Functional relations in conformal field theory and their solutions

Dissertation with the aim of achieving a doctoral degree

at the Faculty of Mathematics, Computer Science and Natural Sciences

Department of Mathematics

Universität Hamburg

submitted by Lorenz Hilfiker

Hamburg, 2018

The following evaluators recommend the admission of the dissertation:

Prof. Dr. Ingo Runkel (thesis supervisor)
Prof. Dr. Gleb Arutyunov

Day of oral defense: September 13th, 2018

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den 30. 4. 2018

Lorenz Hilfiker

Publications

Chapters 1 and 2 as well as Appendix A of this thesis are based on the following preprint:

L. Hilfiker, I. Runkel, *Existence and uniqueness of solutions to Y-systems and TBA equations*, [arXiv:1708.00001 [math-ph]].

Contents

Introduction

This thesis is motivated by mathematical questions that arise in the study of 2-dimensional integrable quantum field theories (IQFTs) as well as 2-dimensional conformal field theories (CFTs).

The notion of IQFT is not a mathematically rigorous one. Loosely speaking, it is used to describe all quantum field theories which exhibit sufficient symmetry to make them amenable to exact solution. A characteristic feature of IQFTs is the existence of an infinite number of independent observables I_j which commute both with the Hamiltonian H of the theory as well as among one another:

$$[H, I_j] = 0 \ , \qquad [I_i, I_j] = 0$$

The first equation says that the I_j are conserved charges (or *integrals of motion*), and the second one states that they share eigenstates. In integrability jargon, these equations describe a set of *integrals of motion in involution*. One of the first steps in “solving” an integrable quantum field theory is to identify such integrals of motion, and to compute their spectrum. We call this the *spectral problem* of IQFT.

CFTs are quantum field theories which are invariant under conformal transformations. They contain a subclass of quantum field theories known as topological quantum field theories (TFTs), which are invariant under homeomorphisms. The situation regarding mathematical rigour is different for these types of QFTs than for IQFTs. The notion of TFT is well-defined, and since its first axiomatisation by Atiyah [At] the study of TFT has become a branch of pure mathematics. The notion of CFT has no axiomatic description to date which includes all known examples, even though several (rather complicated) sets of axioms have been proposed [Seg, Gaw]. Nevertheless, many aspects of CFT have found a rigorous mathematical description, building on close connections with fields such as vertex algebras and tensor categories.

Integrable quantum field theories generalise topological quantum field theories and conformal field theories,

$$TFT \ \subset \ CFT \ \subset \ IQFT \ \subset \ QFT \ .$$

In fact, a rich source for generating IQFTs is to start with a CFT and deform it in a way that preserves as many of the integrals of motion as possible. This approach was pioneered in [Za1], where the integrability of certain deformations of Virasoro minimal CFTs was first corroborated. From a mathematical point of view, this means that non-conformal IQFTs are not completely out of reach: it may be possible to rigorously define them as deformations of CFT that one has well under control, and to apply some of the machinery of CFT to deal with them.

This thesis has two related parts:

- Part I is a rather general contribution to the spectral problem of IQFT. We study solutions of a certain class of functional difference equations (*Y-systems*) and integral equations (*TBA equations*) which are central to the spectral problem in many IQFTs. The main result is an existence and uniqueness theorem for a particular type of solutions.

- Part II is about the explicit construction of integrals of motion in involution in CFT, whereby known results are generalised to the case of CFT with twisted boundary conditions. This construction is illustrated in the free compactified boson CFT, where a specific set of integrals of motion (known as *T-operators*) is studied. The investigation of their spectrum uses results from Part I.

It should be noted that while we mostly take our motivation from the study of quantum field theory, there is a very close relationship between IQFTs and the statistical mechanics of solvable lattice models as well as spin chains. Some of the results of this thesis, especially in part I, should connect to this context as well.

Before we outline in more detail part I and II, let us give a brief heuristic and historical introduction into the over-arching topic of functional relations and nonlinear integral equations in IQFT. We do not in any way aim for a comprehensive review of the vast physics literature on the topic.

Functional relations and nonlinear integral equations in IQFT

Functional relations are ubiquitous in IQFT. Most notably, perhaps, the additional symmetries force the scattering matrix of an IQFT to satisfy the so-called Yang-Baxter functional equation [Do]. A very similar equation with the same name is satisfied by the R-matrix of a quasi-triangular Hopf algebra (also known as *quantum group*) [Kas]. The similarity is no coincidence. It indicates the intimate connection between integrability and quantum groups. We will have a taste of that connection in part II.

The functional relations most relevant to the spectral problem, however, are of a different kind: In a nutshell, the large number of integrals of motion in IQFTs typically contains a distinguished subset of integrals of motion, each depending on a complex *spectral parameter* λ, which satisfy a (not necessarily closed) system of non-linear finite difference equations as functions of λ. If those integrals of motion are in involution, then all their eigenvalues satisfy the same set of equations.

The power of such functional difference equations in solving the spectral problem was first discovered in the context of 2-dimensional solvable lattice models by Baxter [Bax, Ch. 9]. The protagonists of his functional relations were mutually commuting transfer matrices (the "t" in "transfer" giving rise to the name "T-operator") or so-called Baxter Q-operators, or both at once. The respective functional relations themselves are today known as *T-systems*, *Q-systems*, and *Baxter TQ-relations*.

In a parallel development in statistical mechanics, the so-called Thermodynamic Bethe Ansatz was developed to study thermodynamic properties of systems such as a gas of particles moving on a circle [YY] or spin chains [Tak]. This approach yields non-linear integral equations (*TBA equations*) for the free energy of the vacuum state as a function of temperature, which can easily be solved numerically (by iteration) to great accuracy. It was also realised that the solutions of TBA equations extend to analytic solutions of functional difference equations which we today call *Y-systems*.

In [Za2], the TBA approach was adapted to study the thermodynamics of relativistic particles in IQFTs with known scattering matrix. The reformulation of these TBA equations as a Y-system was first described in [Za3]. At about the same time, the intimate relationship between T-systems and Y-systems (and thus TBA equations) began to emerge. A relationship was also established between a different type of non-linear integral equations (called *Destri de Vega equations* [DdV]) and TQ-relations [BLZ2].

While originally the non-linear integral equations only described the eigenvalues of observables in the vacuum state, it became clear that eigenvalues of excited states solve modified TBA or DdV equations [KP, KM, Fe1, DT, BLZ4, FMQR], all of which correspond to one and the same set of functional difference equations. The solutions for different states exhibit the same asymptotic behaviours dictated by physical considerations, but have distinct patterns of roots inside some distinguished domain of the complex plane. The positions of the roots in this domain enter as parameters into the equations, and they satisfy a complex set of constraints which make numerical solution much less trivial than for the vacuum. In this way, the spectral problem can be condensed to the following problem:

Find all solutions (with specified asymptotic behaviour and with specified analytic properties such as being entire and real on the real axis) to a set of functional finite difference equations.

There is typically no hope of a closed form solution. The problem should be considered solved, for a given model, if for any state in the spectrum it is possible to express its eigenvalues (of T-operators or Q-operators) as solutions of a set of equations that

1. can be shown to have a unique solution,

2. can easily be solved numerically to great accuracy.

In the past decade, the second aim has been achieved in a number of important IQFTs by exploiting a relation between Y-systems and the Hirota equation [KLWZ, GKV], and by a related method known as the quantum spectral curve approach which is based on an elaborate manipulation of Q-systems [Gr]. In contrast, the first goal has not been a major focus in the physics literature. The main results of Part I of the present thesis are a contribution towards the first goal.

Before we turn our attention to the specific Y-systems and TBA equations that we will study, let us mention that the same type of functional relations also appear in higher-dimensional field theories which exhibit features of integrability, most notably in super

Yang-Mills theory. Due to its central role in understanding the so-called AdS/CFT correspondence between gauge and string theories, the Y-system in $N = 4$ super Yang-Mills theory (which is a generalisation of the Y-systems we treat in this thesis) has received a particularly staggering amount of attention during the past decade. Note that the solutions in this case have quite non-trivial analytic properties. This underlines the broader relevance of the functional relations approach.

Part I: Y-systems and TBA equations

There seems to be no precise definition of what constitutes a Y-system. Rather, the term comprises a a number of second order difference equations of a similar structure which appear in the context of integrable models. Moreover, Y-systems are intimately related with the representation theory of certain quantum groups, and arise also in other mathematical contexts such as cluster algebras. A comprehensive review of Y-systems and their applications can be found for example in [KNS]. In this thesis we consider a special, one might say fundamental, class of Y-systems.

The Y-system which we focus on is the set of equations

$$Y_j(z+i)Y_j(z-i) = \prod_{k=1}^{N}(1+Y_k(z))^{G_{jk}}$$

for complex functions $Y_j(z)$, $j = 1, ..., N$ (defined at least in the closure of the domain $\{z \in \mathbb{C} | -1 < \mathrm{Im}(z) < 1\}$), where G_{jk} are the non-negative entries of a matrix $\mathbf{G}$ which comes with the model. We allow for greater generality than previously considered by allowing the matrix elements G_{jk} to be real rather than just integers.

If $\mathbf{G}$ is the adjacency matrix of an ADE Dynkin diagram or a tadpole graph ($T_N = A_{2N}/\mathbb{Z}_2$), there exist known IQFTs in whose spectral problem this Y-system appears, the so-called *ADE scattering theories* [KM, Za3]. To take two well-known examples: the Y-system for the E_8 case corresponds to the effective field theory which describes the critical Ising model perturbed by a magnetic field; the T_1 Y-system ($N = 1$, $G_{11} = 1$) corresponds to the only integrable deformation of the Yang-Lee minimal CFT.

As outlined earlier, different solutions of the same Y-system are described by different associated TBA equations. The TBA equation corresponding to a solution of the above Y-system which has no roots in the strip $\{z \in \mathbb{C} | -1 < \mathrm{Im}(z) < 1\}$ is the following set of coupled non-linear integral equations for bounded functions $f_n(x)$, $n = 1, \dots, N$:

$$f_n(x) \;=\; \sum_{m,k=1}^{N} \int_{-\infty}^{\infty} [\Phi_{\mathbf{C}}]_{nm}\,(x-y)\left(G_{mk}\log(e^{-a_k(y)} + e^{f_k(y)}) - C_{mk} f_k(y)\right) dy$$

Here, C_{jk} are the matrix elements of an arbitrary real diagonalisable matrix $\mathbf{C}$ with spectral radius less than 2, and $\Phi_{\mathbf{C}}(x)$ is a matrix-valued function which can be obtained explicitly and which depends only on $\mathbf{G}$ and $\mathbf{C}$. The functions $a_n(x)$ are bounded from below and they are supposed to satisfy $f_n(x) = \log Y_n(x) - a_n(x)$. They are the asymptotic data which come as part of a specific model.

The TBA equations corresponding to solutions with roots in the strip are similar, but the root locations enter the equations as parameters.

Our results in Part I can be summarised as follows:

- We prove a statement (Proposition 1.2.1) establishing the correspondence between the TBA equations and the Y-system. We note that while our proof follows the standard steps, we are not aware of a previous proof in the literature, in the sense that all analytic questions are carefully addressed. The result is crucial to most proofs in the remainder of part I. Moreover, we also think that the consideration of a whole family of TBA equations (for every $\mathbf{C}$ rather than just for $\mathbf{C} = 0$ or $\mathbf{C} = \mathbf{G}$) is new. The freedom to choose $\mathbf{C}$ at discretion is a crucial cornerstone in the proof of our next results:

- We prove existence and uniqueness of a real bounded continuous solution to the above set of TBA equations for all $\mathbf{G}$ with spectral radius smaller than 2, under very mild assumptions on $a_n(x)$ (Theorem 2.1.2). The $\mathbf{G}$ which fit our assumptions include in particular the adjacency matrix of finite Dynkin diagrams of types A_N–G_N or tadpole graphs – these are called *Dynkin TBAs* in [RTV]. We will find Dynkin TBAs of D_N and B_N type in part II of the thesis.

- Under slightly more restrictive conditions on $a_n(x)$, the above unique TBA solution extends to a unique solution of the Y-system which is real on the real axis, which has no roots in the strip $\{z \in \mathbb{C} | -1 < \mathrm{Im}(z) < 1\}$, and whose asymptotics is captured by $a_n(x)$ (Theorem 2.1.1).

- We study asymptotic properties of the unique TBA solution as $x \to \infty$ (Lemma 3.1.1), and we find and prove a general asymptotic expansion formula (Proposition 3.1.5) which has been conjectured to contain relevant information on local integrals of motion.

- Finally, we introduce the TBA equations associated to solutions of Y-systems with roots in the domain $\{z \in \mathbb{C} | -1 < \mathrm{Im}(z) < 1\}$ (Proposition 3.2.8), and we study some properties (Lemma 3.2.1 and Lemma 3.2.10) of such solutions which will be useful in part II.

The main mathematical results are clearly the existence and uniqueness Theorems 2.1.1 and 2.1.2, since they fill, as we believe, a significant gap in the literature. The only previously existing rigorous proof that we are aware of is for the above TBA equation with $N = 1$, $G_{11} = 1$ and $a_n(x) \sim \cosh(x)$ [FKS], which we briefly discuss in section 2.2.5.

Part I is organised as follows. In Chapter 1 the Y-system and the TBA equations are introduced, and the relationship between them is proven. Chapter 2 contains the uniqueness theorems and their proofs. In Chapter 3 we collect some additional results on asymptotic properties of the solutions, and on more general solutions with roots. Finally, Appendix A contains the proofs of some results used in Chapter 1.

Part II: Twisted T-operators from topological defects in conformal field theory

Although 2-dimensional CFTs clearly have enough symmetry to be considered integrable, it is not immediately clear how to construct families of integrals of motion in involution. Finding integrals of motion is not hard, but ensuring mutual commutativity is. This task was investigated in a series of papers by Bazhanov, Lukyanov and Zamolodchikov [BLZ1, BLZ2, BLZ3, BLZ4]. The authors give explicit prescriptions for T-operators, Q-operators and Y-functions in Virasoro minimal CFTs, the compactified free boson CFT, as well as in some integrable perturbations of these. Their construction is based on a somewhat heuristic quantisation of monodromy matrices of the integrable Korteweg-de Vries partial differential equation, from which T-operators and Q-operators emerge as associated to representations of certain quantum groups.

The T-operators of these authors were identified as perturbed topological defect operators by Runkel [Ru1, Ru2], making use of the mathematically rigorous TFT-approach to CFT developed in [FRS1, FRS3, FRS4, FRS5, FjFRS] and topological defects in particular [FrFRS]. This discovery was followed by the development, together with Manolopoulos, of a categorical framework to study perturbed defects [MaRu]. In a paper together with Buecher [BR], the formalism was pushed further to include integrable perturbations of CFTs.

The main result of [MaRu] was that perturbed topological defects in CFT are governed by a rigid tensor category $\mathcal{D}_F$ (where $\mathcal{D}$ stands for the category of topological defects of the CFT and F indicates the perturbation applied to the defects), and that their composition and superposition is governed by the Grothendieck ring $K_0(\mathcal{D}_F)$. Hence, a family of objects in $\mathcal{D}_F$ whose tensor products mutually commute (up to isomorphism) corresponds to a family of perturbed defect operators in involution. Moreover, short exact sequences in $\mathcal{D}_F$ with a tensor product in the middle provide a way to rewrite the composition of perturbed defect operators as a superposition of other perturbed defect operators. In this way, one can obtain functional relations such as T-systems and Q-systems.

In practise, computing the Grothendieck ring in question is usually hard. Yet it is also unnecessary to know $K_0(\mathcal{C})$ completely. In [BR], the precise relationship between representations of the q-deformed loop algebra of $\mathfrak{sl}_2$ and T-operators in the uncompactified sine-Gordon model (an integrable deformation of the uncompactified free boson CFT) was established rigorously. This relationship yields at least some short exact sequences in $\mathcal{D}_F$, which were enough for the authors to identify the T-operators and find the infinite T-system of that model.

Part II of this thesis builds on these results. The main generalisations are the following:

- We study perturbed *twisted* defect operators. These are perturbed defect operators which act on states with non-periodic boundary conditions. We investigate conditions under which they are integrals of motion (Lemmas 6.2.8 and 6.2.9 and Corollary

6.2.10), and we find that their composition and superposition is still controlled by $K_0(\mathcal{C}_F)$ (Propositions 6.2.11 and 6.2.12).

- In chapter 7 we study *twisted* T-operators in the free boson *compactified* on a cylinder. We reproduce the twisted T-system of [BLZ4] (Section 7.3.1). For rational compactification radii, the correspondence with the q-deformed loop algebra of $\mathfrak{sl}_2$ is for q a root of unity. This allows a simplification of the T-system, reducing it to a finite set of equations. We obtain the Y-system of D_N Dynkin type and use results from part I to derive TBA equations (section 7.3.2).

- We perform a continuous deformation in the twist parameter to obtain numerical solutions not only for the vacuum state, but also for some low lying excited states, both in the twisted and untwisted theory (Section 7.4).

This thesis completes the series of works which reproduce the BLZ results regarding T-operators from the point of view of the TFT-approach to CFT. The challenge to achieve the same for Q-operators is open. Candidates for Q-operators among the perturbed defects have been proposed in [BR], but they are quite clearly not the BLZ Q-operators and moreover lack some crucial features.

The need for Q-operators is particularly obvious considering the fact that they have often shown to be more powerful in tackling the spectral problem numerically. The T-operators and Y-systems are more universal in the sense that they seem to exist in more models, but, as our investigations for the free boson show, the associated non-linear integral equations are pretty hard to solve except for very few states in the spectrum. It was Q-operators which were at the heart of new approaches in the physics literature in recent years (most notably the ODE-IM correspondence [DDT] and the quantum spectral curve approach [Gr]) which have lead to major advances in understanding the spectral problem in many IQFTs.

Part II is organised as follows. Chapter 4 contains some necessary background about tensor categories and algebraic structures that arise in and from them. In Chapter 5 we give a brief review of 2-dimensional CFT and the TFT-approach to rational CFT. Chapter 6 introduces defect operators and contains our results on twisted perturbed defect operators. In Chapter 7, the example of T-operators in the free compactified boson is treated.

Part I.

Y-systems and TBA equations

1. From Y-systems to TBA equations

1.1. A class of Y-systems and their associated $\mathbf{C}$-dependent TBA equations

In this section we describe the class of Y-systems which we will consider throughout part I of this thesis (and in the later parts of chapter 7). Afterwards we introduce families of TBA equations which depend on a real matrix $\mathbf{C}$, and sketch the argument connecting the Y-system to these TBA equations.

Let us start by fixing some notation which we need to formulate the results and which we will use throughout.

Notations 1.1.1.

- Let $\mathbb{K}$ stand for $\mathbb{R}$ or $\mathbb{C}$. We denote by

  $$BC(\mathbb{R}, \mathbb{K})$$

  the functions from $\mathbb{R}$ to $\mathbb{K}$ which are continuous and bounded, and by

  $$BC_{-}(\mathbb{R}, \mathbb{R})$$

  the continuous real-valued functions on $\mathbb{R}$ which are bounded from below.

- For $a > 0$ we denote by $\mathbb{S}_a := \{z \in \mathbb{C} | -a < \mathrm{Im}(z) < a\}$ the open horizontal strip in $\mathbb{C}$ of height $2a$, and by $\overline{\mathbb{S}}_a$ its closure. We define the spaces of functions

  $$\mathcal{A}(\mathbb{S}_a) \supset \mathcal{BA}(\mathbb{S}_a) ,$$

  where $\mathcal{A}(\mathbb{S}_a)$ is the space of $\mathbb{C}$-valued functions which are analytic on $\mathbb{S}_a$ and have a continuous extension to $\overline{\mathbb{S}}_a$, and $\mathcal{BA}(\mathbb{S}_a)$ consists of those functions in $\mathcal{A}(\mathbb{S}_a)$ which are in addition bounded on $\overline{\mathbb{S}}_a$.

- We fix the constants

  $$N \in \mathbb{Z}_{>0} \quad , \quad s \in \mathbb{R}_{>0} .$$

- We denote by

  $$\mathrm{Mat}_{<2}(N) \subset \mathrm{Mat}(N, \mathbb{R})$$

  the subset of real-valued $N \times N$ matrices which can be diagonalised over the real numbers, and whose eigenvalues lie in the open interval $(-2, 2)$.

The Y-system

For $\mathbf{G} \in \mathrm{Mat}(N, \mathbb{R})$ (with entries G_{nm}) and $Y_1, \ldots, Y_N \in \mathcal{A}(\mathbb{S}_s)$, the Y-system which we consider throughout part I of this thesis is the set of functional equations

$$Y_n(x+is)Y_n(x-is) = \prod_{m=1}^{N} (1+Y_m(x))^{G_{nm}} \quad , \tag{Y}$$

for $n = 1, \ldots, N$ and all $x \in \mathbb{R}$. If $\mathbf{G}$ is not integer valued, one needs to give a prescription how to deal with the multi-valuedness of the right hand side. We will later do this, in chapter 2 by demanding Y_n to be positive and real-valued on the real axis, and in section 3.2 by making a specific choice.

Valid asymptotics

If $Y_n \in \mathcal{A}(\mathbb{S}_s)$ has no zeros, we may pick $h_n \in \mathcal{A}(\mathbb{S}_s)$ such that $Y_n(z) = e^{h_n(z)}$ for all $z \in \overline{\mathbb{S}}_s$. Denote $h_n(z) = \log Y_n(z)$. We can think of a function $a_n \in \mathcal{A}(\mathbb{S}_s)$ as capturing the asymptotic behaviour of $Y_n(z)$ if $\log Y_n(z) - a_n(z)$ is bounded on $\overline{\mathbb{S}}_s$. This condition is independent of the branch choice for the logarithm. To formulate many of our results, we need to single out a certain type of asymptotic behaviour.

Definition 1.1.2. We call $\mathbf{a} \in \mathcal{A}(\mathbb{S}_s)^N$ (with components $a_n(z)$) a *valid asymptotics for* (Y) if

1. for $n = 1, ..., N$, a_n is real valued on $\mathbb{R}$ and the functions $e^{-a_n(x)}$ and $\frac{d}{dx}e^{-a_n(x)}$ are bounded on $\mathbb{R}$,

2. $\mathbf{a}$ satisfies
$$\mathbf{a}(x+is) + \mathbf{a}(x-is) = \mathbf{G} \cdot \mathbf{a}(x) \qquad \text{for all } x \in \mathbb{R} \ . \tag{1.1.1}$$

One important valid asymptotics for (Y) is simply $\mathbf{a} = 0$, in which case the Y_n themselves are bounded. We will see in Corollary 2.1.3 below that then in fact the Y_n are constant. The Perron-Frobenius eigenvector $\mathbf{w}$ of $\mathbf{G}$ provides a whole family of valid asymptotics. By our assumptions on $\mathbf{G}$, $\mathbf{w}$ can be chosen to have positive entries and its eigenvalue lies strictly between 0 and 2 (see Theorem 2.2.9). Then for any choice of $\gamma \in \mathbb{R}_{>0}$ such that $\mathbf{G} \cdot \mathbf{w} = 2\cos(\gamma)\mathbf{w}$, the functions

$$\mathbf{a}(x) = e^{\gamma x/s}\, \mathbf{w} \qquad \text{and} \qquad \mathbf{a}(x) = e^{-\gamma x/s}\, \mathbf{w} \tag{1.1.2}$$

are valid asymptotics for (Y). We can also take linear combinations with positive coefficients; in particular the symmetric choice $\mathbf{a}(x) = r\cosh(\gamma x/s)\, \mathbf{w}$ is considered frequently in the context of massive relativistic IQFT, where $\mathbf{w}$ takes the role of the mass vector and $r > 0$ represents the volume.

The TBA equations

Let $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$ and consider the following Fourier transform of a matrix-valued function, for $x \in \mathbb{R}$,

$$\Phi_{\mathbf{C}}(x) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\left(2\cosh(sk)\mathbf{1}-\mathbf{C}\right)^{-1} dk\ . \tag{1.1.3}$$

The integral is well defined since by the condition on the eigenvalues of $\mathbf{C}$, the components of the integrand are Schwartz-functions. Then the components of $\Phi_{\mathbf{C}}$ are also Schwartz-functions which moreover are real and even. See Section 1.2.2 for more details on $\Phi_{\mathbf{C}}$. For $\mathbf{a} \in BC_{-}(\mathbb{R},\mathbb{R})^N$, $\mathbf{G} \in \mathrm{Mat}(N,\mathbb{R})$, and $\mathbf{C}$ as above, the TBA equation is the following nonlinear integral equation for a vector-valued function $\mathbf{f} \in BC(\mathbb{R},\mathbb{R})^N$:

$$\mathbf{f}(x) \;=\; \int_{-\infty}^{\infty} \Phi_{\mathbf{C}}(x-y)\cdot\Big(\mathbf{G}\cdot\log\left(e^{-\mathbf{a}(y)}+e^{\mathbf{f}(y)}\right)-\mathbf{C}\cdot\mathbf{f}(y)\Big)\,dy\ . \tag{TBA}$$

Here we used the short-hand notation $\log\left(e^{-\mathbf{a}(y)}+e^{\mathbf{f}(y)}\right)$ to denote the function $\mathbb{R}\to\mathbb{R}^N$ with entries

$$\left[\log\left(e^{-\mathbf{a}(y)}+e^{\mathbf{f}(y)}\right)\right]_j := \log\left(e^{-a_j(y)}+e^{f_j(y)}\right)\ . \tag{1.1.4}$$

The integral (TBA) is well-defined because the components of $\mathbf{a}$ are bounded from below, $\mathbf{f}$ is bounded and the components of $\Phi_{\mathbf{C}}$ are Schwartz-functions.

Connection between Y-systems and TBA equations

We will now outline the transformations linking the Y-system and the TBA equation, see [Za3] and e.g. [RTV, DDT, vTo].

Rewrite Y_n in (Y) as

$$Y_n(z) = e^{f_n(z)+a_n(z)}\ , \tag{1.1.5}$$

where $f_n(z)$ are bounded functions and $a_n(z)$ are valid asymptotics for (Y). Upon taking the logarithm, one verifies that the $a_n(z)$ cancel out and one remains with the set of finite difference equations

$$f_n(x+is)+f_n(x-is) = \sum_{m=1}^{N} G_{nm}\log\left(e^{-a_m(x)}+e^{f_m(x)}\right)\ . \tag{1.1.6}$$

Even though it looks like a trivial modification of the above equation, it will be crucial for us to add a linear term in $\mathbf{f}$ to both sides (we switch to vector notation to avoid too many indices, recall also the convention (1.1.4)),

$$\mathbf{f}(x+is)+\mathbf{f}(x-is)-\mathbf{C}\cdot\mathbf{f}(x) = \mathbf{G}\cdot\log\left(e^{-\mathbf{a}(y)}+e^{\mathbf{f}(y)}\right)-\mathbf{C}\cdot\mathbf{f}(x)\ . \tag{1.1.7}$$

To get rid of the nonlinearity for a while, we replace the $\mathbf{f}$-dependent function on the right hand side simply by a suitable function $\mathbf{g}$,

$$\mathbf{f}(x+is)+\mathbf{f}(x-is)-\mathbf{C}\cdot\mathbf{f}(x) = \mathbf{g}(x)\ . \tag{1.1.8}$$

The difference equation can now be solved by a Greens-function-like approach. Namely, the function $\Phi_{\mathbf{C}}$ from (1.1.3) gives rise to a representation of the Dirac δ-distribution (see Lemma 1.2.16 for details):

$$\Phi_{\mathbf{C}}(x+is) + \Phi_{\mathbf{C}}(x-is) - \mathbf{C}\cdot\Phi_{\mathbf{C}}(x) = \delta(x)\mathbf{1}_N \ . \tag{1.1.9}$$

This allows one to write a solution to the functional equation (1.1.8) as an integral,

$$\mathbf{f}(x) = \int_{-\infty}^{\infty} \Phi_{\mathbf{C}}(x-y)\cdot\mathbf{g}(y)\, dy \ . \tag{1.1.10}$$

Substituting the right hand side of (1.1.7) for $\mathbf{g}$ in (1.1.10) produces (TBA).

Remark 1.1.3. In the case $\mathbf{C} = 0$, the matrix $\Phi_{\mathbf{C}}(x)$ is proportional to the identity matrix and corresponds to the *standard kernel* (often denoted by s) which produces the *universal* or *simplified TBA equations* of the physics literature. The case $\mathbf{C} = \mathbf{G}$ yields the canonical TBA equations which emphasise the relation to the relativistic scattering matrix $\mathbf{S}(x)$ if such is available (see Remark 1.2.18). Specifically, we have (see e.g. [DDT])

$$[\Phi_{\mathbf{G}}(x)\cdot\mathbf{G}]_{nm} = \frac{i}{2\pi}\frac{d}{dx}\log\left(S_{nm}(x)\right) \ , \tag{1.1.11}$$

More details and references can be found e.g. in [Za3, RTV, vTo]. Note that our Green's function $\Phi_{\mathbf{G}}(x)$ has to be multiplied with $\mathbf{G}$ to obtain the canonical kernel used in the physics TBA-literature. In this thesis, we consistently treat the factor $\mathbf{G}$ not as part of the kernel, but absorb it in the function $\mathbf{g}(x)$. This is a natural convention for $\mathbf{C} = 0$, and we preserve this convention for general $\mathbf{C}$.

In the rest of this chapter we give a detailed statement and proof of the above claim that (1.1.10) solves (1.1.8), see Proposition 1.2.1. Such results have been widely used in the physics literature, but the only rigorous proof that we are aware of is [TW, Lem. 2], which treats the case $N = 1$ and $\mathbf{C} = 0$ and imposes a decay condition on $\mathbf{f}(x)$ for $x \to \pm\infty$. Therefore, in this chapter we give a proof in the generality we require.

1.2. Solution to a set of difference equations

For two functions $\mathbf{F} : \mathbb{C} \to \mathrm{Mat}(N, \mathbb{C})$ with components F_{nm} and $\mathbf{g} : \mathbb{R} \to \mathbb{C}^N$ with components g_n let us formally denote by

$$(\mathbf{F} \star \mathbf{g})(z) := \int_{-\infty}^{\infty} \mathbf{F}(z-t)\cdot\mathbf{g}(t)\, dt \tag{1.2.1}$$

the convolution product, where the components of the integrand are $[\mathbf{F}(z-t)\cdot\mathbf{g}(t)]_n = \sum_{m=1}^{N} F_{nm}(z-t)g_m(t)$. In this section we will prove the following important proposition which will allow us to relate finite difference equations and integral equations. Recall from

Notations 1.1.1 the definition of the subset $\mathrm{Mat}_{<2}(N) \subset \mathrm{Mat}(N,\mathbb{R})$, and that we fixed constants $N \in \mathbb{Z}_{>0}$ and $s \in \mathbb{R}_{>0}$. Recall also the definition of $\Phi_{\mathbf{C}}$ from (1.1.3).

Recall also that a function $f : \mathbb{R} \to \mathbb{K}$ is called *Hölder continuous* if there exist $0 < \alpha \leq 1$ and $C > 0$, such that

$$|f(x) - f(y)| \leq C\,|x-y|^{\alpha} \qquad \text{for all} \ \ x, y \in \mathbb{R} \ . \tag{1.2.2}$$

If $\alpha = 1$, then f is called *Lipschitz continuous.*

The remainder of this chapter will be devoted to proving the following statement:

Proposition 1.2.1. *Let* $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$*,* $\mathbf{f} : \mathbb{R} \to \mathbb{C}^N$ *and* $\mathbf{g} \in BC(\mathbb{R},\mathbb{C})^N$*. Consider the following two statements:*

1. $\mathbf{f}$ *is real analytic and can be analytically continued to a function* $\mathbf{f} \in \mathcal{BA}(\mathbb{S}_s)^N$ *satisfying the functional equation*

$$\mathbf{f}(x+is) + \mathbf{f}(x-is) - \mathbf{C}\cdot\mathbf{f}(x) = \mathbf{g}(x) \qquad \textit{for all} \ \ x \in \mathbb{R} \ . \tag{1.2.3}$$

2. $\mathbf{f}$ *and* $\mathbf{g}$ *are related via the convolution*

$$\mathbf{f}(x) = (\Phi_{\mathbf{C}} \star \mathbf{g})\,(x) \qquad \textit{for all} \ \ x \in \mathbb{R} \ . \tag{1.2.4}$$

We have that 1 implies 2. If the components of $\mathbf{g}$ *are in addition Hölder continuous, then 2 implies 1.*

The basic reasoning of our proof is the same as in the physics literature, as outlined in the previous section. We take care to prove all the required analytical properties, which to our knowledge has not been done before in this generality. We also note that the observation that Proposition 1.2.1 applies to all $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$ (rather than just $\mathbf{C} = 0$ and adjacency matrices of certain graphs) seems to be new. The freedom to choose arbitrary $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$ will be crucial in chapter 2 in the proof of the existence and uniqueness theorems stated there.

The proof relies on a number of ingredients developed in Sections 1.2.1–1.2.3. The proof of Proposition 1.2.1 itself is given in Section 1.2.4.

1.2.1. The Green's function $\phi_d(z)$

In this subsection we introduce a family of meromorphic functions $\phi_d(z)$, parametrized by a real number

$$d \in (-2, 2) \ . \tag{1.2.5}$$

In this section we adopt the convention that whenever the parameter d appears, it is understood that it is chosen from the above range.

The function ϕ_d will be central to our problem since it plays, in analogy with the theory of differential equations, the role of a Green's functions for the difference operator

$$D[f](x) \ := \ f(x+is) + f(x-is) - d\cdot f(x) \ . \tag{1.2.6}$$

We start by defining the function ϕ_d on the real axis in terms of a Fourier integral representation.

Definition 1.2.2. The function $\phi_d : \mathbb{R} \to \mathbb{R}$ is defined by

$$\phi_d(x) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{2\cosh(sk) - d}\, dk. \tag{1.2.7}$$

Note that $\phi_d(x)$ is real and even, since it is the Fourier transform of a real and even function. Moreover, $(2\cosh(sk) - d)^{-1}$ is in the Schwartz space of rapidly decaying functions; thus also $\phi_d(x)$ is a Schwartz function, and Fourier inversion holds.

Example 1.2.3. Consider the case $d = 0$. The Fourier integral can then be computed explicitly, with the result

$$\phi_0(z) = \frac{1}{4s\cosh\left(\frac{\pi}{2s}z\right)}\ . \tag{1.2.8}$$

This function is called the *universal kernel* or *standard kernel* in the physics literature. It has a meromorphic continuation to the whole complex plane, with poles of first order in $z = is(1 + 2\mathbb{Z})$.

Explicit expressions for ϕ_d in terms of hyperbolic functions for other specific values of d can be found e.g. in [DDT, App. D] and [BLZ4, Eqn. 4.22]. A general expression is given in Remark 1.2.15 below. Here we will proceed by deducing the analytic structure of $\phi_d(z)$ from its definition in (1.2.7).

Recall that smoothness of a function is related to the rate of decay of its Fourier transform. If the decay is exponential, analytic continuation is possible; the faster the exponential decay, the further one can analytically continue:

Lemma 1.2.4. *Let $f(x)$ be a function on $\mathbb{R}$ whose Fourier transform $\hat{f}(k)$ exists. Suppose there exist constants $A, a > 0$ such that $|\hat{f}(k)| \le A\exp(-a|k|)$ for all $k \in \mathbb{R}$. Then $f(x)$ has an analytic continuation to the strip $\mathbb{S}_a$.*

For a proof see for example [SS, Ch. 4, Thm. 3.1]. In particular, $\phi_d(x)$ can be analytically continued to the strip $\mathbb{S}_s$. In fact, it can be continued to a meromorphic function with poles of order ≤ 1 in $is\mathbb{Z}$ (Lemma 1.2.8 below). To get there we need some preparation. We start with:

Lemma 1.2.5. *$\phi_d(z)$ has an analytic continuation to the complex plane with two cuts, $\mathbb{C} \setminus (i\mathbb{R}_{\ge s} \cup i\mathbb{R}_{\le -s})$.*

Proof. Take any $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and consider the function

$$\tilde{\phi}_d(z) = e^{-i\theta} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ike^{-i\theta}z} \left(2\cosh\left(ske^{-i\theta}\right) - d\right)^{-1} dk\ . \tag{1.2.9}$$

By Lemma 1.2.4, this integral is analytic in $z \in e^{i\theta}\mathbb{S}_{s\cdot\cos\theta}$, a strip in the z-plane tilted by the angle θ. We claim that $\tilde{\phi}_d(z)$ and $\phi_d(z)$ coincide in the intersection $e^{i\theta}\mathbb{S}_{s\cdot\cos\theta} \cap \mathbb{S}_s$ of their respective analytic domains. This can be checked via contour deformation. We will first show that for $z \in e^{i\theta}\mathbb{S}_{s\cdot\cos\theta} \cap \mathbb{S}_s$ we have

$$\phi_d(z) = \frac{1}{2\pi}\int_{e^{-i\theta}\mathbb{R}} e^{ikz}(2\cosh(sk) - d)^{-1} dk \ . \tag{1.2.10}$$

Since for $|d| < 2$, the integrand $k \mapsto e^{ikz}(2\cosh(sk) - d)^{-1}$ has no poles away from the imaginary axis, rotating the contour by $-\theta$ does not pick up any residues. It remains to verify that there are no contributions from infinity. We express $z = x + iy$ and $k = u + iv$ in real coordinates, in terms of which the absolute value of the integrand can be written as

$$\left|\frac{e^{ikz}}{2\cosh(sk) - d}\right| = \left|\frac{e^{-(uy+vx)}}{e^{su}e^{isv} + e^{-su}e^{-isv} - d}\right| \ . \tag{1.2.11}$$

On the circular contour components one can parametrise $v = -u\tan(\tau)$, with τ running from 0 to θ. When $u \to \pm\infty$, the right hand side of (1.2.11) approaches

$$\left|e^{-u(y\pm s)-vx}\right| = \left|e^{-u(y\pm s-\tan(\tau)x)}\right| \ . \tag{1.2.12}$$

Thus, if the inequalities

$$\begin{aligned} y &> \tan(\tau)x - s \\ y &< \tan(\tau)x + s \end{aligned} \tag{1.2.13}$$

are satisfied for all τ between 0 and θ, then the two circular integrals do indeed vanish when the radius is taken to infinity. But these inequalities just describe the strips $e^{i\tau}\mathbb{S}_{s\cdot\cos\tau}$ in the z-plane, and their intersection for all $\tau \in [0, \theta]$ is precisely $\mathbb{S}_s \cap e^{i\theta}\mathbb{S}_{s\cdot\cos\theta}$.

Now, substituting $k' = e^{i\theta}k$ in (1.2.10) shows that $\phi_d(z) = \tilde{\phi}_d(z)$ on the intersection of their domains, and hence $\tilde{\phi}_d(z)$ is the analytic continuation of $\phi_d(z)$ to the strip $e^{i\theta}\mathbb{S}_{s\cdot\cos\theta}$. Consequently, $\phi_d(z)$ has an analytic continuation to the union of all of these strips,

$$\bigcup_{\theta\in(-\frac{\pi}{2},\frac{\pi}{2})} e^{i\theta}\mathbb{S}_{s\cdot\cos\theta} = \mathbb{C} \setminus (i\mathbb{R}_{\geq s} \cup i\mathbb{R}_{\leq -s}) \ . \tag{1.2.14}$$

□

In order to understand the behaviour of $\phi_d(z)$ on the whole imaginary axis, it is natural to start with the case $d = 0$ whose analytic structure we know explicitly (Example 1.2.3). When comparing $\phi_d(z)$ to $\phi_0(z)$ we will need to control the derivative $\frac{\partial}{\partial d}\phi_d(z)$.

Lemma 1.2.6. *For all $z \in \mathbb{C} \setminus (i\mathbb{R}_{\geq s} \cup i\mathbb{R}_{\leq -s})$, the partial derivative $\frac{\partial}{\partial d}\phi_d(z)$ exists and is an analytic function on $\mathbb{C} \setminus (i\mathbb{R}_{\geq 2s} \cup i\mathbb{R}_{\leq -2s})$. For $z \in e^{i\theta}\mathbb{S}_{2s\cdot\cos\theta}$ it has the integral representation*

$$\frac{\partial}{\partial d}\phi_d(z) = e^{-i\theta}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ike^{-i\theta}z}\left(2\cosh(ske^{-i\theta}) - d\right)^{-2} dk \ . \tag{1.2.15}$$

Proof. For any $z \in e^{i\theta}\mathbb{S}_{s\cdot\cos\theta}$, consider $\phi_d(z)$ given by the integral representation (1.2.9). Write $ke^{-i\theta} = u + iv$ with $u, v \in \mathbb{R}$. Assuming $u \geq 0$, one then estimates

$$\begin{aligned}&\left|2\cosh\left(ske^{-i\theta}\right) - d\right| \;=\; \left|e^{su}e^{isv} + e^{-su}e^{-isv} - d\right| \\ \geq\; &\left|e^{su}e^{isv}\right| - \left|e^{-su}e^{-isv}\right| - |d| \;\geq\; e^{su} - 1 - |d| \;\geq\; e^{s\cos\theta\,|k|} - 3\;. \end{aligned} \tag{1.2.16}$$

The same overall estimate applies for $u < 0$ as well. For $|k|$ large enough, the last expression on the right hand side becomes bigger than $\frac{1}{2}e^{s\cos\theta\,|k|}$ and we can estimate

$$\left|2\cosh\left(ske^{-i\theta}\right) - d\right|^{-1} \;\leq\; 2\,e^{-s\cos\theta\,|k|} \qquad \text{(for } |k| \text{ large enough)}\;. \tag{1.2.17}$$

Next, writing $z = e^{i\theta}(x + iy)$ with $x \in \mathbb{R}$ and $y \leq s\cos\theta$ we obtain, for all $k \in \mathbb{R}$,

$$\left|e^{ike^{-i\theta}z}\right| \;=\; e^{-ky} \;\leq\; e^{s\cos\theta\,|k|}\;. \tag{1.2.18}$$

For $|k|$ large enough and $z \in e^{i\theta}\mathbb{S}_{s\cdot\cos\theta}$ we can now estimate

$$\begin{aligned}\left|\frac{\partial}{\partial d}e^{ike^{-i\theta}z}\left(2\cosh(ske^{-i\theta}) - d\right)^{-1}\right| &= \left|e^{ike^{-i\theta}z}\right| \cdot \left|2\cosh(ske^{-i\theta}) - d\right|^{-2} \\ &\leq 4\,e^{-s\cos\theta\,|k|}\;. \end{aligned} \tag{1.2.19}$$

One can choose $k_0 > 0$ large enough, such that this estimate applies for all $d \in (-2, 2)$ and $|k| \geq k_0$.

Let $\varepsilon > 0$, and define $D := 2 - \varepsilon$. Since for any $k_0 > 0$, the integrand of (1.2.15) is continuous (and finite) as a function of (k, d) in the compact region $[-k_0, k_0] \times [-D, D]$, it is in particular bounded. One can therefore find a constant $A > 0$ such that the integrand of (1.2.15) is bounded by $A\,e^{-s\cos\theta\,|k|}$ for all $k \in \mathbb{R}$ and $d \in [-D, D]$.

The integrand is thus majorised by the integrable function $A\,e^{-s\cos\theta\,|k|}$ for all $d \in [-D, D]$ and therefore integration and d-derivative can be swapped, establishing (1.2.15) for all $d \in [-D, D]$ and $z \in e^{i\theta}\mathbb{S}_{s\cdot\cos\theta}$. Since $\varepsilon > 0$ was arbitrary, this extends to all $d \in (-2, 2)$, proving the first part of the claim. Moreover, according to Lemma 1.2.4, the integral on the right-hand-side of (1.2.15) is actually analytic for $z \in e^{i\theta}\mathbb{S}_{2s\cdot\cos\theta}$. By uniqueness of the analytic continuation it must also coincide with $\frac{\partial}{\partial d}\phi_d(z)$ on this larger domain. □

One consequence of the integral representation of the d-derivative is the following functional relation for ϕ_d.

Lemma 1.2.7. *For all $x \in \mathbb{R} \setminus \{0\}$ we have*

$$\phi_d(x + is) + \phi_d(x - is) - d\phi_d(x) = 0\;. \tag{1.2.20}$$

Proof. Fix $x \in \mathbb{R}$, $x \neq 0$, and write

$$\mathcal{L}_x(d) := \phi_d(x + is) + \phi_d(x - is) - d\phi_d(x)\;. \tag{1.2.21}$$

We will show that $\mathcal{L}_x(d)$ solves the initial value problem

$$\mathcal{L}_x(0) = 0 \quad , \quad \frac{\partial}{\partial d}\mathcal{L}_x(d) = 0 \qquad \text{for all } d \in (-2, 2) \ . \tag{1.2.22}$$

The initial condition in (1.2.22) is straightforward to check by recalling from Example 1.2.3 that $\phi_0(z) = \left(4s\cosh\left(\frac{\pi}{2s}z\right)\right)^{-1}$. In order to prove the differential equation in (1.2.22), we use the integral representation (1.2.15) for $\theta = 0$ and $z \in \mathbb{S}_{2s}$:

$$\begin{aligned} \frac{\partial}{\partial d}\mathcal{L}_x(d) &= \frac{\partial}{\partial d}\phi_d(x+is) + \frac{\partial}{\partial d}\phi_d(x-is) - \phi_d(x) - d\frac{\partial}{\partial d}\phi_d(x) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\frac{F(k)}{(2\cosh(sk)-d)^2}dk, \end{aligned} \tag{1.2.23}$$

where

$$F(k) = e^{-ks} + e^{ks} - (2\cosh(sk) - d) - d = 0 \ . \tag{1.2.24}$$

Hence, (1.2.22) follows. □

By Lemma 1.2.5, the function $x \mapsto \mathcal{L}_x(d)$ has an analytic continuation to all $z \in \mathbb{C} \setminus i\mathbb{R}$. The functional relation (1.2.20) thus extends to this domain:

$$\phi_d(z+is) + \phi_d(z-is) - d\phi_d(z) = 0 \qquad \text{for all} \quad z \in \mathbb{C} \setminus i\mathbb{R} \ . \tag{1.2.25}$$

Using this, we now show:

Lemma 1.2.8. *$\phi_d(z)$ has a meromorphic continuation to the whole complex plane which satisfies:*

i) The poles are all of first order and form a subset of $is\mathbb{Z} \setminus \{0\}$.

ii) For $z \in \mathbb{C} \setminus is\mathbb{Z}$ we have $\phi_d(z+is) + \phi_d(z-is) - d\phi_d(z) = 0$.

iii) For $z \in \mathbb{C} \setminus is\mathbb{Z}$ and $n \in \mathbb{Z}$ we have

$$\phi_d(z+ins) = \frac{\sin(n\gamma)}{\sin(\gamma)}\phi_d(z+is) - \frac{\sin((n-1)\gamma)}{\sin(\gamma)}\phi_d(z) \ , \tag{1.2.26}$$

where $\gamma \in \mathbb{R}$ satisfies $d = 2\cos(\gamma)$.

Proof.
• *Relation* (1.2.26) *holds on* $\mathbb{C} \setminus i\mathbb{R}$: Writing $p_n(z) := \phi_d(z+ins)$, we can rewrite (1.2.25) as $p_{n+1} + p_{n-1} - dp_n = 0$. It is straight-forward to check that this recursion relation is solved by

$$p_n = \frac{\sin(n\gamma)}{\sin(\gamma)}p_1 - \frac{\sin((n-1)\gamma)}{\sin(\gamma)}p_0 \ . \tag{1.2.27}$$

• *ϕ_d has an analytic continuation to $\mathbb{C}$ minus the points $is\mathbb{Z} \setminus \{0\}$*: The right hand side of (1.2.26) is actually analytic in $\{z \in \mathbb{C} | -s < \mathrm{Im}(z) < 0\}$. Hence the same holds for the left hand side, that is, ϕ_d is analytic on the shifted strips $\{z \in \mathbb{C} | sn < \mathrm{Im}(z) < (n+1)s\}$ for all $n \in \mathbb{Z}$. Combining this with Lemma 1.2.5, which states that $\phi_d(z)$ is analytic on $\mathbb{C} \setminus (i\mathbb{R}_{\geq s} \cup i\mathbb{R}_{\leq -s})$, we obtain the claim.

In particular, (1.2.25) and (1.2.26) hold on $\mathbb{C} \setminus is\mathbb{Z}$, showing parts ii) and iii) of the lemma.

• *ϕ_d is Lipschitz continuous in d on $[-D, D]$ for $D < 2$:* Consider again the case $d = 0$, given by $\phi_0(z) = \left(4s\cosh\left(\frac{\pi}{2s}z\right)\right)^{-1}$: this is a meromorphic function with poles of first order located at $is(1+2\mathbb{Z})$. We are now going to show that in the strip $\mathbb{S}_{2s}$ the pole structure of $\phi_d(z)$ for general d coincides with the pole structure of $\phi_0(z)$.

Recall the derivative $\frac{\partial}{\partial d}\phi_d(z)$ defined inside $\mathbb{S}_{2s}$ by the integral representation (1.2.15). We write $y = \mathrm{Im}(z)$, where $y \in (-2s, 2s)$, and obtain

$$\left|\frac{\partial}{\partial d}\phi_d(z)\right| \leq \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ky}}{(2\cosh(sk) - d)^2} dk \ . \tag{1.2.28}$$

Let $\varepsilon > 0$ be arbitrary so that $D := 2 - \varepsilon > 0$ and $Y := 2s - \varepsilon > 0$. For $d \leq D$ and $|y| \leq Y$, the integrand can then be further estimated as follows:

$$\frac{e^{-ky}}{(2\cosh(sk) - d)^2} \leq \frac{2\cosh(yk)}{(2\cosh(sk) - d)^2} \leq \frac{2\cosh(Yk)}{(2\cosh(sk) - D)^2} \tag{1.2.29}$$

But

$$B := \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{2\cosh(Yk)}{(2\cosh(sk) - D)^2} dk \tag{1.2.30}$$

still converges (as $|Y| < 2s$), and hence

$$\left|\frac{\partial}{\partial d}\phi_d(z)\right| \leq B \ . \tag{1.2.31}$$

Put differently, B is a Lipschitz constant for $\phi_d(z)$ understood as a function of d on the interval $[-D, D]$. The Lipschitz condition reads

$$|\phi_d(z) - \phi_0(z)| \leq B \cdot d \ . \tag{1.2.32}$$

• *The pole order of ϕ_d at $is\mathbb{Z} \setminus \{0\}$ is 0 or 1*: By Lipschitz-continuity in d, we know that $|\phi_d(z) - \phi_0(z)| \leq B \cdot d$ for all $z \in \mathbb{S}_{2s} \setminus \{\pm is\}$. Since ϕ_0 has first order poles at $\pm is$, it follows that so does ϕ_d. Since (1.2.26) holds on $\mathbb{C} \setminus is\mathbb{Z}$, the pole structure is as claimed. This finally proves part i) of the lemma.

We remark that the Lipschitz-continuity in d does not extend beyond the strip $\mathbb{S}_{2s}$. Indeed, ϕ_0 has no pole at $\pm 2is$, whereas for $d \neq 0$, (1.2.20) forces $\phi_d(z)$ to have a pole there. □

Next we turn to the growth properties of ϕ_d. We need the following result from complex analysis (see e.g. [SS, Ch. 4, Thm. 3.4] for a proof).

Theorem 1.2.9 (Phragmén-Lindelöf)**.** *Let f be a holomorphic function in the wedge $W = \{z \in \mathbb{C} \,|\, -\frac{\pi}{2\beta} < \arg(z) < \frac{\pi}{2\beta}\}$, $\beta > \frac{1}{2}$, which is continuous on the closure of W. Suppose that $|f(z)| \leq 1$ on the boundary of W and that there are $A, B > 0$ and $0 < \alpha < \beta$ such that $|f(z)| \leq Ae^{B|z|^\alpha}$ for all $z \in W$. Then $|f(z)| \leq 1$ on W.*

With the help of this theorem we can establish the following boundedness properties.

Lemma 1.2.10. *Let $\Theta \in [0, \frac{\pi}{2})$. Then for all $m, n \in \mathbb{Z}_{\geq 0}$ the function $z^m \frac{d^n}{dz^n}\phi_d(z)$ is bounded in the wedges $|\arg(z)| \leq \Theta$ and $|\arg(z) - \pi| \leq \Theta$.*

Proof. Consider the lines $z = e^{\pm i\Theta}\mathbb{R}$, which constitute the boundary of the wedges we are interested in. The integral representation (1.2.9) of $\phi_d(z)$, when restricted to these lines, yields functions in the Schwartz space, because $\left(2\cosh\left(ske^{\mp i\Theta}\right) - d\right)^{-1}$ are functions in the Schwartz space. By definition of the Schwartz space, the function $z^m \frac{d^n}{dz^n}\phi_d(z)$ is bounded on these two lines as well. Moreover, it is analytic in the interior of the wedges. We will now show that the growth of $z^m \frac{d^n}{dz^n}\phi_d(z)$ is less than exponential in the interior of the wedges. The statement of the lemma then follows from Theorem 1.2.9.

It suffices to show that $\frac{d^n}{dz^n}\phi_d(z)$ is bounded in the wedges $|\arg(z)| < \Theta$ and $|\arg(z) - \pi| < \Theta$. Let $x \in \mathbb{R}$, $\theta \in (-\Theta, \Theta)$, and use the integral representation (1.2.9) to obtain the x-independent estimate

$$\left|\frac{d^n}{dz^n}\phi_d(xe^{i\theta})\right| \leq \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\frac{k^n}{2\cosh\left(ske^{-i\theta}\right) - d}\right| dk \;. \tag{1.2.33}$$

Next we estimate the integrand by a θ-independent integrable function. Recall from (1.2.17) that for $|k|$ large enough we can estimate

$$\left|2\cosh\left(ske^{-i\theta}\right) - d\right|^{-1} \;\leq\; 2\,e^{-s\cos\theta\,|k|} \;\leq\; 2\,e^{-s\cos\Theta\,|k|} \;. \tag{1.2.34}$$

But for $|k|$ large enough, we also have $|k^n| \leq e^{\frac{s}{2}\cos\Theta\,|k|}$. In other words, there exists a $k_0 > 0$, independent of θ, such that for all $|k| \geq k_0$,

$$\left|\frac{k^n}{2\cosh\left(ske^{-i\theta}\right) - d}\right| \leq 2e^{-\frac{s}{2}\cos\Theta\,|k|} \;. \tag{1.2.35}$$

Meanwhile, the function $w \mapsto |2\cosh(sw) - d|$ is continuous in $\{w \in \mathbb{C} \,|\, |\arg(w)| \leq \Theta, |\mathrm{Re}(w)| \leq k_0\}$ and in the corresponding wedge with $|\arg(-w)| \leq \Theta$, and has no zeros in this bow tie shaped compact subset of the complex plane. Hence, it is bounded from below by a strictly positive number $M > 0$. Consequently, for all $|k| < k_0$,

$$\left|\frac{k^n}{2\cosh\left(ske^{-i\theta}\right) - d}\right| \leq \frac{k_0^n}{M} \;. \tag{1.2.36}$$

Plugging (1.2.35) and (1.2.36) into (1.2.33) yields the bound

$$\left|\frac{d^n}{dz^n}\phi_d(z)\right| \leq \frac{1}{2\pi}\left(\int_{-k_0}^{k_0}\frac{k_0^n}{M}dk + 2\int_{k_0}^{\infty} 2e^{-\frac{s}{2}\cos\Theta\,k}dk\right) \;, \tag{1.2.37}$$

valid for all z in the two wedges defined by Θ, and where the right hand side is finite and independent of θ. □

Corollary 1.2.11. *For $y \in \mathbb{R}$, let $\partial^n \phi_d^{[y]}(x) := \frac{d^n}{dz^n}\phi_d(x+iy)$ be the restrictions of $\frac{d^n}{dz^n}\phi_d(z)$ to horizontal lines. If $y \notin s\mathbb{Z} \setminus \{0\}$, then $(x+iy)^m \partial^n \phi_d^{[y]}(x) \in L_1(\mathbb{R})$ for all $n, m \in \mathbb{Z}_{\geq 0}$.*

Corollary 1.2.12. *For any $a, b \in \mathbb{R}$ and for all $n, m \in \mathbb{Z}_{\geq 0}$,*

$$\lim_{x \to \pm\infty} \int_a^b (x+it)^m \frac{d^n}{dz^n}\phi_d(x+it)\, dt = 0 \ . \tag{1.2.38}$$

To understand at which points of $is\mathbb{Z} \setminus \{0\}$ the function ϕ_d has a first order pole and at which points the singularity can be lifted, we compute the residues.

Lemma 1.2.13. *For $n \in \mathbb{Z}$, the residue of $\phi_d(z)$ in $z = isn$ is given by*

$$\mathrm{Res}_{isn}(\phi_d) = \frac{1}{2\pi i} \frac{\sin(n\gamma)}{\sin(\gamma)} \ , \tag{1.2.39}$$

where $\gamma \in \mathbb{R}$ satisfies $d = 2\cos(\gamma)$.

Proof. We start by computing the residue at is:

$$\begin{aligned}
2\pi i \,\mathrm{Res}_{is}(\phi_d) &\overset{(a)}{=} \int_{\mathbb{R}+\frac{1}{2}is} \phi_d(z)\, dz - \int_{\mathbb{R}+\frac{3}{2}is} \phi_d(z)\, dz \\
&= \int_{-\infty}^{\infty} \phi_d\left(x + \tfrac{is}{2}\right) dx - \int_{-\infty}^{\infty} \phi_d\left(x + \tfrac{3is}{2}\right) dx \\
&\overset{(b)}{=} \int_{-\infty}^{\infty} \phi_d\left(x + \tfrac{is}{2}\right) dx + \int_{-\infty}^{\infty} \phi_d\left(x - \tfrac{is}{2}\right) dx - d \int_{-\infty}^{\infty} \phi_d\left(x + \tfrac{is}{2}\right) dx \\
&\overset{(c)}{=} (2-d) \int_{-\infty}^{\infty} \phi_d(x) dx = (2-d) \frac{1}{2\cosh(sk) - d}\Big|_{k=0} = 1 \ .
\end{aligned} \tag{1.2.40}$$

Here, all integrals exist by Corollary 1.2.11. In step (a), the circular contour is deformed to two horizontal infinite lines, making use of Corollary 1.2.12 to ensure that no contribution is picked up when pushing the vertical parts of the contour to infinity. Step (b) is the functional relation in Lemma 1.2.8 ii). In step (c) all contours are moved to the real axis, using that ϕ_d is analytic in $\mathbb{S}_s$ (Lemma 1.2.8) and that by Corollary 1.2.12 there are no contributions from infinity.

But from (1.2.26) we see that

$$\mathrm{Res}_{isn}\phi_d = \frac{\sin(n\gamma)}{\sin(\gamma)} \mathrm{Res}_{is}\phi_d - \frac{\sin((n-1)\gamma)}{\sin(\gamma)} \mathrm{Res}_0 \phi_d \ . \tag{1.2.41}$$

Since $\mathrm{Res}_0 \phi_d = 0$ we obtain the statement of the lemma. □

We are now in a position to justify the notion that $\phi_d(z)$ is a Green's function for the difference operator (1.2.6):

Lemma 1.2.14. *$\phi_d(x)$ gives rise to a representation of the Dirac δ-distribution on $BC(\mathbb{R},\mathbb{C})$ via*

$$\lim_{y\nearrow s}\left(\phi_d(x+iy)+\phi_d(x-iy)-d\phi_d(x)\right)=\delta(x)\ . \tag{1.2.42}$$

Before we turn to the proof, we note that for $|y|<s$,

$$\begin{aligned}\phi_d(x+iy)+\phi_d(x-iy)-d\phi_d(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{ikx-yk}+e^{ikx+yk}-de^{ikx}}{2\cosh(sk)-d}dk\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ikx}\frac{2\cosh(yk)-d}{2\cosh(sk)-d}dk\ .\end{aligned} \tag{1.2.43}$$

In the limit $y\nearrow s$, the integrand on the right hand side approaches e^{ikx} pointwise. The usual exchange-of-integration-order argument proves that one obtains a Dirac δ-distribution on L_1-functions whose Fourier-transformation is also L_1. To show that we obtain a δ-distribution on $BC(\mathbb{R},\mathbb{C})$, we follow a different route.

Proof of Lemma 1.2.14. Since $\phi_d(z)$ has simple poles at $z=\pm is$ of residue $\pm\frac{1}{2\pi i}$ (see Lemma 1.2.13), we can write

$$\phi_d(z)=\pm\frac{1}{2\pi i}\frac{1}{(z\mp is)}+r_\pm(z)\ , \tag{1.2.44}$$

where $r_\pm(z)$ is now analytic at $z=\pm is$. In particular, by Lemma 1.2.10 $r_+(z)$ is bounded in the upper half of $\mathbb{S}_s$ and $r_-(z)$ is bounded in the lower half of $\mathbb{S}_s$. Then

$$\begin{aligned}\delta_y(x) &:= \phi_d(x+iy)+\phi_d(x-iy)-d\phi_d(x)\\ &= \frac{1}{2\pi i}\left(\frac{1}{(x+iy-is)}-\frac{1}{(x-iy+is)}\right)+r_+(x+iy)+r_-(x-iy)-d\phi_d(x)\\ &= \tilde{\delta}_y(x)+u_y(x)\ ,\end{aligned} \tag{1.2.45}$$

where

$$\tilde{\delta}_y(x):=\frac{1}{\pi}\frac{s-y}{x^2+(s-y)^2} \tag{1.2.46}$$

and $u_y(x):=r_+(x+iy)+r_-(x-iy)-d\phi_d(x)$.

Now suppose $f\in BC(\mathbb{R},\mathbb{C})$. We have to show that

$$\lim_{y\nearrow s}\int_{-\infty}^{\infty}\delta_y(x)f(x)dx=f(0). \tag{1.2.47}$$

In order to do that, let $\varepsilon>0$ be arbitrary and split the integral,

$$\int_{-\infty}^{\infty}\delta_y(x)f(x)dx=I_1(y)+I_2(y)+I_3(y)\ , \tag{1.2.48}$$

where

$$I_1(y) = \int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \delta_y(x) f(x) dx \ , \tag{1.2.49}$$

$$I_2(y) = \int_{-\varepsilon}^{\varepsilon} \tilde{\delta}_y(x) f(x) \ dx, \qquad I_3(y) = \int_{-\varepsilon}^{\varepsilon} u_y(x) f(x) dx \ . \tag{1.2.50}$$

The *Lorentz functions* $\tilde{\delta}_y(x)$ are a well-known representation of the Dirac δ-distribution on $BC(\mathbb{R}, \mathbb{C})$ as $y \to s$, so independently of ε we have

$$\lim_{y \nearrow s} I_2(y) = f(0) \ . \tag{1.2.51}$$

The functions $u_y(x)$ are uniformly bounded for $y \in [0, s]$, say by C. Hence,

$$|I_3(y)| \leq 2\varepsilon C \, \|f\|_\infty \ . \tag{1.2.52}$$

Finally, consider $\delta_y(x)$ on $\mathbb{R} \setminus (-\varepsilon, \varepsilon)$. Due to the functional equation (1.2.20), it converges pointwisely to zero on this domain as $y \to s$. But Lemma A.3.1 in connection with Lemma 1.2.10 even ensures uniform convergence, and this is still true for the function $x^2\delta_y(x)$. By means of the variable transformation $t = 1/x$ we can recast $I_1(y)$ as an integral over a finite interval:

$$I_1(y) = \int_{-\frac{1}{\varepsilon}}^{\frac{1}{\varepsilon}} \frac{1}{t^2} \delta_y(\tfrac{1}{t}) f(\tfrac{1}{t}) dt \ , \tag{1.2.53}$$

By what we just said, the integrand converges uniformly to zero on $[-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]$. Thus, integral and limit can be swapped, which results in

$$\lim_{y \nearrow s} I_1(y) = 0 \ . \tag{1.2.54}$$

We conclude that

$$\left| f(0) - \lim_{y \nearrow s} \int_{-\infty}^{\infty} \delta_y(x) f(x) dx \right| \leq 2\varepsilon C \, \|f\|_\infty \ . \tag{1.2.55}$$

As $\varepsilon > 0$ was arbitrary, the statement follows. □

Remark 1.2.15. One can actually give a simple explicit expression for $\phi_d(z)$ for arbitrary $d \in (-2, 2)$:

$$\phi_d(z) \ = \ \frac{1}{2s\sin(\gamma)} \cdot \frac{\sinh\left(\frac{\pi-\gamma}{s} z\right)}{\sinh\left(\frac{\pi}{s} z\right)} \ , \tag{1.2.56}$$

where $\gamma \in (0, \pi)$ is defined via $d = 2\cos(\gamma)$. This can be seen via a contour deformation argument using the analytical properties of ϕ_d established in this section. The explicit integral can also be found in tables, see e.g. [Er, 1.9 (6)]. However, we were unable to obtain it in a more straight-forward fashion, circumventing the analysis carried out in this section.

1.2.2. The N-dimensional Green's function $\Phi_{\mathbf{C}}(z)$

Now let us investigate an N-dimensional version of the Green's function $\phi_d(z)$. Recall from Notations 1.1.1 the definition of the subset $\mathrm{Mat}_{<2}(N) \subset \mathrm{Mat}(N,\mathbb{R})$, as well as from (1.1.3) the definition of $\Phi_{\mathbf{C}} : \mathbb{R} \to \mathrm{Mat}(N,\mathbb{R})$ for $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$.

Lemma 1.2.16. *$\Phi_{\mathbf{C}}(x)$ has the following properties:*

i) $\Phi_{\mathbf{C}}(x)$ and $\mathbf{C}$ are simultaneously diagonalisable for all $x \in \mathbb{R}$.

ii) Any matrix element of $\Phi_{\mathbf{C}}(x)$ can be written as a linear combination

$$[\Phi_{\mathbf{C}}]_{nm}(x) = \sum_{j=1}^{N} \Omega^j_{nm} \phi_{d_j}(x) \tag{1.2.57}$$

of one-dimensional Green's functions, with $d_j \in (-2,2)$ given by the eigenvalues of $\mathbf{C}$, and Ω^j_{nm} some real coefficients.

iii) $\Phi(x)$ gives rise to a representation of the Dirac δ-distribution on $BC(\mathbb{R},\mathbb{C})^N$:

$$\delta(x)\mathbf{1} = \lim_{y \nearrow s} \left(\Phi_{\mathbf{C}}(x+iy) + \Phi_{\mathbf{C}}(x-iy) - \mathbf{C}\cdot\Phi_{\mathbf{C}}(x)\right) \tag{1.2.58}$$

Proof. First of all, note that the matrix inverse in the definition is well-defined since $\mathbf{C}$ has spectral radius smaller than 2.

i) Let $\mathbf{D} \in \mathrm{Mat}(N,\mathbb{R})$ be a diagonal matrix such that $\mathbf{D} = \mathbf{T}^{-1}\mathbf{C}\mathbf{T}$ for some invertible $\mathbf{T} \in \mathrm{Mat}(N,\mathbb{R})$. Then

$$\begin{aligned}\mathbf{T}^{-1}\Phi_{\mathbf{C}}(x)\mathbf{T} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\, \mathbf{T}^{-1}(2\cosh(sk)\mathbf{1} - \mathbf{C})^{-1}\mathbf{T}\, dk \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\, (2\cosh(sk)\mathbf{1} - \mathbf{D})^{-1}\, dk \;=\; \Phi_{\mathbf{D}}(x)\end{aligned} \tag{1.2.59}$$

is also diagonal.

ii) Write $\mathbf{D} = \mathrm{diag}(d_1,...,d_N)$. Then the matrix elements can be written as

$$[\Phi_{\mathbf{C}}]_{nm}(x) = \sum_{j=1}^{N} T_{nj}[T^{-1}]_{jm}[\Phi_D]_{jj}(x) \tag{1.2.60}$$

where $\Omega^j_{nm} = T_{nj}[T^{-1}]_{jm}$ are real constants, and $[\Phi_D]_{jj}(x) = \phi_{d_j}(x)$. Since $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$, $|d_j| < 2$ holds for all $j = 1,...,N$.

iii) Using i) and ii) and applying the Green's function property of $\phi_d(z)$ (Lemma 1.2.14), one computes

$$\begin{aligned}
&\lim_{y\nearrow s}\left[\Phi_{\mathbf{C}}(x+iy)+\Phi_{\mathbf{C}}(x-iy)-\mathbf{C}\cdot\Phi_{\mathbf{C}}(x)\right]_{nm}\\
&\quad=\sum_{j=1}^{N}T_{nj}[T^{-1}]_{jm}\lim_{y\nearrow s}\left(\phi_{d_j}(x+iy)+\phi_{d_j}(x-iy)-d_j\phi_{d_j}(x)\right)\\
&\quad=\sum_{j=1}^{N}T_{nj}[T^{-1}]_{jm}\delta(x)=\delta_{nm}\delta(x)\ . \qquad (1.2.61)
\end{aligned}$$

This completes the proof. □

From the definition of ϕ_d in (1.2.7) and from Lemma 1.2.16 ii) we know that all components of $\Phi_{\mathbf{C}}(x)$ are Schwartz functions on $\mathbb{R}$ (cf. Lemma 1.2.10). Hence the Fourier transformation of $\Phi_{\mathbf{C}}(x)$ reproduces the integrand in (1.1.3). In particular, for $k=0$ we obtain the following integral, which we will need later:

$$\int_{-\infty}^{\infty}\Phi_{\mathbf{C}}(x)\,dx \ = \ \left(2\cdot\mathbf{1}-\mathbf{C}\right)^{-1}\ . \qquad (1.2.62)$$

Lemma 1.2.17. *Suppose* $\mathbf{C}$ *is non-negative. Then the matrix* $\Phi_{\mathbf{C}}(x)$ *is non-negative for all* $x\in\mathbb{R}$*.*

Proof. The integrand can be expanded into a Neumann series,

$$\begin{aligned}
(2\cosh(sk)\mathbf{1}-\mathbf{C})^{-1}&=(2\cosh(sk))^{-1}\left(\mathbf{1}-\frac{\mathbf{C}}{2\cosh(sk)}\right)^{-1}\\
&=\sum_{j=0}^{\infty}\frac{\mathbf{C}^j}{(2\cosh(sk))^{j+1}}\ , \qquad (1.2.63)
\end{aligned}$$

which converges absolutely since for $k\in\mathbb{R}$ all eigenvalues of $(2\cosh(sk))^{-1}\,\mathbf{C}$ are strictly smaller than 1. Fubini's theorem (with counting measure on $\mathbb{Z}_{\geq 0}$ and Lebesgue measure on $\mathbb{R}$) then justifies pulling the sum out of the Fourier integral, and we find that

$$\Phi_{\mathbf{C}}(x)=\sum_{j=0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ikx}\,(2\cosh(sk))^{-j-1}\,dk\right)\mathbf{C}^j\ . \qquad (1.2.64)$$

In Appendix A.1 it is shown that

$$\int_{-\infty}^{\infty}e^{ikx}\,(2\cosh(sk))^{-j-1}\,dk=\frac{\pi}{2^{j+1}j!s}\left(\prod_{\substack{l=j-1\\ \text{step}-2}}^{1}\left(\frac{x^2}{s^2}+l^2\right)\right)\cdot\begin{cases}\frac{1}{\cosh\left(\frac{\pi}{2s}x\right)} & \text{if } j \text{ even}\\ \frac{x}{s\sinh\left(\frac{\pi}{2s}x\right)} & \text{if } j \text{ odd}\end{cases}\ , \qquad (1.2.65)$$

which is a strictly positive function of $x\in\mathbb{R}$. Furthermore, $\mathbf{C}^j$ is a non-negative matrix. Hence, $\Phi_{\mathbf{C}}(x)$ is non-negative for all $x\in\mathbb{R}$. □

Remark 1.2.18. Suppose $\mathbf{G} \in \mathrm{Mat}_{<2}(N)$ is non-negative and irreducible. One of the equivalent ways to characterise irreducibility is that for each i, j there is an $m > 0$ such that $[(\mathbf{G})^m]_{ij} \neq 0$. Together with non-negativity of $\mathbf{G}$ and strict positivity of (1.2.65), this implies that $\Phi_{\mathbf{G}}(x)$ has *strictly positive* entries for all $x \in \mathbb{R}$. By Corollary 1.2.11 and Lemma 1.2.16 ii), the components of $\Phi_{\mathbf{G}}(x)$ are integrable, and so we can choose $\Psi_{\mathbf{G}} \in BC(\mathbb{R}, \mathrm{Mat}(N, \mathbb{R}))$ such that $\Psi_{\mathbf{G}}(x)$ has positive entries bounded away from zero and satisfies $\frac{d}{dx}\Psi_{\mathbf{G}}(x) = \Phi_{\mathbf{G}}(x)$. Comparing to Remark 1.1.3, we see that with the above assumption on $\mathbf{G}$, it is always possible to find an $\mathbf{S} \in BC(\mathbb{R}, \mathrm{Mat}(N, \mathbb{C}))$ such that (1.1.11) holds.

