockwise or anticlockwise path is 'pre-determined' for each branch point pair between programs based on the results of where the breakdown regions for those branch points occur! In other words, one is forced to take an anticlockwise path if by taking a clockwise path they will enter a region where the program has already broken down and vice versa, see 8.7.

8.3 Numerical Results

8.3.1 Sinh-Gordon Energy vs r

For this test, we use the 0, 1, and 2-singularity TBA programs We then modify our programs for the sinh-Gordon analytic continuation, see Appendix D, to plot the energy as a function of r versus the continuation parameter r for real r, as can be seen in Fig. 8.8. It turns out that the 0-singularity program contains no particles, the 1-singularity program contains two particles, and the 2-singularity program contains four particles. This can be directly seen from Fig. 8.8. This pattern continues for the higher singularity programs.

In this figure, the 2-particle and the 4-particle program's curves are associated with the Bethe parameters N_i, which corresponds to the level in the solution of the 2-particle and 4-particle asymptotic Bethe ansatz, see (5.5.3). For the 4-particle program, there are two Bethe parameters, $N0$ and $N1$. For this, $N0$-values were selected identically to the 2-particle program, while we set $N1 = N0 + 1$.

The only parameter in the sinh-Gordon TBA is r, and here each point along a single-curve is given by a different value of r. As we can clearly see, as we move towards the right, there is a clustering of particles in all programs. These particles

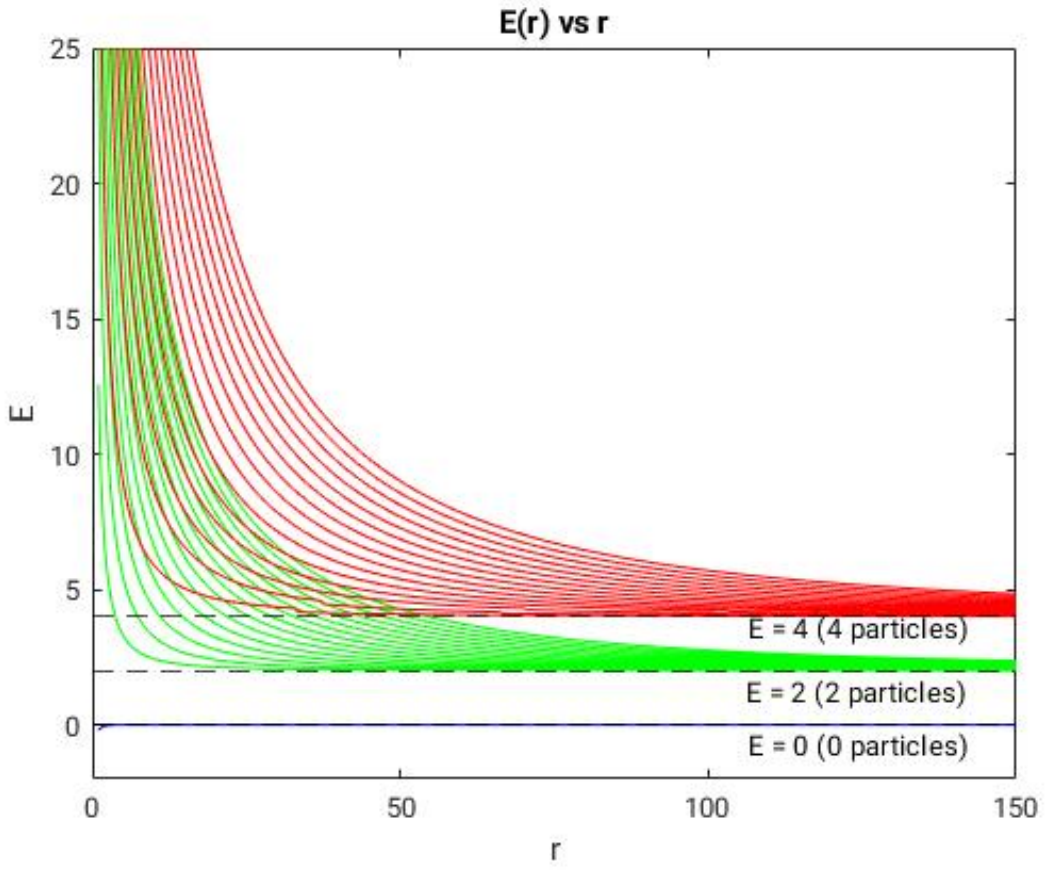


Figure 8.8: Plot of $E(r)$ versus the continuation parameter r for the 0-particle, 2-particle, and 4-particle TBA. Note all r-lines are expected to be horizontal at $E = 0$ for the 0-singularity program, as it contains no particles. As $p \to 0$, r-lines tend to the asymptotic lines commensurate with the number of particles they have, so $E = 2$ for the 1-singularity program and $E = 4$ for the 2-singularity program. This is exactly what is seen in this graph.

are moving slower here and should converge towards a value that represents the number of particles in the program we are dealing with. We see in Figure 8.8 that the energy is zero for the entire run, which makes sense as this program contains no particles. Thus they cannot be moving faster at the left and slower at the right - $E(r)$ here must always be zero. For the 2-singularity program, this is no longer the case. We expect that as r increases, the particles move slower. This is explained by equation (5.4.2), where in our derivation of the TBA as $L \to \infty$ we then took the space of states, which was along R of the torus, and then placed it in the L-direction.

Thus for our problem we have

$$p_1 = \frac{2\pi k}{R},\ k \epsilon \mathbb{Z}. \tag{8.3.1}$$

When $R \to \infty$, clearly the momentum p_1 is reduced. Hence the particles move slower in this limit.

Returning our attention to the right of the plot we see that $E(r) = 2$ because this is the number of particles in the program. We have found a similar confirmation for our 4-particle program in Figure 8.8. These results are all expected and represent a very good indication that our understanding and our numerical programs are correct.

Note that for the excited state program, we must fix θ_0. Thus the 1-singularity program will have to have θ_0 fixed. For the next excited state, i.e. the 2-singularity program, θ_0 and θ_1 must be fixed. This pattern continues for higher singularity programs.

There are many more interesting and fascinating aspects about the sinh-Gordon model that we can study from its Riemann surface, so much so that this could take up a whole other treatise on the subject. During the course of our investigations of the sinh-Gordon model with this technique, we happened to stumble upon a situation that is of immense importance to this technique and curiously lead to a series of fundamental insights that shed light on a deep relationship between eigenvalues, boundary conditions, and complex analysis. We begin part of this exploration in the next section, but the full treatment will have to wait until the following chapter.

8.3.2 Branch Points & Free Boson Limit

We have extensively made use of the scan programs, see Appendix C for the 1-particle scan programs, to locate the branch points that enable us to access the other levels of the Riemann surface. This is a very time-intensive process and requires patience to set the correct parameters for the scan to obtain accurate locations for the branch points. We note that the locations and number of branch points in each $2\pi mi$ interval depend on which program one is running, i.e. the 0-singularity, 1-singularity, etc. The exact locations of the branch points also change depending on the value of the parameter p, as we will demonstrate. Selected data for branch point locations for different programs and for different values of p will now be presented below.

Our first example set of data comes from the 0-singularity program in the first $2\pi mi$ region. Note that regions are defined via the imaginary part of the branch point location. Each region is defined by the quantity $2\pi mi$, where $m = 1, 2, 3, ...$ and corresponds to the first, second, etc. region. The first region lies within the

range $0 < r \leq 2\pi i$, the second region within the range $2\pi i < r \leq 4\pi i$, and so on and so forth. Note that we denote this 0-singularity program and region as BPm, where m corresponds to the region number. We will look at the first region for the 0-singularity program, i.e. $BP1$. This data is shown in *Table* 8.1.

BP1						
p = 0.5	-0.40301+4.16370i		p = 0.09	-0.35082 + 5.00757i	p = 0.009	-0.14042 + 5.91579i
p = 0.45	-0.39355 + 4.17220i		p = 0.08	-0.34087 + 5.07473i	p = 0.008	-0.13212 + 5.94087i
p = 0.4	-0.39490 + 4.19810i		p = 0.07	-0.32891 + 5.14862i	p = 0.007	-0.12317 + 5.96739i
p = 0.35	-0.39665 + 4.24295i		p = 0.06	-0.31439 + 5.23058i	p = 0.006	-0.11344 + 5.99564i
p = 0.3	-0.39820 + 4.30945i		p = 0.05	-0.29656 + 5.32253i	p = 0.005	-0.10274 + 6.02601i
p = 0.25	-0.39840 + 4.40220i		p = 0.04	-0.27423 + 5.42732i	p = 0.004	-0.09080 + 6.05908i
p = 0.2	-0.39510 + 4.52870i		p = 0.03	-0.24543 + 5.54952i	p = 0.003	-0.07717 + 6.09581i
p = 0.15	-0.38470 + 4.70173i		p = 0.02	-0.20646 + 5.69765i	p = 0.002	-0.06103 + 6.13796i
p = 0.1	-0.35914 + 4.94609i		p = 0.01	-0.14817 + 5.89195i	p = 0.001	-0.040355 + 6.189745i

Table 8.1: Location data for branch point 1 for 0-singularity program.

BP2				
p = 0.5	-0.3639+10.0767i		p = 0.09	-0.40350 + 11.00752i
p = 0.45	-0.3659+10.0849i		p = 0.08	-0.39502 + 11.08726i
p = 0.4	-0.3717+10.1099i		p = 0.07	-0.38402 + 11.17544i
p = 0.35	-0.3809+10.1542i		p = 0.06	-0.36982 + 11.27376i
p = 0.3	-0.3929+10.2218i		p = 0.05	-0.35146 + 11.38462i
p = 0.25	-0.4061+10.3191i		p = 0.04	-0.32744 + 11.51157i
p = 0.2	-0.4177+10.4565i		p = 0.03	-0.29532 + 11.66034i
p = 0.15	-0.42258 + 10.65113i		p = 0.02	-0.25046 + 11.84155i
p = 0.1	-0.40996 + 10.93493i		p = 0.01	-0.25046 + 11.84155i

Table 8.2: Location data for branch point 2 for 0-singularity program.

We note that for the maximum value we are able to select for p, i.e. $p = 0.5$, the imaginary part of complex number specifying the branch point falls within the range for region 1. As we send $p \to 0$, we note that the real part of the branch point approaches zero, and the imaginary part of the location value approaches $2\pi i$, i.e. the uppermost value of the first region. In fact for $BP2$, see Table 8.2, and all higher branch points, as we send $p \to 0$, we notice that the real part of the branch point tends to zero and the imaginary part tends to $2\pi mi$ where m refers to the particular region of the branch point. Thus in the small p limit ($p \to 0$), our branch points form a vertical line at the origin, where they are equally spaced apart in intervals of $2\pi i$, as can be seen from Figure 8.9.

The branch point locations for the small-p limit actually coincides with the location of the branch points for the free boson model, and thus we have shown numerically

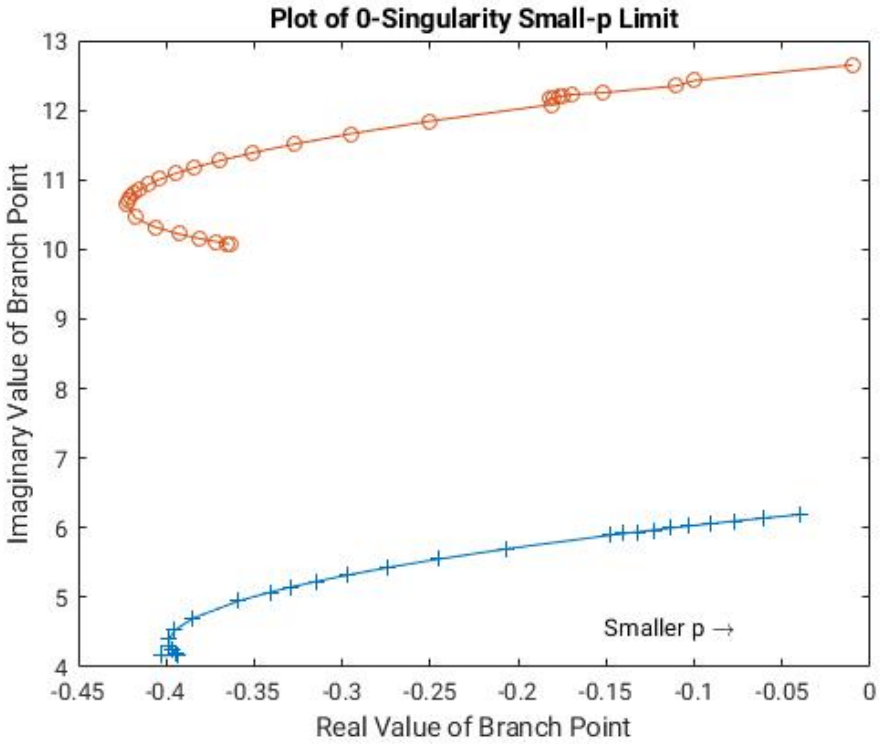


Figure 8.9: Plot of the branch point locations as p varies from large to small for the first and second branch points. The points plotted are those found in Table 8.1 for the first branch point (lower line) and in Table 8.2 for the second branch point (upper line). Notice that as p gets smaller, the branch point lines up at zero and is located roughly at $2\pi i$.

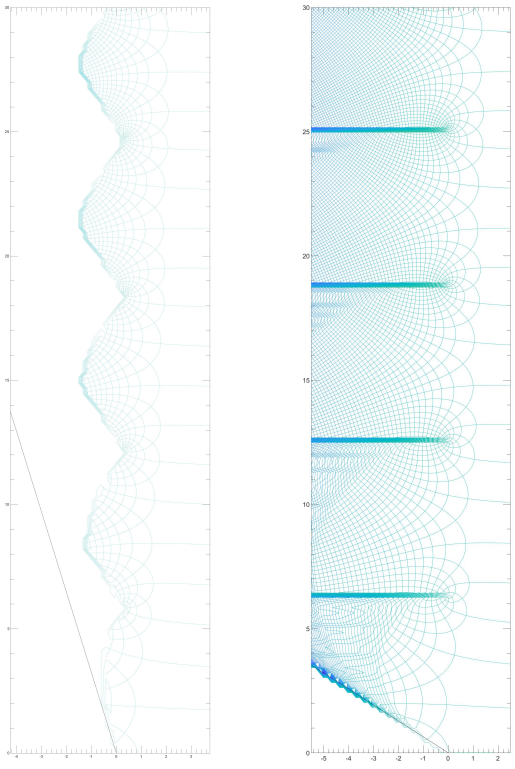
that there exists a mapping between the sinh-Gordon model and the free boson model, see Fig. 8.10.

We next enlist the 1-singularity scan program and focus on region 1. We now have three branch points, respectively given by Tables 8.3, 8.4, and 8.5.

BP1_1a							
p = 0.5	-0.23+2.595i		p = 0.09	-0.236+4.05i		p = 0.009	-0.11291 + 5.64627i
p = 0.45	-0.23+2.61i		p = 0.08	-0.233+4.168i		p = 0.008	-0.10675 + 5.69022i
p = 0.4	-0.23+2.655i		p = 0.07	-0.228+4.297i		p = 0.007	-0.10002 + 5.73668i
p = 0.35	-0.233+2.73i		p = 0.06	-0.222+4.441i		p = 0.006	-0.09262 + 5.78611i
p = 0.3	-0.236+2.845i		p = 0.05	-0.212+4.602i		p = 0.005	-0.08439 + 5.83918i
p = 0.25	-0.24+3.002i		p = 0.04	-0.201+4.787i		p = 0.004	-0.07507 + 5.89688i
p = 0.2	-0.244+3.22i		p = 0.03	-0.183+5.002i		p = 0.003	-0.06428 + 5.96083i
p = 0.15	-0.245+3.517i		p = 0.02	-0.158+5.263i		p = 0.002	-0.05128 + 6.03401i
p = 0.1	-0.239+3.943i		p = 0.01	-0.11861 + 5.60444i		p = 0.001	-0.03430 + 6.12354i

Table 8.3: Location data for branch point 1a for 1-singularity scan program.

For large p-values, these branch points are all distinct. It has been verified that



(a) sinh-Gordon at $p = 0.001$ (ground-state)

(b) free boson model (p - irrelevant)

Figure 8.10: Sinh-Gordon and free boson comparision as $p \to 0$.

BP1_1b						
p = 0.5	-0.40301+4.16370i	p = 0.09	-0.35082 + 5.00757i	p = 0.009	-0.14042 + 5.91579i
p = 0.45	-0.39355 + 4.17220i	p = 0.08	-0.34087 + 5.07473i	p = 0.008	-0.13212 + 5.94087i
p = 0.4	-0.39490 + 4.19810i	p = 0.07	-0.32891 + 5.14862i	p = 0.007	-0.12317 + 5.96739i
p = 0.35	-0.39665 + 4.24295i	p = 0.06	-0.31439 + 5.23058i	p = 0.006	-0.11344 + 5.99564i
p = 0.3	-0.39820 + 4.30945i	p = 0.05	-0.29656 + 5.32253i	p = 0.005	-0.10274 + 6.02601i
p = 0.25	-0.39840 + 4.40220i	p = 0.04	-0.27423 + 5.42732i	p = 0.004	-0.09080 + 6.05908i
p = 0.2	-0.39510 + 4.52870i	p = 0.03	-0.24543 + 5.54952i	p = 0.003	-0.07717 + 6.09581i
p = 0.15	-0.38470 + 4.70173i	p = 0.02	-0.20646 + 5.69765i	p = 0.002	-0.06103 + 6.13796i
p = 0.1	-0.35914 + 4.94609i	p = 0.01	-0.14817 + 5.89195i	p = 0.001	-0.040355 + 6.189745i

Table 8.4: Location data for branch point 1b for 1-singularity scan program. Notice that this branch point is identical to the branch point BP1 from the 0-singularity scan program

certain branch points link up with one another on different levels of the Riemann surface, as is to be expected. Note that branch point *BP1_1b* (the branch point for region 1 branch point *b* of the 1-singularity scan program) is actually the same branch point found in region 1 of the 0-singularity scan program, and we will use *BP1_1b* to describe the locations of the other two branch points relative to it. Branch point *BP1_1a* is found beneath it, and *BP1_1c* is found above it. The interesting behavior

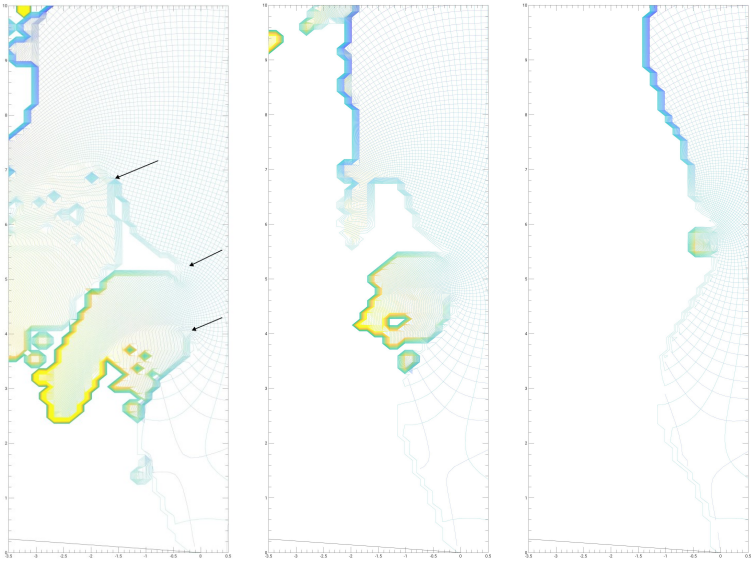
BP1_1c							
p = 0.5	-3.6735 + 10.8136i		p = 0.09	-1.7670 + 6.9600i		p = 0.009	-0.42371 + 6.37007i
p = 0.46	-3.714 + 10.7765i		p = 0.08	-1.6432 + 6.8860i		p = 0.008	-0.39341 + 6.36195i
p = 0.4	-3.929 + 10.513i		p = 0.07	-1.5135 + 6.8130i		p = 0.007	-0.36163 + 6.35369i
p = 0.36	-4.1055 + 10.039i		p = 0.06	-1.3765 + 6.7410i		p = 0.006	-0.32803 + 6.34526i
p = 0.3	-3.851 + 9.002i		p = 0.05	-1.2302 + 6.6695i		p = 0.005	-0.29222 + 6.33665i
p = 0.26	-3.476 + 8.488i		p = 0.04	-1.0722 + 6.5985i		p = 0.004	-0.25359 + 6.32772i
p = 0.2	-2.916 + 7.87505i		p = 0.03	-0.8978 + 6.5270i		p = 0.003	-0.21107 + 6.31846i
p = 0.15	-2.43 + 7.435i		p = 0.02	-0.6980 + 6.4542i		p = 0.002	-0.16278 + 6.30865i
p = 0.1	-1.8860 + 7.0350i		p = 0.01	-0.45273 + 6.37806i		p = 0.001	-0.10411 + 6.297895i

Table 8.5: Location data for branch point 1c for 1-singularity scan program.

again comes from the small p-limit. In the small p-limit, the branch points actually coincide. *BP1_1c* moves downward toward *BP1_1b*, while *BP1_1a* moves upward toward *BP1_1b*.

Eventually in the limit $p \to 0$, the three branch points merge to form a single branch point - the branch point at $2\pi i$! We have found that in the small p-limit, the 1-singularity program for the sinh-Gordon model reduces to the free boson model. This actually holds true regardless of the number of particles we employ in our computations of the sinh-Gordon model. In Fig. 8.11 the following series of plots of the 1-singularity scan program visually show the branch points merging in the small-p limit.

We also include a plot of the three branch points above for the 1-singularity program in Figure 8.12. In this figure we can clearly see that the branch points coalesce as p gets smaller. We also notice that once again in the first branch point region, these points merge with the imaginary value roughly at $2\pi i$, similarly to the 0-singularity program for the same region.



(a) $p = 0.1$ (b) $p = 0.05$ (c) $p = 0.01$

Figure 8.11: Scans of 1-singularity branch points as $p \to 0$. Notice that we observe multiple distinct branchpoints in this region at $p = 0.1$ which merge into a single branchpoint as p gets smaller. Note that the black arrows in (a) give the positions of the three branch points initially at $p = 0.1$.

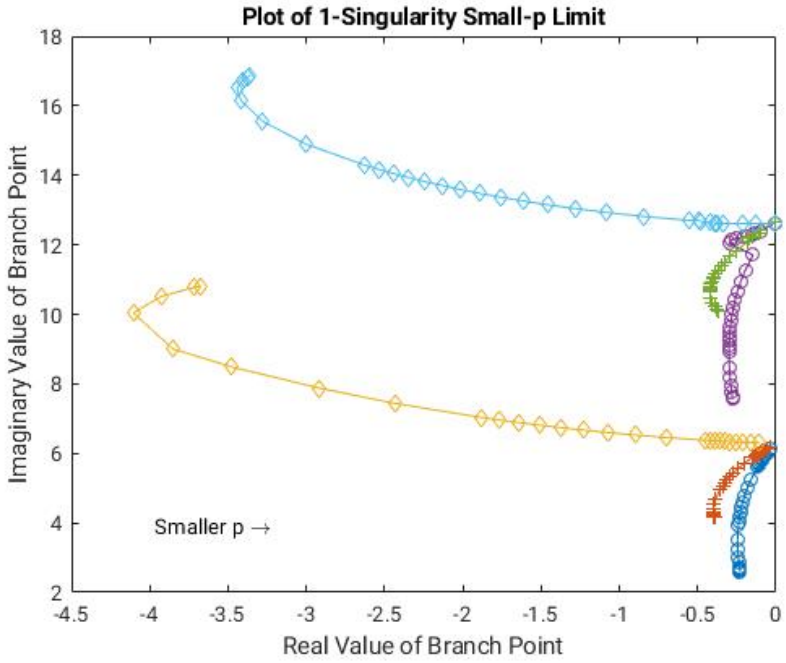


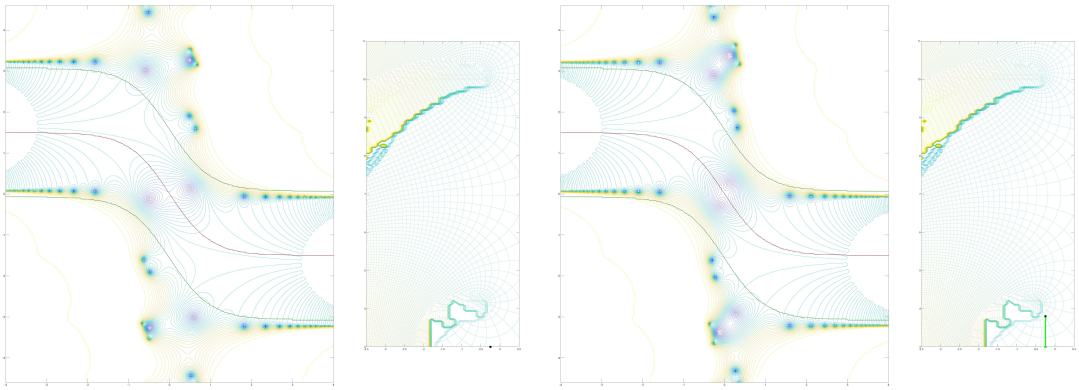
Figure 8.12: Plot of the branch point locations as $p \to 0$ for the first branch point region. The branch points plotted are those found in Table 8.3, 8.4, and 8.5. Notice that as p gets smaller, the three branch points merge together with a real-part of zero and with their imaginary parts at multiples of $2\pi i n$, where n refers to the branch point in question. Note that the purple line looks discontinuous, but this is only because the super computer was not available at the time to run the program on a finer grid, unlike the data for the others.

8.4 Singularities Sequencing Pattern

The important points for the sinh-Gordon model are the points where $Y = -1$, recall equation (7.1.10). These are the singularities of the model. Having said this, one look at the analytic continuation window in Fig. 8.4, for example, one would notice many singularities. How do we know which singularities will form the pinch singularity when we round the branch point? Answering this question is the subject of this section.

We will begin by focusing on the 0-singularity program for simplicity. The same pattern, though more complex, holds for the higher singularity programs. The idea here is to use a straight path consisting of 10 separate steps that passes close to the branch point, but does not actually encircle it. Recall here that we are simply trying to demonstrate what happens with the singularities in the sinh-Gordon model.

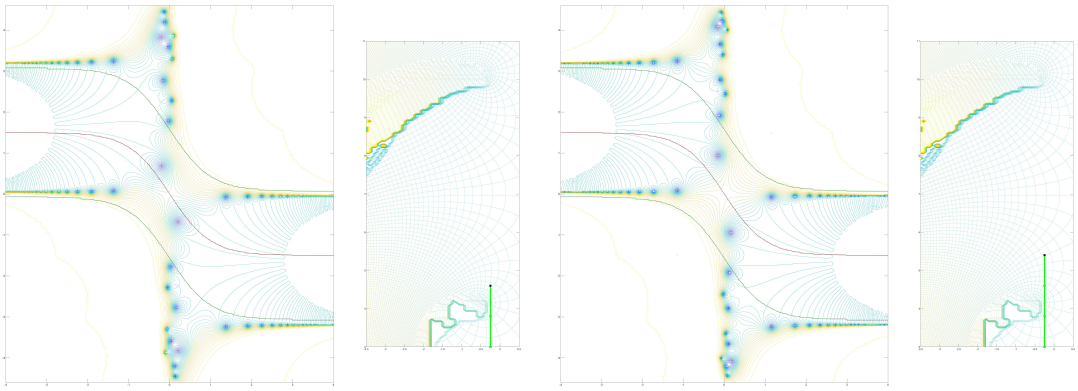
We will start with a path that passes past the first branch point but remains in its vicinity. One can verify this path, the fact that we are dealing with the first singularity, and the behavior of the contour in the following figures.



(a) Step 1 (b) Step 2

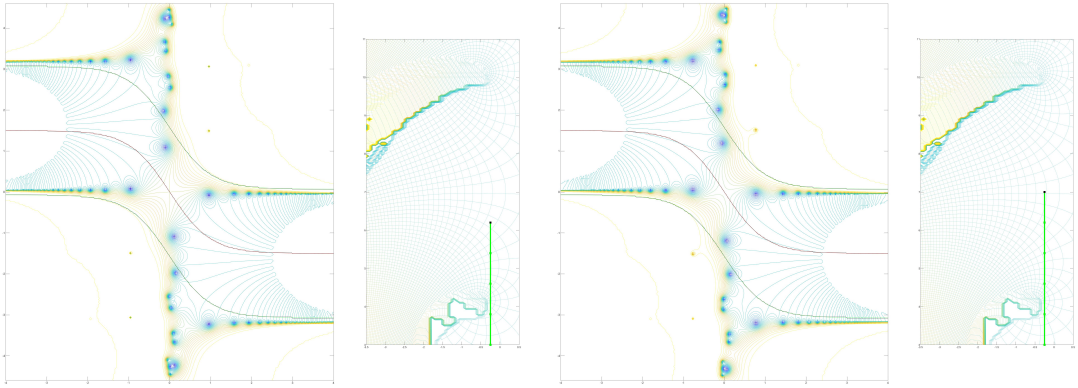
Figure 8.13: Set one of images showing the sequencing pattern of singularities for the 0-particle program. For each plot a) and b), the plot on the right is the continuation path. In this plot one can trace the path we take for our analytic continuation. Note here it starts at the point $-0.25 + 3.0i$. The larger figure on the left is the plot of the Y-system at $p = 0.5$. As we transverse our path, the singularities in the Y-system plot will move (see other steps).

We notice that throughout each of the ten steps the first singularity from the right (for the part of the contour above the distorted contour (thick black line in Figs. 8.13 - 8.17)) and the first singularity from the left (for the part of the contour below the



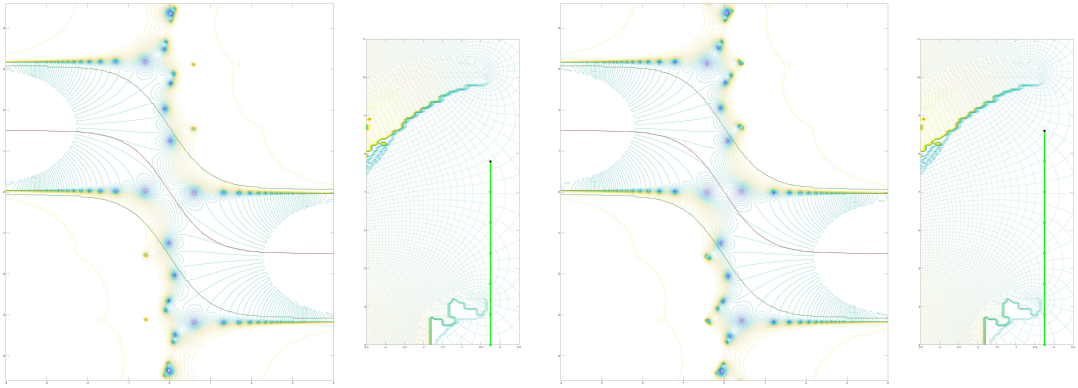
(a) Step 3 (b) Step 4

Figure 8.14: Set two of images showing the sequencing pattern of singularities for the 0-particle program.



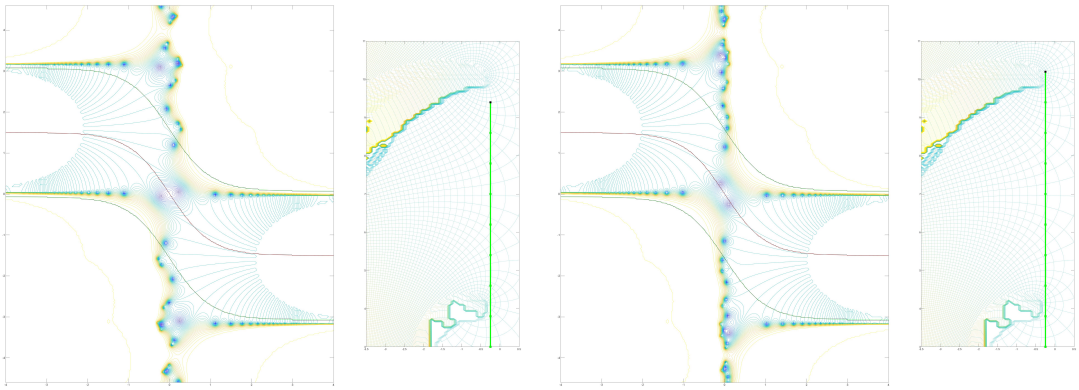
(a) Step 5 (b) Step 6

Figure 8.15: Set three of images showing the sequencing pattern of singularities for the 0-particle program.



(a) Step 7 (b) Step 8

Figure 8.16: Set four of images showing the sequencing pattern of singularities for the 0-particle program.



(a) Step 9 (b) Step 10

Figure 8.17: Set five of images showing the sequencing pattern of singularities for the 0-particle program.

distorted contour) take part in the little 'dance around the contour' and ultimately trap the contour to create the pinch singularity that marks out the first branch point. Thus the first singularity from the right and the first singularity from the left take part to create the first branch point.

Let us move on to the second branch point. To do so, we approach the second singularity from the same starting point as we did for the first branch point. The only difference is the distance covered, but we will continue to have ten steps for this path. We see this in Figs. 8.13 - Figs. 8.17.

For the path approaching the second singularity, we notice that as we pass the first singularity, as verified from the continuation path window in Figs. 8.13 - Figs. 8.17, the first singularity from the right and the first singularity from the left are skipped. Now the second singularities from both the left and the right participate in the "little dance" that creates the pinch singularity that ultimately gives rise to the second branch point. We notice a pattern here!

It turns out that the branch points regions are separated by a distance of $2\pi im$. Thus the pattern is that for each m, where $m = 1...n$, the nth horizontal singularities above and below the contour are the singularities that create the branch point for the nth branch point.

What about the higher order programs? For the 1-singularity program, the branch point regions contain multiple branch points per region, as seen above in the data and figures. In the small-p limit, the branch points for the higher singularity programs, the multiple branch points seen per region coalesce into a single branch point and the separation of branch points by $2\pi im$ is easily verified. Since much has been mentioned about the small-p limit, it is now prudent to analyze it in the next section.

8.5 Small-p Limit

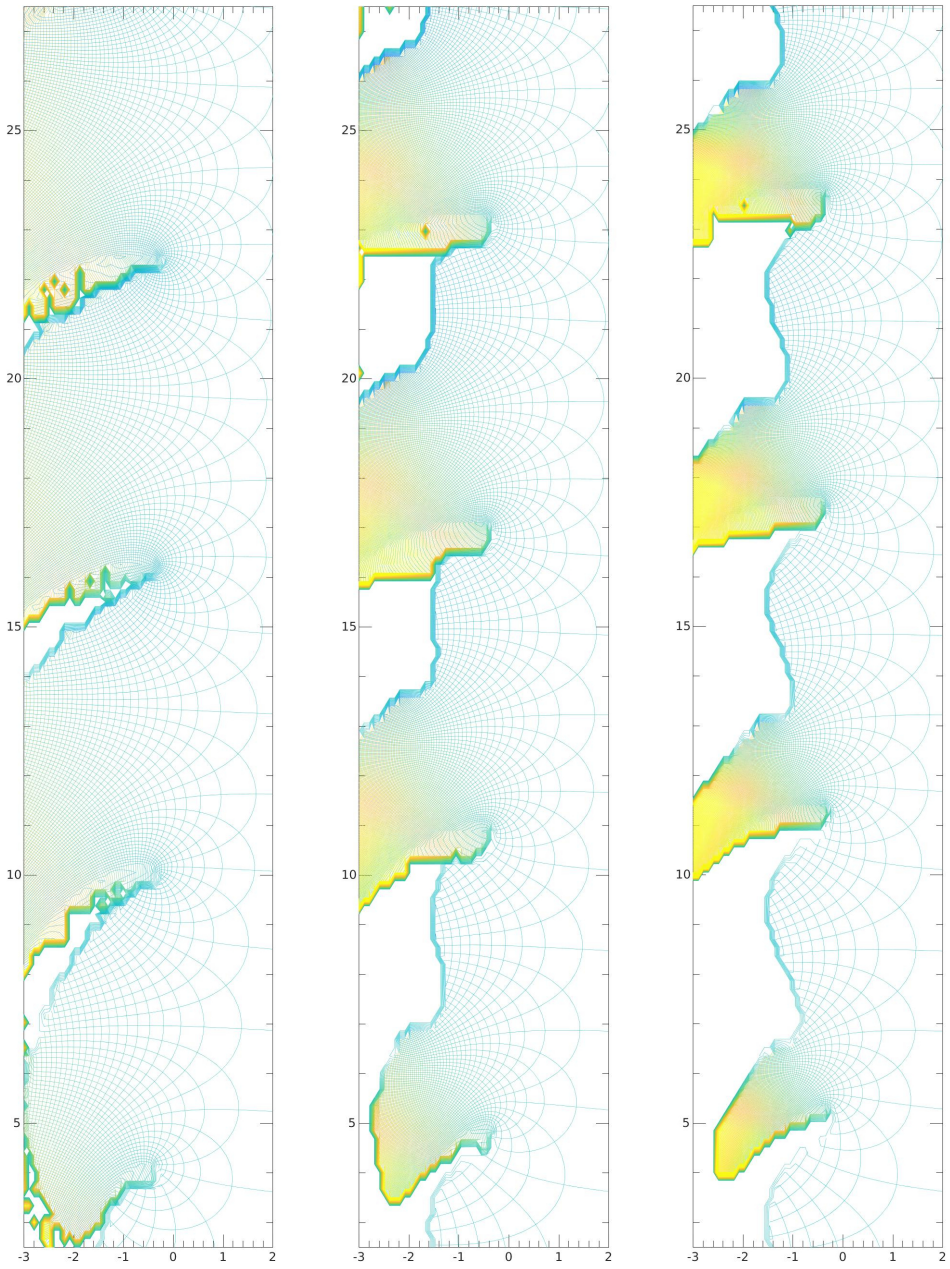
In previous parts we have discussed phenomena, such as the joining together of the branch points in the 1-singularity case, when $p \to 0$. Recall that p is the redefined sinh-Gordon parameter, and it can influence the sinh-Gordon solution to quite an extent. A more thorough look at the small-p limit is warranted to better understand what happens with the sinh-Gordon model and to predict general behaviors. We attempt to accomplish this feat in this section.

8.5.1 Behavior of X-System and Y-System in the Small-p Limit

In the previous analytical account of the small-p limit, we focused on the central charge we would need to compute the energy of the system as well as the Y-system. Given the fact that for the sinh-Gordon model we also have an X-system that we must extend throughout the grid and solve for the Y-system from it, we should also explore what happens to the X-system in such a limit. We will start by doing this visually. Note that for this section we had to use an unusually high precision factor, including up to 500 digits of precision for our smallest values of p. MATLAB natively does not possess this capability, so we made use of an add-on package called Advanpix which allows for variable precision[1].

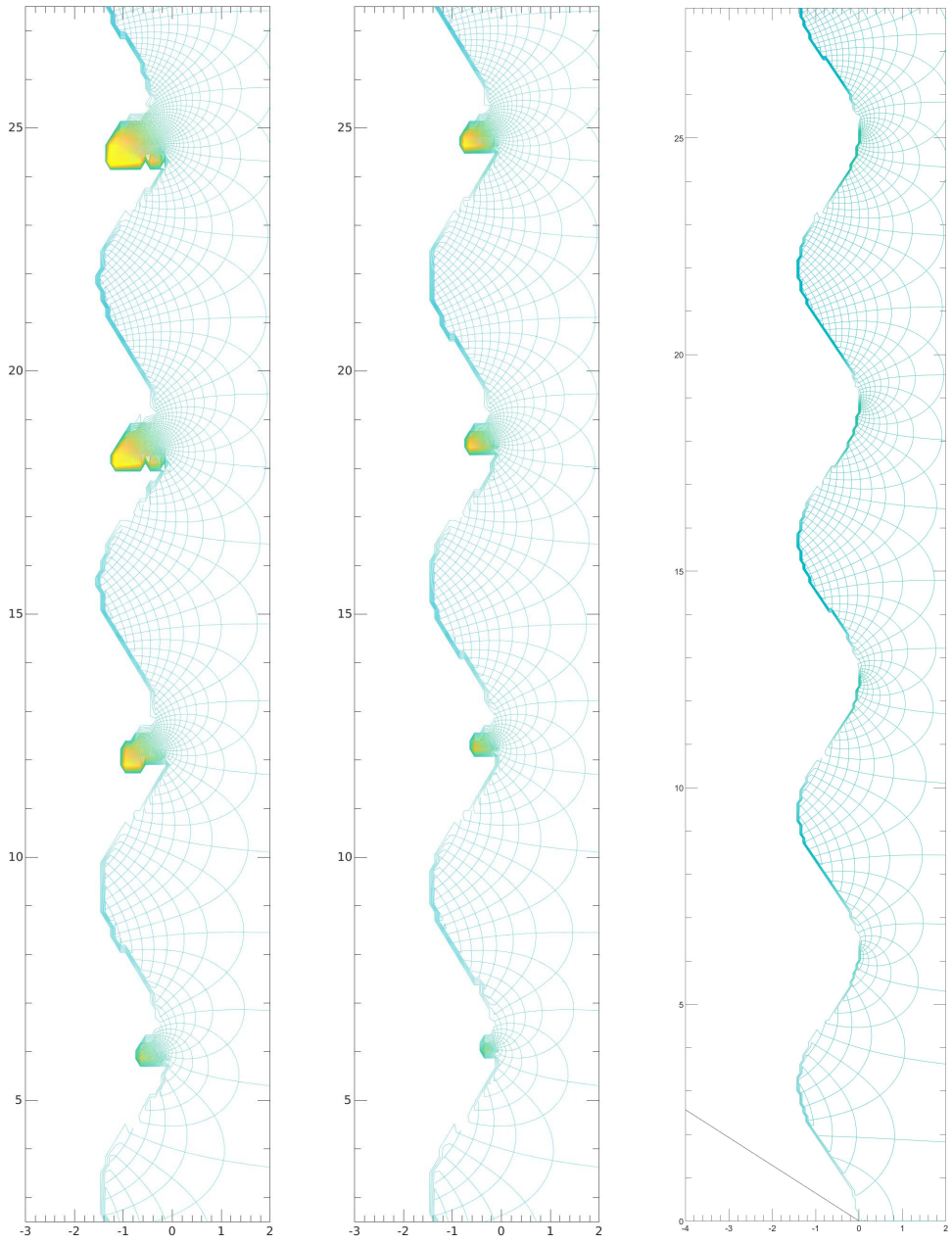
Small-p Limit of Sinh-Gordon Grid

Recall that when we seek to discover the locations of the branch points for the sinh-Gordon model, we create a separate program to compute the numerical grid and then we scan it. We have seen the branch points line up. We wish to know, in the small-p limit, the shape of the underlying function governing the convergence in these scans. In this limit, the function that governs these locations is clearly seen to have a sinusoidal-like form, as can be seen in the following sequence of plots:



(a) $p = 0.5$ (b) $p = 0.1$ (c) $p = 0.05$

Figure 8.18: Set one of images showing the emergence of a sinusoidal-like underlying function governing the convergence of the sinh-Gordon branch points revealed as $p \to 0$.

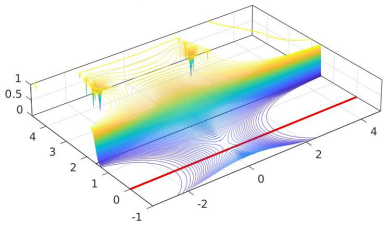


(a) $p = 0.01$ (b) $p = 0.005$ (c) $p = 0.0$

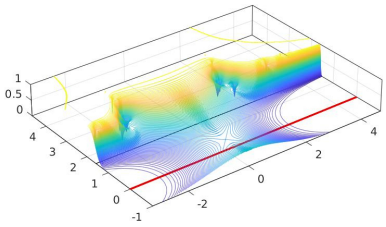
Figure 8.19: Set two of images showing the emergence of a sinusoidal-like underlying function governing the convergence of the sinh-Gordon branch points revealed as $p \to 0$.

Bird's Eye View of X and Y Function Limits

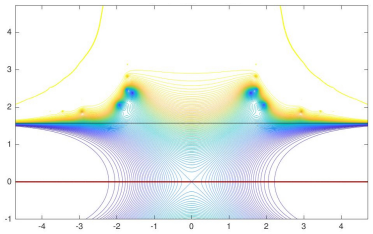
As we decrease the value of p and plot the X-function, as well as the Y-function for comparison, we observe the plot of the singularities flattening out in the middle-region of the graph of the X-function, as can be seen from the following selective sequence of plots:



(a) 3D X-function

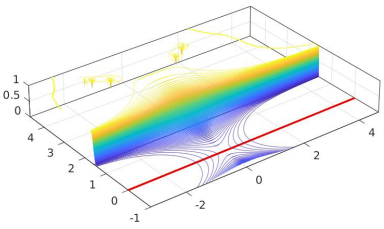


(b) 3D Y-function

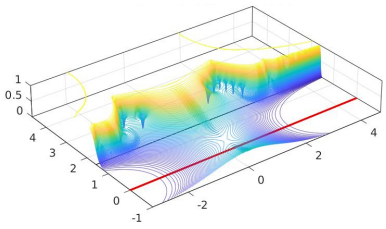


(c) 2D Y-function

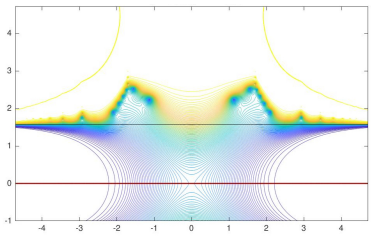
Figure 8.20: 2-D and 3-D plots of X and Y functions when $p = 0.1$.



(a) 3D X-function

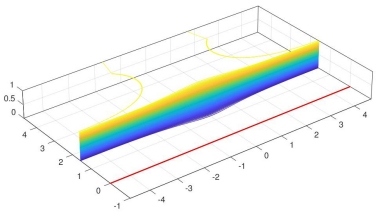


(b) 3D Y-function

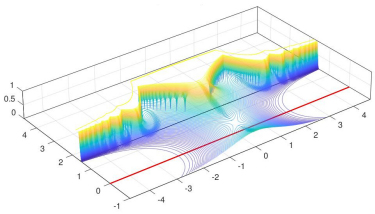


(c) 2D Y-function

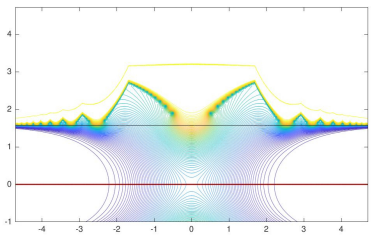
Figure 8.21: 2-D and 3-D plots of X and Y functions when $p = 0.05$.