1.2.3. Convolution integrals involving $\phi_d(z)$

In this section we adopt again the convention (1.2.5) that the parameter d will always take values in the range

$$d \in (-2, 2) \ . \tag{1.2.66}$$

Just as in the case of differential equations, the Green's function approach to difference equations will eventually express solutions in terms of convolution integrals involving the Green's function. For $g \in BC(\mathbb{R}, \mathbb{C})$, the convolution with $\phi_d(z)$ is defined by

$$F_d[g](z) := \int_{-\infty}^{\infty} \phi_d(z - t) g(t) \ dt \ . \tag{1.2.67}$$

Due to Corollary 1.2.11, this function is well-defined on $\mathbb{S}_s$. As we will see in Section 1.2.4, it is important to understand the properties of such integrals as functions in z. That is the subject of this section.

The first question to ask is whether $F_d[g](z)$ is analytic. More generally: does the integration of a parameter-dependent analytic function preserve analyticity? The following lemma gives a criterion:

Lemma 1.2.19. *Let $D \subseteq \mathbb{C}$ be a complex domain. Suppose $f : D \times \mathbb{R} \to \mathbb{C}$ is a function with the following properties:*

1. *for every $t_0 \in \mathbb{R}$, the function $f(z, t_0)$ is analytic in D.*
2. *for every $z_0 \in D$, the function $f(z_0, t)$ is continuous on $\mathbb{R}$.*
3. *for every $z_0 \in D$ there exists a neighbourhood U and an $L_1(\mathbb{R})$-integrable function $M(t)$, such that $|f(z, t)| \leq M(t)$ for all $z \in U$ and all $t \in \mathbb{R}$.*

Then the function

$$F(z) = \int_{-\infty}^{\infty} f(z, t) \ dt \tag{1.2.68}$$

is analytic in D.

Proof. Let $z \in D$. Note that $F(z)$ is well-defined since the integrand is continuous (condition 2) and dominated by an integrable function (condition 3). Now take an arbitrary closed triangular contour Γ inside D. Define the function

$$L(z) := \int_{-\infty}^{\infty} |f(z,t)| \; dt \; . \tag{1.2.69}$$

Since by condition 3, $f(z,t)$ is locally dominated by an integrable function, $L(z)$ is continuous on D. Thus, on the compact contour Γ the function $L(z)$ is bounded and the integral

$$\oint_\Gamma \left(\int_{-\infty}^{\infty} |f(z,t)| \, dt \right) dz \tag{1.2.70}$$

is finite. This warrants the application of Fubini's theorem, followed by analyticity (condition 1):

$$\oint_\Gamma F(z)dz = \oint_\Gamma \int_{-\infty}^{\infty} f(z,t)dt \; dz = \int_{-\infty}^{\infty} \oint_\Gamma f(z,t)dz \; dt = 0 \; . \tag{1.2.71}$$

By Morera's theorem, the claim follows. □

This lemma can be applied to the convolution integral $F_d[g](z)$. Set $D = \mathbb{S}_s$, and for any given $z_0 \in \mathbb{S}_s$ set $U = B_\varepsilon(z_0)$ (the open ball with radius ε and center z_0) with some sufficiently small ε. By Lemma 1.2.10, a dominating integrable function $M(t)$ can be found by taking it to be a constant $B > 0$ for $|t - z_0| < T$ and $B|t - z_0|^{-2}$ else. B and T are to be chosen sufficiently large. We have shown:

Corollary 1.2.20. *For every $g \in BC(\mathbb{R}, \mathbb{C})$, the function $F_d[g](z)$ is analytic in $\mathbb{S}_s$.*

Note that in the same fashion as for $\phi_d(z-t)g(t)$, one can also use Lemma 1.2.10 to construct integrable dominating functions for $\frac{d^n}{dz^n}\phi_d(z-t)g(t)$. Hence, we are allowed to differentiate inside the integral:

$$\frac{d^n}{dz^n} F_d[g](z) = \int_{-\infty}^{\infty} \frac{d^n}{dz^n} \phi_d(z-t) \; g(t) \; dt \; . \tag{1.2.72}$$

More can be said about the nature of $F_d[g](z)$ and its derivatives. The following lemma provides a stepping stone.

Lemma 1.2.21. *Let $n \in \mathbb{Z}_{\geq 0}$. The function $y \mapsto \left\| \partial^n \phi_d^{[y]} \right\|_{L_1}$ is bounded in the compact interval $[-Y, Y]$ for every $0 < Y < s$.*

Proof. The function is well-defined due to Corollary 1.2.11. Now fix a $\Theta \in [0, \frac{\pi}{2})$ and $0 < Y < s$. Due to Lemma 1.2.10, there exists a $C > 0$ such that

$$\left| (x+iy)^2 \partial^n \phi_d^{[y]}(x) \right| \leq C \tag{1.2.73}$$

for all $x, y \in \mathbb{R}$ with $|\frac{y}{x}| < \tan\Theta$, and consequently

$$\left|\partial^n \phi_d^{[y]}(x)\right| \leq \begin{cases} \frac{C}{|x+iy|^2} \leq \frac{C}{x^2} & \text{for } |x| > \frac{Y}{\tan\Theta} \\ \max_{z\in\mathbb{S}_Y} |\partial^n \phi_d(z)| & \text{else} \end{cases} \tag{1.2.74}$$

for all $y \in [-Y, Y]$. The right hand side is in $L_1(\mathbb{R})$ and independent of y. Its integral over $\mathbb{R}$ provides a bound for $\left\|\partial^n \phi_d^{[y]}\right\|_{L_1}$ in the interval $[-Y, Y]$. □

Lemma 1.2.22. *Let $n \in \mathbb{Z}_{\geq 0}$. For every $g \in BC(\mathbb{R}, \mathbb{C})$, the function $\frac{d^n}{dz^n} F_d[g](z)$ is bounded in $\mathbb{S}_Y$ for all $0 < Y < s$.*

Proof. With 1.2.72, one has

$$\left|\frac{d^n}{dz^n} F_d[g](z)\right| \leq \int_{-\infty}^{\infty} \left|\frac{d^n}{dz^n}\phi_d(z-t)\right| |g(t)|\, dt \leq \sup_{t\in\mathbb{R}} |g(t)| \left\|\partial^n \phi_d^{[\mathrm{Im}(z)]}\right\|_{L_1} . \tag{1.2.75}$$

According to Lemma 1.2.21, the right-hand side is bounded for $|\mathrm{Im}(z)| \leq Y$. Hence, $F_d[g](z)$ is bounded in $\mathbb{S}_Y$. □

Since $\phi_d(z)$ has poles in $z = \pm is$, there is no obvious way to extend the domain of $F_d[g](z)$ beyond $\mathbb{S}_s$. Lemma 1.2.19 thus provides no information regarding the behaviour of this convolution integral as z approaches the boundary $\partial\mathbb{S}_s = \mathbb{R} \pm is$. Moreover, Lemma 1.2.22 can only be used to prove boundedness of $F_d[g](z)$ in a strip which is strictly contained in $\mathbb{S}_s$. In the remainder of this section, we will show that for g is Hölder continuous, $F_d[g](z)$ can be extended to $\overline{\mathbb{S}}_s$ as a bounded and continuous function. To this end, we need another result from complex analysis.

Let us relax the analyticity condition in Lemma 1.2.19: suppose $f(z, t)$ is analytic everywhere except in $z = t$, where it shall have a pole of first order. Consider a contour γ in D, and integrate over it:

$$F(z) = \int_\gamma f(z, t)\, dt . \tag{1.2.76}$$

The pole of $f(z, t)$ at $z = t$ causes $F(z)$ to have a branch cut along γ. Theorems describing this behaviour often go by the name of Sokhotski-Plemelji [Gak]. The next proposition is an instance of this for $\gamma = \mathbb{R}$, and it follows from a more more general statement proven in Appendix A.2.

Proposition 1.2.23. *Let $a > 0$ and let $h : \mathbb{S}_a \to \mathbb{C}$ be an analytic function such that both $zh(z)$ and $\frac{d}{dz}h(z)$ are bounded in $\mathbb{S}_a$. Moreover, let $g : \mathbb{R} \to \mathbb{C}$ be a bounded Hölder continuous function. Then*

$$F(z) = \int_{-\infty}^{\infty} \frac{h(z-t)}{z-t} g(t)\, dt \tag{1.2.77}$$

is analytic in $\mathbb{S}_a \setminus \mathbb{R}$. Moreover, the limits

$$F^{\pm}(x) := \lim_{y \searrow s} F(x \pm iy) \qquad (x \in \mathbb{R}) \tag{1.2.78}$$

exist and are uniform in x. The functions $F^{\pm}(x)$ are bounded on $\mathbb{R}$ and provide continuous extensions of $F(z)$ from the upper/lower half-plane to the real axis, related by

$$F^{+}(x) - F^{-}(x) \;=\; 2i\pi\, h(0)g(x) \;. \tag{1.2.79}$$

Now let us apply this result to convolution integrals involving $\phi_d(z)$: let $a < s$ and define the functions $h_{\pm}(z) = z\phi_d(z \pm is)$. By Lemma 1.2.8, $h_{\pm}(z)$ are analytic in $\mathbb{S}_a$. Due to Lemma 1.2.10, both $zh_{\pm}(z)$ and $\frac{d}{dz}h_{\pm}(z)$ are bounded in $\mathbb{S}_a$. Hence, $h_{\pm}(z)$ satisfy the conditions of Proposition 1.2.23 for any $a < s$. Specifically, this clarifies the behaviour of our convolution integral as we approach the boundary of the strip:

Corollary 1.2.24. *Let $g : \mathbb{R} \to \mathbb{C}$ be bounded and Hölder continuous. Then the function $F_d[g](z)$ has a continuous extension to $\overline{\mathbb{S}}_s$, and this extension is bounded on $\partial\mathbb{S}_s = \mathbb{R} \pm is$.*

Lastly, combining this result with the case $n = 0$ of Lemma 1.2.22, we obtain:

Lemma 1.2.25. *Let $g : \mathbb{R} \to \mathbb{C}$ be bounded and Hölder continuous. Then the function $F_d[g](z)$ is bounded in $\mathbb{S}_s$.*

Proof. By Corollary 1.2.24, there is some constant $B > 0$ such that $|F_d[g](z)| \leq B$ for all $z \in \partial\mathbb{S}_s$. According to Proposition 1.2.23, $F_d[g](x \pm iy) \to F_d[g](x \pm is)$ uniformly as $y \nearrow s$. Thus, for any given $\varepsilon > 0$ there exists a $\delta > 0$ such that $|F_d[g](z)| \leq B + \varepsilon$ for all $z \in \mathbb{S}_s \setminus \mathbb{S}_{s-\delta}$. In other words, $F_d[g](z)$ is bounded in $\mathbb{S}_s \setminus \mathbb{S}_{s-\delta}$. But on the other hand, by Lemma 1.2.22 $F_d[g](z)$ is also bounded in $\mathbb{S}_{s-\delta}$. □

The results of this section can now be summed up as follows: if g is bounded and Hölder continuous, then $F_d[g] \in \mathcal{BA}(\mathbb{S}_s)$.

1.2.4. Proof of Proposition 1.2.1

1 ⇒ 2: For $y \in \mathbb{R}$, define the family of continuous functions $\mathbf{f}^{[y]}(x) := \mathbf{f}(x+iy)$. Continuity of $\mathbf{f}$ on the closure of the strip $\mathbb{S}_s$ guarantees pointwise convergence $\mathbf{f}^{[y]} \to \mathbf{f}^{[s]}$ as $y \nearrow s$. By boundedness of $\mathbf{f}$, the components $f_m^{[y]}$ of $\mathbf{f}^{[y]}$ are uniformly bounded by some constant M. It follows that, for any fixed value of $b \in \mathbb{R}$ and any $d \in (-2, 2)$,

$$\left|\phi_d(b-x) f_m^{[y]}(x)\right| \leq M|\phi_d(b-x)| \qquad \text{for all} \;\; x \in \mathbb{R} \;. \tag{1.2.80}$$

The function on the right-hand-side is in $L_1(\mathbb{R})$ according to Corollary 1.2.11. Thus, by Lebesgue's dominated convergence theorem we can write

$$\int_{-\infty}^{\infty} \phi_d(x-t) f_m^{[s]}(t)dt = \lim_{y\nearrow s} \int_{-\infty}^{\infty} \phi_d(x-t) f_m^{[y]}(t)dt \;. \tag{1.2.81}$$

By a simple change of variables followed by contour deformation (which is now allowed because for $y < s$ the contour lies inside the analytic domain) one can transfer the appearance of y from f_m to ϕ_d:

$$\int_{-\infty}^{\infty} \phi_d(x-t) f_m(t+iy)dt = \int_{-\infty}^{\infty} \phi_d(x+iy-t) f_m(t)dt \tag{1.2.82}$$

Note that the integrals over the vertical parts of the contour vanish when pushed to infinity (see Corollary 1.2.12). Plugging (1.2.82) into (1.2.81), and making use of Lemma 1.2.16 ii) to write $[\Phi_{\mathbf{C}}]_{nm}$ in terms of the one-dimensional Green's functions ϕ_d, gives rise to the identity

$$\Phi_{\mathbf{C}} \star \mathbf{f}^{[\pm s]}(x) = \lim_{y \nearrow s} \Phi_{\mathbf{C}}^{[\pm y]} \star \mathbf{f}(x) \qquad \text{for all} \quad x \in \mathbb{R} \; . \tag{1.2.83}$$

Taking into account $[\mathbf{C}, \Phi_{\mathbf{C}}(x)] = 0$ due to Lemma 1.2.16 i) and distributivity of the convolution, (1.2.83) directly implies

$$\Phi_{\mathbf{C}} \star \left(\mathbf{f}^{[+s]} + \mathbf{f}^{[-s]} - (\mathbf{C} \cdot \mathbf{f})\right)(x) = \lim_{y \nearrow s} \left(\Phi_{\mathbf{C}}^{[+y]} + \Phi_{\mathbf{C}}^{[+y]} - \mathbf{C} \cdot \Phi_{\mathbf{C}}\right) \star \mathbf{f}(x) \tag{1.2.84}$$

for all $x \in \mathbb{R}$. On the left-hand-side we can substitute the functional relation (1.2.3), and on the right-hand-side apply Lemma 1.2.16 iii). This results in (1.2.4).

2 $\Rightarrow$ 1: According to Lemma 1.2.16 ii), the components of $\mathbf{f}(x) = (\Phi_{\mathbf{C}} \star \mathbf{g})(x)$ are given by real linear combinations of the form

$$\sum_{d,m} c_{d,m} F_d[g_m](x) \; , \tag{1.2.85}$$

where the $g_m(x)$ are bounded and Hölder continuous by assumption and $F_d[g_m](x)$ are the convolution integrals discussed in Section 1.2.3. According to Corollary 1.2.20, Corollary 1.2.24 and Lemma 1.2.25, these integrals have analytic continuations $F_d[g_m] \in \mathcal{BA}(\mathbb{S}_s)$. This shows that $\mathbf{f} \in \mathcal{BA}(\mathbb{S}_s)^N$ and

$$\mathbf{f}^{[y]}(x) = \left(\Phi_{\mathbf{C}}^{[y]} \star \mathbf{g}\right)(x) \; . \tag{1.2.86}$$

To obtain the functional equation (1.2.3) we basically reverse the above reasoning,

$$\begin{aligned}
\left(\mathbf{f}^{[+s]} + \mathbf{f}^{[-s]} - (\mathbf{C} \cdot \mathbf{f})\right)(x) &\overset{(a)}{=} \lim_{y \nearrow s} \left(\mathbf{f}^{[+y]} + \mathbf{f}^{[-y]} - (\mathbf{C} \cdot \mathbf{f})\right)(x) \\
&\overset{(1.2.86)}{=} \lim_{y \nearrow s} \left(\left(\Phi_{\mathbf{C}}^{[+y]} + \Phi_{\mathbf{C}}^{[-y]} - (\mathbf{C} \cdot \Phi_{\mathbf{C}})\right) \star \mathbf{g}\right)(x) \\
&\overset{(b)}{=} \mathbf{g}(x) \; .
\end{aligned} \tag{1.2.87}$$

Here (a) follows from pointwise convergence $\mathbf{f}^{[\pm y]}(x) \to \mathbf{f}^{[\pm s]}(x)$ due to continuity of $\mathbf{f}$ on $\overline{\mathbb{S}}_s$, and step (b) is the δ-function property from Lemma 1.2.16 iii).

This completes the proof of Proposition 1.2.1.

2. Existence and uniqueness of solution without roots

In this chapter, we continue to work with the Y-system (Y) and the associated TBA equations (TBA) introduced in chapter 1. We tackle the question of existence and uniqueness of certain solutions to these equations.

2.1. Statement of Theorems

Recall from Perron-Frobenius theory that a real $N \times N$ matrix is called *non-negative* if all its entries are ≥ 0, and *irreducible* if there is no permutation of the standard basis which makes it block-upper-triangular. The first main result of this chapter is the following existence and uniqueness statement.

Theorem 2.1.1. *Let* $\mathbf{G} \in \mathrm{Mat}_{<2}(N)$ *be non-negative and irreducible, and* $\mathbf{a} \in \mathcal{A}(\mathbb{S}_s)^N$ *a valid asymptotics for* (Y). *Then there exists a solution* $Y_1, \ldots, Y_N \in \mathcal{A}(\mathbb{S}_s)$ *to* (Y) *which satisfies, for* $n = 1, \ldots, N$ *,*

1. $Y_n(\mathbb{R}) \subseteq \mathbb{R}_{>0}$ *,* (real & positive)
2. $Y_n(z) \neq 0$ *for all* $z \in \overline{\mathbb{S}}_s$ *.* (no roots)
3. $\log Y_n(z) - a_n(z) \in \mathcal{BA}(\mathbb{S}_s)$*.* (asymptotics)

Moreover, this solution is the unique one in $\mathcal{A}(\mathbb{S}_s)$ *which satisfies properties 1–3.*

Recall that the logarithm in property 3 exists on all of $\overline{\mathbb{S}}_s$ as by condition 2, Y_n has no zeros, and that property 3 is not affected by the choice of branch for the logarithm.

Even though the unique solution $Y_1, \ldots, Y_N$ is initially only defined on $\overline{\mathbb{S}}_s$, using (Y) and property 2, it is easy to see that Y_n can be analytically continued at least to $\overline{\mathbb{S}}_{3s}$.

The second main result of this chapter is:

Theorem 2.1.2. *Let* $\mathbf{G} \in \mathrm{Mat}_{<2}(N)$ *be non-negative and irreducible,* $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$*, and* $\mathbf{a} \in BC_{-}(\mathbb{R}, \mathbb{R})^N$ *such that the components of* $e^{-\mathbf{a}}$ *are Hölder continuous. Then the following holds:*

i) The TBA equation (TBA) *has a unique solution* $\mathbf{f}_\star \in BC(\mathbb{R}, \mathbb{R})^N$*. The function* $\mathbf{f}_\star$ *is independent of the choice of* $\mathbf{C}$*.*

ii) $\mathbf{f}_\star$ *can be extended to a function in* $\mathcal{BA}(\mathbb{S}_s)^N$*, which we denote also by* $\mathbf{f}_\star$*. If* $\mathbf{a}$ *can be extended to a valid asymptotics, then the functions* $Y_n(z) = e^{a_n(z)+f_{\star,n}(z)}$*, for* $z \in \overline{\mathbb{S}}_s$ *and* $n = 1, \dots, N$*, provide the unique solution to* (Y) *with the properties as stated in Theorem 2.1.1.*

In the case $N = 1$, $\mathbf{G} = \mathbf{C} = 1$ and $\mathbf{a} = r\cosh(x)$, the existence and uniqueness of $\mathbf{f}_\star$ was already shown in [FKS] (see discussion in Section 2.2.5). Apart from that, not many results in this direction are known in the literature. Existence of a solution to TBA equations has in some cases been argued constructively. In [YY] and [La] solutions to some specific TBA equations (albeit with $\Phi_{\mathbf{C}}(x)$ replaced by functions substantially different from ours) are constructed from a specific starting function by iterating the equations and showing uniform convergence. A different approach to existence and uniqueness of solutions to certain TBA equations (which are a subset of ours) is suggested in a footnote in [KM, Sec. 3.2], where the authors propose to use a fixed point theorem due to Leray, Schauder and Tychonoff. A detailed proof is, however, not provided. Various methods to solve so-called Hammerstein integral equations (of which TBA as well as DdV equations are a subclass) are discussed in [AC], including several fixed point theorems. In particular their example 1 is similar in spirit to ours. We may also note that existence and uniqueness of solutions to an equation of Destri-de-Vega type in the XXZ model has been investigated in [Koz].

Let us specialise Theorems 2.1.1 and 2.1.2 to the case $\mathbf{a} = 0$. From the proofs of these theorems we get the following corollary.

Corollary 2.1.3. *For* $\mathbf{a} = 0$*, the unique solution* $\mathbf{f}_\star$ *from Theorem 2.1.2 is constant, and correspondingly, the unique solution* $Y_1, \dots, Y_N$ *from Theorem 2.1.1 is constant.*

Remark 2.1.4.

i) The constant case is interesting in itself. The functional equations (Y) turn into the *constant Y-system*

$$Y_n^2 = \prod_{m=1}^{N} (1 + Y_m)^{G_{nm}} \quad , \tag{2.1.1}$$

for $Y_1, \dots, Y_N \in \mathbb{C}$. Suppose $\mathbf{G} \in \mathrm{Mat}_{<2}(N)$ is non-negative and irreducible as in Theorem 2.1.1. Since a real and positive solution to (2.1.1) also solves (Y) and satisfies conditions 1–3 in Theorem 2.1.1 (for $\mathbf{a} = 0$), by Corollary 2.1.3 the constant Y-system *has a unique positive solution*. This extends a result of [NK, IIKKN], where symmetric and positive definite $\mathbf{G}$ were considered, as well as adjacency matrices of finite Dynkin diagrams.

ii) If $\mathbf{G}$ is the adjacency matrix of a finite Dynkin diagram, explicit trigonometric expressions for the unique positive solution to the constant Y-system (and more general versions thereof) are known or conjectured, see [Ki] and [KNS, Sec. 14].

iii) If $\mathbf{G}$ has spectral radius ≥ 2, it is shown in [RTV, Sec. 4] that the constant Y-system (2.1.1) does not possess a real positive solution at all. This shows that for $\mathbf{a} = 0$, the condition on the spectral radius in Theorems 2.1.1 and 2.1.2 is sharp.

Theorems 2.1.1 and 2.1.2 as well as Corollary 2.1.3 will be proved in Section 2.3. The proofs are based on an auxiliary result which we prove in section 2.2, as well as on Proposition 1.2.1.

2.2. Unique solution to a family of integral equations

In this section we give a criterion for functional equations of the form (TBA) in the introduction to have a unique solution. Specifically, we will prove a special case of Theorem 2.1.2 where we choose $\mathbf{C} = \frac{1}{2}\mathbf{G}$.

Proposition 2.2.1. *Let $\mathbf{G} \in \mathrm{Mat}_{<2}(N)$ be non-negative and irreducible, and let $\mathbf{a} \in BC_{-}(\mathbb{R}, \mathbb{R})^N$. Then the system of nonlinear integral equations*

$$\mathbf{f}(x) = \int_{-\infty}^{\infty} \Phi_{\frac{1}{2}\mathbf{G}}(x-y) \cdot \mathbf{G} \cdot \left(\log\left(e^{-\mathbf{a}(y)} + e^{\mathbf{f}(y)}\right) - \tfrac{1}{2}\mathbf{f}(y)\right) dy \tag{2.2.1}$$

has exactly one bounded continuous solution, $\mathbf{f}_\star \in BC(\mathbb{R}, \mathbb{R})^N$.

The proof will consist of verifying that the Banach Fixed Point Theorem can be applied and will be given in Section 2.2.3. In Sections 2.2.1 and 2.2.2 we lay the groundwork by discussing a type of integral equations called Hammerstein integral equations, and by applying the general results there to TBA-type equations.

After the proof, in Section 2.2.4 we comment on the special case that $\mathbf{G}$ is the adjacency matrix of a graph – the case most commonly considered in applications – and in Section 2.2.5 we look at the case $N = 1$ in more detail.

The only previous proof of Theorem 2.1.2 we know of concerns the case $N = 1$, $\mathbf{G} = \mathbf{C} = 1$ and $\mathbf{a} \sim \cosh(x)$, and can be found in [FKS]. Their argument also uses the Banach Fixed Point Theorem but is different from ours (we rely on being able to choose $\mathbf{C}$ different from $\mathbf{G}$) and we review it in Section 2.2.5.

2.2.1. Hammerstein integral equations as contractions

In this section we take $\mathbb{K}$ to stand for $\mathbb{R}$ or $\mathbb{C}$. We use the abbreviation $BC(\mathbb{R}) := BC(\mathbb{R}, \mathbb{K})$. Similarly, we write $BC(\mathbb{R}^m)^N$ for $BC(\mathbb{R}^m, \mathbb{K}^N)$, which we think of either as $\mathbb{K}^N$-valued functions, or as N-tuples of $\mathbb{K}$-valued functions.

Consider the nonlinear integral equation

$$f(x) = \int_{-\infty}^{\infty} K(x,y) L(y, f(y)) dy \ , \tag{2.2.2}$$

where $K : \mathbb{R} \times \mathbb{R} \to \mathbb{K}$ and $L : \mathbb{R} \times \mathbb{K} \to \mathbb{K}$ are some functions continuous in both arguments, and where it is understood that the integral is well-defined for $f(x)$ in some suitable class of functions on $\mathbb{R}$. Integral equations of this form are commonly referred to as *Hammerstein equations*, see e.g. [Kr, I.3] and [PM, Ch. 16]. A function $f(x)$ solves this equation if and only if it is a fixpoint of the corresponding integral operator

$$A[f](x) := \int_{-\infty}^{\infty} K(x,y)L(y,f(y))dy \; . \tag{2.2.3}$$

When does such a map have a unique fixpoint? We will try to bring Banach's Fixed Point Theorem to bear on this question, which we now briefly recall.

Definition 2.2.2. Let X be a metric space. A map $A : X \to X$ is called a *contraction* if there exists a positive real constant $\kappa < 1$ such that

$$d_X(A(x), A(y)) \leq \kappa \, d_X(x,y) \tag{2.2.4}$$

for all $x, y \in X$. If the condition is satisfied for $\kappa = 1$, then A is called *non-expansive*.

Theorem 2.2.3 (Banach)**.** *Let X be a complete metric space and $A : X \to X$ a contraction. Then A has a unique fixpoint $x_\star \in X$. Furthermore, for every $x_0 \in X$, the recursive sequence $x_n := A(x_{n-1})$ converges to $x_\star$.*

We now describe a general principle which facilitates the application of Banach's Theorem to integral operators of Hammerstein type (see e.g. [PM, 16.6-1, Thm. 3]). Suppose there exists a constant $\rho > 0$ such that

$$\int_{-\infty}^{\infty} |K(x,y)| \, dy \leq \rho \qquad \text{for all} \quad x \in \mathbb{R} \; . \tag{2.2.5}$$

Suppose also that L is Lipschitz continuous in the second variable, i.e. there exists a constant $\sigma > 0$ such that

$$|L(x,t) - L(x,s)| \leq \sigma \, |t - s| \qquad \text{for all} \quad x \in \mathbb{R}, \; s, t \in \mathbb{K} \; . \tag{2.2.6}$$

Provided $A[f]$ defines a map $BC(\mathbb{R}) \to BC(\mathbb{R})$, we can use this to compute for any $f, g \in BC(\mathbb{R})$:

$$\begin{aligned} \|A[f] - A[g]\|_\infty &= \sup_{x\in\mathbb{R}} \left| \int_{-\infty}^{\infty} K(x,t) \left(L(t,f(t)) - L(t,g(t)) \right) dt \right| \\ &\leq \sup_{x\in\mathbb{R}} \int_{-\infty}^{\infty} |K(x,t)| \, |L(t,f(t)) - L(t,g(t))| \, dt \\ &\leq \sigma \sup_{x\in\mathbb{R}} \int_{-\infty}^{\infty} |K(x,t)| \, |f(t) - g(t)| \, dt \;\leq\; \sigma\rho \, \|f - g\|_\infty \end{aligned} \tag{2.2.7}$$

If $\kappa := \sigma\rho < 1$, then $A[f]$ is a contraction with respect to the metric induced by the supremum norm $\|\cdot\|_\infty$. Recall that $BC(\mathbb{R})$ together with the norm $\|\cdot\|_\infty$ is a Banach space, and so the Banach Theorem 2.2.3 applies.

Consider now N coupled nonlinear integral equations of Hammerstein type:

$$\mathbf{f}(x) = \int_{-\infty}^{\infty} \mathbf{K}(x,y) \cdot \mathbf{L}(y, \mathbf{f}(y))\, dy \ , \tag{2.2.8}$$

where $\mathbf{K} : \mathbb{R} \times \mathbb{R} \to \mathrm{Mat}(N, \mathbb{K})$ and $\mathbf{L} : \mathbb{R} \times \mathbb{K}^N \to \mathbb{K}^N$ are continuous in both arguments. Our arguments depend crucially on the right choice of norm for the functions $\mathbf{f} : \mathbb{R} \to \mathbb{K}^N$.

For $1 \le p \le \infty$, we equip the space $BC(\mathbb{R})^N$ with the norm $\|\cdot\|_{\infty_p}$ given by

$$\|\mathbf{f}\|_{\infty_p} := \sup_{x\in\mathbb{R}} \|\mathbf{f}(x)\|_p = \sup_{x\in\mathbb{R}} \left(\sum_{i=1}^{N} |f_i(x)|^p \right)^{\frac{1}{p}} \ . \tag{2.2.9}$$

The normed space $\left(BC(\mathbb{R})^N, \|\cdot\|_{\infty_p}\right)$ is a Banach space by the following standard lemma, which we state without proof.

Lemma 2.2.4. *Let X be a metric space and $(Y, \|\cdot\|_Y)$ a Banach space. Then also $(BC(X,Y), \|\cdot\|_\infty)$ is a Banach space.*

We now give the N-component version of the general principle outlined above. The proof is the same, just with heavier notation.

Lemma 2.2.5. *Let $\mathbf{K} \in \mathrm{Mat}(N, BC(\mathbb{R}^2))$ and $\mathbf{L} : \mathbb{R} \times \mathbb{K}^N \to \mathbb{K}^N$ a function. Suppose that*

- $\mathbf{K}$ *has bounded integrals in the second variable, in the sense that*

$$\rho_{ij} := \sup_{x\in\mathbb{R}} \int_{-\infty}^{\infty} |K_{ij}(x,y)|\, dy \ < \ \infty \qquad , \quad i,j = 1,...,N \ . \tag{2.2.10}$$

- *all the components of $\mathbf{L}$ are Lipschitz continuous in the second variable in the sense that there are constants $\sigma_j \ge 0$, $j = 1, \dots, N$, such that*

$$|L_j(y, \mathbf{v}) - L_j(y, \mathbf{w})| \le \sigma_j \|\mathbf{v} - \mathbf{w}\|_p \quad \textit{for all} \ \ \mathbf{v}, \mathbf{w} \in \mathbb{K}^N, \ y \in \mathbb{R} \ . \tag{2.2.11}$$

- *the matrix $\boldsymbol{\rho} = (\rho_{ij})_{i,j=1,\dots,N}$ and the vector $\boldsymbol{\sigma} = (\sigma_i)_{i=1,\dots,N}$ are such that $\kappa := \|\boldsymbol{\rho} \cdot \boldsymbol{\sigma}\|_p < 1$.*

Suppose that the following integral operator defines a map $\mathbf{A} : BC(\mathbb{R})^N \to BC(\mathbb{R})^N$,

$$\mathbf{A}[\mathbf{f}](x) := \int_{-\infty}^{\infty} \mathbf{K}(x,y) \cdot \mathbf{L}(y, \mathbf{f}(y))\, dy \ . \tag{2.2.12}$$

Then $\mathbf{A}$ is a contraction on $\left(BC(\mathbb{R})^N, \|\cdot\|_{\infty_p}\right)$.

Proof. Let $\mathbf{f}, \mathbf{g} \in BC(\mathbb{R})^N$. Then

$$\begin{aligned}
\left(\|\mathbf{A}[\mathbf{f}] - \mathbf{A}[\mathbf{g}]\|_{\infty_p}\right)^p &= \sup_{x\in\mathbb{R}} \sum_{i=1}^{N} |A_i[\mathbf{f}] - A_i[\mathbf{g}]|^p \\
&= \sup_{x\in\mathbb{R}} \sum_{i=1}^{N} \left| \sum_{j=1}^{N} \int_{-\infty}^{\infty} K_{ij}(x,y)\left(L_j(y,\mathbf{f}(y)) - L_j(y,\mathbf{g}(y))\right) dy \right|^p \\
&\leq \sup_{x\in\mathbb{R}} \sum_{i=1}^{N} \left| \sum_{j=1}^{N} \int_{-\infty}^{\infty} |K_{ij}(x,y)| \left|L_j(y,\mathbf{f}(y)) - L_j(y,\mathbf{g}(y))\right| dy \right|^p \\
&\leq \left(\|\mathbf{f} - \mathbf{g}\|_{\infty_p}\right)^p \sum_{i=1}^{N} \left| \sum_{j=1}^{N} \rho_{ij}\sigma_j \right|^p \\
&= \left(\|\mathbf{f} - \mathbf{g}\|_{\infty_p}\right)^p \left(\|\boldsymbol{\rho}\cdot\boldsymbol{\sigma}\|_p\right)^p = \left(\kappa \|\mathbf{f} - \mathbf{g}\|_{\infty_p}\right)^p .
\end{aligned} \tag{2.2.13}$$

Since $\kappa < 1$, $\mathbf{A}$ is a contraction. □

2.2.2. Unique solution to TBA-type equations

In this section we specialise the results of the previous section to integral equations of the form (TBA). We will restrict ourselves to the case $\mathbb{K} = \mathbb{R}$.

Let us call a function $f : \mathbb{R}^N \to \mathbb{R}$ *p-Lipschitz-continuous*, if it satisfies the Lipschitz condition with respect to $\|\cdot\|_p$. In fact, this is a slightly redundant denomination: equivalence of all p-norms ensures that f is Lipschitz continuous either with respect to all or none of the p-norms. However, the optimal Lipschitz constants differ, which is important in view of the third condition in Lemma 2.2.5. For differentiable functions that satisfy a Lipschitz condition, we can characterise the p-Lipschitz constant in terms of the gradient as follows:

Lemma 2.2.6. *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a continuous function whose gradient $\nabla f : \mathbb{R}^N \to \mathbb{R}^N$ is also a continuous function, and $1 \leq p \leq \infty$. Then the following inequality holds:*

$$|f(\mathbf{v}) - f(\mathbf{w})| \leq \left(\sup_{\mathbf{u}\in\mathbb{R}^N} \|\nabla f(\mathbf{u})\|_q\right) \|\mathbf{v} - \mathbf{w}\|_p \qquad \textit{for all} \;\; \mathbf{v}, \mathbf{w} \in \mathbb{R}^N , \tag{2.2.14}$$

where q is defined via the relation $\frac{1}{p} + \frac{1}{q} = 1$. In particular, f is Lipschitz continuous if the gradient is bounded.

Proof. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^N$. By the mean value theorem there exists a $t \in [0,1]$ such that

$$|f(\mathbf{v}) - f(\mathbf{w})| = |\nabla f(\mathbf{x}) \cdot (\mathbf{v} - \mathbf{w})| \tag{2.2.15}$$

for $\mathbf{x} = \mathbf{v} + t(\mathbf{w} - \mathbf{v})$. Applying first the triangle and then the Hölder inequality to the right hand side, one obtains

$$\begin{aligned}
|f(\mathbf{v}) - f(\mathbf{w})| &\leq \sum_{i=1}^{N} |\nabla_i f(\mathbf{x})(v_i - w_i)| \\
&\leq \|\nabla f(\mathbf{x})\|_q \|\mathbf{v} - \mathbf{w}\|_p \\
&\leq \left(\sup_{\mathbf{u}\in\mathbb{R}^N} \|\nabla f(\mathbf{u})\|_q \right) \|\mathbf{v} - \mathbf{w}\|_p \ ,
\end{aligned} \tag{2.2.16}$$

which means that $\sup_{\mathbf{u}\in\mathbb{R}^N} \|\nabla f(\mathbf{u})\|_q$, if bounded, is a p-Lipschitz constant for f. It is easy to see that it is the optimal p-Lipschitz constant. □

Lemma 2.2.7. *Let $A \geq 0$ and $a, b, c \in \mathbb{R}$. The function $L : \mathbb{R} \to \mathbb{R}$ defined by*

$$L(x) = a \cdot \log\left(A + e^{bx}\right) - cx \tag{2.2.17}$$

is Lipschitz continuous with Lipschitz constant $\sigma_L = \max(|c|, |ab - c|)$.

Proof. If any one of a, b, A is zero, the statement is clear. Suppose $a, b, A \neq 0$. The derivative

$$\frac{d}{dx}L(x) = \frac{ab}{Ae^{-bx} + 1} - c \tag{2.2.18}$$

interpolates between $-c$ (for $bx \to -\infty$) and $ab - c$ (for $bx \to \infty$). It is also a monotonous function, since

$$\frac{d^2}{dx^2}L(x) = a \cdot \frac{Ab^2 e^{bx}}{(A + e^{bx})^2} \ , \tag{2.2.19}$$

which is ≥ 0 for $a \geq 0$ and ≤ 0 for $a \leq 0$. It thus follows that $\left|\frac{d}{dx}L(x)\right| \leq \max(|c|, |ab - c|)$. Lemma 2.2.6 completes the proof. □

Proposition 2.2.8. *Let $1 \leq p, q \leq \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$ holds. For $i, j = 1, ..., N$, let $\phi_{ij} \in BC(\mathbb{R}, \mathbb{R}) \cap L_1(\mathbb{R})$, $a_j \in BC_-(\mathbb{R}, \mathbb{R})$ and $G_{ij}, C_{ij} \in \mathbb{R}$, $w_i \in \mathbb{R}_{>0}$. Furthermore, set $M_{ij} := \max\left(|C_{ij}|, |G_{ij} - C_{ij}|\right)$ and define*

$$\sigma_i := \left(\sum_{j=1}^{N} (M_{ij} w_j)^q \right)^{\frac{1}{q}} \ , \qquad \rho_{ij} := \frac{1}{w_i} \int_{-\infty}^{\infty} |\phi_{ij}(y)| \, dy \ . \tag{2.2.20}$$

If $\kappa := \|\boldsymbol{\rho} \cdot \boldsymbol{\sigma}\|_p < 1$, then the system of nonlinear integral equations given by

$$f_i(x) = \sum_{j=1}^{N} \int_{-\infty}^{\infty} \phi_{ij}(x - y) \sum_{k=1}^{N} \left[G_{jk} \log\left(e^{-a_k(y)} + e^{f_k(y)}\right) - C_{jk} f_k(y) \right] dy \tag{2.2.21}$$

has exactly one bounded continuous solution, $(f_{\star,1}, \ldots, f_{\star,N}) \in BC(\mathbb{R}, \mathbb{R})^N$.

Proof. We start by rewriting (2.2.21) in terms of the rescaled functions $g_i(x) = f_i(x)/w_i$:

$$g_i(x) = \sum_{j=1}^{N} \int_{-\infty}^{\infty} \frac{1}{w_i}\phi_{ij}(x-y) \sum_{k=1}^{N} \left[G_{jk} \log\left(e^{-a_k(y)} + e^{w_k g_k(y)}\right) - C_{jk} w_k g_k(y)\right] dy \quad (2.2.22)$$

The lower bound on the $a_k(y)$ ensures that the corresponding integral operator is a well-defined map $BC(\mathbb{R},\mathbb{R})^N \to BC(\mathbb{R},\mathbb{R})^N$. Consider the functions

$$L_i(y, \mathbf{v}) = \sum_{k=1}^{N} \left[G_{ik} \log\left(e^{-a_k(y)} + e^{w_k v_k}\right) - C_{ik} w_k v_k\right] \ . \quad (2.2.23)$$

The individual summands, as functions in v_k for fixed y, are of the form considered in Lemma 2.2.7, from which one concludes

$$\left|\frac{\partial}{\partial v_j} L_i(y, \mathbf{v})\right| \le w_j \ \max\left(|C_{ij}|, |G_{ij} - C_{ij}|\right) \ . \quad (2.2.24)$$

According to Lemma 2.2.6, the functions $L_i(y, \mathbf{v})$ are thus p-Lipschitz continuous in the second variable, with Lipschitz constants

$$\begin{aligned} \sigma_i &= \sup_{\substack{\mathbf{v}\in\mathbb{R}^N \\ y\in\mathbb{R}}} \left\|\nabla_{\mathbf{v}} L_i(y, \mathbf{v})\right\|_q = \sup_{\substack{\mathbf{v}\in\mathbb{R}^N \\ y\in\mathbb{R}}} \left(\sum_{j=1}^{N} \left|\frac{\partial}{\partial v_j} L_i(y, \mathbf{v})\right|^q\right)^{\frac{1}{q}} \\ &= \left(\sum_{j=1}^{N} \sup_{\substack{\mathbf{v}\in\mathbb{R}^N \\ y\in\mathbb{R}}} \left|\frac{\partial}{\partial v_j} L_i(y, \mathbf{v})\right|^q\right)^{\frac{1}{q}} = \left(\sum_{j=1}^{N} \left(w_j \ \max\left(|C_{ij}|, |G_{ij} - C_{ij}|\right)\right)^q\right)^{\frac{1}{q}} \ . \end{aligned} \quad (2.2.25)$$

An application of Lemma 2.2.5 completes the proof. □

We note that the bound on κ used in Proposition 2.2.8 is actually independent of the functions a_j.

2.2.3. Proof of Proposition 2.2.1

The proof makes use of the Perron-Frobenius Theorem in the following form (see e.g. [BH, Theorem 2.2.1]):

Theorem 2.2.9 (Perron-Frobenius)**.** *Let* $\mathbf{A}$ *be a non-negative real-valued irreducible* $N \times N$ *matrix. Then the largest eigenvalue* λ_{PF} *of* $\mathbf{A}$ *is real and has geometric and algebraic multiplicity 1. Its associated eigenvector can be chosen to have strictly positive components and is the only eigenvector with that property.*

We now turn to the proof of Proposition 2.2.1.

Set $p = \infty$ and $q = 1$. In terms of Proposition 2.2.8 we have $\mathbf{C} = \frac{1}{2}\mathbf{G}$, so that $\mathbf{M} = \frac{1}{2}\mathbf{G}$. By the Perron-Frobenius theorem, $\mathbf{G}$ has an eigenvector $\mathbf{w}$ with strictly positive components $w_i > 0$ associated to its largest eigenvalue λ_{PF}. With this choice of w_i the constant vector $\boldsymbol{\sigma}$ in Proposition 2.2.8 is given by

$$\boldsymbol{\sigma} = \tfrac{1}{2}\mathbf{G}\mathbf{w} = \tfrac{1}{2}\lambda_{\mathrm{PF}}\mathbf{w} \ . \tag{2.2.26}$$

Let us abbreviate $\Phi(x) := \Phi_{\frac{1}{2}\mathbf{G}}(x)$ with components $\phi_{ij}(x)$. Due to Lemma 1.2.17 we know that $|\phi_{ij}(x)| = \phi_{ij}(x)$ for all $x \in \mathbb{R}$ and $i, j = 1, \dots, N$. Combining this with (1.2.62) one computes the matrix $\boldsymbol{\rho}$ in Proposition 2.2.8 to be

$$\rho_{ij} = \frac{1}{w_i}\int_{-\infty}^{\infty} \phi_{ij}(y)dy = \frac{1}{w_i}\big[(2\,\mathbf{1} - \tfrac{1}{2}\mathbf{G})^{-1}\big]_{ij} \ . \tag{2.2.27}$$

Since $\mathbf{w}$ is an eigenvector of $(2\,\mathbf{1} - \frac{1}{2}\mathbf{G})^{-1}$ we find

$$\boldsymbol{\rho}\cdot\mathbf{w} = \frac{1}{2 - \frac{\lambda_{\mathrm{PF}}}{2}} \cdot (1, 1, \dots, 1) \ . \tag{2.2.28}$$

Hence the contraction constant κ in Proposition 2.2.8 is given by

$$\kappa = \|\boldsymbol{\rho}\cdot\boldsymbol{\sigma}\|_\infty = \frac{\lambda_{\mathrm{PF}}}{2}\max_{i=1,\dots,N}\left|\sum_{j=1}^{N}\rho_{ij}w_j\right| = \frac{\lambda_{\mathrm{PF}}}{2}\left|\frac{1}{2 - \frac{\lambda_{\mathrm{PF}}}{2}}\right| = \left|\frac{\lambda_{\mathrm{PF}}}{4 - \lambda_{\mathrm{PF}}}\right| \ . \tag{2.2.29}$$

It follows that $\kappa < 1$ if and only if $\lambda_{\mathrm{PF}} < 2$.

By Proposition 2.2.8 there is a unique solution to (2.2.1), completing the proof of Proposition 2.2.1.

2.2.4. Adjacency matrices of graphs

In Proposition 2.2.1, the matrix $\mathbf{G}$ may have non-negative real entries. This is in itself interesting because it makes the solution $\mathbf{f}_\star$ depend on an additional set of continuous parameters. However, in the application to integrable quantum field theory, $\mathbf{G}$ is usually the *adjacency matrix* of some (suitably generalised) graph. Irreducibility is then equivalent to the corresponding generalised graph being strongly connected.

If $\mathbf{G}$ is symmetric and has entries $\{0, 1\}$ with zero on the diagonal, then by definition it is the adjacency matrix of a simple (undirected, unweighted) graph whose nodes i and j are connected if and only if $G_{ij} = 1$. In this case, strongly connected and connected are equivalent. The only connected simple graphs with $\lambda_{\mathrm{PF}} < 2$ are the graphs associated to the ADE Dynkin diagrams (while their affine versions are the sole examples satisfying $\lambda_{\mathrm{PF}} = 2$), and their adjacency matrix is diagonalisable over $\mathbb{R}$ [BH, Thm. 3.1.3].

Consider now generalised graphs. If we allow for loops ($G_{ii} \neq 0$, edges connecting a node to itself) and multiple edges between the same nodes (where the entry G_{ij} is the number of edges connecting nodes i and j) we additionally get the tadpole $T_N = A_{2N}/\mathbb{Z}_2$ (defined as the adjacency matrix of A_N with additional entry $G_{11} = 1$). If, moreover, the symmetry requirement is dropped ($G_{ij} \neq G_{ji}$), we may still associate to $\mathbf{G}$ a mixed multigraph (with some edges now being replaced by arrows). The $BCFG$ Dynkin diagrams provide examples of this type for which $\lambda_{\mathrm{PF}} < 2$ hold. However, in contrast with the undirected (symmetric) case, they are not exhaustive at all. For instance, a directed graph with $\lambda_{\mathrm{PF}} > 2$ can be turned into a new directed graph with $\lambda_{\mathrm{PF}} < 2$ by subdividing all its edges often enough. If $\mathbf{G}$ is not even assumed to be integer-valued, then every non-example becomes an example after appropriate rescaling.

To summarise, we have the following special cases in which Proposition 2.2.1 applies:

Corollary 2.2.10. *If $\mathbf{G}$ is the adjacency matrix of a finite Dynkin diagram (A_N, B_N, C_N, D_N, E_6, E_7, E_8, F_4 or G_2) or of the tadpole T_N, then the system* (2.2.1) *has exactly one bounded continuous solution.*

2.2.5. The case of a single TBA equation

Consider a single integral equation of TBA-type ($N = 1$ with $G_{11} = g \in (0, 2)$, $C_{11} = c \in (-2, 2)$):

$$f(x) = \int_{-\infty}^{\infty} \phi_c(x-y) \left[g \log\left(e^{-a(y)} + e^{f(y)}\right) - cf(y) \right] dy \tag{2.2.30}$$

This case is instructive: we do not have to worry about the choice of p and q, nor does a rescaling $f(x) \to f(x)/w$ influence the contraction constant. We compute the quantities in Proposition 2.2.8 to be

$$\begin{aligned} \rho &= \int_{-\infty}^{\infty} \phi_c(y) dy = \frac{1}{2-c} \quad , \qquad \sigma = M = \max(|c|, |g-c|) \ , \\ \kappa(c) &= \rho\sigma = \frac{\max(|c|, |g-c|)}{2-c} \ . \end{aligned} \tag{2.2.31}$$

The most important case is $g = 1$ (this case arises for example in the Yang-Lee model). The contraction constant for this case is shown in Figure 2.1.

For $c < 1$, our estimate guarantees a contraction. The canonical TBA equation with $c = g = 1$ corresponds to the marginal case $\kappa = 1$, whereas the universal TBA with $c = 0$ yields $\kappa = \frac{1}{2}$. The “sweet spot” is $c = \frac{1}{2} = \frac{1}{2}g$ where our estimate for the contraction constant attains its minimum $\kappa = \frac{1}{3}$.

If we chose a different value $g \in (0, 2)$, the sweet spot shifts, but the overall picture remains the same (the region of assured contraction is $c < 1$). However, for $g \geq 2$, κ is larger or equal to one everywhere, and no region of assured contraction exists.

Remark 2.2.11. It is noteworthy that κ has virtually no practical bearing on the speed of convergence of the iterative numerical solution of equation (2.2.30). We solved it numerically for $g = 1$, $a(x) = r\cosh(x)$ (the massive Yang-Lee model in volume r) for different values of $c \in [0, 1]$ and $r \in (0, 1]$, and the speed of convergence (as measured by the number

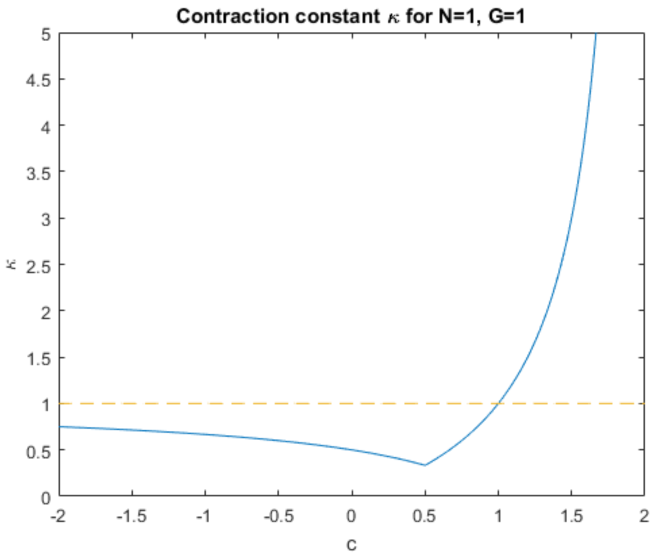


Figure 2.1.: Contraction constant for a single TBA equation at different values of c.

of iterations required to obtain a certain accuracy) increases almost linearly in c, instead of being governed by κ. If κ were the optimal contraction constant, we would instead expect to see the fastest convergence for $c = \frac{1}{2}$.

We now describe in more detail the proof given in [FKS, Sec. 5], which concerns the case $g = c = 1$ and $a(x) = r\cosh(x)$ with $r > 0$, in the above example. As described above, our bound on κ in this case is 1, so that Proposition 2.2.8 does not apply. We circumvent this by using $c = \frac{1}{2}$ and proving (in Section 2.3, see Theorem 2.1.2 from the introduction) that the fixed point is independent of c and unique in the range $c \in (-2, 2)$. The argument of [FKS] also uses the Banach Theorem but proceeds differently.

Namely, they exhaust the space of bounded continuous functions by subspaces with specific bounds, $BC(\mathbb{R}, \mathbb{R}) = \bigcup_{q\in[e^{-r},1)} D_{q,r}$, where

$$D_{q,r} := \left\{ f \in BC(\mathbb{R}, \mathbb{R}) \,\middle|\, \|f\|_\infty \leq \log(\tfrac{q}{1-q}) + r \right\} . \tag{2.2.32}$$

Using convexity of $D_{q,r}$ it is shown that for $q \geq e^{-r}$ the corresponding integral operator maps $D_{q,r}$ to itself, and that the associated contraction constant is $\kappa = q$. Hence, as expected one obtains $\kappa \to 1$ when $q \to 1$ (so that $\log(\frac{q}{1-q}) \to \infty$), but on each $D_{q,r}$ we have $\kappa < 1$ and thus a proper contraction. This implies a unique solution on the whole space $BC(\mathbb{R}, \mathbb{R})$.

It is also stated in [FKS] that the generalisation to higher N should be straightforward. This is less clear to us. In the examples from finite Dynkin diagrams we need all the extra freedom we introduce in Proposition 2.2.8 in an optimal way in order to press our bound κ under 1. If we take, for example, $\mathbf{G} = \mathbf{C}$ to be the adjacency matrix of the A_N Dynkin diagram, the quantities in Proposition 2.2.8 take the values

$$M_{ij} = G_{ij} \quad , \quad \rho_{ij} = \left[(A_N)^{-1}\right]_{ij} \quad , \quad \sigma_i = \begin{cases} 1 & ; i = 1, N \\ 2^{1/q} & ; i = 2, \dots, N-1 \end{cases} \quad , \tag{2.2.33}$$

where A_N denotes the Cartan matrix $2\mathbf{1} - \mathbf{G}$. With some more work, from this one can estimate that for N large enough one has $\|\boldsymbol{\rho} \cdot \boldsymbol{\sigma}\|_p > \frac{1}{8}N^2$ (independent of p and q). Hence, as opposed to what happened at $N = 1$, for larger N the bound κ is not 1 but grows at least quadratically with N. It is not obvious to us how to obtain drastically better estimates by choosing subsets analogous to $D_{q,r}$ of $BC(\mathbb{R}, \mathbb{R}^N)$.

However, it might be possible – at least in the massive case, cf. Remark 2.1.4 iii) – that one can combine the freedom to choose $\mathbf{C}$ and w_i that we introduce with the method of [FKS] to extend our results to the case of spectral radius 2 or to infinitely many coupled TBA equations (such as the $N \to \infty$ limit of classical finite Dynkin diagrams, where the spectral radius approaches 2). The idea of choosing subsets $D_{q,r}$ as above might push κ strictly below 1. We hope to return to these points in the future.

2.3. Uniqueness of solution to the Y-system

For $\mathbf{G} \in \mathrm{Mat}(N, \mathbb{R})$, $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$ and $\mathbf{a} \in BC_-(\mathbb{R}, \mathbb{R})^N$ we define the map $\mathbf{L}_\mathbf{C} : BC(\mathbb{R}, \mathbb{R})^N \to BC(\mathbb{R}, \mathbb{R})^N$ as (recall the convention in (1.1.4))

$$\mathbf{L}_\mathbf{C}[\mathbf{f}](x) := \mathbf{G} \cdot \log\left(e^{-\mathbf{a}(x)} + e^{\mathbf{f}(x)}\right) - \mathbf{C} \cdot \mathbf{f}(x) \ . \tag{2.3.1}$$

With this notation, the TBA equation (TBA) reads

$$\mathbf{f}(x) = \left(\Phi_\mathbf{C} \star \mathbf{L}_\mathbf{C}[\mathbf{f}]\right)(x) \ . \tag{2.3.2}$$

From Proposition 2.2.1 we know that (2.3.2) has a unique bounded continuous solution for one special choice of $\mathbf{C}$, namely $\mathbf{C} = \frac{1}{2}\mathbf{G}$. In this section we will apply the results of section 1.2 in order to translate that statement to other choices of $\mathbf{C}$, as well as to the associated Y-system. In particular, we will prove Theorems 2.1.1, 2.1.2 and Corollary 2.1.3 from the introduction.

2.3.1. Independence of the choice of C

In this subsection we fix

$$\mathbf{G} \in \mathrm{Mat}(N, \mathbb{R}) \quad , \qquad \mathbf{C} \in \mathrm{Mat}_{<2}(N) \quad , \qquad \mathbf{a} \in BC_-(\mathbb{R}, \mathbb{R})^N \ . \tag{2.3.3}$$

We stress that for the moment, we make no further assumptions on $\mathbf{G}$ (as opposed to Theorems 2.1.1 and 2.1.2).

We will later need to apply Proposition 1.2.1 to (2.3.2). To this end we now provide a criterion for the components of $\mathbf{L}_\mathbf{C}[\mathbf{f}]$ to be Hölder continuous.

Lemma 2.3.1. *Let* $\mathbf{f} \in BC(\mathbb{R},\mathbb{R})^N$*. If the components of* $\mathbf{f}$ *and of* $e^{-\mathbf{a}}$ *are Hölder continuous, then the components of* $\mathbf{L_C}[\mathbf{f}]$ *are Hölder continuous.*

Proof. It is easy to see that the composition of Hölder continuous functions is again Hölder continuous, as is the sum of bounded Hölder continuous functions. Therefore, and since $x \mapsto e^x$ is Hölder continuous on any compact subset of $\mathbb{R}$ – in particular on the images of the bounded functions f_m – the functions $x \mapsto u_m(x) := e^{-a_m(x)} + e^{f_m(x)}$ are Hölder continuous. The bounds on a_m and f_m ensure that the image of u_m is contained in some interval $[x_0, x_1]$ with $x_0, x_1 > 0$. But $x \mapsto \log(x)$ is Hölder continuous on $[x_0, x_1]$, and so the functions $x \mapsto l_m(x) := \log(u_m(x))$ are bounded and Hölder continuous. From this it follows that the components of $\mathbf{L_C}[\mathbf{f}](x)$,

$$x \longmapsto \left[\, \mathbf{L_C}[\mathbf{f}]\,\right]_n(x) = \sum_{m=1}^{N} \left(G_{nm} l_m(x) - C_{nm} f_m(x)\right) \ , \tag{2.3.4}$$

are Hölder continuous. □

We are careful not to make too strong assumptions on $\mathbf{a}$ here, namely we do not require the components of $\mathbf{a}$ to be Hölder continuous. For instance, the relevant example $a_m(x) \sim e^{\gamma x/s}$ from (1.1.2) is only locally Hölder continuous. Meanwhile, the Hölder condition on $\mathbf{f}$ is, in fact, obtained from the TBA equation for free:

Lemma 2.3.2. *Suppose* $\mathbf{f} \in BC(\mathbb{R},\mathbb{R})^N$ *is a solution of the TBA equation* (2.3.2)*. Then the components of* $\mathbf{f}$ *are Lipschitz continuous.*

Proof. Using Lemma 1.2.16 ii) one quickly verifies that one can write the components of $\mathbf{f}$ as real linear combinations of the form

$$f_n(x) = \sum_{d,m} c^{(n)}_{d,m} F_d\big[\,[\,\mathbf{L_C}[\mathbf{f}]\,]_m\,\big](x) \ , \tag{2.3.5}$$

where $F_d[-]$ is the convolution functional defined in (1.2.67), see also (1.2.85). Lemma 1.2.22 then implies that the derivatives $\frac{d}{dx} f_n(x)$ are bounded. This shows the claim (cf. Lemma 2.2.6). □

Now we are in the position to apply Proposition 1.2.1 to (2.3.2). The $\mathbf{C}$-independence will boil down to the following simple observation on the functional equation (1.2.3): the $\mathbf{C}$-dependence on the left and right hand side of

$$\mathbf{f}(x + is) + \mathbf{f}(x - is) - \mathbf{C} \cdot \mathbf{f}(x) = \mathbf{L_C}[\mathbf{f}](x) \tag{2.3.6}$$

simply cancels, see (2.3.1). Thus (2.3.6) is in particular equivalent to

$$\mathbf{f}(x + is) + \mathbf{f}(x - is) = \mathbf{L_0}[\mathbf{f}](x) \tag{2.3.7}$$

Proposition 2.3.3. *Suppose that the components of* $e^{-\mathbf{a}}$ *are Hölder continuous and that there exists* $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$*, such that the TBA equation* (2.3.2) *has a unique solution* $\mathbf{f}_\star$ *in* $BC(\mathbb{R},\mathbb{R})^N$*. Then:*

i) $\mathbf{f}_\star$ *is real analytic and can be continued to a function in* $\mathcal{BA}(\mathbb{S}_s)^N$*, which we also denote by* $\mathbf{f}_\star$*. It is the unique solution to the functional equation*

$$\mathbf{f}(x+is)+\mathbf{f}(x-is)=\mathbf{L_0}[\mathbf{f}](x) \quad , \qquad x\in\mathbb{R} \ , \tag{2.3.8}$$

in the space $\mathcal{BA}(\mathbb{S}_s)^N$ *which also satisfies* $\mathbf{f}(\mathbb{R})\subset\mathbb{R}^N$*.*

ii) For any $\mathbf{C}'\in\mathrm{Mat}_{<2}(N)$*,* $\mathbf{f}_\star$ *is the unique solution to the TBA equation*

$$\mathbf{f}(x)=\left(\Phi_{\mathbf{C}'}\star\mathbf{L}_{\mathbf{C}'}[\mathbf{f}]\right)(x) \tag{2.3.9}$$

in the space $BC(\mathbb{R},\mathbb{R})^N$*.*

Proof. i) By definition, the components of $\mathbf{g}_\star(x):=\mathbf{L_C}[\mathbf{f}_\star](x)$ are bounded real functions. Moreover, Lemma 2.3.2 ensures that they are Lipschitz continuous, so in particular Hölder continuous. It follows from direction $2\Rightarrow 1$ in Proposition 1.2.1 that $\mathbf{f}_\star\in\mathcal{BA}(\mathbb{S}_s)^N$, and that $\mathbf{f}_\star$ satisfies the functional relation

$$\mathbf{f}_\star(x+is)+\mathbf{f}_\star(x-is)-\mathbf{C}\cdot\mathbf{f}_\star(x)=\mathbf{L_C}[\mathbf{f}_\star](x) \tag{2.3.10}$$

for all $x\in\mathbb{R}$, which, as we just said, is equivalent to (2.3.8).

Now suppose there is another solution $\mathbf{f}'_\star\in\mathcal{BA}(\mathbb{S}_s)^N$ to (2.3.8), or, equivalently, (2.3.10), which satisfies $\mathbf{f}'_\star(\mathbb{R})\subset\mathbb{R}^N$. By direction $1\Rightarrow 2$ of Proposition 1.2.1, the restriction $\mathbf{f}'_\star\big|_\mathbb{R}$ is also a solution to the TBA equation (2.3.2). By our uniqueness assumption, we must have $\mathbf{f}'_\star\big|_\mathbb{R}=\mathbf{f}_\star\big|_\mathbb{R}$, and by uniqueness of the analytic continuation also $\mathbf{f}'_\star=\mathbf{f}_\star$ on $\overline{\mathbb{S}}_s$.

ii) For any choice of $\mathbf{C}'\in\mathrm{Mat}_{<2}(N)$, (2.3.10) can be rewritten as

$$\mathbf{f}_\star(x+is)+\mathbf{f}_\star(x-is)-\mathbf{C}'\cdot\mathbf{f}_\star(x)=\mathbf{L}_{\mathbf{C}'}[\mathbf{f}_\star](x) \ . \tag{2.3.11}$$

Direction $1\Rightarrow 2$ of Proposition 1.2.1 shows that $\mathbf{f}_\star$ satisfies (2.3.9). Suppose $\mathbf{f}'_\star\in BC(\mathbb{R},\mathbb{R})^N$ is another solution to (2.3.9). Then by Lemma 2.3.2, $\mathbf{f}'_\star$ is Hölder continuous, and by direction $2\Rightarrow 1$ of Proposition 1.2.1 it satisfies (2.3.11). But (2.3.11) is equivalent to (2.3.8), whose solution is unique and equal to $\mathbf{f}_\star$ by part i). □

2.3.2. Proofs of Theorems 2.1.1, 2.1.2 and of Corollary 2.1.3

We now turn to the proof of the main results of this chapter. We start with part of Theorem 2.1.2, as this is used in the proof of Theorem 2.1.1. Then we show Theorem 2.1.1 and subsequently the missing part of Theorem 2.1.2. Finally, we give the proof of Corollary 2.1.3.

Proof of Theorem 2.1.2, less uniqueness in part ii)

Part i): First note that the assumptions in Proposition 2.2.1 are satisfied. Thus there exists a unique solution $\mathbf{f}_\star$ to (TBA) in $BC(\mathbb{R},\mathbb{R})^N$ for the specific choice $\mathbf{G}=\frac{1}{2}\mathbf{C}$. Therefore, the conditions of Proposition 2.3.3 are satisfied and part ii) of that proposition establishes existence and uniqueness of a solution $\mathbf{f}_\star \in BC(\mathbb{R},\mathbb{R})^N$ to (TBA) for any choice of $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$, as well as $\mathbf{C}$-independence of that solution.

Part ii) (without uniqueness): By Proposition 2.3.3 i), $\mathbf{f}_\star$ can be analytically continued to a function in $\mathcal{BA}(\mathbb{S}_s)^N$, which we will also denote by $\mathbf{f}_\star$. We define

$$\mathbf{Y}(z) := \exp\big(\mathbf{a}(z) + \mathbf{f}_\star(z)\big) \ . \tag{2.3.12}$$

Let us denote the components of $\mathbf{f}_\star$ and $\mathbf{Y}$ by $f_{\star,n}$ and Y_n, respectively. Note that $Y_n \in \mathcal{A}(\mathbb{S}_s)$.

It is immediate that the Y_n satisfy properties 1–3 of Theorem 2.1.1 (for property 1 note that $f_{\star,n}$ and a_n are real-valued on the real axis). Furthermore, as a consequence of the functional relations (1.1.1) and (2.3.8) for $\mathbf{a}$ and $\mathbf{f}_\star$ respectively, the Y_n solve (Y):

$$\begin{aligned} Y_n(x+is)Y_n(x-is) &= e^{a_n(x+is)+a_n(x-is)} e^{f_{\star,n}(x+is)+f_{\star,n}(x-is)} \\ &= e^{\sum_m G_{nm} a_m(x)} e^{\sum_m G_{nm} \log\left(e^{-a_m(x)} + e^{f_{\star,m}(x)}\right)} \\ &= \prod_m e^{G_{nm} a_m(x)} \left(e^{-a_m(x)} + e^{f_{\star,m}(x)}\right)^{G_{nm}} \\ &= \prod_m \left(1 + Y_m(x)\right)^{G_{nm}} \ . \end{aligned} \tag{2.3.13}$$

This proves that $\mathbf{Y}$ is a solution to (Y) which lies in $\mathcal{A}(\mathbb{S}_s)^N$ and satisfies properties 1–3 in Theorem 2.1.1. It remains to show that $\mathbf{Y}$ is the unique such solution. This will be done below as an immediate consequence of Theorem 2.1.1, whose proof we turn to now.

Proof of Theorem 2.1.1

Existence of a $\mathbf{Y} \in \mathcal{A}(\mathbb{S}_s)^N$ which solves (Y) and satisfies properties 1–3 has just been proven above. The solution $\mathbf{Y}$ is given via (2.3.12) in terms of the valid asymptotics $\mathbf{a}$ and the unique solution $\mathbf{f}_\star$ to (TBA) obtained in Theorem 2.1.2 i). It remains to show uniqueness of $\mathbf{Y}$.