(a) 3D X-function

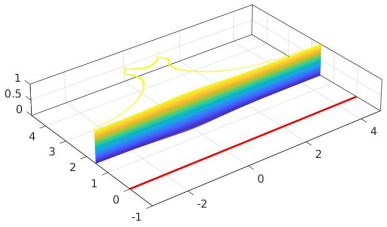


(b) 3D Y-function

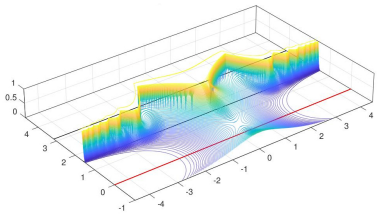


(c) 2D Y-function

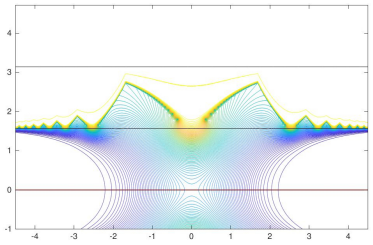
Figure 8.22: 2-D and 3-D plots of X and Y functions when $p = 0.01$.



(a) 3D X-function

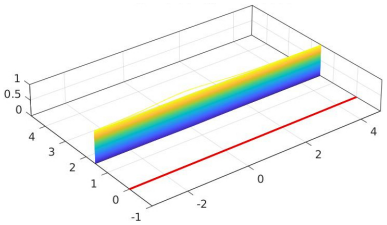


(b) 3D Y-function

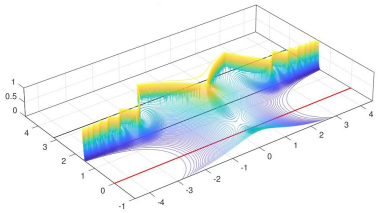


(c) 2D Y-function

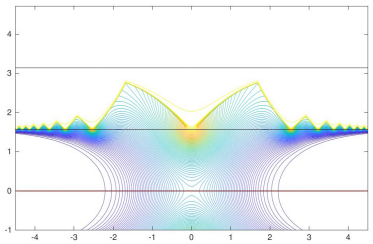
Figure 8.23: 2-D and 3-D plots of X and Y functions when $p = 0.005$.



(a) 3D X-function

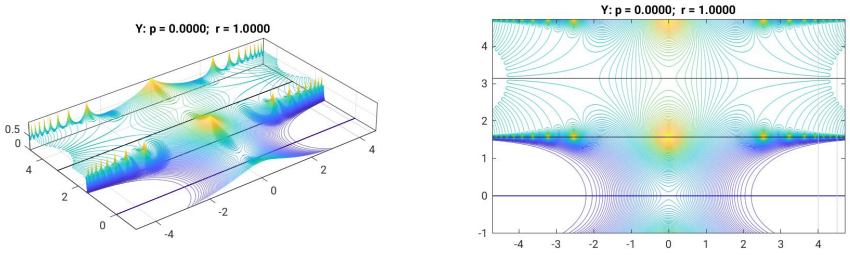


(b) 3D Y-function



(c) 2D Y-function

Figure 8.24: 2-D and 3-D plots of X and Y functions when $p = 0.001$.



(a) 3D Y-function (b) 2D Y-function

Figure 8.25: 2-D and 3-D plots of the Y functions when $p = 0.0$.

As $p \to 0$, the X-system clearly appears to assume a constant value. In the limit $p \to 0$, we also notice that the Y-system flattening out a bit. For $p = 0$, the singularities condense and line up in a line. In fact, we know the exact form of the Y-function at $p = 0$, as is detailed in the next section.

8.5.2 Analysis

Ground State Limit of the Y-Function

To begin, the seemingly most pertinent question one would have about the behavior of the sinh-Gordon model in the small-p limit would likely be what happens to the solution? To investigate this, we need to focus on the Y-system and the central charge.

As we send $p \to 0$, the kernel $\phi(\theta)$ becomes increasingly concentrated around $\theta = 0$. If $L(\theta')$ isn't varying too rapidly near $\theta = 0$, a decent approximation is to replace $\phi(\theta - \theta')$ with $2\pi\delta(\theta - \theta')$ in the convolution term of the TBA equation, i.e. (5.5.28). This yields

$$\epsilon(\theta) = r\cosh(\theta) - \int_{\mathcal{R}} \delta(\theta - \theta')L(\theta')d\theta' = r\cosh(\theta) - L(\theta). \tag{8.5.1}$$

Recall that for the Y-function we have

$$Y(\theta) = e^{-\epsilon(\theta)}, \tag{8.5.2}$$

which is just the negative exponential of the solution to the sinh-Gordon TBA (recall that the S-matrix alone tailors the TBA to a particular model).

We proceed by substituting the result in (8.5.1) into (8.5.2) to obtain

$$Y(\theta) = \exp(-r\cosh(\theta) + L(\theta)) \tag{8.5.3}$$

$$= \exp(-r\cosh(\theta)) * \exp(\log(1+Y(\theta))) \tag{8.5.4}$$

$$= \exp\left(-r\cosh(\theta)\right)(1+Y(\theta))\,, \tag{8.5.5}$$

where we utilized the fact that that $L \equiv \log(1+Y(\theta))$.

Now solving for $Y(\theta)$ yields

$$Y(\theta) = \frac{1}{e^{r\cosh\theta} - 1}. \tag{8.5.6}$$

Now substituting this same information into (6.3.43), the equation for $c(r)$, we find

$$c(r) = -\frac{3}{\pi^2}\int_{-\infty}^{\infty} r\cosh\theta \log\left(1 - e^{-r\cosh\theta}\right) d\theta. \tag{8.5.7}$$

With these last two equations, we now have analytical answers for what happens to our solutions for the ground state sinh-Gordon Y-system and the central charge when $p \to 0$.

8.5.3 Limit of the X and X-Y Systems

In the sinh-Gordon model, we are already aware that one cannot use the equation

$$Y(\theta) = e^{-\epsilon(\theta)} \tag{8.5.9}$$

right 'off the bat' since $\epsilon(\theta)$ is not defined outside the original analyticity strip in which the ground state TBA equation is given. Recall that this is due to the sinh-Gordon kernel function ϕ, which depends upon the model-specific S-matrix. There is a singularity on the boundary of the original analyticity strip, and thus we need the X-system to analytically continue ϵ to all of $\mathbb{C}$. Given the fact that we ultimately obtain the Y-system from the X-system, we must also study the X and the X-Y systems to ultimately understand the behavior of the sinh-Gordon system in the small-p limit.

As a reminder, the X-Y system equations are given by

$$X\left(\theta + \frac{ia\pi}{2}\right) X\left(\theta - \frac{ia\pi}{2}\right) = Y(\theta) \tag{8.5.10}$$

and

$$X\left(\theta + \frac{i\pi}{2}\right) X\left(\theta - \frac{i\pi}{2}\right) = 1 + Y(\theta). \tag{8.5.11}$$

Now if one was to take a look at them closely, we seemingly have a problem. Recall that $a \equiv 1 - 2p$. Thus at $p = 0$, $a = 1$. Substituting this into both equations above

lead to an apparent inconsistency. The left-hand sides of both equations are now equal, however, the right-hand sides remain different. How could these equations then be correct?

Recall that the X-system is given by

$$X(\theta) = \exp\left(-\frac{r}{2\sin(\pi p)}\cosh(\theta) + \frac{1}{2\pi}\int_{\mathbb{R}} \frac{1}{\cosh(\theta-\theta')} L(\theta')d\theta'\right). \tag{8.5.12}$$

In the $p \to 0$ limit, the factor

$$-\frac{r}{2\sin(\pi p)}\cosh(\theta) \tag{8.5.13}$$

dominates so that

$$X(\theta) \to \begin{cases} 0 & |\mathrm{Im}\theta| < \pi/2 \\ \infty & \pi/2 < |\mathrm{Im}\theta| < 3\pi/2. \end{cases} \tag{8.5.14}$$

It is this realization that explains the apparent disagreement before between the two X-Y system equations. They are fine because the left-hand sides in (8.5.10) and (8.5.11) cease to be well-defined in the $p \to 0$ limit. In fact, it is actually possible to gain finer control over this limit and obtain an X-Y system which works for the whole range of p-values.

To do this, we define

$$\tilde{X}(\theta) \equiv \exp\left(\frac{1}{2\pi}\int_{\mathbb{R}} \frac{1}{\cosh(\theta-\theta')L(\theta')d\theta'}\right) \tag{8.5.15}$$

for $|\mathrm{Im}\theta| < \frac{\pi}{2}$. This new definition allows us to write the original X-system equation as

$$X(\theta) = e^{\left(-\frac{r}{2\sin(\pi p)}\cosh(\theta)\right)}\tilde{X}(\theta). \tag{8.5.16}$$

Using this equation the X and X-Y systems become the $\tilde{X}$ and $\tilde{X}$-Y systems. We begin by writing down the $\tilde{X}$ system equation

$$\tilde{X}\left(\theta+\frac{i\pi}{2}\right)\tilde{X}\left(\theta-\frac{i\pi}{2}\right) = 1 + e^{-r\cosh(\theta)}\tilde{X}\left(\theta+\frac{ia\pi}{2}\right)\tilde{X}\left(\theta-\frac{ia\pi}{2}\right). \tag{8.5.17}$$

Now solving for the $\tilde{X}$ product on the right-hand side of (8.5.17) yields

$$\tilde{X}\left(\theta+\frac{ia\pi}{2}\right)\tilde{X}\left(\theta-\frac{ia\pi}{2}\right) = e^{r\cosh(\theta)}Y(\theta). \tag{8.5.18}$$

Substituting (8.5.18) into (8.5.17) yields the $\tilde{X}$-Y system,

$$\tilde{X}\left(\theta+\frac{i\pi}{2}\right)\tilde{X}\left(\theta-\frac{i\pi}{2}\right) = 1 + Y(\theta). \tag{8.5.19}$$

Now that we have the new $\tilde{X}$ and $\tilde{X}$-Y system equations, we shall endeavour to

show what happens to these equations when $p = 0$. As before, $p = 0$ forces $a = 1$. Substituting this into the $\tilde{X}$ system equation yields

$$\tilde{X}\left(\theta + \frac{i\pi}{2}\right)\tilde{X}\left(\theta - \frac{i\pi}{2}\right) = 1 + e^{-r\cosh(\theta)}\tilde{X}\left(\theta + \frac{i\pi}{2}\right)\tilde{X}\left(\theta - \frac{i\pi}{2}\right). \tag{8.5.20}$$

Now (8.5.18) becomes

$$\tilde{X}\left(\theta + \frac{i\pi}{2}\right)\tilde{X}\left(\theta - \frac{i\pi}{2}\right) = e^{r\cosh(\theta)}Y(\theta), \tag{8.5.21}$$

and substituting this into (8.5.20) yields

$$\tilde{X}\left(\theta + \frac{i\pi}{2}\right)\tilde{X}\left(\theta - \frac{i\pi}{2}\right) = 1 + Y(\theta). \tag{8.5.22}$$

As one can clearly see, (8.5.19) and (8.5.22) are in fact the same. Thus we have found an X-Y system that is consistent when $p \to 0$. Now if we equate the right-hand sides of (8.5.21) and (8.5.20) we find

$$Y(\theta) = \frac{1}{e^{r\cosh(\theta)} - 1}. \tag{8.5.23}$$

This exact result was found in the previous section when analyzing the Y-system via different means. Thus the previous X-Y system, although not well-defined for the small-p limit, works just fine. Note that (8.5.20) implies

$$\tilde{X}\left(\theta + \frac{i\pi}{2}\right)\tilde{X}\left(\theta - \frac{i\pi}{2}\right) = \frac{1}{1 - e^{-r\cosh(\theta)}}, \tag{8.5.24}$$

but note that this too could be found from either (8.5.21) or (8.5.22) knowing $Y(\theta)$ in this limit.

8.5.4 Iteration in the Small-p Limit

Taking a look at the sinh-Gordon plot in Fig. 8.10, we see a rather simple plot. How now could such a simple plot in the small-p limit be produced by such a seemingly complicated set of functions and iterations? We answer this question by taking a closer look at the TBA equation in the small-p limit. As before, the TBA equation is given by

$$\epsilon(\theta) = r\cosh(\theta) - \frac{1}{2\pi}\int \phi(\theta - \theta')\log\left(1 + e^{-\epsilon(\theta)}\right)d\theta'. \tag{8.5.25}$$

As we take $p \to 0$, we know from above that the term $\frac{1}{2\pi}\phi(\theta - \theta') \to \delta(\theta - \theta')$. Solving for the special case of $\theta = 0$, which implies that $r\cosh(\theta) = r$, our TBA equation becomes

$$e^{\epsilon} = e^{r} - 1, \tag{8.5.26}$$

where we originally solved this problem in (8.5.23). Since we know this exact result from the analytic properties above, we now wish to see whether the same result can be recovered via iterating the TBA.

To solve iteratively, the TBA equation can be written as $\epsilon = F[\epsilon]$ where

$$F[\epsilon] = r\cosh(\theta) - \frac{1}{2\pi}\int_{\mathcal{R}} \phi(\theta-\theta')\log\left(1+e^{-\epsilon(\theta')}\right)d\theta'. \tag{8.5.27}$$

One can immediately notice that if one substitutes equation (8.5.27) into $\epsilon = F[\epsilon]$, we have exactly the TBA equation. Thus so far so good. We then may solve the TBA numerically via an iteration by using $\epsilon_{n+1} = F[\epsilon_n]$. Since this is an iteration, we must worry about convergence. In order to aid convergence, we introduced a damping factor λ in the program with the condition that $0 < \lambda \leq 1$. With this damping factor, we then iterate the equation $\epsilon = G_\lambda[\epsilon]$ where

$$G_\lambda[\epsilon] = (1-\lambda)\epsilon + \lambda F[\epsilon], \tag{8.5.28}$$

with $F[\epsilon]$ defined as before. Unfortunately for us, this new scheme doesn't always converge either! We see this explicitly in our program where we cannot analytically continue in certain regions of the complex plane. What is going on?

Taking the limit when $p \to 0$, we have

$$F[\epsilon] = r\cosh(\theta) - \log\left(1+e^{-\epsilon(\theta)}\right), \tag{8.5.29}$$

as once again $\frac{1}{2\pi}\phi(\theta-\theta') \to \delta(\theta-\theta')$. The iterations at different values of θ decouple. The lack of convergence can be understood by first redefining a term. We proceed by setting $r\cosh(\theta) = \tilde{r}$. For both the F and G iterations the fixed point $\epsilon = \epsilon_\infty$, as we found before, where $e^{\epsilon_\infty} = e^{\tilde{r}} - 1$. For this to be stable for $\epsilon_{n+1} = G_\lambda[\epsilon_n]$ we must have $|G'_\lambda(\epsilon_\infty)| < 1$, i.e.

$$|1-\lambda+\lambda e^{-\tilde{r}}| < 1. \tag{8.5.30}$$

If we identify $\tilde{r} = x + iy$, then we have

$$|1-\lambda+\lambda e^{-x-iy}|^2 = (1-\lambda)^2 + 2\lambda(1-\lambda)e^{-x}\cos(y) + \lambda^2 e^{-2x} < 1, \tag{8.5.31}$$

and condition implies

$$\lambda e^{-2x} + 2(1-\lambda)e^{-x}\cos(y) - 2 + \lambda < 0. \tag{8.5.32}$$

We notice in (8.5.32) that for the equality in (8.5.31) to hold, then e^x must solve a quadratic equation. By making a substitution like $x' = e^{-x}$ and $x'^2 = e^{-2x}$, we can turn (8.5.32) into a quadratic equation of x by taking the logarithm of both sides of

the equation and after setting $x' = x$ to obtain

$$x(y) = -\log\left(\frac{-(1-\lambda)\cos(y) + \sqrt{(1-\lambda)^2\cos^2(y) + (2-\lambda)\lambda}}{\lambda}\right) \tag{8.5.33}$$

This is the solution for our quadratic equation and is the region of the $x-y$ plane to the right of the curve. The point $\theta = 0$, which corresponds to the point $\hat{r} = r$, always lies on the integration contour and so for convergence r must always lie within this region. The smaller the value of λ, the stronger the damping, and the larger the convergence region becomes. However, for $\cos(y) = 0$ and $\lambda \to 0$ we have two limits. Now for the following, we are ultimately going to take the Taylor expansion of the square root. First we need to get the square root into a better form to perform this feat. We will now factor out a term

$$\frac{(1-\lambda)|\cos(y)|}{\lambda} \tag{8.5.34}$$

from the square root. Note that we took the absolute value of $\cos(y)$ because this will allow us to more easily analyze the limiting cases of equation (8.5.32) which involves a multi-valued transedental function. From this factorisation, we obtain

$$x(y) = -\log\left(\frac{(1-\lambda)|\cos(y)|}{\lambda}\left(\sqrt{1 + \frac{(2-\lambda)\lambda}{(1-\lambda)^2\cos^2(y)}} - \sigma\right)\right) \tag{8.5.35}$$

where σ represents the first term in equation (8.5.32) that was not under the square root. Given the fact that we took the absolute value of $\cos(y)$, we note that σ, the sign of the first term, takes one of two values, i.e. $\sigma = \pm 1$. Thus for each case of σ, we have the following approximate behaviors after taking the Taylor expansion of the square root:

$$x(y) \sim \begin{cases} \log(\cos(y)) & \sigma = +1 \\ \log(\lambda/2) - \log|\cos(y)| & \sigma = -1, \end{cases} \tag{8.5.36}$$

where one can view the plots of the functions in Figures 8.26, 8.27, and 8.28. Note again that λ is the damping factor and its range is $0 < \lambda \leq 1$ where $\lambda = 1$ is no damping. From the plots it is abundantly clear that as we decrease the damping factor λ we get a bigger and bigger region of convergence.

The limiting curve for $\sigma = +1$ is independent of λ, meaning that in the $p \to 0$ limit there exist 'forbidden zones' that lie within the strips $|\mathrm{Im}(r) - 2n\pi| < \frac{\pi}{2}, n \in \mathbb{Z}$. In this region the iterative algorithm does not converge despite any attempt to make the damping stronger.

One aspect to note is that we have very recently been discovered that if one makes

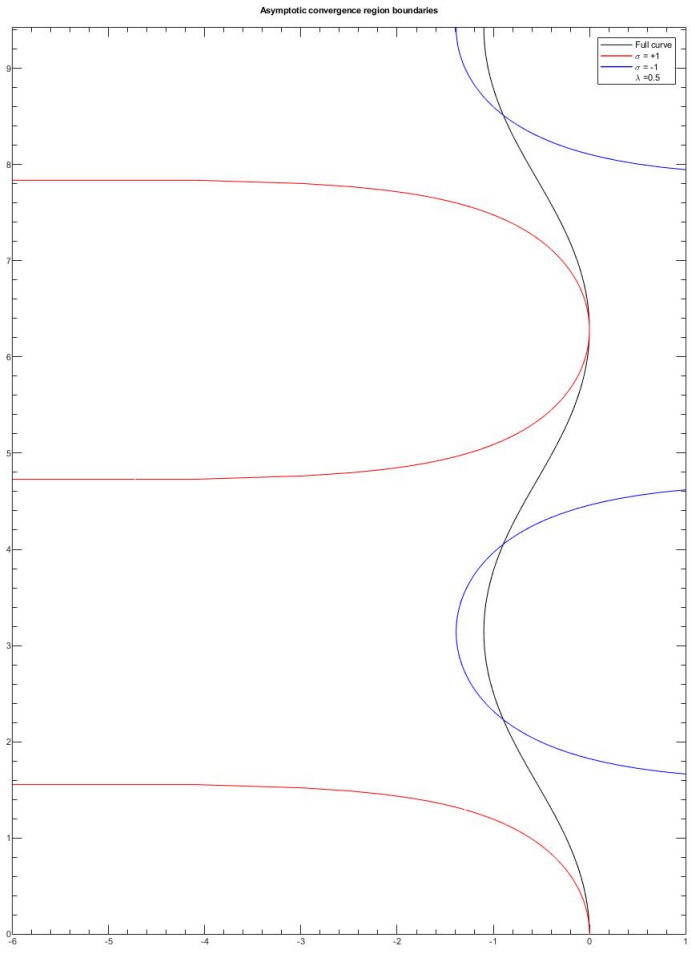


Figure 8.26: Combined convergence plot for $\lambda = 0.5$.

the highly unorthodox assumption that the damping itself is complex, the region of convergence can be improved. Of course given this recent realization, it will have to wait for future work for a proper exploration to be undertaken.

There are many more aspects of the sinh-Gordon model one can study with this method. This will likely be the subject of future work. In the remaining chapter of this thesis, we turn our attention to the 'negative energy problem' and its explanation.

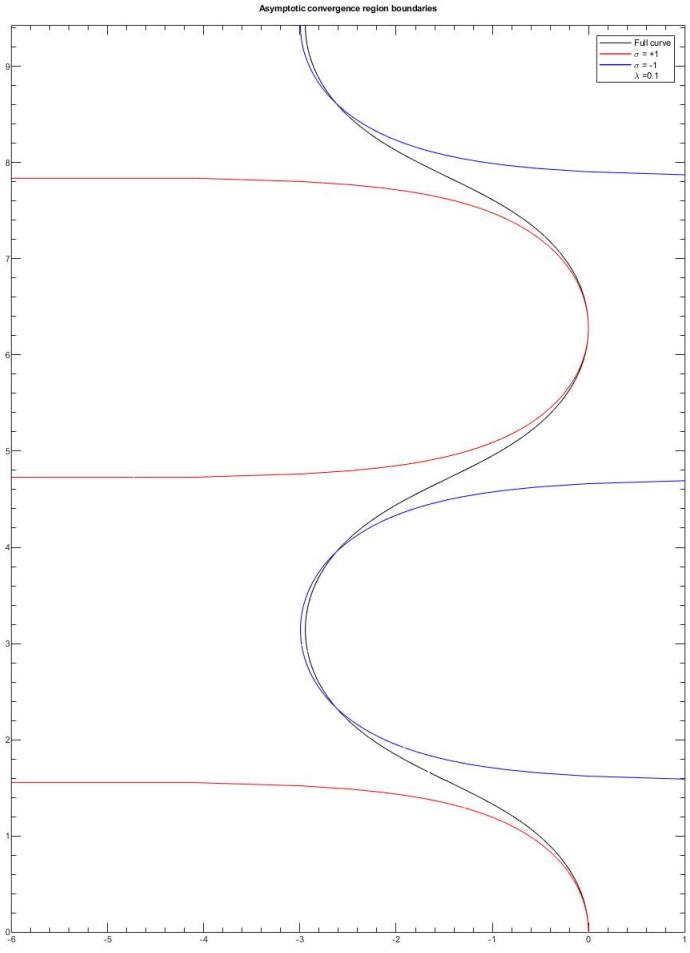


Figure 8.27: Combined convergence plot for $\lambda = 0.1$.