Suppose there is another function $\mathbf{Y}' \in \mathcal{A}(\mathbb{S}_s)^N$ with the properties 1–3. Since $\mathbb{S}_s$ is a simply connected domain and the components $Y'_n(z)$ have no roots in $\overline{\mathbb{S}}_s$ (property 2), there exists a function $\mathbf{h} \in \mathcal{A}(\mathbb{S}_s)^N$, such that $\mathbf{Y}'(z) = \exp(\mathbf{h}(z))$ for all $z \in \overline{\mathbb{S}}_s$. In fact, property 1 (real & positive) allows one to choose this function in such a way that $\mathbf{h}(\mathbb{R}) \subseteq \mathbb{R}^N$. Consequently, the function $\mathbf{f}'_\star(z) := \mathbf{h}(z) - \mathbf{a}(z)$ is in $\mathcal{A}(\mathbb{S}_s)^N$ and satisfies $\mathbf{f}'_\star(\mathbb{R}) \subseteq \mathbb{R}^N$. Due to property 3 (asymptotics), $\mathbf{f}'_\star \in \mathcal{BA}(\mathbb{S}_s)^N$. As a consequence of the Y-system (Y) and the functional relation (1.1.1) for the asymptotics, we have

$$\begin{aligned} e^{f'_{\star,n}(x+is)+f'_{\star,n}(x-is)} &= e^{-a_n(x+is)-a_n(x-is)}\, Y'_n(x+is)Y'_n(x-is) \\ &= \prod_m e^{-G_{nm} a_m(x)} \left(1 + Y'_m(x)\right)^{G_{nm}} \\ &= \prod_m \left(e^{-a_m(x)} + e^{f'_{\star,m}(x)}\right)^{G_{nm}} \ . \end{aligned} \tag{2.3.14}$$

Hence, due to continuity of $\mathbf{f}'_\star(x+is)+\mathbf{f}'_\star(x-is)-\mathbf{L_0}[\mathbf{f}'_\star](x)$ there exists $\mathbf{v}\in\mathbb{Z}^N$, such that

$$\mathbf{f}'_\star(x+is)+\mathbf{f}'_\star(x-is)+2\pi i\mathbf{v}=\mathbf{L_0}[\mathbf{f}'_\star](x)\ . \tag{2.3.15}$$

But since $\mathbf{f}'_\star(\mathbb{R})\subseteq\mathbb{R}^N$, the Schwarz reflection principle $\mathbf{f}'_\star(x-is)=\overline{\mathbf{f}'_\star(x+is)}$ allows (2.3.15) to be rewritten as

$$2\,\mathrm{Re}\,\mathbf{f}'_\star(x+is)+2\pi i\mathbf{v}=\mathbf{L_0}[\mathbf{f}'_\star](x)\ . \tag{2.3.16}$$

Since the right hand side is real, it follows that $\mathbf{v}=0$. Thus $\mathbf{f}'_\star$ solves (2.3.8).

As in the previous proof, by Proposition 2.2.1 we can apply Proposition 2.3.3. Part i) of the latter proposition states that $\mathbf{f}'_\star$ is the unique solution to (2.3.8) in $\mathcal{BA}(\mathbb{S}_s)^N$ which maps $\mathbb{R}$ to $\mathbb{R}^N$. Part ii) states that $\mathbf{f}'_\star$ solves the TBA equation (2.3.9) for any choice of $\mathbf{C}'$, so in particular it solves (TBA). But from Theorem 2.1.2 i) we know that $\mathbf{f}_\star$ is the unique solution to (TBA) in $BC(\mathbb{R},\mathbb{R})^N$, and hence $\mathbf{f}'_\star=\mathbf{f}_\star$ on $\mathbb{R}$ (and therefore, by uniqueness of the analytic continuation, also on $\overline{\mathbb{S}}_s$).

This completes the proof of Theorem 2.1.1.

Proof of Theorem 2.1.2, uniqueness in part ii)

This is now immediate from Theorem 2.1.1, as the solution to the Y-system in $\mathcal{A}(\mathbb{S}_s)^N$ satisfying 1–3 is unique.

This completes the proof of Theorem 2.1.2.

Proof of Corollary 2.1.3

We only need to show that for $\mathbf{a}=0$, the unique solution $\mathbf{f}_\star$ from Theorem 2.1.2 is constant. By part 2 of that theorem, $\mathbf{Y}$ is then constant, too.

By Theorem 2.1.2 i), the solution $\mathbf{f}_\star$ is independent of $\mathbf{C}$, and in particular is equal to the unique solution found in Proposition 2.2.1 for $\mathbf{C}=\frac{1}{2}\mathbf{G}$. In the proof of Proposition 2.2.1 in Section 2.2.3 it was verified that the integral operator $I: BC(\mathbb{R},\mathbb{R})^N\to BC(\mathbb{R},\mathbb{R})^N$ from (2.2.22) is a contraction. Explicitly,

$$I_i[\mathbf{g}]:=\sum_{j=1}^{N}\int_{-\infty}^{\infty}\frac{1}{w_i}\phi_{ij}(x-y)\sum_{k=1}^{N}G_{jk}\left[\log\left(e^{-a_k(y)}+e^{w_kg_k(y)}\right)-\tfrac{1}{2}w_kg_k(y)\right]dy \tag{2.3.17}$$

where ϕ_{ij} are the entries of $\Phi_{\frac{1}{2}\mathbf{G}}$ and $\mathbf{w}$ is the Perron-Frobenius eigenvector of $\mathbf{G}$. The relation between $\mathbf{g}$ and the functions $\mathbf{f}$ in the TBA equation is $f_i(x)=w_ig_i(x)$.

By assumption we have $\mathbf{a}=0$. Clearly, if also $\mathbf{g}$ is constant, so is $I[\mathbf{g}]$. This shows that the operator I preserves the space of constant functions (for $\mathbf{a}=0$). Hence the unique fixed point $\mathbf{g}_\star$ of I (and hence the unique solution $\mathbf{f}_\star$ to (TBA)) must be a constant function.

This completes the proof of Corollary 2.1.3.

2.4. Discussion and outlook

In this section, we will make additional comments on the existence and uniqueness results in this chapter and discuss possible further investigations, in particular with reference to the physical background.

First of all, let us mention that the existence of a unique solution to Y-systems or TBA equations, even if the latter arise from a physical context, is by no means clear:

- *Existence*: as already mentioned earlier, the constant Y-system (2.1.1) has no non-negative solution if the spectral radius of $\mathbf{G}$ is larger or equal than 2 [RTV]. It seems likely that this generalises to solutions with asymptotics $\mathbf{a} = \mathbf{w}e^{\pm\gamma x/s}$, cf. (1.1.2).
- *Uniqueness*: when relaxing the reality condition $\mathbf{Y}(\mathbb{R}) \subset \mathbb{R}$, uniqueness generally fails to hold (see [DT, Fig. 1] for the case $N = 1$). Moreover, there exist Y-systems of some more general form for which stability investigations show that the associated TBA equation for constant functions is not a contraction, and in fact may display chaotic behaviour upon iteration [CF].[1]

We will conclude this chapter with a brief outline of possible future investigations. Our results already cover a significant class of integrable models [Za3, KM]. In physical terms, the state corresponding to the unique solution with no zeros in $\overline{\mathbb{S}}_s$ is the ground state. Two main directions in which to extend our results are to a) consider excited states, and b) treat situations associated with more general models. We will briefly comment on both of these.

a) Including excited states

To include excited states one has to allow for Y-functions which have roots in $\overline{\mathbb{S}}_s$. In this case, the TBA equation involves an additional term which depends on the positions of these roots [KP, DT, BLZ4]. We will treat this situation in Chapter 3.2. It seems possible that existence and uniqueness of a solution to the TBA equation for a generic set of root positions can be established with similar techniques as in the present chapter. However, to produce a solution to the Y-system with a sufficiently far analytic continuation, one needs to impose additional constraints (see Lemma 3.2.10). It would be very interesting to understand if some general statements about solutions to these constraints can be made.

In examples, these solutions are parametrised by a discrete set of "quantum numbers". In the asymptotic limit $r \to \infty$ of Y-systems with $\mathbf{a}(x) = r\cosh(\gamma x/s)\,\mathbf{w}$ (relativistic scattering theories in volume r) these quantum numbers are expected to coincide with the Bethe-Yang quantum numbers [YY] which parametrise solutions of the Bethe ansatz

[1] Note that our results do not imply that the TBA integral operators are contracting, except in the specific case in Section 2.2.3 to which the Banach Theorem is applied. However, our bound on the contraction constant κ is certainly not optimal (see Remark 2.2.11) and the TBA integral operator may well be contracting even if our bound yields $\kappa > 1$. The results of [CF] show that sometimes it is not contracting.

equations (see also Remark 3.2.11). In the limit $r \to 0$, quantisation conditions in terms of Virasoro states have been conjectured for the Yang-Lee model ($N = 1, \mathbf{G} = 1$), see [BDP]. A related $N = 1$ model, albeit with a slightly deformed Y-system (see below), is the Sinh-Gordon model for which a conjecture on the classification of states (for all r) in terms of solutions to the Y-system has been given in [Te].

b) More general Y-systems

It is fair to ask if the approach presented in this thesis is flexible enough to make contact with a larger number of physical models. This would require us to consider different conditions on $\mathbf{G}$ as well as Y-systems or TBA equations of a more general form.

There are several generalisations which still fit the form (Y):

- As mentioned in the end of Section 2.2.5, with the method used in [FKS] it may be possible to treat cases of slightly more general $\mathbf{G}$ giving rise to $\kappa = 1$, where our results no longer apply.
- Giving up the reality requirement in a controlled way would for example allow for a treatment of Y-systems with chemical potentials, where a constant imaginary vector is added to $\mathbf{a}$ (see e.g. [KM, Fe1]). The twisted Y-system which we encounter in part II of this thesis is of this type.

There are also more general forms of (Y), which would be of interest. For example:

- Y-systems with a second shift parameter t such as

$$Y_n(x+is)Y_n(x-is) = \frac{(1+Y_n(x+it))\,(1+Y_n(x-it))}{\prod_{m=1}^{N}\left(1+\frac{1}{Y_m(x)}\right)^{H_{nm}}} \tag{2.4.1}$$

  where $\mathbf{H}$ is the adjacency matrix of a finite Dynkin diagram. This specific type of Y-system appears in the context of Affine Toda Field Theories [Ma, FKS] whose simplest representative, the Sinh-Gordon model, has $\mathbf{H} = 0$ and bears some resemblance to (Y).

- The case of two simple Lie algebras giving rise to a Y-system of the form

$$Y_{n,m}(x+is)Y_{n,m}(x-is) = \frac{\prod_{k=1}(1+Y_{k,m}(x))^{G_{nk}}}{\prod_{l=1}\left(1+\frac{1}{Y_{n,l}(x)}\right)^{H_{ml}}}. \tag{2.4.2}$$

  via their Dynkin diagrams $\mathbf{G}$ and $\mathbf{H}$. In applications, often many of the $Y_{n,m}$ are required to be trivial. One example, albeit with an infinite number of Y-functions, is the famous "T-hook" of the AdS/CFT Y-system (see e.g. [Baj] for more details and references). The physically relevant solutions in this case have, however, rather complicated analytical properties involving also branch cuts.

3. Further results

3.1. Asymptotic behaviour of solutions to TBA equations

In this section we compile a few results on the behaviour of solutions to (TBA) as $x \to \infty$.

Lemma 3.1.1. *Let* $\mathbf{G} \in \mathrm{Mat}(N, \mathbb{R})$*,* $\mathbf{C} \in \mathrm{Mat}_{<2}(N)$*,* $\mathbf{a} \in BC_{-}(\mathbb{R}, \mathbb{R})^N$*, and* $\mathbf{f} \in BC(\mathbb{R}, \mathbb{R})^N$ *a solution of* (TBA)*. Then the following holds:*

1. *If* $a_n(x) \xrightarrow{x\to\infty} +\infty$ *for all* $n = 1, \ldots, N$*, then* $f_n(x) \xrightarrow{x\to\infty} 0$ *for all* $n = 1, \ldots, N$*.*
2. *If there exist* $A, X > 0$ *such that* $a_n(x) > Ax$ *for all* $x > X$ *and all* $n = 1, \ldots, N$*, then there exist* $C, c > 0$ *such that* $|f_n(x)| < Ce^{-cx}$ *for all* $x \in \mathbb{R}$ *and all* $n = 1, \ldots, N$*.*

Proof. We treat case 1 and 2 both at once. Define

$$\tilde{\mathbf{f}} := (\Phi_{\mathbf{C}} \star (\mathbf{G} - \mathbf{C}) \cdot \mathbf{f})\,(x)\ . \tag{3.1.1}$$

The components of $\mathbf{f}$ are Hölder continuous according to Lemma 2.3.2. It then follows easily that the components of $\tilde{\mathbf{f}}$ are also Hölder continuous. Hence, by Proposition 1.2.1 we have $\tilde{\mathbf{f}} \in \mathcal{BA}(\mathbb{S}_s)$ and

$$\tilde{\mathbf{f}}(x + is) + \tilde{\mathbf{f}}(x - is) - \mathbf{C} \cdot \tilde{\mathbf{f}}(x) = (\mathbf{G} - \mathbf{C}) \cdot \mathbf{f}(x) \tag{3.1.2}$$

for all $x \in \mathbb{R}$. We reorganise this equation as follows:

$$\tilde{\mathbf{f}}(x + is) + \tilde{\mathbf{f}}(x - is) - \mathbf{G} \cdot \tilde{\mathbf{f}}(x) = (\mathbf{G} - \mathbf{C}) \cdot \Big(\mathbf{f}(x) - \tilde{\mathbf{f}}(x)\Big) \tag{3.1.3}$$

and use the inverse direction of Lemma 1.2.1 to obtain

$$\tilde{\mathbf{f}}(x) = (\Phi_{\mathbf{G}} \star (\mathbf{G} - \mathbf{C}) \cdot \mathbf{g})\,(x) \tag{3.1.4}$$

where, by (3.1.1) and (TBA),

$$\mathbf{g}(x) = \mathbf{f}(x) - \tilde{\mathbf{f}}(x) = \int_{-\infty}^{\infty} \Phi_{\mathbf{C}}(x - y) \cdot \mathbf{G} \cdot \log\left(e^{-\mathbf{a}(y) - \mathbf{f}(y)} + 1\right)\, dy\ . \tag{3.1.5}$$

We are now going to show that $\mathbf{g}(x)$ vanishes appropriately when $x \to \infty$. From section 1.2 it is clear that there exist $B, \gamma > 0$ such that $0 < [\Phi_{\mathbf{C}}(x)]_{nm} < Be^{-\gamma x/s}$ for all $x \in \mathbb{R}$ and all $n, m = 1, \ldots, N$. Furthermore, we have

$$0 < \log\left(e^{-a_n(x) - f_n(x)} + 1\right) < e^{-a_n(x) - f_n(x)} \leq Se^{-a_n(x)}\ , \tag{3.1.6}$$

where $S := e^{-\inf_{n,x}(f_n(x))} > 0$. Define also $L := \sup_{n,x} \log\left(e^{-a_n(x)-f_n(x)} + 1\right) > 0$, $M(x) := \sup_{n,y\geq x} e^{-a_n(y)}$, $P := \max_{n,m,x} |[\Phi_{\mathbf{C}}(x)]_{nm}|$, $R := \max_{n,m} \rho_{nm} > 0$ (where ρ_{nm} is as defined in section 2.2) and $G := \max_{n,m} |G_{nm}|$. Let us split the integral (3.1.5) and estimate each part separately and component-wise. For any $x \in \mathbb{R}$ and $n, m, k = 1, \ldots, N$ we have:

$$\left| \int_{-\infty}^{x/2} [\Phi_{\mathbf{C}}(x-y)]_{nm}\, G_{mk} \log\left(e^{-a_k(y)-f_k(y)} + 1\right)\, dy \right|$$
$$< LGBe^{-\frac{\gamma}{s}x} \int_{-\infty}^{x/2} e^{\frac{\gamma}{s}y}\, dy \;=\; \frac{LGBs}{\gamma} e^{-\frac{\gamma}{2s}x} \tag{3.1.7}$$

$$\left| \int_{x/2}^{\infty} [\Phi_{\mathbf{C}}(x-y)]_{nm}\, G_{mk} \log\left(e^{-a_k(y)-f_k(y)} + 1\right)\, dy \right|$$
$$< GSM(\tfrac{x}{2}) \int_{x/2}^{\infty} \left| [\Phi_{\mathbf{C}}(x-y)]_{nm} \right| dy \;<\; GSRM(\tfrac{x}{2}) \tag{3.1.8}$$

Since $M(x) \to 0$ as $x \to \infty$, we conclude $\mathbf{g}(x) \to 0$ as $x \to \infty$. If we impose additionally at least linear growth on $a_k(x)$ for $x > X$ (case 2), then we have $M(\frac{x}{2}) < e^{-Ax}$ for $x > 2X$. Hence, in this case $\mathbf{g}(x)$ exhibits exponential decay.

Having established the appropriate decay of $\mathbf{g}(x)$ it is now an easy exercise to conclude the appropriate decay of $\tilde{\mathbf{f}}(x)$ from (3.1.4) and, by virtue of $\mathbf{f}(x) = \mathbf{g}(x) + \tilde{\mathbf{f}}(x)$, the appropriate decay of $\mathbf{f}(x)$. □

Lemma 3.1.2. *Let $a, f \in C(\mathbb{R}, \mathbb{R})$ and consider the function $F(x) := \log(e^{-a(x)} + e^{f(x)})$. Suppose $f(x)$ has an asymptotic expansion*

$$f(x) \sim \sum_{j=0}^{\infty} f^{(j)} e^{-g_j x} \qquad (x \to \infty)\,, \tag{3.1.9}$$

where $\{g_j\}_{j\in\mathbb{Z}_{\geq 0}}$ is a strictly increasing sequence of real non-negative numbers. If $a(x)$ grows strictly faster than linearly as $x \to \infty$, then (3.1.9) is also an asymptotic expansion for $F(x)$.

Proof. Consider the function $\Delta(x) := F(x) - f(x) = \log(e^{-a(x)-f(x)} + 1)$. If the recursively defined limits

$$\delta^{(j)} := \lim_{x\to\infty} e^{g_j x} \left(\Delta(x) - \sum_{k=0}^{j-1} \delta^{(k)} e^{-g_k x} \right) \tag{3.1.10}$$

exist, then

$$\Delta(x) \sim \sum_{j=0}^{\infty} \delta^{(j)} e^{-g_j x} \qquad (x \to \infty) \tag{3.1.11}$$

is the (unique) asymptotic expansion of $\Delta(x)$ with respect to the asymptotic sequence $\{e^{-g_j x}\}_{j\in\mathbb{Z}_{\geq 0}}$. Clearly, $\delta^{(0)} = \lim_{x\to\infty}\Delta(x) = 0$. Suppose now that $\delta^{(0)},\dots,\delta^{(j-1)} = 0$. Then we have also

$$\begin{aligned}\delta^{(j)} = \lim_{x\to\infty} e^{g_j x}\Delta(x) &= \lim_{x\to\infty}\log\left(\left(1+\frac{1}{e^{a(x)+f(x)}}\right)^{e^{g_j x}}\right)\\ &= \log\left(\lim_{x\to\infty}\left(1+\frac{1}{e^{a(x)+f(x)}}\right)^{e^{g_j x}}\right) = \log(1) = 0,\end{aligned} \tag{3.1.12}$$

where the last limit is evaluated using the fact that $a(x)$ grows faster than linearly. Since asymptotic expansions respect addition, we conclude that

$$F(x) \sim \sum_{j=0}^{\infty} f^{(j)} e^{-g_j x} \qquad (x\to\infty)\,, \tag{3.1.13}$$

which proves the claim. $\square$

Next we give a series representation for the kernel:

Lemma 3.1.3. *For $x \neq 0$, the function $\phi_d(x)$ can be written as*

$$\phi_d(x) \;=\; \frac{1}{2s\sin(\gamma)}\left(e^{\frac{\pi-\gamma}{s}x} - e^{-\frac{\pi-\gamma}{s}x}\right)\sum_{j=0}^{\infty} e^{-\frac{\pi}{s}(2j+1)|x|}\,, \tag{3.1.14}$$

where $\gamma\in(0,\pi)$ is given by $d = 2\cos(\gamma)$.

Proof. Rewrite (1.2.56) in terms of a geometric series. $\square$

Notations 3.1.4.

- If $\mathbf{G}$ is diagonalisable, we denote by $\gamma_k \in [0,\pi]$ the number such that $\lambda_k = 2\cos(\gamma_k)$ is the k-th eigenvalue (with ordering $\lambda_{\mathrm{PF}} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$). Furthermore, we denote by $\mathbf{v}^k$ the k-th right eigenvector and by $\mathbf{w}^k$ the k-th left eigenvector of $\mathbf{G}$, such that $\mathbf{D} = \mathbf{T}^{-1}\mathbf{G}\mathbf{T} = \mathrm{diag}(\lambda_1,\dots,\lambda_N)$, where $\mathbf{T} = (\mathbf{v}^1\ \dots\mathbf{v}^N)$ and $\mathbf{T}^{-1} = (\mathbf{w}^1\ \dots\mathbf{w}^N)^\top$.
- For every $j\in\mathbb{Z}_{>0}$, define the bracket

$$[j] := \begin{cases} j \bmod 2N & \text{if } j \bmod 2N = 1,\dots,N \\ (2N+1-j) \bmod 2N & \text{if } j \bmod 2N = N+1,\dots,2N \end{cases}. \tag{3.1.15}$$

Moreover, we define

$$g_j := \begin{cases} 2\pi\left\lfloor\frac{j}{2N+1}\right\rfloor + \gamma_{[j]} & \text{if } j \bmod 2N = 1,\dots,N \\ 2\pi\left\lfloor\frac{j}{2N+1}\right\rfloor + 2\pi - \gamma_{[j]} & \text{if } j \bmod 2N = N+1,\dots,2N \end{cases}. \tag{3.1.16}$$

- We denote the standard scalar product on $\mathbb{R}^N$ by $\langle\cdot,\cdot\rangle$.

With these notations, we have the following

Proposition 3.1.5. *Let* $\mathbf{G} \in \mathrm{Mat}_{<2}(N)$ *and* $\mathbf{a} \in BC_-(\mathbb{R},\mathbb{R})^N$ *such that all its components* $a_n(x)$ *grow strictly faster than linearly as* $x \to \infty$*. Suppose* $\mathbf{f} \in BC(\mathbb{R},\mathbb{R})^N$ *is a solution of* (TBA) *for* $\mathbf{C} = \mathbf{G}$*. Then it has the asymptotic expansion*

$$\mathbf{f}(x) \;\sim\; \frac{1}{s}\sum_{j=1}^{\infty} \frac{\langle \mathbf{w}^{[j]}, \mathbf{X}^j\rangle}{\tan(g_j)}\, \mathbf{v}^{[j]}\, e^{-\frac{1}{s}g_j x} \qquad (x\to\infty)\ , \tag{3.1.17}$$

where

$$\mathbf{X}^j = \int_{-\infty}^{\infty} e^{\frac{1}{s}g_j y} \log\left(e^{-\mathbf{a}(y)-\mathbf{f}(y)}+1\right)\, dy\ . \tag{3.1.18}$$

Proof. First note that the integrals $\mathbf{X}^j$ exist: With Lemma 3.1.1 we conclude that the leading terms of the components of $\mathbf{f}(x)$ are exponential. Thus, all the components $f_n(x)$ have asymptotic expansions of the form (3.1.9). But this implies with Lemma 3.1.2 that for all $j \in \mathbb{Z}_{\geq 0}$ there exists a $X > 0$ such that

$$\log\left(e^{-a_n(x)-f_n(x)}+1\right) = \log\left(e^{-a_n(x)}+e^{f_n(x)}\right) - f_n(x) < e^{-\frac{1}{s}g_j x} \tag{3.1.19}$$

for all $x > X$. Hence, the integrands in (3.1.18) decay exponentially on both sides.

Now we turn to (TBA) with $\mathbf{C} = \mathbf{G}$, and rewrite it as

$$\mathbf{f}(x) = \mathbf{T}\cdot\mathbf{D}\cdot\int_{-\infty}^{\infty} \Phi_{\mathbf{D}}(x-y)\cdot\mathbf{T}^{-1}\cdot\log\left(e^{-\mathbf{a}(y)-\mathbf{f}(y)}+1\right)\, dy\ , \tag{3.1.20}$$

with $\mathbf{D} = \mathbf{T}^{-1}\mathbf{G}\mathbf{T} = 2\,\mathrm{diag}(\cos(\gamma_1),\ldots,\cos(\gamma_N))$, where the left and right eigenvectors of $\mathbf{G}$ appear through $\mathbf{T} = (\mathbf{v}^1\ \ldots\mathbf{v}^N)$ and $\mathbf{T}^{-1} = (\mathbf{w}^1\ \ldots\mathbf{w}^N)^\top$. With the help of Lemma 3.1.1 it is easy to see that the integrals

$$\int_x^{\infty} \phi_d(x-y)\log\left(e^{-a_n(y)-f_n(y)}+1\right)\, dy \tag{3.1.21}$$

and also (for any $g \in \mathbb{R}$)

$$\int_x^{\infty} e^{gy}\log\left(e^{-a_n(y)-f_n(y)}+1\right)\, dy \tag{3.1.22}$$

decay faster than asymptotically as $x \to \infty$. Hence, we may replace the kernel $\Phi_{\mathbf{D}}(x-y)$ in (3.1.20) with the expansion (3.1.14) (using the expansion valid for positve arguments on the whole domain); the resulting function $\tilde{\mathbf{f}}(x)$ must have the same asymptotic expansion as $\mathbf{f}(x)$ in terms of exponentials.

We will compute $\tilde{\mathbf{f}}(x)$ in components. Let us denote $Q_j = \frac{\pi-\gamma_j}{\pi}$ and

$$\chi_n^{(j,m)\pm} = \int_{-\infty}^{\infty} e^{\frac{\pi}{s}(2j+1\pm Q_m)y}\log\left(e^{-a_n(y)-f_n(y)}+1\right)\, dy\ . \tag{3.1.23}$$

Then we have:

$$\tilde{f}_n(x) = \frac{1}{s}\sum_{j=0}^{\infty} e^{-\frac{\pi}{s}(2j+1)x} \sum_{m=1}^{N}\sum_{k=1}^{N} \frac{T_{nm}(T^{-1})_{mk}}{\tan(\gamma_m)} \left(e^{Q_m \frac{\pi}{s}} \chi_k^{(j,m)^-} - e^{-Q_m \frac{\pi}{s}} \chi_k^{(j,m)^+} \right) \tag{3.1.24}$$

$$= -\frac{1}{s}\sum_{\epsilon=\pm} \epsilon \sum_{j=0}^{\infty}\sum_{m=1}^{N} e^{-\frac{\pi}{s}(2j+1+\epsilon Q_m)x} \frac{T_{nm}}{\tan(\gamma_m)} \sum_{k=1}^{N} (T^{-1})_{mk} \chi_k^{(j,m)^\epsilon} \tag{3.1.25}$$

$$= \frac{1}{s}\sum_{j=0}^{\infty}\sum_{m=1}^{N} e^{-\frac{1}{s}(2\pi j+\gamma_m)x} \frac{(\mathbf{v}^m)_n}{\tan(\gamma_m)} \sum_{k=1}^{N} (\mathbf{w}^m)_k \chi_k^{(j,m)^-}$$

$$- \frac{1}{s}\sum_{j=0}^{\infty}\sum_{m=1}^{N} e^{-\frac{1}{s}(2\pi(j+1)-\gamma_m)x} \frac{(\mathbf{v}^m)_n}{\tan(\gamma_m)} \sum_{k=1}^{N} (\mathbf{w}^m)_k \chi_k^{(j,m)^+} \tag{3.1.26}$$

Merging the sums over m and j, the motivation for Notations 3.1.4 becomes clear, and we obtain precisely (3.1.17). □

Remark 3.1.6. In the context of integrable models, the asymptotic expansion (3.1.17) is of interest since sometimes the scalar products $I_j = \langle \mathbf{w}^{[j]}, \mathbf{X}^j \rangle$ are the eigenvalues of local integrals of motion, see for example [BLZ4, Fe2, Ni, MMR].

3.2. Solutions with roots in the strip

In this section, we will extend some of the previous discussion to solutions of (Y) which do have roots in $\overline{\mathbb{S}}_s$.

Let us be even more general and drop the requirement that solutions of the Y-system are real on the real axis. In case $\mathbf{G}$ is not an integer-valued matrix, this requires a branch choice on the right hand side of (Y). We fix this by setting

$$(1+Y_m(x))^{G_{nm}} := \begin{cases} e^{G_{nm}\log(1+Y_m(x))} & \text{if } Y_m(x) \neq -1 \\ \lim_{t\to 0} t^{G_{nm}} & \text{else} \end{cases}, \tag{3.2.1}$$

where we agree to choose that branch of the logarithm which for $x = 0$ coincides with the standard branch and is continuous from there.[1]

Roots on $\partial\mathbb{S}_s$

Let us first look at the case where some of the Y-functions have roots on the boundary $\partial\mathbb{S}_s$ of the strip.

[1]Other choices are possible. It might be more natural to choose the branch such that it coincides with the standard branch as $x \to \infty$ or $x \to -\infty$. However, this choice is tricky due to the fact that a priori $\mathrm{Im}(Y_m(x))$ has no well-defined limit.

Lemma 3.2.1. *Suppose $Y_1, \ldots, Y_N \in \mathcal{A}(\mathbb{S}_s)$ solve the Y-system* (Y). *Then the following holds:*

i) If $Y_m(z_0) = 0$ for some $z_0 \in \partial\mathbb{S}_s$ and $m \in \{1, \ldots, N\}$, then there exists $n \in \{1, \ldots, N\}$ such that $Y_n(\mathrm{Re}(z_0)) = -1$.

ii) Suppose $\mathbf{G}$ is non-negative and irreducible. If $Y_n(x_0) = -1$ for some $x_0 \in \mathbb{R}$ and $n \in \{1, \ldots, N\}$, then there exists $m \in \{1, \ldots, N\}$ such that $Y_m(x_0 + is) = 0$ or $Y_m(x_0 - is) = 0$.

Proof. i) Let $Y_m(z_0) = 0$. Then from (Y) we get

$$\prod_{k=1}^{N} (1 + Y_k(\mathrm{Re}(z_0)))^{G_{mk}} = 0 \ . \tag{3.2.2}$$

At least for one $n \in \{1, \ldots, N\}$ the factor $(1 + Y_n(\mathrm{Re}(z_0)))^{G_{mn}}$ is zero.

ii) Let $Y_n(x_0) = -1$. Since $\mathbf{G}$ is irreducible, there exists $m \in \{1, \ldots, N\}$ such that $G_{mn} \neq 0$. From non-negativity of $\mathbf{G}$ thus follows that

$$\prod_{k=1}^{N} (1 + Y_k(x_0))^{G_{mk}} = 0 \ . \tag{3.2.3}$$

Substituting (Y) yields $Y_m(x_0 + is)Y_m(x_0 - is) = 0$, from which the claim follows. □

Remark 3.2.2.

i) If $\mathrm{Re}(z_0) \to \pm\infty$, then non-negativity of $\mathbf{G}$ is already needed in i) to conclude the argument.

ii) If we require that the solution of (Y) satisfies $\mathbf{Y}(\mathbb{R}) \subset \mathbb{R}^N$, then $Y_m(x_0 + is) = 0$ and $Y_m(x_0 - is) = 0$ *both* follow in Lemma 3.2.1.ii) due to the Schwarz reflection principle. This tends to have consequences for the global bahaviour of the functions $Y_n(z)$, which we now illustrate for the case $N = 1$, $\mathbf{G} = 1$. The Y-system is

$$Y(x + is)Y(x - is) = 1 + Y(x) \ , \tag{3.2.4}$$

and we assume $Y \in \mathcal{A}(\mathbb{S}_s)$ is real on the real axis and $Y(x) > 0$ if $x \in \mathbb{R}$ is large enough (positive or negative). Suppose we have $Y(x_0) = -1$ for some $x_0 \in \mathbb{R}$. In other words, there is a $k \in \mathbb{Z}_{>0}$ such that $1 + Y(z)$ has a root of order k in x_0. Then it follows from ii) together with $Y(\mathbb{R}) \subset \mathbb{R}$ that $Y(z)$ has roots of order $\frac{k}{2}$ in $x_0 + is$ and $x_0 - is$. This leaves two options:

- k is odd: in that case it is impossible to analytically continue $Y(z)$ to $x_0 \pm is$ without creating a square root branch cut.
- k is even: $1 + Y(x)$ has a root of even order in x_0. But since $Y(x) \to +\infty$ as $x \to \infty$, it follows that $Y(x) \geq -1$ for all $x \in \mathbb{R}$.

The last remark indicates the practical relevance of Lemma 3.2.1: complex-conjugation-symmetric Y-functions (with asymptotics $Y_n(z) \to \infty$ as $x \to \pm\infty$) are, even if they have roots in the strip, typically larger than -1 all along the real line. If such a solution to (Y) is analytically continued in some parameter and some component $Y_n(x)$ suddenly dips below -1 on the real line, this indicates the entry of new roots across the boundary into the strip.

TBA equations

As is well-known, TBA equations can also be obtained for solutions of the Y-system with roots in the strip. The new TBA equations will, however, depend on extra parameters, namely the positions of the roots in $\overline{\mathbb{S}}_s$. We will make the following restrictions for the rest of this section: we only consider solutions to (Y) whose components $Y_n(z)$

1. do *not* have any root on the boundary $\partial\mathbb{S}_s$,
2. have a *finite* number of roots inside $\mathbb{S}_s$.

Definition 3.2.3. A *finite root configuration in* $\mathbb{S}_s$, denoted by ϱ, is given by the following data: for each $n = 1, \dots, N$ an integer $M_n \in \mathbb{Z}_{\geq 0}$ and an unordered M_n-tuple $(z_1^{(n)}, \dots, z_{M_n}^{(n)})$ of (possibly mutually equal) points $z_j^{(n)} \in \mathbb{S}_s$. We say that a function $\mathbf{F} \in \mathcal{A}(\mathbb{S}_s)^N$ has root configuration ϱ if $F_n(z)$ has precisely M_n roots in $\mathbb{S}_s$ (with multiplicity), and $z_1^{(n)}, \dots, z_{M_n}^{(n)}$ are the locations of these roots.

Notations 3.2.4. We write $\mathcal{R}_N$ for the set of all finite root configurations in $\mathbb{S}_s$. Furthermore, we denote by $\mathcal{R}_N^{\mathrm{cc}} \subset \mathcal{R}_N$ and the subset of those finite root configurations which are invariant under complex conjugation.

Lemma 3.2.5. *Let* $\mathbf{F} \in \mathcal{A}(\mathbb{S}_s)^N$ *be such that all its components* $F_n(z)$ *have a finite number of roots in* $\mathbb{S}_s$*. Then there exists a unique* $\varrho \in \mathcal{R}_N$*, such that* $\mathbf{F}$ *has root configuration* ϱ*. If* $\mathbf{F}(\mathbb{R}) \subseteq \mathbb{R}^N$*, then* $\varrho \in \mathcal{R}_N^{\mathrm{cc}}$*.*

Proof. The first statement is trivial. The second statement follows from the Schwartz reflection principle. □

Definition 3.2.6. Let $\varrho \in \mathcal{R}_N$. We call a function $\mathbf{B}^\varrho \in \mathcal{BA}(\mathbb{S}_s)$ with components $B_n^\varrho(z)$ a *root detector for* ϱ if

i. $\mathbf{B}^\varrho$ has root configuration ϱ, and $B_n^\varrho(z) \neq 0$ for all $z \in \partial\mathbb{S}_s$ and $n = 1, \dots, N$,

ii. for all $n = 1, \dots, N$ and $y \in [-s, s]$: $\lim_{x\to\pm\infty} B_n^\varrho(x + iy) \in \mathbb{C}^\times$

To every root detector $\mathbf{B}^\varrho \in \mathcal{BA}(\mathbb{S}_s)$ there exists a unique function $\mathbf{d}^\varrho \in BC(\mathbb{R}, \mathbb{C})$ whose components $d_n^\varrho(x)$, for all $n = 1, \dots, N$, are defined by the equation

$$B_n^\varrho(x + is)B_n^\varrho(x - is) = e^{d_n^\varrho(x)} , \tag{3.2.5}$$

together with the condition that in the limit $x \to \infty$ the component $d_n^\varrho(x)$ coincides with the standard branch of the logarithm of $B_n^\varrho(x+is)B_n^\varrho(x-is)$. The main example of a root detector later used in this thesis is the following:

Example 3.2.7. For $\varrho \in \mathcal{R}_N$, the function $\mathbf{B}^\varrho : \mathbb{S}_s \to \mathbb{C}^N$ with components

$$B_n^\varrho(z) = \prod_{j=1}^{M_n} \tanh\left(\tfrac{\pi}{4s}(z - z_j^{(n)})\right) \tag{3.2.6}$$

is in $\mathcal{BA}(\mathbb{S}_s)^N$ and has root configuration ϱ. It is a root detector for ϱ with $d_n^\varrho(x) = 0$ for all $n = 1, \ldots, N$ and $x \in \mathbb{R}$.

However, in the rest of this section we use the general notion of root detector rather than only this particular example, since there are many examples in the literature where other root detectors are more natural (for instance root detectors defined using the scattering matrix, if such is available).

Consider a finite root configuration $\rho \in \mathcal{R}_N$ in $\mathbb{S}_s$. Let $Y_1, \ldots, Y_N \in \mathcal{A}(\mathbb{S}_s)$ be a solution to (Y) with finite root configuration $\varrho \in \mathcal{R}_N$ and asymptotics $\mathbf{a} \in \mathcal{A}(\mathbb{S}_s)^N$. We now describe the associated TBA equations. Fix a root detector $\mathbf{B}^\varrho \in \mathcal{BA}(\mathbb{S}_s)^N$ for ϱ and denote by $\mathbf{d}^\varrho \in BC(\mathbb{R}, \mathbb{C})^N$ the function associated to $\mathbf{B}^\varrho$ as defined above. We can then write the Y-functions in the form

$$Y_n(z) = B_n^\varrho(z) e^{a_n(z) + f_n(z)} \ , \tag{3.2.7}$$

for some $\mathbf{f} \in \mathcal{A}(\mathbb{S}_s)^N$ which is bounded if a_n captures the asymptotics of Y_n and $Y_n(z)$ has no roots on $\partial\mathbb{S}_s$. Furthermore, fix a matrix $\mathbf{C} \in \mathrm{Mat}_{<2}(N, \mathbb{R})$, and let $\mathbf{k} \in \mathbb{Z}^N$ be a constant integer vector. The corresponding TBA equation reads

$$\mathbf{f}(x) + 2\pi i \,(2\mathbf{1} - \mathbf{C})^{-1} \cdot \mathbf{G} \cdot \mathbf{k} = \left(\Phi_{\mathbf{C}} \star \mathbf{L}_{\mathbf{C}}^\varrho[\mathbf{f}]\right)(x) \ , \tag{TBA$^\varrho$}$$

where

$$\mathbf{L}_{\mathbf{C}}^\varrho[\mathbf{f}](x) = \mathbf{G} \cdot \log\left(e^{-\mathbf{a}(x)} + \mathbf{B}^\varrho(x) e^{\mathbf{f}(x)}\right) - \mathbf{C} \cdot \mathbf{f}(x) - \mathbf{d}^\varrho(x) \ . \tag{3.2.8}$$

Here, the logarithm is again understood to be taken elementwise:

$$\left[\log\left(e^{-\mathbf{a}(y)} + \mathbf{B}^\varrho(y) e^{\mathbf{f}(y)}\right)\right]_j := \log\left(e^{-a_j(y)} + B_j(y) e^{f_j(y)}\right) \tag{3.2.9}$$

Moreover, we again choose that branch of the logarithm which for $y = 0$ coincides with the standard branch. Of course, existence of the logarithm is not a priori guaranteed. This problem goes away once reality on the real line is imposed.

Proposition 3.2.8. *Let $\mathbf{G} \in \mathrm{Mat}(N, \mathbb{R})$ be non-negative and irreducible, and $\mathbf{a} \in \mathcal{A}(\mathbb{S}_s)^N$ a function which satisfies* (1.1.1) *and whose restriction to the real line has real parts bounded from below. Let $Y_1, \ldots, Y_N \in \mathcal{A}(\mathbb{S}_s)$ be a solution to* (Y) *with finite root configuration $\varrho \in \mathcal{R}_N$ in $\mathbb{S}_s$, which satisfies, for all $n = 1, \ldots, N$,*

1. *$Y_n(z) \neq 0$ for all $z \in \partial\mathbb{S}_s$,*

2. *$\log Y_n(z) - a_n(z) \in \mathcal{BA}(\mathbb{S}_s)$.*

Then we can write Y_n in the form (3.2.7) where $\mathbf{f} \in \mathcal{BA}(\mathbb{S}_s)^N$, when restricted to the real axis, is a solution to (TBA$^\varrho$) for some $\mathbf{k} \in \mathbb{Z}^N$.

Proof. First of all, due to condition 2 and the definition of root detectors, there exists $\mathbf{f} \in \mathcal{BA}(\mathbb{S}_s)^N$, such that the Y-functions can be written in the form (3.2.7). Now let us show that $\mathbf{f}$ satisfies (TBA$^\varrho$).

Taking into account relations (1.1.1) and (3.2.5), the Y-system (Y) implies, for $x \in \mathbb{R}$,

$$e^{f_n(x+is)+f_n(x-is)} = e^{-d_n^\varrho(x)} \prod_{m=1}^{N} e^{-G_{nm}a_m(x)} \left(1+Y_m(x)\right)^{G_{nm}} \quad . \tag{3.2.10}$$

Consider the product on the right hand side. By condition 1, none of the Y_n has a root on $\partial\mathbb{S}_s$. Since $\mathbf{G}$ is non-negative and irreducible, Lemma 3.2.1 ii) applies. By its transposition, the factors $1+Y_m(x)$ are all nonzero. Moreover, since $\mathbf{G}$ is non-negative and $\mathrm{Re}(a_m(x))$ is bounded from below, the factors $e^{-G_{nm}a_m(x)}$ are finite. Hence, it is possible to take the logarithm (recall also (3.2.1)): there exists $\mathbf{k} \in \mathbb{Z}^N$, such that

$$\mathbf{f}(x+is) + \mathbf{f}(x-is) = \mathbf{G} \cdot \log\left(e^{-\mathbf{a}(x)} + \mathbf{B}^\varrho(x) e^{\mathbf{f}(x)}\right) - 2\pi i\, \mathbf{G} \cdot \mathbf{k} - \mathbf{d}^\varrho(x) \ , \tag{3.2.11}$$

where the branch of the logarithm is such that it coincides with the standard branch for $x = 0$ and is continuous from there. Upon subtracting the term $\mathbf{C} \cdot \mathbf{f}(x)$ on both sides, we apply Lemma 1.2.1 and recall (1.2.62). This results in (TBA$^\varrho$). □

Remark 3.2.9.

i) Of course, the correct choice of $\mathbf{k}$ depends on the branch choice we made for the logarithm and for the factors $(1+Y_m(x))^{G_{nm}}$.

ii) If $\mathbf{Y}(\mathbb{R}), \mathbf{a}(\mathbb{R}) \subset \mathbb{R}$, we may choose a root detector such that $\mathbf{B}^\varrho(\mathbb{R}) \subset \mathbb{R}$. In this case, we necessarily have $\mathbf{k} = 0$ due to the Schwarz reflection principle, and $\mathbf{f}$ is a real solution to (TBA$^\varrho$).

Entire solutions

Let $\varrho \in \mathcal{R}_N$ be a generic root configuration. Suppose $Y_1, \ldots, Y_N \in \mathcal{A}(\mathbb{S}_s)$ is a solution to (Y) with root configuration ϱ. A natural question to ask is whether the functions $Y_n(z)$ can be analytically continued beyond the strip $\mathbb{S}_s$. In particular, one is often interested in entire solutions.

Typically, solutions to (Y) cannot be continued to entire functions. There are nontrivial necessary conditions:

Lemma 3.2.10. *Suppose $Y_1, \ldots, Y_N$ extend to an entire solution of (Y). Then, for all roots $z_j^{(n)}$ of ϱ (with $n = 1, \ldots, N$ and $j = 1, \ldots, M_n$), the pair of conditions*

$$\prod_{m=1}^{N} \left(1 + Y_m(z_j^{(n)} \pm is)\right)^{G_{nm}} = 0 \ . \tag{3.2.12}$$

Proof. Suppose that $Y_n(z) = 0$ for some $z \in \mathbb{C}$ and some $1 \leq n \leq N$. Since Y_n is entire, we have $Y_n(z \pm 2is) \in \mathbb{C}$. Hence, the Y-system (Y) implies $0 = Y_n(z)Y_n(z \pm 2is) = \prod_{m=1}^{N} (1 + Y_m(z \pm is))^{G_{nm}}$. $\square$

Remark 3.2.11. In the context of relativistic massive integrable quantum field theories on a cylinder of radius r (with $\mathbf{G}$ of Dynkin type and $\mathbf{a}(x) = r\cosh(x)$), the conditions (3.2.12) have a clear physical interpretation. Roots in $\mathbb{S}_s$ correspond to repidities (relativistic momenta) of particle excitations (if $Y_n(z_0) = 0$ for some $z_0 \in \mathbb{S}_s$, then there is a particle of species n with rapidity z_0 running around the cylinder). Imaginary excitations correspond to "virtual particles". Since momenta of quantum particles on a circle are quantised, not every root configuration $\varrho \in \mathcal{R}_N$ can lead to a physical solution of (Y). One expects the allowed physical root configurations to be quantised. Indeed, the quantisation conditions are precisely (3.2.12). In the limit $r \to \infty$ these are known to reduce to *Bethe ansatz equations* for the rapidities $z_j^{(n)}$, see e.g. [BLZ4, DT, Te].

Question 3.2.12. It is an interesting question to ask if (3.2.12) are also sufficient conditions for $Y_1, \ldots, Y_N$ to be entire.[2] For example, for $N = 1$ and $\mathbf{G} = 1$ this is the case (indepedently of ϱ): the corresponding Y-system implies the periodicity $Y(z+5is) = Y(z)$ (for all $z \in \mathbb{C}$) of the only Y-function. Thus, it is enough to show that $Y(z)$ extends to an analytic function on $\mathbb{S}_{\frac{5}{2}s}$. It is not hard to see that conditions (3.2.12) suffice for that purpose. Similar but more complicated periodicities are known to hold for Y-systems of Dynkin type. Perhaps, they are also enough to turn (3.2.12) into sufficient conditions?

[2] In case $Y_1, \ldots, Y_N$ have roots on $\partial\mathbb{S}_s$, we should include the corresponding conditions for them as well. Only then is there a chance that the set of conditions (3.2.12) is sufficient.

A. Analysis

A.1. Fourier transformation of $\cosh(x)^{-m}$

Lemma A.1.1. *Let $m \in \mathbb{Z}_{\geq 1}$, and $f_m(x) = \cosh(x)^{-m}$. The Fourier transformation of $f_m(x)$ is given by*

$$\hat{f}_m(k) := \int_{-\infty}^{\infty} e^{-ikx} f_m(x)\, dx = \frac{\pi}{(m-1)!} \left(\prod_{\substack{l=m-2 \\ \text{step } -2}}^{1} (k^2 + l^2) \right) \cdot \begin{cases} \frac{1}{\cosh\left(\frac{\pi}{2}k\right)} & \text{if } m \text{ odd} \\ \frac{k}{\sinh\left(\frac{\pi}{2}k\right)} & \text{if } m \text{ even} \end{cases} . \tag{A.1.1}$$

Proof. It is convenient to first get rid of the infinite number of poles of the integrand. This is achieved by the variable transformation $r = e^{2x}$, which results in

$$\hat{f}_m(k) = 2^{m-1} \int_0^{\infty} \frac{\left(r^{1/2}\right)^{m-2-ik}}{(1+r)^m}\, dr . \tag{A.1.2}$$

This transformation comes at the expense of single-valuedness: the new integrand has only one simple pole at $r = -1$, but also a branch cut connecting the origin and infinity via, say, the negative imaginary axis. Since the integrand decays fast enough for $|r| \to \infty$, it is possible to revolve the integration contour once around the origin like the big hand of a watch. If we do this *counter-clockwise*, the whole integral picks up a residue, and the square-root acquires a monodromy of -1, which changes the numerator to

$$\left(-r^{1/2}\right)^{m-2-ik} = (-1)^m e^{k\pi} \left(r^{1/2}\right)^{m-2-ik} . \tag{A.1.3}$$

This contour manipulation yields the equation

$$\int_0^{\infty} \frac{\left(r^{1/2}\right)^{m-2-ik}}{(1+r)^m}\, dr = 2\pi i \operatorname{Res}_{z=-1} \left(\frac{\left(z^{1/2}\right)^{m-2-ik}}{(1+z)^m} \right) + (-1)^m e^{k\pi} \int_0^{\infty} \frac{\left(r^{1/2}\right)^{m-2-ik}}{(1+r)^m}\, dr . \tag{A.1.4}$$

This can now be solved for the integral:

$$\hat{f}_m(k) = \frac{1}{1-(-1)^m e^{k\pi}} 2^m i\pi \operatorname{Res}_{z=-1} \left(\frac{\left(z^{1/2}\right)^{m-2-ik}}{(1+z)^m} \right) . \tag{A.1.5}$$

The residue can be computed as the coefficient of the Laurent-expansion

$$\left(-z^{1/2}\right)^a = \sum_{n=0}^{\infty}(-1)^n i^a (1+z)^n \binom{\frac{a}{2}}{n} \tag{A.1.6}$$

for $n = m-1$, namely

$$\mathrm{Res}_{z=-1}\left(\frac{\left(z^{1/2}\right)^{m-2-ik}}{(1+z)^m}\right) = e^{\frac{\pi}{2}(k-im)}\binom{\frac{1}{2}(m-2-ik)}{m-1}. \tag{A.1.7}$$

Plugging this into (A.1.5) gives

$$\hat{f}_m(k) = 2^{m-1}\pi e^{i\frac{3}{2}\pi(1-m)}\binom{\frac{1}{2}(m-2-ik)}{m-1}\cdot\begin{cases}\frac{1}{\cosh\left(\frac{\pi}{2}k\right)} & \text{if } m \text{ odd}\\ \frac{k}{\sinh\left(\frac{\pi}{2}k\right)} & \text{if } m \text{ even}\end{cases}. \tag{A.1.8}$$

Finally, the following identity is straight-forward to prove by induction $m \to m+2$:

$$\prod_{j=1}^{m-1}(m-2j-ik) = e^{i\frac{3}{2}\pi(m-1)}\left(\prod_{\substack{l=m-2\\ \text{step } -2}}^{1}(k^2+l^2)\right)\cdot\begin{cases}1 & \text{if } m \text{ odd}\\ k & \text{if } m \text{ even}\end{cases} \tag{A.1.9}$$

Substituting (A.1.9) into (A.1.8) results in (A.1.1). □

A.2. Sokhotski integrals

The following proposition is the key ingredient in the proof of Proposition 1.2.23 in Appendix A.3. The proposition and its proof are adapted from [Gak, Ch. 1, §4], where a version of this theorem with contours of general shape but finite length is treated, and where the functions $\varphi(z,t)$ below are constant in z.

Proposition A.2.1. *Let $D \subseteq \mathbb{C}$ be a complex domain such that $\mathbb{S}_a \subseteq D$ for some $a > 0$. Let $\varphi : D \times \mathbb{R} \to \mathbb{C}$ be a function with the following properties:*

1. *(Analyticity) For every $t_0 \in \mathbb{R}$, the function $z \mapsto \varphi(z,t_0)$ is analytic in D.*
2. *(Hölder-continuity) There exist $0 < \alpha \leq 1$ and $C > 0$, such that for every $z_0 \in D$ the function $t \mapsto \varphi(z_0,t)$ is α-Hölder continuous with Hölder constant C.*
3. *(Decay) There exist $\mu > 0$ and $T > 0$, such that $|\varphi(z,t)| \leq |t-z|^{-\mu}$ for all $z \in D$ and $t \in \mathbb{R}$ with $|t-z| \geq T$.*
4. *(Local majorisation) For every $z_0 \in D \setminus \mathbb{R}$ there exist a neighbourhood $U \subseteq D \setminus \mathbb{R}$ and a function $M \in L_1(\mathbb{R})$, such that $|\varphi(z,t)| \leq |z-t| M(t)$ for all $z \in U$.*

5. *(Uniform convergence) The convergence* $\varphi(x \pm iy, t) \xrightarrow{y \searrow 0} \varphi(x,t)$ *is uniform in* $(x,t) \in \mathbb{R}^2$.

6. *(Boundedness)* $\sup_{(x,t)\in\mathbb{R}^2} |\varphi(x,t)| < \infty$.

Then the function

$$F(z) = \int_{-\infty}^{\infty} \frac{\varphi(z,t)}{t-z}\, dt \tag{A.2.1}$$

is analytic in $D \setminus \mathbb{R}$, *and there exist limiting functions* $F^{\pm} : \mathbb{R} \to \mathbb{C}$ *such that*

$$F(x \pm iy) \to F^{\pm}(x) \tag{A.2.2}$$

uniformly as $y \searrow 0$. *The functions* $F^{\pm}(x)$ *are bounded and satisfy*

$$F^{+}(x) - F^{-}(x) = 2i\pi\varphi(x,x) \qquad \text{for all} \quad x \in \mathbb{R} \ . \tag{A.2.3}$$

Proof.
• *$F(z)$ is analytic on $D \setminus \mathbb{R}$:* Conditions 2 and 3 ensure that the integrand in (A.2.1) is always in $L_1(\mathbb{R})$. Thus, $F(z)$ is well-defined. Analyticity of $F(z)$ in $D \setminus \mathbb{R}$ follows directly from lemma 1.2.19 together with condition 4.

• *The auxiliary function ψ:* Below we make frequent use of the following simple integral. Let $x, y, \eta, L \in \mathbb{R}$, $\eta \geq 0$, $\eta \pm x \leq L$, and suppose that $y \neq 0$ in case $\eta = 0$. Denote $B_\eta(x) = (x-\eta, x+\eta)$. Then

$$\int_{[-L,L]\setminus B_\eta(x)} \frac{1}{t-x-iy}\, dt \;=\; \log\frac{L-x-iy}{L+x+iy} \;+\; \begin{cases} \log\frac{\eta+iy}{\eta-iy} & ;\eta > 0 \\ i\pi\,\mathrm{sgn}(y) & ;\eta = 0 \end{cases} \ . \tag{A.2.4}$$

Here, the branch cut of the logarithm is placed along the negative real axis.

We now investigate the $y \searrow 0$ limit of $F(x \pm iy)$. To do so, we split $F(z)$ into two integrals by adding and subtracting a term in the integrand. Namely, for $z \in D \setminus \mathbb{R}$ we have

$$F(z) = \lim_{L\to\infty} \int_{-L}^{L} \frac{\varphi(z,t) - \varphi(z,\mathrm{Re}(z))}{t-z}\, dt + \varphi(z,\mathrm{Re}(z)) \lim_{L\to\infty} \int_{-L}^{L} \frac{1}{t-z}\, dt \ . \tag{A.2.5}$$

The improper integral

$$\psi(z) := \lim_{L\to\infty} \int_{-L}^{L} \frac{\varphi(z,t) - \varphi(z,\mathrm{Re}(z))}{t-z}\, dt \tag{A.2.6}$$

exists because by (A.2.4) the limit in the second summand of (A.2.5) exists. We obtain, for $z \in D \setminus \mathbb{R}$,

$$F(z) = \psi(z) + i\pi\, \varphi(z,\mathrm{Re}(z))\, \mathrm{sgn}(\mathrm{Im}(z)) \ . \tag{A.2.7}$$

Since both $F(z)$ and $\varphi(z,\mathrm{Re}(z))$ are continuous in $D \setminus \mathbb{R}$, $\psi(z)$ is also continuous in $D \setminus \mathbb{R}$.

We now claim that the integral and limit defining $\psi(z)$ in (A.2.6) also exist for $z \in \mathbb{R}$. To see this, first note that due to the Hölder condition (condition 2)

$$\left| \frac{\varphi(x,t) - \varphi(x,x)}{t-x} \right| \leq \frac{C}{|t-x|^{1-\alpha}} \qquad \forall x, t \in \mathbb{R} \ . \tag{A.2.8}$$

Hence, the integral

$$\int_{x-1}^{x+1} \frac{\varphi(x,t) - \varphi(x,x)}{t-x} \, dt \tag{A.2.9}$$

exists. On the other hand, by (A.2.4) and for $|x| + 1 < L$,

$$\int_{[-L,L]\setminus B_1(x)} \frac{\varphi(x,t) - \varphi(x,x)}{t-x} \, dt = \int_{[-L,L]\setminus B_1(x)} \frac{\varphi(x,t)}{t-x} \, dt - \varphi(x,x) \log \frac{L-x}{L+x} \ . \tag{A.2.10}$$

The integral in the first summand has a well-defined $L \to \infty$ limit by condition 3 Adding (A.2.9) and (A.2.10) shows that the limit and integral in (A.2.6) exist also for $z \in \mathbb{R}$, so that altogether ψ is defined on all of D.

• *Uniform convergence of ψ:* Next we study the continuity properties of ψ on D. Let us restrict $\psi(z)$ to lines parallel to the real axis. Namely, for $|y| < a$ we define $\psi^{[y]} : \mathbb{R} \to \mathbb{C}$ by $\psi^{[y]}(x) := \psi(x + iy)$. We will show that $\psi^{[y]}(x)$ converges to $\psi^{[0]}(x)$ uniformly as $y \to 0$ (from both sides). To do so, it is convenient to define the family of functions

$$\Delta^{[y]}(x,t) := \varphi(x+iy,t) - \varphi(x,t) \ . \tag{A.2.11}$$

Due to the uniform convergence condition on $\varphi(x+iy,t)$ (condition 5), $\Delta^{[y]}(x,t)$ converges to 0 uniformly in $(x,t) \in \mathbb{R}^2$ as $y \to 0$. In particular, for $|y|$ small enough, $\Delta^{[y]}(x,t)$ is bounded. Moreover, $\Delta^{[y]}(x,t)$ inherits the decay property (condition 3) from $\varphi(z,t)$. These properties will be used later in the proof.

Choose $\eta > 0$ such that $\eta < T$ and split the integration over the interval $[-L,L]$ (w.l.o.g. $|x|+\eta < L$) into the interval $B_\eta(x)$ and its complement $[-L,L] \setminus B_\eta(x)$. A straightforward computation yields

$$\begin{aligned}
&\psi^{[y]}(x) - \psi^{[0]}(x) \\
&= iy \int_{B_\eta(x)} \frac{\varphi(x,t) - \varphi(x,x)}{(t-x)(t-x-iy)} \, dt + iy \lim_{L\to\infty} \int_{[-L,L]\setminus B_\eta(x)} \frac{\varphi(x,t) - \varphi(x,x)}{(t-x)(t-x-iy)} \, dt \\
&\quad + \int_{B_\eta(x)} \frac{\Delta^{[y]}(x,t) - \Delta^{[y]}(x,x)}{t-x-iy} \, dt + \lim_{L\to\infty} \int_{[-L,L]\setminus B_\eta(x)} \frac{\Delta^{[y]}(x,t) - \Delta^{[y]}(x,x)}{t-x-iy} \, dt \ .
\end{aligned} \tag{A.2.12}$$

We will now show that all four integrals and limits exist and at the same time provide estimates for them. For the first three we compute, where (*) refers to the use of α-Hölder continuity (condition 2) and (**) to boundedness (condition 6) – we set $S := \sup_{(x,t)\in\mathbb{R}^2} |\varphi(x,t)|$,

$$\left| iy \int_{B_\eta(x)} \frac{\varphi(x,t) - \varphi(x,x)}{(t-x)(t-x-iy)} \, dt \right| \leq \int_{B_\eta(x)} \left| \frac{\varphi(x,t) - \varphi(x,x)}{t-x} \right| \frac{|y|}{|t-x-iy|} \, dt$$

$$\overset{(*)}{\leq}\ C \int_{B_\eta(x)} |t-x|^{\alpha-1}\, dt \ = \ 2C \int_0^\eta r^{\alpha-1}\, dr \ = \ \frac{2C\eta^\alpha}{\alpha}\ , \tag{A.2.13}$$

$$\begin{aligned}\left|\int_{B_\eta(x)} \frac{\Delta^{[y]}(x,t)-\Delta^{[y]}(x,x)}{t-x-iy}\, dt\right| & \\ \leq\ & \int_{B_\eta(x)} \frac{|\varphi(x+iy,t)-\varphi(x+iy,x)|+|\varphi(x,t)-\varphi(x,x)|}{|t-x|}\, dt \\ \overset{(*)}{\leq}\ & 2C \int_{B_\eta(x)} |t-x|^{\alpha-1}\, dt \ = \ 4C \int_0^\eta r^{\alpha-1}\, dr = \frac{4C\eta^\alpha}{\alpha}\ ,\end{aligned} \tag{A.2.14}$$

$$\begin{aligned}\left| iy \lim_{L\to\infty} \int_{[-L,L]\setminus B_\eta(x)} \frac{\varphi(x,t)-\varphi(x,x)}{(t-x)(t-x-iy)}\, dt\right| \ \leq\ |y| \int_{\mathbb{R}\setminus B_\eta(x)} \frac{|\varphi(x,t)|+|\varphi(x,x)|}{|t-x|^2}\, dt & \\ \overset{(**)}{\leq}\ 4S|y|\eta^{-1}\ . &\end{aligned} \tag{A.2.15}$$

Now let us turn to the fourth integral, which is slightly more involved. With the help of the decay condition on $\Delta^{[y]}(x,t)$ and (A.2.4) we can rewrite it as

$$\begin{aligned}\lim_{L\to\infty} \int_{[-L,L]\setminus B_\eta(x)} & \frac{\Delta^{[y]}(x,t)-\Delta^{[y]}(x,x)}{t-x-iy}\, dt \\ &= \int_{\mathbb{R}\setminus B_\eta(x)} \frac{\Delta^{[y]}(x,t)}{t-x-iy}\, dt \ -\ \Delta^{[y]}(x,x) \log\frac{\eta+iy}{\eta-iy}\ .\end{aligned} \tag{A.2.16}$$

To estimate the integral over $\mathbb{R}\setminus B_\eta(x)$, we split it as follows, for $R>T$,

$$\int_{\mathbb{R}\setminus B_\eta(x)} \frac{\Delta^{[y]}(x,t)}{t-x-iy}\, dt = \int_{\mathbb{R}\setminus B_R(x)} \frac{\Delta^{[y]}(x,t)}{t-x-iy}\, dt + \int_{B_R(x)\setminus B_\eta(x)} \frac{\Delta^{[y]}(x,t)}{t-x-iy}\, dt\ . \tag{A.2.17}$$

We now estimate the two integrals separately:

$$\begin{aligned}\left|\int_{\mathbb{R}\setminus B_R(x)} \frac{\Delta^{[y]}(x,t)}{t-x-iy}\, dt\right| &\leq \int_{\mathbb{R}\setminus B_R(x)} \left|\frac{\Delta^{[y]}(x,t)}{t-x}\right|\, dt \\ &\leq \int_{\mathbb{R}\setminus B_R(x)} \frac{1}{|t-x|^{1+\mu}}\, dt = 2\int_R^\infty \frac{1}{r^{1+\mu}}\, dr = \frac{2}{\mu R^\mu}\end{aligned} \tag{A.2.18}$$

$$\begin{aligned}\left|\int_{B_R(x)\setminus B_\eta(x)} \frac{\Delta^{[y]}(x,t)}{t-x-iy}\, dt\right| &\leq \int_{B_R(x)\setminus B_\eta(x)} \left|\frac{\Delta^{[y]}(x,t)}{t-x}\right|\, dt \leq \frac{1}{\eta}\int_{x-R}^{x+R} |\Delta^{[y]}(x,t)|\, dt \\ &\leq \frac{2R}{\eta} \sup_{t\in\mathbb{R}} |\Delta^{[y]}(x,t)| \leq \frac{2R}{\eta} \sup_{t,x\in\mathbb{R}} |\Delta^{[y]}(x,t)|\end{aligned} \tag{A.2.19}$$

Finally, we remark that with our choice of branch cut for the logarithm,

$$\left|\log\tfrac{\eta+iy}{\eta-iy}\right| = 2\,|\arg(\eta+iy)| \leq \pi\ . \tag{A.2.20}$$

Assembling all of the above estimates, we obtain:

$$\left|\psi^{[y]}(x) - \psi^{[0]}(x)\right| \leq \frac{6C\eta^\alpha}{\alpha} + \frac{4S|y|}{\eta} + \frac{2}{\mu R^\mu} + \left(\frac{2R}{\eta} + \pi\right) \sup_{x,t\in\mathbb{R}} |\Delta^{[y]}(x,t)| \tag{A.2.21}$$

To establish uniform convergence, we need to show that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $|y| < \delta$ and all $x \in \mathbb{R}$ we have $|\psi^{[y]}(x) - \psi^{[0]}(x)| \leq \varepsilon$. To find δ, we choose η and R in the above estimate appropriately.

Choose η such that the first term in (A.2.21) equals $\varepsilon/4$:

$$\eta = \left(\frac{\alpha\varepsilon}{24C}\right)^{\frac{1}{\alpha}} . \tag{A.2.22}$$

The second term is smaller than $\varepsilon/4$ provided $|y| < \delta_1$, where

$$\delta_1 = \frac{\eta\,\varepsilon}{16\,S} = \frac{1}{16S}\left(\frac{\alpha}{24C}\right)^{\frac{1}{\alpha}} \varepsilon^{1+\frac{1}{\alpha}} . \tag{A.2.23}$$

The third term is smaller than $\varepsilon/4$ if we set

$$R = \left(\frac{8}{\mu\varepsilon}\right)^{\frac{1}{\mu}} . \tag{A.2.24}$$

Finally, we remember that $\Delta^{[y]}(x,t) \xrightarrow{y\to 0} 0$ uniformly in x and t. Hence, there exists a $\delta_2 > 0$ such that for all $|y| < \delta_2$,

$$\sup_{x,t\in\mathbb{R}} |\Delta^{[y]}(x,t)| \leq \frac{\varepsilon}{4}\left(\frac{2R}{\eta} + \pi\right)^{-1} = \frac{\varepsilon}{4}\left(2\left(\frac{24C}{\alpha\varepsilon}\right)^{\frac{1}{\alpha}}\left(\frac{8}{\mu\varepsilon}\right)^{\frac{1}{\mu}} + \pi\right)^{-1} . \tag{A.2.25}$$

This makes the last term smaller than $\varepsilon/4$. Setting $\delta := \min(\delta_1, \delta_2)$, this proves uniform convergence $\psi^{[y]} \xrightarrow{y\to 0} \psi^{[0]}$.

- *Uniform convergence of F and relation of $F^\pm$:* The claim of uniform convergence of (A.2.2) and formula (A.2.3) now both follow from (A.2.7).

- *Boundedness of $F^\pm$:* It is enough to provide a bound for $\psi^{[0]}(x)$. This can be achieved as follows. Split the integral:

$$\begin{aligned}\int_{-L}^{L} \frac{\varphi(x,t) - \varphi(x,x)}{t-x}\,dt \;=\; &\int_{B_1(x)} \frac{\varphi(x,t) - \varphi(x,x)}{t-x}\,dt \;+\; \int_{[-L,L]\setminus B_1(x)} \frac{\varphi(x,t)}{t-x}\,dt \\ &-\; \varphi(x,x)\int_{[-L,L]\setminus B_1(x)} \frac{1}{t-x}\,dt\end{aligned} \tag{A.2.26}$$

The first summand can be estimated using the Hölder inequality (condition 2)

$$\left|\int_{B_1(x)} \frac{\varphi(x,t) - \varphi(x,x)}{t-x}\,dt\right| \leq C\int_{-1}^{1} |r|^{\alpha-1}\,dr = \frac{2C}{\alpha} \tag{A.2.27}$$

For the second integral, we make use of both boundedness (condition 6, where as above we denote $S := \sup_{(x,t)\in\mathbb{R}^2} |\varphi(x,t)|$) and the decay property (condition 3, w.l.o.g. $1 < T < L$):

$$\begin{aligned}\left|\int_{[-L,L]\setminus B_1(x)} \frac{\varphi(x,t)}{t-x}\, dt\right| &\leq \int_{B_T(x)\setminus B_1(x)} \left|\frac{\varphi(x,t)}{t-x}\right| dt + \int_{[-L,L]\setminus B_T(x)} \left|\frac{\varphi(x,t)}{t-x}\right| dt \\ &\leq \int_{B_T(x)\setminus B_1(x)} |\varphi(x,t)|\ dt + \int_{[-L,L]\setminus B_T(x)} |t-x|^{-\mu-1}\ dt \\ &\leq 2TS + 2\int_1^\infty r^{-\mu-1}\, dr = 2TS + \frac{1}{\mu} \end{aligned} \tag{A.2.28}$$

By (A.2.4), the third integral is simply bounded as follows:

$$\left|-\varphi(x,x)\int_{[-L,L]\setminus B_1(x)} \frac{1}{t-x}\, dt\right| \leq S\left|\log\left(\frac{L-x}{L+x}\right)\right| \tag{A.2.29}$$

Altogether, we obtain the bound

$$\left|\int_{-L}^{L} \frac{\varphi(x,t)-\varphi(x,x)}{t-x}\, dt\right| \leq \frac{2C}{\alpha} + \left(2T + \left|\log\left(\frac{L-x}{L+x}\right)\right|\right) S + \frac{1}{\mu}\ . \tag{A.2.30}$$

Since this bound itself converges as $L \to \infty$, we obtain a bound for $\left|\psi^{[0]}(x)\right|$. □

Write $D^+ := \{z \in D | \mathrm{Im}(z) > 0\}$, $D^- := \{z \in D | \mathrm{Im}(z) < 0\}$ and $\tilde{D}^\pm := D^\pm \cup \mathbb{R}$. Consider the functions

$$\tilde{F}^\pm : \tilde{D}^\pm \to \mathbb{C} \quad , \quad \tilde{F}^\pm(z) := \begin{cases} F(z) & ;\ \mathrm{Im}(z) \neq 0 \\ F^\pm(z) & ;\ \mathrm{Im}(z) = 0 \end{cases}\ . \tag{A.2.31}$$

Corollary A.2.2. *$\tilde{F}^\pm$ is a continuous extension of F from $D^\pm$ to $\tilde{D}^\pm$.*

Proof. It suffices to show that $\psi(z)$ is continuous in D. We know already that it is continuous in $D \setminus \mathbb{R}$. Now let us show that it is continuous in $x_0 \in \mathbb{R}$.

Let $\varepsilon > 0$. By uniform convergence $\psi^{[y]}(x) \to \psi^{[0]}(x)$ there is a $\delta_1 > 0$ such that for all $|y| < \delta_1$ and all $x \in \mathbb{R}$ we have $|\psi^{[y]}(x) - \psi^{[0]}(x)| < \varepsilon/2$. By continuity of $F^\pm$ on $\mathbb{R}$ there is $\delta_2 > 0$ such that for all x with $|x - x_0| < \delta_2$ we have $|\psi^{[0]}(x) - \psi^{[0]}(x_0)| < \varepsilon/2$.

Take $\delta := \min(\delta_1, \delta_2)$. For $z = x + iy \in D$ with $|z - x_0| < \delta$ we have

$$|\psi(z) - \psi(x_0)| = \left|\psi^{[y]}(x) - \psi^{[0]}(x_0)\right| \leq |\psi^{[y]}(x) - \psi^{[0]}(x)| + |\psi^{[0]}(x) - \psi^{[0]}(x_0)| \leq \varepsilon. \tag{A.2.32}$$

□

A.3. Proof of Proposition 1.2.23

Here we prove Proposition 1.2.23 as a special case of Proposition A.2.1 and Corollary A.2.2. Recall the setting of Proposition 1.2.23:

- $D = \mathbb{S}_a$ for some $a > 0$.
- $\varphi(z,t) = h(z-t)g(t)$ where $h : \mathbb{S}_a \to \mathbb{C}$ is analytic and $g : \mathbb{R} \to \mathbb{C}$ is bounded and Hölder continuous.
- $zh(z)$ and $\frac{d}{dz}h(z)$ are bounded in $\mathbb{S}_a$.

Let us check the conditions on $\varphi(z,t)$ one by one. Denote by G the bound of $g(t)$ and by H the bound of $zh(z)$.

Condition 1 This is obvious since $h(z)$ is analytic.

Condition 2 Let C, α be constants expressing the Hölder-continuity of g:

$$\forall t, t' \in \mathbb{R} \ : \ |g(t') - g(t)| \le C|t-t'|^\alpha \ . \tag{A.3.1}$$

Boundedness of $\frac{d}{dz}h(z)$ implies that in particular the derivative $\frac{d}{dx}h(x+iy)$ in the real direction is bounded for every $y < (-a,a)$. As a consequence, for any given $z_0 \in \mathbb{S}_a$ the function $t \mapsto h(z_0 - t)$ is Lipschitz-continuous with Lipschitz constant $L := \sup_{z\in\mathbb{S}_a}(\frac{d}{dz}h(z))$. Since $zh(z)$ is bounded, we also have boundedness of h on $\mathbb{S}_a$: $|h(z)| \le B$ for some $B \ge 0$. Accordingly,

$$\begin{aligned}
|\varphi(z_0,t) - \varphi(z_0,t')| &= |h(z_0-t)g(t) - h(z_0-t')g(t')| \\
&\le |h(z_0-t) - h(z_0-t')|\,|g(t)| + |h(z_0-t')|\,|g(t)-g(t')| \\
&\le BC\,|t-t'|^\alpha + \begin{cases} GL\,|t-t'| & ; \ |t-t'| \le 1 \\ 2BG & ; \ |t-t'| > 1 \end{cases} \ .
\end{aligned} \tag{A.3.2}$$

Note that B, G, L are all independent of z_0. Hence there is an $M > 0$, independent of z_0, such that for all $t, t' \in \mathbb{R}$:

$$|\varphi(z_0,t) - \varphi(z_0,t')| \le M|t-t'|^\alpha \ . \tag{A.3.3}$$

Condition 3 For all $z_0 \in \mathbb{S}_a$, we have the inequality

$$|\varphi(z_0,t)| \le G|h(z_0-t)| \le GH\,|z_0-t|^{-1} \ . \tag{A.3.4}$$

Fix $0 < \mu < 1$. There exists a $T > 0$, independent of z_0, such that $GH\,|z_0-t|^{-1} \le |z_0-t|^{-\mu}$ whenever $|z_0 - t| > T$.

Condition 4 For given $z_0 \in D \setminus \mathbb{R}$, set $\rho := \frac{1}{2}\mathrm{Im}(z_0)$ and $U := \mathbb{S}_a \cap B_\rho(z_0)$, where $B_\rho(z_0)$ is the open ball of radius ρ with center z_0. Then for all $z \in U$, we have

$$|\varphi(z,t)| \le G|h(z-t)| \le GH\frac{1}{|z-t|} = |z-t|\frac{GH}{|z-t|^2} \ . \tag{A.3.5}$$

Now set

$$M(t) := \begin{cases} \frac{GH}{|\mathrm{Re}(z_0)-\rho-t|^2} & \text{if } t < \mathrm{Re}(z_0) - 2\rho \\ \frac{GH}{\rho^2} & \text{if } t \in [\mathrm{Re}(z_0) - 2\rho, \mathrm{Re}(z_0) + 2\rho] \\ \frac{GH}{|\mathrm{Re}(z_0)+\rho-t|^2} & \text{if } t > \mathrm{Re}(z_0) + 2\rho \end{cases} . \tag{A.3.6}$$

Then $M \in L_1(\mathbb{R})$ and one quickly checks that

$$\frac{GH}{|z-t|^2} \leq M(t) . \tag{A.3.7}$$

Condition 5 Since $\varphi(x \pm iy, t) = h(x \pm iy - t)g(t)$, the condition is satisfied if $h(x+iy) \xrightarrow{y\to 0} h(x)$ uniformly. This is easily established with the following lemma.

Lemma A.3.1. *Let $a > 0$, $b \in \mathbb{R}$, and set $D := \mathbb{S}_a \cap \{z \in \mathbb{C} | \mathrm{Re}(z) > b\}$. Let $f : D \to \mathbb{C}$ be an analytic function such that $zf(z)$ and $\frac{d}{dz}f(z)$ are bounded. Then $f(x+iy) \xrightarrow{y\to Y} f(x+iY)$ uniformly on $[t, \infty)$ for any $Y \in (-a, a)$ and any $t > b$.*

Proof. Pointwise convergence is clear by continuity. Now we claim that the convergence is uniform. Let $\varepsilon > 0$. Set $x_0 := \frac{2B}{\varepsilon}$, where $B > 0$ is the bound of $zf(z)$. Without loss of generality, assume $b < t < x_0$. Then for all $x \geq x_0$ one has

$$\begin{aligned} |f(x+iy) - f(x+iY)| &\leq |f(x+iy)| + |f(x+iY)| \\ &\leq \frac{B}{|x+iy|} + \frac{B}{|x+iY|} \leq \frac{2B}{|x|} \leq \frac{2B}{x_0} = \varepsilon \end{aligned}$$

for all $|y| < a$.

Thus, it remains to be shown that convergence on the compact interval $[t, x_0]$ is uniform. Since $f(z)$ is bounded, the family of functions $f^{[y]}(x) := f(x+iy)$ on the interval $[t, x_0]$ is uniformly bounded. Moreover, boundedness of $\frac{d}{dz}f(z)$ means that $\frac{d}{dx}f^{[y]}(x)$ are uniformly bounded. But this implies that $f^{[y]}(x)$ are equicontinuous. Thus, we can apply the Arzela-Ascoli theorem: for every sequence $y_n \to Y$, the sequence of functions $f^{[y_n]}(x)$ has a uniformly convergent subsequence. Now assume that $f^{[y]}(x)$ do not converge uniformly on $[t, x_0]$. Then there exists a sequence $u_n \to Y$ and a sequence x_n of points in $[t, x_0]$ such that $\left|f^{[u_n]}(x_n) - f(x_n)\right| \geq \varepsilon$ for all n. But then $f^{[u_n]}(x)$ has no uniformly convergent subsequence, which is a contradiction. □

Condition 6 This is again obvious, since both $h(z)$ and $g(t)$ are bounded.

This completes the proof of Proposition 1.2.23.

Part II.

Twisted topological defect operators in conformal field theory

4. Categorical and algebraic preliminaries

4.1. Tensor categories

For an introduction to general category theory, we refer to [McL]. This chapter provides a quick introduction to categories with a tensor product.

Definition 4.1.1. A *monoidal category* is a category $\mathcal{C}$ equipped with a bifunctor

$$\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C} \ , \qquad \textit{(the tensor product)}$$

with a distinguished object

$$\mathbf{1} \in \mathcal{C} \ , \qquad \textit{(the tensor unit)}$$

and with natural isomorphisms

$$\begin{aligned} &\alpha : ((-) \otimes (-)) \otimes (-) \longrightarrow (-) \otimes ((-) \otimes (-)) && \textit{(the associator)} \\ &\lambda : \mathbf{1} \otimes (-) \longrightarrow (-) && \textit{(the left unitor)} \\ &\rho : (-) \otimes \mathbf{1} \longrightarrow (-) && \textit{(the right unitor)} \end{aligned}$$

such that the *pentagon diagram*

$$\begin{tikzcd} & ((U \otimes V) \otimes W) \otimes T \arrow[dl, "\alpha_{U,V,W} \otimes \mathrm{id}_T"'] \arrow[dr, "\alpha_{U \otimes V,W,T}"] & \\ (U \otimes (V \otimes W)) \otimes T \arrow[d, "\alpha_{U,V \otimes W,T}"'] & & (U \otimes V) \otimes (W \otimes T) \arrow[d, "\alpha_{U,V,W \otimes T}"] \\ U \otimes ((V \otimes W) \otimes T) \arrow[rr, "\mathrm{id}_U \otimes \alpha_{V,W,T}"] & & U \otimes (V \otimes (W \otimes T)) \end{tikzcd} \tag{4.1.1}$$

and the triangle diagram

$$\begin{tikzcd} (U \otimes \mathbf{1}) \otimes V \arrow[rr, "\alpha_{U,\mathbf{1},V}"] \arrow[dr, "\rho_U \otimes \mathrm{id}_V"'] & & U \otimes (\mathbf{1} \otimes V) \arrow[dl, "\mathrm{id}_U \otimes \lambda_V"] \\ & U \otimes V & \end{tikzcd} \tag{4.1.2}$$

commute for all $U, V, W, T \in \mathcal{C}$. A monoidal category whose coherence maps (associators and unitors) are identity natural transformations is called *strict*.

A *monoidal functor* between monoidal categories $\mathcal{C}$ and $\mathcal{D}$ is a functor $G : \mathcal{C} \to \mathcal{D}$ together with a natural isomorphism $\mu : G(-) \otimes_{\mathcal{D}} G(-) \to G(- \otimes_{\mathcal{C}} -)$ and an isomorphism $\varepsilon : \mathbf{1}_{\mathcal{C}} \to G(\mathbf{1}_{\mathcal{D}})$, which are compatible with associator and unitors in the obvious way.

Remark 4.1.2.
1. According to the McLane coherence theorem (see e.g. [EGNO, Ch. 2.9]), every monoidal category is monoidally equivalent to a strict one. Hence, there is no loss of generality in assuming that $\mathcal{C}$ is strict right from the beginning. We will usually do that, but keep the coherence maps explicit in this section.
2. Monoidal categories are sometimes also called *tensor categories*. However, some authors distinguish between the two notions. In this case, tensor categories are for example assumed to be abelian or $\mathbb{K}$-linear (in addition to the monoidal structure). We will not make this distinction.

Example 4.1.3.
1. The category $\mathrm{Vect}_{\mathbb{K}}$ of finite dimensional vector spaces over a field $\mathbb{K}$ together with the ordinary tensor product over $\mathbb{K}$ is a monoidal category. The tensor unit in this case is simply $\mathbf{1} = \mathbb{K}$, the associator $\alpha_{U,V,W}$ acts, for $u \in U$, $v \in V$ and $w \in W$, by $(u \otimes v) \otimes w \mapsto u \otimes (v \otimes w)$, and the unitors λ_U and ρ_U act, for $x \in \mathbb{K}$, by $x \otimes u \mapsto xu$ and $u \otimes x \mapsto xu$. More generally, the category $\mathrm{Rep}(\mathfrak{g})$ of finite dimensional representations of a Lie algebra $\mathfrak{g}$ over $\mathbb{K}$, together with the tensor product over $\mathfrak{g}$, is a monoidal category with tensor unit the one-dimensional representation with vanishing $\mathfrak{g}$-action.
2. Let $(G, \circ)$ be a group. The category Vect^G of finite-dimensional G-graded vector spaces has objects of the form $V = \oplus_{g \in G} V_g$ (where $V_g \in \mathrm{Vect}_{\mathbb{K}}$ are called the *homogeneous subspaces*) and morphisms $\mathrm{Hom}_{\mathrm{Vect}^G}(U, V) = \oplus_{g \in G} \mathrm{Hom}_{\mathbb{K}}(U_g, V_g)$ grade-preserving linear maps. It is a monoidal category with tensor product $(U \otimes V)_g = \oplus_{h \in G} U_h \otimes V_{h^{-1} \circ g}$.
3. The category 2Cob with objects closed oriented 1-dimensional manifolds and morphisms 2-dimensional diffeomorphism classes of oriented cobordisms with boundary parametrisation between them (for a precise definition see [Koc]) is a strict monoidal category with tensor product the disjoint union and tensor unit the empty manifold. It is not a tensor category in the above sense.

Definition 4.1.4. Let $\mathcal{C}$ be an abelian category. Its *Grothendieck group* $K_0(\mathcal{C})$ is the free group over the isomorphism classes $[U]$ of objects $U \in \mathcal{C}$, modulo the relation $[U] + [W] = [V]$ for every short exact sequence $0 \to U \to V \to W \to 0$. Now let $\mathcal{C}$ be additionally monoidal. If the tensor product is exact, then the multiplication $[U] \cdot [V] := [U \otimes V]$ defines a ring structure on $K_0(\mathcal{C})$, which is called the *Grothendieck ring*.

In a monoidal category $\mathcal{C}$ one can also define the opposite tensor product bifunctor $\otimes^{\mathrm{op}} : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ by the assignment $(U, V) \mapsto V \otimes U$. The following definition provides a notion of generalised commutativity of the tensor product $\otimes$.

Definition 4.1.5. A monoidal category $\mathcal{C}$ is called *braided* if it comes equipped with a natural isomorphism

$$c : \otimes \longrightarrow \otimes^{\mathrm{op}} \,, \tag{\textit{the braiding}}$$

such that the *hexagon diagrams*

$$\begin{tikzcd}
& U \otimes (V \otimes W) \arrow[r, "c_{U,V\otimes W}"] & (V \otimes W) \otimes U \arrow[dr, "\alpha_{V,W,U}"] & \\
(U \otimes V) \otimes W \arrow[ur, "\alpha_{U,W,W}"] \arrow[dr, "c_{U,V} \otimes \mathrm{id}_W"'] & & & V \otimes (W \otimes U) \\
& (V \otimes U) \otimes W \arrow[r, "\alpha_{V,U,W}"] & V \otimes (U \otimes W) \arrow[ur, "\mathrm{id}_V \otimes c_{U,W}"'] &
\end{tikzcd} \tag{4.1.3}$$

$$\begin{tikzcd}
& U \otimes (V \otimes W) \arrow[r, "c^{-1}_{V\otimes W,U}"] & (V \otimes W) \otimes U \arrow[dr, "\alpha_{V,W,U}"] & \\
(U \otimes V) \otimes W \arrow[ur, "\alpha_{U,W,W}"] \arrow[dr, "c^{-1}_{V,U} \otimes \mathrm{id}_W"'] & & & V \otimes (W \otimes U) \\
& (V \otimes U) \otimes W \arrow[r, "\alpha_{V,U,W}"] & V \otimes (U \otimes W) \arrow[ur, "\mathrm{id}_V \otimes c^{-1}_{W,U}"'] &
\end{tikzcd} \tag{4.1.4}$$

commute for all $U, V, W \in \mathcal{C}$.

A monoidal functor $G : \mathcal{C} \to \mathcal{D}$ between braided monoidal categories is called *braided* if it commutes with the braiding.

Example 4.1.6.

1. Both $\mathrm{Vect}_{\mathbb{K}}$ and $\mathrm{Rep}(\mathfrak{g})$ are braided. The braiding $c_{U,V} : U \otimes V \to V \otimes U$ simply acts by $u \otimes v \mapsto v \otimes u$ on elementary tensors. In fact, these two categories are actually *symmetric monoidal*, which is to say that $c_{U,V} = c^{-1}_{V,U}$ for all $U, V \in \mathcal{C}$.

2. If G is an abelian group, then the category Vect^G can be equipped with a braiding. In fact, every bi-character $\chi : G \times G \to \mathbb{K}^\times$ gives rise to a braiding, which acts on homogeneous elements $v \in V_g$ and $u \in U_h$ by $u \otimes v \mapsto \chi(g,h) v \otimes u$ [Pa1].

3. Also the category 2Cob is braided, whereby the braiding of a circle with another circle is given by two disjoint cylinders whose opposite boundary components are parametrised by the first and the second circle, and the second and the first circle, respectively. Since these cobordisms are not embedded anywhere, there is no notion of over- or undercrossing of the cylinders. Thus, 2Cob is symmetric monoidal.

For a fixed object $U \in \mathcal{C}$ and a monoidal subcategory $\mathcal{D} \subseteq \mathcal{C}$, denote by $U \otimes (-)\big|_{\mathcal{D}} : \mathcal{D} \to \mathcal{C}$ the functor defined by the assignment $V \mapsto U \otimes V$ for all $V \in \mathcal{D}$. Similarly, define $(-)\big|_{\mathcal{D}} \otimes U$ by $V \mapsto V \otimes U$. The following definition can be found for example in [Scha].

Definition 4.1.7. Let $\mathcal{C}$ be a monoidal category, and $\mathcal{D} \subseteq \mathcal{C}$ a monoidal subcategory. The *monoidal centraliser of* $\mathcal{D}$ *in* $\mathcal{C}$, denoted by $\mathcal{Z}(\mathcal{D};\mathcal{C})$, is given by

- *objects*: pairs $(U, \sigma_{U,-})$, where $U \in \mathcal{C}$ and
$$\sigma_{U,-} : U \otimes (-)\big|_{\mathcal{D}} \longrightarrow (-)\big|_{\mathcal{D}} \otimes U \qquad \textit{(the half-braiding)}$$
is a natural isomorphism such that the first hexagon diagram (4.1.3), with c replaced by σ, commutes for all $V, W \in \mathcal{D}$.
- *morphisms*: $f \in \mathrm{Hom}_{\mathcal{C}}(U, V)$, such that the diagram
$$\begin{tikzcd} U \otimes W \arrow[r, "\sigma_{U,W}"] \arrow[d, "f\otimes \mathrm{id}_W"'] & W \otimes U \arrow[d, "\mathrm{id}_W \otimes f"] \\ V \otimes W \arrow[r, "\sigma_{V,W}"] & W \otimes V \end{tikzcd} \tag{4.1.5}$$
commutes for all $W \in \mathcal{D}$.

The special case $\mathcal{Z}(\mathcal{C}) := \mathcal{Z}(\mathcal{C};\mathcal{C})$ is called the *monoidal center of* $\mathcal{C}$.

Definition 4.1.8. Let $\mathcal{C}$ be a monoidal category and $U \in \mathcal{C}$. A *right dual* to U is an object $U^\vee \in \mathcal{C}$ together with morphisms
$$\mathrm{ev}_U : U^\vee \otimes U \longrightarrow \mathbf{1} \qquad \textit{(the evaluation)}$$
$$\mathrm{coev}_U : \mathbf{1} \longrightarrow U \otimes U^\vee \qquad \textit{(the coevaluation)}$$
such that the compositions
$$U \xrightarrow{\lambda_U^{-1}} \mathbf{1} \otimes U \xrightarrow{\mathrm{coev}_U \otimes \mathrm{id}_U} (U \otimes U^\vee) \otimes U \xrightarrow{\alpha_{U,U^\vee,U}} U \otimes (U^\vee \otimes U) \xrightarrow{\mathrm{id}_U \otimes \mathrm{ev}_U} U \otimes \mathbf{1} \xrightarrow{\rho_U} U$$
$$U^\vee \xrightarrow{\rho_{U^\vee}^{-1}} U^\vee \otimes \mathbf{1} \xrightarrow{\mathrm{id}_{U^\vee} \otimes \mathrm{coev}_U} U^\vee \otimes (U \otimes U^\vee) \xrightarrow{\alpha^{-1}_{U^\vee,U,U^\vee}} (U^\vee \otimes U) \otimes U^\vee \xrightarrow{\mathrm{ev}_U \otimes \mathrm{id}_{U^\vee}} \mathbf{1} \otimes U^\vee \xrightarrow{\lambda_{U^\vee}} U^\vee \tag{4.1.6}$$
are equal to id_U and $\mathrm{id}_{U^\vee}$, respectively. A left dual ${}^\vee U$ is defined similarly, with maps $\widetilde{\mathrm{ev}}_U : U \otimes {}^\vee U \longrightarrow \mathbf{1}$ and $\widetilde{\mathrm{coev}}_U : \mathbf{1} \longrightarrow {}^\vee U \otimes U$.

Duals are unique up to a unique isomorphism which is compatible with evaluation and coevaluation. In a braided category, existence of right duals implies the existence of left duals, and they are automatically isomorphic.

Definition 4.1.9. A monoidal category $\mathcal{C}$ is called *rigid* if every object in $\mathcal{C}$ has a left and right dual.

In a rigid category, composition with evaluation and coevaluation provides canonical isomorphisms

$$\mathrm{Hom}(U \otimes V, W) = \mathrm{Hom}(U, W \otimes V^\vee) \ , \quad \mathrm{Hom}(U, V \otimes W) = \mathrm{Hom}(V^\vee \otimes U, W) \ . \quad (4.1.7)$$

For any morphism $f : U \to V$, this yields a right dual morphism $f^\vee : V^\vee \to U^\vee$ given by $f^\vee = \lambda_{U^\vee} \circ (\mathrm{ev}_V \otimes \mathrm{id}_{U^\vee}) \circ \alpha^{-1}_{V^\vee V U^\vee} \circ (\mathrm{id}_{V^\vee} \otimes f \otimes \mathrm{id}_{U^\vee}) \circ (\mathrm{id}_{V^\vee} \otimes \mathrm{coev}_U) \circ \rho^{-1}_{V^\vee}$. Similarly, there is a left dual morphism ${}^\vee f : {}^\vee V \to {}^\vee U$.

Moreover, if $\mathcal{C}$ is rigid abelian, then using the isomorphisms (4.1.7) it can be shown that the tensor product functor is exact, i.e. preserves exact sequences.

Example 4.1.10. The categories in Example 4.1.3 are rigid, and their left and right duals agree:

1. The dual of a finite dimensional vector space $V \in \mathrm{Vect}_\mathbb{K}$ is its ordinary dual space $V^\star$. In terms of a basis b_i of V and the dual basis φ_i of $V^\star$, the evaluation is given by $\varphi_i \otimes b_j \mapsto \delta_{ij}$, and the coevaluation by $1 \mapsto \sum_i b_i \otimes \varphi_i$. (Obviously, there is a problem when trying to define the coevaluation for infinite dimensional vector spaces. In fact, the category of infinite dimensional vector spaces is not rigid.) More generally, the dual of a finite-dimensional Lie algebra representation $R \in \mathrm{Rep}(\mathfrak{g})$ is its contragradient representation.

2. The dual of a G-graded vector space $V \in \mathrm{Vect}^G$ is defined to be the object $V^\vee = \oplus_{g \in G} V_g^\star$, called the *graded dual* of V. Evaluation and coevaluation are given by the grade-wise evaluation and coevaluation from $\mathrm{Vect}_\mathbb{K}$.

3. The dual of an oriented 1-dimensional manifold $(M, +) \in 2\mathrm{Cob}$ is the same manifold with opposite orientation $(M, -)$. Evaluation and coevaluation are cylinders connecting $(M, +)$ and $(M, -)$, once with two incoming and once with two outgoing boundaries.

4.2. Algebraic structures in tensor categories

Algebras, co-algebras, Hopf algebras and Frobenius algebras

Recall that an (associative, unital) *algebra* over a field $\mathbb{K}$ is a vector space A together with $\mathbb{K}$-linear maps $\mu : A \otimes A \to A$ (the multiplication) and $\eta : \mathbb{K} \to A$ (the unit), such that $\mu \circ (\mu \otimes \mathrm{id}_A) = \mu \circ (\mathrm{id}_A \otimes \mu)$ (associativity) and $\mu \circ (\eta \otimes \mathrm{id}_A) = \mathrm{id}_A = \mu \circ (\mathrm{id}_A \otimes \eta)$ (unitality) hold. A *co-algebra* C over $\mathbb{K}$ is defined in the same way, with $\mathbb{K}$-linear maps $\Delta : C \to C \otimes C$ (the co-multiplication) and $\varepsilon : C \to \mathbb{K}$ (the co-unit) satisfying analogous relations.

A (co-)algebra homomorphism is a $\mathbb{K}$-linear map between (co-)algebras which is compatible with the (co-)multiplications and (co-)units in the obvious way. If B is both an algebra as well as a co-algebra, and the maps Δ and ε are algebra homomorphisms, then B

is called a *bi-algebra*. The category $\mathrm{Rep}(B)$ of representations of B is naturally a monoidal category, with the tensor product induced via the co-multiplication Δ.

If H is a bi-algebra and $S : H \to H$ a $\mathbb{K}$-linear map satisfying $\mu \circ (S \otimes \mathrm{id}_H) \circ \Delta = \eta \circ \varepsilon = \mu \circ (\mathrm{id}_H \otimes S) \circ \Delta$, then H is called a *Hopf algebra*. It can be shown that S (the antipode) is an anti-algebra homomorphism, and that it is unique. The category of finite-dimensional representations $\mathrm{Rep}(H)$ is naturally rigid monoidal (but not necessarily braided). For more on Hopf algebras, see for example [Kas, CP1].

Example 4.2.1. The universal envelopping algebra $U(\mathfrak{g})$ of a Lie algebra is a Hopf algebra with co-multiplication $\Delta(x) = 1 \otimes x + x \otimes 1$, co-unit $\varepsilon(x) = 0$ and antipode $S(x) = -x$ for all $x \in \mathfrak{g}$. Moreover, various kinds of q-deformations thereof are also Hopf algebras (called *quantum groups*). Unlike $\mathrm{Rep}(\mathfrak{g})$, the representation category of a quantum group is usually not symmetric monoidal anymore. It may not even be braided. There is a special class of Hopf algebras, called quasi-triangular, which give rise to a braiding [Kas, Ch. VIII.2].

The notion of bi-algebra is just one way in which co-algebra and algebra structure can be compatible with one another.

If A is both an algebra and a co-algebra, such that the multiplication and co-multiplication satisfy $(\mathrm{id}_A \otimes \mu) \circ (\Delta \otimes \mathrm{id}_A) = \Delta \circ \mu = (\mu \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes \Delta)$, then A is called a *Frobenius algebra*. Equivalently, A is a finite dimensional algebra with a non-degenerate bilinear form $\beta : A \otimes A \to \mathbb{K}$ satisfying $\beta \circ (\mathrm{id}_A \otimes \mu) = \beta \circ (\mu \otimes \mathrm{id}_A)$. For details, see [Koc].

Example 4.2.2. Any matrix algebra over $\mathbb{K}$ together with the bilinear form $\beta(X, Y) = \mathrm{tr}(XY)$ is a Frobenius algebra.

If A is a Frobenius algebra, such that the maps $\mu \circ \Delta$ and $\epsilon \circ \eta$ are the identity (or non-zero scalar multiples thereof), then A is called a *special Frobenius algebra* [FRS1].

Algebras in tensor categories

Let $\mathcal{C}$ be a monoidal category. An algebra in $\mathcal{C}$ is an object $A \in \mathcal{C}$ together with morphisms $\mu \in \mathrm{Hom}(A \otimes A, A)$ and $\eta \in \mathrm{Hom}(\mathbf{1}, A)$ satisfying associativity and unitality as above (with associator and unitors included in the obvious way, see for example [Pa2, Def. 3.2.15]). If $\mathcal{C}$ is the category of vector spaces, we recover the definition of an algebra over $\mathbb{K}$. In the same way, the notions of a co-algebra and of a (special) Frobenius algebra over a field can easily be generalised to co-algebra and (special) Frobenius algebra in a general monoidal category. Moreover, the notions of bi-algebra and Hopf algebra make sense in any braided monoidal category.

Example 4.2.3. The circle with positive orientation $(S^1, +) \in 2\mathrm{Cob}$ defines a commutative Frobenius algebra, whereby the multiplication is given by the pair of pants $(S^1, +) \sqcup (S^1, +) \to (S^1, +)$ with two incoming and one outgoing boundary, the co-multiplication by the pair of pants $(S^1, +) \to (S^1, +) \sqcup (S^1, +)$ with one in-going and two outgoing boundaries, the unit by the disk $\emptyset \to (S^1, +)$ with one outgoing boundary, and the counit by the disk $(S^1,+) \to \emptyset$ with one incoming boundary. Evidently, this Frobenius algebra is not a special algebra, since cobordisms of different genus are not diffeomorphic.

Modules and bimodules

Let $\mathcal{C}$ be a strict monoidal category and $A \in \mathcal{C}$ an algebra. Just as for ordinary algebras over $\mathbb{K}$, we have the notion of modules and bi-modules over A: a *left A-module* in $\mathcal{C}$ is an object $M \in \mathcal{C}$ together with a morphism $\rho^L_M \in \mathrm{Hom}(A \otimes M, M)$ (the left A-*action*) which is compatible with the algebra structure in the sense that $\rho^L_M \circ (\mu \otimes \mathrm{id}_M) = \rho^L_M \circ (\mathrm{id}_A \otimes \rho^L_M)$ holds. A morphism between left A-modules $M \in \mathcal{C}$ and $N \in \mathcal{C}$ is a morphisms $f : M \to N$ in $\mathcal{C}$ which satisfies $f \circ \rho^L_M = \rho^L_N \circ f$. A *right A-module* in $\mathcal{C}$ is defined analogously, with a right action $\rho^R_M \in \mathrm{Hom}(M \otimes A, M)$ instead. An *A-bimodule* in $\mathcal{C}$ is an object $M \in \mathcal{C}$ which is both a left and a right A-module, such that $\rho^L_M \circ (\mathrm{id}_A \otimes \rho^R_M) = \rho^R_M \circ (\rho^L_M \otimes \mathrm{id}_A)$. Morphisms of A-bimodules are morphisms in $\mathcal{C}$ which commute with both the left and right A-action.

We denote the category of A-bimodules in $\mathcal{C}$ by ${}_A\mathcal{C}_A$. The set of morphisms between bimodules M and N is denoted by $\mathrm{Hom}_{A|A}(M,N) \subseteq \mathrm{Hom}(M,N)$. If $\mathcal{C}$ is abelian, then the category ${}_A\mathcal{C}_A$ is monoidal, with the tensor product $M \otimes_A N$ over A defined as the cokernel of $\rho^R_M \otimes \mathrm{id}_N - \mathrm{id}_M \otimes \rho^L_N$, and with tensor unit the algebra A as a bimodule over itself.

The functors $\alpha^\pm : \mathcal{C} \to {}_A\mathcal{C}_A$

Let $A \in \mathcal{C}$ be an algebra in a strict monoidal category. For any object $U \in \mathcal{C}$ one obtains canonically a left A-module structure on $A \otimes U$ by setting $\rho^L_{A\otimes U} = \mu \otimes \mathrm{id}_U$. This is called the *induced module*. If $\mathcal{C}$ is braided, then one obtains also two canonical right module structures on $A \otimes U$ by setting either $\rho^{R+}_{A\otimes U} = (\mu \otimes \mathrm{id}_U) \circ (\mathrm{id}_A \otimes c_{U,A})$ or $\rho^{R-}_{A\otimes M} = (\mu \otimes \mathrm{id}_U) \circ (\mathrm{id}_A \otimes c^{-1}_{A,U})$. By associativity of the multiplication, these are compatible with the left induced module action. Hence, this gives rise to two types of *induced bimodules*, denoted by

$$\alpha^\pm(U) := \left(A \otimes U,\, \rho^L_{A\otimes U}, \rho^{R\pm}_{A\otimes U}\right) \ . \tag{4.2.1}$$

In fact, this yields two faithful functors

$$\alpha^\pm : \mathcal{C} \to {}_A\mathcal{C}_A \tag{4.2.2}$$

which send an object $U \in \mathcal{C}$ to $\alpha^\pm(U)$ and a morphism $f : U \to V$ to $\mathrm{id}_A \otimes f$, respectively.

The category $\mathcal{C}_U$

The following category plays a central role in the treatment of perturbed defects.

Definition 4.2.4. Let $\mathcal{C}$ be a monoidal category and $U \in \mathcal{C}$. The category $\mathcal{C}_U$ has

- objects (V, f): pairs consisting of an object $V \in \mathcal{C}$ together with a morphism $f \in \mathrm{Hom}_{\mathcal{C}}(U \otimes V, V)$.

- morphisms $a : (V, f) \to (W, g)$: morphisms $a \in \mathrm{Hom}_{\mathcal{C}}(V, W)$ satisfying

$$a \circ f = g \circ (\mathrm{id}_U \otimes a) \ . \tag{4.2.3}$$

If $\mathcal{C}$ is additive and U is equipped with a lift to the monoidal center of $\mathcal{C}$, i.e. if there is a family of morphisms $\varphi_{U,V} : U \otimes V \to V \otimes U$ such that $(U, \varphi) \in \mathcal{Z}(\mathcal{C})$, then the category $\mathcal{C}_U$ can be equipped with a monoidal structure given by the tensor product

$$\begin{aligned} (V, f) \otimes_U (W, g) \ &:= \ ((V \otimes W), T(f, g)) \\ T(f, g) \ &:= \ \big(f \otimes \mathrm{id}_W + (\mathrm{id}_V \otimes g) \circ \alpha_{V,U,W} \circ (\varphi_{U,V} \otimes \mathrm{id}_W)\big) \circ \alpha^{-1}_{U,V,W} \ . \end{aligned} \tag{4.2.4}$$

The associator and unit maps of $\mathcal{C}_U$ are identical with those of $\mathcal{C}$.

Notations 4.2.5. Denote by $G : \mathcal{Z}(\mathcal{C}) \to \mathcal{C}$ the forgetful functor. Let us agree on the following convention: whenever we write $\mathcal{C}_Z$ for some $Z \in \mathcal{Z}(\mathcal{C})$, we mean the category $\mathcal{C}_{G(Z)}$ with the monoidal structure fixed as in (4.2.4).

Remark 4.2.6. The morphism f suggests to think of (V, f) as a "left U-module" (with action f) in some sense. That intuition is not so far off: under suitable conditions one can show [BR, Prop. 2.6] that $\mathcal{C}_U \cong T(U)\text{-Mod}_{\mathcal{C}}$, where $T(U) = \oplus_{n\in\mathbb{Z}} U^{\otimes n}$ is the tensor algebra (with the algebra structure defined by the associator) .

4.3. Ribbon categories

Definition 4.3.1. A *ribbon category* is a rigid braided monoidal category $\mathcal{C}$ equipped with a natural transformation

$$\theta : \mathrm{id} \longrightarrow \mathrm{id} \ , \tag{\textit{the twist}}$$

such that $\theta_{U^\vee} = \theta_U^\vee$ holds, and the diagram

$$\begin{tikzcd} U \otimes V \arrow[r, "\theta_{U\otimes V}"] \arrow[d, "c_{U,V}"'] & U \otimes V \\ V \otimes U \arrow[r, "\theta_V \otimes \theta_U"] & V \otimes U \arrow[u, "c_{V,U}"'] \end{tikzcd} \tag{4.3.1}$$

commutes for all $U, V \in \mathcal{C}$.

Such a category is in particular also *pivotal*: there exists a monoidal natural isomorphism from the identity functor to the double dual functor, which is compatible with the monoidal structure, and left and right duals agree.

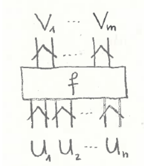
Ribbon categories are particularly nice because they admit a simple graphical interpretation of morphisms in terms of ribbons[1] as follows. Without loss of generality, let us assume that $\mathcal{C}$ is strict. Any object $U \in \mathcal{C}$ is represented by a ribbon *labelled* by U. The

[1] A ribbon is a framed oriented line, with additional choice of front/backside, embedded in $\mathbb{R}^3$.

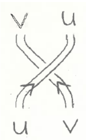
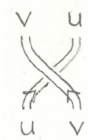
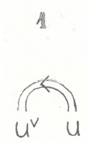
tensor unit is represented by an invisible ribbon. The dual object is represented by the ribbon with opposite direction:

$\mathrm{id}_U =$  , $\mathrm{id}_{U^\vee} =$  (4.3.2)

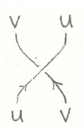
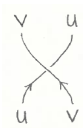
A morphism $f \in \mathrm{Hom}(\otimes_{j=1}^n U_j, \otimes_{k=1}^m V_k)$ is represented by a coupon with n ordered incoming ribbons labelled by $U_1, \ldots, U_n$ and m ordered outgoing ribbons labelled by $V_1, \ldots, V_m$:

$f =$  (4.3.3)

These coupons are to be understood as black boxes. The black boxes of the braiding, (co)evaluation and twist are explicitly given by

$c_{U,V} =$  , $c_{U,V}^{-1} =$  , $\mathrm{ev}_U =$  , $\mathrm{coev}_U =$  , $\theta_U =$  . (4.3.4)

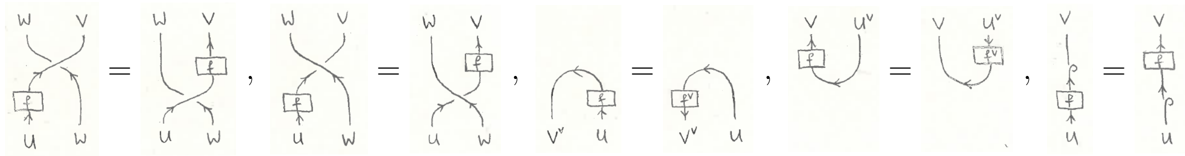
By stacking together ribbons and coupons, any morphism can be represented by a so-called *ribbon graph* embedded in a box in $\mathbb{R}^3$. Hereby, the ribbons representing the incoming objects are attached to the floor of the box along a line, while the ribbons representing the outgoing objects are attached to the top of the box along a line. Two morphisms are equal if their corresponding ribbon graphs are isotopic (whereby the ribbon ends attached to the top and bottom remain fixed). In fact, the twist is the only instance where a ribbon is distinguishable from a line. Thus, one usually represents above morphisms as follows:

$c_{U,V} =$  , $c_{U,V}^{-1} =$  , $\mathrm{ev}_U =$  , $\mathrm{coev}_U =$  , $\theta_U =$  . (4.3.5)

Rather than by a ribbon graph, a morphism is now represented by what is called a *string diagram*. The loops representing the twist keep track of the monodromy of the ribbons with respect to the plane in which the diagrams are drawn. If two string diagrams can be manipulated into one another by using only the second and third Reidemeister move of knot theory (see e.g. [Kau]), then the corresponding morphisms are equal.[2] Moreover,

[2]The first Reidemeister move corresponds to "untwisting" a string, which is not allowed here.

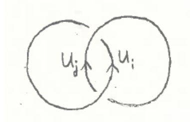
functoriality of the special morphisms (4.3.5) implies that for all $U, V, W \in \mathcal{C}$ and all $f : U \to V$ the following identities hold:



(4.3.6)

Example 4.3.2. $\mathrm{Vect}_{\mathbb{K}}$ is a ribbon category with trivial twist. More generally, the finite-dimensional representations of a suitable q-deformation of $U(\mathfrak{g})$ yield a ribbon category $\mathrm{Rep}(U_q(\mathfrak{g}))$. Ribbon graphs labelled by the 2-dimensional representation of $U_q(\mathfrak{sl}_2)$ are closely related to the Jones polynomial.

Definition 4.3.3. Let $\mathcal{C}$ be an abelian $\mathbb{C}$-linear ribbon category, and denote by $\mathcal{I}$ the set indexing its isomorphism classes of simple objects. Pick a representative U_j for each isomorphism class $j \in \mathcal{I}$. The category $\mathcal{C}$ is called *modular* if it is semi-simple, $\mathbf{1}$ is simple, $|\mathcal{I}|$ is finite, and the matrix $(S_{ij})_{i,j\in\mathcal{I}}$ with entries

$$S_{ij} \;=$$  (4.3.7)

is non-degenerate.

Example 4.3.4. $\mathrm{Vect}_{\mathbb{K}}$ is a modular category with only one simple object, namely the tensor unit $\mathbf{1}$. In contrast, $\mathrm{Rep}(\mathfrak{g})$ is not modular. For example, $\mathfrak{sl}_2$ has a simple representation for every dimension.

Remark 4.3.5. Modular categories are, in particular, finite abelian categories, on which the natural tensor product operation $\boxtimes$ (*Deligne product*) can be defined. If $\mathcal{C}$ is a modular category and $\overline{\mathcal{C}}$ denotes the same category with reverse braiding, then there is a braided monoidal equivalence

$$\begin{aligned} \iota : \mathcal{C} \boxtimes \overline{\mathcal{C}} &\xrightarrow{\;\cong\;} \mathcal{Z}(\mathcal{C}) \\ U \otimes_{\mathbb{C}} V &\mapsto (U \otimes V, \sigma_{U\otimes V,-}) \ , \end{aligned} \tag{4.3.8}$$

where $\sigma_{U\otimes V,-} := (c_{U,-} \otimes \mathrm{id}_V) \circ (\mathrm{id}_U \otimes c^{-1}_{-,V})$ (see [Mu] or [EGNO, Ch. 8.6] for details).

4.4. Topological field theory

A *topological (quantum) field theory* (TQFT, or TFT for short) is a monoidal functor from some category of cobordisms to a category of vector spaces (possibly with additional structure). The paradigmatic example are 2-dimensional closed topological field theories, which are braided monoidal functors

$$\mathrm{tft} : 2\mathrm{Cob} \longrightarrow \mathrm{Vect}_{\mathbb{K}} \ . \tag{4.4.1}$$

Recall that $(S^1, +)$ is a commutative Frobenius algebra in 2Cob. Its image under tft is a commutative Frobenius algebra over $\mathbb{K}$. It is a classic result that this Frobenius algebra uniquely determines the TFT [Koc].

In chapter 5, we will refer to a more complicated class of TFTs known under the name of *Reshetikin-Turaev*, which we briefly introduce now. The datum of a Reshetikin-Turaev TFT is a modular category $\mathcal{C}$. This datum defines a category of (extended) cobordisms $3\mathrm{Cob}_{\mathcal{C}}$ as well as a braided monoidal functor

$$\mathrm{tft}_{\mathcal{C}} : 3\mathrm{Cob}_{\mathcal{C}} \longrightarrow \mathrm{Vect}_{\mathbb{K}} \ . \tag{4.4.2}$$

It is beyond the scope of this thesis to give the full definition of both this cobordism category and the TFT functor, but let us give a sketch of their construction. For full details we refer to [Tu], and also [BK, Ch. 4].

Definition 4.4.1. A *smooth $\mathcal{C}$-marked surface* $\widehat{\Sigma}$ is an oriented compact smooth 2-manifold Σ with a finite number of marked points $p_1, \ldots, p_n \in \Sigma$, whereby each point p_i comes equipped with a tangent vector $t_i \in T_{p_i}\Sigma$, a sign $\epsilon_i \in \{\pm 1\}$ ("out" or "in") and an object $W_i \in \mathcal{C}$. A *smooth $\mathcal{C}$-marked 3-manifold* $\widehat{M}$ is an oriented smooth 3-manifold M with boundary, together with an embedded $\mathcal{C}$-labelled ribbon graph whose open strands are attached to the boundary. Smooth $\mathcal{C}$-marked 3-manifolds function as cobordisms between smooth $\mathcal{C}$-marked surfaces: each point p_i coincides with the end of an open strand of the ribbon graph labelled by W_i (the ribbon is out- or in-going depending on ϵ_i, and the tangent vector t_i determines the alignment of the base of the ribbon on the boundary surface).

Let us denote the category of smooth $\mathcal{C}$-marked surfaces and 3-manifolds by $3\mathrm{Cob}'_{\mathcal{C}}$. One can define a standard connected smooth $\mathcal{C}$-marked surface $\widehat{\Sigma}_{g,n}$ of genus g with n marked points as follows: take the sphere and fix the points $p_1 < \cdots < p_n < q_1 < \cdots < q_g$ on the equator. Equip $p_1, \ldots, p_n$ with tangent vectors parallel to the equator in positive direction, and replace $q_1, \ldots, q_g$ with little handles.

There is now a map $\mathrm{tft}'_{\mathcal{C}} : 3\mathrm{Cob}'_{\mathcal{C}} \to \mathrm{Vect}_{\mathbb{K}}$ defined as follows:

- The positively oriented connected smooth $\mathcal{C}$-marked standard surface $\widehat{\Sigma}_{g,n}$ with marked points labelled by objects $W_1, \ldots, W_n$ and signs $\epsilon_1, \ldots, \epsilon_n$, is mapped to the vector space $\mathrm{tft}'_{\mathcal{C}}(\widehat{\Sigma}_{g,n}) := \mathrm{Hom}_{\mathcal{C}}(W_1^{\epsilon_1} \otimes \cdots \otimes W_n^{\epsilon_n} \otimes H^{\otimes g}, \mathbf{1})$, where $H = \bigoplus_{j \in \mathcal{I}} U_j \otimes U_j^{\vee}$ and $W_i^+ = W_i$, $W_i^- = W_i^{\vee}$. The same $\mathcal{C}$-marked surface with negative orientation is mapped to $\mathrm{Hom}_{\mathcal{C}}(\mathbf{1}, H^{\otimes g} \otimes W_n^{-\epsilon_n} \otimes \cdots \otimes W_1^{-\epsilon_1})$.

- A smooth $\mathcal{C}$-marked 3-manifold M connecting an incoming boundary ∂M_- of genus g_{in} (with points marked by $V_1, \ldots, V_n$ and signs $\nu_1, \ldots, \nu_n$) to an outgoing boundary ∂M_+ of genus g_{out} (with points marked by $W_1, \ldots, W_m$ and signs $\epsilon_1, \ldots, \epsilon_m$) is sent to the linear map $\mathrm{tft}'_{\mathcal{C}}(\widehat{M}) : \mathrm{tft}'_{\mathcal{C}}(\partial_- M) \to \mathrm{tft}'_{\mathcal{C}}(\partial_+ M)$ induced by a morphism $V_1^{\nu_1} \otimes \cdots \otimes V_n^{\nu_n} \otimes H^{\otimes g_{\mathrm{in}}} \to W_1^{\nu_1} \otimes \cdots \otimes W_m^{\nu_m} \otimes H^{\otimes g_{\mathrm{out}}}$ corresponding to a certain ribbon graph. This ribbon graph consists of three disjoint ribbon graphs: a) the ribbon graph which comes with $\widehat{M}$, b) loops $\sum_{j\in\mathcal{I}} \mathrm{ev}_{U_j}$ or $\sum_{j\in\mathcal{I}} \mathrm{coev}_{U_j}$ for every handle of ∂M_- and ∂M_-, respectively, c) a ribbon link L (ribbon graph without open strands) such that M can be obtained from S^3 by Dehn surgery [BK, Ch. 4.1] along L.

One can show that this assignment is compatible with the braiding of $\mathcal{C}$. Take, for instance, the positively oriented standard sphere with 2 marked points p_1, p_2. If we label p_1 by W_1 and p_2 by W_2 (all signs positive), then this is mapped to $\mathrm{Hom}(W_1 \otimes W_2, \mathbf{1})$. If p_1 is labelled by W_2 and p_2 by W_1, then the result is $\mathrm{Hom}(W_2 \otimes W_1, \mathbf{1})$. But the second $\mathcal{C}$-marked sphere can be obtained from the first one by composing with the $\mathcal{C}$-marked 3-manifold $S^2 \times [0,1]$, whose embedded ribbon graph represents the braiding c_{W_1,W_2} (or the inverse braiding). Compatibility with the braiding means that it makes no difference whether we apply $\mathrm{tft}_{\mathcal{C}}$ directly to the second sphere, or apply $\mathrm{tft}_{\mathcal{C}}$ separately to the first sphere and the cylinder and then compose the resulting morphisms.

Using cylinders with embedded braided ribbon graphs as above, one can extend the assignment $\mathrm{tft}'_{\mathcal{C}}$ to any smooth $\mathcal{C}$-marked surface. Finally, one can show that the assignment $\mathrm{tft}'_{\mathcal{C}}$ does in fact not depend on the smooth structure of the $\mathcal{C}$-marked surfaces and 3-manifolds. Instead, it acts on *topological $\mathcal{C}$-marked surfaces* and *topological $\mathcal{C}$-marked 3-manifolds*, which are obtained from the smooth versions by forgetting the smooth structure and identifying homeomorphic cobordisms. In this procedure, the tangent vectors get replaced by arc germs together with orientations.

The assignment $\mathrm{tft}'_{\mathcal{C}}$ is compatible with the monoidal and braided structures, but it is not quite a functor: $\mathrm{tft}'_{\mathcal{C}}$ commutes with composition of morphisms only up to a constant. This problem does not occur at genus zero, which means that it is somewhat irrelevant for the purposes of this thesis. In any case, it can be remedied by adding certain data [Tu, Ch. IV.3] (which are trivial at genus zero) to the objects and morphisms in $3\mathrm{Cob}'_{\mathcal{C}}$, resulting in the category denoted by $3\mathrm{Cob}_{\mathcal{C}}$. With these adjustments, the assignment $\mathrm{tft}'_{\mathcal{C}}$ can then be turned into a proper braided monoidal functor $\mathrm{tft}_{\mathcal{C}} : 3\mathrm{Cob} \to \mathrm{Vect}_{\mathbb{K}}$ as claimed in (4.4.2), see [Tu, Ch. IV.9].

5. Review of 2-dimensional CFT

5.1. 2d CFT without defects

In this section, we will review some cornerstones of 2-dimensional CFT. Throughout this thesis, whenever we speak of CFT we mean 2-dimensional euclidean CFT. Conformal symmetry in 2 dimensions is quite special, since the algebra of local conformal tranformations in 2 dimensions is much larger than in higher dimensions. This enhanced symmetry yields ample constraints to make 2-dimensional CFT the best understood class of quantum field theories beyond TFT.

The material treated in this section can be found in many introductory books on CFT, such as [Scho]. For a review of axiomatic approaches, see [Kon]. We will not attempt to give a comprehensive definition of CFT, since there is at present no axiomatic framework that covers everything one could reasonably call a CFT. Instead, we introduce one possible set of defining data for a CFT, and discuss some of the conditions and properties of these data. Particular attention will be given to a specific class of CFTs, namely *rational* CFTs (RCFT). These include many of the most physically relevant CFTs (including, for example, the Virasoro minimal models and WZW models), and they are among the mathematically best understood examples. In particular, the TFT-approach to CFT, reviewed in Section 5.4 and used in Chapter 6, has been worked out completely for this class of CFTs. However, RCFT excludes many interesting theories such as Liouville theory or logarithmic CFTs.

5.1.1. Data of a full closed CFT without defects

Let (Σ, γ) be a smooth, compact, oriented, closed 2-manifold Σ equipped with a Riemannian metric γ of Euclidean signature. Consider its tangent bundle, defined as a set by

$$T\Sigma = \{(p,t)|p \in \Sigma, t \in T_p\Sigma\} \ . \tag{5.1.1}$$

The cartesian product $(T\Sigma)^{\times n}$ is the set of n-tuples of points in Σ equipped with tangent vectors. We will need the subset of such tuples where none of the points coincide,

$$T\Sigma^n_\Delta := \left\{(p_1,t_1) \times \cdots \times (p_n,t_n) \in (T\Sigma)^{\times n} | p_k \neq p_l \text{ for all } 1 \leq k \neq l \leq n\right\} \ . \tag{5.1.2}$$

Elements in $T\Sigma^n_\Delta$ may be thought of as the choice of n ordered distinct points on Σ, together with a choice of isothermal local coordinates[1] around the points, whereby the

[1] An isothermal coordinate system in a neighbourhood U of $p \in \Sigma$ is a conformal map $\varphi : U \to \mathbb{R}^2$ into the Euclidean plane. In dimension 2, local isothermal coordinates always exist. Coordinate transformations between isothermal coordinates are conformal.

tangent vector stands for the direction of the real axis of the local coordinate system.

A 2-dimensional (Euclidean) full closed[2] CFT (without defects) consists of the following data:

- a vector space $\mathcal{H}$ (*the field space*) with special vectors $\Omega \in \mathcal{H}$ (the *vacuum*) and $T, \bar{T} \in \mathcal{H}$ (*stress tensor*).
- a collection of maps (*n-point correlators*)

$$\begin{aligned} C^n_{(\Sigma,\gamma)} : \mathcal{H}^{\otimes n} \times T\Sigma^n_\Delta &\longrightarrow \mathbb{C} \\ (v_1 \otimes \cdots \otimes v_n, \tilde{p}_1, \ldots, \tilde{p}_n) &\mapsto \langle v_1(\tilde{p}_1) \ldots v_n(\tilde{p}_n) \rangle_\Sigma \end{aligned} \tag{5.1.3}$$

  one for each each $n \in \mathbb{Z}_{\geq 0}$ and each smooth, closed, compact, oriented Riemann surface (Σ, γ). These maps have to satisfy certain conditions individually, as well as consistency conditions among each other.

Remark 5.1.1. Before we discuss these data, let us point out how this characterisation of CFT relates to the framework of functorial QFT. From the latter point of view, correlators are linear maps

$$C(\widehat{\Sigma}_n, \gamma) : \mathcal{H}^{\otimes n} \longrightarrow \mathbb{C} \tag{5.1.4}$$

where $(\widehat{\Sigma}_n, \gamma)$ is a smooth, compact, oriented Riemann surface with n disjoint parametrised boundary components.[3] One should think of $\widehat{\Sigma}_n$ as being obtained from the closed Riemann surface (Σ, γ) with n marked points $p_1, \ldots, p_n$ by cutting out little discs around the marked points. The boundary parametrisations arise naturally from the choice of local coordinates around the points p_j. To each boundary component one associates a copy of $\mathcal{H}$. One would like to interpret $C(\widehat{\Sigma}_n, \gamma)$ as the image of a cobordism $(S^1)^{\sqcup n} \to \emptyset$ with conformal structure (modulo conformal equivalence) under a monoidal functor C. An axiomatisation of CFT in this spirit has been proposed by Segal [Seg]. In this approach, it is natural to call $\mathcal{H}$ (or, in fact, an appropriate completion of it) the *state space* rather than field space. The assumption that this point of view is in some way equivalent to ours goes under the name of *state-field correspondence*. While not completely rigorous, it often serves as good intuition. For example, it prompts us to think of a 2-point correlator at genus 0 as a linear map associated to a cylinder.

Let us now discuss the data in more detail.

5.1.2. The field space

The characteristic feature of 2d CFT is that the space of fields $\mathcal{H}$ carries a representation of $\mathfrak{vir} \times \mathfrak{vir}$, where $\mathfrak{vir}$ is the *Virasoro algebra*. This is the complex Lie algebra spanned by

[2] The qualifiers refer to the types of surfaces on which the theory is defined. “Full” means that it is consistently defined on surfaces of any genus, while “closed” means that we restrict to surfaces without a boundary.

[3] Note that these boundaries are not boundaries in the sense of open-closed CFT.

generators $\{L_n | n \in \mathbb{Z}\} \cup \{c\}$ with relations

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{1}{12}(m^3 - m)\delta_{n+m,0}\, c \ , \tag{5.1.5}$$

$$[L_m, c] = 0 \ . \tag{5.1.6}$$

The Lie algebra $\mathfrak{vir}$ is the unique nontrivial one-dimensional central extension of the Witt algebra, the Lie algebra of (anti-)holomorphic polynomial vector fields on the Riemann sphere minus two points. Holomorphic and anti-holomorphic polynomial vector fields generate together the local conformal transformations [Scho, Ch. 5], which explains the double appearance of $\mathfrak{vir}$ as a symmetry of $\mathcal{H}$. The generators of the second (anti-holomorphic) copy of $\mathfrak{vir}$ are denoted by $\bar{L}_n$ to distinguish them from the generators of the holomorphic copy. The subalgebras $\langle L_{-1}, L_0, L_1\rangle$ and $\langle \bar{L}_{-1}, \bar{L}_0, \bar{L}_1\rangle$ are of special relevance, since they generate local conformal transformations which extend to global conformal transformations on the Riemann sphere (Möbius transformations). In particular, the combination $L_0 + \bar{L}_0$ generates scalings, $L_0 - \bar{L}_0$ generates rotations, and $L_{-1}, \bar{L}_{-1}$ generate translations in the direction z and $\bar{z}$, respectively. Accordingly, the representing endomorphisms of these generators on $\mathcal{H}$ correspond to the Hamiltonian, the momentum, and the derivatives with respect to z and $\bar{z}$.

Let V be a representation of the Virasoro algebra on which L_0 acts diagonalisably.[4] A vector $\phi \in V$ with the property

$$L_n.\phi = 0 \qquad \forall n > 0 \quad , \qquad L_0.\phi = \Delta L_0 \quad \text{for some } \Delta \in \mathbb{R} \ , \tag{5.1.7}$$

is called a *Virasoro heighest weight vector*[5] or chiral *Virasoro primary field* of weight Δ. (Anti-chiral primary fields are defined with respect to generators $\bar{L}_n$ instead.) Vectors of the form

$$L_{-n_1}L_{-n_2}\cdots L_{-n_k}.\phi \qquad\qquad (\text{with } k \in \mathbb{Z}_{>0}\, ;\, n_1, \ldots, n_k \in \mathbb{Z}_{>0}) \tag{5.1.8}$$

are called *descendants* of ϕ. If V is generated by acting with $\mathfrak{vir}$ on ϕ in this way, then V is called a Virasoro heighest weight representation with conformal weight Δ. It is not hard to see from (5.1.5) that the eigenvalues of L_0 acting on such a representation lie in the set $\Delta + \mathbb{Z}_{\geq 0}$, with the descendant (5.1.8) sitting in the eigenspace associated to the eigenvalue $\Delta + \sum_{l=0}^{k} n_l$. In this way, the action of L_0 induces a $\mathbb{Z}$-grading on the representation V.

If $V, \bar{V}$ are two Virasoro representations, then $V \otimes_{\mathbb{C}} \bar{V}$ is a representation of $\mathfrak{vir} \times \mathfrak{vir}$. If the vector $\phi \in V$ is a chiral and the vector $\bar{\phi} \in \bar{V}$ an anti-chiral Virasoro primary field, of weight Δ and $\bar{\Delta}$ respectively, then we say that the combination $\phi \otimes \bar{\phi}$ is a (non-chiral) Virasoro primary field of weight $(\Delta, \bar{\Delta})$. If V and $\bar{V}$ are highest weight representations, then we say that $V \otimes_{\mathbb{C}} \bar{V}$ is a highest weight representation with conformal weights $(\Delta, \bar{\Delta})$.

It is normally assumed that the field space $\mathcal{H}$ is spanned by a collection of Virasoro primary fields and their descendants. If it contains a finite number of Virasoro primaries,

[4] This is always the case for rational CFTs and the free boson, which we consider in this thesis. It is wrong for logarithmic (non-semisimple) CFTs.

[5] "highest weight" is a conventional misnomer which, in the context of Virasoro representations, should always be read as "lowest weight"

then there exists a lower bound on the possible eigenvalues of $L_0 + \bar{L}_0$ that any vector in $\mathcal{H}$ can have. This reflects the physical condition that the energy must be bounded from below. In the simplest cases, $\mathcal{H}$ is simply a direct sum of some Virasoro highest weight representations of $\mathfrak{vir} \times \mathfrak{vir}$.

However, the Virasoro action only describes the minimal symmetry of $\mathcal{H}$. In general, $\mathcal{H}$ carries an action of $\mathcal{V} \times \mathcal{V}$, where $\mathcal{V}$ is an algebraic structure containing $\mathfrak{vir}$ as well as extended symmetries, loosely called the *chiral algebra*. Chiral algebras are axiomatised by the notion of *vertex operator algebras (VOAs)*, see for example [Kac]. An introductory treatment of VOAs more geared towards CFT can be found in [Scho]. In this framework, $\mathcal{V}$ is a $\mathbb{Z}$-graded vectorspace equipped with a map

$$Y(-,z) : \mathcal{V} \longrightarrow \mathrm{End}(\mathcal{V})\,[[z^{\pm 1}]] \ , \tag{5.1.9}$$

which assigns to each $v \in \mathcal{V}$ a formal power series $Y(v,z)$ (the *vertex operator*) with coefficients in $\mathrm{End}(\mathcal{V})$. This map must satisfy a weak version of commutativity and associativity. The Virasoro action on $\mathcal{V}$ is implemented by a distinguished homogeneous vector $T \in \mathcal{V}$ in degree 2 with the property that the coefficients of its vertex operator are the representing endomorphisms of the generators of $\mathfrak{vir}$:

$$Y(T,z) = \sum_{n=-\infty}^{\infty} L_n z^{-n-2} \tag{5.1.10}$$

The endomorphism L_0 counts the degree, i.e. $L_0 v = \deg(v) v$ for any homogeneous vector v, while the endomorphism L_{-1} determines a derivative of vertex operators via $\frac{d}{dz} Y(v,z) = Y(L_{-1}v, z)$. Moreover, there is a distinguished element $|0\rangle \in \mathcal{V}$ in degree 0 such that $Y(|0\rangle, z) = \mathrm{id}_{\mathcal{V}}$ and $Y(v,z)|0\rangle \to v$ as $z \to 0$.

Like for usual associative algebras over a field, there exists a notion of modules over the VOA $\mathcal{V}$, and of intertwiners between them. A module R over $\mathcal{V}$ is a $\mathbb{C}$-graded vectorspace equipped with a map

$$Y_R(-,z) : \mathcal{V} \to \mathrm{End}(R)[[z^{\pm}]] \tag{5.1.11}$$

satisfying properties analogous to the ones of $Y(-,z)$. The coefficients of $Y_R(T,z)$ implement the action of $\mathfrak{vir}$ on R in analogy with (5.1.10). An intertwiner from R to S is a linear map $\psi \in \mathrm{Hom}(R,S)$ such that $\psi(Y_R(v,z)r) = Y_S(v,z)\psi(r)$ for all $r \in R$ and $v \in \mathcal{V}$. Modules and intertwiners of $\mathcal{V}$ form together the category $\mathrm{Rep}(\mathcal{V})$. There is a faithful forgetful functor from $\mathrm{Rep}(\mathcal{V})$ to the category of representations of $\mathfrak{vir}$.

For a given VOA, we will denote the (isomorphism classes of) simple objects by U_j where j runs over some index set $\mathcal{I}$. The field space $\mathcal{H}$ is a module over $\mathcal{V} \otimes_{\mathbb{C}} \mathcal{V}$, and thus an object in $\mathrm{Rep}(\mathcal{V}) \boxtimes \mathrm{Rep}(\mathcal{V})$. There exists a class of VOAs whose representation categories are semisimple and which have only a finite number $|\mathcal{I}|$ of (isomorphism classes of) simple modules. These VOAs are called *rational* (see [Hu] for a precise definition), and they are used to define *rational CFTs*. In this setting, all modules have duals, and there exists an involution $\mathcal{I} \to \mathcal{I}$, $j \mapsto \bar{j}$, with the property that $U_{\bar{j}} \cong U_j^{\vee}$. The field space of a rational CFT is of the form

$$\mathcal{H} = \bigoplus_{i,j \in \mathcal{I}} (U_i \otimes_{\mathbb{C}} U_j)^{\oplus Z_{ij}} \ , \tag{5.1.12}$$

with $Z_{ij} \in \mathbb{Z}_{\geq 0}$ some multiplicities not directly determined by $\mathcal{V}$. The case $Z_{ij} = \delta_{i\bar{j}}$ is referred to as the *charge-conjugate* or *Cardy case*.

The VOA $\mathcal{V}$ itself is assumed to be a simple module (called *vacuum module*) which we identify with U_0. The vector $|0\rangle$ is a Virasoro heighest weight vector with weight $\Delta_0 = 0$, and which also satisfies $L_{-1}|0\rangle = 0$. A vector in $\mathcal{H}$ which vanishes under the action of all generators of Möbius transformations is called *vacuum*. Setting $Z_{00} = 1$ in (5.1.12) ensures existence and uniqueness of a vacuum. This unique vacuum is given by the vector $\Omega = |0\rangle \otimes_{\mathbb{C}} |0\rangle$.

The action of L_0 induces an $\mathbb{R}$-grading $U_j = \bigoplus_{\Delta\in\mathbb{R}} (U_j)_\Delta$ with countably many non-trivial and finite-dimensional homogeneous subspaces $(U_j)_\Delta := \{u \in U_j | L_0.u = \Delta u\}$. Correspondingly, the action of the Hamiltonian $L_0 + \bar{L}_0$ defines an $\mathbb{R}$-grading on the field space, given by

$$\mathcal{H} = \bigoplus_{w\in\mathbb{R}} \mathcal{H}_w \qquad \text{with} \qquad \mathcal{H}_w = \bigoplus_{i,j\in\mathcal{I}} \bigoplus_{\substack{\Delta,\bar{\Delta}\in\mathbb{R} \\ \Delta+\bar{\Delta}=w}} \left((U_i)_\Delta \otimes_{\mathbb{C}} (U_j)_{\bar{\Delta}}\right)^{\oplus Z_{ij}} . \tag{5.1.13}$$

Since $|\mathcal{I}| < \infty$ for rational CFTs, the homogeneous subspaces $\mathcal{H}_w$ are finite-dimensional. Moreover, since there are a finite number of Virasoro highest-weight vectors, there exists a $w' \in \mathbb{R}_{<0}$ such that $\mathcal{H}_w = \{0\}$ for all $w \leq w'$.

5.1.3. The correlators

We will discuss the construction of the correlators in more detail in the following sections. In this section, we will briefly outline some of the conditions that they have to satisfy.

Among the basic properties of correlators are the following: $\langle v_1(\tilde{p}_1) \dots v_n(\tilde{p}_n)\rangle_{\Sigma,\gamma}$ must be linear in the fields v_j, smooth in the marked points p_j as well as the tangent vectors t_j, and continuous in the metric γ. In the limit of coinciding points $p_i \to p_j$ (for $i \neq j$), singularities can occur. In rational CFTs, these are poles whose order is governed by the weights of v_i and v_j. Moreover, for any permutation $\sigma \in \mathfrak{S}_n$, we have

$$\left\langle v_{\sigma(1)}(\tilde{p}_{\sigma(1)}) \dots v_{\sigma(n)}(\tilde{p}_{\sigma(n)})\right\rangle_{\Sigma,\gamma} = \langle v_1(\tilde{p}_1) \dots v_n(\tilde{p}_n)\rangle_{\Sigma,\gamma} . \tag{5.1.14}$$

The less trivial conditions on the correlators of a full CFT are twofold:

(C1) (Conformal symmetry) They must be invariant under conformal transformations of Σ, and be invariant under Weyl transformations of the metric γ.[6]

(C2) (Factorisation) They must satisfy consistency conditions among each other, the *sewing relations*. (We will give an example of these below.)

For an explicit general formulation of these conditions, we refer to [RFjFS]. The sewing relations relate correlators with different numbers of marked points and different genus

[6] up to an anomalous factor depending on the Virasoro central charge c (see e.g. [Gaw]), which drops out if one considers suitably normalised correlators.

of Σ to one another. In particular, they allow every correlator to be reduced to genus 0 correlators. In the functorial formulation (see Remark 5.1.1) the sewing relations amount to functoriality of C.

In the main part of this thesis, we only deal with genus 0 correlators. We will now illustrate some of the above conditions for this subset of correlators. It is standard practice in this case to work on the complex plane instead of the sphere. Denote by $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ the Riemann sphere. As is standard practice, we consider $\overline{\mathbb{C}}$ together with the metric induced from the Euclidean metric on the plane, which of course becomes singular at infinity. Let $z_1, \ldots, z_n \in \overline{\mathbb{C}}$. Fix some standard local coordinates by giving tangent vectors $t_1, \ldots, t_n \in \mathbb{R}^2$. The usual choice is for all tangent vectors to be of length 1 and point in positive real direction. If any other convention is used, we will specifically mention this. We will now abbreviate

$$\langle v_1(z_1) \cdots v_n(z_n) \rangle \;:=\; \langle v_1(\tilde{z}_1) \ldots v_n(\tilde{z}_n) \rangle_{\overline{\mathbb{C}}} \;, \tag{5.1.15}$$

where $\tilde{z}_j := (z_j, t_j)$. This set of correlators with fixed metric then satisfies conformal covariance instead of invariance. This includes for example translation and rotation invariance as well as covariance under scaling $z \mapsto \lambda z$ (where $\lambda > 0$),

$$\langle v_1(z_1) \cdots v_n(z_n) \rangle \;=\; \lambda^{\Delta_1 + \cdots + \Delta_n} \lambda^{\bar{\Delta}_1 + \cdots + \bar{\Delta}_n} \langle v_1(\lambda z_1) \cdots v_n(\lambda z_n) \rangle \;. \tag{5.1.16}$$

In fact, conformal covariance is so restrictive that 3-point correlators of Virasoro primary fields ϕ_i of weight $(\Delta_i, \bar{\Delta}_i)$ are bound to be of the form

$$\langle \phi_1(z_1)\phi_2(z_2)\phi_3(z_3) \rangle = C(\phi_1, \phi_2, \phi_3) \prod_{\substack{1 \le i < j \le 3 \\ k \neq i,j}} (z_i - z_j)^{\Delta_k - \Delta_i - \Delta_j} \, (\bar{z}_i - \bar{z}_j)^{\bar{\Delta}_k - \bar{\Delta}_i - \bar{\Delta}_j} \;. \tag{5.1.17}$$

The quantities $C(\phi_1, \phi_2, \phi_3) \in \mathbb{C}$ are called *structure constants*. Note that 0-point, 1-point and 2-point correlators of primary fields can be obtained from (5.1.17) by setting some of the fields ϕ_1, ϕ_2, ϕ_3 to be the vacuum vector. Moreover, it can be shown that all 3-point correlators for descendants of ϕ_1, ϕ_2, ϕ_3 are determined by (5.1.17).

The most prominent incarnation of the sewing conditions is the *operator product expansion (OPE)*, which relates n-point correlators to $(n-1)$-point correlators. It postulates that any two fields $v_i, v_j \in \mathcal{H}$ inserted at nearby marked points z and w can be replaced by a superposition of fields as follows: if $|z - w|$ is smaller than the distance of any of the other marked points from z, then

$$\begin{aligned} \langle \cdots v_i(z) \cdots v_j(w) \cdots \rangle \;&=\; \\ &\sum_{\chi} C(v_i, v_j, \chi)\; (z - w)^{\Delta_\chi - \Delta_i - \Delta_j} (\bar{z} - \bar{w})^{\bar{\Delta}_\chi - \bar{\Delta}_i - \bar{\Delta}_j} \langle \cdots \chi(z) \cdots \rangle \;, \end{aligned} \tag{5.1.18}$$

where the sum runs over a basis of $\mathcal{H}$, and $C(u, v, w)$ are constants which coincide with the structure constants if u, v, w are primary, and otherwise are determined by them. The sum has, in fact, finite radius of convergence in $|z - w|$. Successive application of (5.1.18) makes it possible to reduce all correlators to 3-point correlators. In this fashion, the structure constants entirely determine the full CFT.