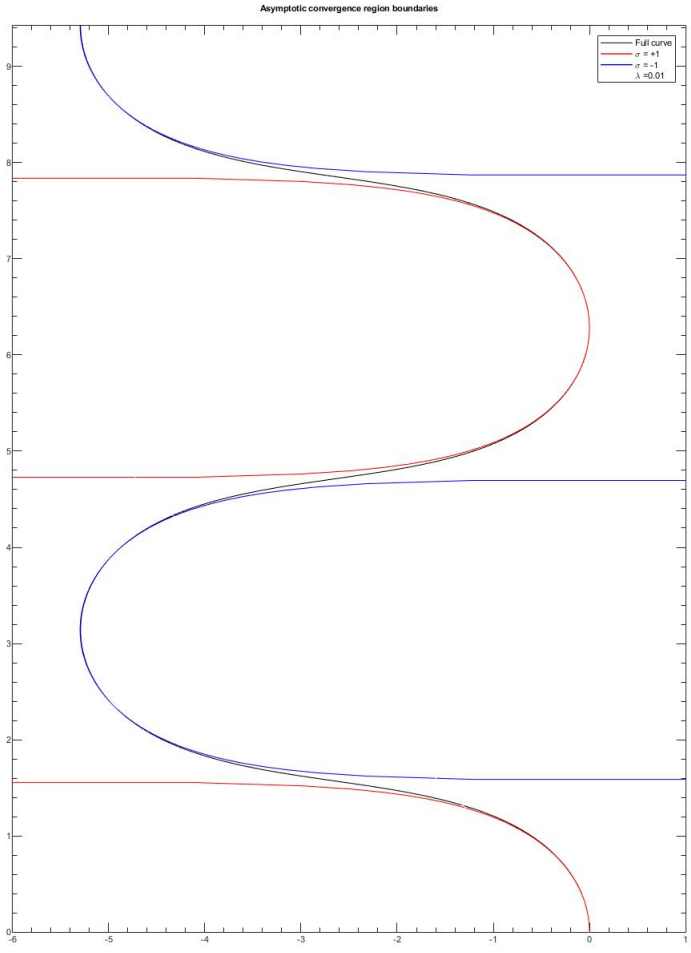


Figure 8.28: Combined convergence plot for $\lambda = 0.01$.

Chapter 9

The Negative Energy Problem

The basic question we set out to answer in this thesis was that supposing we know the ground-state energy of some system, does this information alone yield the energies of the excited states? The answer to that question is often affirmative, that is, if we are able to analytically continue within a parameter of the problem. One can reference [6] for the quantum mechanical case. One must note that this demonstration by Bender and Wu utilizes methods that are of "unknown validity" [29][1]. The point of this thesis was to obtain a similar understanding of this phenomena for a quantum field theory. Since we discovered in Chapter 8 that the sinh-Gordon model reduces to the free boson model in the limit $p \to 0$, it is only natural to apply this technique to that model.

9.1 Ising Field Theory

We return to the free fermion Ising model and the ideas we introduced in Chapter 6. To illustrate where we are going, it will be useful to quickly review this model once more. Recall that the central charge of the Ising field theory is given by

$$c^{Ising}(r) = \frac{1}{2} - \frac{3r^2}{2\pi^2}\left[\log\frac{1}{r} + \frac{1}{2} + \ln\pi - \gamma_E\right] + \frac{6}{\pi}\sum_{k=1}^{\infty}\left(\sqrt{r^2+(2k-1)^2\pi^2} - (2k-1)\pi - \frac{r^2}{2(2k-1)\pi}\right), \tag{9.1.1}$$

and $E_{bulk} \sim R^2 \log R$. For $c(0) = \frac{1}{2}$, note that we have

$$E_0(R \to 0) = \frac{2\pi}{R}\left(d + \bar{d} - \frac{c}{12}\right), \tag{9.1.2}$$

[1] In this paper, Simon and Dicke rigorously demonstrated the results obtained by Bender and Wu actually do hold.

which is correct for the ground state where $d+\bar{d}=0$ in the $c_{1/2}$ CFT, which is found as $R \to 0$.

Now referring to (9.1.1), we start at the ground state and take a path around one of the branch point, similar to Fig. 6.3. For the Ising field theory, the branch points are located at $(2k-1)\pi i$, and they are of square root form. Thus if we take a path around k_1, k_2, etc., this action flips the signs of the square roots of E_0 from positive to negative. Recall upon our return to the real axis, from which we left to analytically continue around the branch point, we obtain

$$E_{k_1,k_2,\ldots} = E_0(m,R) + \frac{2}{R}\sum_{i=1}^{n}\sqrt{r^2+(2k-1)^2\pi^2}. \tag{9.1.3}$$

Note that (9.1.3) is an excited state. So our method does appear to work for the Ising case. Now let us see what happens to the free boson field theory.

9.2 Free Boson Field Theory

We will focus our attention to the effective central charge c_0 for the free boson case. For a free massive boson, the effective central charge is given in [23] and appears as

$$\begin{aligned} c_0(r) = 1 &- \frac{3r}{\pi} + \frac{3r^2}{2\pi^2}\left[\ln\frac{1}{r} + \frac{1}{2} + \ln 4\pi - \gamma_E\right] \\ &- \frac{6}{\pi}\sum_{k=1}^{\infty}\left(\sqrt{(2k\pi)^2+r^2} - 2k\pi - \frac{r^2}{4k\pi}\right). \end{aligned} \tag{9.2.1}$$

In the calculation of $c(r)$, some of the square roots of the infinite sum of square roots flip when encircling a branch point. The infinite sum of square roots for the case of the free massive boson, equation (9.2.1), is *negative*. This is in contrast to the Ising field theory. Rounding a branch point will cause some of these to flip their sign, just like before. Since we started with the opposite sign in front of the sum of square roots, the flip that caused an overall increase in the energy for the Ising field theory will cause an overall decrease in energy for the free boson model.

Any attempt to analytically continue the TBA for the free boson type of particle to obtain the excited states yields an 'excited-energy' that actually has appears to have less energy than the ground state, i.e. yields what we hereafter call '*negative energies*'. This is obviously an issue as the excited state cannot be less energetic than the ground state. The question then naturally becomes why does this technique work for the free fermionic model, but fail for the free boson model? There appears to be a more complex geometric structure to these type of problems than first realized. This section attempts to explore and provide an answer to this question.

9.3 The Coupled Harmonic Oscillator

Studies of the quantum anharmonic oscillator have shown that singularities in the coupling-constant plane are the cause of the divergence in perturbation theory. These singularities are usually square-root branch points [5]. Studies regarding coupling-constant analyticity have divulged a generic phenomenon - the eigenvalues that belong to the spectrum of the Hamiltonian are analytic continuations of one another as functions of the complex coupling constant, as we have mentioned previously. Thus the energy levels of a quantum system, which are discrete when the coupling constant is real and positive, are actually smooth continuations of each other in the complex-coupling-constant plane. With this knowledge, we endeavor to use these branch points to obtain the other eigenvalues of the system.

9.3.1 Solving for the Energy

We proceed by studying the analytic structure of coupled quantum systems. The purpose of this analysis is to show that there exists a complex geometric structure between eigenvalues and their analytic continuations, one not immediately apparent through our usual methods of solution. The model studied by Bender et al. in [5] consists of a pair of simple harmonic oscillators coupled to one another. Its Hamiltonian has the form

$$H = p^2 + \nu^2 x^2 + q^2 + \omega^2 y^2 + gxy, \tag{9.3.1}$$

where p, q are the momenta of the two oscillators respectively, ν, ω are the frequencies of the two oscillators respectively, x, y are the positions of the two oscillators respectively, and g is the coupling constant between the pair of oscillators.

Its eigenfunction has the form

$$\psi(x, y) = e^{-ax^2/2 - by^2/2 + cxy}, \tag{9.3.2}$$

where a,b, and c are constants.

In [5], the eigenfunction is substituted into the Hamiltonian and solved for the constants. It is then substituted back into the original eigenvalue equation, solving for the energy E to obtain

$$E(g) = \left[\nu^2 + \omega^2 + \left[4\nu^2\omega^2 - g^2\right]^{\frac{1}{2}}\right]^{\frac{1}{2}}. \tag{9.3.3}$$

A clearer picture than that given in [5] is found by first diagonalizing the Hamiltonian.

We begin by writing (9.3.1) in the form

$$H = \sum_{j=1}^{2} \left(\hat{p}_j^2 + \nu_j^2 + g\hat{x}_{j+1}\hat{x}_j \right). \tag{9.3.4}$$

Note here that in this form we have the frequencies given by ν_j and g is the coupling between the two 'nearest-neighbor' oscillators x_{j+1} and x_j. This form makes it a bit easier to directly apply the Fourier transforms to diagonalize the problem.

Next we diagonalize (9.3.4) by taking the Fourier transform,

$$\hat{x}_j = \frac{1}{\sqrt{2}} \sum_{k}^{2} \tilde{\hat{x}}_k e^{ikja} \tag{9.3.5}$$

and

$$\hat{p}_j = \frac{1}{\sqrt{2}} \sum_{k}^{2} \tilde{\hat{p}}_k e^{ikja}. \tag{9.3.6}$$

Since $\hat{p}_k$ and $\hat{x}_j$ are Hermitian, this means $\hat{p}_k^\dagger = \hat{p}_{-k}$ and $\hat{x}_k^\dagger = \hat{x}_{-k}$. Now we define creation and annihilation operators which are given by

$$\hat{a}_k = \sqrt{\frac{\omega_k}{4}} \left(\hat{x}_k + \frac{i}{\omega_k} \hat{p}_k \right), \tag{9.3.7}$$

and

$$\hat{a}_k^\dagger = -i\sqrt{\frac{\omega_k}{4}} \left(\hat{x}_{-k} - \frac{i}{\omega} \hat{p}_{-k} \right), \tag{9.3.8}$$

From these annihilation and creation operators we can solve for the operators $\hat{x}_k$ and $\hat{p}_k$

$$\hat{x}_k = \sqrt{\frac{1}{4\omega_k}} \left(\hat{a}_k + \hat{a}_{-k}^\dagger \right), \tag{9.3.9}$$

and

$$\hat{p}_k = -i\sqrt{\frac{\omega_k}{4}} \left(\hat{a}_k - \hat{a}_{-k}^\dagger \right), \tag{9.3.10}$$

we are ultimately able to obtain the diagonalized form of Bender's Hamiltonian,

$$H = \sum_{k}^{2} \omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2} \right). \tag{9.3.11}$$

with $w_k^2 = \nu_1^2 + \nu_2^2 + g e^{ika}$. The energy eigenvalues are given by

$$E = \sum_{k}^{2} \omega_k \left(n_k + \frac{1}{2} \right), \tag{9.3.12}$$

where the n_k are the different energy levels.

Bender's original problem had two oscillators, so we allow $k = 0, 1$ here. We hold the frequencies constant, setting them to $\nu_1 = 1$ and $\nu_2 = 2$. We then select appropriate factors to simplify the form, and interpreting $a = \pi$ to treat the exponential as a phase term, we obtain for the energy the system

$$E = \frac{\pm\sqrt{1+\frac{g}{2}} \pm \sqrt{1-\frac{g}{2}}}{\sqrt{2}}. \tag{9.3.13}$$

This approach yields the same Riemann surface as Bender's original form does, but it is now quite clear what is going on. As we analytically continue the coupling constant around an appropriate branch point, we flip one of the signs in front of the square roots. In doing so, we will have transversed onto a different sheet of a Riemann surface. Since there are four choices overall, there are four sheets in this Riemann surface, as can be seen via Fig. 9.1. We find that the four sheets of the Riemann surface correspond to four distinct spectral phases of this coupled oscillator system. Thus four distinct values for the energy can be obtained. The single conventional spectrum is the one with both roots positive. The three other spectra are the unconventional spectra.

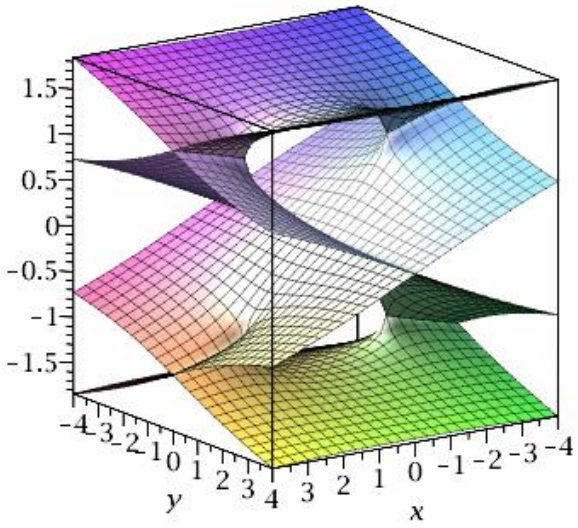


Figure 9.1: Riemann surface of Bender's original problem. The Riemann surface has four sheets, and rounding a branch point allows one to transverse the surface. Note here that this figure is a plot of the energy as a function of (Re(g), Im(g)).

This means that the coupled oscillator actually has a conventional spectrum and

three unconventional ones. As we have just seen, transversing the entire Riemann surface has allowed us access to both conventional and unconventional phases of the coupled harmonic oscillator. Note that this held true even though we had fixed the two frequency parameters ν and ω in Bender's original problem (9.3.1). These spectra are both real and are related by analytic continuation in the natural frequency of the oscillator. Thus in principle via analytically continuing the coupling constant of a coupled quantum theory, we can end up in states that have a lower energy than the ground state of a given theory! Is this what is happening with the free boson model? We must investigate this further.

9.3.2 The Free Boson Revisted

Modified Bender Problem for Field Theory

The "Bender problem" described in the previous section was certainly illuminating. However his problem was more quantum mechanical and not field theoretic. Thus to proceed with our analysis and figure out what is happening with the free boson model, we must proceed by making the couple harmonic oscillator problem of Bender's a bit more field theory friendly.

From the Bender paper, we have

$$-\Psi_{xx}+\nu^2x^2\Psi-\Psi_{yy}+\omega^2y^2\Psi+gxy\Psi=E\Psi, \tag{9.3.14}$$

with

$$\Psi(x,y)=e^{-\frac{ax^2}{2}-\frac{by^2}{2}+cxy}. \tag{9.3.15}$$

Now substituting (9.3.15) into (9.3.14) yields

$$a(2cxy+1)-c^2y^2-a^2x^2+b(2cxy+1)-c^2x^2-b^2y^2+\nu^2x^2+\omega^2y^2+gxy=E. \tag{9.3.16}$$

From (9.3.16), for a field theory we now make the change,

$$gxy\to g(x-y)^2 \quad \text{and} \quad \nu=\omega, \tag{9.3.17}$$

which yields

$$a(2cxy+1)-c^2y^2-a^2x^2+b(2cxy+1)-c^2x^2-b^2y^2+\nu^2(x^2+y^2)+g(x-y)^2=E. \tag{9.3.18}$$

Now we equate the coefficients of the x^2, y^2, xy, and x^0y^0 terms to obtain

$$x^2: \quad -a^2-c^2+\nu^2+g=0 \tag{9.3.19}$$
$$y^2: \quad -b^2-c^2+\nu^2+g=0 \tag{9.3.20}$$

$$xy: \qquad 2ac + 2bc - 2g = 0 \tag{9.3.21}$$
$$x^0y^0: \qquad a + b = E \tag{9.3.22}$$

With $\nu = \omega$, the single change from the Bender system with $\nu \neq \omega$ is (9.3.20) above, which now when subtracted from (9.3.19) above yields

$$a^2 = b^2, \tag{9.3.23}$$

where we take the principle square root, as not doing so will yield $c = \frac{0}{0}$ and $g = 0$ in the set of four coefficient equations above. Thus we have

$$a = b \tag{9.3.24}$$

Using (9.3.24) in (9.3.22) yields

$$a = \frac{E}{2} \tag{9.3.25}$$

Using (9.3.24) again, we arrive at

$$b = \frac{E}{2}. \tag{9.3.26}$$

For c, we substitute (9.3.22) into (9.3.21) to obtain

$$c = \frac{g}{E}. \tag{9.3.27}$$

Now we plug-in a, b, and c into (9.3.18) and simplify to obtain

$$E_{\pm}^2 = 2(\nu^2 + g) \pm 2[\nu^2(\nu^2 + 2g)]^{\frac{1}{2}}. \tag{9.3.28}$$

Taking the square-root of the above yields

$$E = \pm[2(\nu^2 + g) \pm 2[\nu^2(\nu^2 + 2g)]^{\frac{1}{2}}]^{\frac{1}{2}}. \tag{9.3.29}$$

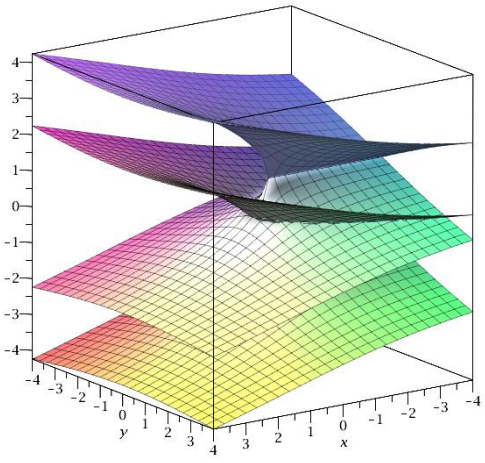
Note that one can complete a similar analysis for this situation as we did for Bender's original problem. Instead we will analyze this case slightly differently than the unmodified version in order to properly visualize the Riemann surface.

We analyzed the original Bender system while holding the frequencies constant. In this case, we set all frequencies equal. We will now set our frequency equal to unity, as the shape and structure of the Riemann surface will still be preserved, and yet it will be simpler to analyze. Setting $\nu = 1$, we have

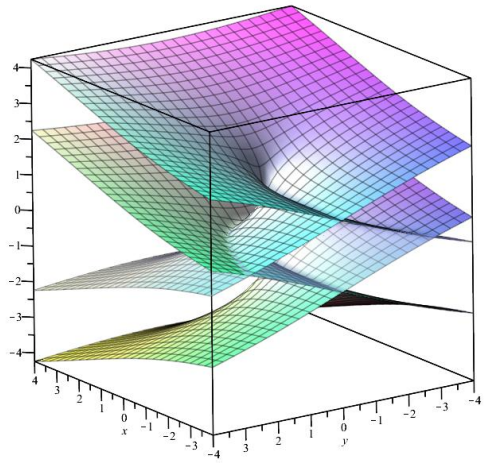
$$E = \pm\sqrt{2 + 2g \pm 2\sqrt{1 + 2g}}. \tag{9.3.30}$$

We can then individually plot each of the four sheets in (9.3.30) and combine them to form the full Riemann surface, and we obtain Fig. 9.2. We once again have a connected Riemann surface with four sheets, as expected from (9.3.30).

Just as in the former case for the unmodified coupled harmonic oscillator, we will



(a)



(b)

Figure 9.2: Riemann surface of the modified Bender problem which is closer to field theory. Observe a connected Riemann surface that has four sheets. The associated Hamiltonian has three unconventional spectra. The same Riemann surface is seen in both pictures here to show the full shape as the surface is rotated.

have three unconventional spectra for the Hamiltonian in this case as well. Thus by analytically continuing the coupling constant, we can end up with an energy in the excited state that is lower that the ground state energy.

This is all beautiful, but do the changes we made to the Bender problem actually model the free boson model? We turn to the coupled oscillator chain for further

analysis.

Coupled Oscillator Chain

Given that the free boson model is similar to the model for the coupled harmonic oscillator chain, we shall now explore this problem.

The problem consists of the following information:

- N - number of coupled oscillators
- a - distance between the oscillators
- g - coupling parameter
- ja - position of oscillators
- Dealing with QFT, so all frequencies equal and should have form $K(\hat{x}_{j+1}-\hat{x}_j)^2$
- We will include a coupling term $\omega^2\hat{x}_j^2$

We have the coupled harmonic oscillator chain Hamiltonian given via

$$\hat{H} = \sum_j \frac{\hat{p}_j^2}{2m} + \omega^2\hat{x}_j^2 + \frac{K}{2}\left(\hat{x}_{j+1} - \hat{x}_j\right)^2 \tag{9.3.31}$$

where K here is the coupling constant. This is done for very specific reasons, which will be apparent as the problem is solved.

We will diagonalize the problem as we did for the Bender's Hamiltonian earlier. We once again begin by Fourier transforming the coordinates p_j and x_j in (9.3.31) via

$$\hat{x}_j = \frac{1}{\sqrt{N}}\sum_k \tilde{x}_k e^{ikja} \tag{9.3.32}$$

$$\hat{p}_j = \frac{1}{\sqrt{N}}\sum_k \tilde{p}_k e^{ikja} \tag{9.3.33}$$

We substititute the Fourier transforms for p_j and x_j into (9.3.31) and following the same procedure as before, only this time using

$$\hat{x}_k = \sqrt{\frac{1}{2m\omega_k}}\left(\hat{a}_k + \hat{a}_{-k}^\dagger\right), \tag{9.3.34}$$

and

$$\hat{p}_k = -i\sqrt{2m\omega_k}\left(\hat{a}_k - \hat{a}_{-k}^\dagger\right). \tag{9.3.35}$$

These relations ultimately yields the same diagonalized Hamiltonian as above, i.e.

$$H = \sum_k \omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2} \right). \tag{9.3.36}$$

As before, acting on the eigenstates via the Schrödinger equation allows us to obtain the energy

$$E = \sum_k \omega_k \left(n_k + \frac{1}{2} \right), \tag{9.3.37}$$

with $\omega_n^2 = \omega^2 + \frac{4K}{m} \sin^2 \left(\frac{ka}{2} \right)$ and where n_k refers to the number of oscillators in each state, which we obtain due to acting on the number operator $\hat{N} = a_k^\dagger a_k$ with the Schrödinger equation.

To continue our analysis and turn this into a field theory, we consider a system of two particles, so we sum over $k = 0, 1$ to obtain

$$E = \omega_0 \left(n_0 + \frac{1}{2} \right) + \omega_1 \left(n_1 + \frac{1}{2} \right). \tag{9.3.38}$$

We wish to work with the ground state, so we have for both $n_0, n_1 = 0$,

$$E = \omega_0 \left(0 + \frac{1}{2} \right) + \omega_1 \left(0 + \frac{1}{2} \right). \tag{9.3.39}$$

Now we have for ω_0

$$\omega_0 = \left(\omega^2 + \frac{4K}{m} \sin^2 \left(\frac{(0)a}{2} \right) \right)^{\frac{1}{2}} = \omega, \tag{9.3.40}$$

where we used the fact that $\sin(0) = 0$.

We also have for ω_1

$$\omega_1 = \left(\omega^2 + \frac{4K}{m} \sin^2 \left(\frac{(1)a}{2} \right) \right)^{\frac{1}{2}} \approx \left(\omega^2 + \frac{4K}{m} \frac{a^2}{4} \right)^{\frac{1}{2}} = \left(\omega^2 + \frac{K}{m} a^2 \right)^{\frac{1}{2}}, \tag{9.3.41}$$

where the small-angle approximation was used for sine.