There is another important condition, which we will need in chapter 6: for every pair $z_1, z_2 \in \mathbb{C}$, the 2-point correlators

$$\begin{aligned} \langle -(z_1) - (z_2) \rangle : \mathcal{H} \otimes \mathcal{H} &\longrightarrow \mathbb{C} \\ v_1 \otimes v_2 &\mapsto \langle v_1(z_1) v_2(z_2) \rangle \end{aligned} \tag{5.1.19}$$

define a bilinear form on $\mathcal{H}$. One demands of this pairing that it is non-degenerate. However, the pairing is, in general, not compatible with the grading of $\mathcal{H}$. Yet for $z_1 \to \infty$, $z_2 = 0$, and after a suitable change of local coordinates around z_1 which removes the singularity at infinity, this can be turned into a pairing

$$(-,-) : \mathcal{H} \otimes \mathcal{H} \longrightarrow \mathbb{C} \tag{5.1.20}$$

which is both non-degenerate *and* compatible with the grading (see also Remark 5.2.2). The new local coordinate around z_1 is $z_1 \mapsto 1/z_1$ which has the effect of reversing the tangent vector at this marked point.

5.2. Chiral CFT and conformal blocks

The construction of a consistent set of conformal correlators from a given chiral algebra $\mathcal{V}$ is commonly carried out in two distinct steps, an approach which goes back to [BPZ]:

1. Obtain an ansatz by imposing local conformal symmetry as well as extended symmetries coming from the VOA.

2. Solve remaining consistency conditions

The outcome of step 1 is called a *chiral* CFT. It satisfies the local part of condition (C1) (conformal symmetry). Step 2 amounts to combining two chiral CFTs (which we take to be identical) to a consistent full CFT whose correlators satisfy condition (C2) (factorisation) as well as modular invariance (which is the remaining non-local part of condition (C1)). In this section we briefly review step 1, while step 2 will be covered in section 5.4.2.

Holomorphic factorisation Let us explain the idea of step 1 for the n-point correlator $\langle v_1(z_1) \dots v_n(z_n) \rangle$ on the Riemann sphere. The covariance of this correlator under each infinitesimal conformal (or extended) symmetry can be expressed in terms of a so-called conformal Ward identity. These are linear differential equations in z_k or $\bar{z}_k$, respectively, depending on whether one considers a holomorphic (chiral) or anti-holomorphic (anti-chiral) generator of the symmetry.

The solutions to the chiral n-point Ward identities are functionals on the n-fold tensor product of $\mathcal{H}$. For fixed $v_1, \dots, v_n \in \mathcal{H}$, denote by $F_\alpha(z_1, \dots, z_n)$ a basis of solutions to the chiral n-point Ward identites, and by $\overline{F}_\beta(\bar{z}_1, \ldots, \bar{z}_n)$ a basis of solutions to the anti-chiral Ward identities. The n-point correlator can then be expanded in either basis. Combining the two expansions, one obtains an expression of the form

$$\langle v_1(z_1) \ldots v_n(z_n) \rangle \;=\; \sum_{\alpha,\beta} a_{\alpha,\beta} F_\alpha(z_1, \ldots, z_n) \overline{F}_\beta(\bar{z}_1, \ldots, \bar{z}_n) \;, \tag{5.2.1}$$

with $a_{\alpha\beta} \in \mathbb{C}$ some constants still to be determined. This ansatz is known as *holomorphic factorisation*. It carries over to correlators of any genus, where a choice of complex structure provides us with the notion of complex conjugation locally.

To summarise, step 1 boils down to understanding the space of solutions to the Ward identities. Step 2 amounts to finding a collection of coefficients $a_{ij} \in \mathbb{C}$ for *all* correlators on all surfaces, such that (5.2.1) defines a consistent full CFT.

The space of conformal blocks Let Σ^c be a smooth, closed, compact Riemann surface of genus g with n marked points $p_1, \ldots, p_n$, such that p_k is marked by the representation R_k of the chiral algebra, and with a choice of local holomorphic coordinates in each point. The vector space spanned by the solutions to the chiral Ward identities is called the *space of n-point conformal blocks*[7], which we denote by

$$V_n\left(\Sigma^c\right) \quad \subset \quad \left(R_1 \otimes_{\mathbb{C}} \cdots \otimes_{\mathbb{C}} R_n\right)^\star . \tag{5.2.2}$$

The conformal blocks depend locally holomorphically on the marked points and on the complex structure. They form a vector bundle over the moduli space $\hat{\mathcal{M}}_{g,n}$ of Riemann surfaces of genus g with n marked points and local holomorphic coordinates, the *bundle of n-point conformal blocks*. The fibers are given by (5.2.2).

The bundle of n-point conformal blocks comes equipped with a projectively flat connection, the Knizhnik-Zamolodchikov connection [FBZ, Ch. 17]. Parallel transport thus yields an isomorphism

$$U_\Gamma : V_n(\Sigma^c) \xrightarrow{\cong} V_n(\tilde{\Sigma}^c) \tag{5.2.3}$$

for every path $\Gamma : [0,1] \to \hat{\mathcal{M}}_{g,n}$ between $\Gamma(0) = \Sigma^c$ and $\Gamma(1) = \tilde{\Sigma}^c$. The moduli space $\hat{\mathcal{M}}_{g,n}$ is connected (see [BK, Thm. 6.1.6.]), and hence it suffices to understand the conformal blocks $V(\Sigma)$ for a standard choice of moduli. The functions $F_\alpha(z_1, \ldots, z_n)$ in (5.2.1) are local flat sections of the bundle, evaluated at specific vectors in $R_1, \ldots, R_n$. Typically, these sections cannot be extended to global ones, which means that the $F_\alpha(z_1, \ldots, z_n)$ are multivalued functions with branch cuts instead of meromorphic functions. The final correlator (5.2.1), however, must be single-valued.

Example 5.2.1. Consider chiral primary fields ϕ_1, ϕ_2, ϕ_3 with weights $\Delta_1, \Delta_2, \Delta_3$. When evaluated on these three vectors, any 3-point conformal block on the Riemann sphere is, up to a scalar, given by

$$F_1(z_1, z_2, z_3) \;=\; (z_1 - z_2)^{\Delta_3 - \Delta_1 - \Delta_2} \, (z_1 - z_3)^{\Delta_2 - \Delta_1 - \Delta_3} \, (z_2 - z_3)^{\Delta_1 - \Delta_2 - \Delta_3} \;. \tag{5.2.4}$$

[7] The spaces of chiral and anti-chiral conformal blocks are of course trivially isomorphic.

This is in accordance with (5.1.17). For generic values of $\Delta_1, \Delta_2, \Delta_3$, (5.2.4) is indeed a multivalued function. However, the correlator

$$\langle \phi_1(z_1)\phi_2(z_2)\phi_3(z_3)\rangle \; = \; a_{11} F_1(z_1, z_2, z_3) F_1(\bar{z}_1, \bar{z}_2, \bar{z}_3) \; , \tag{5.2.5}$$

defined by the number $a_{11} = C(\phi_1, \phi_2, \phi_3)$, is single valued.

Conformal blocks and TFT Ward identities are generally hard to solve explicitly. Yet this does not mean that the quest to construct correlators is doomed. Even without access to explicit solutions, the global behaviour of conformal blocks can be understood by an approach which goes back to [MS1, MS2]. From this point of view, the space of conformal blocks of a surface arises as a state space of a three-dimensional Reshetikin-Turaev TFT (see Section 4.4). We will now review this connection for the case of RCFT [MS2, BK].

Let us set, once and for all,

$$\mathcal{C} = \mathrm{Rep}(\mathcal{V}) \; . \tag{5.2.6}$$

If $\mathcal{V}$ is a rational VOA, then $\mathcal{C}$ is a modular category [Hu]. Recall that any modular category defines a Reshetikin-Turaev 3d TFT functor, denoted $\mathrm{tft}_{\mathcal{C}}$ (see section 4.4).

Let Σ^c be a smooth, closed, compact, oriented Riemann surface with n marked points $p_1, \dots, p_n$, marked by representations $R_1, \dots, R_n \in \mathcal{C}$. In order to make contact with topological field theory, Σ^c must be stripped of any non-topological data. This means that we forget the complex and smooth structure as well as the local coordinates around the marked points. In each marked point, an arc germ and an orientation are the only remnants of the local coordinates. We call this (topological) $\mathcal{C}$-marked surface Σ, and we set

$$V_n(\Sigma) = \mathrm{tft}_{\mathcal{C}}(\Sigma) \; . \tag{5.2.7}$$

Hence, if Σ is connected, we have

$$V_n(\Sigma) = \mathrm{Hom}_{\mathcal{C}}\left(R_1 \otimes \cdots \otimes R_n \otimes H^{\otimes g}, \mathbf{1}\right) \; . \tag{5.2.8}$$

In this way, any vector in $V_n(\Sigma)$ is represented by a $\mathcal{C}$-labeled ribbon graph with open ends $R_1, \dots, R_n$ (embedded in a 3-manifold bounded by Σ). In the case $g = 0$, the ribbon graph is simply a coupon labelled by an element of $\mathrm{Hom}_{\mathcal{C}}(R_1 \otimes \cdots \otimes R_n, \mathbf{1})$, embedded in a solid 3-ball in such a way that the ribbons emanating from the coupon are attached to the sphere which bounds the ball.

It should be remembered that $\mathrm{tft}_{\mathcal{C}}(\Sigma)$ was defined using a specific standard smooth $\mathcal{C}$-marked surface Σ_0^s. We may add a complex structure to get a $\mathcal{C}$-marked Riemann surface $\Sigma_0^c \in \hat{\mathcal{M}}_{n,g}$, for which we then have $V_n(\Sigma_0^c) = V_n(\Sigma)$. In order to recover the space of conformal blocks $V_n(\Sigma^c)$ from the TFT-state space $V_n(\Sigma)$, one has to use the isomorphism U_Γ (see (5.2.3)) along a path Γ from Σ_0^c to Σ^c. In this way, interpretation of the additional structure of $\mathcal{C}$ becomes clear (see [MS2] for an abundance of pictorial explanations):

- The tensor product $R \otimes S$ corresponds to bringing two points marked by R and S next to one another. It is often called the fusion product, as it models the fusion of two fields into one.

- The braiding reflects the fact that the position of two marked points can be exchanged smoothly by analytic continuation of the respective conformal blocks (by integrating them along the corresponding path in $\hat{\mathcal{M}}_{g,n}$). Since the conformal blocks are typically not meromorphic, the braiding is not symmetric.

- Reversing the orientation of a marked point corresponds to replacing the associated object by its dual. The dual of a module R over $\mathcal{V}$ is called its contragradient module and is denoted by $R^\vee$. As a graded vector space, it is the graded dual of R. The evaluation map ev_R provides a basis for 2-point conformal blocks at genus 0.

- The ribbon twist describes rotation of the local coordinates (i.e. of the tangent vector) in a given marked point on Σ^c. It acts by $e^{2\pi i L_0}$, which corresponds to a full 2π-rotation of the coordinate frame. Unless the L_0-eigenvalue of the field inserted at that point is integer, the twist is not the identity.

- Modularity: the twist and the braiding together yield a projective action of the modular group $SL(2,\mathbb{Z})$ (the mapping class group of the torus) on $\mathrm{Hom}_{\mathcal{C}}(H,\mathbf{1})$ (the 0-point conformal blocks on the torus). This feature extends, for all surfaces, to an action of their mapping class group on the conformal blocks.

The interpretation of conformal blocks in terms of a monoidal functor provides a natural starting point for Step 2 of the construction of a consistent set of correlators. It is evidently well-suited to address condition (C2) (factorisation). Moreover, modular invariance of a correlator boils down to invariance under the mapping class group action on the conformal blocks involved.

Remark 5.2.2. For any module R there is a non-degenerate pairing $R^\vee \otimes R \to \mathbb{C}$ which is compatible with the grading [FHL]. It should be thought of as the chiral half of a 2-point correlator on the Riemann sphere with punctures at the north and south pole: two copies of this pairing yield together the pairing (5.1.20). For simple objects, we shall denote this pairing by

$$(-,-)_{\bar{i}i} : U_{\bar{i}} \otimes U_i \to \mathbb{C}\ . \tag{5.2.9}$$

5.3. CFT with defects

Before we proceed with Step 2, let us extend the notion of a full closed CFT to a full closed CFT with topological defects. For a general discussion of field theories with defects, see [DKR]. In the case of CFT with defects, we have the following additional data:

- A rigid monoidal category $\mathcal{D}$ (the *category of topological defects*).

- For each $X \in \mathcal{D}$ a vector space $\mathcal{H}_X$ (the *spaces of defect fields*).

Let (Σ, γ) be a smooth, closed, compact, oriented Riemannian 2-manifold as in section 5.1. In CFT with defects, Σ can be decorated with defect lines. A *defect line* on Σ is the oriented image of a smooth map $[0,1] \to \Sigma$, labelled by an object $X \in \mathcal{D}$, which does not intersect itself. Both its endpoints lie on marked points, except if they are identical (in which case the defect line is a loop, and we demand no marked point on it). The orientation of the line may be reversed by simultaneously replacing X by its dual $X^\vee$.

A *topological defect line* is a defect line which can be moved freely on Σ as long as it does not intersect any marked point or other defect line, and as long as its endpoints as well as the tangent vectors in the endpoints remain fixed. Two topological defect lines labelled by X and Y which run parallel to each other may be fused into one line. The resulting defect line is labelled by the tensor product $X \otimes Y$ in $\mathcal{D}$. Any defect line labelled by the tensor unit $\mathbf{1}_\mathcal{D}$ is called *invisible* or *trivial*. It can be added to or removed from Σ at will.

A *defect network* on Σ is a collection of non-intersecting defect lines, which may be joined to each other in their endpoints. A point in which $m \in \mathbb{Z}_{>0}$ defect lines meet is called an m-vertex of the defect network. Each m-vertex joining incoming topological defects $X_1, \ldots, X_m \in \mathcal{D}$ (in this cyclic order) is labelled by a morphism in $\mathrm{Hom}_\mathcal{D}(X_1 \otimes \cdots \otimes X_m, \mathbf{1}_\mathcal{D})$.

A full closed CFT with $\mathcal{D}$-type topological defects is an extension of full closed CFT without defects in the following sense: it does not only associate a correlator to every surface with marked points, but also to every surface with marked points and a defect network. An object in $\mathcal{D}$ should be thought of as labelling a two-sided boundary condition. The correlators must be smooth in the marked points *except* when a marked point crosses a defect line.

More precisely, marked points on Σ may either be located in the *bulk* (between defects) or somewhere on the defect network. A field inserted at a marked point in the bulk is called a *bulk field*, and it is an element in the space $\mathcal{H}$ as before. A field inserted at a marked point on the defect network is called a *defect field*. In the latter case, we assume without loss of generality that the marked point coincides with a vertex of the network. (This has the advantage that the local coordinates provide a natural order of the incoming defect lines.) If the marked point $p \in \Sigma$ is located at an m-vertex with incoming defects $X_1, \ldots, X_m$, then a field inserted at this point is an element of $\mathcal{H}_p := \mathcal{H}_{X_1 \otimes \cdots \otimes X_m}$. Hence, the defect field spaces are cyclic in the tensor factors $X_1, \ldots, X_m$, and they satisfy $\mathcal{H}_{X_1 \otimes \cdots \otimes X_{k-1} \mathbf{1}_\mathcal{D} \otimes X_{k+1} \otimes \cdots \otimes X_m} \cong \mathcal{H}_{X_1 \otimes \cdots \otimes X_{k-1} \otimes X_{k+1} \otimes \cdots \otimes X_m}$ and $\mathcal{H}_{\mathbf{1}_\mathcal{D}} = \mathcal{H}$.

Instead of presenting correlators of defect CFT in the way we introduced them in (5.1.3) (by treating everything as variables), it is more convenient to introduce the notion of a *worldsheet*. A worldsheet is a smooth, closed, compact, oriented Riemann surface (Σ, γ) together with a defect network and a number of marked points with associated tangent vectors. We denote such a decorated surface by $\widehat{\Sigma}$, to distinguish it from the undecorated surface. A closed CFT with topological defects associates to every worldsheet $\widehat{\Sigma}$ (whose marked points are denoted by $p_1, \dots, p_n$) a map (the correlator)

$$\begin{aligned} C_{\widehat{\Sigma}} : \bigotimes_{k=1}^{n} \mathcal{H}_{p_k} &\longrightarrow \mathbb{C} \\ (v_1, \dots, v_n) &\mapsto \langle v_1(p_1) \dots v_n(p_n) \rangle_{\widehat{\Sigma}} \end{aligned} \tag{5.3.1}$$

As before, we will write $\langle v_1(z_1) \dots v_n(z_n) \rangle$ for a correlator of a worldsheet based on the Riemann sphere (with both tangent vectors and defect network suppressed in the notation).

For worldsheets without visible defects, the maps (5.3.1) must reduce to the correlators of a CFT without defects. Correlators of worldsheets with defects must satisfy much the same constraints as correlators without defects (smoothness in the marked points, conformal covariance, factorisation, etc.). The only significant difference is, as mentioned, the appearance of discontinuities when marked points in the bulk cross a defect line. A more detailed description of all the constraints in the case of RCFT can be found in [FjFS].

There is also a non-degeneracy condition for 2-point correlators with defects: the 2-point correlator of the Riemann sphere with marked points at the north and south pole, and with one single defect line labelled by $P \in \mathcal{D}$ connecting the two marked points, must give rise to a grading-compatible non-degenerate pairing

$$(-,-)_P : \ \mathcal{H}_P \otimes \mathcal{H}_{P^\vee} \longrightarrow \mathbb{C} \tag{5.3.2}$$

in the same way as (5.1.19) does.[8] In particular, in RCFT this yields an isomorphism $\mathcal{H}_P \cong \mathcal{H}_{P^\vee}$. More generally, we have isomorphisms $\mathcal{H}_{X_1 \otimes \dots \otimes X_n} \cong \mathcal{H}_{X_n^\vee \otimes \dots \otimes X_1^\vee}$.

Let us also explicitely give one factorisation constraint for genus 0 worldsheets (a generalisation of the OPE (5.1.18)). Consider a defect network on the complex plane with n marked points, and denote the corresponding worldsheet by $\widehat{\Sigma}$. Draw a circle around the origin which does not intersect any marked points. Order the marked points in such a way that $z_1, \dots, z_k$ lie outside the circle, and $z_{k+1}, \dots, z_n$ inside. Denote by $X_1, \dots, X_m \in \mathcal{D}$ (in this cyclic order) the defects which intersect the circle (direction considered from inside to outside). Define the worldsheets $\widehat{\Sigma}_{\text{in}}$ and $\widehat{\Sigma}_{\text{out}}$ as follows:

- $\widehat{\Sigma}_{\text{out}}$ coincides with $\widehat{\Sigma}$ outside the circle; inside the circle it has exactly one marked point in the origin with local coordinates induced by the complex plane, and exactly m defect lines labelled by $X_1, \dots, X_m$, which connect the respective defect intersection points on the circle to the origin without intersecting and without winding.

- $\widehat{\Sigma}_{\text{in}}$ coincides with $\widehat{\Sigma}$ inside the circle; outside the circle it has exactly one marked point in ∞ with local coordinate $1/z$ and exactly m defect lines labelled by $X_1, \dots, X_m$, which connect the respective defect intersection points on the circle to ∞ without intersecting and without winding.

[8] It turns out that for CFTs constructed from the TFT-formalism described in the next section, this is not a condition but a result which follows from non-degeneracy of the 2-point correlator without defects [FjFS].

Note that $\widehat{\Sigma}_{\text{out}}$ has $k+1$ marked points, while $\widehat{\Sigma}_{\text{in}}$ has $n-k+1$ marked points. The factorisation constraints demand that the correlator of $\widehat{\Sigma}$ can be expressed in terms of the correlators of $\widehat{\Sigma}_{\text{in}}$ and $\widehat{\Sigma}_{\text{out}}$ as follows. Let $\{\chi\}$ be a basis of $\mathcal{H}_{X_1 \dots X_m}$, and denote by $\{\tilde{\chi}\}$ the basis of $\mathcal{H}_{X_m^\vee \dots X_1^\vee}$ which is dual to $\{\chi\}$ with respect to the pairing (5.3.2). Then we have

$$\langle v_1(z_1) \dots v_n(z_n) \rangle_{\widehat{\Sigma}} = \sum_{\chi} \langle v_1(z_1) \dots v_k(z_k) \chi(0) \rangle_{\widehat{\Sigma}_{\text{out}}} \langle \tilde{\chi}(\infty) v_1(z_{k+1}) \dots v_n(z_n) \rangle_{\widehat{\Sigma}_{\text{in}}} , \quad (5.3.3)$$

with local coordinates in ∞ given by $z \mapsto 1/z$.

5.4. 3d TQFT description of 2d RCFT

5.4.1. Algebraic data of full RCFT

Starting with the description of conformal blocks in terms of a the monoidal functor $\text{tft}_{\mathcal{C}}$, one can try to go on to Step 2 and construct full CFTs. In the rational case, this program has been carried out in the series of works [FRS1, FRS3, FRS4, FRS5, FjFRS, FrFRS, FjFS], leading to a complete classification of full rational open-closed CFTs: they are given by pairs $(\mathcal{C}, A)$, where $\mathcal{C} = \text{Rep}(\mathcal{V})$ is the modular category of representations of some rational VOA $\mathcal{V}$ (i.e. a chiral CFT) and $A \in \mathcal{C}$ is a special symmetric Frobenius algebra in $\mathcal{C}$. This result is very much in the spirit of the classification of closed 2d TFTs by commutative Frobenius algebras mentioned in section 4.4.[9]

This classification of RCFTs is also a rare example of a fully rigorous incarnation of the holograpic principle (here between 2d CFT and 3d TFT, see [KS] for a discussion of the results from this point of view): the proof is based on an explicit construction of all correlators in terms of Reshetikin-Turaev 3d TFT, which can then be shown to obey all necessary factorisation constraints and modular invariance. In this way, calculations with correlators essentially boil down to manipulations of $\mathcal{C}$-labelled ribbon graphs, which makes the 3dTFT construction of correlators a handy tool to deal with such computations.

Before we turn to the construction, we review the algebraic data which describe defects and fields in the full CFT $(\mathcal{C}, A)$, and which are the building blocks for the construction.

Category $\mathcal{D}$ of topological defects Topological defects in the full RCFT $(\mathcal{C}, A)$ are labelled by A-A-bimodules. Recall the category $\mathcal{D}$ introduced in section 5.3.

If $(\mathcal{C}, A)$ is a full RCFT, then $\mathcal{D} = {}_A\mathcal{C}_A$ is the category of its topological defects.

The defect category $\mathcal{D}$ is rigid with duals and (co)evaluation inherited from $\mathcal{C}$. Moreover, it is monoidal with the tensor product $\otimes_A$ of bimodules. This tensor product describes fusion of two defect lines into one. The tensor unit is $\mathbf{1}_{\mathcal{D}} = A$ as a bimodule over itself,

[9] In fact, the construction yields a closed 2d TFT if one sets $\mathcal{C} \cong \text{Vect}$ (i.e. the chiral CFT contains only the vacuum module) and takes any special symmetric Frobenius algebra $A \in \text{Vect}$. Its center $Z(A)$ is the commutative Frobenius algebra which uniquely defines the 2d TFT as described in section 4.4.

which corresponds to the invisible defect. Note that the fused defects $X \otimes_A Y$ and $Y \otimes_A X$ are in general not isomorphic, which means that $\mathcal{D}$ is typically not braided.

The morphism space

$$\mathcal{H}^{\text{top}}_{XY} := \mathrm{Hom}_{\mathcal{D}}(X, Y) \tag{5.4.1}$$

describes all fields that change the defect condition from X to Y, and which are topological in the sense that they can be moved freely on the worldsheet. More generally, due to the interpretation of the tensor product as defect fusion, the morphism space $\mathrm{Hom}_{\mathcal{D}}(\otimes_{n=1}^{N} X_n, \otimes_{m=1}^{M} Y_m)$ describes topological junctions of $N + M$ topological defects, N of which are incoming, and M are outgoing.

Converting fields into defects Recall the monoidal functors $\alpha^{\pm} : \mathcal{C} \to \mathcal{D}$ defined in (4.2.2). Similarly, there is a braided monoidal functor [Scha]

$$\begin{aligned} \alpha : \mathcal{Z}(\mathcal{C}) &\longrightarrow \mathcal{Z}(\mathcal{D}) \\ (U, \sigma_{U,-}) &\mapsto (\alpha(U), \varphi_{\alpha(U),-}) \ . \end{aligned} \tag{5.4.2}$$

Here, $\alpha(U) := A \otimes U$ is the induced A-bimodule determined by the half-braiding σ in $\mathcal{Z}(\mathcal{C})$: the left action is given by $\rho^L_{\alpha(U)} = \mu \otimes \mathrm{id}_U$ and the right action by $\rho^R_{\alpha(U)} = (\mu \otimes \mathrm{id}_U) \circ (\mathrm{id}_A \otimes \sigma_{U,A})$. The half-braiding $\varphi_{\alpha(U),-}$ is the family of morphisms in $\mathcal{D}$ induced via the universal property of $\otimes_A$ from

$$\tilde{\varphi}_{\alpha(U),-} = (\mathrm{id}_- \otimes \eta \otimes \mathrm{id}_U) \circ (\rho^l \otimes \mathrm{id}_U) \circ (\mathrm{id}_A \otimes \sigma_{U,-}) \ . \tag{5.4.3}$$

Under certain conditions elaborated in [Scha], α is an equivalence.

Recall also the functor $\iota : \mathcal{C} \boxtimes \overline{\mathcal{C}} \to \mathcal{Z}(\mathcal{C})$ from (4.3.8) which is a braided monoidal equivalence if $\mathcal{C}$ is modular. Composing it with α and the forgetful functor $U : \mathcal{Z}(\mathcal{D}) \to \mathcal{D}$, we obtain the sequence of functors

$$\mathcal{C} \boxtimes \overline{\mathcal{C}} \xrightarrow{\iota \,=\, \iota^+ \otimes \iota^-} \mathcal{Z}(\mathcal{C}) \xrightarrow{\alpha} \mathcal{Z}(\mathcal{D}) \xrightarrow{U} \mathcal{D} \ . \tag{5.4.4}$$

It is easy to check that $U \circ \alpha \circ \iota = \alpha^+ \otimes \alpha^-$.

The relevance of the map (5.4.4) for CFT derives from the fact that all field spaces are objects in the category $\mathcal{C} \boxtimes \overline{\mathcal{C}}$. Hence, the sequence (5.4.4) allows us to assign to every field a defect in a natural way. Since this assignment factors through $\mathcal{Z}(\mathcal{D})$, we are also given a natural prescription how to braid the resulting defect past any other defect.

Fields as morphisms In an RCFT $(\mathcal{C}, A)$ with topological defects, the most general point where a field can be inserted is on the junction between two topological defects $X, Y \in \mathcal{D}$. This includes the case of fields which do not change the defect condition ($X = Y$), *twist* (or *disorder*) *fields* which sit at the end of a defect line ($X \neq A$, $Y = A$ or $X = A$, $Y \neq A$), as well as *bulk fields*, which are regarded as fields on the invisible defect ($X = Y = A$). Any junction of more than two topological defects can be reduced to a junction of two defects by fusion.

We denote the space of fields that change a defect condition from X to Y by $\mathcal{H}_{XY}$. As a special case, we have the space of bulk fields $\mathcal{H} = \mathcal{H}_{AA}$. Moreover, we will occasionally use the notation $\mathcal{H}_X = \mathcal{H}_{AX}$ for the space of twist fields which create a defect X. As explained in section 5.1, $\mathcal{H}$ carries a representation of two copies of the chiral algebra $\mathcal{V}$; in other words, it is an object in $\mathcal{C} \boxtimes \overline{\mathcal{C}}$. Since X and Y are topological, the same must also hold for $\mathcal{H}_{XY}$. In fact, it is given by [FRS4]

$$\mathcal{H}_{XY} \;=\; \bigoplus_{i,j\in\mathcal{I}} U_i \otimes_{\mathbb{C}} U_j \otimes_{\mathbb{C}} \mathrm{Hom}_{\mathcal{D}}\big(\alpha^+(U_i) \otimes X \otimes \alpha^-(U_j), Y\big) \;. \tag{5.4.5}$$

Equivalently, this may be written as

$$\mathcal{H}_{XY} \;=\; \bigoplus_{i,j\in\mathcal{I}} (U_i \otimes_{\mathbb{C}} U_j)^{Z_{XY,ij}} \;\in\; \mathcal{C} \boxtimes \overline{\mathcal{C}} \;, \tag{5.4.6}$$

where

$$Z_{XY,ij} \;=\; \dim \mathrm{Hom}_{\mathcal{D}}\big(\alpha^+(U_i) \otimes X \otimes \alpha^-(U_j), Y\big) \;. \tag{5.4.7}$$

Example 5.4.1.

i) For $X = Y = A$ the expression (5.4.6) yields the space of bulk fields 5.1.12, with the constants $Z_{ij} = Z_{XY,ij}$ now determined by the special symmetric Frobenius algebra A.

ii) Since $\alpha^\pm$ is monoidal, we have that $\alpha^\pm(\mathbf{1}) = A$. Hence, restricting the sum (5.4.5) to the vacuum module ($i = j = 0$) yields precisely the subspace of topological defect changing fields $\mathcal{H}^{\mathrm{top}}_{XY} \;=\; \mathbf{1} \otimes_{\mathbb{C}} \mathbf{1} \otimes_{\mathbb{C}} \mathrm{Hom}_{\mathcal{D}}(X, Y)$ as in (5.4.1).

5.4.2. TFT-construction of CFT correlators

We have now assembled the building blocks needed for the TFT-construction of full closed rational CFT with topological defects. The starting point is the pair $(\mathcal{C}, A)$, consisting of a modular category $\mathcal{C}$ (the chiral CFT) and a special symmetric Frobenius algebra $A \in \mathcal{C}$. Recall also the principle of holomorphic factorisation and the interpretation of conformal blocks in terms of state spaces of the functor $\mathrm{tft}_{\mathcal{C}}$ as described in section 4.4. The TFT-construction determines each correlator by the action of $\mathrm{tft}_{\mathcal{C}}$ on a particular $\mathcal{C}$-marked 3-manifold $\widehat{M}$ (a manifold with embedded $\mathcal{C}$-labelled ribbon graph). The underlying manifold M depends only on the topology of the world sheet without field insertions or defects, while the ribbon graph contains the information about the fields and defects which decorate the worldsheet.

In this section, we will briefly review this construction. We start with a worldsheet Σ, decorated with marked points labelled by objects in $\mathcal{C} \boxtimes \overline{\mathcal{C}}$ and topological defect lines labelled by objects in $\mathcal{D} = {}_A\mathcal{C}_A$.

The 3-manifold M The manifold M (so far without any embedded ribbon graph) is simply a thickening of the underlying worldsheet:

$$M := \Sigma \times [-1,1] \tag{5.4.8}$$

The worldsheet is naturally embedded in M by the map $p \mapsto (p,0)$. The surface $\partial M = \Sigma \times \{-1,1\}$ is called the *complex double* of the worldsheet. Its two disjoint components are taken to have opposite orientation, in such a way that the orientation of the component $\Sigma \times \{1\}$ agrees with the orientation of the embedded worldsheet.

The $\mathcal{C}$**-marked surface** $\partial\widehat{M}$ The complex double ∂M becomes a $\mathcal{C}$-marked surface in the following way:

- Let $p \in \Sigma$ be a marked point on the worldsheet where a field in representation $R \otimes \bar{R} \in \mathcal{C} \boxtimes \overline{\mathcal{C}}$ is inserted. Then $(z,+1)$ is a marked point on the first component of ∂M, labelled by the chiral representation R, while $(z,-1)$ is a marked point on the second component of ∂M, labelled by the anti-chiral representation $\bar{R}$. Moreover, the tangent vector in p is canonically inherited by both points (flipped on the second component, in order to comply with the opposite orientation of that boundary component).
- defect lines have no effect on $\partial\widehat{M}$.

The $\mathcal{C}$**-marked 3-manifold** $\widehat{M}$ The 3-manifold M with $\mathcal{C}$-marked boundary surface $\partial\widehat{M}$ is turned into a $\mathcal{C}$-marked manifold by defining a ribbon graph embedded in M, whose free ends connect to the $\mathcal{C}$-labelled marked points on $\partial\widehat{M}$. The construction of this ribbon graph is described in detail in [FRS4]. We give here just the main points (which will be followed up by an example):

- Triangulate the embedded world sheet in M by a connected planar graph, such that
  - all defect lines are covered by edges,
  - all marked points are covered by vertices.
- On all edges of this graph, place a ribbon (facing up with respect to the orientation of the worldsheet).
  - on edges coinciding with a defect line labelled by $X \in \mathcal{D}$, place a ribbon labelled by the underlying object of X in $\mathcal{C}$. The direction of the ribbon is determined by the orientation of the defect line.
  - on all remaining edges, place a ribbon labelled by A. For the conventions regarding their orientation we refer to [FRS4].
- On all vertices of the graph which do not coincide with a marked point, place a coupon representing a morphism in $\mathcal{D}$ connecting the adjacent ribbons. In particular,

– on trivalent vertices where an A-ribbon hits an X-ribbon from the left (right), insert the left (right) module action of A on X,

– on m-valent vertices where an X-ribbon is converted into a Y-ribbon, insert a bimodule intertwiner $X \to Y$.

- On a vertex converting an X-ribbon into a Y-ribbon, which additionally coincides with a marked point labelled by $R \otimes_{\mathbb{C}} \bar{R} \in \mathcal{C} \boxtimes \overline{\mathcal{C}}$, insert the morphism in $\mathrm{Hom}_{\mathcal{D}}(\alpha^+(R) \otimes X \otimes \alpha^-(\bar{R}), Y)$ which represents the field in this marked point according to (5.4.5). This creates two open ends of the ribbon graph, labelled by R and $\bar{R}$, respectively.
- Finally, attach the $2n$ open ends of this (so far) planar ribbon graph to the marked points on $\partial\widehat{M}$, creating a 3-dimensional ribbon graph inside M: the ribbon coloured by the chiral representation R emanating from the marked point $(z, 0)$ is attached to the point $(z, 1)$, while the ribbon coloured by the anti-chiral representation $\bar{R}$ emanating from the same marked point $(z, 0)$ is attached to the point $(z, -1)$. The intersection of the ribbons with $\partial\widehat{M}$ in the marked points $(z, \pm 1)$ coincides with the tangent vectors at these points. For the exact conventions, we refer again to [FRS4].

This completes the description of the $\mathcal{C}$-marked manifold $\widehat{M}$. The construction of the ribbon graph involves several rather arbitrary choices (the triangulation, the orientations of the ribbons, as well as some of the morphisms involving only A). The construction of the correlators is independent of these choices.

The correlator Consider the space of $2n$-point conformal blocks $V_{2n}(\partial\widehat{M})$ associated to the $\mathcal{C}$-marked surface $\partial\widehat{M}$, which according to the TFT axioms factorises into chiral and anti-chiral blocks as follows:

$$\begin{aligned} V_{2n}(\partial\widehat{M}) &= \mathrm{tft}_{\mathcal{C}}(\partial\widehat{M}) = \mathrm{tft}_{\mathcal{C}}(\partial_+\widehat{M} \sqcup \partial_-\widehat{M}) = \mathrm{tft}_{\mathcal{C}}(\partial_+\widehat{M}) \otimes \mathrm{tft}_{\mathcal{C}}(\partial_-\widehat{M}) \\ &= V_n(\partial_+\widehat{M}) \otimes V_n(\partial_-\widehat{M}) \end{aligned} \tag{5.4.9}$$

According to (5.2.8), this space is thus, restricted to representations $R_1, \ldots, R_n \in \mathcal{C}$ and $\bar{R}_1, \ldots, \bar{R}_n \in \mathcal{C}$, given by

$$V_{2n}(\partial\widehat{M}) = \mathrm{Hom}_{\mathcal{C}}(R_1 \otimes \cdots \otimes R_n \otimes H^{\otimes g}, \mathbf{1}) \otimes \mathrm{Hom}_{\mathcal{C}}(\bar{R}_1 \otimes \cdots \otimes \bar{R}_n \otimes H^{\otimes g}, \mathbf{1}) \ . \tag{5.4.10}$$

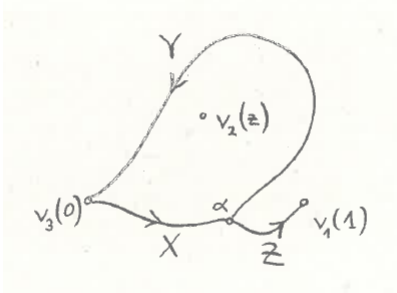
Recall from section 5.2 that any n-point correlator (irrespective of defect lines) for fields in representations $R_1 \otimes_{\mathbb{C}} \bar{R}_1, \ldots, R_n \otimes_{\mathbb{C}} \bar{R}_n$ is an element in this space. The correlator for the specific worldsheet at hand is constructed from the $\mathcal{C}$-marked 3-manifold $\widehat{M}$ as follows.

Regard $\widehat{M}$ as a cobordism $\widehat{M} : \emptyset \to \partial\widehat{M}$, and apply the 3d TFT functor to it. This yields a linear map $\mathrm{tft}_{\mathcal{C}}(\widehat{M}) : \mathbb{C} \to V_{2n}(\partial\widehat{M})$ which can be evaluated in 1, defining the multilinear form

$$C_{\hat{\Sigma}}[R_1, \ldots, R_n; \bar{R}_1, \ldots \bar{R}_n] \ := \ \mathrm{tft}_{\mathcal{C}}(\widehat{M})\, 1 \ \in \ V_{2n}(\partial\widehat{M}) \ . \tag{5.4.11}$$

Evaluating this form on the fields yields a number, which is precisely the correlator.

Example 5.4.2. The construction is best understood by looking at an example. Let us compute a 3-point correlator $\langle v_1(1) v_2(z) v_3(0) \rangle$ on the Riemann sphere decorated with marked points and defects as follows:



(5.4.12)

The trivalent defect vertex α is given by a morphism

$$\alpha \in \mathrm{Hom}_{\mathcal{D}}(X, Y \otimes Z) \tag{5.4.13}$$

The fields $v_1 \in \mathcal{H}_{ZA}$, $v_2 \in \mathcal{H}$ and $v_3 \in \mathcal{H}_{YX}$ are given by the following data:

- representations $R_1, R_2, R_3, \bar{R}_1, \bar{R}_2, \bar{R}_3 \in \mathcal{C}$
- for $k \in \{1, 2, 3\}$, vectors $v_k^L \in R_k$ and $v_k^R \in \bar{R}_k$ (we assume that v_k are elementary tensors).
- morphisms

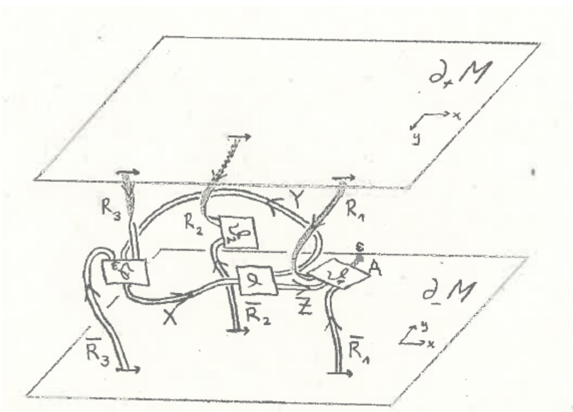
$$\vartheta_1 \in \mathrm{Hom}_{\mathcal{D}}\big(\alpha^+(R_1) \otimes Z \otimes \alpha^-(\bar{R}_1), A\big) \tag{5.4.14}$$
$$\vartheta_2 \in \mathrm{Hom}_{\mathcal{D}}\big(\alpha^+(R_2) \otimes \alpha^-(\bar{R}_2), A\big) \tag{5.4.15}$$
$$\vartheta_3 \in \mathrm{Hom}_{\mathcal{D}}\big(\alpha^+(R_3) \otimes Y \otimes \alpha^-(\bar{R}_3), X\big) \tag{5.4.16}$$

Put together, we have

$$v_k = v_k^L \otimes v_k^R \otimes \vartheta_k \quad , \qquad\qquad k = 1, 2, 3 \ . \tag{5.4.17}$$

The $\mathcal{C}$-marked 3-manifold with embedded ribbon graph looks as follows:[10]

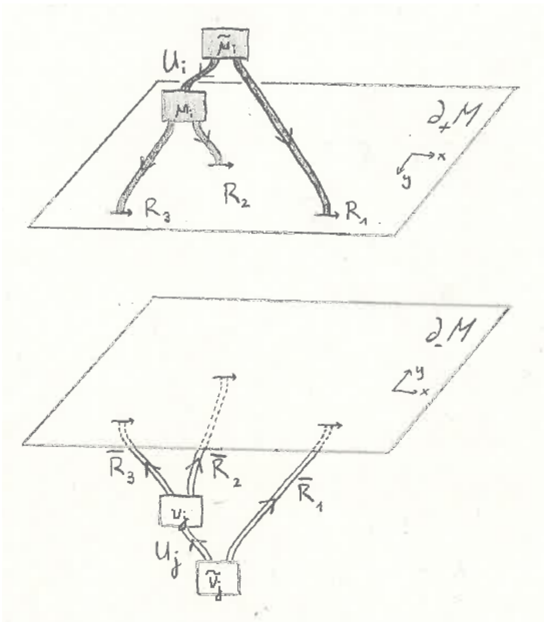
$\widehat{M} =$



(5.4.18)

[10] Note that we used the special properties of A to reduce the number of A-ribbons to a mininum.

We would like to express the correlator $\langle v_1(1)v_2(z)v_3(0)\rangle$ in terms of conformal blocks. Pick a basis of $V_3(\overline{\mathbb{C}}) \otimes V_3(\overline{\mathbb{C}})$, for example the one obtained as the image under $\mathrm{tft}_{\mathcal{C}}$ of the following ribbon graphs embedded in the complement of $\widehat{M}$ with respect to S^3:



(5.4.19)

Here, the indices i and j run over $\mathcal{I}$, while μ_i and $\tilde{\mu}_i$ run over bases of $\mathrm{Hom}(R_3^\vee \otimes R_2^\vee, U_i^\vee)$ and $\mathrm{Hom}(U_i^\vee \otimes R_1^\vee, \mathbf{1})$, and ν_j and $\tilde{\nu}_j$ run over bases of $\mathrm{Hom}(\bar{R}_2^\vee \otimes \bar{R}_3^\vee, U_j^\vee)$ and $\mathrm{Hom}(\bar{R}_1^\vee \otimes U_j^\vee, \mathbf{1})$, respectively. Let us combine the indices i, μ_i and $\tilde{\mu}_i$ into a multi-index α, and the indices j, ν_j and $\tilde{\nu}_j$ into a multi-index $\bar{\alpha}$. Denote by $b_\alpha \in V_3(\overline{\mathbb{C}})$ and $b_{\bar{\alpha}} \in V_3(\overline{\mathbb{C}})$ the conformal blocks such that $b_\alpha \otimes b_{\bar{\alpha}} = \mathrm{tft}_{\mathcal{C}}((5.4.19))$. They define indeed a basis of $V_3(\overline{\mathbb{C}}) \otimes V_3(\overline{\mathbb{C}})$. Write

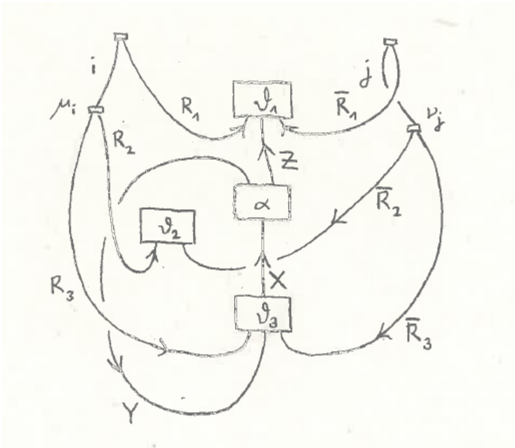
$$F_\alpha(1,z,0) = b_\alpha(v_1^L, v_2^L, v_3^L) \quad \text{and} \quad F_{\bar{\alpha}}(1,\bar{z},0) = b_{\bar{\alpha}}(v_1^R, v_2^R, v_3^R) \ . \tag{5.4.20}$$

Finally, the correlator can be expressed in this basis as follows:

$$\langle v_1(1)v_2(z)v_3(0)\rangle \; = \; \sum_{\alpha,\bar{\alpha}} \Omega_{\alpha,\bar{\alpha}} F_\alpha(1,z,0) F_{\bar{\alpha}}(1,\bar{z},0) \tag{5.4.21}$$

where $\Omega_{\alpha,\bar{\alpha}} \in \mathbb{C}$ is the “matrix element” given by the following morphism $\mathbf{1} \to \mathbf{1}$ in $\mathcal{C}$

(recall that $\mathrm{End}_{\mathcal{C}}(\mathbf{1}) \cong \mathbb{C}$):

$\Omega_{\alpha,\bar{\alpha}} =$  (5.4.22)

6. Perturbed twisted defect operators in RCFT

6.1. Twisted defect operators

In this chapter, we study defect operators for RCFT. We fix a full RCFT as described in section 5.4.1, namely given by the modular category $\mathcal{C} = \mathrm{Rep}(\mathcal{V})$ of representations of a rational chiral algebra, together with a special symmetric Frobenius algebra $A \in \mathcal{C}$. Throughout the chapter, we denote by $\mathcal{D} = {}_A\mathcal{C}_A$ the category of topological defects of the theory.

Operators from pairings

Let us recall some facts about bilinear pairings of graded vector spaces, and the operators that arise naturally from comparison of two such pairings. Let U, V be graded vector spaces. We call a linear map $f : U \to V$ *degree-preserving*, if $\deg(f(u)) = \deg(u)$ for all homogeneous $u \in U$. Moreover, we call a bilinear pairing $b : U \times V \to \mathbb{C}$ *compatible with the grading*, if $b(u, v) = 0$ whenever $\deg(u) \neq \deg(v)$.

Lemma 6.1.1. *Let $U = \bigoplus_{z\in\mathbb{C}} U_z$ and $V = \bigoplus_{z\in\mathbb{C}} V_z$ be $\mathbb{C}$-graded vector spaces with finite-dimensional homogeneous components. Let*

$$b_1, b_2 : U \times V \to \mathbb{C} \tag{6.1.1}$$

be bilinear pairings compatible with the grading. If b_1 is non-degenerate in both arguments, then there exists a unique linear map $B : V \to V$, such that

$$b_2(u, v) = b_1(u, Bv) \qquad \forall\ u \in U\ ,\ \ v \in V\ . \tag{6.1.2}$$

Moreover, B is degree-preserving.

Proof. For $j = 1, 2$, consider the linear map

$$\beta_j : V \to U^\star\ , \qquad (\beta_j(v))(u) = b_j(u, v) \tag{6.1.3}$$

Since the pairings are compatible with the grading, we have

$$(\beta_j)_z\ :=\ \beta_j\Big|_{V_z} \subseteq U_z^\star \tag{6.1.4}$$

for all $z \in \mathbb{C}$, and we can define

$$\hat{\beta}_j : V \to U^\vee \ , \qquad \hat{\beta}_j(v) = \beta_j(v) \ . \tag{6.1.5}$$

Moreover, the pairing b_1 restricts to a non-degenerate pairing on $U_z \times V_z$. Since U_z and V_z are finite-dimensional, it follows that $(\beta_1)_z$ are, in fact, isomorphisms. Consequently, the map $\hat{\beta}_1$ is invertible, and the degree-preserving map

$$B = \hat{\beta}_1^{-1} \circ \hat{\beta}_2 \tag{6.1.6}$$

is the unique linear map satisfying $\hat{\beta}_1 \circ B = \hat{\beta}_2$. The claim follows immediately from the definition of β_j. □

Remark 6.1.2. It may seem as though compatibility of b_2 with the grading is not needed for the proof to work, since we do not use the inverse of β_2. This is only partially true. If we drop this condition, then the image of β_2 can be larger than the domain of $\hat{\beta}^{-1}$, making (6.1.6) ill-defined. This issue can be fixed by using the degree-wise extension of β_1 to the completion of V, denoted by $\overline{\beta}_1 : \overline{V} \longrightarrow U^\star$, and defining

$$B = \overline{\beta}_1^{-1} \circ \hat{\beta}_2 : \ V \longrightarrow \overline{V} \ . \tag{6.1.7}$$

Since b_1 is non-degenerate, it extends trivially to a pairing $\bar{b}_1 : U \times \overline{V} \to \mathbb{C}$. The map (6.1.7) satisfies the condition $b_2(u, v) = \bar{b}_1(u, Bv)$. Notice that B is not degree-preserving.

Twisted 2-point correlator

Consider the Riemann sphere with 2 marked points connected by a topological defect line $P \in \mathcal{D}$. We represent this situation in the complex z-plane, with the marked points sitting at $z = 0$ and $z = L \in \mathbb{R}_+$. Since P is a topological defect, it may be deformed freely except at its endpoints. The tangent vectors in $z = 0$ and $z = L$ remain fixed, which ensures that the total winding number of P remains the same. By convention, we take P to run along the real axis from $z = 0$ to $z = L$, with tangent vectors in positive real direction. We call this worldsheet the *P-twisted cylinder*, and P the *twist defect*. The *twist fields* sitting at the marked points live in the space $\mathcal{H}_{AP}$ and $\mathcal{H}_{PA}$, respectively. According to (5.4.5), these field spaces may be written as

$$\mathcal{H}_{AP} \ = \ \bigoplus_{i,j \in \mathcal{I}} (U_i \otimes_{\mathbb{C}} U_j) \otimes_{\mathbb{C}} \mathrm{Hom}_{\mathcal{D}} \big(\alpha^+(U_i) \otimes \alpha^-(U_j), P\big) \ , \tag{6.1.8}$$

$$\mathcal{H}_{PA} \ = \ \bigoplus_{i,j \in \mathcal{I}} (U_i \otimes_{\mathbb{C}} U_j) \otimes_{\mathbb{C}} \mathrm{Hom}_{\mathcal{D}} \big(\alpha^+(U_i) \otimes P \otimes \alpha^-(U_j), \mathbf{1}\big) \ . \tag{6.1.9}$$

Let $u \in \mathcal{H}_{PA}$ and $v \in \mathcal{H}_{AP}$. Without loss of generality, we will write these vectors as elementary tensors as follows

$$u = \sum_{i,j \in \mathcal{I}} (u_i \otimes u_j) \otimes \tilde{\phi}_{ij} \qquad\qquad v = \sum_{i,j \in \mathcal{I}} (v_i \otimes v_j) \otimes \phi_{ij} \ , \tag{6.1.10}$$

where

$$u_i \otimes u_j \in (U_i \otimes_{\mathbb{C}} U_j) \ , \qquad \tilde{\phi}_{ij} \in \mathrm{Hom}_{\mathcal{D}}\big(\alpha^+(U_i) \otimes P \otimes \alpha^-(U_j), \mathbf{1}\big) \ , \tag{6.1.11}$$

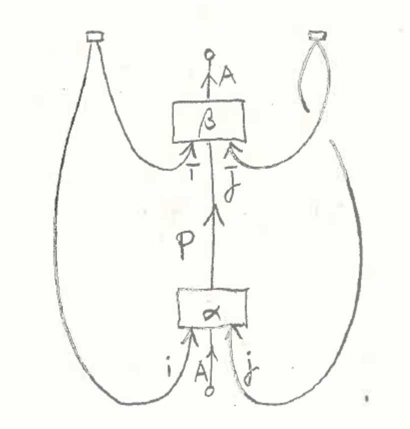
$$v_i \otimes v_j \in (U_i \otimes_{\mathbb{C}} U_j) \ , \qquad \phi_{ij} \in \mathrm{Hom}_{\mathcal{D}}\big(\alpha^+(U_i) \otimes \alpha^-(U_j), P\big) \ . \tag{6.1.12}$$

According to the TFT-construction, the correlator of the P-twisted cylinder is given by

$$\langle u(0)v(L)\rangle_P \ = \ \sum_{i,j\in\mathcal{I}} \Omega^P_{ij}\left[\tilde{\phi}_{\bar{i}\bar{j}}, \phi_{ij}\right] \cdot F_{\bar{i}i}(0,L) F_{\bar{j}j}(0,L) \ , \tag{6.1.13}$$

where $F_{\bar{i}i}$ and $F_{\bar{j}j}$ are the chiral and anti-chiral 2-point conformal blocks for the pairs $(u_{\bar{i}}, v_i)$ and $(u_{\bar{j}}, v_j)$, respectively, and $\Omega^P_{ij}[\beta,\alpha] \in \mathbb{C}$ is the ribbon invariant obtained from the following morphism in $\mathcal{C}$ (recall that $\mathrm{End}_{\mathcal{C}}(\mathbf{1}) \cong \mathbb{C}$):

$$\Omega^P_{ij}[\beta,\alpha] \ = \tag{6.1.14}$$



The linear map

$$\Omega^P_{ij} : \ \mathrm{Hom}_{\mathcal{D}}\big(U_{\bar{i}} \otimes P \otimes U_{\bar{j}}, \mathbf{1}\big) \times \mathrm{Hom}_{\mathcal{D}}(U_i \otimes U_j, P) \longrightarrow \mathbb{C}$$
$$(\alpha,\beta) \ \longmapsto \ \Omega^P_{ij}[\alpha,\beta] \tag{6.1.15}$$

defines a non-degenerate pairing [FjFS, App .A4]. Recall also from chapter 5 (see Remark 5.2.2 and (5.3.2)) that that the 2-point conformal blocks for $L \to \infty$ give rise to non-degenerate pairings

$$(-,-)_{\bar{i}i} : U_{\bar{i}} \otimes U_i \longrightarrow \mathbb{C} \ . \tag{6.1.16}$$

Combination of these facts ensures that the following is a non-degenerate pairing:

$$(-,-)_P : \ \ \mathcal{H}_{PA} \otimes \mathcal{H}_{AP} \longrightarrow \mathbb{C}$$
$$(u,v) \ \longmapsto \ \sum_{i,j\in\mathcal{I}} \Omega^P_{ij}\left[\tilde{\phi}_{\bar{i}\bar{j}}, \phi_{ij}\right] \cdot (u_{\bar{i}}, v_i)_{\bar{i}i} (u_{\bar{j}}, v_j)_{\bar{j}j} \tag{6.1.17}$$

This pairing yields the structure constants of P-twisted 2-point correlators. It is compatible with the grading induced by $L_0 + \bar{L}_0$, and it will be our standard pairing (b_1 in the context of Lemma 6.1.1) to which other pairings $\mathcal{H}_{PA} \otimes \mathcal{H}_{AP} \to \mathbb{C}$ will be compared by means of a linear map $\mathcal{H}_{AP} \to \mathcal{H}_{AP}$.

Notations 6.1.3. We will abbreviate

$$\mathcal{H}_P := \mathcal{H}_{AP} \ . \tag{6.1.18}$$

Twisted defect operators

Now we add to the P-twisted cylinder another topological defect line $X \in \mathcal{D}$, which is wrapped once around the marked point $z = 0$ in clockwise direction. By convention, we place the new defect on the unit circle, such that the defects X and P cross each other in $z = 1$. We choose a morphism

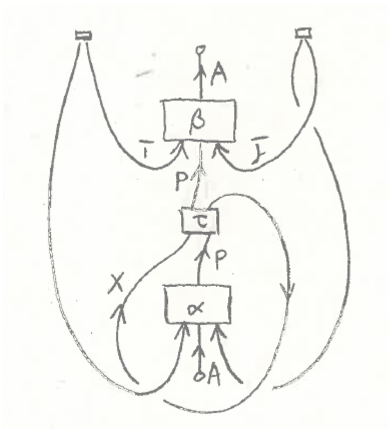
$$\tau \in \mathrm{Hom}_{\mathcal{D}}(X \otimes P, P \otimes X) \tag{6.1.19}$$

representing this crossing. Since everything in $\mathcal{D}$ is topological, we can of course deform X and move τ up and down the twist-defect as we see fit. In the same way as above, the 2-point correlator for this worldsheet gives rise to a non-degenerate pairing

$$(-,-)_{P,X}: \quad \mathcal{H}_{PA} \otimes \mathcal{H}_{AP} \longrightarrow \mathbb{C} \tag{6.1.20}$$

The only difference with (6.1.17) is that the pairing Ω^{P}_{ij} is generalised to the pairing Ω^{PX}_{ij}, which is defined via the following morphism in $\mathcal{C}$:

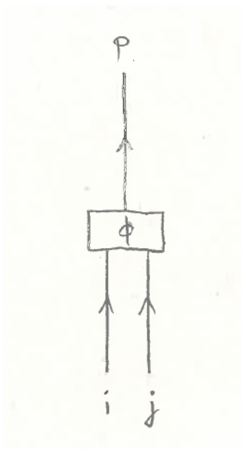
$$\Omega^{PX}_{ij}[\beta,\alpha] \ = \tag{6.1.21}$$



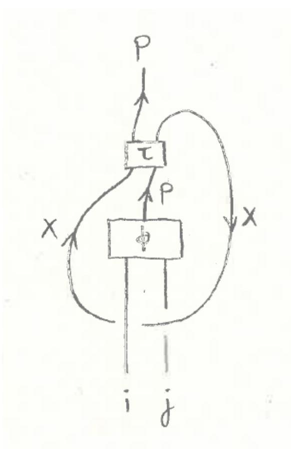
Applying Lemma 6.1.1 to the pairings $b_1 = (-,-)_P$ and $b_2 = (-,-)_{P,X}$, we obtain a unique linear map

$$D[X|P,\tau] : \mathcal{H}_P \longrightarrow \mathcal{H}_P \ , \tag{6.1.22}$$

which in case of the trivial defect $X = A$ reduces to the identity. We call (6.1.22) the *P-twisted defect operator of X*. The action of this operator on $\phi_{ij} \in \mathrm{Hom}_{\mathcal{D}}(U_i \otimes U_j, P)$ is



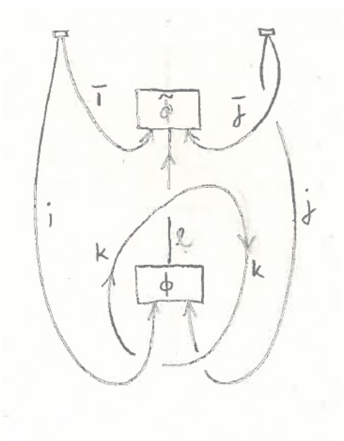
$$\mapsto \tag{6.1.23}$$



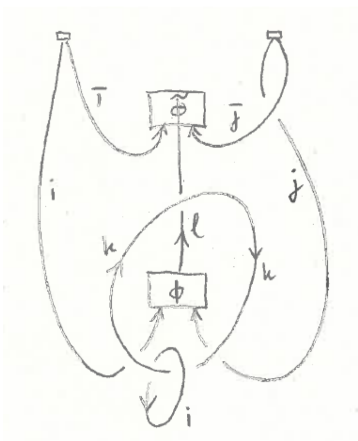
Example 6.1.4. Consider a charge-conjugate RCFT (i.e. the Cardy case $A = \mathbf{1}$, from which follows $\mathcal{D} = \mathcal{C}$). Let $U_k, U_l \in \mathcal{C}$ be simple objects. Set $X = U_k$ and $P = U_l$. If we take the crossing τ to equal the braiding c_{kl} of U_k and U_l, then the P-twisted defect operator of X has the explicit form

$$D[U_k|U_l, c_{kl}] \;=\; \sum_{i,j\in\mathcal{I}} \frac{S_{\bar{k}i}}{S_{0i}}\, \mathrm{id}_{U_i\otimes U_j\otimes \mathrm{Hom}(U_i\otimes U_j, U_k)} \;, \tag{6.1.24}$$

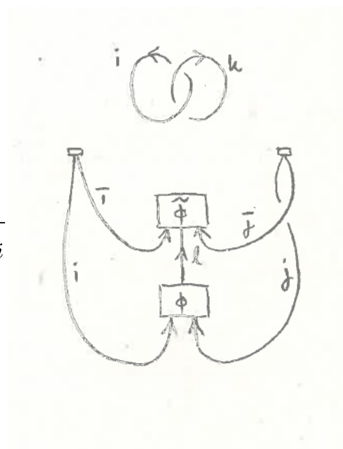
where S_{mn} is the modular matrix of $\mathcal{C}$. This follows from straight-forward comparison of the morphisms $\Omega_{ij}^{U_l U_k}$ and $\Omega_{ij}^{U_l}$:



$$= \frac{1}{S_{0i}}$$



$$= \frac{1}{S_{0i}}$$



(6.1.25)

Here, we used the identity $(\mathrm{id}_i \otimes \widetilde{\mathrm{ev}}_k) \circ ((c_{ki} \circ c_{ik}) \otimes \mathrm{id}_{\bar{k}}) \circ (\mathrm{id}_i \otimes \mathrm{coev}_k) = \frac{S_{ik}}{S_{0i}} \mathrm{id}_i$ which holds in any modular tensor category [BK, Lem. 3.1.4]. If τ is given by the inverse braiding instead, the result is

$$D[U_k|U_l, c_{kl}^{-1}] \;=\; \sum_{i,j\in\mathcal{I}} \frac{S_{kj}}{S_{0j}}\, \mathrm{id}_{U_i\otimes U_j\otimes \mathrm{Hom}(U_i\otimes U_j, U_k)} \;. \tag{6.1.26}$$

Composition and superposition of defect operators

Now let us discuss how some of the structure of the category $\mathcal{D}$ translates to the set of twisted defect operators.

Consider two defects $X, Y \in \mathcal{D}$, together with crossings $\tau_{X,P} : X \otimes P \to P \otimes X$ and $\tau_{Y,P} : Y \otimes P \to P \otimes Y$. The composition of the corresponding twisted defect operators is again a defect operator. As an immediate consequence of (6.1.23), it is compatible with the tensor product in $\mathcal{D}$:

Lemma 6.1.5. *We have*

$$D[X|P, \tau_{X,P}]\, D[Y|P, \tau_{Y,P}] \;=\; D[X \otimes Y|\tau_{X\otimes Y,P}] \;, \tag{6.1.27}$$

where

$$\tau_{X\otimes Y,P} \;=\; (\tau_{X,P} \otimes \mathrm{id}_Y) \circ (\mathrm{id}_X \otimes \tau_{Y,P}) \;. \tag{6.1.28}$$

Now consider the superposition of twisted defect operators. It is compatible with the direct sum in $\mathcal{D}$ as follows. Denote by $\iota_X : X \to X \oplus Y$ and $\iota_Y : Y \to X \oplus Y$ the canonical injections and by $\pi_X : X \oplus Y \to X$ and $\pi_Y : X \oplus Y \to Y$ the canonical projections.

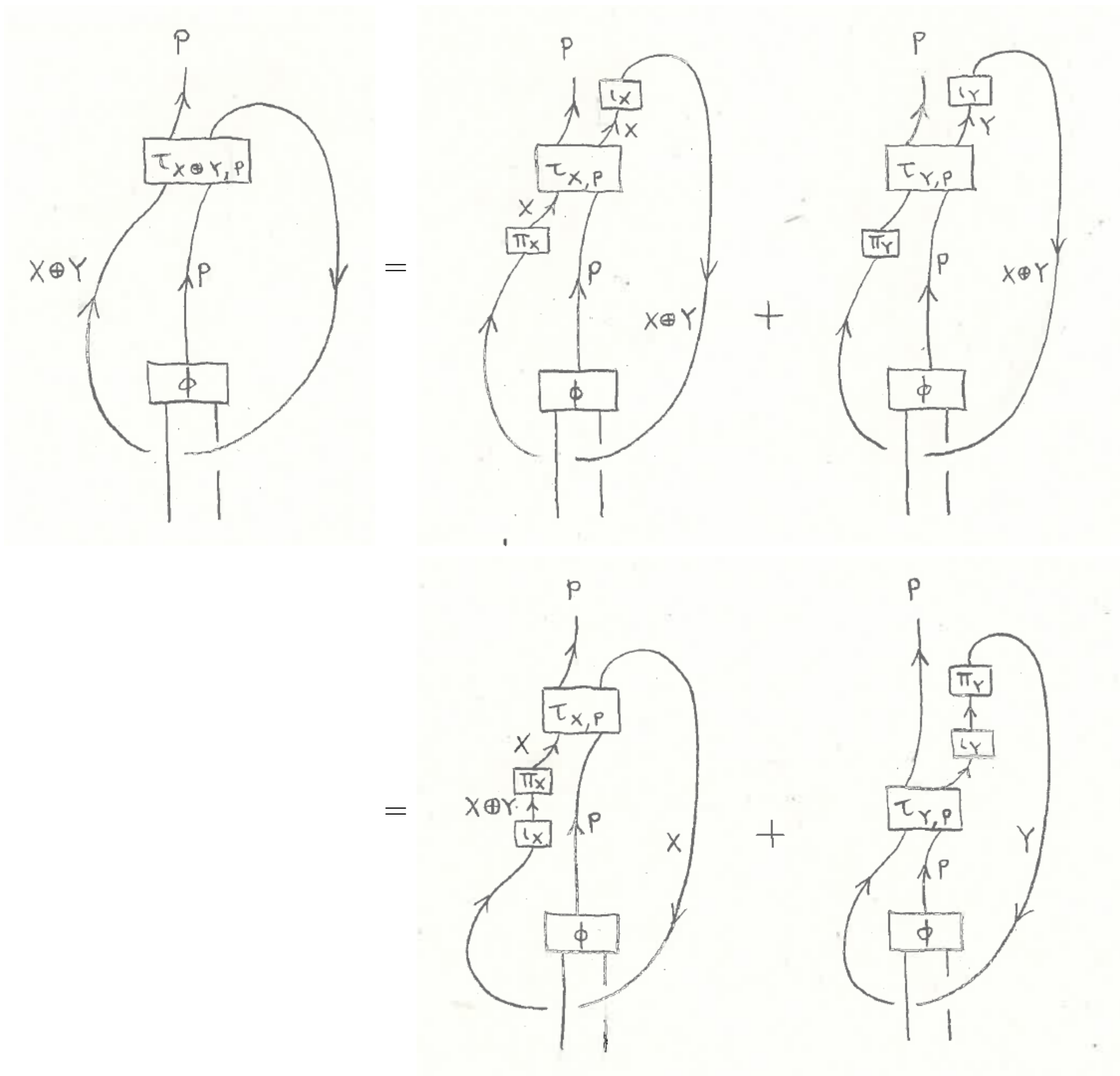
Lemma 6.1.6. *We have*

$$D[X|P, \tau_{X,P}] + D[Y|P, \tau_{Y,P}] \;=\; D[X \oplus Y|P, \tau_{X\oplus Y,P}] \tag{6.1.29}$$

where

$$\tau_{X\oplus Y,P} \;=\; (\mathrm{id}_P \otimes \iota_X) \circ \tau_{X,P} \circ (\pi_X \otimes \mathrm{id}_P) \;+\; (\mathrm{id}_P \otimes \iota_Y) \circ \tau_{Y,P} \circ (\pi_Y \otimes \mathrm{id}_P)\;. \tag{6.1.30}$$

Proof. We start with the pairing Ω_{ij}^{PX} for the right-hand-side of (6.1.29).



(6.1.31)

In the first equality, (6.1.30) was substituted for $\tau_{X\oplus Y,P}$. In the second equality, we used rigidity to move ι_X and π_Y around the defect loop. Finally, the projections and inclusions can be eliminated from the last line by using $\pi_X \circ \iota_X = \mathrm{id}_X$ and $\pi_Y \circ \iota_Y = \mathrm{id}_Y$. This results in the left hand side of (6.1.29). □

Example 6.1.7. A natural choice for $\tau_{X,P}$ is the braiding, if available. In the Cardy case ($\mathcal{D} = \mathcal{C}$) we can set $\tau_{X,P} = c_{X,P}$ and $\tau_{Y,P} = c_{Y,P}$, from which we obtain $\tau_{X\otimes Y,P} = c_{X\otimes Y,P}$ and $\tau_{X\oplus Y,P} = c_{X\oplus Y,P}$. We could also take the inverse braiding instead.

6.2. Perturbed twisted defect operators

We will now introduce a perturbed version of the X-defect on the P-twisted cylinder. This will give rise to a continuous family of twisted defect operators, which in some cases remain integrals of motion.

Fix a subsector $F \subseteq \mathcal{H}_{XX}$ and a defect field $\psi \in F$. Let $f \in \mathrm{Hom}_{\mathcal{D}}(\alpha\iota(F) \otimes X, X)$ represent this defect field in accordance with (5.4.5). Let $L \in \mathbb{R}$ be large and fix some $\theta \in [0, 2\pi)$. Consider the following expression for $\lambda \in \mathbb{C}$ (the perturbation strength):

$$\Big\langle\, u(Le^{i\theta})\; \mathcal{P}\left\{e^{\lambda \int_X \psi(z)\,dz}\right\}\; v(0)\Big\rangle_{P,X}$$
$$:= \sum_{n=0}^{\infty} \lambda^n \int_0^{2\pi} d\theta_1 \int_{\theta_1}^{2\pi} d\theta_2 \cdots \int_{\theta_{n-1}}^{2\pi} d\theta_n \,\big\langle\, u(Le^{i\theta})\, \psi(e^{i\theta_1}) \ldots \psi(e^{i\theta_n})\, v(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)} \tag{6.2.1}$$

The correlators on the right-hand-side are defined as follows. The defect X is placed on the unit circle. The twist-defect P runs along a straight ray from 0 to $Le^{i\theta}$, crossing X in the point $e^{i\theta}$ by means of the crossing $\tau_{X,P}$. For the local coordinates around the marked points we choose the following convention: the tangent vectors in the points $e^{i\theta_k}$ are tangent to the defect X, and the tangent vectors in 0 and $Le^{i\theta}$ are tangent to the defect P.

We assume that all the integrals in (6.2.1) exist and that the sum has a finite radius of convergence in λ. For $\lambda = 0$, the expression clearly reduces to the 2-point correlator $\langle u(Le^{i\theta})v(0)\rangle_{X,P}$. In this chapter we explore the idea of producing from the perturbed twisted defect correlator (6.2.1) a perturbed twisted defect operator along the same lines as in the previous section, by letting $L \to \infty$ and comparing it to the pairing (6.1.17).