As we wish to deal with a field theory, we may relate the mass and the coupling constant K to the lattice as $m \equiv \rho a$ and $K = \frac{2g}{a}$, where a is the spacing between the oscillator sites in the harmonic chain, ρ is the mass density, and g is the new coupling constant. This yields

$$\omega_1 = \left(\omega^2 + \frac{1}{a\rho} \frac{2g}{a} a^2 \right)^{\frac{1}{2}} = \left(\omega^2 + \frac{1}{\rho} 2g \right)^{\frac{1}{2}} = \left(\omega^2 + 2g \right)^{\frac{1}{2}}. \tag{9.3.42}$$

for the mass density $\rho = 1$. Now using (9.3.40) and (9.3.42) in (9.3.39) to obtain

$$E = \frac{1}{2}\omega + \frac{1}{2} \left(\omega^2 + 2g \right)^{\frac{1}{2}}. \tag{9.3.43}$$

Now we want to square (9.3.43), and we obtain

$$E^2 = \frac{1}{4}\left(\omega^2 + 2\omega\left(\omega^2 + 2g\right)^{1/2} + \left(\omega^2 + 2g\right)\right) \tag{9.3.44}$$

$$= \frac{1}{4}\left(2(\omega^2 + g) + 2[\omega^2(\omega^2 + 2g)]^{1/2}\right). \tag{9.3.45}$$

Now taking square-root yields

$$E = \frac{1}{2}\left(2(\omega^2 + g) + 2[\omega^2(\omega^2 + 2g)]^{1/2}\right)^{1/2}. \tag{9.3.46}$$

Multiplying arbitrarily by a constant to rid ourselves of the leading factor leaves us with

$$E = \left(2(\omega^2 + g) + 2[\omega^2(\omega^2 + 2g)]^{1/2}\right)^{1/2}. \tag{9.3.47}$$

Now comparing the above form to the case where $\nu = \omega$ and $g(x-y)^2$ in (9.3.29) shows an exact match. Note that we also arrived at these equations following two separate methods of analysis. Because of this exact match in the form of our energy, this means that the unusual results for the modified Bender problem, i.e. the unconventional states, may indeed be accessed by the free boson model as well! This still doesn't tell us exactly why fermionic models work and the free boson fails. Thus we must mathematically answer this conundrum.

9.4 Stokes Sectors

How could there even be a negative spectrum at all? In order to understand what is going on we will have to turn to an idea that has its origin quantum mechanics, but first we shall introduce the theory of the Stokes phenomenon needed for a full understanding.

Conventional and Unconventional Spectra

So returning to our initial question of how any Hamiltonian could have more than one spectrum, the answer boils down to how we impose our boundary conditions. The positive spectrum is obtained when one imposes the boundary conditions in a pair of Stokes wedges which have a center about the positive and negative real axes. This is the *conventional* one we are used to in quantum theory and QFT. What about the negative one? We obtain the *unconventional*, or negative, spectrum by imposing the boundary conditions in a pair of Stokes wedges centered along the upper and lower imaginary axes.

In the case of the harmonic oscillator, these Stokes wedges have an angular opening of $\pi/2$. In order to make sense of the configuration of the wedges, we must turn to an old approximation method familiar from quantum mechanics.

JWKB Approximation

As was introduced in single-particle quantum mechanics, the JWKB approximation is a method applied to the Schrödinger equation (a linear differential equation) to find the solutions (the wavefunctions) when one has spatially varying coefficients. It assumes the potentials are slowly varying and is the necessary approximation when we wish to know the behavior as $|x| \to \infty$.

Essentially the JWKB approximation helps us construct solutions to the static Schrödinger equation. We will attempt to apply this logic to the equation for the harmonic oscillator. In one-dimension the Schrödinger equation is given via

$$\left(-\frac{d^2}{dx^2} + P(x)\right)\psi(x) = 0, \tag{9.4.1}$$

where $P(x)$ is a function that depends on the problem being considered. For example, the function $P(x)$ in the oscillator problem discussed in [17] is given by

$$P(x) = (\nu x)^{2N} - E, \tag{9.4.2}$$

where $N = 1$ corresponds to the harmonic oscillator.

Traditionally in quantum mechanics, solutions to this equation are only considered along the positive and negative real axes, but we will consider more general options.

Normally we wish for our solutions to decay at $|x| \to \infty$ and thus we expect the set of permitted energies to quantize.

The JWKB approximation states that as $|x| \to \infty$,

$$\psi(x) \sim \frac{1}{P(x)^{1/4}}\, e^{\pm \int^x \sqrt{P(x)}\, dx}. \tag{9.4.3}$$

Given the form of (9.4.3) above we note that for an oscillator problem

$$\pm \int^x \sqrt{P(x)} dx = \pm \int^x \sqrt{(\nu x)^{2N} - E}\, ds \tag{9.4.4}$$

$$\approx \pm \int^x \nu^N (x)^N dx \tag{9.4.5}$$

$$= \pm \frac{\nu^N x^{N+1}}{N+1}. \tag{9.4.6}$$

This implies that our wavefunction in (9.4.3) is now approximated by

$$\psi(x) \sim \frac{1}{P(x)^{1/4}} \exp\left\{ \pm \frac{\nu^N x^{N+1}}{N+1} \right\}, \tag{9.4.7}$$

as $|x| \to \infty$ for $x \in \mathbb{C}$.

We now have an approximate wavefunction, but it tells us nothing so far about the boundary conditions of the problem, nor does it tell us why we access the negative energies for the free boson case. To continue the analysis the key is to examine the behaviour of solutions as $|x| \to \infty$ along a general ray in the complex plane. This is, of course, in spite of the fact that we initially only defined our problem along two rays (the negative and positive real axes).

If we want to obtain any ray in the complex plane, the form x must take is obviously

$$x = \rho e^{i\theta}. \tag{9.4.8}$$

To continue, we will substitute (9.4.8) into (9.4.7) to obtain

$$\psi_{\pm}(x) \sim \frac{1}{P(x)^{1/4}} e^{\pm \frac{e^{(N+1)i\theta} \rho^{(N+1)} \nu^N}{N+1}}. \tag{9.4.9}$$

Now the term we are most interested in is the nested exponential in (9.4.9). This term is

$$e^{(N+1)i\theta}, \tag{9.4.10}$$

and for general θ it has a non-zero real-part. Thus one solution out of $\psi(x)_{\pm}$ grows and the other shrinks as $\rho \to 0$. The growing solution is the *dominant* solution and the decaying one is the *subdominant* solution. If we want the solution to be bounded, we must ensure that our solution only contains the subdominant part of the solution.

However if the condition

$$\mathrm{Re}\left(e^{(N+1)i\theta}\right) = 0, \tag{9.4.11}$$

holds, both solutions, ψ_+ and ψ_-, will oscillate as $\rho \to \infty$ and neither solution will dominate the other. This condition (9.4.11) defines the so-called anti-Stokes lines. Note that the Stokes lines are given by

$$\mathrm{Im}\left(e^{(N+1)i\theta}\right) = 0, \tag{9.4.12}$$

though we mention this for completeness and it will hereafter not concern us for our purposes. The anti-Stokes lines specifically are what we are concerned about. Note that when we cross an anti-Stokes line, the dominant and subdominant solutions swap. The solution that was once dominant is now subdominant and vice versa. Note that this swap will change the nature of the eigenvalue problem, thus in any analytic

continuation it is absolutely imperative that the rays along which the eigenfunctions are defined do not cross anti-Stokes lines.

One must note here an ambiguity in the terminology. Depending on the author, what we defined as the Stokes and anti-Stokes lines are oppositely defined, such as in works by Bender. Thus in such works what we called the Stokes lines would be called the anti-Stokes lines and vice-versa. We follow the convention used by Stokes himself.

Looking at (9.4.2), we note that setting $N = 1$ yields the correct form of the harmonic oscillator. Thus for the harmonic oscillator (9.4.11) becomes

$$\mathrm{Re}\left(e^{2i\theta}\right) = 0. \tag{9.4.13}$$

Analyzing (9.4.13), we see that the equation is satisfied when $2\theta = \pm\frac{\pi}{2}, \pm\frac{3\pi}{2}$, etc. Solving this yields

$$\theta = \pm\frac{\pi}{4}, \pm\frac{3\pi}{4}, ... \tag{9.4.14}$$

Thus these are the locations of the anti-Stokes lines for the harmonic oscillator, see Fig. 9.3.

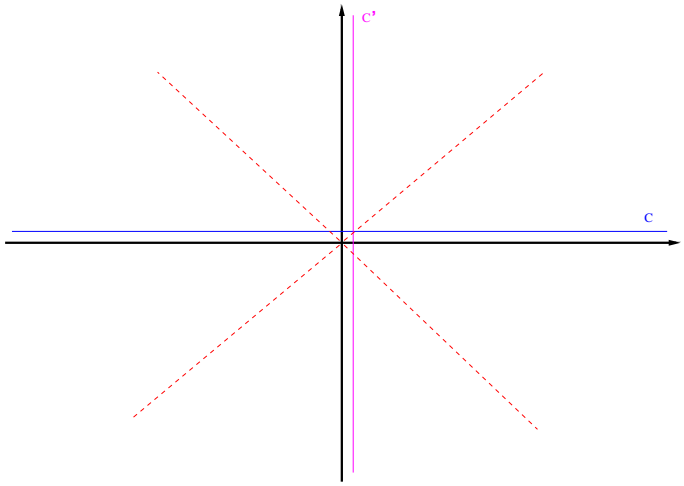


Figure 9.3: Diagram of the anti-Stokes lines for the harmonic oscillator. The broken red lines are the the anti-Stokes lines. The blue line is the original good contour C on which our solution lies. The magenta line is an alternative contour C' along which one obtains negative solutions.

This is great, but recall that the issue with the free boson was found only after we analytically continued our parameter in the complex plane, so we are still missing something. To continue our analysis, we will make use of analytic continuation for a particular parameter of the problem. Note that for the sinh-Gordon the only

parameter in the problem, and the parameter for which we continued the TBA, is r. For Bender's initial problem, the parameter he continued was the coupling constant g between two oscillators in his 2-body problem. For our problem ν is the parameter we use to analytically continue in the complex plane. Let us consider the substitution for continuation given by

$$\nu = e^{-i\gamma}. \tag{9.4.15}$$

We will define a path $\gamma \in [0, \pi]$. Now let us analyze the case at the start of our path where $\gamma = 0$, i.e.

$$\nu(\gamma) = \nu(0) = e^{i(0)} = 1. \tag{9.4.16}$$

Since ν is squared in (9.4.2), this means we must square the result of (9.4.16) to obtain

$$\nu(0)^2 = 1. \tag{9.4.17}$$

Now at the end of our path we have $\theta = \pi$, so this implies

$$\nu(\gamma) = \nu(\pi) = e^{i(\pi)} = -1. \tag{9.4.18}$$

Since ν is squared as before, this means we must square the result of (9.4.18) to obtain

$$\nu(\pi)^2 = 1. \tag{9.4.19}$$

So as γ is continued from 0 to π in the harmonic oscillator, (9.4.1) returns to itself after the continuation. So far so good, as this is what we would normally hope for when performing an analytic continuation in a parameter. But there is this other business regarding the Stokes theory introduced. The very same continuation path in ν rotates the Stokes sectors by only a quarter. Now how did we arrive at this fact? The example of the anti-Stokes lines above was for when ν was real, so γ was fixed and equal to zero.

Once we do vary γ in our parameter ν, we must modify (9.4.9), which now has the form

$$\psi(x)_{\pm} \sim \frac{1}{P(x)^{1/4}} e^{\pm \frac{e^{(N+1)i\theta} \rho^{(N+1)} e^{-i\gamma N}}{N+1}}. \tag{9.4.20}$$

Now when we attempt to solve for the anti-Stokes lines, our condition becomes

$$\mathrm{Re}\left(e^{(N+1)i\theta} e^{-Ni\gamma}\right) = 0. \tag{9.4.21}$$

For the harmonic oscillator we have $N = 1$ and the condition is

$$\mathrm{Re}\left(e^{i(2\theta-\gamma)}\right) = 0. \tag{9.4.22}$$

Let us set $\alpha = \theta - \frac{\gamma}{2}$ which implies $2\alpha = 2\theta - \gamma$. Using this substitution in (9.4.22),

we obtain

$$\alpha = \pm\frac{\pi}{4}, \pm\frac{3\pi}{4}, ... \tag{9.4.23}$$

Hence the directions of the anti-Stokes lines satisfy

$$\theta - \frac{\gamma}{2} = \pm\frac{\pi}{4}, \pm\frac{3\pi}{4}, ... \tag{9.4.24}$$

Let us now restrict ourselves to the initial term on the right to write down an equation for the anti-Stokes lines θ,

$$\theta = \pm\frac{\pi}{4} + \frac{\gamma}{2} \tag{9.4.25}$$

Thus to obtain the anti-Stokes lines for varying γ, we solve the above equation for θ for a given γ. As above we use $\gamma = 0, \pi$ to obtain

$$\theta = \pm\frac{\pi}{4} + \frac{\gamma}{2} = \pm\frac{\pi}{4} + \frac{0}{2} = \pm\frac{\pi}{4} \tag{9.4.26}$$

and

$$\theta = \pm\frac{\pi}{4} + \frac{\gamma}{2} = \pm\frac{\pi}{4} + \frac{\pi}{2} = \frac{\pi}{4}, \frac{3\pi}{4} \tag{9.4.27}$$

where we restricted ourselves to the initial term, as it determines the entire pattern henceforth. Note that for (9.4.27) $\frac{3\pi}{4}$ is part of the same series as in (9.4.22). Thus we see quite clearly that we have quarter rotation for the anti-Stokes lines for the harmonic oscillator for the path $\gamma \in [0, \pi]$. Due to the addition of γ on the RHS of (9.4.25), our anti-Stokes lines will rotate in the counterclockwise direction (if we were subtracting, then the rotation would occur in the clockwise direction). Thus even though the Stokes sectors are rotated, they do *not* return to their same initial positions. This is in contrast to the form of the Schrödinger equation in (9.4.1) returning to itself after the same continuation in ν. This is the heart of what is happening in the free boson case.

In quantum mechanics we normally ask that the solution decays, i.e. $\psi \to 0$, as $|x| \to \infty$ along C, see Fig. 9.3. For the simple harmonic oscillator, this requirement yields the positive eigenvalue. We could have equally asked for our solutions to have decayed along C' instead. It turns out, via a change of variable, that the eigenvalues are negative along C'.

In our case, as we mentioned before, (9.4.1) returns to itself with a continuation path in ν where $\gamma \in [0, \pi]$. According to (9.4.25), the anti-Stokes lines, are rotated as in Fig. 9.4. Note in this figure γ is an intermediate value between 0 and π. Since we have chosen a path $\gamma \in [0, \pi]$, γ will assume all values between 0 and π. Thus the Stokes sectors change their positions with every γ. Since our full path for ν ended with $\gamma = \pi$, this means that the final full rotation for the anti-Stokes lines are a

quarter as seen in Fig. 9.5, and the contour C must also rotate.

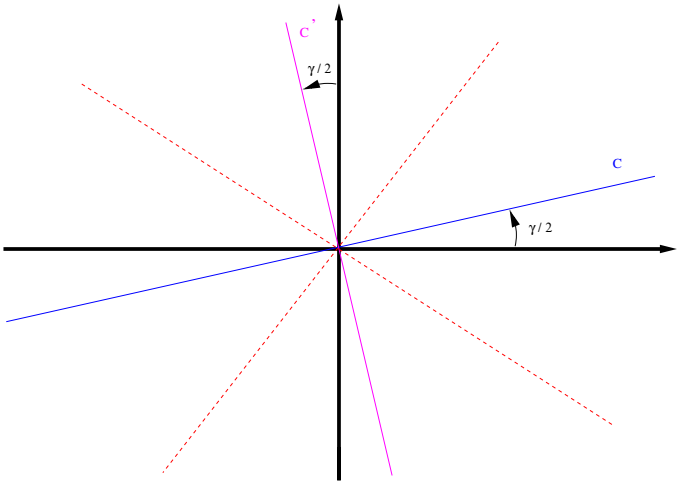


Figure 9.4: Diagram of the rotated anti-Stokes lines for the harmonic oscillator. The broken red lines are the the anti-Stokes lines. The blue line is the original good contour C on which our solution lies. The positions of the anti-Stokes lines are at an intermediate value of γ between 0 and π.

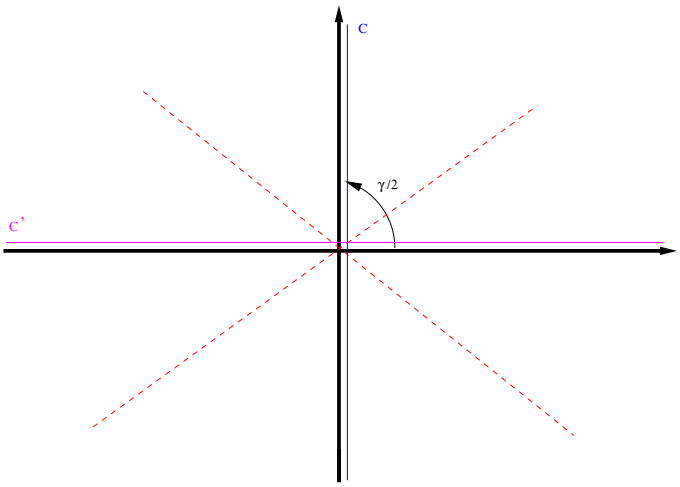


Figure 9.5: Diagram of the rotated anti-Stokes lines for the harmonic oscillator at $\gamma = \pi$. Now C has rotated around to where C' was in Fig. 9.4.

Since the rotation is only a quarter, the solutions now decay along C' instead of C. Solutions that decay along C' are negative and this is the reason why we can analytically continue our parameter and reach a new phase of the harmonic oscillator whose spectrum is negative and unbounded below.

Normally when dealing with a finite matrix eigenvalue problem, if one analytically continues the problem, one winds up with another eigenvalue of the problem. We obtain the negative spectrum due to an interplay with the Stokes phenomenon and the boundary conditions of the problem at hand. Thus for quantum mechanics and quantum field theory, the differential operator alone does not govern the solution, the boundary conditions play an equally important role as well.

One must note that there are actually two eigenvalue problems for (9.4.1) with $\nu^2 x^2 - E$. Depending on the choice of Stokes sectors for the quantization contour, i.e. either $\psi \in L^2(\mathbf{R})$ or $\psi \in L^2(\mathbf{iR})$. Thus even if a continuation leaves the Hamiltonian invariant, we must check what it does with the boundary conditions, as these could be rotated and the Stokes sectors shifted.

Sinh-Gordon and the Free Boson

The example provided above explains the case for a single harmonic oscillator. The case for the free boson can be represented as an oscillator chain, i.e. we will have a coupled series of oscillators. It turns out the Stokes analysis for even the two coupled oscillator problem is even more complex. A full and complete answer would consist of working on the Stokes analysis for the coupled oscillator problem and then generalize it for the infinite oscillator case, but this in itself has proven to be a major undertaking.

There are quite a few unanswered questions and more work is needed to be confident that a correct and full explanation has been achieved. Although we have made a major headway, unfortunately the rest (and more) of this work must be left for the future. However direct continuation of the eigenvalues for the coupled harmonic oscillator problem, as discussed above, shows that in this case the continuation leading to negative eigenvalues is indeed associated with a flip of the boundary conditions and so we can be confident that this explains the emergence of negative energy states in the free boson example.

So we determined that indeed this method does fail for the free boson model and explained that the reason for this is due to a change of boundary conditions under analytic continuation for the harmonic oscillator case. Why then are we not worried that it fails somewhere for the sinh-Gordon model? The simple fact is that the growth of the potential $\cosh(\beta\phi)$ at each point x is much stronger than the growth of other terms in the sinh-Gordon Hamiltonian. Thus the boundary conditions should also be governed by this dominant potential as well and the continuation in r, which does not affect the potential, should leave the boundary conditions untouched. This idea is further supported by the plethora of numerical work detailed in Chapter

8 for the sinh-Gordon, where we observed this technique to hold for various limits. However more work will be needed for a direct confirmation of this explanation.

Chapter 10

Conclusion

In this thesis we have shown that armed with only the exact S-matrix of a QFT, one can derive an integral equation using the TBA equation to compute the ground state energy of an integrable model. One can then analytically continue the TBA equation along an arbitrary path in the complex plane of the given model, parameterized by r, which implicitly depends upon the coupling-constant. Thus we are analytically continuing via the coupling constant. Upon encircling a branch point of the model in the complex plane, we can obtain another energy level. This corresponds to another sheet on the Riemann surface of the integrable QFT model in question. On this sheet an excited state TBA equation holds. Further excited states are then found by repeating this process.

Finally we also covered the so-called 'negative-energy problem'. In particular cases of quantum field theories, the method in this thesis can yield an unphysical, negative energy spectrum. This is due to a flip in the boundary conditions under analytic continuation. We additionally indicated why for the sinh-Gordon model we can expect that the method outlined in this thesis works, while it fails for the case of the free boson.

It should be noted that the method outlined in this thesis is by no means a 'cute' trick. It is a serious and very general method for obtaining the spectrum of an integrable, interacting quantum field theory. Furthermore this method allows one to study an integrable model topologically via its Riemann surface. This technique not only can reproduce information found via other means, it can produce knowledge elusive in other techniques. For example, it is far less likely we would have stumbled upon the existence of the Stokes phenomenon in QFT via another means. This technique, being exceptionally general, should allow for the study of a whole host of integrable models that might not be accessible via other methods of solution. Another paper by Dorey and Tateo [19] explored analytic continuation in some other simple perturbed

conformal field theories. The regularities we observed in the sinh-Gordon model in the $p \to 0$ limit are interesting as they offer hope of obtaining a complete picture of the model, while in the earlier papers only a few levels were considered.

Since this method has now been successfully demonstrated for the Yang-Lee model and other perturbed conformal field theories in previous work, see [18] and [19], and the sinh-Gordon model in this thesis, future work should include the study of an integrable model previously unsolved by another method. All we would require is the S-matrix of the model, and the rest of the procedure follows directly. It may also be interesting to study previously solved integrable QFT models with this method, as topologically studying such models may yield insights neglected by the original techniques used to solve them. It would also be very interesting to compare the viability of our method with other methods of obtaining solutions to integrable systems. Given the generality of analytically continuing the TBA equations to obtain the excited states and solve for the full spectrum, it would be interesting to research the Sklyanin's separation of variables method and compare with ours, as it too is highly general.

Appendix A

Matlab Dependency Programs

This appendix contains the necessary dependency programs to run the scan and analytic continuation programs in Appendix B. These dependencies are common to all programs irrespective of the value of the singularity for the n-singularity programs.