The category of perturbed defects

In the situation above, the defect perturbation is encapsulated in the following data:[1]

- The defect $X \in \mathcal{D}$.
- The representation $F \in \mathcal{C} \boxtimes \overline{\mathcal{C}}$ from which the perturbing field is taken.

[1] The complete data include of course also the field $\psi \in F$. This is the datum which the categorical approach disregards.

- The morphism $f : \alpha\iota(F) \otimes X \to X$ by which F perturbs X.

This data extends to a monoidal category of topological defects perturbed by fields in representation F: the category $\mathcal{D}_{\alpha\iota F}$ of Definition 4.2.4 (recall also Notations 4.2.5).

Notations 6.2.1. We take the liberty to abbreviate $\mathcal{D}_F := \mathcal{D}_{\alpha\iota F}$. We refer to $\mathcal{D}_F$ as the *category of F-perturbed defects* and we call F the *defect perturbation* for short.

The morphisms in $\mathcal{D}_F$ correspond to topological defect changing fields which remain topological under the perturbation. This is ensured by condition 4.2.3, which demands that morphisms representing defect changing fields commute with the "action" f of the perturbation.

In the literature, typically one considers defect perturbations of the form

$$F \;=\; F^+ \otimes \mathbf{1} \;\oplus\; \mathbf{1} \otimes F^- \;, \tag{6.2.2}$$

where $F^\pm \in \mathcal{C}$. In this case, the morphism f splits into a sum of morphisms $f^+ : F^+ \otimes X \to X$ and $f^- : F^- \otimes X \to X$. If $F^- = 0$, we call F a *chiral perturbation.*

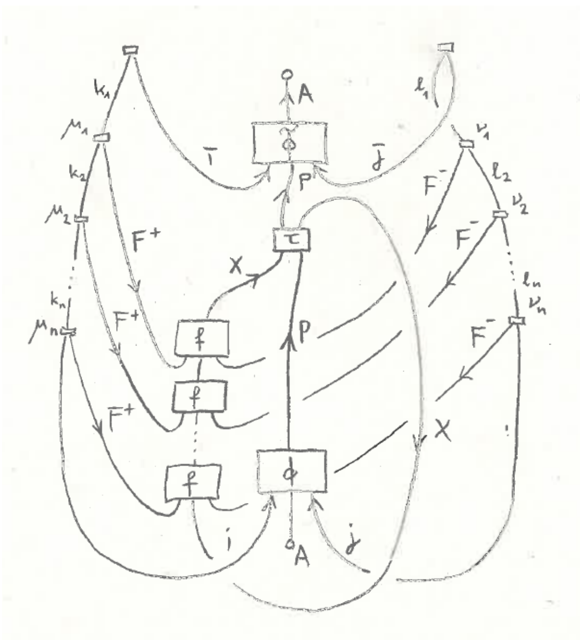
The integrand

Let us now return to (6.2.1). We will spend some time studying the integrand of (6.2.1). By the TFT-construction, we have

$$\begin{aligned}
&\big\langle\, u(Le^{i\theta})\, \psi(e^{i\theta_1}) \ldots \psi(e^{i\theta_n})\, v(0)) \big\rangle^{\tau_{X,P}}_{P,(X,f)} \\
&\quad = \sum_{i,j \in \mathcal{I}} \sum_{\alpha,\bar\alpha} \Omega^{P(X,f)}_{ij\alpha\bar\alpha} \left[\tilde{\phi}_{\bar{i}\bar{j}}, \phi_{ij} \right] \cdot F_{\bar{i}i\alpha}(Le^{i\theta}, e^{i\theta_1}, \ldots, e^{i\theta_n}, 0) F_{\bar{j}i\bar\alpha}(Le^{-i\theta}, e^{-i\theta_1}, \ldots, e^{-i\theta_n}, 0)
\end{aligned} \tag{6.2.3}$$

where $\Omega^{P(X,f)}_{ij\alpha\bar\alpha}$ is the morphism

$$\Omega^{P(X,f)}_{ij\alpha\bar\alpha} \;=$$



$$. \tag{6.2.4}$$

The multi-index α runs over all n-tuples of representations $k_1, \dots, k_n \in \mathcal{I}$ and basis elements $\mu_1, \dots, \mu_n$ in the corresponding Hom-spaces, while $\bar{\alpha}$ runs over representations $l_1, \dots, l_n \in \mathcal{I}$ and basis elements $\nu_1, \dots, \nu_n$. The functions $F_{\bar{i}i\alpha}$ and $F_{\bar{j}i\bar{\alpha}}$ are the corresponding conformal blocks.[2]

Note that the correlator (6.2.3) is only well-defined if $\theta, \theta_1, \dots, \theta_n$ are all distinct. If some of the angles coincide, we may set its value to zero. In this way, we obtain for each $L > 1$ and each θ a function

$$\begin{aligned} &[0, 2\pi]^n \longrightarrow \mathbb{C} \\ &(\theta_1, \dots, \theta_n) \mapsto \big\langle\, u(Le^{i\theta})\, \psi(e^{i\theta_1}) \dots \psi(e^{i\theta_n})\, v(0)\big\rangle_{P,(X,f)}^{\tau_{X,P}} \end{aligned} \tag{6.2.5}$$

which is piecewise smooth in all θ_k. From now on, we will assume the following condition to hold:

Condition 6.2.2. *Fix $\psi \in F$. Let $L \in \mathbb{R}_{>1}$ and $\theta \in [0, 2\pi)$. The function (6.2.5) is $L_1\left([0, 2\pi]^n\right)$-integrable for all choices of $u \in \mathcal{H}_{PA}$ and $v \in \mathcal{H}_{AP}$.*

Remark 6.2.3.

i) Condition 6.2.2 is expected to be satisfied if the sum of the chiral and anti-chiral scaling dimensions Δ_ψ and $\bar{\Delta}_\psi$ of ψ is less than 1, since the correlator typically exhibits a singularity of order $\sim |\theta_i - \theta_j|^{-\Delta_\psi - \bar{\Delta}_\psi}$ when $\theta_i \to \theta_j$.[3] More generally, for $\Delta_\psi + \bar{\Delta}_\psi < 2$ (corresponding to relevant defect perturbations), one would expect that the integral can be regularized in some way.

ii) The following statement is certainly *not* true: if Condition 6.2.2 holds for $\psi \in F$, then it holds for any $\psi' \in F$. In fact, most fields in F will have large enough scaling dimensions to spoil convergence. This makes Condition 6.2.2 particularly unappealing to deal with from a categorical point of view. On the level of abstraction of $\mathcal{D}_F$, we do not distinguish between different elements of F.

Definition of perturbed twisted defect operators

Let us introduce the shorthand notation

$$\begin{aligned} &\int \big\langle\, u(Le^{i\theta})\, \psi^{(n)}\, v(0)\big\rangle_{P,(X,f)}^{\tau_{X,P}} \\ &\qquad := \int_0^{2\pi} d\theta_1 \int_{\theta_1}^{2\pi} d\theta_2 \cdots \int_{\theta_{n-1}}^{2\pi} d\theta_n \;\big\langle\, u(Le^{i\theta})\, \psi(e^{i\theta_1}) \dots \psi(e^{i\theta_n})\, v(0)\big\rangle_{P,(X,f)}^{\tau_{X,P}} \end{aligned} \tag{6.2.6}$$

[2] We stress that $F_{\bar{i}i\alpha}$ and $F_{\bar{j}i\bar{\alpha}}$ are the conformal blocks with respect to the choice of local coordinates as previously defined for the corresponding correlator. Hence, they differ from the standard definition of genus 0 conformal blocks in the literature, where the canonical local coordinates of $\mathbb{C}$ are often assumed.

[3] For perturbations of the special form 6.2.2, the condition should even hold if $\Delta_\psi, \Delta_{\bar{\psi}} < 1$.

This correlator defines, for any value of L, a bilinear pairing $\mathcal{H}_{AP} \times \mathcal{H}_{PA} \to \mathbb{C}$. By rotation invariance, it actually does not depend on θ. In general, there is no reason for (6.2.6) to be compatible with the grading induced by $L_0 + \bar{L}_0$.

Let us also denote by

$$\int \big(u(\infty)\, \psi^{(n)}\, v(0) \big)_{P,(X,f)}^{\tau_{X,P}} : \mathcal{H}_{AP} \times \mathcal{H}_{PA} \to \mathbb{C} \tag{6.2.7}$$

the pairing obtained from (6.2.6) by taking $L \to \infty$ and changing the local coordinates at infinity to $z \mapsto 1/z$. By Remark 6.1.2, comparison of this pairing with the pairing $(-,-)_P$ yields for every $n \in \mathbb{Z}_{\geq 0}$ a linear map

$$D\left[(X,f)|P,\tau\right]^{(n)} \;:\; \mathcal{H}_P \longrightarrow \overline{\mathcal{H}}_P\ . \tag{6.2.8}$$

Finally, to reproduce the path ordered exponential (6.2.1), we would like to sum over the number of defect fields n. We define the *P-twisted perturbed defect operator of* $(X,f) \in \mathcal{D}_F$ by

$$\mathsf{D}[(X,f)|P,\tau] \;:=\; \sum_{n=0}^{\infty} D[(X,f)|P,\tau]^{(n)}\ . \tag{6.2.9}$$

Remark 6.2.4.

i) Note that formally, by linearity of correlators, we have

$$\mathsf{D}[(X,\lambda f)|P,\tau] \;=\; \sum_{n=0}^{\infty} \lambda^n D[(X,f)|P,\tau]^{(n)}\ . \tag{6.2.10}$$

The parameter $\lambda \in \mathbb{C}$ controls the strength of the perturbation, with

$$\lim_{\lambda\to 0} \mathsf{D}[(X,\lambda f)|P,\tau] = D[X|P,\tau] \tag{6.2.11}$$

reducing to the unperturbed defect operator.

ii) If our procedure of perturbing the defect describes any real physical situation, the sum (6.2.9) must of course have a finite radius of convergence in λ. If it converges for $\lambda \in \mathbb{C}$, we will say that the defect operator $\mathsf{D}[(X,\lambda f)|P,\tau]$ is convergent.[4] If this is not the case, one may instead consider (6.2.9) as a formal power series. This view was taken in [MaRu]. It may also be possible to rescale $f \mapsto \mu f$ with $\mu \in \mathbb{R}_{\geq 0}$ small enough to make the sum convergent.

[4] We think of convergence in the strong operator topology. This assumes that we are dealing with maps between Hilbert spaces. If the summands $D[(X,f)|P,\tau]^{(n)}$ are degree-preserving with respect to $L_0 + \bar{L}_0$, we are dealing with operators from $\mathcal{H}_P$ to itself. In the case of unitary RCFT, $\mathcal{H}_P$ can always be given a Hilbert space completion using the non-degenerate bilinear hermitean Shapovalov form. However, in the present thesis we want to leave this aspect of the story aside and concentrate on the algebraic properties of the defect operators.

Compatibility of the crossing τ with the perturbation F

Recall that we have $(\alpha\iota F, \varphi_{\alpha\iota F,-}) \in \mathcal{Z}(\mathcal{D})$. We say that the crossing $\tau : X \otimes P \to P \otimes X$ is compatible with the perturbation $f : \alpha\iota(F) \otimes X \to X$ if it satisfies the relation

$$\tau \circ (f \otimes \mathrm{id}_P) \;=\; (\mathrm{id}_P \otimes f) \circ (\varphi_{\alpha\iota F,P} \otimes \mathrm{id}_X) \circ (\mathrm{id}_{\alpha\iota F} \otimes \tau)\ . \tag{6.2.12}$$

The importance of this relation lies in the following fact: if it is satisfied, then the correlator (6.2.3) as a function of θ_k ($k = 1, \dots, n$) is smooth in the point $\theta_k = \theta$. In other words, the defect field $\psi \in F$ does not feel the twist defect P, but instead can be freely moved past it. In particular, the correlator is independent of θ up to a change of local coordinates:

Lemma 6.2.5. *Let $n \in \mathbb{Z}_{\geq 0}$. If the crossing is compatible with the perturbation (relation (6.2.12)), then we have, for all $\theta \in [0, 2\pi)$, that*

$$\big\langle\, u(Le^{i\theta})\, \psi(e^{i\theta_1}) \dots \psi(e^{i\theta_n})\, v(0) \big\rangle^{\tau_{X,P}}_{P,(X,f)} \;=\; \big\langle\, u'(L)\, \psi(e^{i\theta_1}) \dots \psi(e^{i\theta_n})\, v'(0) \big\rangle^{\tau_{X,P}}_{P,(X,f)}\ , \tag{6.2.13}$$

where $u' = e^{i\theta(L_0 - \bar{L}_0)} u$ and $v' = e^{i\theta(L_0 - \bar{L}_0)} v$.

Proof. This is clear from looking at (6.2.3): if the crossing is compatible with the perturbation, then the morphisms $\Omega^{P(X,f)}_{ij\alpha\bar{\alpha}}$ for the left hand side and the right-hand-side of (6.2.13) are equal. Meanwhile, the conformal blocks experience a rotation of the local coordinates at the endpoints of the P-defect, resulting in a chiral twist $e^{i\theta L_0}$ or an anti-chiral twist $e^{-i\theta \bar{L}_0}$ acting on the corresponding vectors, respectively. $\square$

By differentiating (6.2.13), we immediately get the following

Corollary 6.2.6. *Under the conditions of Lemma 6.2.5, we have*

$$\frac{\partial}{\partial\theta} \big\langle\, u(Le^{i\theta})\, \psi(e^{i\theta_1}) \dots \psi(e^{i\theta_n})\, v(0) \big\rangle^{\tau_{X,P}}_{P,(X,f)} \Big|_{\theta=0} \;= \tag{6.2.14}$$
$$\big\langle\, u''(L)\, \psi(e^{i\theta_1}) \dots \psi(e^{i\theta_n})\, v(0) \big\rangle^{\tau_{X,P}}_{P,(X,f)} \;+\; \big\langle\, u(L)\, \psi(e^{i\theta_1}) \dots \psi(e^{i\theta_n})\, v''(0) \big\rangle^{\tau_{X,P}}_{P,(X,f)}\ ,$$

where $u'' = i(L_0 - \bar{L}_0)u$ and $v'' = i(L_0 - \bar{L}_0)v$.

Remark 6.2.7.

i) If we take $L \to \infty$ and take the local coordinates at infinity to be $z \mapsto 1/z$, as we do when we compare correlators to the pairing $(-,-)_P$, then this results in a reversal of the respective tangent vector and the rotation in Lemma 6.2.5 is the other way around. Hence, in that case we need to replace $u'' \mapsto -u''$.

ii) It is clear that if Condition 6.2.2 holds, then the derivative (6.2.14) is also $L_1([0, 2\pi]^n)$-integrable (since we may assume u and v to be homogeneous, and so u'' and v'' are just proportional to u and v, respectively).

We also note, that rotation invariance of correlators demands

$$\big\langle\, u(Le^{i\theta})\,\psi(e^{i\theta_1})\dots\psi(e^{i\theta_n})\,v(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)} \;=\; \big\langle\, u(L)\,\psi(e^{i(\theta_1-\theta)})\dots\psi(e^{i(\theta_n-\theta)})\,v(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)} \tag{6.2.15}$$

for all $\theta \in [0, 2\pi)$.

If the crossing τ is compatible with the perturbation f, then (6.2.8) preserves the grading induced by $L_0 - \bar{L}_0$ (note the minus sign!), as the following Lemma shows:

Lemma 6.2.8. *Let $n \in \mathbb{Z}_{\geq 0}$. If the crossing is compatible with the perturbation (in the sense of relation (6.2.12)), then the pairing (6.2.7) is compatible with the grading induced by the action of $L_0 - \bar{L}_0$.*

Proof. By rotation invariance (6.2.15) and Corollary 6.2.6 we have that

$$\begin{aligned}
&\frac{\partial}{\partial\theta}\int_0^{2\pi} d\theta_1 \int_{\theta_1}^{2\pi} d\theta_2 \cdots \int_{\theta_{n-1}}^{2\pi} d\theta_n\; \big\langle\, u(L)\,\psi(e^{i(\theta_1-\theta)})\dots\psi(e^{i(\theta_n-\theta)})\,v(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)}\Big|_{\theta=0} \\
&\overset{(6.2.15)}{=} \frac{\partial}{\partial\theta}\int_0^{2\pi} d\theta_1 \int_{\theta_1}^{2\pi} d\theta_2 \cdots \int_{\theta_{n-1}}^{2\pi} d\theta_n\; \big\langle\, u(Le^{i\theta})\,\psi(e^{i\theta_1})\dots\psi(e^{i\theta_n})\,v(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)}\Big|_{\theta=0} \\
&\overset{\text{Remark 6.2.7.}ii)}{=} \int_0^{2\pi} d\theta_1 \int_{\theta_1}^{2\pi} d\theta_2 \cdots \int_{\theta_{n-1}}^{2\pi} d\theta_n\; \frac{\partial}{\partial\theta}\big\langle\, u(Le^{i\theta})\,\psi(e^{i\theta_1})\dots\psi(e^{i\theta_n})\,v(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)}\Big|_{\theta=0} \\
&\overset{(6.2.14)}{=} -\int \big\langle\, u''(L)\,\psi^{(n)}\,v(0)\big)\big\rangle_{P,(X,f)} \;+\; \int \big\langle\, u(L)\,\psi^{(n)}\,v''(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)}\;.
\end{aligned} \tag{6.2.16}$$

The left hand side clearly vanishes due to translation invariance of the integral. Taking $L \to \infty$ on the right hand side and performing the change of local coordinates at infinity (which kills the minus in front of the first integral due to Remark 6.2.7.i)), this thus results in

$$0 \;=\; \int \big(\, u''(\infty)\,\psi^{(n)}\,v(0)\big)\big)^{\tau_{X,P}}_{P,(X,f)} \;+\; \int \big(\, u(L)\,\psi^{(n)}\,v''(0)\big)\big)^{\tau_{X,P}}_{P,(X,f)}\;. \tag{6.2.17}$$

Hence, if u and v are homogeneous vectors with L_0- and $\bar{L}_0$- eigenvalues Δ_u, $\bar{\Delta}_u$, Δ_v, $\bar{\Delta}_v$, it follows that

$$0 \;=\; (\Delta_u - \bar{\Delta}_u - \Delta_v + \bar{\Delta}_v)\int \big(\, u(\infty)\,\psi^{(n)}\,v(0)\big)\big)^{\tau_{X,P}}_{P,(X,f)}\;. \tag{6.2.18}$$

If the integral is nonzero, we must have

$$\Delta_u - \bar{\Delta}_u \;=\; \Delta_v - \bar{\Delta}_v\;. \tag{6.2.19}$$

This proves the claim. □

Unfortunately, the grading induced by $L_0 - \bar{L}_0$ is somewhat less useful than the one induced by $L_0 + \bar{L}_0$, since the homogeneous subspaces it creates are not finite dimensional. However, in some cases we have more:

Corollary 6.2.9. *Assume the conditions of Lemma 6.2.8. If F is a chiral perturbation, then the pairing (6.2.7) is compatible with the grading induced by L_0 and $\bar{L}_0$ individually (and hence with the one induced by $L_0 + \bar{L}_0$).*

Proof. If ψ is a chiral field, then $\bar{\Delta}_u = \bar{\Delta}_v$ follows from the fact that the anti-chiral 2-point conformal block is compatible with the $\bar{L}_0$-grading. Combining this with (6.2.18), we conclude also $\Delta_u = \Delta_v$. □

If the pairing is compatible with the grading induced by both L_0 and $\bar{L}_0$, then the corresponding linear map $D\left[(X,f)|P,\tau\right]^{(n)}$ is an honest endomorphism of $\mathcal{H}_P$ which preserves L_0- and $\bar{L}_0$-eigenspaces. These are particularly interesting for us. Not only does it mean that such operators can be composed (which otherwise is tricky since their image may exceed their domain), but more importantly it means that

$$\left[D\left[(X,f)|P,\tau\right]^{(n)}, L_0 + \bar{L}_0\right] \;=\; 0\ . \tag{6.2.20}$$

Recalling that $L_0 + \bar{L}_0$ is the Hamiltonian, this is telling us that $D\left[(X,f)|P,\tau\right]^{(n)}$ is an integral of motion.[5]

To conclude this section, let us give a sufficient criterium for a crossing $\tau : X \otimes P \to P \otimes X$ to be compatible with the perturbation $f : \alpha\iota(F) \otimes X \to X$:

Lemma 6.2.10. *Let $\mathcal{P} \subseteq \mathcal{D}$ be a monoidal subcategory, such that $P \in \mathcal{P}$. If $((X, \tau_{X,-}), f) \in \mathcal{Z}(\mathcal{P};\mathcal{D})_F$, then $\tau_{X,P}$ is compatible with f.*

Proof. Recall that $(\alpha\iota F, \varphi_{\alpha\iota F,-}) \in \mathcal{Z}(\mathcal{D})$. By restricting the half-braiding $\varphi_{\alpha\iota F,Y}$ to objects $Y \in \mathcal{P}$, the pair $(\alpha\iota F, \varphi_{\alpha\iota F,-})$ becomes an object in $\mathcal{Z}(\mathcal{P};\mathcal{D})$. Recall that $\mathcal{Z}(\mathcal{P};\mathcal{D})$ is a monoidal category. The half-braiding of $(\alpha\iota F, \varphi_{\alpha\iota F,-}) \otimes (X, \tau_{X,-})$ is, by definition, given by $(\varphi_{\alpha\iota F,-} \otimes \mathrm{id}_X) \circ (\mathrm{id}_{\alpha\iota F} \otimes \tau_{X,-})$. Since f is a morphism in $\mathcal{Z}(\mathcal{P};\mathcal{D})$, it commutes with the half-braiding. But that is precisely the relation (6.2.12). □

Although this may not be the only way in which compatibility can be achieved, Lemma 6.2.10 establishes $\mathcal{Z}(\mathcal{P},\mathcal{D})_F$ as a natural setting in which to study F-perturbed defects twisted by a defect in $\mathcal{P}$.

Composition and superposition

In this section, we want to study the behaviour of perturbed defect operators under superposition and composition. Just as in the unperturbed case, we will find that both, if well-defined, are very naturally encoded in the structure of the category $\mathcal{D}_F$. These results generalise the conclusions of [MaRu] to the twisted case.

Let $(X,f), (Y,g) \in \mathcal{D}_F$ be perturbed defects whose associated operators (6.2.9) converge. Fix crossings $\tau_{X,P} : X \otimes P \to P \otimes X$ and $\tau_{Y,P} : Y \otimes P \to P \otimes Y$. Their sum is well-defined

[5] Due to its definition as a multiple integral, it is, in fact, a highly non-local integral of motion.

and is given by the convergent series

$$\begin{aligned}\mathsf{D}[(X,f)|P,\tau_{X,P}] \;+\; \mathsf{D}[(Y,g)|P,\tau_{Y,P}] \\ = \sum_{n=0}^{\infty}\big(D[(X,f)|P,\tau_{X,P}]^{(n)} + D[(X,f)|P,\tau_{Y,P}]^{(n)}\big)\ .\end{aligned} \tag{6.2.21}$$

Due to the splitting Lemma, any split exact sequence $0 \to X \xrightarrow{a} Z \xrightarrow{b} Y \to 0$ in an abelian category is isomorphic to the split exact sequence $0 \to X \xrightarrow{\iota_X} X \oplus Y \xrightarrow{\pi_Y} Y \to 0$. Denote by $\varphi : Z \to X \oplus Y$ the corresponding isomorphism. Recall the definition of $\tau_{X\oplus Y,P}$ from Lemma 6.1.6, and set

$$\tau_{Z,P} = (\mathrm{id}_P \otimes \varphi^{-1}) \circ \tau_{X\oplus Y,P} \circ (\varphi \otimes \mathrm{id}_P)\ . \tag{6.2.22}$$

We have the following sum decomposition result which generalises Lemma 6.1.6:

Proposition 6.2.11. *If $0 \to (X,f) \xrightarrow{a} (Z,h) \xrightarrow{b} (Y,g) \to 0$ is an exact sequence in $\mathcal{D}_F$ such that the underlying sequence in $\mathcal{D}$ splits, then we have*

$$\mathsf{D}[(X,f)|P,\tau_{X,P}] + \mathsf{D}[(Y,g)|P,\tau_{Y,P}] \;=\; \mathsf{D}[(Z,h)|P,\tau_{Z,P}]\ . \tag{6.2.23}$$

Proof. Without loss of generality, we set $Z = X \oplus Y$, $a = \iota_X$ and $b = \pi_Y$. We follow the proof of Lemma 6.1.6 in order to show that

$$\begin{aligned}\big\langle u(\infty)\,\psi(e^{i\theta_1})\ldots\psi(e^{i\theta_n})\,v(0)\big)\big\rangle^{\tau_{X,P}}_{P,(X,f)} \;+\; \big\langle u(\infty)\,\psi(e^{i\theta_1})\ldots\psi(e^{i\theta_n})\,v(0)\big)\big\rangle^{\tau_{Y,P}}_{P,(Y,g)} \\ =\; \big\langle u(\infty)\,\psi(e^{i\theta_1})\ldots\psi(e^{i\theta_n})\,v(0)\big)\big\rangle^{\tau_{X\oplus Y,P}}_{P,(Z,h)}\ .\end{aligned} \tag{6.2.24}$$

Compared to (6.1.31) we face the following additional obstruction: there are now perturbing morphisms $h : \alpha\iota(F) \otimes (X \oplus Y) \to X \oplus Y$ attached to the defect loop labelled by $X \oplus Y$, and one has to be able commute ι_X and π_Y past h when pulling them around the loop. But this is by definition possible, since they are morphisms in $\mathcal{D}_F$.[6] □

Unfortunately, composition requires a little more care: if perturbed defect operators are not degree-preserving with respect to $L_0 + \bar{L}_0$, then their domain and image do not match, and composition is ill-defined. Yet if at least (Y,g), together with the crossing $\tau_{Y,P}$, yields a degree-preserving convergent twisted defect operator (for example if F is chiral and the crossing is compatible with the perturbation), then precomposition with $\mathsf{D}[Y_n|P,\tau_{Y,P}]$ is well-defined and given by the convergent series

$$\mathsf{D}[X_m|P,\tau_{X,P}]\ \mathsf{D}[Y_n|P,\tau_{Y,P}] \;:=\; \sum_{k=0}^{\infty}\sum_{l=0}^{k} D[X_m|P,\tau_{X,P}]^{(l)} D[Y_n|P,\tau_{X,P}]^{(k-l)}\ . \tag{6.2.25}$$

[6] Note that, while in the unperturbed case there was a choice to be made whether to pull ι_X or π_X around the defect loop (and ι_Y or π_Y in the second summand), this choice is no longer possible here as π_X and ι_Y are not morphisms in $\mathcal{D}_F$.

We assume again that both defect operators are convergent.

Before we formulate a generalisation of Lemma 6.1.27 to the perturbed case, let us introduce some technicalities. Let $u \in \mathcal{H}_{PA}$ and $v \in \mathcal{H}_{AP}$. For $j, k \in \mathbb{Z}_{\geq 0}$, $\underline{\theta} \in [0, 2\pi)^k$, $\underline{\omega} \in [0, 2\pi)^j$ and $r \in (0, 1)$, denote by

$$c^{XY}_{k,j}(r, \underline{\theta}, \underline{\omega})[u, v] \;:=\; \big\langle\, u(\infty)\, \psi^f(e^{i\theta_1}) \dots \psi^f(e^{i\theta_k}) \psi^g(re^{i\omega_1}) \dots \psi^g(re^{i\omega_j})\, v(0) \big\rangle^{\tau_{X,P}, \tau_{Y,P}}_{P;(X,f),(Y,g)} \tag{6.2.26}$$

the correlator defined as follows: the defect X lies on the unit circle as usual, and the defect Y on a circle of radius r. The defect P runs from 0 to ∞, crossing X with $\tau_{X,P}$ and Y with $\tau_{Y,P}$. The fields $\psi \in F$ which sit on X act by the morphism f and the fields on Y act by g. The local coordinates are as usual. In particular, $c^{XA}_{n,0}(r, \underline{\theta}, \emptyset)$ is the usual defect correlator (6.2.3) with one defect X and n fields ψ acting on it by f.

We will assume not only that Condition 6.2.2 (L_1-integrability) holds for the functions

$$\begin{aligned} &[0, 2\pi]^{k+j} \longrightarrow \mathbb{C} \\ &(\theta_1, \dots, \theta_k, \omega_1, \dots, \omega_j) \mapsto c^{XY}_{k,j}(r, \underline{\theta}, \underline{\omega})[u, v]\ , \end{aligned} \tag{6.2.27}$$

but also that we can find an L_1-integrable function which majorises them for all r sufficiently close to 1.

Under these assumptions, we have the following result:

Proposition 6.2.12. *Let $(X, f), (Y, g) \in \mathcal{D}_F$, and let $\tau_{X,P} \in \mathrm{Hom}_{\mathcal{D}}(X \otimes P, P \otimes X)$, $\tau_{Y,P} \in \mathrm{Hom}_{\mathcal{D}}(Y \otimes P, P \otimes Y)$. If $\mathsf{D}[(Y, g)|P, \tau_{Y,P}]$ is degree-preserving with respect to $L_0 + \bar{L}_0$, then we have*

$$\mathsf{D}[(X, f)|P, \tau_{X,P}]\ \mathsf{D}[(Y, g)|P, \tau_{Y,P}] \;=\; \mathsf{D}[(X, f) \otimes_F (Y, g)|P, \tau_{X \otimes Y,P}]\ . \tag{6.2.28}$$

Proof. Let $u \in \mathcal{H}_{PA}$ and $v \in \mathcal{H}_{AP}$. Let $\mathcal{B}$ be a basis of $\mathcal{H}_P$. For every $\chi \in \mathcal{B}$ denote by $\hat{\chi} \in \mathcal{H}_{PA}$ the vector such that $(\hat{\chi}, \chi)_P = 1$. By the factorisation property (5.3.3) we have the following decomposition of (6.2.26) :

$$c^{XY}_{k,j}(r, \underline{\theta}, \underline{\omega})[u, v] \;=\; \sum_{\chi \in \mathcal{B}} c^{XA}_{k,0}(r, \underline{\theta}, \emptyset)[u, \chi] c^{AY}_{0,j}(r, \emptyset, \underline{\omega})[\hat{\chi}, v] \tag{6.2.29}$$

For every $m \in \mathbb{Z}_{\geq 0}$, define

$$A_m \;:=\; \sum_{l=0}^{m} D[(X, f)|P, \tau_{X,P}]^{(l)} D[(Y, g)|P, \tau_{Y,P}]^{(m-l)}\ , \tag{6.2.30}$$

$$B_m \;:=\; D[(X, f) \otimes_F (Y, g)|P, \tau_{X \otimes Y,P}]^{(m)}\ . \tag{6.2.31}$$

It suffices to prove that $A_m = B_m$ for all m. The proof proceeds in two steps: for all $u \in \mathcal{H}_{PA}$ and $v \in \mathcal{H}_{AP}$ we show

$$(u, A_m v)_P \;\overset{\text{Step 1}}{=}\; \lim_{r \nearrow 1} \sum_{l=0}^{m} \int_{\theta_1 < \dots < \theta_l} \int_{\omega_1 < \dots < \omega_{m-l}} c^{XY}_{l,m-l}(r, \underline{\theta}, \underline{\omega})[u, v]\ d\underline{\omega}\, d\underline{\theta} \;\overset{\text{Step 2}}{=}\; (u, B_m v)_P\ . \tag{6.2.32}$$

• *Step 1:* For $r \in \mathbb{R}_{>0}$, we define

$$D_r\left[(X,f)|P,\tau\right]^{(n)} : \mathcal{H}_P \longrightarrow \overline{\mathcal{H}}_P \tag{6.2.33}$$

in the same way as $D\left[(X,f)|P,\tau\right]^{(n)}$, but with the defect X placed on a circle of radius r instead of 1. Since correlators are smooth functions of the field insertion points, $D_r\left[(X,f)|P,\tau\right]^{(n)}$ is, in any case, continuous in r, and we can write

$$(u, A_m v)_P \;=\; \lim_{r\nearrow 1} \sum_{l=0}^{m} \big(u, D[(X,f)|P,\tau_{X,P}]^{(l)} D_r[(Y,g)|P,\tau_{Y,P}]^{(m-l)} v\big)_P \tag{6.2.34}$$

Let us compute the summands. We have[7]

$$\begin{aligned}
&\big(u, D[(X,f)|P,\tau_{X,P}]^{(l)} D_r[(Y,g)|P,\tau_{Y,P}]^{(m-l)} v\big)_P \\
&\quad= \int_{\theta_1<\cdots<\theta_l} c^{XA}_{l,0}(r,\underline{\theta},\emptyset)\,\big[u, D_r[(Y,g)|P,\tau_{Y,P}]^{(m-l)} v\big]\ d\underline{\theta} \\
&\quad= \int_{\theta_1<\cdots<\theta_l} \sum_{\chi\in\mathcal{B}} c^{XA}_{l,0}(r,\underline{\theta},\emptyset)[u,\chi]\ c^{AA}_{0,0}(r,\emptyset,\emptyset)\,\big[\hat{\chi}, D_r[(Y,g)|P,\tau_{Y,P}]^{(m-l)} v\big]\ d\underline{\theta} \\
&\quad\overset{(\star)}{=} \sum_{\chi\in\mathcal{B}} c^{AA}_{0,0}(r,\emptyset,\emptyset)\,\big[\hat{\chi}, D_r[(Y,g)|P,\tau_{Y,P}]^{(m-l)} v\big] \int_{\theta_1<\cdots<\theta_l} c^{XA}_{l,0}(r,\underline{\theta},\emptyset)[u,\chi]\ d\underline{\theta} \\
&\quad= \sum_{\chi\in\mathcal{B}} \left(\int_{\theta_1<\cdots<\theta_l} c^{XA}_{l,0}(r,\underline{\theta},\emptyset)\,[u,\chi]\ d\underline{\theta}\right)\cdot\left(\int_{\omega_1<\cdots<\omega_{m-l}} c^{AY}_{0,m-l}(r,\emptyset,\underline{\omega})\,[\hat{\chi},v]\ d\underline{\omega}\right) \\
&\quad= \sum_{\chi\in\mathcal{B}} \int_{\theta_1<\cdots<\theta_l}\int_{\omega_1<\cdots<\omega_{m-l}} c^{XA}_{l,0}(r,\underline{\theta},\emptyset)\,[u,\chi]\ c^{AY}_{0,m-l}(r,\emptyset,\underline{\omega})\,[\hat{\chi},v]\ d\underline{\omega}\,d\underline{\theta} \\
&\quad\overset{(\star)}{=} \int_{\theta_1<\cdots<\theta_l}\int_{\omega_1<\cdots<\omega_{m-l}} \sum_{\chi\in\mathcal{B}} c^{XA}_{l,0}(r,\underline{\theta},\emptyset)\,[u,\chi]\ c^{AY}_{0,m-l}(r,\emptyset,\underline{\omega})\,[\hat{\chi},v]\ d\underline{\omega}\,d\underline{\theta} \\
&\quad= \int_{\theta_1<\cdots<\theta_l}\int_{\omega_1<\cdots<\omega_{m-l}} c^{XY}_{l,m-l}(r,\underline{\theta},\underline{\omega})[u,v]\ d\underline{\omega}\,d\underline{\theta}
\end{aligned} \tag{6.2.35}$$

The limit $r \nearrow 1$ on the left hand side exists. Hence, so does the limit on the right hand side, and we have proved the first equality in (6.2.32).

• *Step 2:* This step is combinatorial. The key ingredients is that we can decompose the perturbing morphism on $X \otimes Y$ into a sum of morphisms on X and Y via $T(f,g) = T(f,0) + T(0,g)$. We spell out the argument as explained in [MaRu]:

$$(u, B_m v)_P \;=\; \int_{\alpha_1<\cdots<\alpha_m} \big\langle\, u(\infty)\,\psi^{T(f,g)}(e^{i\alpha_1})\dots\psi^{T(f,g)}(e^{i\alpha_m})\,v(0)\big\rangle^{\tau_{X\otimes Y,P}}_{P,(X\otimes Y,T(f,g))}\ d\underline{\alpha}$$

[7] Since $D_r[(Y,g)|P,\tau_{Y,P}]^{(m-l)}$ is degree-preserving, the sum over $\mathcal{B}$ has finitely many nonzero terms. Hence, in $(\star)$ the sum and integral can be swapped.

$$
\begin{aligned}
&= \int_{\alpha_1<\cdots<\alpha_m} \sum_{(h_1,\dots,h_m)\in\{f,g\}^{\times m}} \big\langle u(\infty)\,\psi^{h_1}(e^{i\alpha_1})\dots\psi^{h_m}(e^{i\alpha_m})\,v(0)\big\rangle_{P,X\otimes Y}^{\tau_{X\otimes Y,P}}\;d\underline{\alpha}\\
&= \frac{1}{m!}\int_{[0,2\pi]^m} \sum_{(h_1,\dots,h_m)\in\{f,g\}^{\times m}} \big\langle u(\infty)\,\psi^{h_1}(e^{i\alpha_1})\dots\psi^{h_m}(e^{i\alpha_m})\,v(0)\big\rangle_{P,X\otimes Y}^{\tau_{X\otimes Y,P}}\;d\underline{\alpha}\\
&= \frac{1}{m!}\sum_{l=0}^{m}\binom{m}{l}\int_{[0,2\pi]^m}\lim_{r\nearrow 1} c^{XY}_{l,m-l}(r,\underline{\theta},\underline{\omega})[u,v]\;d\underline{\omega}\,d\underline{\theta}\\
&= \sum_{l=0}^{m}\frac{1}{l!}\int_{[0,2\pi]^l}\frac{1}{(m-l)!}\int_{[0,2\pi]^{m-l}}\lim_{r\nearrow 1} c^{XY}_{l,m-l}(r,\underline{\theta},\underline{\omega})[u,v]\;d\underline{\omega}\,d\underline{\theta}\\
&= \sum_{l=0}^{m}\int_{\theta_1<\cdots<\theta_l}\int_{\omega_1<\cdots<\omega_{m-l}}\lim_{r\nearrow 1} c^{XY}_{l,m-l}(r,\underline{\theta},\underline{\omega})[u,v]\;d\underline{\omega}\,d\underline{\theta}
\end{aligned}
\tag{6.2.36}
$$

Under the assumptions made, there exists an $L_1([0,2\pi]^m)$-integrable function which majorises $c^{XY}_{l,m-l}(r,\underline{\theta},\underline{\omega})[u,v]$ for all r sufficiently close to 1. Thus, by Lebesgue majorised convergence, the limit and integral can be swapped. □

In the untwisted case $P = A$, Propositions 6.2.11 and 6.2.12 together with Corollary 6.2.9 can be elegantly synthesised as follows. Denote by $\mathrm{End}_{L_0,\bar{L}_0}(\mathcal{H}_P)$ the space of endomorphisms which preserve the degree of L_0 and $\bar{L}_0$.

Corollary 6.2.13. *If F is a chiral perturbation and $\mathsf{D}[(X,f)|A,\mathrm{id}_X]$ converges for all $(X,f)\in\mathcal{D}_F$, then the assignment $\mathsf{D}[-|A,\mathrm{id}_X] : \mathrm{Ob}(\mathcal{D}_F)\to\mathrm{End}_{L_0,\bar{L}_0}(\mathcal{H}_A)$ induces a ring homomorphism $K_0(\mathcal{D}_F)\to\mathrm{End}_{L_0,\bar{L}_0}(\mathcal{H}_A)$.*

Of course, the same statement is true for any other perturbation F which produces only degree-preserving operators.

A similar but more complicated statement could be made for the twisted case, where additionally one has to ensure compatibility of the crossings with the perturbing morphisms. This could be achieved for example by incorporating the result of Lemma 6.2.10.

7. Perturbed twisted defect operators in the compactified free boson CFT

In this chapter, we present a concrete example of a CFT, namely the compactified free boson, in which twisted perturbed defect operators as introduced in section 6 can be constructed. Using the theory developed there, one finds that a specific class of these perturbed defect operators (called *T-operators*) commute among themselves as well as with the Hamiltonian, thus making them integrals of motion in involution. Composition and superposition of T-operators is governed by the quantum group $U_\hbar(L\mathfrak{sl}_2)$. As a result, all the eigenvalues of the T-defects (or rather: some simple combinations thereof) satisfy Y-system functional relations (Y) of Dynkin type D_N, with real solutions in case of the untwisted ($P = \mathbf{1}$) theory (and real solutions up to a constant phase otherwise).

Making an ansatz for the asymptotics of the solutions (in the limit of large perturbation), we can then apply the results from chapter 1 to compute the eigenvalues of the T-defects (numerically, as the unique solution of a TBA integral equation) for the vacuum state at any strength of the perturbation. Moreover, we exploit the fact that the T-operators exist for a continuous family of twists P which interpolate between untwisted sectors in a non-trivial fashion. Continuous deformation in the twist allows us to access solutions of the Y-system with roots in the strip $\mathbb{S}_s$, corresponding to excited states.

This chapter is organised as follows: in section 7.1 we set the stage by introducing the free boson compactified at radius $r \in \mathbb{R}$ in terms of its chiral category $\mathcal{C}$ and the algebra $A_r \in \mathcal{C}$. We also introduce the simple topological defects (simple objects in ${}_{A_r}\mathcal{C}_{A_r}$) of the theory. In section 7.2, we introduce the T-defects, and we study a particular chiral perturbation of these defects. The perturbed T-defects are identified with representations of $U_\hbar(L\mathfrak{sl}_2)$, and this identification is used to derive their behaviour under composition. This gives rise, in section 7.3, to the twisted BLZ T-system functional relations for the eigenvalues of the perturbed T-operators. Also the corresponding Y-system and TBA equations are derived. In section 7.4 we discuss the numerical solution of these TBA equations for the specific case $N = 3$.

7.1. The chiral massless free boson

The chiral free boson CFT is defined by the Heisenberg vertex operator algebra $\mathfrak{H}$ (see e.g. [FBZ, Ch. 2.1]). This chiral algebra describes fields which, apart from the mandatory $\mathfrak{vir}$-symmetry, carry an action of the Heisenberg Lie algebra (the algebra of creation and annihilation operators of a quantum mechanical harmonic oscillator). Since the formalism used here only depends on the resulting representation category $\mathrm{Rep}(\mathfrak{H})$, we will allow ourselves to take a shortcut by describing this category without going through the rather cumbersome definition of $\mathfrak{H}$.

The Heisenberg algebra

Consider the one-dimensional abelian Lie algebra $\mathfrak{h} := \mathfrak{u}(1)$ over $\mathbb{C}$. As a vector space, we have $\mathfrak{h} = \mathbb{C}$. Hence, the objects of the category $\mathrm{Rep}(\mathfrak{h})$ are $\mathbb{C}$-vector spaces together with an arbitrary endomorphism. Denote by $\mathrm{Rep}^{\Delta}(\mathfrak{h}) \subset \mathrm{Rep}(\mathfrak{h})$ the full subcategory for which the endomorphism is diagonalisable. Simple objects in $\mathrm{Rep}^{\Delta}(\mathfrak{h})$ are necessarily one-dimensional, with the endomorphism given by a complex number.

The chiral algebra of interest in describing a single free boson is the affinisation of $\mathfrak{h}$, namely the infinite-dimensional Heisenberg algebra $\hat{\mathfrak{h}} := \hat{\mathfrak{u}}(1)$ over $\mathbb{C}$. This is the Lie algebra generated by $\{a_n | n \in \mathbb{Z}\} \cup K$ with relations

$$[a_m, a_n] = m\delta_{m+n,0}K \ , \qquad [a_m, K] = 0 \ . \tag{7.1.1}$$

Each module of $\mathfrak{H}$ is a $\hat{\mathfrak{h}}$-representation, but the converse is not true. In fact, one has $\mathrm{Rep}(\mathfrak{H}) \cong \mathrm{Rep}^{\Delta}_{\flat,1}(\hat{\mathfrak{h}})$ [Ru3], where $\mathrm{Rep}^{\Delta}_{\flat,1}(\hat{\mathfrak{h}}) \subset \mathrm{Rep}(\hat{\mathfrak{h}})$ is the full subcategory of $\hat{\mathfrak{h}}$-representations defined as follows: $R \in \mathrm{Rep}(\hat{\mathfrak{h}})$ is in $\mathrm{Rep}^{\Delta}_{\flat,1}(\hat{\mathfrak{h}})$, if

- $\flat$: R is bounded from below in the following sense: for any vector $v \in R$ there exists an $N_v > 0$, such that for all $L > 0$ we have $a_{m_1} \cdots a_{m_L} v = 0$ whenever $m_1 + \cdots + m_L > N_v$.
- 1: K acts as the identity on R.
- Δ: a_0 acts diagonalisably on R.

One can show that $\mathrm{Rep}^{\Delta}_{\flat,1}(\hat{\mathfrak{h}}) \cong \mathrm{Rep}^{\Delta}(\mathfrak{h})$ as $\mathbb{C}$-linear categories (see e.g. [Ru3]). The trick is to forget about the actions of all generators except the diagonal action of a_0, and only keep the subspace of a_0-heighest weight vectors. This defines the action of the single generator of $\mathfrak{h}$. Since $\mathfrak{h} = \mathbb{C}$ and the action is diagonalisable, we simply have $\mathrm{Rep}^{\Delta}(\mathfrak{h}) \cong \mathrm{Vect}^{\mathbb{C}}$, the category of $\mathbb{C}$-graded vector spaces. Hence, in this chapter we set, once and for all,

$$\mathcal{C} := \mathrm{Vect}^{\mathbb{C}} \ . \tag{7.1.2}$$

This is the category which, together with additional structure to be discussed, defines the free boson as a chiral CFT.

Denote by $V_\alpha \in \mathrm{Rep}^{\Delta}_{\flat,1}(\hat{\mathfrak{h}})$ a highest-weight-representation with heighest weight vector $v \in V_\alpha$, such that $a_0.v = \alpha v$. By construction of the Heisenberg VOA, V_α is also a highest weight representation of the Virasoro algebra with heighest-weight vector v, such that $L_0.v = \frac{1}{2}\alpha^2 v$ (the factor $\frac{1}{2}$ is just one possible convention).

The category $\mathcal{C}$

The category $\mathcal{C}$ is semisimple. The (isomorphism classes of) simple objects will be denoted by $\mathbb{C}_\alpha$ with $\alpha \in \mathbb{C}$, and the unit in $\mathbb{C}_\alpha$ will be denoted by 1_α. The object $\mathbb{C}_\alpha$ corresponds to the heighest-weight representation V_α. The category $\mathcal{C}$ is monoidal with tensor unit $\mathbf{1} = \mathbb{C}_0$, and the tensor product of simple objects satisfies

$$\mathbb{C}_\alpha \otimes \mathbb{C}_\beta \cong \mathbb{C}_{\alpha+\beta} \ . \tag{7.1.3}$$

By Schur's Lemma, we thus have $\dim\mathrm{Hom}(\mathbb{C}_\alpha \otimes \mathbb{C}_\beta, \mathbb{C}_\gamma) = \delta_{\alpha+\beta,\gamma}$. We define the morphism

$$\begin{aligned} {}_\alpha^\gamma\lambda_\beta : \mathbb{C}_\alpha \otimes \mathbb{C}_\beta &\longrightarrow \mathbb{C}_\gamma \\ 1_\alpha \otimes 1_\beta &\mapsto \delta_{\alpha+\beta,\gamma}\ 1_\gamma \ , \end{aligned} \tag{7.1.4}$$

which for $\alpha+\beta = \gamma$ gives a basis of the one-dimensional space $\mathrm{Hom}(\mathbb{C}_\alpha \otimes \mathbb{C}_\beta, \mathbb{C}_\gamma)$. Similarly, we define the morphism

$$\begin{aligned} {}^\beta\curlyvee_\alpha^\gamma : \mathbb{C}_\alpha &\longrightarrow \mathbb{C}_\beta \otimes \mathbb{C}_\gamma \\ 1_\alpha &\mapsto \delta_{\alpha,\beta+\gamma}\ 1_\beta \otimes 1_\gamma \end{aligned} \tag{7.1.5}$$

which gives a basis of $\mathrm{Hom}(\mathbb{C}_\alpha, \mathbb{C}_\beta \otimes \mathbb{C}_\gamma)$ if $\alpha = \beta + \gamma$.

The category $\mathcal{C}$ can be equipped with additional structure: first of all, there exists a one-parameter family $c^{(\kappa)}$ ($\kappa \in \mathbb{C}^\times$) of braidings on $\mathcal{C}$, given on simple objects by

$$\begin{aligned} c^{(\kappa)}_{\alpha,\beta} : \ \mathbb{C}_\alpha \otimes \mathbb{C}_\beta &\longrightarrow \mathbb{C}_\beta \otimes \mathbb{C}_\alpha \\ 1_\alpha \otimes 1_\beta &\mapsto e^{\kappa\alpha\beta} 1_\beta \otimes 1_\alpha \ . \end{aligned} \tag{7.1.6}$$

This is the braiding which comes from the braiding of $\mathrm{Rep}(\mathfrak{H})$ (which describes analytic continuation of conformal blocks in the marked points as discussed in section 5.2). The value of κ depends on the conventions chosen there. We will fix it later according to our convenience.

Moreover, all finite dimensional objects in $\mathcal{C}$ have duals. The dual of $\mathbb{C}_\alpha$ is given by $\mathbb{C}_{-\alpha}$. Evaluation and coevaluation on simple objects are given by

$$\mathrm{ev}_\alpha = {}_{-\alpha}^{\ 0}\lambda_\alpha \qquad \mathrm{coev}_\alpha = {}^\alpha\curlyvee_0^{-\alpha} \ . \tag{7.1.7}$$

Finally,

$$\Theta_\alpha = e^{\kappa\alpha^2}\mathrm{id}_{\mathbb{C}_\alpha} \quad , \qquad S_{\alpha\beta} = e^{2\kappa\alpha\beta} \tag{7.1.8}$$

define a ribbon twist and a modular S-matrix on simple objects. The full subcategory of finite dimensional vector spaces in $\mathrm{Vect}^{\mathbb{C}}$ thus becomes a ribbon category.

7.1.1. The compactified free boson CFT

Recall from section 5.4 that a full rational CFT can be built from a chiral RCFT by choosing a special symmetric Frobenius algebra inside the corresponding modular category. In case of the boson, the chiral algebra is not rational; the category $\mathcal{C}$ is semisimple, but has an infinite number of isomorphism classes of simple objects. Nevertheless, it is possible to construct a family of CFTs from it, which we will introduce now. On one hand, there is the CFT which describes a free uncompactified boson; it is is given by $A = \mathbf{1}$. On the other hand, the CFT associated to a free boson compactified on a cylinder of radius $r \in \mathbb{R}$ is described by the object

$$A_r = \bigoplus_{k\in\mathbb{Z}} \mathbb{C}_{kr} \tag{7.1.9}$$

in $\mathrm{Vect}^{\mathbb{C}}$. We will focus on this case. The object A_r can be turned into an algebra with multiplication $\mu : A_r \otimes A_r \to A_r$ defined by

$$\mu \ := \sum_{m,n\in\mathbb{Z}} \iota_{(m+n)r} \circ {}^{(m+n)r}_{\ \ mr}\lambda_{nr} \circ \left(\pi_{mr} \otimes \pi_{nr}\right) , \tag{7.1.10}$$

where $\pi_z : A_r \to \mathbb{C}_z$ denotes the canonical projection and $\iota_z : \mathbb{C}_z \to A_r$ the canonical injection. This algebra clearly cannot be Frobenius, since the underlying object A_r is not dualisable. This is the point in which our example deviates from the formalism developped for RCFTs. However, this deviation is relatively minor and we are confident that much of the formalism for RCFT works analogously here. A more detailed investigation why this is also justified in the non-rational case will be left to future work.

Topological defects

Recall that the topological defects of a full rational CFT are described by bimodules over the Frobenius algebra which defines the CFT. Assuming that this result remains true for the (in general non-rational) free boson CFT given by the pair $(\mathcal{C}, A_r)$, the category of topological defects is

$$\mathcal{D} := {}_{A_r}\mathcal{C}_{A_r} \ . \tag{7.1.11}$$

Before we study this category in more detail, let us fix some notations:

Notations 7.1.1. for any $\beta \in \mathbb{C}$ we denote by $[\beta] \in \mathbb{C}/r\mathbb{Z}$ the coset of $r\mathbb{Z} \subset \mathbb{C}$ with respect to β. Furthermore, we will denote by $\overline{\beta} \in [0, r) \times i\mathbb{R}$ the representative of $[\beta]$ with smallest positive real part.

Let $\beta \in \mathbb{C}$ and $\xi \in \mathbb{C}^\times$. We will define an object $X_{[\beta],\xi} \in \mathcal{D}$ as follows: as a $\mathbb{C}$-graded vector space, it is given by

$$X_{[\beta],\xi} \ = \bigoplus_{n\in\mathbb{Z}} \mathbb{C}_{nr+\beta} \ . \tag{7.1.12}$$

The A_r-actions $\rho^l : A_r \otimes X_{[\beta],\xi} \to X_{[\beta],\xi}$ and $\rho^r : X_{[\beta],\xi} \otimes A_r \to X_{[\beta],\xi}$ are defined by

$$\rho^l \;=\; \sum_{m,n\in\mathbb{Z}} {}^{(m+n)r+\beta}_{\quad mr} \curlywedge_{nr+\beta} \;, \qquad \rho^r \;=\; \sum_{m,n\in\mathbb{Z}} \xi^m \; {}^{(m+n)r+\beta}_{\quad nr+\beta} \curlywedge_{mr} \;. \tag{7.1.13}$$

Note that, despite its definition in terms of β, $X_{[\beta],\xi}$ indeed only depends on $[\beta]$. We prefer to make this very explicit in the notation, since some quantities below depend on β rather than $[\beta]$.

As the next lemma shows, the objects $X_{[\beta],\xi}$ are precisely the simple A_r-bimodules (see e.g. [BR, Rem. 6.1]):

Lemma 7.1.2.

1. *Let $X \in \mathcal{D}$ be a simple object. Then there exist $\beta \in \mathbb{C}$ and $\xi \in \mathbb{C}^\times$, such that $X \cong X_{[\beta],\xi}$.*

2. *We have $X_{[\beta],\xi} \cong X_{[\beta'],\xi'}$ if and only if $[\beta] = [\beta']$ and $\xi = \xi'$.*

Note that, as objects in $\mathcal{C}$, we have $X_{[\beta],\xi} \cong A_r \otimes \mathbb{C}_\beta$. Two special cases deserve attention: Firstly, A_r as a bimodule over itself is given by $X_{[0],1}$. As a defect, this corresponds to the invisible defect. More generally, a quick comparison of the actions (7.1.13) with the braiding (7.1.6) reveals that the α-induced bimodules (4.2.1) are simple:

$$\alpha^{\pm}(\mathbb{C}_\beta) \;\cong\; X_{[\beta],e^{\pm\kappa r\beta}} \;. \tag{7.1.14}$$

The category $\mathcal{D}$ is equipped with the tensor product $\otimes_{A_r}$ of A_r-bimodules. The tensor product between simple objects is additive in β and multiplicative in ξ [BR]:

$$X_{[\beta],\xi} \otimes_{A_r} X_{[\beta'],\xi'} \;\cong\; X_{[\beta+\beta'],\xi\xi'} \tag{7.1.15}$$

Remark 7.1.3. According to the relation (7.1.15) it is, depending on the context, more natural to parametrise the simple modules by $(e^{\frac{2\pi i}{r}\beta}, \xi) \in \mathbb{C}^\times \times \mathbb{C}^\times$. This is done for example in [FGRS].

It should be pointed out that knowing the simple objects of $\mathcal{D}$ is not quite enough to have full control of this category, since $\mathcal{D}$ is not semi-simple (see [BR] for examples of indecomposable reducible bimodules). For our purposes, however, it will suffice to restrict attention to the full semi-simple subcategory

$$\mathcal{D}_{\mathrm{ss}} := \langle X(\beta,\xi)\rangle_{\beta\in\mathbb{C},\xi\in\mathbb{C}^\times} \subset \mathcal{D} \tag{7.1.16}$$

whose objects are isomorphic to finite direct sums of simple objects in $\mathcal{D}$. Due to (7.1.15), $\mathcal{D}_{\mathrm{ss}}$ still closes under the tensor product.

Twisted field space

The situation we want to describe is the free boson on a twisted cylinder. Let us fix a simple object

$$P := X_{[\beta],\xi} \quad \in \mathcal{D} \tag{7.1.17}$$

to describe the twist defect. It it is convenient to fix an $x \in \mathbb{C}$ such that $\xi = e^{\kappa r x}$. According to (5.4.6), the space of twist fields is then given by

$$\mathcal{H}_P = \bigoplus_{p,q\in\mathbb{C}} (V_p \otimes_{\mathbb{C}} V_q)^{M_{pq}([\beta],\xi)} \tag{7.1.18}$$

with

$$\begin{aligned} M_{pq}([\beta],\xi) &= \dim\mathrm{Hom}_{\mathcal{D}}(\alpha^+(\mathbb{C}_p) \otimes \alpha^-(\mathbb{C}_q), P) \\ &= \dim\mathrm{Hom}_{\mathcal{D}}(X_{[p],e^{\kappa rp}} \otimes X_{[q],e^{-\kappa rq}}, X_{[\beta],\xi}) \\ &= \dim\mathrm{Hom}_{\mathcal{D}}(X_{[p+q],e^{\kappa r(p-q)}}, X_{[\beta],\xi}) \\ &= \sum_{\mu,\varpi\in\mathbb{Z}} \delta_{p+q,\beta+r\varpi}\ \delta_{p-q,x+\frac{2\pi i}{\kappa r}\mu} \ . \end{aligned} \tag{7.1.19}$$

Here, (7.1.14) and (7.1.15) were used. The points in the pq-plane which solve $M_{pq}(\beta,\xi) = 1$ define an affine rectangular lattice, which is called the charge lattice (with $\mu \in \mathbb{Z}$ the momentum quantum number and $\varpi \in \mathbb{Z}$ the winding number). Variation of the twist defect $X_{[\beta],\xi}$ with respect to the variables β and ξ generates translations of the charge lattice. Consider, say, the continuous family of twist defects

$$X_{[0],e^{2\pi i t}} \qquad , \ t \in [0,1] \tag{7.1.20}$$

of twist defects. For $t = 0$, this describes the invisible defect, and so $\mathcal{H}_{X_{[0],1}}$ is the state space of the untwisted theory. The point $(0,0)$ in the pq-plane is then an element of the charge lattice; it has quantum numbers $\mu = 0$ and $\varpi = 0$ and stands for the summand of (7.1.18) in which the ground state (vacuum) is contained. If t is increased, this lattice point shifts continuously. At $t = 1$, the twist defect is again invisible, so that we have returned to the initial untwisted theory. However, the point with quantum numbers $\mu = 0$ and $\varpi = 0$ is now located at $\frac{\pi i}{\kappa r}(1,-1)$, the point in the lattice which was initially described by the quantum numbers $\mu = 1$ and $\varpi = 0$. One could say that the twist defect helps us generate a discrete symmetry of the field space in a continuous fashion. We will exploit this idea in section 7.4.

Defect crossings and $\mathcal{Z}(\mathcal{D}_{\mathrm{ss}})$

In this section, we study crossings of semisimple defects. Recall that morphisms in

$$\mathrm{Hom}_{\mathcal{D}}(X_{[\beta],\xi} \otimes X_{[\alpha],\zeta}, X_{[\alpha],\zeta} \otimes X_{[\beta],\xi}) \tag{7.1.21}$$

correspond to topological crossings of the defects $X_{[\beta],\xi}$ and $X_{[\alpha],\zeta}$. To keep in touch with the theory developed in Chapter 6, we should ideally concentrate on those morphisms in (7.1.21) which extend to a half-braiding on $X_{[\beta],\xi}$ with respect to some monoidal subcategory of $\mathcal{D}$ containing P. In our example, the relevant subcategory will be $\mathcal{D}_{\mathrm{ss}}$. The aim of this section is to study how to lift a simple object $X_{[\beta],\xi}$ to $\mathcal{Z}(\mathcal{D}_{\mathrm{ss}})$.

Let $\alpha, \beta \in \mathbb{C}$ and $\xi, \zeta \in \mathbb{C}^\times$. We define $\chi_{\substack{\alpha,\zeta\\ \beta,\xi}} \in \mathrm{Hom}_{\mathcal{D}}(X_{[\beta],\xi} \otimes X_{[\alpha],\zeta}, X_{[\beta+\alpha],\xi\zeta})$ by the underlying morphism

$$\chi_{\substack{\alpha,\zeta\\ \beta,\xi}} := \sum_{m,n\in\mathbb{Z}} \xi^n \;{}^{(m+n)r+\beta+\alpha}_{\quad mr+\beta}\curlywedge_{nr+\alpha} \tag{7.1.22}$$

in $\mathcal{C}$. One easily checks that this is a morphism in $\mathcal{D}$. Moreover, one finds that, for any fixed $m \in \mathbb{Z}$,

$$\chi^{-1}_{\substack{\alpha,\zeta\\ \beta,\xi}} = \sum_{n\in\mathbb{Z}} \xi^{m-n} \;{}_{mr+\beta}\curlyvee^{(n-m)\alpha}_{nr+\alpha+\beta} \tag{7.1.23}$$

satisfies $\chi^{-1}_{\substack{\alpha,\zeta\\ \beta,\xi}} \circ \chi_{\substack{\alpha,\zeta\\ \beta,\xi}} = \mathrm{id}$, is a morphism in $\mathcal{D}$ as well, and as such is independent of the choice of m. A straight-forward calculation yields

Lemma 7.1.4. *Let $\beta, \alpha, \alpha' \in \mathbb{C}$ and $\xi, \zeta, \zeta' \in \mathbb{C}^\times$. The following identity of morphisms $X_{[\beta],\xi} \otimes X_{[\alpha],\zeta} \otimes X_{[\alpha'],\zeta'} \to X_{[\alpha+\alpha'],\zeta\zeta'} \otimes X_{[\beta],\xi}$ holds:*

$$\begin{aligned}\left(\chi_{\substack{\alpha',\zeta'\\ \alpha,\zeta}} \otimes \mathrm{id}_{[\beta],\xi}\right) \circ \left(\mathrm{id}_{[\alpha],\zeta} \otimes \chi^{-1}_{\substack{\beta,\xi\\ \alpha',\zeta'}} \circ \chi_{\substack{\alpha',\zeta'\\ \beta,\xi}}\right) \circ \left(\chi^{-1}_{\substack{\beta,\xi\\ \alpha,\zeta}} \circ \chi_{\substack{\alpha,\zeta\\ \beta,\xi}} \otimes \mathrm{id}_{[\alpha'],\zeta'}\right)&\\ = \chi^{-1}_{\substack{\beta,\xi\\ \alpha+\alpha',\zeta\zeta'}} \circ \chi_{\substack{\alpha+\alpha',\zeta\zeta'\\ \beta,\xi}} \circ \left(\mathrm{id}_{[\beta],\xi} \otimes \chi_{\substack{\alpha',\zeta'\\ \alpha,\zeta}}\right)&\end{aligned} \tag{7.1.24}$$

Notice that both morphisms $\chi_{\substack{\alpha,\zeta\\ \beta,\xi}}$ and $\chi^{-1}_{\substack{\alpha,\zeta\\ \beta,\xi}}$ depend on α, not just $[\alpha]$. A quick calculation shows:

Lemma 7.1.5. *We have $\chi_{\substack{\alpha+Mr,\zeta\\ \beta+Nr,\xi}} = \xi^{-M}\chi_{\substack{\alpha,\zeta\\ \beta,\xi}}$ and $\chi^{-1}_{\substack{\alpha+Mr,\zeta\\ \beta+Nr,\xi}} = \xi^{M}\chi^{-1}_{\substack{\alpha,\zeta\\ \beta,\xi}}$ for all $M, N \in \mathbb{Z}$.*

Notations 7.1.6. Since the morphisms only depend on $[\beta]$, it makes sense to emphasise this in the notation:

$$\chi_{\substack{\alpha,\zeta\\ [\beta],\xi}} = \chi_{\substack{\alpha,\zeta\\ \beta,\xi}} \qquad \chi^{-1}_{\substack{\alpha,\zeta\\ [\beta],\xi}} = \chi^{-1}_{\substack{\alpha,\zeta\\ \beta,\xi}} \tag{7.1.25}$$

We also set

$$\chi_{\substack{[\alpha],\zeta\\ [\beta],\xi}} := \chi_{\substack{\overline{\alpha},\zeta\\ [\beta],\xi}} \qquad \chi^{-1}_{\substack{[\alpha],\zeta\\ [\beta],\xi}} = \chi^{-1}_{\substack{\overline{\alpha},\zeta\\ [\beta],\xi}}\ , \tag{7.1.26}$$

but contrary to (7.1.25) this involves a choice of representative.

The isomorphisms (7.1.26) will be our preferred bases of the one-dimensional spaces $\mathrm{Hom}_{\mathcal{D}}(X_{[\beta],\xi} \otimes X_{[\alpha],\zeta}, X_{[\beta+\alpha],\xi\zeta})$ and $\mathrm{Hom}_{\mathcal{D}}(X_{[\beta+\alpha],\xi\zeta}, X_{[\alpha],\zeta} \otimes X_{[\beta],\xi})$, respectively. Any morphism

$$\Lambda_{[\beta],\xi;[\alpha],\zeta} \in \mathrm{Hom}_{\mathcal{D}}(X_{[\beta],\xi} \otimes X_{[\alpha],\zeta}, X_{[\alpha],\zeta} \otimes X_{[\beta],\xi}) \tag{7.1.27}$$

can be written in the form

$$\Lambda_{[\beta],\xi;[\alpha],\zeta} \;=\; \lambda_{[\beta],\xi;[\alpha],\zeta}\; \chi^{-1}_{\substack{[\beta],\xi\\[\alpha],\zeta}} \circ \chi_{\substack{[\alpha],\zeta\\[\beta],\xi}} \tag{7.1.28}$$

with $\lambda_{[\beta],\xi;[\alpha],\zeta} \in \mathbb{C}$ some constant.

Let us now fix β and ξ. Moreover, let $D \in \mathcal{D}_{\mathrm{ss}}$, and denote by $N_{[\alpha],\zeta}$ the multiplicity of $X_{[\alpha],\zeta}$ in D. We define

$$\Lambda_{[\beta],\xi;D} := \sum_{\substack{\alpha\in\mathbb{C}\\ \zeta\in\mathbb{C}^\times}} \sum_{k=1}^{N_{[\alpha],\zeta}} (\iota^k_{[\alpha],\zeta} \otimes \mathrm{id}_{[\beta],\xi}) \circ \Lambda_{[\beta],\xi;[\alpha],\zeta} \circ (\mathrm{id}_{[\beta],\xi} \otimes \pi^k_{[\alpha],\zeta})\ , \tag{7.1.29}$$

where $\pi^k_{[\alpha],\zeta} : D \to X_{[\alpha],\zeta}$ and $\iota^k_{[\alpha],\zeta} : X_{[\alpha],\zeta} \to D$ denote the canonical projection and injection onto and from the k-th copy of $X_{[\alpha],\zeta}$, respectively. Under which conditions does the family of morphisms $\Lambda_{[\beta],\xi;-}$ define a half-braiding on $X_{[\beta],\xi}$ relative to $\mathcal{D}_{\mathrm{ss}}$?

Proposition 7.1.7. *Let $\beta \in \mathbb{C}$ and $\xi \in \mathbb{C}^\times$. The following are equivalent:*

1. $\big(X_{[\beta],\xi}, \Lambda_{[\beta],\xi;-}\big) \in \mathcal{Z}(\mathcal{D}_{\mathrm{ss}})$.
2. *There exist group homomorphisms $\varphi_1, \varphi_2 : \mathbb{C}^\times \to \mathbb{C}^\times$, and a number $x \in \mathbb{C}$ satisfying $e^x = \xi$, such that*
$$\lambda_{[\beta],\xi;[\alpha],\zeta} \;=\; e^{\frac{\overline{\alpha}}{r}x}\; \varphi_1(e^{\frac{2\pi i}{r}\alpha})\, \varphi_2(\zeta) \tag{7.1.30}$$
for all $\alpha \in \mathbb{C}$ and all $\zeta \in \mathbb{C}^\times$.