A.1 Branchpoints.m

```
% File to store data on branch points
%
% Branch point locations as a function of p (to do: add similar lists for
% BP2, BP3 etc)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
BP1 = [...
        0.5  -0.39310+4.16370i ; ... % s = 50 for cntr = 1.0,0.5 (for v close scans cntr = 1.0,0.2 also OK)
        0.45 -0.39355 + 4.17220i ; ...
        ];


BP1_1a = [...
        0.5  -0.23+2.595i ; ...
        0.45 -0.23+2.61i ; ...
        ];

BP1_1b = BP1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

A.2 MakeContourPlots.m

```
% makes and saves contour plots of c

% some of the variables that need to be in place
%
% plottitle  - title of plot to be shown on picture
% filename   - root of filename for saving plots
% contourspacing - spacing of contours
% rmn, rmx   - min and max real values of r on grid
% imn, imx   - min and max imaginary values of r on grid
% rvals      - real values of r on grid
% ivals      - imaginary values of r on grid
% itervals   - the set of iteration numbers
```

```
% c  - matrix of central charge values on grid

%%%%%%%%%% plot contour lines %%%%%%%%%%

titlefontsize = 12;
axesfontsize = 12;

cntrs = -500:contourspacing:500;
icntrs = linspace(0,max(iter_vals,[],'all'),40);

figure(1); clf; hold on; box on;
contour(rvals,ivals,real(c),cntrs,'LineWidth',0.25);
plot([0,-imx*tan(contour_height)],[0,imx],'-k');
if showbranchpoint
    plot(real(scancentres(1,it_p)),imag(scancentres(1,it_p)),'.k','MarkerSize',11);
end
title(plottitle,'FontSize',titlefontsize);
axis equal
axis([rmn rmx imn imx])
ax1 = gca;
ax1.XMinorTick = 'on';
ax1.YMinorTick = 'on';
ax1.FontSize = axesfontsize;
set(gcf,'PaperPosition', [0 0 outputwidth outputheight]); % plot left hand corner w/ given width and height.
set(gcf,'PaperSize', [outputwidth outputheight]);         % set the paper width and height.
set(gca,'LooseInset',[0.075 lmrg 0.05 umrg]);

figure(2); clf; hold on; box on;
contour(rvals,ivals,imag(c),cntrs,'LineWidth',0.25);
plot([0,-imx*tan(contour_height)],[0,imx],'-k');
if showbranchpoint
    plot(real(scancentres(1,it_p)),imag(scancentres(1,it_p)),'.k','MarkerSize',11);
end
title(plottitle,'FontSize',titlefontsize);
axis equal
axis([rmn rmx imn imx])
ax2 = gca;
ax2.XMinorTick = 'on';
ax2.YMinorTick = 'on';
ax2.FontSize = axesfontsize;
set(gcf,'PaperPosition', [0 0 outputwidth outputheight]); % put plot at left hand corner with given width and height.
set(gcf,'PaperSize', [outputwidth outputheight]);         % set the paper width and height.
set(gca,'LooseInset',[0.075 lmrg 0.05 umrg]);
%
figure(3); clf; hold on; box on;
caxis('manual');caxis([-40 40]); % may need to be changed depending on scan range and level
contour(rvals,ivals,real(c),cntrs,'LineWidth',0.25);
contour(rvals,ivals,imag(c),cntrs,'LineWidth',0.25);
plot([0,-imx*tan(contour_height)],[0,imx],'-k');
if showbranchpoint
    plot(real(scancentres(1,it_p)),imag(scancentres(1,it_p)),'.k','MarkerSize',11); hold on
end
title(plottitle,'FontSize',titlefontsize);
axis equal
axis([rmn rmx imn imx])
ax3 = gca;
ax3.XMinorTick = 'on';
ax3.YMinorTick = 'on';
ax3.FontSize = axesfontsize;
set(gcf,'PaperPosition', [0 0 outputwidth outputheight]); % put plot at left hand corner with given width and height.
set(gcf,'PaperSize', [outputwidth outputheight]);         % set the paper width and height.
set(gca,'LooseInset',[0.075 lmrg 0.05 umrg]);
%
figure(4); clf; hold on; box on;
[c_y,c_x] = gradient(c);
c_z = (c_x/rstepsize - 1i*c_y/istepsize);
pltz = 1./(c_z.^2);
contour(rvals,ivals,real(pltz),dcntrs,'LineWidth',0.25);
contour(rvals,ivals,real(pltz),[0 0],'LineWidth',0.25,'LineColor','r');
contour(rvals,ivals,real(exp(1i*pi/4)*pltz),[0 0],'LineWidth',0.25,'LineColor','k');
contour(rvals,ivals,imag(pltz),dcntrs,'LineWidth',0.25);
contour(rvals,ivals,imag(pltz),[0 0],'LineWidth',0.25,'LineColor','r');
contour(rvals,ivals,imag(exp(1i*pi/4)*pltz),[0 0],'LineWidth',0.25,'LineColor','k');
plot([0,-imx*tan(contour_height)],[0,imx],'-k');
if showbranchpoint
    plot(real(scancentres(1,it_p)),imag(scancentres(1,it_p)),'.k','MarkerSize',11);
end
title(plottitle,'FontSize',titlefontsize);
axis equal
axis([rmn rmx imn imx])
ax4 = gca;
```

```
ax4.XMinorTick = 'on';
ax4.YMinorTick = 'on';
ax4.FontSize = axesfontsize;
set(gcf,'PaperPosition', [0 0 outputwidth outputheight]); % put plot at left hand corner with given width and heig
set(gcf,'PaperSize', [outputwidth outputheight]);         % set the paper width and height.
set(gca,'LooseInset',[0.075 lmrg 0.05 umrg]);
%
figure(5); clf; hold on; box on;
contour(rvals,ivals,iter_vals,icntrs,'LineWidth',0.25);
plot([0,-imx*tan(contour_height)],[0,imx],'-k');
if showbranchpoint
    plot(real(scancentres(1,it_p)),imag(scancentres(1,it_p)),'.k','MarkerSize',11); hold on
end
title(plottitle,'FontSize',titlefontsize);
axis equal
axis([rmn rmx imn imx])
ax5 = gca;
ax5.XMinorTick = 'on';
ax5.YMinorTick = 'on';
ax5.FontSize = axesfontsize;
set(gcf,'PaperPosition', [0 0 outputwidth outputheight]); % put plot at left hand corner with given width and heig
set(gcf,'PaperSize', [outputwidth outputheight]);         % set the paper width and height.
set(gca,'LooseInset',[0.075 lmrg 0.05 umrg]);
%
if savejpg
  % save as jpg:
  print(1,'-painters','-djpeg',strcat(filename,'r','.jpg'));
  print(2,'-painters','-djpeg',strcat(filename,'i','.jpg'));
  print(3,'-painters','-djpeg',strcat(filename,'b','.jpg'));
  print(4,'-painters','-djpeg',strcat(filename,'z','.jpg'));
  print(5,'-painters','-djpeg',strcat(filename,'it','.jpg'));
end

if savepdf
  % save as pdf:
  print(1,'-painters','-dpdf',strcat(filename,'r','.pdf'));
  print(2,'-painters','-dpdf',strcat(filename,'i','.pdf'));
  print(3,'-painters','-dpdf',strcat(filename,'b','.pdf'));
  print(4,'-painters','-dpdf',strcat(filename,'z','.pdf'));
  print(5,'-painters','-dpdf',strcat(filename,'it','.pdf'));
end
```

A.3 SetPaths.m

```
%%%% Code to pre-define a bunch of continuation paths %%%%

%%%% Standard r-path building blocks:
straight20c = 0:0.05:1.0;
halfscale20c = (1-0.25*straight20c);

% Straight line path in complex plane:
x = [0.4:0.01:1.5];
y = (-1.55/1.1).*x + (1.55*1.5/1.1);
complex = x + 1i*y;
Complex = fliplr(complex);  % Complex path

%%%%
```

A.4 multiprecision.m

```
function r = num_t(expression)
    global class_t;
    if (nargin > 0)
        if(strcmpi(class_t,'mp')), r = mp(expression);
        else
            if isnumeric(expression)
                r = expression;
            else
```

```
                r = eval(expression);
            end;
        end;
    else
        r = class_t;
    end;
end
```

Appendix B

1-Particle TBA Programs

This appendix contains the dependencies specific to the 1-singularity scan and analytic continuation programs, as well as the scan and analytic continuation programs for the 1-singularity themselves.

B.1 Solve1TBAh.m

```
% iterates 1 singularity TBA equation

% this is the half-line version

%%%%%%%%%%%%%%%

exp_minuslogphase = exp(-logphase);
exp_minussourcephase = exp(-sourcephase);
exp_minusb_0phase = exp(-b_0phase);
exp_halfb_0phase = exp(b_0phase/2);

for mm = 1:mxit
    %%%%%%%%%% epsilon iteration:
    source = log(exp_minussourcephase*S(bh_contour-b_0)./S(bh_contour+b_0)) + sourcephase;
    epsilonh_old_checkpoints = epsilonh(checkpoints);
    for m = 1:10
       Lh = log(exp_minuslogphase*(one1+exp(-epsilonh)))+logphase;
       epsilonh = (one1-epsdamping).*epsilonh + epsdamping.*(rcoshthetah + source - Phidthhbar*Lh);
    end
    %%%%%%%%%% root iteration:
    last_b_0 = b_0;
    % compute convolution at theta = b_0 and b_1:
    % choose particular log here via the 2 pi i N0 factor:
    cosh_next_b_0 = ((phi(b_0-bh_contour)+phi(b_0+bh_contour)).'*(Lh.*dthhbar_contour) + ...
        log(S(two2*b_0)/exp_minussourcephase) - sourcephase + twoi*N0*pi_p)/r;
    % choose branch cut of sqrt here using exp_minusbphase:
    next_b_0 = log(cosh_next_b_0+sqrt(exp_minusb_0phase*(cosh_next_b_0^2-one1))*exp_halfb_0phase);
    b_0 = (one1-rootdamping)*b_0+rootdamping*next_b_0;
    if mm >= minit && (abs(norm(epsilonh(checkpoints)-epsilonh_old_checkpoints)) + ...
            abs(b_0-last_b_0) <= epsprec)
        break;
    end
end

failed = (mm == mxit); % flag for failed iteration

if ~failed %%%%% iteration converged successfully so update phases for logs and sqrts:
    ztocheck = one1 + exp(-epsilonh);
    angles = angle(exp_minuslogphase*ztocheck)+imag(logphase);
    maxangle = max(angles); minangle = min(angles);
    logphase = onei * (maxangle+minangle)/two2;
    Stocheck = S(bh_contour-b_0)./S(bh_contour+b_0);
    anglesS = angle(exp_minussourcephase*Stocheck)+imag(sourcephase);
```

```
    maxangleS = max(anglesS); minangleS = min(anglesS);
    sourcephase = onei * (maxangleS+minangleS)/two2;
    b_0phase = onei*angle(exp_minusb_0phase*(cosh(b_0)^2-one1)) + b_0phase;
    exp_minuslogphase = exp(-logphase);
    exp_minussourcephase = exp(-sourcephase);
    exp_minusb_0phase = exp(-b_0phase);
    exp_halfb_0phase = exp(b_0phase/2);
    Lh = log(exp_minuslogphase*(1+exp(-epsilonh)))+logphase;
end

%%%%%%%%%%%%%%%
```

B.2 ShG__1singularity__Scan__lr.m

```
% Solves the 1-singularity sinh-Gordon TBA and scans a grid of r values.
%
% NB this one uses half-line storage of pseudoenergies etc
%
% Solves the TBA equation on a contour
% This one scans left-right first, and tries to look sideways when downward
% scan fails, and saves time by using theta -> -theta symmetry
%

clear
cd('C:\Users\Research\Documents\MATLABtest\Plots\1sing\');

advanpix = false;  % flag as to whether to use advanpix (or not)
global class_t;

if advanpix
    dgits = 34;
    mp.Digits(dgits);
    class_t = 'mp';
    format longG;
else
    dgits = 0;
    class_t = 'double';
end

one1 = num_t('1.0');
two2 = num_t('2.0');
onei = num_t('1i');
twoi = num_t('2i');
pi_p = num_t('pi');

%%%%%%%%%% levels in solution of asymptotic 2-particle BAE:
%
Nvalues = [ 1 ];

%%%%%%%%%% input branch point locations as a function of p:
%
BranchPoints

%%%%%%%%%% Point from which to continue before starting to scan
%
FirstPoint = 3.0 + 0.1i;

%%%%%%%%%% scan steps and region:
%
rsteps = 30; % was 100 then 45
isteps = 90; % was 140
scanwidth = -2.0;
scanheight = -6.0;

rstepsize = scanwidth/(rsteps-1);
istepsize = -scanheight/(isteps-1);

fixedscanwindow = true;

if fixedscanwindow
    pvalues = ...
          [0.1];
    stripstepslist = ...
          [50];
    scancentre = -0.5+4.0i;
    scancentres = scancentre.*ones(size(pvalues));
    showbranchpoint = false;  % don't show the current best guess for the BP location
else
```

```
    BP = BP1;
    pvalues = BP(end:1:end,1).';
    scancentres = BP(end:1:end,2).';
    %
    stripsteps = 50; %number of integration grid points per Yshift (was 100)
    stripstepslist = stripsteps.*ones(size(pvalues)); % could also be made variable
    %
    showbranchpoint = true;  % show the current best guess for the BP location
end

cvalues = cell(1,length(pvalues));
b_0values = cell(1,length(pvalues));

%%%%%%%%%% contour shape parameters:
%
contour_height = num_t('0.5');
contour_width = num_t('1.0');

%%%%%%%%%% countour spacings for plots of Re(c) and Im(c):
%
contourspacing = 0.5; % for big fixed scans with 0 to 8 or 8 to 14 imaginary part

%%%%%%%%%% countour spacing for plot of 1/(dc/dr)^2:
%
dcntrs = -0.02:0.0005:0.02; % for wide scans

outputwidth = 10; % width on output page
outputheight = outputwidth * abs(scanheight)/abs(scanwidth);
lmrg = 0.03; % was 0 for 3-15 scans
umrg = 0.07; % was 0 for 3-15 scans

%%%%%%%%%% vertical offset and shift for initial approach path and flags
%%%%%%%%%% for other plots:
%
vertshift = 0.1; % -min(abs(scanheight/2),0.005)*sign(scanheight);
roundthecorner = true; upshift = 1;
%
plotscanregions = true;
plotprogress = false; % show progress on the scanregions plot (NB: slows program significantly if grid is fine)
plotlogplot = false;
plotbranchpoints = false;

%%%%%%%%%% numerics parameters:
%
epsprec = num_t('10^(-10)');
epsdamping = num_t('0.5'); %iteration damping. 1 = no damping, -->0 = maximal damping (was 0.2)
rootdamping = num_t('0.3'); %iteration damping. 1 = no damping, -->0 = maximal damping (was 0.2)
minit = 2;        % minimum number of iterations
maxit = 1000;    % maximum number of iterations
maxit_approach = maxit + 10000;   % maximum number of iterations on approach path
maxit2 = 30;     % max number of iterations after first fail
maxtries = 1;     % maximum number of tries on a given row
maxtries_lr = 3;  % maximum number of tries on an outer loop

%%%%%%%%%% what to save at the end of the run:
%
savepdf = false;
savejpg = true;

%%%%%%%%%% integration details:
%
Yshift =  pi_p/two2;         % basic Y-system shift
th_max = num_t('3');         % upper integration limit in units of Yshift
hYshift = Yshift/two2;       % half Yshift (used for splitting up contour)
theta_max = th_max*Yshift;

%%%%%%%%%% loop through chosen values of N:
%
t_scans_all = tic;

for Ns = Nvalues.'

N0 = Ns(1);

fprintf('\nN0 = %s \n',num2str(N0))

%%%%%%%%%% main loop in sinh-Gordon parameter:
%
t_scans_N = tic;
it_p = 0;

for p = pvalues
clk = clock;
fprintf('Scan starting; time =
%s:%s:%s\n',num2str(clk(4)),num2str(clk(5),'%02d'),num2str(round(clk(6)),'%02d'))
```

```
fprintf('p = %s\n',num2str(p))
t_scan = tic;

it_p = it_p+1;
stripsteps = stripstepslist(it_p);
%
if mod(stripsteps,2) ~= 0
    error('stripsteps should be even')
end
hstripsteps = stripsteps/two2;

%%%%%%%%%% set up deformed contour:
%
contourfn  = @(theta) -contour_height*tanh(theta/contour_width);
dcontourfn = @(theta) -contour_height./contour_width./((cosh(theta/contour_width)).^two2);

stripsteplist = round(sqrt(dcontourfn((0:(th_max*2-1))*hYshift).^2+one1).*hstripsteps);
                                    % halfline stripsteps adjusted for slope of contour
dthlist = hYshift./stripsteplist;   % values of dth on each subinterval
dthhalflist = dthlist/two2;
sh = sum(stripsteplist);            % total number of integration points on half line
checkpoints = round(linspace(1,sh,5));   % choose 5 points for epsilon-convergence check

bh = [];
dthh_contour = [];
for i = 1:2*th_max
    next_b_piece = linspace((i-1)*hYshift+dthhalflist(i),i*hYshift-dthhalflist(i),stripsteplist(i));
    bh = [ bh next_b_piece ];
    dthh_contour = [ dthh_contour dthlist(i)*(one1 + onei*dcontourfn(next_b_piece)) ];
end

bh = bh.';                                  % transpose to get column vector
dthh_contour = dthh_contour.';              % transpose to get column vector
bh_contour = bh + onei*contourfn(bh);       % half-line complex contour as rapidity vector
dthhbar_contour = dthh_contour./(two2*pi_p);
coshbh_contour = cosh(bh_contour);

%%%%%%%%%% grid ranges for scan:
%
rmin = real(scancentres(1,it_p)) - scanwidth/two2;
rmax = real(scancentres(1,it_p)) + scanwidth/two2;
imin = imag(scancentres(1,it_p)) - scanheight/two2;
imax = imag(scancentres(1,it_p)) + scanheight/two2;

imx = max(imin,imax); imn = min(imin,imax);
rmx = max(rmin,rmax); rmn = min(rmin,rmax);

rvals = linspace(rmin,rmax,rsteps);
ivals = linspace(imin,imax,isteps);

%%%%%%%%%% kernel function and S matrix in TBA equation:
%
sphi = num_t('4.0')*sin(pi_p*p);
cphi = cos(two2*pi_p*p);
phi = @(theta) (sphi*cosh(theta))./(cosh(two2*theta)-cphi);
S = @(theta) (sinh(theta)-onei*sin(pi_p*p))./(sinh(theta)+onei*sin(pi_p*p));

%%%%%%%%%% grid and kernel for TBA iteration: (note, the kernel includes the dthbar = dth/(2pi) factor)
%
[Pkh,Plh]=ndgrid(1:sh,1:sh);

Phidthhbar=(phi(bh_contour(Pkh)-bh_contour(Plh))+phi(bh_contour(Pkh)+bh_contour(Plh))).*dthhbar_contour(Plh);

%%%%%%%%%% space for storage of results:
%
c = zeros(isteps,rsteps); %values of c on grid
b_0vals = zeros(isteps,rsteps); % values of b_0 on grid
iter_vals = zeros(isteps,rsteps); % number of iterations on grid

%%%%%%%%%% initial and carry-over values:
%
b_0_start = num_t('1+1.5i');        % random starting value (would be better to use asympt BA value for each r)
logphase_start = num_t('0.0');
sourcephase_start = num_t('0.0');   % phase for branch of log in log(S/S) source term
b_0phase_start = onei*pi_p/two2;  % phase for branch of sqrt in equation to find b_0

lastepsilonh = zeros(rsteps,sh,class_t);
lastlogphase = zeros(rsteps,1,class_t);
lastb_0 = zeros(rsteps,1,class_t);
lastsourcephase = zeros(rsteps,1,class_t);
lastb_0phase = zeros(rsteps,1,class_t);
numfails = zeros(rsteps,1,'int8');

%%%%%%%%%% continue in to start of grid: %%%%%%%%%%
```

```
if rvals(1) > 0.01
    % Continue in from FirstPoint to StartPoint:
    StartPoint = rvals(1) + 1i * ivals(1);
    r_startpath = FirstPoint + [0:0.05:1.0] * (StartPoint - FirstPoint);
else
    % initial continuation path options (could be changed if need be)
    StartPoint = 0.015 + 1i*(ivals(1)+vertshift) % first continue in from FirstPoint to StartPoint
    numpoints_startpath = max(ceil(abs(0.015-rvals(1)+1i*vertshift)/0.02),16);
    if roundthecorner  %%% "round the corner" path for lr then du scans
        vertshift = abs(vertshift);
        numpoints_startpath2 = max(ceil(abs(scanheight)/0.02),24);
        r_startpath = [0.015+1i*(ivals(end)+vertshift+upshift*scanheight) + ...
            linspace(0,1,numpoints_startpath)*(rvals(1)-0.015-1i*vertshift),...
            rvals(1) + 1i*(ivals(end)+upshift*scanheight) - ...
            linspace(0,1,numpoints_startpath2)*(upshift+1)*1i*scanheight];
    else               %%% make path a straight line from StartPoint:
        r_startpath = ...
            [ FirstPoint + [0:0.1:0.9] * (StartPoint - FirstPoint) ,
              StartPoint + linspace(0,1,numpoints_startpath)*(rvals(1) + 1i*ivals(1)- StartPoint)];
        numpoints_startpath2 = 0;
    end
    fprintf("Continuing in from r = %s using %s steps\n",num2str(r_startpath(1)),...
        num2str(numpoints_startpath+numpoints_startpath2));
end

if plotscanregions
    figure(10); clf; hold on; box on;
    plot(real(r_startpath),imag(r_startpath),'b');
    plot([rmn rmx rmx rmn rmn],[imn imn imx imx imn],'r');
    title(['Scan area and approach path' 10],'FontSize',9);
    axis equal;shg;
end

r_start = r_startpath(1);
epsilonh = r_start * coshbh_contour;
b_0 = b_0_start;
logphase = logphase_start;
sourcephase = sourcephase_start;
b_0phase = b_0phase_start;

startstep = 0;
mxit = maxit_approach;

for r = r_startpath
    if plotscanregions
        figure(10); plot(real(r),imag(r),'*b'); hold on; shg;
    end
    startstep = startstep + 1;
    if mod(startstep,100) == 0
        fprintf('\n');
    end
    rcoshthetah = r*coshbh_contour;
    Solve1TBAh
    if startstep == 1
        fprintf('b_start initial value = %s\n',num2str(b_0));
    elseif failed
        fprintf(':')
    else
        fprintf('.')
    end
end

%%%%%%%%%% scan grid in r %%%%%%%%%%%

fprintf('\n Starting scan \n');

epsilonh_start = epsilonh;
b_0_start = b_0;
logphase_start = logphase;
sourcephase_start = sourcephase;
b_0phase_start = b_0phase;
nummaxit_lr = 0;

for istep = 1:isteps

    fprintf('istep = %d ',istep)

    epsilonh = epsilonh_start;
    b_0 = b_0_start;
    logphase = logphase_start;
    sourcephase = sourcephase_start;
    b_0phase = b_0phase_start;

    most_its = [0,0];
    OKsofar = true;
    for rstep = 1:rsteps
```

```
    r = rvals(rstep) + 1i * ivals(istep);
    if angle(r) > pi/2 + contour_height % outside wedge where TBA works
        %%%%%%%%%% set c to zero and skip iteration:
        c(istep,rstep) = NaN; % 0;
        fprintf('/');
        %%%%%%%%%%
    elseif abs(imag(r)) < 0.001 % avoid real axis for excited states!
        %%%%%%%%%% set c to value 1 above and skip iteration:
        c(istep,rstep) = c(istep-1,rstep);
        fprintf('/');
        %%%%%%%%%%
    else
        if OKsofar || numfails(rstep) < maxtries
            mxit = maxit;
        else
            mxit = maxit2;
        end
        rcoshthetah = r*coshbh_contour;
        Solve1TBAh

        %%%%%%%%%% resulting value of c:
        %
        c(istep,rstep) = 12i*r*sinh(b_0)/pi + 6*rcoshthetah.'*(Lh.*dthh_contour)./((pi)^2);
        b_0vals(istep,rstep) = b_0;
        iter_vals(istep,rstep) = mm;

        if failed %%%% didn't converge so reset epsilon and logphase to last scan values
            fprintf(':');
            numfails(rstep) = numfails(rstep) + 1;
            OKsofar = false;
            if rstep ~= rsteps
                epsilonh = lastepsilonh(rstep+1,:).';
                b_0 = lastb_0(rstep+1);
                logphase = lastlogphase(rstep+1);
                sourcephase = lastsourcephase(rstep+1);
                b_0phase = lastb_0phase(rstep+1);
            end
            c(istep,rstep) = NaN;
            b_0vals(istep,rstep) = NaN;
            iter_vals(istep,rstep) = maxit;
            if plotprogress
                figure(10); plot(real(r),imag(r),'.r'); hold on; shg;
            end
            if most_its(2) < mm
                most_its=[most_its(2),mm];
            end
            % This bit is to speed up for the rest of the current loop after maxtries tries
            % (useful if too slow):
            if rstep == 1
                nummaxit_lr = nummaxit_lr + 1;
            end
        else %%%% did converge so save values in case of failure next time:
            fprintf('.');
            OKsofar = true;
            numfails(rstep) = 0;
            sorted_its = sort([mm,most_its]);
            most_its = sorted_its(2:3);
            %
            if plotprogress
                figure(10); plot(real(r),imag(r),'.b'); hold on; shg;
            end
            %
            lastepsilonh(rstep,:) = epsilonh.';
            lastb_0(rstep) = b_0;
            lastlogphase(rstep) = logphase;
            lastsourcephase(rstep) = sourcephase;
            lastb_0phase(rstep) = b_0phase;
        end

    end
    if rstep == 1
        epsilonh_start = epsilonh;
        b_0_start = b_0;
        logphase_start = logphase;
        sourcephase_start = sourcephase;
        b_0phase_start = b_0phase;
        if nummaxit_lr == maxtries_lr
            break
        end
    end
end
fprintf('\n most iterations = %d, %d; b_start value = %s\n',...
    most_its(1),most_its(2),num2str(b_0_start));
if nummaxit_lr == maxtries_lr
   break
```

```
    end
end

%%%%%%%%%% tidy up c a bit %%%%%%%%%%%%%%%

for istep = 2:isteps-1
    for rstep = 2:rsteps-1
        if (isnan(c(istep,rstep-1)) && isnan(c(istep,rstep+1))) || ...
                (isnan(c(istep-1,rstep)) && isnan(c(istep+1,rstep)))
            c(istep,rstep) = NaN;
        end
    end
end

%%%%%%%%%% plot contour lines %%%%%%%%%%

plottitle = ['N0 = ' num2str(N0) '';  p = ' num2str(p) 10];

filename = strcat('1scan',num2str(N0),'_p_',num2str(double(p),'%.4f'),'_lr_',num2str(rmin),
'_',num2str(rmax),'_',...
    num2str(imin),'_',num2str(imax),'_cntr_',num2str(contour_width,'%.1f'),'_',
    num2str(contour_height,'%.2f'),...
    '_s_',num2str(stripsteps,'%04i'),'_a_',num2str(th_max),'_dgts_',num2str(dgits,'%3i'),...
    '_edamp_',num2str(epsdamping),'_rdamp_',num2str(rootdamping),'_maxit_',num2str(maxit));

MakeContourPlots

%%%%%%%%%% save data for later %%%%%%%%%%%

save(strcat(filename,'_data.mat'),'c','b_0vals','iter_vals','rvals','ivals');

time=round(toc(t_scan));
t_hms = datevec(time./(60*60*24));
hrs = strcat(num2str(t_hms(4))," hours ");
if t_hms(4) == 0
    hrs = "";
elseif t_hms(4) == 1
    hrs = " 1 hour ";
end
fprintf(strcat("Scan finished; time taken = ",hrs,'%s minutes %s
seconds\n\n'),num2str(t_hms(5)),num2str(t_hms(6)))

end % (of scan in p)

time=round(toc(t_scans_N));
t_hms = datevec(time./(60*60*24));

if t_hms(4) == 0
    hrs = "";
elseif t_hms(4) == 1
    hrs = " 1 hour ";
else
    hrs = strcat(num2str(t_hms(4)),," hours ");
end

if t_hms(3) == 0
    dys = "";
elseif t_hms(3) == 1
    dys = " 1 day ";
else
    dys = strcat(num2str(t_hms(3)),," days ");
end

clk = clock;
fprintf('\nScans finished; time =
%s:%s:%s\n',num2str(clk(4)),num2str(clk(5),'%02d'),num2str(round(clk(6)),'%02d'))
fprintf(strcat("\nTime for scans = ",dys,hrs,'%s minutes %s
seconds\n'),num2str(t_hms(5)),num2str(t_hms(6)))

end % (of scan in Ns)

time=round(toc(t_scans_all));
t_hms = datevec(time./(60*60*24));

if t_hms(4) == 0
    hrs = "";
elseif t_hms(4) == 1
    hrs = " 1 hour ";
else
    hrs = strcat(num2str(t_hms(4)),," hours ");
end

if t_hms(3) == 0
    dys = "";
```

```
elseif t_hms(3) == 1
    dys = " 1 day ";
else
    dys = strcat(num2str(t_hms(3))," days ");
end

clk = clock;
fprintf('\nAll scans finished; time =
%s:%s:%s\n',num2str(clk(4)),num2str(clk(5),'%02d'),num2str(round(clk(6)),'%02d'))
fprintf(strcat("\nTime for all scans = ",dys,hrs,'%s minutes %s
seconds\n'),num2str(t_hms(5)),num2str(t_hms(6)))

% end of program

function r = num_t(expression)
    global class_t;
    if (nargin > 0)
        if(strcmpi(class_t,'mp')), r = mp(expression);
        else
            if isnumeric(expression)
                r = expression;
            else
                r = eval(expression);
            end
        end
    else
        r = class_t;
    end
end
```

B.3 ShG__1singularity__AC.m

```
clear
cd('C:\Users\Research\Documents\MATLAB\Plots\1sing\');

advanpix = false;  % flag as to whether to use advanpix (or not)
global class_t;

if advanpix
    dgits = 34;
    mp.Digits(dgits);
    class_t = 'mp';
    format longG;
else
    dgits = 0;
    class_t = 'double';
end

one1 = num_t('1.0');
two2 = num_t('2.0');
onei = num_t('1i');
twoi = num_t('2i');
pi_p = num_t('pi');

%%%%%%%%%% flags to decide what to plot:
%
whattoplot = 'Y';   % can be Y, X, T or Tt (for T-tilde)
drawYlines = true;  % include asymptotic Y lines in Y plot
drawXlines = true;  % include asymptotic X lines in X plot
plotrpath = true;   % plot path of r
ZoomedView = false; % show zoomed-in view of contour plot
savejpg = false;    % save main plot as jpg

%%%%%%%%%% sinh-Gordon parameter:
%
p = num_t('0.1');

%%%%%%%%%% level in solution of asymptotic 2-particle BAE:
%
N0 = 1;

%%%%%%%%%% contour shape parameters:
%
contour_height = num_t('0.5');
contour_width = num_t('1.0');

%%%%%%%%%% numerics parameters for this program:
```

```
%
Yshift =  num_t('pi')/two2;       % basic Y-system shift
stripsteps = num_t('50');        % number of integration steps per horizontal Yshift (must be even)
th_max = num_t('3');              % upper integration limit in units of Yshift
epsdamping = num_t('0.5');        % TBA iteration damping; 1 = no damping, ->0 = maximal damping
rootdamping = num_t('0.3');       % root iteration damping; 1 = no damping, ->0 = maximal damping
epsprec = num_t('10^(-10)');      % precision at which to terminate iteration                    %%% MODIFIED
maxit = 1000;                % maximum number of iterations
minit = 2;                    % minimum number of iterations

%%%%%%%%%% theta plane plot resolution:
%
plotstripsteps_r = num_t('100');  % number of plot steps per Yshift in real direction
plotstripsteps_i = num_t('100');  % number of plot steps per Yshift in imaginary direction

%%%%%%%%%% theta plane plot range:
%
if ZoomedView
    plot_min_r = -2.0;
    plot_max_r =  2.0;
    plot_min_i = -1.3;
    plot_max_i =  3.3;
else
    plot_min_r = -4.0;
    plot_max_r =  4.0;
    plot_min_i = -4.6;
    plot_max_i =  4.6;
end

%%%%%%%%%% scan program parameters relevant for inset plot:
%
contourspacing = 0.5;
scanwidth = -2.0;
scanheight = -6.0;
scancentre = -0.5+4.0i;
%%% TBA iteration params used in scan program: (inferred from the value of p)
dgits_scan = 0;
if p <= 0.1
    p_index = 11-round(p*100);
    p_scan_values = [50  50   50   50   50   50   60   80   120  240];
    stripsteps_scan = p_scan_values(p_index);
else
    stripsteps_scan = 50;
end
contour_height_scan = contour_height;
contour_width_scan = contour_width;
p_scan = double(p);
epsdamping_scan = double(epsdamping);
rootdamping_scan = double(rootdamping);
maxit_scan = 1000;
th_max_scan = double(th_max);

%%%%%%%%%% import scan plot data:
%
rmin = real(scancentre) - scanwidth/2;
rmax = real(scancentre) + scanwidth/2;
imin = imag(scancentre) - scanheight/2;
imax = imag(scancentre) + scanheight/2;

imx = max(imin,imax); imn = min(imin,imax);
rmx = max(rmin,rmax); rmn = min(rmin,rmax);

filename = strcat('1scan',num2str(N0),'_p_',num2str(p_scan,'%.4f'),'_lr_',...
    num2str(rmin),'_',num2str(rmax),'_',num2str(imin),'_',num2str(imax),...
    '_cntr_',num2str(contour_width_scan,'%.1f'),'_',num2str(contour_height_scan,'%.2f'),'_s_',...
    num2str(stripsteps_scan,'%04i'),'_a_',num2str(th_max_scan),'_dgts_',num2str(dgits_scan,'%3i'),...
    '_edamp_',num2str(epsdamping_scan),'_rdamp_',num2str(rootdamping),'_maxit_',num2str(maxit_scan));
load(strcat('datafiles/',filename,'_data.mat'),'c','rvals','ivals');

%%%%%%%%%% size/shape and filename for output plots:
%
outputwidth = 26;
outputheight = 20;

if ZoomedView
    flg='_zoomed_';
else
    flg='_';
end

filename_saveplots = strcat('0continuation',flg,'p_',num2str(double(p),'%.4f'),...
    '_cntr_',num2str(contour_width,'%.1f'),'_',num2str(contour_height,'%.2f'),'_s_',...
    num2str(stripsteps,'%04i'),'_dgts_',num2str(dgits,'%3i'),'_step_');
```

```
%%%%%%%%%% input standard set of paths:
%
SetPaths

%%%%%%%%%% other (fixed) parameters:
%
hYshift = Yshift/two2;            % half Yshift (used for splitting up contour)
theta_max = th_max*Yshift;        % upper integration limit
%
if mod(stripsteps,2) ~= 0
    error('stripsteps should be even')
end
hstripsteps = stripsteps/two2;

%%%%%%%%%% shifts needed to extend X and then compute Y, T and T-tilde:
%
a = one1 - two2*p;
Xshift1 = round(2*p*plotstripsteps_i);
Xshift2 = round((2-2*p)*plotstripsteps_i);
Xshift3 = 2*plotstripsteps_i;
XYshift = round(a*plotstripsteps_i);
XtoTshift = round(2*p*plotstripsteps_i);
XtoTtshift = round(2*(1-p)*plotstripsteps_i);

%%%%%%%%%% theta plane plot parameters:
%
dthplot_r = Yshift/plotstripsteps_r;
dthplot_i = Yshift/plotstripsteps_i;

%%%%%%%%%% calculate X and then Y on symmetrical grid big enough to include plot range:
%
plotgrid_max_r = max(-plot_min_r,plot_max_r);
plotgrid_max_i = contour_height + max([-plot_min_i,plot_max_i,Yshift]);
% (third element here is to ensure basic strip lies in grid)

gridsteps_r = ceil(plotgrid_max_r/dthplot_r+num_t('0.5'));

switch whattoplot % add 1 in each case for luck
    case 'Y'
        gridsteps_i = XYshift + ceil(plotgrid_max_i/dthplot_i+num_t('0.5'))+one1;
    case 'X'
        gridsteps_i = ceil(plotgrid_max_i/dthplot_i+num_t('0.5'))+one1;
    case 'T'
        gridsteps_i = XtoTshift + ceil(plotgrid_max_i/dthplot_i+num_t('0.5'))+one1;
    case 'Tt'
        gridsteps_i = XtoTtshift + ceil(plotgrid_max_i/dthplot_i+num_t('0.5'))+one1;
    otherwise
        error('whattoplot variable incorrectly set')
end

grid_max_r = (gridsteps_r-num_t('0.5'))*dthplot_r;
grid_max_i = (gridsteps_i-num_t('0.5'))*dthplot_i;

bre = linspace(-grid_max_r,grid_max_r,2*gridsteps_r);
bim = linspace(-grid_max_i,grid_max_i,2*gridsteps_i);

%%%%%%%%%% array to store X on complex theta plane:
%
X = zeros(2*gridsteps_r,2*gridsteps_i,class_t);

%%%%%%%%%% Checks if 2*p*plotstripsteps_i is a required integer value (ends if not)
%
fprintf('p = %s; stripsteps = %s; plotstripsteps_i = %s; 2*p*plotstripsteps_i = %s\n\n',...
    num2str(double(p)),num2str(stripsteps),num2str(plotstripsteps_i),num2str(double(2*p*plotstripsteps_i)))
if abs(rem(2*p*plotstripsteps_i,1)) > 1e-10
    error('2*p*plotstripsteps_i must be an integer')
end

%%%%%%%%%% set up deformed contours:
%
contourfn  = @(theta) -contour_height*tanh(theta/contour_width);
dcontourfn = @(theta) -contour_height./contour_width./((cosh(theta/contour_width)).^two2);

stripsteplist = round(sqrt(dcontourfn((0:(th_max*2-1))*hYshift).^2+one1).*hstripsteps);
                                        % halfline stripsteps adjusted for slope of contour
dthlist = hYshift./stripsteplist;       % values of dth on each subinterval
dthhalflist = dthlist/two2;
sh = sum(stripsteplist);                % total number of integration points on half line
s = 2*sh;                               % number of points on full line
checkpoints = round(linspace(1,sh,5));  % choose 5 points for epsilon-convergence check
contourshift = num_t('1');              % up shift of contour for calculation of middle strip

bh = [];
dthh_contour = [];
for i = 1:2*th_max
```

```
    next_b_piece = linspace((i-1)*hYshift+dthhalflist(i),i*hYshift-dthhalflist(i),stripsteplist(i));
    bh = [ bh next_b_piece ];
    dthh_contour = [ dthh_contour dthlist(i)*(one1 + onei*dcontourfn(next_b_piece)) ];
end

bh = bh.';                                  % transpose to get column vector
dthh_contour = dthh_contour.';              % transpose to get column vector
bh_contour = bh + onei*contourfn(bh);       % half-line complex contour as rapidity vector

b = [ -flip(bh) ; bh ];
b_contour = [ -flip(bh_contour) ; bh_contour ];                % full line complex contour
b_contour_shifted = b_contour + onei*contourshift*dthplot_i;   % full line upshifted complex contour
dth_contour = [ flip(dthh_contour) ; dthh_contour ];

dthhbar_contour = dthh_contour./(two2*pi_p);
dthbar_contour = dth_contour./(two2*pi_p);

coshbh_contour = cosh(bh_contour);
coshb_contour_shifted = cosh(b_contour_shifted);

contourgridshifts = round(contourfn(bre)/dthplot_i);
b_contour_grid = bre + onei*dthplot_i*contourgridshifts;
bY = b_contour_grid.'; %transpose to get column vector

%%%%%%%%%% phase to determine branches of logs and sqrts: (adjusted dynamically in program)
%
logphase = num_t('0.0');
sourcephase = num_t('0.0');
b_0phase = num_t('0.0');

%%%%%%%%%% kernel function in TBA equation:
%
sphi = num_t('4.0')*sin(pi_p*p);
cphi = cos(two2*pi_p*p);
phi = @(theta) sphi*cosh(theta)./(cosh(two2*theta)-cphi);

%%%%%%%%%% kernel function for X:
%
phiX = @(theta) one1./cosh(theta);

%%%%%%%%%% S matrix:
%
S = @(theta) (sinh(theta)-onei*sin(pi_p*p))./(sinh(theta)+onei*sin(pi_p*p));

%%%%%%%%%% "S"-matrix for continuation of X(theta):
%
SX = @(theta) (one1+onei*exp(theta))./(one1-onei*exp(theta))/onei;

%%%%%%%%%% grid and kernel for TBA equation: (note, the kernel includes the dthbar=dth/(2pi) factor)
%
[Pkh,Plh]=ndgrid(1:sh,1:sh);

Phidthhbar=(phi(bh_contour(Pkh)-bh_contour(Plh))...
+phi(bh_contour(Pkh)+bh_contour(Plh))).*dthhbar_contour(Plh);

%%%%%%%%%% grid and kernel for contour shift:
%
[Pk,Pl] = ndgrid(1:s,1:sh);

Phidthhbarshift=(phi(b_contour_shifted(Pk)-bh_contour(Pl))...
+phi(b_contour_shifted(Pk)+bh_contour(Pl))).*dthhbar_contour(Pl);

%%%%%%%%%% grid and kernel for extending to basic strip in theta plane:
%
[PPre,PPgg,PPii] = ndgrid(1:2*gridsteps_r,1:s,1:plotstripsteps_i);

PhiXdthbar = phiX(b_contour_grid(PPre)+(PPii-num_t('0.5')).*dthplot_i.*onei ...
- b_contour_shifted(PPgg)).*dthbar_contour(PPgg);

%%%%%%%%%% space for storage of results:
%
maxstep = 1000;
vol = NaN(maxstep,1);           % values of r along the path
c_path = zeros(maxstep,1);      % values of casimir energy along the path
F = zeros(maxstep,1);           % values of scaling energy along the path
b_0_path = zeros(maxstep,1);    % values of active root along the path
logarg_path = zeros(maxstep,1); % values of argument to log along the path

%%%%%%%%%% starting values:
%
epsilonh = coshbh_contour;
b_0 = one1+onei;
cosh_next_b_0 = 0;

%%%%%%%%%% set up plot windows:
```

```
%
figure(1)
clf;
set(gcf,'Position',[5 5 1000 700]);

contouraxis = subplot('Position',[0.04 0.04 0.6 0.92]);

if plotrpath
    rpathaxis = subplot('Position',[0.7 0.04 0.28 0.92]);
    hold on; box on;
    caxis('manual');caxis([-40 40]);
    cntrs_inset = -500:contourspacing:500;
    contour(rvals,ivals,real(c),cntrs_inset,'LineWidth',0.25);
    contour(rvals,ivals,imag(c),cntrs_inset,'LineWidth',0.25);
    title('Continuation path');
    axis equal;
    axis([rmn rmx imn imx])
end

%%%%%%%%% main loop, analytically continuing along path in r:
%
step_=0;
first=true;
while step_<maxstep
    prompt = 'Enter path in r as a row vector! ';
    path = input(prompt);
    if path == 0
        epsilonh = epsilonh_last;
        path = r_last;
        vol = vol_last;
        F = F_last;
    end
    if isempty(path)
        break
    end
    epsilonh_last = epsilonh;
    r_last = path(1);
    vol_last = vol;
    F_last = F;
    rl = r_last;

    for step = 1:length(path)
        r = path(step);
        vol(step_+step) = r;
        rcoshthetah = r*coshbh_contour;

        mxit = maxit;
        Solve1TBAh

        % save stuff for diagnostic plots:
        b_0_path(step_+step) = b_0;
        logarg_path(step_+step) = cosh_next_b_0+sqrt(exp_minusb_0phase*(cosh_next_b_0^2-1))*exp_halfb_0phase;

        %%%%%%%%% resulting scaling energy:
        %
        c_path(step_+step) = 12i*r*sinh(b_0)/pi + ...
            6*rcoshthetah.'*(Lh.*dthh_contour)./((pi)^2);
        F(step_+step) = -c_path(step_+step)./12 - (r^2)./(8*sqrt(3)*pi);

        %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% compute X in whole theta plane %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

        %%%%%%%%% compute L on up-shifted contour:
        %
        expsourceup = S(b_contour_shifted-b_0)./S(b_contour_shifted+b_0);
        Lup = log(exp_minuslogphase * ...
                (one1 + exp( -r.*coshb_contour_shifted + Phidthhbarshift*Lh)./expsourceup)) + logphase;

        %%%%%%%%% compute X in basic strip of theta plane:
        %
        for ii = gridsteps_i + 1 : gridsteps_i + plotstripsteps_i
            X(:,ii) = SX(bY+1i*bim(ii)+b_0)./SX(bY+1i*bim(ii)-b_0) .* ...
                      exp(-r.*cosh(bY+onei*bim(ii))./(two2*sin(pi_p*p)) + ...
                      PhiXdthbar(:,:,ii-gridsteps_i)*Lup);
            X(:,2*gridsteps_i+1-ii) = flip(X(:,ii));
        end

        %%%%%%%%% extend to rest of theta plane using the X-system equation:
        %
        for ii = gridsteps_i + plotstripsteps_i + 1 : 2*gridsteps_i
            X(:,ii)=(one1+X(:,ii-Xshift1).*(X(:,ii-Xshift2 )))./X(:,ii-Xshift3);
            X(:,2*gridsteps_i+1-ii) = flip(X(:,ii));
        end

        %%%%%%%%% straighten out X and compute Y:
        %
```

```
for kk = 1 : 2*gridsteps_r
   X(kk,:) = circshift(X(kk,:),contourgridshifts(kk),2);
end

Y = circshift(X,XYshift,2).*circshift(X,-XYshift,2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% compute the function to plot %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

switch whattoplot
    case 'Y'
        Z=abs(one1+Y);
        rel=Z./(Z+one1);
        cntrs = 0:0.025:1;
    case 'X'
        Z=tanh(log(abs(X))/100);
        rel=Z;
        cntrs = -1:0.01:1;
    case 'T'
        T = (circshift(X,XtoTshift,2)+circshift(X,-XtoTshift,2))./X;
        Z=abs(T);
        rel=Z./(Z+one1);
        cntrs = 0:0.025:1;
    case 'Tt'
        Tt = (circshift(X,XtoTtshift,2)+circshift(X,-XtoTtshift,2))./X;
        Z=abs(Tt);
        rel=Z./(Z+one1);
        cntrs = 0:0.025:1;
end

relz=rel';

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% do plots %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(1)

%%%%%%%%%%%%% plot contour lines:
%
contour(contouraxis,bre,bim,relz,cntrs); hold(contouraxis,'on')
title(contouraxis,[whattoplot ':  p = ' num2str(double(p)) ',  r = ' num2str(double(r))]);
axis(contouraxis,'equal');
axis(contouraxis,[plot_min_r plot_max_r plot_min_i plot_max_i]);

%%%%%%%%%%%%% add integration contour and its shifts:
%
plot(contouraxis,real(b_contour),imag(b_contour),'-r');
plot(contouraxis,real(b_contour),imag(b_contour)+pi*p,'-g');
plot(contouraxis,real(b_contour),imag(b_contour)-pi*p,'-g');
plot(contouraxis,real(b_contour_grid),imag(b_contour_grid),'-k');
plot(contouraxis,real(b_contour_grid),imag(b_contour_grid)+Yshift,'-k');
plot(contouraxis,real(b_contour_grid),imag(b_contour_grid)-Yshift,'-k');