Proof. Compatibility of the half-braiding with the tensor product requires

$$\begin{aligned} &\lambda_{[\beta],\xi;[\alpha],\zeta}\lambda_{[\beta],\xi;[\alpha'],\zeta'} \left(\chi_{\substack{[\alpha'],\zeta'\\[\alpha],\zeta}} \otimes \mathrm{id}_{[\beta],\xi}\right) \circ \left(\mathrm{id}_{[\alpha],\zeta} \otimes \chi^{-1}_{\substack{[\beta],\xi\\[\alpha'],\zeta'}} \circ \chi_{\substack{[\alpha'],\zeta'\\[\beta],\xi}}\right) \circ \left(\chi^{-1}_{\substack{[\beta],\xi\\[\alpha],\zeta}} \circ \chi_{\substack{[\alpha],\zeta\\[\beta],\xi}} \otimes \mathrm{id}_{[\alpha'],\zeta'}\right) \\ &\quad= \lambda_{[\beta],\xi;[\alpha+\alpha'],\zeta\zeta'}\; \chi^{-1}_{\substack{[\beta],\xi\\[\alpha+\alpha'],\zeta\zeta'}} \circ \chi_{\substack{[\alpha+\alpha'],\zeta\zeta'\\[\beta],\xi}} \circ \left(\mathrm{id}_{[\beta],\xi} \otimes \chi_{\substack{[\alpha'],\zeta'\\[\alpha],\zeta}}\right) \end{aligned} \tag{7.1.31}$$

Combining Lemma 7.1.5 with Lemma 7.1.4 yields

$$\begin{aligned} &\left(\chi_{\substack{\overline{\alpha'},\zeta'\\[\alpha],\zeta}} \otimes \mathrm{id}_{[\beta],\xi}\right) \circ \left(\mathrm{id}_{[\alpha],\zeta} \otimes \chi^{-1}_{\substack{\overline{\beta},\xi\\[\alpha'],\zeta'}} \circ \chi_{\substack{\overline{\alpha'},\zeta'\\[\beta],\xi}}\right) \circ \left(\chi^{-1}_{\substack{\overline{\beta},\xi\\[\alpha],\zeta}} \circ \chi_{\substack{\overline{\alpha},\zeta\\[\beta],\xi}} \otimes \mathrm{id}_{[\alpha'],\zeta'}\right) \\ &\quad= \chi^{-1}_{\substack{\overline{\beta},\xi\\[\alpha+\alpha'],\zeta\zeta'}} \circ \chi_{\substack{\overline{\alpha+\alpha'},\zeta\zeta'\\[\beta],\xi}} \circ \left(\mathrm{id}_{[\beta],\xi} \otimes \chi_{\substack{\overline{\alpha'},\zeta'\\[\alpha],\zeta}}\right) \\ &\quad= \xi^{\lfloor\frac{\overline{\alpha}+\overline{\alpha'}}{r}\rfloor}\; \chi^{-1}_{\substack{\overline{\beta},\xi\\[\alpha+\alpha'],\zeta\zeta'}} \circ \chi_{\substack{\overline{\alpha}+\overline{\alpha'},\zeta\zeta'\\[\beta],\xi}} \circ \left(\mathrm{id}_{[\beta],\xi} \otimes \chi_{\substack{\overline{\alpha'},\zeta'\\[\alpha],\zeta}}\right) \end{aligned} \tag{7.1.32}$$

Using (7.1.32), the half-braiding condition (7.1.31) reduces to an equation for the constants $\lambda_{[\beta],\xi;[\alpha],\zeta}$:

$$\lambda_{[\beta],\xi;[\alpha],\zeta}\lambda_{[\beta],\xi;[\alpha'],\zeta'} \;=\; \xi^{\lfloor\frac{\overline{\alpha}+\overline{\alpha'}}{r}\rfloor}\, \lambda_{[\beta],\xi;[\alpha+\alpha'],\zeta\zeta'} \tag{7.1.33}$$

Setting $\alpha' = 0$ and $\zeta = 1$, we achieve the separation of variables

$$\lambda_{[\beta],\xi;[\alpha],\zeta} = \lambda_{[\beta],\xi;[\alpha],1} \lambda_{[\beta],\xi;[0],\zeta} \ . \tag{7.1.34}$$

Set $\psi([\alpha]) := \lambda_{[\beta],\xi;[\alpha],1}$ and $\varphi_2(\zeta) := \lambda_{[\beta],\xi;[0],\zeta}$. From (7.1.33) it follows that

$$\psi([\alpha])\psi([\alpha']) = \xi^{\lfloor \frac{\overline{\alpha}+\overline{\alpha'}}{r} \rfloor} \psi([\alpha+\alpha']) \ , \tag{7.1.35}$$

$$\varphi_2(\zeta)\varphi_2(\zeta') = \varphi_2(\zeta\zeta') \ . \tag{7.1.36}$$

Let $x \in \mathbb{C}$, such that $e^x = \xi$. One solution to (7.1.35) is given by $\psi([\alpha]) = e^{\frac{\overline{\alpha}}{r}x} \varphi_1(e^{\frac{2\pi i}{r}\alpha})$, where

$$\varphi_1(e^{\frac{2\pi i}{r}\alpha})\varphi_1(e^{\frac{2\pi i}{r}\alpha'}) = \varphi_1(e^{\frac{2\pi i}{r}(\alpha+\alpha')}) \ . \tag{7.1.37}$$

By (7.1.36) and (7.1.37), $\varphi_1, \varphi_2 : \mathbb{C}^\times \to \mathbb{C}^\times$ must be group homomorphisms. To see that these are all solutions, assume that $\tilde{\psi}([\alpha])$ is another solution to (7.1.35). Set $\Psi([\alpha]) := \frac{\tilde{\psi}([\alpha])}{\psi([\alpha])}$. Using (7.1.35), we find that $\Psi([\alpha])\Psi([\alpha']) = \Psi([\alpha+\alpha'])$. Hence, $\tilde{\varphi}_1(e^{\frac{2\pi i}{r}\alpha}) = \varphi_1(e^{\frac{2\pi i}{r}\alpha})\Psi([\alpha])$ is a group homomorphism, and so $\tilde{\psi}([\alpha]) = e^{\frac{\overline{\alpha}}{r}x}\tilde{\varphi}_1(e^{\frac{2\pi i}{r}\alpha})$ does not yield a new solution. $\square$

Remark 7.1.8. There is, in fact, a monoidal equivalence $\mathcal{D}_{\mathrm{ss}} \cong \mathrm{Vect}^G_\omega$, where $G = \mathbb{C}^\times \times \mathbb{C}^\times$ is the group of simple A_r-bimodules defined by (7.1.15), and ω denotes a nontrivial 3-cocycle arising from the factor $\xi^{\frac{\overline{\alpha}}{r}}$ in (7.1.30), which defines a non-trivial associator in Vect^G. For more details, see [BR, App. A.6].

Remark 7.1.9. Let $U \in \mathcal{C}$. Recall that the α-induced bimodule $\alpha^+(U)$ comes naturally equipped with a half-braiding $\varphi_{\alpha(U)^+,-}$ (the chiral half of the half-braiding (5.4.3)), which is inherited from the braiding in $\mathcal{C}$ via the braided functor (5.4.4). Consider $U = \mathbb{C}_\beta$, which yields the simple induced bimodule (7.1.14). In this case, the induced half-braiding is, as a morphism in $\mathcal{C}$, given by

$$\begin{aligned}
\varphi_{\alpha^+(\mathbb{C}_\beta),X_{[\alpha],\zeta}} &\overset{(5.4.3)}{=} \sum_{m,n\in\mathbb{Z}} \Big(\Big({}^{\alpha+(m+n)r}\curlyvee^{0}_{\alpha+(m+n)r} \circ {}^{\alpha+(m+n)r}_{mr}\curlywedge_{\alpha+nr} \Big) \otimes \mathrm{id}_\beta \Big) \circ \Big(\mathrm{id}_{mr} \otimes c^{(\kappa)}_{\beta,\alpha+nr} \Big) \\
&\overset{(7.1.6)}{=} \sum_{m,n\in\mathbb{Z}} e^{\kappa\beta(\alpha+nr)} \Big({}^{\alpha+(m+n)r}\curlyvee^{\beta}_{\alpha+\beta+(m+n)r} \circ {}^{\alpha+\beta+(m+n)r}_{\beta+mr}\curlywedge_{\alpha+nr} \Big) \\
&= e^{\kappa\beta\alpha} \Big(\sum_{n\in\mathbb{Z}} {}^{\alpha+nr}\curlyvee^{\beta}_{\alpha+\beta+nr} \Big) \circ \Big(\sum_{m,n\in\mathbb{Z}} e^{\kappa\beta nr} \ {}^{\alpha+\beta+(m+n)r}_{\beta+mr}\curlywedge_{\alpha+nr} \Big) \ .
\end{aligned} \tag{7.1.38}$$

Comparing this with (7.1.22) and (7.1.23), we see that it is in fact equal to the half-braiding (7.1.30) with $x = \kappa\beta r$ and $\varphi_1 = \varphi_2 = 1$:

$$\varphi_{\alpha^+(\mathbb{C}_\beta),X_{[\alpha],\zeta}} = e^{\kappa\beta\overline{\alpha}} \ \chi^{-1}_{\substack{[\beta],\xi \\ [\alpha],\zeta}} \circ \chi_{\substack{[\alpha],\zeta \\ [\beta],\xi}} \tag{7.1.39}$$

7.2. T-operators

We come now to the main focus of this chapter, the study of a special class of defects in the free boson, the T-defects. The section is organised as follows. In section 7.2.1, we introdce the T-defects as a subcategory of topological defects. In section 7.2.2, we chirally perturb these defects using the formalism developed in chapter 6. Section 7.2.3 establishes the relationship of these T-operators with representations of the quantum group $U_\hbar(L\mathfrak{sl}_2)$, which is subsequently used to study the bahaviour of T-operators under composition and superposition.

Remember the caveat: The formalism in Chapter 6 only deals with RCFT. We will make the following global assumption for the rest of chapter 7:

Condition 7.2.1. *The results of chapter 6 apply also to the non-rational compactified free boson CFT.*

7.2.1. The T-defects

Fix $\omega \in \mathbb{C}^\times$ and $z \in \mathbb{C}$. Consider the objects

$$T_n := \bigoplus_{k=-\frac{n}{2}}^{\frac{n}{2}} X_{[k\omega],e^{kz}} \quad \in \mathcal{D}_{\mathrm{ss}} \tag{7.2.1}$$

for $n \in \mathbb{Z}_{\geq 0}$. Let us compute the tensor product $T_n \otimes T_m$ using the isomorphisms (7.1.15). We obtain

$$\begin{aligned} T_n \otimes T_m &= \left(\bigoplus_{k=-\frac{n}{2}}^{\frac{n}{2}} X_{[k\omega],e^{kz}} \right) \otimes \left(\bigoplus_{k'=-\frac{m}{2}}^{\frac{m}{2}} X_{[k'\omega],e^{k'z}} \right) \\ &\cong \bigoplus_{s=-\frac{n+m}{2}}^{\frac{n+m}{2}} \bigoplus_{\substack{k+k'=s \\ k\in\{-\frac{n}{2},\dots,\frac{n}{2}\} \\ k'\in\{-\frac{m}{2},\dots,\frac{m}{2}\}}} X_{[(k+k')\omega],e^{(k+k')z}} \\ &\cong \bigoplus_{s=-\frac{n+m}{2}}^{\frac{n+m}{2}} \left(X_{[s\omega],e^{sz}}\right)^{\oplus N_s} , \end{aligned} \tag{7.2.2}$$

where

$$N_s = \begin{cases} -|s| + \frac{m+n}{2} + 1 & \text{if } |s| > \frac{|n-m|}{2} \\ \min(m,n) + 1 & \text{if } |s| \leq \frac{|n-m|}{2} \end{cases} \tag{7.2.3}$$

After a reordering of the last sum in (7.2.2), one finds

$$T_n \otimes T_m \cong \bigoplus_{j=\frac{|n-m|}{2}}^{\frac{n+m}{2}} T_{2j} \tag{7.2.4}$$

But (7.2.4) are precisely the fusion relations for simple finite-dimensional representations of $\mathfrak{sl}_2$, with T_n corresponding to the $n+1$-dimensional representation. Hence, the monoidal subcategory $\mathcal{T} \subset \mathcal{D}$ generated by the objects T_n and $\otimes_{\mathcal{D}}$ is semi-simple and there is a monoidal functor

$$T : \mathrm{Rep}(U(\mathfrak{sl}_2)) \longrightarrow \mathcal{T} \subset \mathcal{D}_{\mathrm{ss}} \ . \tag{7.2.5}$$

We have thus identified a subcategory of semi-simple topological defects in the free boson which behave like $\mathfrak{sl}_2$-representations. They will be the main protagonists of this chapter.

Remark 7.2.2. If $A = \mathbf{1}$ (the case which defines the uncompactified free boson), then $X_{[k\omega],e^{kz}}$ and $X_{[k'\omega],e^{k'z}}$ are non-isomorphic for $k \neq k'$, which makes the functor T injective. However, for the compactified free boson with $A = A_r$, the following may happen. Suppose that $\omega \in r\mathbb{Q}$, and $z \in i\pi\mathbb{Q}$, then there exists a minimal $k_0 \in \mathbb{Z}_{>0}$, such that $\omega k_0 \in r\mathbb{Z}$ and $k_0 z \in 2i\pi\mathbb{Z}$. In this case, we have $X_{[(k+k_0)\omega],e^{(k+k_0)z}} \cong X_{[k\omega],e^{kz}}$ for all k. By exploiting this isomorphism, one obtains for example always the relation $T_{2k_0+1+2j} \cong T_{2k_0-1} \oplus T_{1+2j}$ for $j \geq 0$, which spoils injectivity of the functor (7.2.5).

From now on we set

$$z = \kappa r \omega \ . \tag{7.2.6}$$

In this case, the T-defects are in the image of the functor α (see (7.1.14)), which has the nice consequence that the α-induced half-braiding (see (7.1.39)) yields a half-braiding $\varphi_{T_n,-}$ on T_n, defined by

$$\varphi_{T_n,X} \;=\; \sum_{k=-\frac{n}{2}}^{\frac{n}{2}} (\mathrm{id}_X \otimes \iota_k) \circ \varphi_{\alpha^+(\mathbb{C}_{k\omega}),X} \circ (\pi_k \otimes \mathrm{id}_X) \ , \tag{7.2.7}$$

where $X \in \mathcal{D}$ and $\iota_k : X_{[k\omega],e^{\kappa r k\omega}} \to T_n$ and $\pi_k : T_n \to X_{[k\omega],e^{\kappa r k\omega}}$ denote the canonical injections and projections. The morphisms $\varphi_{T_n,P}$ can be used in order to turn the T-defects into defect operators

$$D[T_n|P, \varphi_{T_n,P}] : \mathcal{H}_P \to \mathcal{H}_P \tag{7.2.8}$$

on the P-twisted cylinder, as described in chapter 6. Since the T-defects are topological and $T_n \otimes T_m \cong T_m \otimes T_n$ holds, the defect operators $D[T_n|P, \varphi_{T_n,P}]$ describe integrals of motion in involution: for all $m, n \in \mathbb{Z}_{\geq 0}$, we have

$$\big[D[T_n|P, \varphi_{T_n,P}], L_0 + \bar{L}_0\big] = 0 \quad , \qquad \big[D[T_n|P, \varphi_{T_n,P}], D[T_m|P, \varphi_{T_m,P}]\big] = 0 \ . \tag{7.2.9}$$

However, on any given subspace $V_p \otimes V_q \subset \mathcal{H}_P$ they act proportional to the identity, and thus their eigenspaces are vastly degenerate. In the next section we will perturb the T-defects, so as to break this degeneracy while preserving them as integrals of motion in involution.

7.2.2. Perturbed T-defects

We will now, once and for all, fix a chiral defect perturbation $F = F^+ \otimes_{\mathbb{C}} \mathbf{1} \in \mathcal{C} \boxtimes \overline{\mathcal{C}}$ by setting

$$F^+ \;:=\; \mathbb{C}_\omega \oplus \mathbb{C}_{-\omega} \;. \tag{7.2.10}$$

As an object in $\mathcal{Z}(\mathcal{D})$, we have

$$\begin{aligned} \alpha\iota F \;&=\; \big(\alpha^+(F^+), \varphi_{F^+,-}\big) \\ &=\; \big(X_{[\omega],e^{\kappa r\omega}}, \varphi_{\alpha^+(\mathbb{C}_\omega),-}\big) \oplus \big(X_{[-\omega],e^{-\kappa r\omega}}, \varphi_{\alpha^+(\mathbb{C}_{-\omega}),-}\big) \end{aligned} \tag{7.2.11}$$

The perturbing morphism $f_n : \alpha\iota F \otimes T_n \to T_n$ in $\mathcal{D}$ is chosen to be

$$f_n\Big|_{X_{[\pm\omega],e^{\pm\kappa r\omega}} \otimes X_{[k\omega],e^{\kappa rk\omega}}} \;:=\; \begin{cases} \chi_{\substack{[k\omega],e^{\kappa rk\omega} \\ [\pm\omega],e^{\pm\kappa r\omega}}} & \text{if } |k \pm 1| \leq \frac{n}{2} \\ 0 & \text{else} \end{cases} \;. \tag{7.2.12}$$

Our aim is to apply this perturbation to the T-defects and study the fusion relations of the resulting perturbed defect operators (as defined in Chapter 6). We denote these, for perturbation strength $\lambda \in \mathbb{C}$, by the shorthand notation

$$\mathsf{T}_n^P(\lambda) \;:=\; \mathsf{D}[(T_n, \lambda f_n)|P; \varphi_{T_n,P}] \;, \tag{7.2.13}$$

and we refer to them as T-operators. We will *assume* that they exist, i.e. that Condition 6.2.2 holds, and that the sum (6.2.9) converges for all λ. This will be corroborated a posteriori by the numerical results in section 7.4.

First, we note the following fact, which follows from Lemma 6.2.10 and Condition 7.2.1:

Lemma 7.2.3. *For all $n \in \mathbb{Z}_{\geq 0}$ and all $P \in \mathcal{D}_{\text{ss}}$, the crossing $\varphi_{T_n,P}$ is compatible with the perturbation f_n (in the sense of (6.2.12)).*

Thanks to this, the conditions of Lemma 6.2.8 and Corollary 6.2.9 are satisfied, which suggests that the perturbed defect operators $\mathsf{T}_n^P(\lambda)$ preserve the grading: for all $n \in \mathbb{Z}_{\geq 0}$ and $\lambda \in \mathbb{C}$, we have

$$\big[\mathsf{T}_n^P(\lambda), L_0\big] = 0 \quad , \qquad \big[\mathsf{T}_n^P(\lambda), \bar{L}_0\big] = 0 \;. \tag{7.2.14}$$

Hence, $\mathsf{T}_n^P(\lambda)$ are integrals of motion.

The next few sections are dedicated to studying the composition and superposition of T-operators. As was outlined in chapter 6, these are governed by the Grothendieck ring $K_0(\mathcal{T}_F)$. Instead of attempting a direct computation of $\mathcal{T}_F$, we will lift (7.2.5) to an exact tensor functor

$$T_{\text{pert}} : \mathrm{Rep}_{H_0-\text{ss}}(U_\hbar(L\mathfrak{sl}_2)) \longrightarrow \mathcal{T}_F \subset \mathcal{D}_F \;, \tag{7.2.15}$$

where $U_\hbar(L\mathfrak{sl}_2))$ is the $\hbar$-deformed loop algebra of $\mathfrak{sl}_2$, and $\mathrm{Rep}_{H_0-\text{ss}}(U_\hbar(L\mathfrak{sl}_2))$ is the category of its finite-dimensional representations where the generator H_0 acts semi-simple (see next section for precise definitions). In this way, we translate the problem of composition and superposition of T-operators into a problem about representations of this quantum group.

7.2.3. The Hopf algebra $U_\hbar(L\mathfrak{sl}_2)$

Let $q \in \mathbb{C}^\times \setminus \{-1, 1\}$. Most standard versions of quantum $\mathfrak{sl}_2$ involve generators E, F, K and K^{-1} satisfying relations (see e.g. [Kas, Def.VI.1.1.]).

$$\begin{gathered} KK^{-1} = K^{-1}K = 1 \ , \\ KEK^{-1} = q^2 E \ , \quad KFK^{-1} = q^{-2}F \ , \qquad [E,F] = \frac{K - K^{-1}}{q - q^{-1}} \ , \end{gathered} \tag{7.2.16}$$

and possibly some additional relations. However, the version of quantum $\mathfrak{sl}_2$ which will be relevant here is one where K and K^{-1} are replaced by a generator H, thought to be related via $K = q^H$.

Let $\hbar \in \mathbb{C}$, such that $q = e^\hbar$. The price to be paid for working with H instead of K is that one has to work a bit in order to make sense of the exponential q^H as an element in the algebra. This involves passing to an appropriate topological completion. One way to do this is to work over the field of formal power series in $\hbar$ (see e.g. [Dr] and [CP1, Def. 6.4.3.]). However, for us it will be important to assign a numerical value to q. In this case one can consider the free algebra $\mathbb{C}[H_1, \dots, H_n; X_1, \dots, X_m\rangle$ of polynomials in n mutually commuting and m non-commuting variables, and define a topological completion $\mathcal{E}(H_1, \dots, H_n; X_1, \dots, X_m)$. This construction is presented in detail in [BR, App. A.3], and it has the property that $e^{\hbar H_i} \in \mathcal{E}(H_1, \dots, H_n; X_1, \dots, X_m)$.

Definition 7.2.4. $U_\hbar(\mathfrak{sl}_2)$ is the topological Hopf algebra obtained as the quotient of $\mathcal{E}(H; X^+, X^-)$ by the closure of the two-sided ideal generated by the relations

$$[H, X^\pm] = \pm X^\pm \ , \qquad [X^+, X^-] = \frac{e^{hH} - e^{-hH}}{e^h - e^{-h}} \ , \tag{7.2.17}$$

with coproduct, antipode and counit given by

$$\begin{gathered} \Delta(H) = H \otimes 1 + 1 \otimes H \ , \\ \Delta(X^+) = X^+ \otimes e^{hH} + 1 \otimes X^+ \ , \qquad \Delta(X^-) = X^- \otimes 1 + e^{-hH} \otimes X^- \ , \\ S(H) = -H \ , \quad S(X^+) = -X^+ e^{-hH} \ , \quad S(X^-) = -e^{hH} X^- \ , \\ \varepsilon(H) = \varepsilon(X^\pm) = 0 \ . \end{gathered} \tag{7.2.18}$$

Moreover, we denote by $U_\hbar(\mathfrak{sl}_2)^\pm$ the positive/negative (topological) Borel subalgebra generated by $\langle H, X^+\rangle$ and $\langle H, X^-\rangle$, respectively.

The corresponding loop quantum group is defined as follows:

Definition 7.2.5. $U_\hbar(L\mathfrak{sl}_2)$ is the topological Hopf algebra obtained as the quotient of $\mathcal{E}(H_0, H_1; X_0^+, X_1^+, X_0^-, X_1^-)$ by the closure of the two-sided ideal generated by the relations $(i, j = 0, 1)$

$$\begin{gathered} H_0 = -H_1 \ , \qquad [H_i, X_i^\pm] = \pm 2X_i^\pm \ , \qquad [X_i^+, X_j^-] = \delta_{ij} \frac{e^{\hbar H_i} - e^{-\hbar H_i}}{e^\hbar - e^{-\hbar}} \ , \\ (X_i^\pm)^3 X_j^\pm - [3]_q (X_i^\pm)^2 X_j^\pm X_i^\pm + [3]_q X_i^\pm X_j^\pm (X_i^\pm)^2 - X_j^\pm (X_i^\pm)^3 \ = \ 0 \end{gathered} \tag{7.2.19}$$

with coproduct, antipode and counit given by

$$\begin{gathered}\Delta(H_i) = H_i \otimes 1 + 1 \otimes H_i \ , \\ \Delta(X_i^+) = X_i^+ \otimes e^{\hbar H_i} + 1 \otimes X_i^+ \ , \qquad \Delta(X_i^-) = X_i^- \otimes 1 + e^{-\hbar H_i} \otimes X_i^- \ , \\ S(H_i) = -H_i \ , \quad S(X_i^+) = -X_i^+ e^{-\hbar H_i} \ , \quad S(X_i^-) = -e^{\hbar H_i} X_i^- \ , \\ \varepsilon(H_i) = \varepsilon(X_i^\pm) = 0 \ . \end{gathered} \tag{7.2.20}$$

Moreover, we denote by $U_\hbar(L\mathfrak{sl}_2)^\pm$ the positive/negative (topological) Borel subalgebra generated by $\langle H_0, H_1, X_0^+, X_1^+\rangle$ and $\langle H_0, H_1, X_0^-, X_1^-\rangle$, respectively.

For $n \in \mathbb{Z}_{\geq 1}$, we denote by $V_n := (\mathbb{C}^{n+1}, \rho_n)$ the $U_\hbar(\mathfrak{sl}_2)$-representation defined by

$$\rho_n(H) \ := \begin{pmatrix} n & & & & \\ & n-2 & & & \\ & & \ddots & & \\ & & & -n+2 & \\ & & & & -n \end{pmatrix} , \tag{7.2.21}$$

$$\rho_n(X^+) \ := \begin{pmatrix} 0 & [n]_q & & & \\ & 0 & [n-1]_q & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} , \qquad \rho_n(X^-) \ := \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & & \ddots & & \\ & & [n-1]_q & 0 & \\ & & & [n]_q & 0 \end{pmatrix} .$$

Moreover, for all $w \in \mathbb{C}$ which satisfy $e^{\hbar w} = e^{-\hbar w}$ we have the one-dimensional representation $V_0^{(w)} := (\mathbb{C}, \rho_0^{(w)})$ defined by

$$\rho_0^{(w)}(H) := w \ , \qquad \rho_0^{(w)}(X^+) := 0 \ , \qquad \rho_0^{(w)}(X^-) := 0 \ . \tag{7.2.22}$$

These representations give also rise to families of $U_\hbar(L\mathfrak{sl}_2)$-representations as follows. For every $z \in \mathbb{C}^\times$, there is a homomorphism $\mathrm{ev}_z : U_\hbar(L\mathfrak{sl}_2) \to U_\hbar(\mathfrak{sl}_2)$, given by

$$\begin{aligned} \mathrm{ev_z}(H_1) &= H \qquad & \mathrm{ev_z}(X_1^+) &= X^+ \qquad & \mathrm{ev_z}(X_1^-) &= X^- \\ \mathrm{ev_z}(H_0) &= -H \qquad & \mathrm{ev_z}(X_0^+) &= q^{-1} z X^- \qquad & \mathrm{ev_z}(X_0^-) &= q z^{-1} X^+ \ . \end{aligned} \tag{7.2.23}$$

Pulling the $U_\hbar(\mathfrak{sl}_2)$-representation V_n back along ev_z (i.e. $g \in U_\hbar(L\mathfrak{sl}_2)$ acts on $\mathbb{C}^{n+1}$ by the endomorphism $\rho_n(\mathrm{ev}_z(g))$) gives rise to a representation of $U_\hbar(L\mathfrak{sl}_2)$, which is denoted by $V_n(z)$. Note that the one-dimensional representations $V_0^{(w)}(z)$ are independent of z.

Denote by $\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2))$ the category of finite-dimensional representations. In [BR, Thm. 5.5] it was shown that the Grothendieck ring $K_0(\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2)))$ is commutative if q is not a root of unity. Thus, one has $V_n(z) \otimes V_m(w) \cong V_m(w) \otimes V_n(z)$ for all $n, m \in \mathbb{Z}_{\geq 0}$ and $z, w \in \mathbb{C}^\times$. It is also known in this case [CP2, BR] that there exists a short exact sequence

$$0 \longrightarrow V_{n-1}(q^{n+1} z) \longrightarrow V_1(q^{-1} z) \otimes V_n(q^n z) \longrightarrow V_{n+1}(q^{n-1} z) \longrightarrow 0 \tag{7.2.24}$$

for all $n \in \mathbb{Z}_{\geq 1}$ and all $z \in \mathbb{C}^\times$. The underlying sequence in $\mathrm{Rep}(U_\hbar(\mathfrak{sl}_2))$ is of course just the generalisation of the well-known Clebsch-Gordan decomposition of tensor products in $\mathrm{Rep}(\mathfrak{sl}_2)$ (see e.g. [Kas, Ch. VII.7]).

Now let us discuss the case where q is a root of unity. First of all, one observes that tensor products of evaluation representations still commute in this case. Moreover, there is a new short exact sequence:

Proposition 7.2.6. *Let $q \notin \{-1,1\}$ be a root of unity, and N the smallest positive integer such that $q^N \in \{-1,1\}$. For every $z \in \mathbb{C}^\times$ we have a short exact sequence*

$$0 \longrightarrow V_{N-2}(z) \longrightarrow V_N(z) \longrightarrow V_0^{(+N)} \oplus V_0^{(-N)} \longrightarrow 0 \tag{7.2.25}$$

in $\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2))$.

Proof. Let $n \in \mathbb{Z}_{\geq 2}$. Denote by $\{\mathbf{e}_i^{(n)}\}_{i=1,\dots,n+1}$ the canonical basis of V_n (in terms of which the endomorphisms (7.2.21) are defined), and by ι_n the canonical embedding $\mathbf{e}_i^{(n-2)} \mapsto \mathbf{e}_{i+1}^{(n)}$ of V_{n-2} into V_n. Generically, the image of ι_n is not a submodule of V_n. Recall, however, that we have $[\pm N]_q = 0$. As a consequence, $\iota_N(V_{N-2}) \subset V_N$ is a submodule, resulting in the short exact sequence

$$0 \longrightarrow V_{N-2} \xrightarrow{\iota_N} V_N \xrightarrow{\pi_N} V_0^{(+N)} \oplus V_0^{(-N)} \longrightarrow 0 \tag{7.2.26}$$

with π_N the projection onto the subspace generated by $\{\mathbf{e}_1, \mathbf{e}_{N+1}\}$. This is a sequence in the category $\mathrm{Rep}(U_\hbar(\mathfrak{sl}_2))$. It can be lifted to a sequence in $\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2))$,

$$0 \longrightarrow V_{N-2}(z) \xrightarrow{\iota_N} V_N(w) \xrightarrow{\pi_N} V_0^{(+N)} \oplus V_0^{(-N)} \longrightarrow 0 \ , \tag{7.2.27}$$

for some $z, w \in \mathbb{C}^\times$, provided that ι_N and π_N intertwine the action of $U_\hbar(L\mathfrak{sl}_2)$. For π_N this is clearly the case independently of the choice of w. Moreover, ι_N intertwines the action of generators H_0, H_1 and $X_1^\pm$ for all z and w, since these generators act like the generators of $U_\hbar(\mathfrak{sl}_2)$ (compare (7.2.23)) for which ι_N is an intertwiner already. Hence, ι_N is an $U_\hbar(L\mathfrak{sl}_2)$-intertwiner if and only if it intertwines the action of the two remaining generators $X_0^\pm$, namely

$$\iota_N(X_0^\pm.v) \stackrel{!}{=} X_0^\pm.\iota_N(v) \qquad \forall v \in V_{N-2} \ . \tag{7.2.28}$$

Using (7.2.23), linearity of ι_N and again the fact that ι_N is a $U_\hbar(\mathfrak{sl}_2)$-intertwiner, we see that the left hand side equals $q^{\mp1}z^{\pm1}X^\mp.\iota_N(v)$ while the right hand side equals $q^{\mp1}w^{\pm1}X^\mp.\iota_N(v)$. Conditions (7.2.28) thus read

$$(z^{\pm1} - w^{\pm1})X^\mp.u \stackrel{!}{=} 0 \qquad \forall u \in \mathrm{Im}(\iota_N) \tag{7.2.29}$$

This simply boils down to the requirement that $z = w$. □

In much the same way, an explicit calculation shows that the sequence (7.2.24) persists in the the root of unity case, but the computations are rather lengthy and tedious. One takes the Clebsch-Gordan sequence in $U_\hbar(\mathfrak{sl}_2)$ and promotes each module V_m to an evaluation representation $V_m(w)$ of $U_\hbar(L\mathfrak{sl}_2)$ with arbitrary parameter w. The intertwiner conditions then reduce the freedom to choose these parameters to the choice of one single parameter.

7.2.4. An exact tensor functor $\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2)) \to \mathcal{T}_F$

Let us fix

$$\hbar = -\frac{\kappa\omega^2}{2} \; . \tag{7.2.30}$$

In this section we establish the relationship between certain representations of the quantum group $U_\hbar(L\mathfrak{sl}_2)$ and perturbed T-defects. The aim is to build a tensor functor $\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2)) \to \mathcal{T}_F$. If this functor is exact, then one can study composition and superposition of perturbed T-operators by studying exact sequences in $\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2))$.

This endeavour has already been carried out in [BR] in a very similar case. However, we have to additionally cater to the special needs of the twist defect. We have seen that in order to deal with the defect crossings it is desirable to have defects which automatically come with a half-braiding. In fact, we construct the following sequence of tensor functors:

$$\widetilde{T}_{\mathrm{pert}} : \mathrm{Rep}_{H_0-\mathrm{ss}}(U_\hbar(L\mathfrak{sl}_2)^+) \xrightarrow{J} \mathcal{C}_F \xrightarrow{\widetilde{\iota^+}} \mathcal{Z}(\mathcal{C})_F \xrightarrow{\widetilde{\alpha}} \mathcal{Z}(\mathcal{D})_F \tag{7.2.31}$$

The functor $\widetilde{T}_{\mathrm{pert}}$ sends the evaluation representation $V_n(\lambda)$ (or rather its Borel part) to the object $((T_n, \varphi_{T_n,-}), \lambda f_n)$, and we will show that it is faithful and exact. Of course, by composing with forgetful functors we do in fact recover a functor $\mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2)) \to \mathcal{T}_F$ as promised.

The functor $J : \mathrm{Rep}_{H_0-\mathrm{ss}}(U_\hbar(L\mathfrak{sl}_2)^+) \to \mathcal{C}_F$

Definition 7.2.7. The (topological) bialgebra A_F is defined as quotient of $\mathcal{E}(h; f^\pm)$ by the closure of the two-sided ideal generated by the relations

$$[h, f^\pm] = \pm\omega f^\pm \; , \tag{7.2.32}$$

with coproduct

$$\Delta(h) = h \otimes 1 + 1 \otimes h \; , \qquad \Delta(f^\pm) = f^\pm \otimes 1 + e^{\pm\kappa\omega h} \otimes f^\pm \; . \tag{7.2.33}$$

The following three Lemmas (7.2.8, 7.2.9 and 7.2.10) are slight adaptations of results in [BR, Sec. 3].

Lemma 7.2.8. *There is a strict monoidal equivalence*

$$I : \mathrm{Rep}_{h-\mathrm{ss}}(A_F) \longrightarrow \mathcal{C}_F \tag{7.2.34}$$

Proof. Define I as follows:

- An A_F-representation M is mapped to $(M, m) \in \mathcal{C}_F$, where the $\mathbb{C}$-grading on M is given by decomposition into eigenspaces of the h-action, and the morphism $m : F \otimes M \to M$ decomposes into two morphisms $m^\pm : \mathbb{C}_{\pm\omega} \otimes M \to M$ which emulate the action of $f^\pm$, respectively: for all $v \in M$ we set

$$m^\pm(1_{\pm\omega} \otimes v) := f^\pm.v \; . \tag{7.2.35}$$

This is well-defined, as m preserves the grading: suppose $v \in M_\alpha$; then

$$h.(f^\pm.u) \;=\; [h, f^\pm].u + f^\pm.(h.u) \;=\; \pm\omega f^\pm.u + f^\pm.(\alpha u) \;=\; (\alpha \pm \omega) f^\pm.u \;, \tag{7.2.36}$$

from which we see that $m^\pm(1_{\pm\omega} \otimes u) \in M_{\alpha\pm\omega}$.

- An intertwiner $\varphi : M \to N$ is mapped to itself under I. It is naturally a morphism of $\mathcal{C}_F$, since it commutes with m:

$$m^\pm(1_{\pm\omega} \otimes \varphi(v)) \;=\; f^\pm.\varphi(v) \;=\; \varphi(f^\pm.v) \;=\; \varphi(m(1_{\pm\omega} \otimes \varphi(v))) \tag{7.2.37}$$

In fact, it is straightforward to see that I is an equivalence (with strict inverse).

Let M, N be A_F-representations. On one hand, we have

$$I(M) \otimes_F I(N) = (M \otimes N, T(m, n)) \tag{7.2.38}$$

where $T(m, n)$ decomposes into $T^\pm(m, n) : \mathbb{C}_{\pm\omega} \otimes M \otimes N \to M \otimes N$ given (for $v \in M$, $w \in N$) by

$$\begin{aligned} T(m,n)^\pm(1_{\pm\omega} \otimes v \otimes w) \;&=\; \big(m^\pm \otimes \mathrm{id_N} + (\mathrm{id}_M \otimes n^\pm) \circ (c_{\pm\omega,M} \otimes \mathrm{id}_N)\big)\,(1_{\pm\omega} \otimes v \otimes w) \\ &=\; (f^\pm.v) \otimes w + (e^{\pm\kappa\omega h}.v) \otimes (f^\pm.w) \;. \end{aligned} \tag{7.2.39}$$

On the other hand, we have

$$I(M \otimes N) = (M \otimes N, o) \tag{7.2.40}$$

where o is the sum of morphisms $o^\pm : \mathbb{C}_{\pm\omega} \otimes M \otimes N \to M \otimes N$ given by

$$o^\pm(1_{\pm\omega} \otimes v \otimes w) \;=\; f^\pm.(v \otimes w) \;=\; \Delta(f^\pm)(v \otimes w) \;=\; (f^\pm.v) \otimes w + (e^{\pm\kappa\omega h}.v) \otimes (f^\pm.w) \;. \tag{7.2.41}$$

Since (7.2.39) and (7.2.41) coincide, I is strict monoidal. □

Lemma 7.2.9. *Let $\xi \in \mathbb{C}^\times$. The map $u : A_F \to U_\hbar(L\mathfrak{sl}_2)^+$, given by*

$$u(h) = \frac{\omega}{2} H_0, \qquad u(f^+) = X_0^+ e^{-\hbar H_0} \;, \qquad u(f^-) = \xi \frac{e^\hbar - e^{-\hbar}}{e^{2\hbar}} X_1^+ e^{-\hbar H_1} \;, \tag{7.2.42}$$

defines a surjective algebra homomorphism. If $\hbar = -\frac{\kappa\omega^2}{2}$, then u is a also a coalgebra homomorphism.

Proof. The map u is compatible with the algebra structure:

$$\begin{aligned} u([h, f^+]) \;&=\; u(\omega f^+) \;=\; \omega X_0^+ e^{-\hbar H_0} \;=\; \frac{\omega}{2}[H_0, X_0^+] e^{-\hbar H_0} \\ &=\; \frac{\omega}{2} H_0 X_0^+ e^{-\hbar H_0} - X_0^+ e^{-\hbar H_0} \frac{\omega}{2} H_0 \;=\; u(h)u(f^+) - u(f^+)u(h) \end{aligned} \tag{7.2.43}$$

$$
\begin{aligned}
u([h,f^-]) &= u(-\omega f^-) \;=\; -\omega\,\xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}X_1^+e^{-\hbar H_0} \\
&= -\frac{\omega}{2}\,\xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}[H_1,X_1^+]e^{-\hbar H_0} \;=\; \frac{\omega}{2}\,\xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}[H_0,X_1^+]e^{-\hbar H_0} \\
&= \frac{\omega}{2}H_0\,\xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}X_1^+e^{-\hbar H_1} - \xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}X_1^+e^{-\hbar H_1}\frac{\omega}{2}H_0 \\
&= u(h)u(f^-)-u(f^-)u(h)
\end{aligned}
\tag{7.2.44}
$$

Denote the coproduct of A_F by Δ^A, and the coproduct of $U_\hbar(L\mathfrak{sl}_2)$ by Δ^L. If we fix $\hbar = -\frac{\kappa\omega^2}{2}$ then the map u is compatible with the coalgebra structure:

$$
(u\otimes u)(\Delta^A(h)) \;=\; (u\otimes u)(h\otimes 1+1\otimes h) \;=\; \frac{\omega}{2}(H_0\otimes 1+1\otimes H_0) \;=\; \Delta^L u(h) \tag{7.2.45}
$$

$$
\begin{aligned}
(u\otimes u)(\Delta^A(f^+)) &= (u\otimes u)(f^+\otimes 1+e^{\kappa\omega h}\otimes f^+) \\
&= X_0^+e^{-\hbar H_0}\otimes 1+e^{\frac{\kappa\omega^2}{2}H_0}\otimes X_0^+e^{-\hbar H_0} \\
&= (X_0^+\otimes e^{\hbar H_0}+1\otimes X_0^+)(e^{-\hbar H_0}\otimes e^{-\hbar H_0}) \\
&= (\Delta^L(X_0^+))(\Delta^L(e^{-\hbar H_0})) \;=\; \Delta^L(X_0^+e^{-\hbar H_0}) \;=\; \Delta^L(u(f^+))
\end{aligned}
\tag{7.2.46}
$$

$$
\begin{aligned}
(u\otimes u)(\Delta^A(f^+)) &= (u\otimes u)(f^+\otimes 1+e^{\kappa\omega h}\otimes f^+) \\
&= \xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}(X_1^+e^{\hbar H_0}\otimes 1+e^{-\frac{\kappa\omega^2}{2}H_0}\otimes X_1^+e^{\hbar H_0}) \\
&= \xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}(X_1^+\otimes e^{-\hbar H_0}+1\otimes X_1^+)(e^{\hbar H_0}\otimes e^{\hbar H_0}) \\
&= \xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}(\Delta^L(X_1^+))(\Delta^L(e^{-\hbar H_1})) \\
&= \xi\frac{e^{\hbar}-e^{-\hbar}}{e^{2\hbar}}\Delta^L(X_1^+e^{-\hbar H_1}) \;=\; \Delta^L(u(f^-))
\end{aligned}
\tag{7.2.47}
$$

The homomorphism u hits all the generators of $U_\hbar(L\mathfrak{sl}_2)^+$:

$$
u(\tfrac{2}{\omega}h)=H_0\ ,\quad u(-\tfrac{2}{\omega}h)=H_1\ ,\quad u(f^+e^{\hbar\frac{2}{\omega}h})=X_0^+\ ,\quad u(\tfrac{e^{2\hbar}}{\xi(e^{\hbar}-e^{-\hbar})}f^-e^{\hbar\frac{2}{\omega}h})=X_1^+ \tag{7.2.48}
$$

Thus u is surjective. □

Let A, B be algebras, and $g : A \to B$ a surjective algebra homomorphism. Let (M,ρ) be a representation of B. The assignment $(M,\rho)\mapsto(M,\rho\circ g)$ defines a fully faithful and exact functor $G : \mathrm{Rep}(B)\to\mathrm{Rep}(A)$. If A and B are bi-algebras (making $\mathrm{Rep}(B)$ and $\mathrm{Rep}(A)$ monoidal) and g is a bi-algebra homomorphism (i.e. respects the co-multiplication which induces the tensor product of representations), then G is monoidal.

Applying these general considerations to Lemma 7.2.9 yields a fully faithful and exact monoidal functor

$$
G : \mathrm{Rep}_{H_0-\mathrm{ss}}(U_\hbar(L\mathfrak{sl}_2)^+) \longrightarrow \mathrm{Rep}_{h-\mathrm{ss}}(\mathcal{A}_F)\ . \tag{7.2.49}
$$

Denote by $\Pi^+ : \mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2)) \to \mathrm{Rep}(U_\hbar(L\mathfrak{sl}_2)^+)$ the functor which forgets the action of the negative Borel half. This functor is faithful, exact and monoidal. Composing (7.2.34) and (7.2.49) with Π^+ we obtain:

Lemma 7.2.10. *The functor*

$$J := I \circ G \circ \Pi^+ : \mathrm{Rep}_{H_0-\mathrm{ss}}(U_\hbar(L\mathfrak{sl}_2)) \longrightarrow \mathcal{C}_F \tag{7.2.50}$$

is faithful, exact, and monoidal.

The functor $\widetilde{\alpha} : \mathcal{Z}(\mathcal{C})_F \to \mathcal{Z}(\mathcal{D})_F$

Let $\mathcal{A}, \mathcal{B}$ be abelian braided monoidal categories, and $G : \mathcal{A} \to \mathcal{B}$ an additive monoidal functor. For any $U \in \mathcal{A}$, consider the functor

$$\widetilde{G} : \mathcal{A}_U \longrightarrow \mathcal{B}_{G(U)} \tag{7.2.51}$$

which maps an object (V, m) to $(G(V), G(m))$ and a morphism $f : (V, m) \to (W, n)$ to $G(f)$. Note that the latter assignment is well-defined : f satisfies $f \circ m = n \circ (\mathrm{id}_U \otimes f)$, from which, by monoidality of G, it follows that also $G(f) \circ \eta_{UV}^{-1} \circ G(m) \circ \eta_{UV} = G(n) \circ (\mathrm{id}_{G(U)} \otimes G(f))$. Thus, $G(f)$ is indeed a morphism in $\mathcal{B}_{G(U)}$.

Under certain conditions, $\widetilde{G}$ inherits good properties from G:

Lemma 7.2.11. *1. If G is braided, then $\widetilde{G}$ is monoidal.*

2. If G is faithful, then also $\widetilde{G}$ is faithful.

3. If G is fully faithful, then also $\widetilde{G}$ is fully faithful.

4. Suppose the tensor products in $\mathcal{A}$ and $\mathcal{B}$ are right-exact. Then the following holds: if G is exact, then also $\widetilde{G}$ is exact.

Proof. 1: Since G is monoidal, the natural isomorphisms $\eta_{VW} : G(V) \otimes G(W) \to G(V \otimes W)$ are compatible with the associators of $\mathcal{A}$ and $\mathcal{B}$. Since G is additive and respects the braiding, we have $T(G(m), G(n)) = \eta_{VW}^{-1} \circ G(T(m, n)) \circ \eta_{U,V\otimes W} \circ (\mathrm{id}_U \otimes \circ \eta_{VW})$. It follows that η provide natural isomorphisms

$$\begin{aligned} \widetilde{G}(V, m) \otimes_{G(U)} \widetilde{G}(W, n) &= (G(V), G(m)) \otimes_{G(U)} (G(W), G(n)) \\ &= (G(V) \otimes G(W), T(G(m), G(n))) \\ &\xrightarrow{\cong} (G(V \otimes W), G(T(m, n))) = \widetilde{G}(V \otimes W, T(m, n)) . \end{aligned} \tag{7.2.52}$$

These isomorphisms are compatible with the associators of $\mathcal{A}_U$ and $\mathcal{B}_{G(U)}$, since the latter are just the associators of $\mathcal{A}$ and $\mathcal{B}$, respectively.

2: This is obvious.

3: Let $h : (G(V), G(m)) \to (G(W), G(n))$ be a morphism in $\mathcal{B}_{G(U)}$. By definition, we have $h \circ \eta_{UV}^{-1} \circ G(m) \circ \eta_{UV} = G(n) \circ (\mathrm{id}_{G(U)} \otimes G(f))$. If G is full, then there exists

$f : V \to W$, such that $G(f) = h$ as a morphism in $\mathcal{B}$. Hence, $G(f) \circ \eta_{UV}^{-1} \circ G(m) \circ \eta_{UV} = G(n) \circ (\mathrm{id}_{G(U)} \otimes G(f))$. Since G is monoidal and additive, this implies that $f \circ m - n \circ (\mathrm{id}_U \otimes f) \in \ker G_{U \otimes V, W}$, where $G_{V,W} : \mathrm{Hom}_{\mathcal{A}}(V, W) \to \mathrm{Hom}_{\mathcal{B}}(G(V), G(W))$ denote the abelian group homomorphisms induced by G. The kernels of all these maps are trivial, since G is faithful. Hence, f lifts to a morphism in $\mathcal{A}_U$, which proves that $\widetilde{G}$ is full. Faithfulness follows from 2.

4: In [MaRu, Thm. 2.3] it was shown that, if the tensor product of $\mathcal{A}$ is right-exact, a sequence in $\mathcal{A}_U$ is exact if and only if the underlying sequence in $\mathcal{A}$ is exact. The same is of course true for sequences in $\mathcal{B}_{G(U)}$. The claim follows from combining this with exactness of G. □

Notations 7.2.12.

1. Denote by $\mathcal{C}_{\mathrm{fd}} \subset \mathcal{C}$ the subcategory of finite-dimensional $\mathbb{C}$-graded vector spaces.
2. In the same way that we used the abbreviation $\mathcal{D}_F = \mathcal{D}_{\alpha\iota(F)}$ (see Notations 6.2.1), we also set $\mathcal{Z}(\mathcal{C})_F := \mathcal{Z}(\mathcal{C})_{\iota(F)}$, $\mathcal{Z}(\mathcal{C}_{\mathrm{fd}})_F := \mathcal{Z}(\mathcal{C}_{\mathrm{fd}})_{\iota(F)}$ and $\mathcal{Z}(\mathcal{D}_{\mathrm{ss}})_F := \mathcal{Z}(\mathcal{D}_{\mathrm{ss}})_{\alpha\iota(F)}$.

Now we would like to specialise (7.2.51) to the case where $G = \alpha : \mathcal{Z}(\mathcal{C}) \to \mathcal{Z}(\mathcal{D})$ and $U = F$, i.e. to

$$\widetilde{\alpha} : \mathcal{Z}(\mathcal{C})_F \to \mathcal{Z}(\mathcal{D})_F \ . \tag{7.2.53}$$

In fact, however, it is sufficient for our purpose to consider the restriction of α to $\mathcal{A} = \mathcal{Z}(\mathcal{C}_{\mathrm{fd}})$, and the corresponding restriction of $\widetilde{\alpha}$. We note that in our specific case (that is, $\mathcal{C}$ being the category of $\mathbb{C}$-graded vector spaces) the image of $\mathcal{Z}(\mathcal{C}_{\mathrm{fd}})$ under α is contained in $\mathcal{B} = \mathcal{Z}(\mathcal{D}_{\mathrm{ss}})$. Recall that the functor α is braided monoidal. Moreover, α is faithful, since the functors $\alpha^{\pm} : \mathcal{C} \to \mathcal{D}$ are faithful. To see that α is exact, note that $\alpha^{\pm}$ are additive functors between semi-simple abelian categories and therefore exact. The tensor product in $\mathcal{Z}(\mathcal{C}_{\mathrm{fd}})$ is (left and right) exact because $\mathcal{C}_{\mathrm{fd}}$ is rigid and the monoidal center of a rigid category is rigid. The tensor product in $\mathcal{Z}(\mathcal{D}_{\mathrm{ss}})$ is exact for the same reason (recall that $\mathcal{D}_{\mathrm{ss}}$ contains only finite direct sums, so one can define duals). Hence, Lemma 7.2.11 yields:

Lemma 7.2.13. *The restriction of $\widetilde{\alpha}$ to $\mathcal{Z}(\mathcal{C}_{\mathrm{fd}})_F$ is monoidal, faithful and exact.*

Likewise, Lemma 7.2.11 can be used to show that the functor $\widetilde{\iota^+} : \mathcal{C}_F \to \mathcal{Z}(\mathcal{C})_F$ induced from $\iota^+ : \mathcal{C} \to \mathcal{Z}(\mathcal{C})$ is monoidal, fully faithful and exact when restricted to $(\mathcal{C}_{\mathrm{fd}})_F$. At last, notice that the image of the functor $J : \mathrm{Rep}_{H_0-\mathrm{ss}}(U_\hbar(L\mathfrak{sl}_2)^+) \to \mathcal{C}_F$ is contained in $(\mathcal{C}_{\mathrm{fd}})_F$ since the domain contains only finite-dimensional representations. Hence, we have successfully constructed all the ingredients for the functor $\widetilde{T}_{\mathrm{pert}}$ advertised in (7.2.31). Altogether, we conclude

Proposition 7.2.14. *The functor $\widetilde{T}_{\mathrm{pert}} = \widetilde{\alpha} \circ \widetilde{\iota^+} \circ J$ is monoidal, faithful and exact.*

If we want the functor to end in $\mathcal{D}_F$, we can lift also the forgetful functor $U : \mathcal{Z}(\mathcal{D}_{\mathrm{ss}}) \to \mathcal{D}$ to a monoidal, faithful and exact functor $\widetilde{U} : \mathcal{Z}(\mathcal{D}_{\mathrm{ss}})_F \to \mathcal{D}_F$. The composition $T_{\mathrm{pert}} = \widetilde{U} \circ \widetilde{T}_{\mathrm{pert}}$ yields the monoidal faithful exact functor (7.2.15).

7.2.5. Fusion relations of perturbed T-defects

Now we can exploit the exact monoidal functor (7.2.31) to transport the exact sequences (7.2.24) and (7.2.25) as well as the isomorphisms $V_n(z) \otimes V_m(w) \cong V_m(w) \otimes V_n(z)$ from $\mathrm{Rep}_{H_0-\mathrm{ss}}(U_\hbar(L\mathfrak{sl}_2))$ to $\mathcal{Z}(\mathcal{D}_{\mathrm{ss}})_F$. We assume that all involved perturbed defect operators converge and that Condition 7.2.1 holds. Propositions 6.2.11 and 6.2.12 then yield the following relations:

- for any $q \neq \pm 1$, and for all $z, w \in \mathbb{C}^\times$ and $n, m \in \mathbb{Z}_{>0}$:

$$\mathsf{T}^P_n(z)\mathsf{T}^P_m(w) \;=\; \mathsf{T}^P_m(w)\mathsf{T}^P_n(z) \tag{7.2.54}$$
$$\mathsf{T}^P_1(q^{-1}z)\mathsf{T}^P_n(q^n z) \;=\; \mathsf{T}^P_{n-1}(q^{n+1}z) + \mathsf{T}^P_{n+1}(q^{n-1}z) \tag{7.2.55}$$

- for $q \neq \pm 1$ a root of unity (with N the smallest positive integer such that $q^N \in \{-1, 1\}$) and for all $z \in \mathbb{C}^\times$:

$$\mathsf{T}^P_N(z) \;=\; \mathsf{T}^P_{N-2}(z) + \mathsf{T}^{(N)P}_0 + \mathsf{T}^{(-N)P}_0 \tag{7.2.56}$$

Lemma 7.2.15. *For all $w \in \mathbb{C}$, we have $\mathsf{T}^{(w)P}_0 = e^{-\kappa\omega w(a_0 \otimes \mathrm{id})}$.*

Proof. The operator $\mathsf{T}^{(w)P}_0$ is the unperturbed defect operator for the induced defect $\alpha(\mathbb{C}_{\frac{\omega}{2}w})$. In the TFT-formalism it is thus easy to see (compare Example 6.1.4) that for $v \in V_p \otimes V_q$ and $u \in V_{-p} \otimes V_{-q}$ we have

$$\left(u, \mathsf{T}^{(w)P}_0 .v\right)_P \;=\; \frac{S_{-\frac{\omega}{2}w,p}}{S_{0,p}}\,(u,v)_P \;=\; e^{-\kappa\omega w p}\,(u,v)_P \ , \tag{7.2.57}$$

where $S_{pq} = e^{2\kappa pq}$ denotes the matrix element of the modular S-matrix (7.1.8) in $\mathcal{C}$ for simple objects $\mathbb{C}_p, \mathbb{C}_q$. Recall that a_0 is the operator which extracts the degree via $(a_0 \otimes \mathrm{id}_{\mathbb{C}_q}).v = pv$ and $(\mathrm{id}_{\mathbb{C}_p} \otimes a_0).v = qv$. From this, the claim follows. □

Remark 7.2.16. Note that the asymmetric combination $(a_0 \otimes \mathrm{id})$ is due to the fact that we chose the crossing τ to be induced by the braiding in $\mathcal{C}$ (rather than the inverse braiding). This choice was necessary in order for the T-operators to remain integrals of motion under chiral perturbations. Had we chosen the anti-chiral perturbation $F^- = \mathbb{C}_\omega \oplus \mathbb{C}_{-\omega}$ instead, τ would need to be defined from the inverse braiding, which would replace $(a_0 \otimes \mathrm{id})$ in the above lemma by $(\mathrm{id} \otimes a_0)$.

Thanks to Lemma 7.2.15, the relation 7.2.56 simplifies to

$$\mathsf{T}^P_N(z) \;=\; \mathsf{T}^P_{N-2}(z) + e^{-\kappa\omega N(a_0 \otimes \mathrm{id})} + e^{\kappa\omega N(a_0 \otimes \mathrm{id})} \ . \tag{7.2.58}$$

Lemma 7.2.17. *Let $q \neq \pm 1$ and $z \in \mathbb{C}^\times$. For all $n \in \mathbb{Z}_{>0}$ the following relation holds:*

$$\mathsf{T}^P_n(qz)\mathsf{T}^P_n(q^{-1}z) \;=\; \mathrm{id} + \mathsf{T}^P_{n-1}(z)\mathsf{T}^P_{n+1}(z) \ , \tag{7.2.59}$$

Proof. The case $n = 1$ is identical to (7.2.55) if we take into account $\mathsf{T}_0^P = \mathrm{id}$ from Lemma 7.2.15. Now let us assume that (7.2.59) holds for some n. Then we compute

$$\begin{aligned}
\mathsf{T}_{n+1}^P(qz)\mathsf{T}_{n+1}^P(q^{-1}z) &= \mathsf{T}_{n+1}^P(qz)\left(\mathsf{T}_1^P(q^{-n-1}z)\mathsf{T}_n^P(z) - \mathsf{T}_{n-1}^P(qz)\right) \\
&\overset{(7.2.55)}{=} \left(\mathsf{T}_n^P(q^2z) + \mathsf{T}_{n+2}^P(z)\right)\mathsf{T}_n^P(z) - \mathsf{T}_{n+1}^P(qz)\mathsf{T}_{n-1}^P(qz) \\
&\overset{(7.2.59)}{=} \mathrm{id} + \mathsf{T}_{n-1}^P(qz)\mathsf{T}_{n+1}^P(qz) + \mathsf{T}_{n+2}^P(z)\mathsf{T}_n^P(z) - \mathsf{T}_{n+1}^P(qz)\mathsf{T}_{n-1}^P(qz) \\
&\overset{(7.2.54)}{=} \mathrm{id} + \mathsf{T}_{n+2}^P(z)\mathsf{T}_n^P(z)\ ,
\end{aligned} \tag{7.2.60}$$

which completes the induction. □

7.3. Spectrum of T-operators

In this section, we attack the problem of computing the spectrum of T-operators. The problem can be formulated in terms of a Y-system of the kind discussed in chapter 1.

Let us fix the convention

$$\kappa = -i\pi \tag{7.3.1}$$

for the rest of this chapter.

7.3.1. The BLZ T-system

Let $v \in \mathcal{H}_P$ be an eigenvector of the Hamiltonian $L_0 + \bar{L}_0$. Since the perturbed T-operators preserve the grading (commute with the Hamiltonian) and commute among themselves, v is also an eigenvector of all the perturbed T-operators. For the corresponding eigenvalues $T_n^P(\lambda)$, defined by

$$\mathsf{T}_n^P(\lambda)\, v = T_n^P(\lambda)\, v\ , \tag{7.3.2}$$

Lemma 7.2.17 implies the relations

$$T_n^P(q\lambda)T_n^P(q^{-1}\lambda) \;=\; 1 + T_{n-1}^P(\lambda)T_{n+1}^P(\lambda)\ . \tag{7.3.3}$$

This set of relations is called the *T-system*. If q is a root of unity, (7.2.58) yields additionally the relation

$$T_N^P(\lambda) \;=\; T_{N-2}^P(\lambda) + 2\cos\left(\pi\omega N p\right)\ . \tag{7.3.4}$$

Remark 7.3.1. The choice of the parameter ω is free; in particular it does not depend on the radius r. If $\omega^2 \in \mathbb{Q}$, then we are in the root of unity case (see (7.2.30)), and the T-system truncates at N the smallest integer such that $N\frac{\omega^2}{2} \in \mathbb{Z}$. There is a natural choice for ω, however: if one considers a deformation ("bulk perturbation") of the free boson (for example the sine-Gordon model) it must take a value $\omega \in \frac{1}{r}\mathbb{Z} \setminus \{0\}$ [BR, Rem. 6.3]. This reduces the radii for which a truncation of the T-system can be achieved to those for which $r^2 \in \mathbb{Q}$.

If we set $\omega = \frac{2}{r}$, the truncation relation (7.3.4) becomes

$$T_N^P(\lambda) \;=\; T_{N-2}^P(\lambda) + 2\cos\left(2\pi \tfrac{N}{r} p\right) \tag{7.3.5}$$

The relations (7.3.3) and (7.3.5) together reproduce exactly the T-system of [BLZ4], see equations (1.2) and (4.2) therein, which those authors obtain in the context of the free boson with a background charge.

7.3.2. The Y-system and TBA equations

From now on, let us fix $r \in \mathbb{R}_{\geq 0}$, such that

$$r^2 \in \mathbb{Q}_{>0} \; . \tag{7.3.6}$$

Set $\omega = \frac{2}{r}$, so $q = e^{\frac{2\pi i}{r^2}}$ is a root of unity. Define, for all $z \in \mathbb{C}$, the exponential $q^z := e^{2\pi i \frac{z}{r^2}}$. Moreover, set

$$\mu := pr \; , \tag{7.3.7}$$

which is just the momentum quantum number as in (7.1.18). Define the following products of T-eigenvalues, for all $\lambda \in \mathbb{C}$ and $n = 0, 1, \ldots, N$:

$$Y_n^P(\lambda) \;=\; \begin{cases} 0 & \text{if } n = 0 \\ T_{n-1}^P(\lambda) T_{n+1}^P(\lambda) & \text{if } n = 1, \ldots, N-2 \\ q^{N\mu}\, T_{N-2}^P(\lambda) & \text{if } n = N-1 \\ q^{-N\mu}\, T_{N-2}^P(\lambda) & \text{if } n = N \end{cases} \tag{7.3.8}$$

Denote by

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & & & & & \\ 1 & 0 & 1 & & & & \\ & 1 & 0 & & & & \\ & & & \ddots & & & \\ & & & & 0 & 1 & 1 \\ & & & & 1 & 0 & 0 \\ & & & & 1 & 0 & 0 \end{pmatrix} \tag{7.3.9}$$

the adjacency matrix (with entries D_{nm}) of the D_N Dynkin diagram.

Proposition 7.3.2. *The following relation holds, for $n = 1, \ldots, N$:*

$$Y_n^P(q\lambda) Y_n^P(q^{-1}\lambda) \;=\; \prod_{m=1}^{N} \left(1 + Y_m^P(\lambda)\right)^{D_{nm}} \tag{7.3.10}$$

Proof. For $n = 1, \ldots, N-3$ as well as $n = N-1, N$, this follows immediately by substituting (7.3.8) and applying (7.3.3). For $n = N-2$ one obtains in the same way

$$Y_{N-2}^P(q\lambda) Y_{N-2}^P(q^{-1}\lambda) \;=\; \left(1 + Y_{N-3}^P(\lambda)\right)\left(1 + T_{N-2}^P(\lambda) T_N^P(\lambda)\;\right) \; . \tag{7.3.11}$$

In order to express the second factor on the right hand side in terms of Y-functions, the truncation relation (7.3.5) is used:

$$\begin{aligned} 1+T^P_{N-2}(\lambda)T^P_N(\lambda) &= 1+T^P_{N-2}(\lambda)\left(T^P_{N-2}(\lambda)+2\cos\left(2\pi\tfrac{N}{r^2}\mu\right)\right) \\ &= 1+\left(q^{N\mu}+q^{-N\mu}\right)\ T^P_{N-2}(\lambda)+\left(T^P_{N-2}(\lambda)\right)^2 \\ &= \left(1+q^{N\mu}\ T^P_{N-2}(\lambda)\right)\left(1+q^{-N\mu}\ T^P_{N-2}(\lambda)\right) \\ &= \left(1+Y^P_{N-1}(\lambda)\right)\left(1+Y^P_N(\lambda)\right) \end{aligned} \tag{7.3.12}$$

Inserting this back into (7.3.11) completes the proof. □

Remark 7.3.3.

i) The N Y-functions $Y^P_n(z)$ are not all independent. The last Y-function is clearly superfluous, with the equations for $n = N-1$ and N being identical. Hence, there are at most $N-1$ independent Y-functions.

ii) If we restrict to the case of invisible twist defect $P = A_r$ (i.e. $\mu \in \mathbb{Z}$), then it is straightforward to see that (7.3.10) reduces to

$$Y^{A_r}_n(q\lambda)Y^{A_r}_n(q^{-1}\lambda) \;=\; \prod_{m=1}^{N-1}\left(1+Y^{A_r}_m(\lambda)\right)^{B_{nm}} \ , \tag{7.3.13}$$

where B_{nm} denote the entries of the adjacency matrix

$$\mathbf{B} = \begin{pmatrix} 0 & 1 & & & & \\ 1 & 0 & 1 & & & \\ & 1 & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & 2 \\ & & & & 1 & 0 \end{pmatrix} \tag{7.3.14}$$

of the B_{N-1} Dynkin diagram. Notice, that for $N = 4$ there are in fact only 2 independent Y-functions and equations, since the adjacency matrix of the B_3 Dynkin diagram has rank 2.

A logarithmic variable transformation brings the equations (7.3.10) and (7.3.13), respectively, into the form (Y) studied in Part I (with additive shift s instead of multiplicative shift q). In keeping with the conventions of [BLZ4], we define uniquely a new variable θ by setting

$$\lambda = e^{t\theta} \ , \qquad \text{where} \quad t = \tfrac{r^2-2}{r^2} \ , \tag{7.3.15}$$

and demanding that $\theta \in \mathbb{R}$ if $\lambda \in \mathbb{R}_{>0}$. Identifying $q = e^{tis}$ then prompts us to set

$$s = \tfrac{2\pi}{r^2-2} \ . \tag{7.3.16}$$

Rewriting the Y-systems (7.3.10) and (7.3.13) in terms of functions $\theta \mapsto Y_n^P(e^{t\theta})$ (which we, by slight abuse of notation, also denote by Y_n^P) yields

$$Y_n^P(\theta + is)Y_n^P(\theta - is) = \prod_{m=1}^{N} \left(1 + Y_m^P(\theta)\right)^{D_{nm}} , \tag{7.3.17}$$

$$Y_n^{A_r}(\theta + is)Y_n^{A_r}(\theta - is) = \prod_{m=1}^{N-1} \left(1 + Y_m^{A_r}(\theta)\right)^{B_{nm}} . \tag{7.3.18}$$

Unperturbed defects and the constant solution

A simple calculation along the lines of Example 6.1.4 suggests that the unperturbed defect operator $\mathsf{T}_n^P(0)$ acts on the subspace $V_p \otimes_{\mathbb{C}} V_q \subset \mathcal{H}_P$ by the number

$$t_n^P = \sum_{k=-\frac{n}{2}}^{\frac{n}{2}} \frac{S_{k\omega,p}}{S_{0p}} = \sum_{k=-\frac{n}{2}}^{\frac{n}{2}} e^{-2\pi i k\omega p} = \sum_{k=-\frac{n}{2}}^{\frac{n}{2}} q^{2k\mu} \tag{7.3.19}$$

Indeed, these eigenvalues satisfy the constant version of the relations (7.3.3) and (7.3.5):

Lemma 7.3.4. *The numbers t_n^P satisfy the relations*

$$\left(t_n^P\right)^2 = 1 + t_{n-1}^P t_{n+1}^P , \tag{7.3.20}$$

$$t_N^P = t_{N-2}^P + 2\cos(2\pi \tfrac{N}{r} p) . \tag{7.3.21}$$

Proof. In the same way as in (7.2.2), one calculates that $\left(t_n^P\right)^2 = \sum_{j=0}^{n} t_{2j}^P$ and $t_{n-1}^P t_{n+1}^P = \sum_{j=1}^{n} t_{2j}^P$, from which relation (7.3.20) follows. Relation (7.3.21) is obvious. $\square$

Consequently, the numbers

$$y_n^P = \begin{cases} 0 & \text{if } n = 0 \\ t_{n-1}^P t_{n+1}^P & \text{if } n = 1, \dots, N-2 \\ q^{N\mu}\, t_{N-2}^P & \text{if } n = N-1 \\ q^{-N\mu}\, t_{N-2}^P & \text{if } n = N \end{cases} \tag{7.3.22}$$

satisfy the constant Y-systems

$$\left(y_n^P\right)^2 = \prod_{m=1}^{N} \left(1 + y_m^P\right)^{D_{nm}} \quad \text{or} \quad \left(y_n^{A_r}\right)^2 = \prod_{m=1}^{N-1} \left(1 + y_m^{A_r}\right)^{B_{nm}} , \tag{7.3.23}$$

respectively.

Remark 7.3.5. For $P = A_r$ and $p = 0$ we obtain the solution $t_n^{A_r} = n + 1$. Hence, the corresponding numbers $y_n^{A_r}$ are positive and real. According to Corollary 2.1.3, this is the only constant positive real solution to the Y-systems (7.3.17) and (7.3.18).

The asymptotics

We would like to derive TBA equations for the solutions of (7.3.17). As we have learned in Part I, in order to find the solutions which correspond to physical eigenvalues of the T-operators, we need another input datum, namely the asymptotics of the Y-functions. On one hand, for $\theta \to -\infty$ we clearly have constant asyptotics, since this limit yields the constant unperturbed ($\lambda = 0$) solution. On the other hand, the asymptotics for $\theta \to \infty$ are a priori undetermined from our approach. From the physics literature, however, it is known that the correct asymptotics should be the ones discussed in Example (1.1.2) with minimal γ.

Let $\mathbf{w} \in \mathbb{R}^N$ be the Perron-Frobenius vector of the D_N-type adjacency matrix. The corresponding Perron-Frobenius eigenvalue is

$$\lambda_{\mathrm{PF}} = 2\cos\left(\tfrac{\pi}{2(N-1)}\right) . \tag{7.3.24}$$

Thus, we choose $\gamma = \frac{\pi}{2(N-1)}$, and together with $s = \frac{\pi\xi}{2}$ we find that the family of functions $\mathbf{a} : \mathbb{C} \to \mathbb{C}^N$, defined as

$$\mathbf{a}(\theta) := L\mathbf{w}\exp\left(\tfrac{\pi}{2s(N-1)}\theta\right) , \tag{7.3.25}$$

with $L \in \mathbb{C}$, are valid asymptotics for (7.3.17).

For the reduction of the asymptotics to the $N-1$ dimensional Y-system (7.3.18) we note that the vector $\mathbf{w}^\vee \in \mathbb{R}^{N-1}$ with components $w_n^\vee := w_n$ ($n = 1, \dots, N-1$) is an eigenvector of the B_{N-1} adjacency matrix with eigenvalue λ_{PF}. Since $\mathbf{w}^\vee$ is strictly positive, it has to be the Perron-Frobenius eigenvector. Hence, the reduced asymptotics compatible with (7.3.25) is simply the function $\mathbf{a}^\vee : \mathbb{C} \to \mathbb{C}^{N-1}$ given by

$$\mathbf{a}^\vee(\theta) := L\mathbf{w}^\vee\exp\left(\tfrac{\pi}{2s(N-1)}\theta\right) . \tag{7.3.26}$$

TBA equations

Let $Y_1, \dots, Y_N \in \mathcal{A}(\mathbb{S}_s)$ be a solution to the D_N-type Y-system (7.3.17) with finite root configuration $\varrho \in \mathcal{R}_N$ in $\mathbb{S}_s$. Suppose it corresponds to a physical eigenstate of the P-twisted compactified free boson. We would like the corresponding eigenvalues of the T-operators to be real if λ is positive real. Hence, we are going to assume that $Y_n(\mathbb{R}) \subset \mathbb{R}$ for $n = 1, \dots, N-2$, and $Y_n(\mathbb{R}) \subset q^{\pm N\mu}\mathbb{R}$ for $n = N-1, N$. But this implies in fact that $\varrho \in \mathcal{R}_N^{\mathrm{cc}}$ (compare Lemma 3.2.5).