%%%%%%%%%%%%% plot singularity locations linked to b_0 and -b_0:
%
plot(contouraxis,...
    [real(b_0) real(b_0) real(b_0)],[imag(b_0) imag(b_0)-Yshift...
    imag(b_0)-2*Yshift],'-k.','MarkerSize',20);
plot(contouraxis,...
    [real(-b_0) real(-b_0) real(-b_0)],[-imag(b_0) -imag(b_0)+Yshift...
    -imag(b_0)+2*Yshift],'-r.','MarkerSize',20);


switch whattoplot
    case 'Y'
        if drawYlines
            for mm = pi*(1-p):pi*(1-p):grid_max_i-pi*p+pi/2
                for nn = pi*p:pi*p:grid_max_i-mm+pi/2
                    plot(contouraxis,[-grid_max_r grid_max_r],[1 1]*(mm+nn-pi/2),'-.k');
                end
            end
        end
    case 'X'
        if drawXlines
            for mm = pi*(1-p):pi*(1-p):grid_max_i-pi*p
                for nn = pi*p:pi*p:grid_max_i-mm
                    plot(contouraxis,[-grid_max_r grid_max_r],[1 1]*(mm+nn),'-.k');
                end
            end
        end
    case 'T'
        plot(contouraxis,[-grid_max_r grid_max_r],[1 1]*(1-p)*pi,'-.k');
        plot(contouraxis,[-grid_max_r grid_max_r],[0 0],'-.k');
    case 'Tt'
        plot(contouraxis,[-grid_max_r grid_max_r],[1 1]*p*pi,'-.k');
```

```
            plot(contouraxis,[-grid_max_r grid_max_r],[0 0],'-.k');
        end

        hold(contouraxis,'off')

        %%%%%%%%%% update path plot:
        if plotrpath
            rpmn = min(rmn,real(rl));rpmx = max(rmx,real(rl));
            ipmn = min(imn,imag(rl));ipmx = max(imx,imag(rl));
            r_l = vol(max(step_+step-1,1));
            plot(rpathaxis,[real(r_l),real(r)],[imag(r_l),imag(r)],'-gs',...
                'LineWidth',2,...
                'MarkerSize',4); hold(rpathaxis,'on')
            plot(rpathaxis,[real(r)],[imag(r)],'-ks',...
                'LineWidth',2,...
                'MarkerSize',4);
            axis(rpathaxis,[rpmn rpmx ipmn ipmx])
            rl = r;
        end

        set(gcf,'PaperPosition', [0 0 outputwidth outputheight]); % plot left corner w/ given width and height.
        set(gcf,'PaperSize', [outputwidth outputheight]);  % set the paper width and height.

        if savejpg
            % save as jpg:
            print(1,'-painters','-djpeg',strcat(filename_saveplots,num2str(step_+step-1,'%03d'),'.jpg'));
        end


        %Print on screen
        fprintf('Iterations = %i; r = %2.3f+%2.3fi; c = %s\n',...
            mm,real(r),imag(r),num2str(c(step_+step),'%2.5f'));
        %fprintf('cosh_next_b_0 = %s; cosh_next_b_0^2-1 = %s; log() = %s; b_0phase/pi = %s\n',...
        %    num2str(cosh_next_b_0),num2str(cosh_next_b_0^2-1),...
        %    num2str(log(cosh_next_b_0+sqrt(cosh_next_b_0^2-1))),num2str(b_0phase/pi));
        %log(cosh_next_b_0+sqrt(cosh_next_b_0^2-1));
        pause(0.05);
    end
    step_=step_+step;

    fprintf('-------------------------------------------------\n');

end

function r = num_t(expression)
    global class_t;
    if (nargin > 0)
        if(strcmpi(class_t,'mp')), r = mp(expression);
        else
            if isnumeric(expression)
                r = expression;
            else
                r = eval(expression);
            end
        end
    else
        r = class_t;
    end
end
```

Appendix C

E(r) vs r Plot Program

This program produces the $E(r)$ vs. r plot found in Chapter 8.

```
clear all;
close all;
clc;

%
cd('~/Documents/MATLAB/Plots/1sing/');
%
%Flags to decide what to plot:
plotFpath = false;  %plot F on path (if false, plot b_0 instead)
plotz = true;
% plotb_0 = true;     %plot complex b_0 trajectory
% b0range = 0.2;      %range of b_0 plot
plotFr    = false;  %plot F as function of Re(r)
plotslice = true;  %plot slice of Y function
showgridlines = false; %show grid lines on contour plot
% showasymp_b_0 = false; %show asymptotic value of b_0 on contour plot
dualscreen = false; %put graphs onto second screen
ZoomedView = false; %show zoomed-in view of contour plot
saveplots = false; %save plots to files

Realpath = 1:1:11;

%%%%% Branchpoints for sinh-Gordon %%%%%
branchpoints = [-0.4, 4.4 ];

%branch points on zero-particle sheet:
branchpoints0 = [-0.4, 4.4 ];

% phase to determine branch of log: (adjusts dynamically in program)
logphase = 0.0;
exp_minuslogphase = 1.0;
% phase to determine branch of sqrt for b_0_next: (adjusts dynamically in program)
bphase = 0.0;
exp_minusbphase = 1.0;
exp_halfbphase = 1.0;

%%%%%%%%%% level in solution of asymptotic 2-particle BAE:
%N = 1;
for N = 1:1:10;

%%%%%%%%%% integration parameters:
Yshift =  pi/2; %basic Y-system shift
a2u = 3;        %upper integration limit in units of Yshift
a1u = -a2u;     %lower integration limit in units of Yshift
a2 = a2u*Yshift;
a1 = a1u*Yshift;
%
stripsteps = 40; %number of integration steps per Yshift
%
s = (a2u-a1u)*stripsteps+1; %total number of integration points
dth = Yshift/stripsteps; %width of integration steps
dthbar = dth/2/pi; %sometimes handy
b = linspace(a1,a2,s); %real axis rapidity vector (row)
b = b.'; %transpose to get column vector

%%%%%%%%%% contour shape parameters:
```

```
contour_height = 1.5;
contour_width = 1.0;
%
contourfn = @(theta) -contour_height*tanh(theta/contour_width);
dcontourfn = @(theta) -contour_height./contour_width./((cosh(theta/contour_width)).^2);
b_contour = b + 1i*contourfn(b); %complex contour as rapidity vector
coshtheta = cosh(b_contour);
dth_contour = dth*(1 + 1i*dcontourfn(b));
dthbar_contour = dth_contour./2./pi;

%%%%%%%%%% numerics parameters:
epsprec = 10^(-14);
rootprec = 10^(-14);
epsdamping = 0.2; %iteration damping. 1 = no damping, -->0 = maximal damping
rootdamping = 0.2;
checkpoints = round(linspace(1,s,10)); %10 points for epsilon-convergence check
minit = 2;      %minimum number of iterations
maxit = 10000;   %maximum number of iterations
maxstep = 1000; %maximum number of steps of the total analytic continuation path

%%%%%%%%%% sinh-Gordon parameter:
p = 0.2;
a = 1-2*p;

fprintf('p = %s; stripsteps = %s; 2*p*stripsteps = %s\n\n',num2str(p),num2str(stripsteps),num2str(2*p*stripsteps))

%%%%%%%%%% Checks if 2*p*stripsteps a required integer value (ends if not)
if abs(rem(2*p*stripsteps,1)) > 1e-10
    error('2*p*stripsteps must be an integer value.')
end

Xshift1 = round(2*p*stripsteps);
Xshift2 = round((2-2*p)*stripsteps);
Xshift3 = 2*stripsteps;
XYshift = round(a*stripsteps);

%%%%%%%%%% kernel function:
phi = @(theta) (4*sin(pi*p)*cosh(theta))./(cosh(2*theta)-cos(2*pi*p));
%%%%%%%%%% S-matrix:
S = @(theta) (sinh(theta)-1i*sin(pi*p))./(sinh(theta)+1i*sin(pi*p));
%%%%%%%%%% "S"-matrix for continuation of X(theta):
SX = @(theta) (1+1i*exp(theta))./(1-1i*exp(theta))/1i;

%%%%%%%%%% kernel phi_11(beta-beta'): (NB: includes the dthbar = dth/(2pi) factor)
[Pk,Pl]=ndgrid(1:s,1:s);    % Original grid for the Y-System equation
[Pxk,Pxl]=ndgrid(1:s,1:s);  % Grid for the X-System equation

Phidthbar=phi((b_contour(Pk)-b_contour(Pl))).*dthbar_contour(Pl); % Y-System Integral
coshdthbar = (dthbar_contour(Pxl))./(cosh(b_contour(Pxk)-b_contour(Pxl))); % X-System Integral

%%%%%%%%%% space for storage of results:
vol = NaN(maxstep,1); %values of r along the path
c = zeros(maxstep,1); %values of casimir energy along the path
F = zeros(maxstep,1); %values of scaling energy along the path
b_0_path = zeros(maxstep,1); %values of active root along the path
logarg_path = zeros(maxstep,1); %values of argument to log along the path
roottemp = zeros(maxit,1); %temporary values of root position after each iteration

%%%%%%%%%%%% complex theta plane:
ustrips = 5; %no of Ystep strips to compute above real axis
lstrips = 5; %no of Ystep strips to compute below real axis
uplotstrips = 3; %no of Ystep strips to plot above real axis
lplotstrips = 3; %no of Ystep strips to plot below real axis
%
leftstrips = 3; %no of Ystep strips to plot left of imaginary axis
rightstrips = 3; %no of Ystep strips to plot right of imaginary axis
contourshift = 4; %up/down shift of contour for middle strip
usteps = ustrips*stripsteps;
lsteps = lstrips*stripsteps;
leftsteps = leftstrips*stripsteps;
rightsteps = rightstrips*stripsteps;
a2im = ustrips*Yshift-dth;
a1im = -lstrips*Yshift;
a2re = rightstrips*Yshift;
a1re = -leftstrips*Yshift;
%%%% left and right limits for (unzoomed) plots:
a2re_plot = 4.1;
a1re_plot = -4.1;
%%%%
sim = lsteps+usteps;
sre = leftsteps+rightsteps+1;
bim = linspace(a1im,a2im,sim);
bre = linspace(a1re,a2re,sre);
contourgridshifts = round(contourfn(bre)/dth);
b_contour_grid = bre + 1i*dth*contourgridshifts;
bY = b_contour_grid.'; %transpose to get column vector
```

```
X = zeros(sre,sim);
Y = zeros(sre,sim);
%%%%%%%%%%%% kernel extended to whole complex plane:
[PPre,PPgg,PPii] = ndgrid(1:sre,1:s,1:2*stripsteps+1);
[PPxre,PPxgg,PPxii] = ndgrid(1:sre,1:s,1:2*stripsteps+1);

PPhidthbar = phi(b_contour_grid(PPre)+(PPii-stripsteps-1).*dth.*1i ...
- b_contour(PPgg)).*dthbar_contour(PPgg);
ccoshdthbar = (dthbar_contour(PPxgg))./(cosh(b_contour_grid(PPxre)+(PPxii-stripsteps-1).*dth.*1i ...
- b_contour(PPxgg)));

Phidthbarup   = phi((b_contour(Pk)+1i*contourshift*dth-b_contour(Pl))).*dthbar_contour(Pl);
Phidthbardown = phi((b_contour(Pk)-1i*contourshift*dth-b_contour(Pl))).*dthbar_contour(Pl);

coshdthbarup   = (dthbar_contour(Pxl))./(cosh(b_contour(Pxk)+1i*contourshift*dth-b_contour(Pxl)));
coshdthbardown = (dthbar_contour(Pxl))./(cosh(b_contour(Pxk)-1i*contourshift*dth-b_contour(Pxl)));

%%%%%%%%%% starting values:
b_0 = 1+1i;
epsilon = coshtheta;
cosh_next_b_0 = 0;

%%%%%%%%%% analytically continue along path in r %%%%%%%%%%%%
step_=0;
first=true;
%while step_<maxstep
    %prompt = 'Enter path in r as a row vector! ';
    path = Realpath; %input(prompt);
    if path == 0
        epsilon = epsilon_last;
        path = r_last;
        vol = vol_last;
        F = F_last;
    end

    epsilon_last = epsilon;
    r_last = path(1);
    vol_last = vol;
    F_last = F;
    figure(1)
    if dualscreen
        set(gcf,'Position',[-1920 -400 1200 1105]);
    else
        set(gcf,'Position',[5 5 1000 700]);
    end
    for step = 1:length(path)
        r = path(step);
        vol(step_+step) = r;
        rcoshtheta = r*coshtheta;

        %%%%%%%%%% TBA iteration:
        for mm=1:maxit
            %%%%%%%%%% epsilon iteration:
            source = log(S(b_contour-b_0)./S(b_contour+b_0));
            epsilon_old_checkpoints = epsilon(checkpoints);
            for m=1:10
               L = log(exp_minuslogphase*(1+exp(-epsilon)))+logphase;
               epsilon = (1-epsdamping).*epsilon + epsdamping.*(rcoshtheta + source - Phidthbar*L);
            end
            %%%%%%%%%% root iteration:
            last_b_0 = b_0;
            % compute convolution at theta = b_0:
            % choose particular log here via the 2 pi i N factor:
            lcnb0=cosh_next_b_0;
            cosh_next_b_0 = (phi(b_0-b_contour).'*(L.*dthbar_contour) + log(S(2*b_0)) + 2i*N*pi)/r;
            % choose branch cut of sqrt here using exp_minusbphase:
            next_b_0 = log(cosh_next_b_0+sqrt(exp_minusbphase*(cosh_next_b_0^2-1))*exp_halfbphase);
            if abs(next_b_0-b_0)>0.1 && step_ > 5
                fprintf('JUMP')
                stop
            end
            b_0 = (1-rootdamping)*b_0+rootdamping*next_b_0;
            if mm >= minit && abs(norm(epsilon(checkpoints)-epsilon_old_checkpoints))...
            + abs(b_0-last_b_0) <= epsprec
                break;
            end
        end

        % save stuff for diagnostic plots:
        b_0_path(step_+step) = b_0;
        logarg_path(step_+step) = cosh_next_b_0+sqrt(exp_minusbphase*(cosh_next_b_0^2-1))*exp_halfbphase;

        % update branch choices:
        ztoplot = 1 + exp(-epsilon);
        angles = angle(exp_minuslogphase*ztoplot)+imag(logphase);
        maxangle = max(angles); minangle = min(angles);
```

```
        logphase = 1i * (maxangle+minangle)/2;
        exp_minuslogphase = exp(-logphase);
        angle_b = angle(exp_minusbphase*(cosh(b_0)^2-1)) + imag(bphase);
        % angle_b = pi/2;
        bphase = 1i * angle_b;
        exp_minusbphase = exp(-bphase);
        exp_halfbphase = exp(bphase/2);

        %%%%%%%%% resulting scaling energy:
        c(step_+step) = 3*r.*cosh(b_contour).'*((log(exp_minuslogphase*(1+exp(-epsilon)))...
                        +logphase).*dth_contour)./((pi)^2) + 12*r*1i*sinh(b_0)./(pi);
        F(step_+step) = -c(step_+step)./12 - (r^2)./(8*sqrt(3)*pi);
        E(step_+step) = -pi*c(step_+step)./(6*r);

        r2 = real(r);

        Nindex = N + 1;

        r3(r2,Nindex) = r2;
        E2(r2,Nindex) = E(step_+step);

        %%
        fileID = fopen('data.txt','a');
        %Print to file
        fprintf(fileID,'Iterations = %i; r = %2.3f+%2.3fi; F = %s; b = %s\n',...
            mm,real(r),imag(r),num2str(F(step_+step),'%2.5f'),num2str(b_0));
        %Print on screen
        fprintf('Iterations = %i; r = %2.3f+%2.3fi; F = %s; b_0 = %s; S(2*b_0) = %s\n',...
            mm,real(r),imag(r),num2str(F(step_+step),'%2.5f'),num2str(b_0),num2str(S(2*b_0)));
        fprintf('cosh_next_b_0 = %s; cosh_next_b_0^2-1 = %s; log() = %s; bphase/pi = %s\n',...
            num2str(cosh_next_b_0),num2str(cosh_next_b_0^2-1),...
            num2str(log(cosh_next_b_0+sqrt(cosh_next_b_0^2-1))),num2str(bphase/pi));
        log(cosh_next_b_0+sqrt(cosh_next_b_0^2-1));
        pause(0.05);
        fclose(fileID);

    end
 end
     %% plot E vs r
        plot(r3,E2,'*b')
        title('E(r) vs r (1-Particle)')
        xlabel('r')
        ylabel('E(r)')

    fprintf('-----------------------------------------------\n');

%end
```

Appendix D

Riemann Surface Calculation Code for the Coupled Harmonic Oscillator

This Maple program produces the Riemann surface for the coupled harmonic oscillator found in Chapter 9.

```
r := 4;
E := g -> sqrt(2*g + 2 + 2*sqrt(1 + 2*g));
E1 := g -> sqrt(2 + 2*g - 2*sqrt(1 + 2*g));
E2 := g -> -sqrt(2 + 2*g - 2*sqrt(1 + 2*g));
E3 := g -> -sqrt(2*g + 2 + 2*sqrt(1 + 2*g));

p := plot3d(Re(E(x + y*I)), x = -r .. r, y = -r .. r);
p1 := plot3d(Re(E1(x + y*I)), x = -r .. r, y = -r .. r);
p2 := plot3d(Re(E2(x + y*I)), x = -r .. r, y = -r .. r);
p3 := plot3d(Re(E3(x + y*I)), x = -r .. r, y = -r .. r);


with(plots)
[animate, animate3d, animatecurve, arrow, changecoords,
complexplot, complexplot3d, conformal, conformal3d,
contourplot, contourplot3d, coordplot, coordplot3d,
densityplot, display, dualaxisplot, fieldplot, fieldplot3d,
gradplot, gradplot3d, implicitplot, implicitplot3d, inequal,
interactive, interactiveparams, intersectplot, listcontplot,
listcontplot3d, listdensityplot, listplot, listplot3d,
loglogplot, logplot, matrixplot, multiple, odeplot, pareto,
plotcompare, pointplot, pointplot3d, polarplot, polygonplot,
polygonplot3d, polyhedra_supported, polyhedraplot, rootlocus,
semilogplot, setcolors, setoptions, setoptions3d, shadebetween,
spacecurve, sparsematrixplot, surfdata, textplot, textplot3d,
tubeplot]


display([p, p1, p2, p3]);
```

Bibliography

[1] *Advanpix - Multiprecision Computing Toolbox*. URL: `https://www.advanpix.com/`.

[2] C. Ahn, G. Delfino and G. Mussardo. 'Mapping between the Sinh-Gordon and Ising Models'. *Physics Letters B* 317.4 (Nov. 1993), pp. 573–580.

[3] M.T. Batchelor. 'The Bethe ansatz after 75 years'. *Physics Today* 60.1 (2007), pp. 36–40.

[4] A.F. Beardon. *A Primer on Riemann Surfaces*. 1st ed. London Mathematical Society Lecture Notes Series, 1984.

[5] C.M. Bender, A. Felski, N. Hassanpour, S. P. Klevansky and A. Beygi. 'Analytic structure of eigenvalues of coupled quantum systems.' *Physica Scripta.* 92.1 (Nov. 2016), p. 015201.

[6] C.M. Bender and T.T. Wu. 'Anharmonic Oscillator'. *Physical Review* 184 (1969), p. 1231.

[7] H. Bethe. 'Zur Theorie Der Metalle'. *Zeitschrift für Physik* 71.3-4 (1931), pp. 205–226.

[8] R. Blumenhagen and E. Plauschinn. *Introduction to Conformal Field Theory With Applications to String Theory*. 1st ed. Springer-Verlag, 2009.

[9] M.L. Boas. *Mathematical Methods in the Physical Sciences*. 3rd ed. John Wiley and Sons, Inc., 2006.

[10] D. Bombardelli. 'Lectures of S-matrices and Integrability'. *Journal of Physics A: Mathematical and Theoretical* 49.23 (2016), p. 323003.

[11] R.V. Churchill and J.W. Brown. *Complex Analysis and Applications*. 8th ed. McGraw-Hill, 2009.

[12] P. D'Francesco, P. Mathieu and D. Sénéchal. *A Mathematical Introduction to Conformal Field Theory*. 1st ed. Springer-Verlag, 1997.

[13] P. Dennery and A. Krzywicki. *Mathematics for Physicists*. Harper & Row, 1967.

[14] S.K. Donaldson. *Riemann Surfaces*. 1st ed. Oxford University Press, 2011.

[15] P.E. Dorey. 'Exact Finite-Size Effects in Relativistic Field Theories'. *Lecture Notes, University of Utrecht* (Aug. 2008).

[16] P.E. Dorey. 'Exact S-matrices.' (May 1998), p. 41. URL: arXiv:hep-th/9810026.

[17] P.E. Dorey, T.C. Dunning and R. Tateo. 'The ODE/IM Correspondence'. *Journal of Physics A: Mathematical and Theoretical* 40.32 (2007), R205.

[18] P.E. Dorey and R. Tateo. 'Excited states by analytic continuation of TBA equations'. *Nuclear Physics B* 482.3 (1996), pp. 639–659.

[19] P.E. Dorey and R. Tateo. 'Excited states in simple perturbed conformal field theories'. *Nuclear Physics B* 515.3 (1998), pp. 575–623.

[20] H. Eckle. *Models of Quantum Matter*. 1st ed. Oxford University Press, 2019.

[21] R.J. Eden, P.V. Landshoff, D.I. Olive and J.C. Polkinghorne. *The Analytic S-Matrix*. 1st ed. Cambridge University Press, 1966.

[22] P. Ginsparg. *Applied Conformal Field Theory*. Sept. 1988.

[23] T.R. Klassen and E. Melzer. 'The thermodynamics of purely elastic scattering theories and conformal perturbation theory.' *Nuclear Physics B* 250.3 (Feb. 1991), pp. 635–689.

[24] F. Levkovich-Maslyuk. 'The Bethe Ansatz'. *Journal of Physics A: Mathematical and Theoretical* 49.32 (July 2016), p. 28.

[25] D. Meier. 'Analytic Continuation of the TBA equation in the sinh-Gordon model'. Diploma Thesis. Durham University and ETH, 2009.

[26] R.K. Pathria and P.D. Beale. *Statistical Mechanics*. 3rd ed. Butterworth-Heinemann, 2011.

[27] J.D. Qualls. 'Lectures on Conformal Field Theory' (Nov. 2015).

[28] M. Schottenloher. *A Mathematical Introduction to Conformal Field Theory*. 2nd ed. Springer-Verlag, 2008.

[29] B. Simon and A. Dicke. 'Coupling constant analyticity for the anharmonic oscillator'. *Annals of Physics* 58 (1970), pp. 76–136.

[30] C. Teleman. *Riemann Surfaces*. 2003. URL: https://math.berkeley.edu/~teleman/math/Riemann.pdf.

[31] S.J. van Tongeren. 'Introduction to the thermodynamic Bethe ansatz'. *Journal of Physics A: Mathematical and Theoretical* 49.32 (2016), p. 323005.

[32] Al.B. Zamolodchikov. 'On the Thermodynamic Bethe Ansatz Equation in Sinh-Gordon Model'. *Journal of Physics A: Mathematical and General* 39.41 (2006), p. 12863.

[33] Al.B. Zamolodchikov. 'Thermodynamic Bethe ansatz for RSOS scattering theories'. *Nuclear Physics B* 358 (1991), pp. 497–523.

[34] Al.B. Zamolodchikov. 'Thermodynamic Bethe ansatz in relativistic models: Scaling 3-state Potts and Lee-Yang models.' *Nuclear Physics B* 342.3 (1990), pp. 695–720.