Let us use the root detector $\mathbf{B}^\varrho : \mathbb{S}_s \to \mathbb{C}^N$ from Example 3.2.7, with components

$$B_n^\varrho(\theta) = \prod_{j=1}^{M_n} \tanh\left(\frac{\pi}{4s}\left(\theta - z_j^{(n)}\right)\right) . \tag{7.3.27}$$

Note that we have $\mathbf{d}^\varrho = 0$ in this case. Due to $\varrho \in \mathcal{R}_N^{\mathrm{cc}}$, we also have $\mathbf{B}^\varrho(\mathbb{R}) \subset \mathbb{R}$, and we can set $\mathbf{k} = 0$ in (TBA$^\varrho$). According to Proposition (3.2.8), we can now write the Y-functions as

$$Y_n^P(\theta) = B_n^\varrho(\theta)e^{a_n(\theta)+f_n(\theta)} , \tag{7.3.28}$$

where $\mathbf{f}(\theta) = (f_1(\theta), \ldots, f_N(\theta))$ satisfies, for any $\mathbf{C} \in \mathrm{Mat}_{<2}(N, \mathbb{R})$, the TBA equation

$$\mathbf{f}(\theta) \; = \; \int_{-\infty}^{\infty} \Phi_{\mathbf{C}}(\theta - \theta') \cdot \left[\mathbf{D} \cdot \log\left(e^{-\mathbf{a}(\theta')} + \mathbf{B}^{\varrho}(\theta') e^{\mathbf{f}(\theta')}\right) - \mathbf{C} \cdot \mathbf{f}(\theta')\right] d\theta' \; , \tag{7.3.29}$$

where $\Phi_{\mathbf{C}}(\theta)$ is the function defined in (1.1.3), and $\mathbf{D}$ is the adjacency matrix (7.3.9) of the D_N Dynkin diagram. Practically, it is often more convenient to solve for the unbounded function

$$\boldsymbol{\varepsilon}(\theta) := \mathbf{f}(\theta) + \mathbf{a}(\theta) \; , \tag{7.3.30}$$

in terms of which the TBA equations (for the standard choices $\mathbf{C} = 0$ and $\mathbf{C} = \mathbf{D}$) read

$$\mathbf{C} = 0 : \qquad \boldsymbol{\varepsilon}(\theta) \; = \; \mathbf{a}(\theta) + \int_{-\infty}^{\infty} \Phi_{\mathbf{0}}(\theta - \theta') \cdot \mathbf{D} \cdot \left[\log\left(1 + \mathbf{B}^{\varrho}(\theta') e^{\boldsymbol{\varepsilon}(\theta')}\right) - \mathbf{a}(\theta')\right] d\theta' \; , \tag{7.3.31}$$

$$\mathbf{C} = \mathbf{D} : \qquad \boldsymbol{\varepsilon}(\theta) \; = \; \mathbf{a}(\theta) + \int_{-\infty}^{\infty} \Phi_{\mathbf{D}}(\theta - \theta') \cdot \mathbf{D} \cdot \log\left(e^{-\boldsymbol{\varepsilon}(\theta')} + \mathbf{B}^{\varrho}(\theta')\right) d\theta' \; . \tag{7.3.32}$$

We will see that in our example the TBA equation with $\mathbf{C} = 0$ is much more well-behaved numerically.

The TBA equations for the reduced B_N-type Y-system are deduced analogously. Let $\varrho^{\vee} \in \mathcal{R}^{\mathrm{cc}}_{N-1}$ be the finite root configuration obtained from ϱ by forgetting the number M_N and the unordered M_N-tuple $(z_1^{(N)}, \ldots, z_{M_N}^{(N)})$ of roots. We can write, for $n = 1, \ldots, N-1$,

$$Y_n^{A_r} = B_n^{\varrho^{\vee}}(\theta) e^{\varepsilon_n^{\vee}(\theta)} \; , \tag{7.3.33}$$

where $\boldsymbol{\varepsilon}^{\vee}(\theta) = (\varepsilon_1(\theta), \ldots, \varepsilon_{N-1}(\theta))$ solves

$$\boldsymbol{\varepsilon}^{\vee}(\theta) \; = \; \mathbf{a}^{\vee}(\theta) + \int_{-\infty}^{\infty} \Phi_{\mathbf{0}}(\theta - \theta') \cdot \mathbf{B} \cdot \left[\log\left(1 + \mathbf{B}^{\varrho^{\vee}}(\theta') e^{\boldsymbol{\varepsilon}^{\vee}(\theta')}\right) - \mathbf{a}^{\vee}(\theta')\right] d\theta' \tag{7.3.34}$$

with $\mathbf{B}$ the adjacency matrix (7.3.14) of the B_{N-1} Dynkin diagram.

7.4. Example: numerical results for $r^2 = 6$

In this section, we compile some numerical results for the case $r^2 = 6$, which yields a T-system that truncates at $N = 3$. Other relevant parameters and relations are then

$$q = e^{\frac{i\pi}{3}} \quad , \qquad s = \tfrac{\pi}{4} \quad , \qquad \lambda = e^{\frac{4\theta}{3}} \quad , \qquad \mathbf{w} = (\sqrt{2}, 1, 1) \; . \tag{7.4.1}$$

We consider only twist defects of the form

$$P = X_{[0], e^{i\varphi}} \qquad , \; \varphi \in [0, 2\pi) \; , \tag{7.4.2}$$

and states $v \in \mathcal{H}_P$ with zero winding number ($v \in \mathbb{C}_p \otimes_{\mathbb{C}} \mathbb{C}_{-p} \subset \mathcal{H}_P$). An eigenvalue of $\mathsf{T}_n^P(\lambda)$ corresponding to a state in $\mathcal{H}_P$ will henceforth be denoted by $T_n^\varphi(\theta) = T_n^P(\theta)$,[1] and similarly we write $Y_n^\varphi(\theta) = Y_n^P(\theta)$ for the corresponding Y-functions. In the special case at hand, the introduction of the Y-functions is somewhat superfluous, since they are simply given by

$$Y_1^\varphi(\theta) = T_2^\varphi(\theta) \qquad , \qquad Y_2^\varphi(\theta) = e^{i\pi\mu} T_1^\varphi(\theta) \qquad , \qquad Y_3^\varphi(\theta) = e^{-i\pi\mu} T_1^\varphi(\theta) \tag{7.4.3}$$

Note that the possible values of $\mu = pr$ are constrained to

$$\mu \;\in\; \tfrac{\varphi}{2\pi} + \mathbb{Z} \tag{7.4.4}$$

by (7.1.19). In the untwisted case ($P = A_r$), μ is an integer and $Y_2(\theta) = Y_3(\theta)$.

The D_3 Y-system is, explicitely,

$$\begin{aligned} Y_1^\varphi(\theta + i\tfrac{\pi}{4}) Y_1^\varphi(\theta - i\tfrac{\pi}{4}) &= (1 + Y_2^\varphi(\theta))(1 + Y_3^\varphi(\theta)) \;, \\ Y_2^\varphi(\theta + i\tfrac{\pi}{4}) Y_2^\varphi(\theta - i\tfrac{\pi}{4}) &= (1 + Y_1^\varphi(\theta)) \;, \\ Y_3^\varphi(\theta + i\tfrac{\pi}{4}) Y_3^\varphi(\theta - i\tfrac{\pi}{4}) &= (1 + Y_1^\varphi(\theta)) \;. \end{aligned} \tag{7.4.5}$$

Getting rid of the last Y-function yields the system

$$\begin{aligned} T_2^\varphi(\theta + i\tfrac{\pi}{4}) T_2^\varphi(\theta - i\tfrac{\pi}{4}) &= (1 + e^{i\pi\mu} T_1^\varphi(\theta))(1 + e^{-i\pi\mu} T_1^\varphi(\theta)) \;, \\ T_1^\varphi(\theta + i\tfrac{\pi}{4}) T_1^\varphi(\theta - i\tfrac{\pi}{4}) &= (1 + T_2^\varphi(\theta)) \;, \end{aligned} \tag{7.4.6}$$

which becomes the B_2 Y-system in the untwisted case. Notice the symmetry $\mu \mapsto -\mu$. The corresponding ($\mathbf{C} = 0$) TBA equations for a solution where $Y_1^\varphi(\theta)$ has M_1 roots $z_1^{(1)}, \ldots, z_{M_1}^{(1)} \in \mathbb{S}_{\frac{\pi}{4}}$ and $Y_2^\varphi(\theta)$ has M_2 roots $z_1^{(2)}, \ldots, z_{M_2}^{(2)} \in \mathbb{S}_{\frac{\pi}{4}}$ are

$$\begin{aligned} \varepsilon_1(\theta) &= L\sqrt{2}\, e^\theta + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{\cosh(2(\theta - \theta'))} \left(L_2(\theta') - 2Le^{\theta'}\right) d\theta' \;, \\ \tilde{\varepsilon}_2(\theta) &= Le^\theta + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{\cosh(2(\theta - \theta'))} \left(L_1(\theta') - L\sqrt{2} e^{\theta'}\right) d\theta' \;, \end{aligned} \tag{7.4.7}$$

where

$$\begin{aligned} L_2(\theta') &= \log\left(\left(e^{i\pi\mu} + e^{\tilde{\varepsilon}_2(\theta')} \prod_{j=1}^{M_2} \tanh(\theta' - z_j^{(2)}) \right) \left(e^{-i\pi\mu} + e^{\tilde{\varepsilon}_2(\theta')} \prod_{j=1}^{M_2} \tanh(\theta' - z_j^{(2)}) \right) \right) \;, \\ L_1(\theta') &= \log\left(1 + e^{\varepsilon_1(\theta')} \prod_{j=1}^{M_1} \tanh(\theta' - z_j^{(1)}) \right) \;. \end{aligned} \tag{7.4.8}$$

The Y-functions are expressed in terms of a solution $(\varepsilon_1(\theta), \tilde{\varepsilon}_2(\theta))$ as follows:

$$Y_1^\varphi(\theta) \;=\; e^{\varepsilon_1(\theta)} \prod_{j=1}^{M_1} \tanh(\theta' - z_j^{(1)}) \quad , \quad Y_2^\varphi(\theta) \;=\; e^{i\pi\mu + \tilde{\varepsilon}_2(\theta)} \prod_{j=1}^{M_2} \tanh(\theta' - z_j^{(2)}) \;. \tag{7.4.9}$$

[1] Just as for Y-functions, by slight abuse of notation it is understood that $T_n(\theta) = T_n(\lambda)$, where the relation between λ and θ is as in (7.4.1).

Remark 7.4.1. TBA equations whose solutions have a complex phase $e^{i\pi\mu}$ have been studied before, see e.g. [KM, Fe1]. From a thermodynamic perspective μ often has the interpretation of a chemical potential.

7.4.1. The vacuum state

Untwisted vacuum It is natural to assume that the vacuum state in $\mathcal{H}$ (i.e. for $\varphi = 0$) corresponds to the unique positive real solution of the Y-system afforded by Theorem 2.1.1 with no roots in $\overline{\mathbb{S}}_s$. It is not hard to solve the corresponding system of TBA equations (7.4.7) numerically to great precision with a Piccard iteration scheme. The convergence is excellent, even if it is unclear whether the TBA equations are in fact a contraction in the sense of Chapter 1. The first two plots in figure 7.2 show the resulting eigenvalues $T_1^0(\theta)$ and $T_2^0(\theta)$ for $L = 1$ plotted along the real line. Notice that the $\theta \to -\infty$ asymptotics $T_1^0(-\infty) = 2$ and $T_2^0(-\infty) = 3$ coincide with the constant solution of Remark 7.3.5.

Let us also note that the asymptotic behaviour of the numerical solution for $\theta \to \infty$ is consistent with Proposition 3.1.5. From the numerical solution, we can extract the recursively defined functions

$$\mathbf{c}_1(\theta) \;=\; e^{\frac{4g_1}{\pi}\theta}\mathbf{f}(\theta) \quad , \qquad \mathbf{c}_j(\theta) \;=\; e^{\frac{4g_j}{\pi}\theta}\Big(\mathbf{c}_{j-1}(\theta) - \lim_{\theta'\to\infty}\mathbf{c}_{j-1}(\theta')\Big) \ . \tag{7.4.10}$$

For $j = 1, 2, 3$ we find indeed

$$\lim_{\theta\to\infty}\mathbf{c}_j(\theta) \propto \mathbf{v}^{[j]} \ , \tag{7.4.11}$$

where $\mathbf{v}^{[j]}$ is the $[j]$-th eigenvector of the B_2 adjacency matrix.[2] [3] This is in agreement with (3.1.17).

It is interesting to study the analytic continuation of the solution in the complex θ-plane.[4] It seems that both $T_1^0(\theta)$ and $T_1^0(\theta)$ have an infinite number of roots, all of which are lined up on the axis $\theta = \frac{3\pi}{4}i + \mathbb{R}$, which corresponds to negative real λ. This is in accordance with similar observations in the literature [BLZ4, DT].

Twisted vacuum The vacuum state in the twisted space $\mathcal{H}_P$ has $\mu = \frac{\varphi}{2\pi}$. Certainly, we expect the corresponding solution $T_n^\varphi(\theta)$ to be a continuous (perhaps analytic) deformation of the untwisted vacuum solution $T_n^0(\theta)$ in the parameter μ. For starters, the positive constant solution $t_1^{A_r} = 2$, $t_2^{A_r} = 3$ of Remark 7.3.5 can be analytically continued in μ, namely according to (7.3.19):

$$\begin{aligned} t_1^P &= 2\cos(\tfrac{\pi}{3}\mu) \ , \\ t_2^P &= 2\cos(\tfrac{2\pi}{3}\mu) + 1 \end{aligned} \tag{7.4.12}$$

[2] See Notations 3.1.4 for the definition of the square bracket and the definition of g_j.

[3] For $j > 3$, our numerical solution $\mathbf{c}_j(\theta)$ exhibits too much noise at $\theta \to \infty$ for a clear limit to be extracted. To some extent, this situation could be improved by a more sophisticated numerical scheme.

[4] Using contour deformation to avoid the singularities of $\Phi_{\mathbf{0}}(\theta)$ in $\theta = \pm\frac{\pi}{4}$, this analytic continuation can be numerically obtained from the TBA equations themselves inside $\mathbb{S}_{\frac{\pi}{4}}$. Beyond the strip, we use the Y-system to continue it to the rest of the complex plane.

For $\mu \in (-1, 1)$, these numbers ought to be the $\theta \to -\infty$ asymptotics of the twisted vacuum solution $T_n^\varphi(\theta)$.

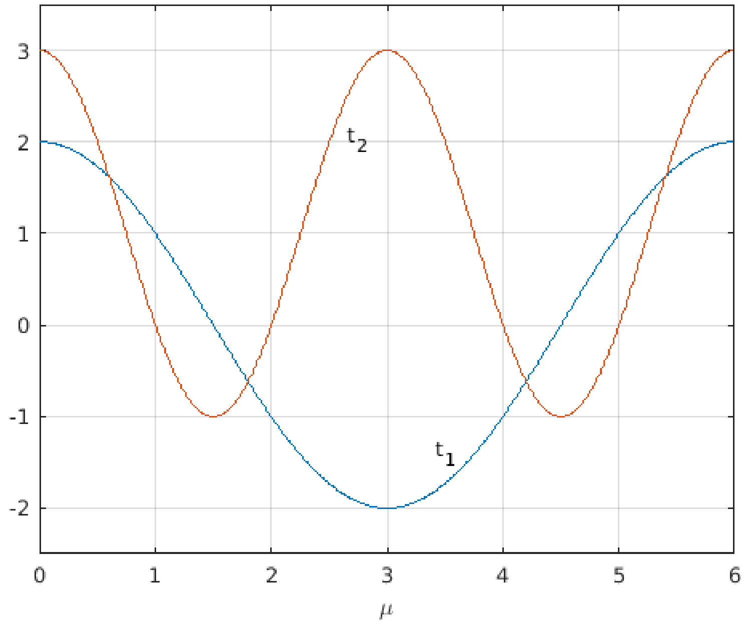


Figure 7.1.: The constant asymptotics t_1^P and t_2^P as functions of the twist parameter μ.

The question arises whether $T_n^\varphi(\theta)$ still has no roots in the strip $\mathbb{S}_{\frac{\pi}{4}}$. Let us ponder this question for a moment. There are only two ways how roots could enter the strip during a continuous deformation in the real parameter μ: either they enter at $\theta \to -\infty$, or they enter sideways as complex-conjugate pairs through the boundary $\partial\mathbb{S}_{\frac{\pi}{4}}$. Both possibilities can be excluded in our case:

- For $|\mu| < 1$ both numbers t_1^P and t_2^P remain positive (and real). Hence, no root enters the strip at $\theta \to -\infty$.

- According to Lemma 3.2.1, any roots entering through $\partial\mathbb{S}_s$ would make themselves known by forcing either $T_1^\varphi(\theta) = -1$ or $T_2^\varphi(\theta) = -1$ for some real θ. But we just argued that both $T_1^\varphi(\theta)$ and $T_1^\varphi(\theta)$ have positive asymptotics at $\theta \to \pm\infty$. Moreover, we have $T_1^\varphi(\mathbb{R}) \subset \mathbb{R}$ and $T_2^\varphi(\mathbb{R}) \subset \mathbb{R}$. Hence, for the function $T_n^\varphi(\theta)$ to drop to -1 anywhere on the real axis, we would need it to have at least 2 roots inside $\mathbb{S}_{\frac{\pi}{4}}$ already *prior* to the new roots entering. This is a contradiction.

Indeed, iteration of the TBA equations for zero roots converges well on the entire interval $\mu \in (-1, 1)$ and yields a solution with constant $\theta \to -\infty$ asymptotics given by (7.4.12). Figure 7.2 shows this solution as a function of real θ for several values of μ. A closer examination of the analytic continuation shows that all the roots remain lined up on the negative real λ-line (which in the θ-plane corresponds to the lines $\theta = \mathbb{R} \pm \frac{3\pi}{4} i$), but they move towards $\lambda = 0$ ($\theta \to -\infty \pm \frac{3\pi}{4} i$) as $|\mu|$ increases.

Notice that as we approach $\mu = 1$ it becomes increasingly crucial that we chose to work with the TBA equations for $\mathbf{C} = 0$. The iteration of the standard TBA equations with $\mathbf{C} = \mathbf{G}$ does not converge near $\mu = 1$, since they involve integrating over $\log(e^{\pm i\pi\mu} + 1/Y_2(\theta))$ instead of $\log(e^{\pm i\pi\mu} + Y_2(\theta))$, and we have $Y_2(\theta) \to 1$ as $\theta \to -\infty$ at $\mu = 1$.

According to Remark 3.1.6 and [BLZ4], the quantities

$$I_j \;=\; \langle \mathbf{w}^{[j]}, \mathbf{X}^j \rangle \;=\; \int_{-\infty}^{\infty} e^{\frac{4g_j}{\pi}\theta} \left\langle \mathbf{w}^{[j]}, \log\left(e^{-\boldsymbol{\varepsilon}(\theta)} + 1\right) \right\rangle d\theta \tag{7.4.13}$$

in the asymptotic expansion (3.1.17) for $\theta \to \infty$ are expected to be the eigenvalues of *local* integrals of motion. In particular, I_1 is expected to yield the energy eigenvalue:

$$E \;=\; -\int_{-\infty}^{\infty} e^{\theta} \left[\sqrt{2} \log\left(e^{-\varepsilon_1(\theta)} + 1\right) + \log\left(e^{i\pi\mu - \tilde{\varepsilon}_2(\theta)} + 1\right) + \log\left(e^{-i\pi\mu - \tilde{\varepsilon}_2(\theta)} + 1\right)\right] d\theta \;. \tag{7.4.14}$$

Figure 7.3 shows a plot of this expression as a function of μ, computed from our numerical solution $\varepsilon_1(\theta), \tilde{\varepsilon}_2(\theta)$ to the TBA equations. The graph is a perfect parabola in μ (with fitting error of the same order of magnitude as the TBA solution), which is precisely what one expects: $p = \frac{\mu}{r}$ is the momentum eigenvalue of the state, and the energy of free particles is quadratic in the momentum.

7.4.2. Excited states by continuous deformation of the twist

In the process of continuously deforming the vacuum solution from $\mu = 0$ to $\mu = 1$, we have clearly picked up a monodromy: the $\mu = 1$ solution corresponds not to the vacuum, but to the highest weight state in the sector $V_{\mu r} \otimes_{\mathbb{C}} V_{-\mu r} \subset \mathcal{H}$. In fact, there is hope that this process of "creating" new states can be continued by increasing μ further.

The $\theta \to -\infty$ asymptotics (7.4.12) are shown in figure 7.1. They are periodic with period $\mu = 6$. However, this does not mean that at $\mu = 6$ we have returned to the vacuum again. Instead, one would expect to reach ever higher excited states. Solutions for integer μ correspond to states in the untwisted theory. Every time we pass through $\mu \in \mathbb{Z}$, we enter what in physics is called a new "Brillouin zone". The vacuum is in the first Brillouin zone.

In this section we discuss TBA solutions for states with $\mu > 1$. What makes the case $\mu \in (1, 5)$ qualitatively very different from $|\mu| < 1$ is, that some of the t_n^P are negative. Hence the corresponding solutions have roots in the strip $\mathbb{S}_{\frac{\pi}{4}}$.[5] For instance, when passing through $\mu = 1$, $T_2^{\varphi}(\theta)$ acquires a root in $\mathbb{S}_{\frac{\pi}{4}}$ through $\theta \to -\infty$.

Given some generic root configuration $\varrho \in \mathcal{R}_2^{\mathrm{cc}}$, the associated TBA equations can easily be solved numerically in the same way as the TBA equation with zero roots.[6] However, to

[5] Of course, if all t_n^P are positive, the solutions may still have roots. For example, this should be the case for $\mu \in (5, 7)$.

[6] In most cases one would expect the solution to be unique, in full analogy with Theorem 2.1.2. It should not be too hard to generalise this theorem accordingly.

determine whether ϱ yields a physical solution (i.e. one that corresponds to an eigenvector of the T-operators) is much harder. The consistency conditions (3.2.12) can be used to optimise ϱ to a root configuration which yields an entire solution to the Y-system. Yet this is only practicable with a relatively small number of roots in $\mathbb{S}_{\frac{\pi}{4}}$. Already 2 or 3 roots can make it a hard problem to tackle numerically. For $\mu > 1$ we are faced with exactly that problem.

The second Brillouin zone Let us start with the solution for μ slightly larger than 1. The number of roots in the strip should be zero for $T_1^\varphi(\theta)$ and one for $T_2^\varphi(\theta)$, and the location of the only root must be on the real axis (by complex conjugation symmetry). We denote the root, as usual, by $z_1^{(1)} \in \mathbb{R}$.[7] For a fixed value of μ, we solve the corresponding TBA equation for different values of $z_1^{(1)}$, and compute from its solution the expressions

$$m_\pm(z_1^{(1)}) = T_1(z_1^{(1)} + i\tfrac{\pi}{4}) + e^{\pm i\pi\mu} \ . \tag{7.4.15}$$

Condition (3.2.12) is equivalent to $m_+(z_1^{(1)}) = 0$ or $m_-(z_1^{(1)}) = 0$. It is clear from a quick look at the numerical values of $m_\pm(z_1^{(1)})$ as a function of $z_1^{(1)}$, that there exists only one value of $z_1^{(1)}$ which satisfies that condition. This value can be found by a simple Newton-Raphson root search.

The whole procedure can be repeated for different values of $\mu \in (1, \frac{3}{2})$. The qualitative picture remains the same, with the root $z_1^{(1)}$ moving to larger θ as μ increases. Table 7.1 compiles the values of $z_1^{(1)}$ for different μ. We must stress that the precision of the obtained numerical values for $z_1^{(1)}$ (und thus of the actual physical solution of the Y-system) is many orders of magnitude less than the precision for the solution of the integral equations. This rather large error comes from the fact that (7.4.15) involves numerical analytic continuation of a discretised function.

At $\mu = \frac{3}{2}$, the situation changes again: yet another root enters the strip at $\theta \to -\infty$, this time a root of $T_1^\varphi(\theta)$. We denote it by $z_1^{(2)}$, and we know that it has to remain on the real axis as well. As before, we employ Newton-Raphson search to find the combination $(z_1^{(1)}, z_1^{(2)}) \in \mathbb{R}^2$ which best satisfies the consistency conditions 3.2.12 (again, a quick glance at the plots of $T_n(z_1^{(n)} + i\frac{\pi}{4})$ leads to the conclusion that there can be only one such combination). The results for different values of μ can again be found in table 7.1.

Into the third Brillouin zone By the time μ approaches 2, the root $z_1^{(1)}$, which entered the strip first, has undergone a change of direction and is now headed on its way back out again. Notice the scenario that is playing out in the rest of the complex plane at the same time (see figure 7.5): the infinite battery of roots on the negative λ-line (i.e. the lines $\theta = \mathbb{R} \pm \frac{3\pi}{4}i$) is still moving towards $\lambda = 0$, about to inject more roots into the strip $\mathbb{S}_{\frac{\pi}{4}}$. At $\mu = 2$ o'clock, catastrophe hits: the outgoing root $z_1^{(2)}$ collides exactly in $\lambda = 0$

[7] Note that $z_1^{(1)}$ is a root of $Y_1^\varphi(\theta)$, hence the superscript (1).

with a root approaching from the opposite side in the λ-plane.[8] The two roots scatter at a 90 degree angle. For μ slightly larger than 2, they become a complex conjugate pair roughly on the imaginary λ-line (see figure 7.6). We can still solve the corresponding TBA equations (together with the usual consistency condition), for which $z_1^{(2)}$ is now the only root in the strip. Some values of $z_1^{(2)}$, for μ close to 2, can be found in Table 7.1. But as μ increases further, the scattered roots start to deviate from the imaginary λ-line and acquire a nonzero real part. Eventually, close to $\mu = 3$, their trajectories in the λ-plane curve so drastically that they seem to re-enter the strip (which in the λ-plane corresponds to the wedge $|\arg(\lambda)| < \frac{\pi}{3}$) sideways through its boundary $\partial\mathbb{S}_{\frac{\pi}{4}}$ (see figure 7.7).

The moment of re-entry itself is not captured by our investigation, since the numerical scheme becomes increasingly unstable. However, the evidence is quite clear: in the whole interval $\mu \in (2, 4)$, the $\theta \to -\infty$ asymptotics of $T_1^{\varphi}(\theta)$ dips below -1. Hence, for all μ in this range, there is a point $z_0 \in \mathbb{R}$ for which $T_1^{\varphi}(z_0) = -1$. For $\mu = 3$, Lemma 3.2.1 then implies that either $T_1^{\varphi}(z_0 + i\frac{\pi}{4}) = 0$ or $T_1^{\varphi}(z_0 - i\frac{\pi}{4}) = 0$. In fact, due to complex conjugation symmetry both need to hold.

Unfortunately, the roots re-entering the strip sideways have almost the same real part (in the θ-plane) as the root $z_1^{(2)}$ on the real line. Hence, the consistency condition for $z_1^{(2)}$ depends on the value of $T_2(\theta)$ very close to the location of the re-entering root, *and* vice versa. This creates numerical instability from which we do not see a straight-forward escape. It seems, however, possible that one could again track down the solution at higher values of μ. Certainly, a qualitative understanding of the root trajectories should be possible by combining numerical and analytical tools. Nevertheless, the issue discussed here indicates why the derivation of Y-systems and TBA equations covers, from the practical point of view, only half the distance towards solving the spectrum of T-operators.

Interestingly, the problems that have befallen us, and that plague other numerical attempts at solving TBA equations for excited states, such as in [DT], do not appear in all cases: the author of [Fe2] considers exactly the same twisted B_2 Y-system (7.4.6) as we discussed here, albeit in a different physical context. In his case, the asymptotics are different: for $\theta \to -\infty$, they coincide with ours, but for $\theta \to \infty$ they decay. This results in a dramatically different behaviour under deformation in μ than what we observed. We have witnessed that roots moving from left to right on the real axis are at some point forced, due to $T_n^{\varphi}(\theta) \to \infty$ as $\theta \to \infty$, to reverse their direction. This causes roots to collide on the (positive or negative) real λ axis sooner or later (we suspect that the one instance we observed is indicative of a general behaviour), making them scatter all over the place and out of control. In the case of decaying asymptotics, however, the roots entering the strip from the left move along the real axis unhindered and later leave the strip at the other end. They never seem to collide with one another, and hence they are bound to remain real at all times.

[8] This was, in fact, entirely predictible: according to the asymptotics 7.4.12, the function $T_1^{\varphi}(\lambda) + 1$ has a double root in $\lambda = 0$ for $\mu = 2$. According to Lemma 3.2.1, this requires $T_2^{\varphi}(\lambda)$ to have a double root in $\lambda = 0$ at the same time.

Brillouin zone	μ	$z_1^{(1)}$	$z_1^{(2)}$
2	1.01	-2.4420	
	1.05	-1.2460	
	1.1	-0.7407	
	1.2	-0.2527	
	1.3	0.0151	
	1.4	0.1885	
	1.45	0.2528	
	1.49	0.2956	
	1.495	0.3005	
	$\vdots$	$\vdots$	
	1.6	0.3793	-0.6151
	1.7	0.4127	-0.0791
	1.8	0.3942	0.2403
	1.9	0.2718	0.4706
	1.95	0.0918	0.5660
	1.98	-0.1895	0.6185
	1.99	-0.4206	0.6353
	1.999	-1.2396	0.6502
	1.9999	-2.0898	0.6517
3	2.0001		0.6520
	2.001		0.6535
	2.01		0.6680
	2.1		0.8018
	2.2		0.9300
	$\vdots$		$\vdots$

Table 7.1.: Location of roots in the strip

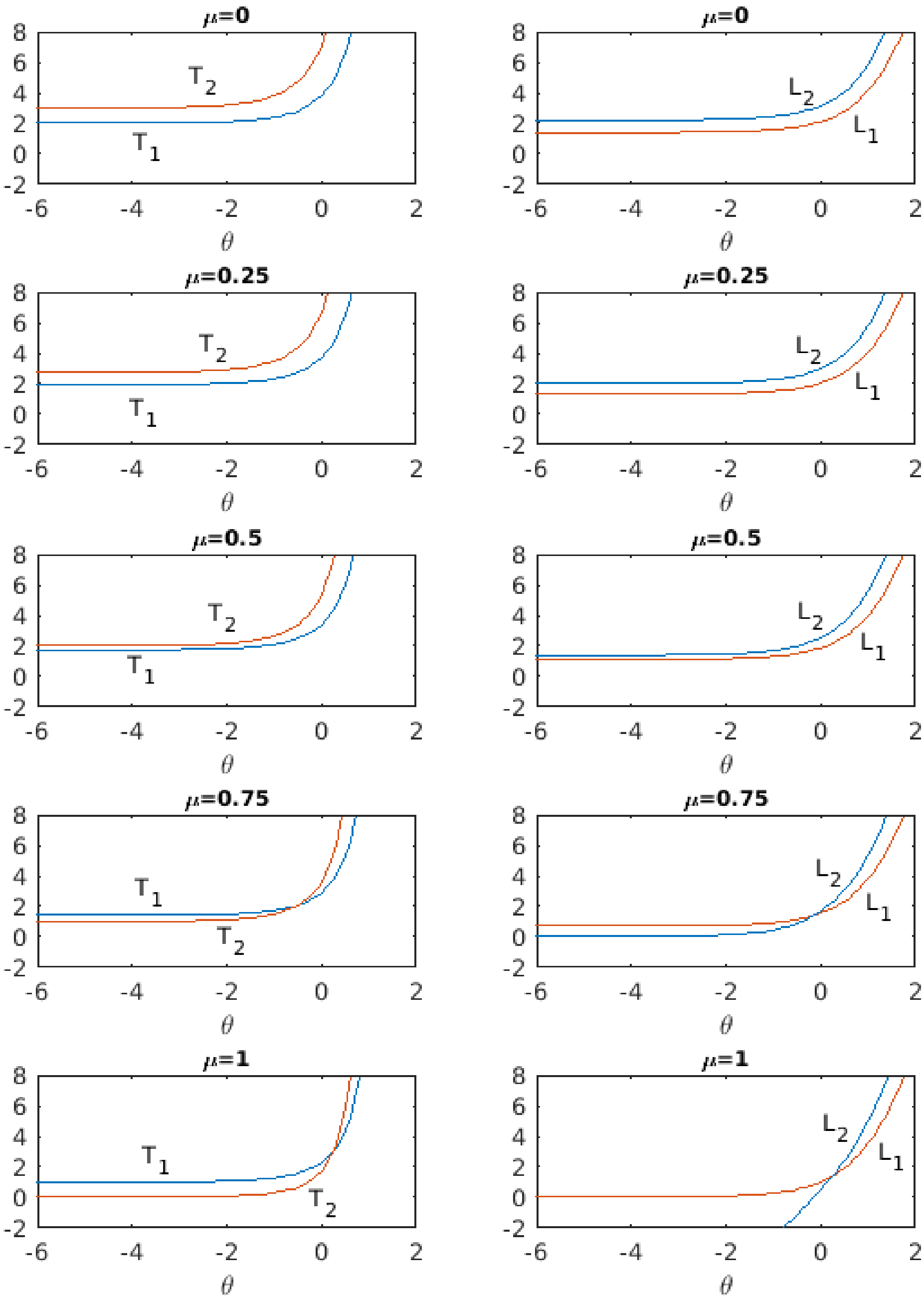


Figure 7.2.: Numerical solutions for the real functions $T_1(\theta)$, $T_2(\theta)$, $L_1(\theta)$ and $L_2(\theta)$, for several values of μ in the first Brillouin zone.

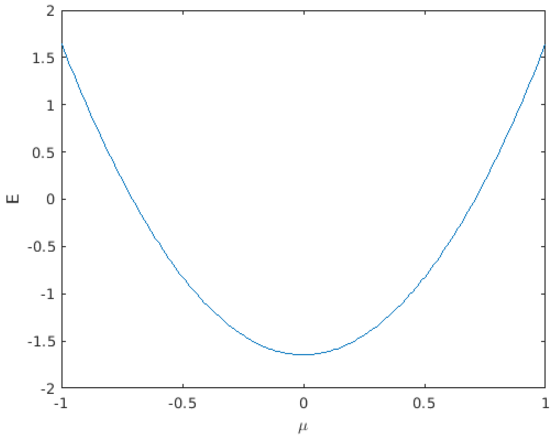


Figure 7.3.: Numerical solution for the conjectured energy eigenvalue E as a function of the twist parameter μ in the first Brillouin zone.

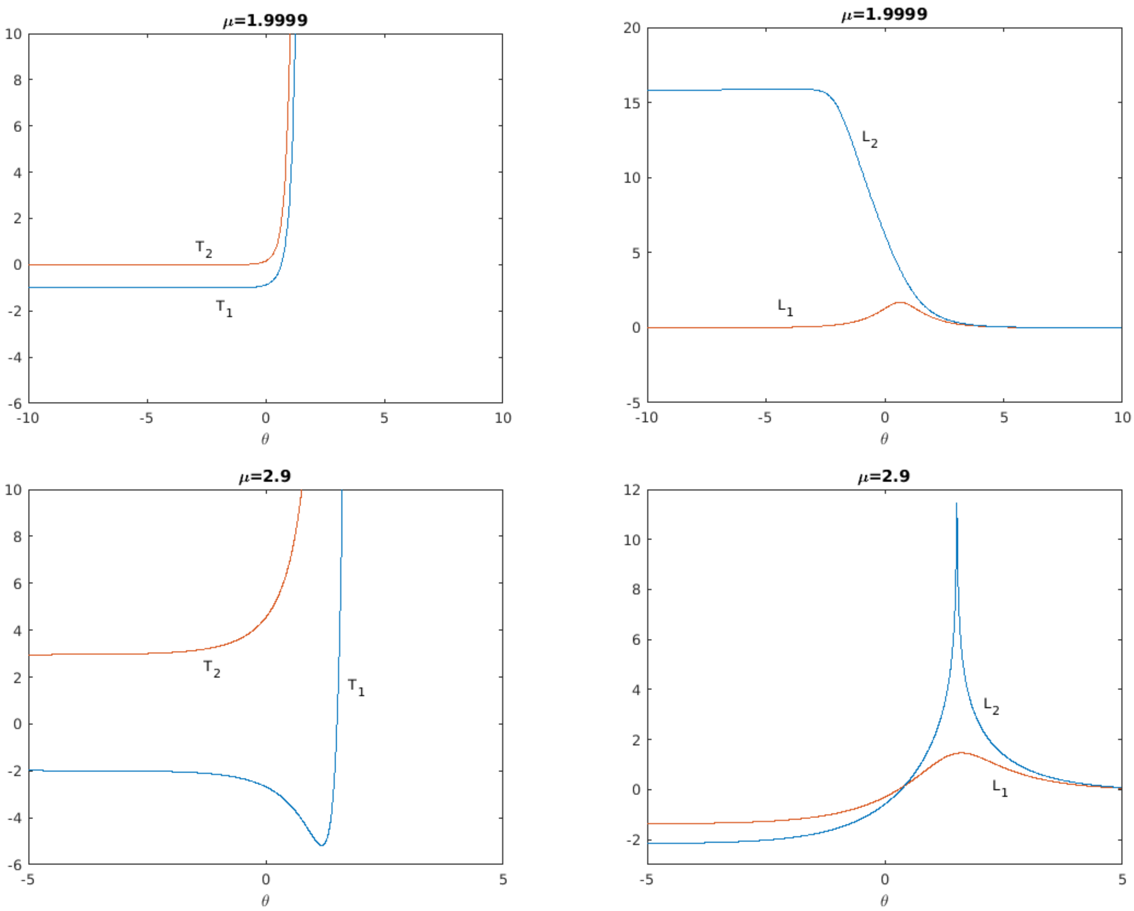


Figure 7.4.: Numerical solutions for the real functions $T_1(\theta)$, $T_2(\theta)$, $L_1(\theta)$ and $L_2(\theta)$, for values of μ near 2 and 3. (Due to the build-up of a singularity in $L_2(\theta)$, the solution for $\mu = 2.9$ acquires a rather significant error towards $\theta \to -\infty$, which is somewhat visible in the lower left plot.)

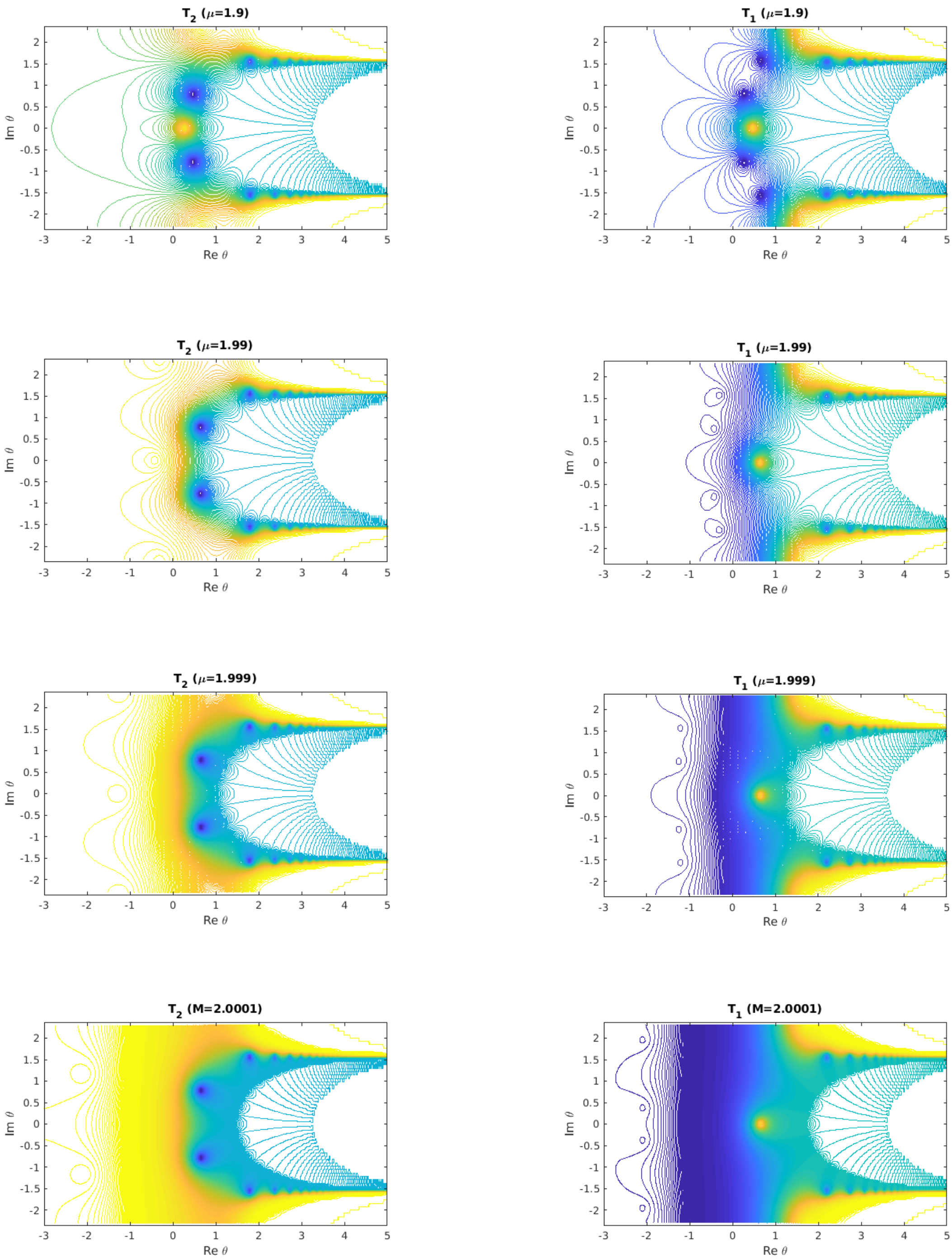


Figure 7.5.: Contour plots of the norm of the functions $T_1^\varphi(\theta)$ and $T_2^\varphi(\theta)$ for values around $\mu = 2$. Yellow corresponds to values close to 0, and blue to values approaching ∞. The distances of the level curves was chosen differently in each figure so as to emphasise the roots and singularities. In the plots on the right-hand side the root $z_1^{(2)}$ can be seen as a yellow dot in the center. In the upper left plot, the root $z_1^{(1)}$ is the yellow dot in the middle. The scattering of this root with another root (coming from the line $\theta = \mathbb{R} \pm \frac{3\pi}{4}$, manifesting itself as a yellow circle) in $\theta \to -\infty$ at $\mu = 2$ is clearly visible in the lower two plots on the left.

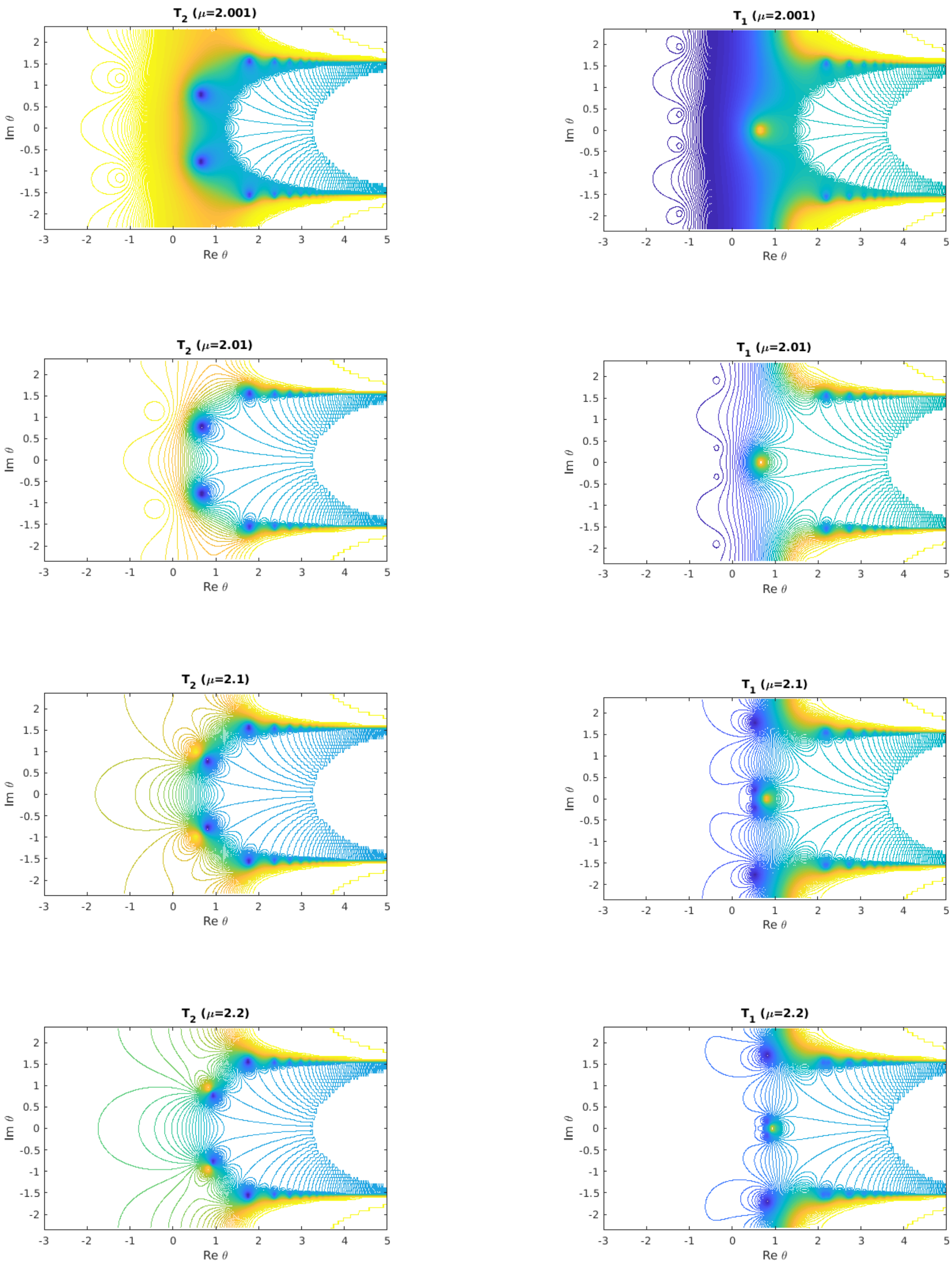


Figure 7.6.: Contour plots of the norm of the functions $T_1^\varphi(\theta)$ and $T_2^\varphi(\theta)$ for values around $\mu > 2$. The scattered roots of $T_2^\varphi(\theta)$ (the two yellow circles in the upper left, or yellow dots in the lower left plots, respectively) can be seen approaching the boundary of the strip $\mathbb{S}_{\frac{\pi}{4}}$ (i.e. the lines $\theta = \mathbb{R} \pm \frac{\pi}{4} i$).

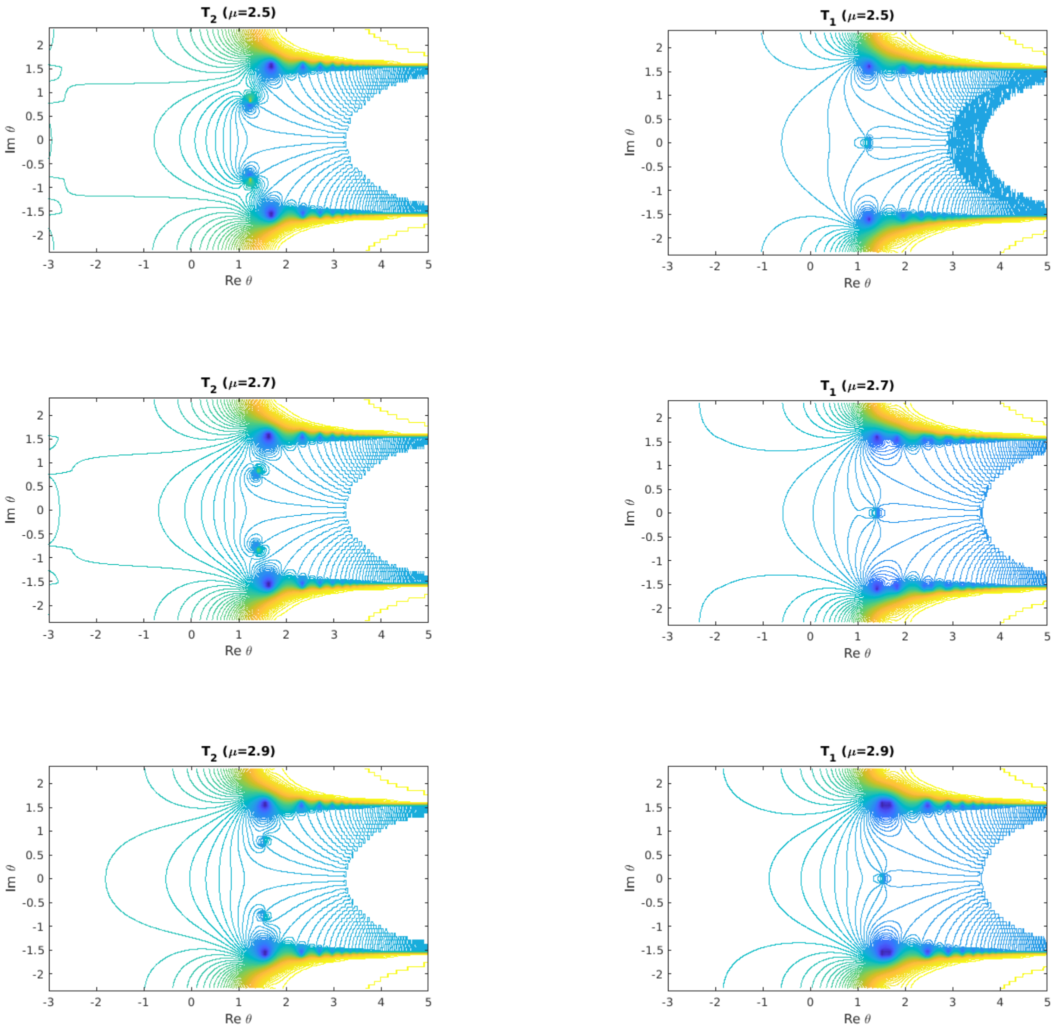


Figure 7.7.: Contour plots of the norm of the functions $T_1^{\varphi}(\theta)$ and $T_2^{\varphi}(\theta)$ for values of μ approaching 3. The scattered roots are further approaching the strip, causing some issues for the numerical scheme used. The numerical solutions in this range are not very accurate anymore, as can be seen at small real values of θ, where the contours in the middle and upper left plots are visibly skewed. Close to $\mu = 3$, all numerical attempts break down.

Acknowledgements

First and foremost I would like to thank Ingo Runkel for his guidance throughout my PhD, for his support as well as his patience during key stages of the endavour, and for carefully reading and commenting on drafts of this thesis. My thanks also go to Gleb Arutyunov who kindly agreed to be the second thesis reviewer. Moreover, I gratefully acknowledge the financial support of the collaborative research center SFB 676 (Particles, Strings and the early Universe).

I am deeply indebted to my former office neighbour Louis-Hadrien Robert for his support and friendship, as well as for his role in bringing so many of us younger colleagues together. I also would like to thank my office mates Vanda Farsad and Tobias Ohrmann for the good time and the many coffee breaks we spent together, as well as Ehud Meir, who not only added his mathematical expertise to those coffee breaks, but also his wonderfully soothing sarcasm. To Rob Klabbers I am grateful for his interest in my work and the opportunity to talk about each other's research. Special thanks go to Lorant Szegedy for being a great PhD brother and flatmate, and for coping so well with this overdose.

Finally, I want to thank my family and friends for their continued love and support, and for reminding me of other things in life.

Bibliography

[AC] J. Appell, C.-J. Chen, *How to Solve Hammerstein Equations*, J. Integral Equations Appl. **18** (2006) 287–296.

[At] M. Atiyah, *Topological quantum field theory*, Publ. math. de l'IHES **68** (1988) 175–186.

[Baj] Z. Bajnok, *Review of AdS/CFT Integrability, Chapter III.6: Thermodynamic Bethe Ansatz*, Lett. Math. Phys. **99** (2012) 299–320, [hep-th/1012.3995].

[BDP] Z. Bajnok, O. el Deeb, P.A. Pearce, *Finite-Volume Spectra of the Lee-Yang Model*, J. High Energy Phys. **73** (2015) 2015:73, [hep-th/1412.8494].

[BK] B. Bakalov, A. A. Kirillov Jr., *Lectures on Tensor Categories and Modular Functors*, University Lecture Series **21**, American Mathematical Society (2001).

[Bax] R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press (1982).

[BLZ1] V.V. Bazhanov, S.L. Lukyanov, A.B. Zamolodchikov, *Integrable Structure of Conformal Field Theory, Quantum KdV Theory and Thermodynamic Bethe Ansatz*, Commun. Math. Phys. **177** (1996) 381–398, [hep-th/9412229].

[BLZ2] V.V. Bazhanov, S.L. Lukyanov, A.B. Zamolodchikov, *Integrable Structure of Conformal Field Theory II. Q-operator and DDV equation*, Commun. Math. Phys. **190** (1997) 247–278, [hep-th/9604044].

[BLZ3] V.V. Bazhanov, S.L. Lukyanov, A.B. Zamolodchikov, *Integrable Structure of Conformal Field Theory III. The Yang-Baxter Relation*, Commun. Math. Phys. **200** (1999) 297–324, [hep-th/9805008].

[BLZ4] V.V. Bazhanov, S.L. Lukyanov, A.B. Zamolodchikov, *Integrable Quantum Field Theories in Finite Volume: Excited State Energies*, Nucl .Phys. **B489** (1997) 487-531, [hep-th/9607099].

[BPZ] A. A. Belavin, A. M. Polyakov, A. B. Zamolodchikov, *Infinite conformal symmetry in two-dimensional quantum field theory*, Nucl. Phys. B **241** (1984) 333–380.

[BH] A.E. Brouwer, W.H. Haemers, *Spectra of graphs*, Springer (2012).

[BR] D. Buecher, I. Runkel, *Integrable perturbations of conformal field theories and Yetter-Drinfeld modules*, J. Math. Phys. **55** (2014) 111705, [1211.4726 [math.QA]].

[CF] O. Castro-Alvaredo, A. Fring, *Chaos in the thermodynamic Bethe ansatz*, Phys. Lett. **A334** (2005) 173–179, [hep-th/0406066].

[CP1] V. Chari, A. Pressley, *A guide to quantum groups*, Cambridge University Press (1994).

[CP2] V. Chari, A. Pressley, *Quantum affine algebras*, Commun. Math. Phys. **142** (1991) 261–283.

[DKR] A. Davydov, L. Kong, I. Runkel, *Field theories with defects and the centre functor*, in: Mathematical Foundations of Quantum Field Theory and Perturbative String Theory, Proceedings of Symposia in Pure Mathematics, American Mathematical Society (2011), [1107.0495 [math.QA]].

[DdV] C. Destri, H. de Vega, *New thermodynamic Bethe ansatz equations without strings*, Phys. Rev. Lett. **69** (1992) 2313–2317.

[DDT] P. Dorey, C. Dunning, R. Tateo, *The ODE/IM Correspondence*, J. Phys. A: Math. Theor. **40** (2007) R205, [hep-th/0703066].

[Do] P. Dorey, *Exact S-matrices*, in: ed. Z. Horvath, L. Palla, Conformal Field Theories and Integrable Models, Lecture Notes in Physics **498**, Springer (1997), [hep-th/9810026].

[DT] P. Dorey, R. Tateo, *Excited states by analytic continuation of TBA equations*, Nucl. Phys. **B482** (1996) 639–659, [hep-th/9607167].

[Dr] V.G. Drinfel'd, *Quantum groups*, in: Proceedings of the ICM 1 (1986).

[EGNO] P. Etingof, S. Gelaki, D. Nikshych, and V. Ostrik, *Tensor categories*, Mathematical Surveys and Monographs **205**, American Mathematical Society (2015).

[Er] A. Erdélyi, W. Magnus, F. Oberhettinger, F. Tricomi, *Tables of Integral Transforms, Vol. 1*, McGraw-Hill (1954).

[Fe1] P. Fendley, *Excited-state thermodynamics*, Nucl. Phys. **B374** (1992) 667–691, [hep-th/9109021].

[Fe2] P. Fendley, *Excited-state energies and supersymmetric indices*, Adv. Theor. Math. Phys. **1** (1997) 210–236, [hep-th/9706161].

[FMQR] D. Fioravanti, A. Mariottini, E. Quattrini, F. Ravanini, *Excited state Destri-De Vega equation for Sine-Gordon and restricted Sine-Gordon models*, Phys. Lett. **B390** (1997) 243–251, [hep-th/9608091].

[FRS5] J. Fjelstad, J. Fuchs, I. Runkel, C. Schweigert, *TFT construction of RCFT correlators V: Proof of modular invariance and factorization*,Theor. Appl. Categor. **16** (2006) 342–433, [hep-th/0503194].

[FjFRS] J. Fjelstad, J. Fuchs, I. Runkel, C. Schweigert, *Uniqueness of open/closed rational CFT with given algebra of open states*, Adv. Theor. Math. Phys. **12** (2008) 1283–1375, [hep-th/0612306].

[FjFS] J. Fjelstad, J. Fuchs, C. Stigner, *RCFT with defects: Factorization and fundamental world sheets* Nucl. Phys. **B863** (2012) 213–259, [hep-th/1202.3929].

[FBZ] E. Frenkel, D. Ben-Zvi, *Vertex Algebras and Algebraic Curves*, Mathematical Surveys and Monographs **88**, American Mathematical Society, 2nd edition (2004).

[FHL] I.B. Frenkel, Y.-Z. Huang, J. Lepowsky, *On Axiomatic Approaches to Vertex Operator Algebras and Modules*, Memoirs of the American Mathematical Society **494** (1991).

[FKS] A. Fring, C. Korff, B.J. Schulz, *The ultraviolet Behaviour of Integrable Quantum Field Theories, Affine Toda Field Theory*, Nucl. Phys. **B549** (1999) 579–612 [hep-th/9902011].

[FrFRS] J. Fröhlich, J. Fuchs, I. Runkel, C. Schweigert, *Duality and defects in rational conformal field theory*, Nucl. Phys. B **763** (2007), 354–430, [hep-th/0607247].

[FGRS] J. Fuchs, M. R. Gaberdiel, I. Runkel, C. Schweigert, *Topological defects for the free boson CFT*, J. Phys. A **40** (2007) 11403, [hep-th/0705.3129].

[FRS1] J. Fuchs, I. Runkel, C. Schweigert, *TFT construction of RCFT correlators I: Partition functions*, Nucl. Phys. B **646** (2002) 353–497, [hep-th/0204148].

[FRS3] J. Fuchs, I. Runkel, C. Schweigert, *TFT construction of RCFT correlators III: Simple currents*, Nucl. Phys. B **694** (2004), 277–353, [hep-th/0403157].

[FRS4] J. Fuchs, I. Runkel, C. Schweigert, *TFT construction of RCFT correlators IV: Structure constants and correlation functions*, Nucl. Phys. B **715** (2005), 539–638, [hep-th/0412290].

[Gak] F.D. Gakhov, *Boundary value problems*, Pergamon Press (1966).

[Gaw] K. Gawedzki, *Lectures on Conformal Field Theory*, in: Quantum Fields and Strings: A Course for Mathematicians, vol. 2, ed. P. Deligne et al., American Mathematical Society (1999).

[GKV] N. Gromov, V. Kazakov, P. Vieira, *Finite volume spectrum of 2D field theories from Hirota dynamics* J. High Energy Phys. 0912 (2009) 060, [hep-th/0812.5091].

[Gr] N. Gromov, *Introduction to the Spectrum of N=2 SYM and the Quantum Spectral Curve*, Les Houches Summer School 2016 lecture notes, preprint, [hep-th/1708.03648].

[Hu] Y.-Z. Huang, *Rigidity and modularity of vertex tensor categories*, Comm. Contemp. Math. **10** (2008), 871–918, [0502533 [math.QA]].

[IIKKN] R. Inoue, O. Iyama, B. Keller, A. Kuniba, T. Nakanishi, *Periodicities of T and Y-systems, dilogarithm identities, and cluster algebras I: Type* B_r, Publ. RIMS **49** (2013) 1–42, [1001.1880 [math.QA]].

[Kac] V. Kac, *Vertex algebras for beginners*, University Lecture Series **10**, American Mathematical Society (1998).

[Kas] C. Kassel, *Quantum Groups*, Graduate Texts in Mathematics **155**, Springer (1995).

[Kau] L. H. Kauffman, *Knots and physics*, Series on knots and everything **1**, World Scientific Publishing (1991).

[Ki] A.N. Kirillov, *Identities for the Rogers dilogarithm function connected with simple Lie algebras*, J. Soviet Math. **47** (1989) 2450–2459.

[KLWZ] I. Krichever, O. Lipan, P. Wiegmann, A. Zabrodin, *Quantum Integrable models and discrete classical Hirota equations* Commun. Math. Phys. **188** (1997), 267–304, [hep-th/960408].

[KM] T.R. Klassen, E. Melzer, *The thermodynamics of purely elastic scattering theories and conformal perturbation theory*, Nucl. Phys. **B350** (1991) 635–689.

[KNS] A. Kuniba, T. Nakanishi, J. Suzuki, *T-systems and Y-systems in integrable systems*, J. Phys. A: Math. Theor. **44** (2011) 103001, [hep-th/1010.1344].

[Koc] J. Kock, *Frobenius algebras and 2D topological quantum field theories*, London Mathematical Society Student Texts **59**, Cambridge University Press (2003).

[Kon] L. Kong, *Conformal field theory and a new geometry*, in: eds. H. Sati, U. Schreiber, Mathematical Foundations of Quantum Field and Perturbative String Theory, Proceedings of Symposia in Pure Mathematics **83**, American Mathematical Society (2011), [1107.3649 [math.QA]].

[Koz] K.K. Kozlowski, *On condensation properties of Bethe roots associated with the XXZ chain*, 1508.05741 [math-ph].

[KP] A. Klümper, P.A. Pearce, *Conformal weights of RSOS lattice models and their fusion hierarchies*, Physica A **183** (1992) 304–350.

[Kr] M.A. Krasnosel'skii, *Topological Methods in the Theory of Nonlinear Integral Equations*, Pergamon Press (1964).

[KS] A. Kapustin, N. Saulina, *Surface operators in 3d Topological Field Theory and 2d Rational Conformal Field Theory*, in: eds. H. Sati, U. Schreiber, Mathematical Foundations of Quantum Field and Perturbative String Theory, Proceedings of Symposia in Pure Mathematics **83**, American Mathematical Society (2011), [hep-th/1012.0911].

[La] C.K. Lai, *Existence of solutions of integral equations in the thermodynamics of one-dimensional fermions with repulsive delta function potential*, J. Math. Phys. **24** (1983) 133–137.

[McL] S. Mac Lane, *Categories for the working mathematician*, 2nd edition, Springer (1998).

[Ma] M.J. Martins, *Exact resonance A-D-E S-matrices and their renormalization group trajectories*, Nucl. Phys. **B394** (1993) 339–355, [hep-th/9208011].

[MaRu] D. Manolopoulos, I. Runkel, *A Monoidal Category for Perturbed Defects in Conformal Field Theory*, Commun. Math. Phys. **295** (2010) 327–362, [hep-th/0904.1122].

[MMR] A. Morin-Duchesne, J. Rasmussen, P. Ruelle, *Integrability and conformal data of the dimer model*, J. Phys. A: Math. Theor. **49** 174002, [hep-th/1507.04193].

[MS1] G. Moore, N. Seiberg, *Classical and quantum conformal field theory*, Comm. Math. Phys. **123** (1989) 177–254.

[MS2] G. Moore, N. Seiberg, *Lectures on RCFT*, in: Spring School on Superstrings, Trieste (1989) 1–129.

[Mu] M. Mueger, *From Subfactors to Categories and Topology II. The quantum double of tensor categories and subfactors*, J. Pure Appl. Alg. **180** (2003) 159–219, [0111205 [math.CT]]

[NK] W. Nahm, S. Keegan, *Integrable deformations of CFTs and the discrete Hirota equations*, hep-th/0905.3776.

[Ni] A. Nigro, *On the integrable structure of the Ising model*, J. Stat. Mech. (2008) P01017, [hep-th/0710.4213].

[Pa1] B. Pareigis, *On Lie algebras in braided categories*, in: Quantum Groups and Quantum Spaces, Banach Center Publ. 40, Polish Acad. Sci. (1997) 139–158, [9612002 [q-alg]].

[Pa2] B. Pareigis, *Lectures on quantum groups and noncommutative geometry*, `www.mathematik.uni-muenchen.de/~pareigis/Vorlesungen/02SS/QGandNCG.pdf` (2012).

[PM] A.D. Polyanin, A.V. Manzhirov, *Handbook of integral equations*, 2nd edition, Chapman&Hall (2008).

[RTV] F. Ravanini, R. Tateo, A. Valleriani, *Dynkin TBA's*, Int. J. Mod. Phys. **A8** (1993) 1707–1728, [hep-th/9207040].

[RFjFS] I. Runkel, J. Fjelstad, J. Fuchs, C. Schweigert, *Topological and conformal field theories as Frobenius algebras*, Contemp. Math. **431** (2007) 225–248, [0512076 [math.CT]].

[Ru1] I. Runkel, *Perturbed Defects and T-systems in Conformal Field Theory*, J. Phys. **A41** (2008) 105401, [hep-th/0711.0102].

[Ru2] I. Runkel, *Non-local conserved charges from defects in perturbed conformal field theory*, J. Phys. A: Math. Theor. **43** (2010) 365206, [hep-th/1004.1909].

[Ru3] I. Runkel, *A braided monoidal category for free super-bosons*, J. Math. Phys. **55** (2014) 041702, [1209.5554 [math.QA]].

[Scha] P. Schauenburg, *The monoidal center construction and bimodules*, J. Pure Appl. Algebra **158** (2001) 325–346.

[Scho] M. Schottenloher, *A Mathematical Introduction to Conformal Field Theory*, Lecture Notes in Physics **759**, Springer (2008).

[Seg] G. Segal, *The definition of conformal field theory*, in: Topology, geometry and quantum field theory, ed. U. Tillmann, London Mathematical Society Lecture Note Series **308**, Cambridge University Press (2004).

[SS] E.M. Stein, R. Shakarchi, *Complex analysis*, Princeton Lectures in Analysis II, Princeton University Press (2003).

[Tak] M. Takahashi, *One-Dimensional Heisenberg Model at Finite Temperature*, Progr. Theor. Phys. **46** (1971).

[Te] J. Teschner, *On the spectrum of the Sinh-Gordon model in finite volume*, Nucl. Phys. **B779** (2008) 403–429, [hep-th/0702214].

[TW] C.A. Tracy, H. Widom, *Proofs of Two Conjectures Related to the Thermodynamic Bethe Ansatz*, Commun. Math. Phys. **179** (1996) 667–680, [solv-int/9509003].

[Tu] V. Turaev, *Quantum invariants of knots and 3-manifolds*, de Gruyter Studies in Mathematics **18** , Walter de Gruyter & Co. (1994).

[vTo] S.J. van Tongeren, *Introduction to the thermodynamic Bethe ansatz*, J. Phys. A: Math. Theor. **49** (2016) 323005, [hep-th/1606.02951].

[YY] C.N. Yang, C.P. Yang, *Thermodynamics of a one-dimensional system of bosons with repulsive delta-function interaction*, J. Math. Phys. **10** (1969) 1115–1122.

[Za1] A.B. Zamolodchikov, *Integrable Field Theory from Conformal Field Theory*, Adv. Stud. Pure Math. **19** (1989) 641.

[Za2] A.B. Zamolodchikov, *Thermodynamic Bethe Ansatz in Relativistic Models. Scaling Three State Potts and Lee-yang Models*, Nucl. Phys. **B342** (1990) 695–720.

[Za3] A.B. Zamolodchikov, *On the thermodynamical Bethe ansatz equations for reflectionless ADE scattering theories*, Phys. Lett. **B253** (1991) 391–394.

Summary

This thesis deals with mathematical questions that arise in the areas of integrable 2-dimensional quantum field theory and 2-dimensional conformal quantum field theory. Such quantum field theories are interesting due to their high degree of symmetry, which opens up the possibility of exact computations and sometimes even mathematically rigorous construction from first principles. Thus, they have been a focal point of mathematical physics for many years.

The thesis consists of two separate but related parts. Part I focuses on Y-systems, which are a type of finite difference equations that play an important role in many integrable quantum field theories, since their solutions encode the spectrum of conserved charges. We first prove a result which helps establish the precise relationship between Y-systems and non-linear so-called Thermodynamic Bethe Ansatz (TBA) integral equations. We then go on to prove existence and uniqueness for positive real solutions to a large class of TBA equations, and as a consequence also obtain an existence and uniqueness theorem for solutions to the corresponding Y-system. Moreover, results about the asymptotic behaviour of these solutions are obtained, and properties of more general solutions are studied.

In Part II, topological defects are used as a tool from which non-local conserved charges in conformal field theory can be constructed. To this end, the mathematical properties of perturbed topological defect operators in rational conformal field theory are analysed, extending the existing formalism and results to the case of twisted boundary conditions. Finally, a special class of twisted perturbed topological defect operators (T-operators) is studied in the conformal field theory which describes the compactified free boson. It is shown that they satisfy the truncated twisted T-system functional relations obtained by Bazhanov, Lukyanov and Zamolodchikov based on a different approach. As a result, the spectrum of the T-operators is encoded in a Y-system of the type studied in Part I. We conclude by a numerical investigation of some solutions for some low-lying states in the spectrum, exploiting in particular the continuum of twisted boundary conditions that exists in the free boson.

Zusammenfassung

Diese Arbeit behandelt mathematische Probleme, welche sich insbesondere auf dem Gebiet 2-dimensionaler integrabler Quantenfeldtheorien sowie 2-dimensionaler konformer Feldtheorien ergeben. Solche Quantenfeldtheorien sind interessant aufgrund ihres hohen Grades an Symmetrie, welcher die Chance auf exakte Berechnungen und manchmal selbst mathematisch rigorose Konstruktion in sich birgt. Sie stehen daher seit langem im Fokus der mathematischen Physik.

Die Arbeit besteht aus zwei getrennten, aber aneinander anknüpfenden Teilen. Teil I handelt von Y-systemen, einem Typus von Differenzengleichungen, denen eine wichtige Rolle in integrablen Quantenfeldtheorien zukommt, da ihre Lösungen das Spektrum erhaltener Ladungen beschreiben. Wir beweisen zuerst ein Resultat, welches uns dabei hilft, die präzise Beziehung zwischen Lösungen von Y-systemen und Lösungen von sogenannten nichtlinearen Thermodynamischen Bethe Ansatz (TBA) Integralgleichungen zu klären. Danach beweisen wir die Existenz und Eindeutigkeit einer positiven reellen Lösung für eine grosse Klasse von TBA Gleichungen, und als Folge davon gewinnen wir auch einen Existenz- und Eindeutigkeitssatz für Lösungen des entsprechenden Y-systems. Überdies werden Resultate zum asymptotischen Verhalten dieser Lösungen hergeleitet, und Eigenschaften allgemeinerer Lösungen werden untersucht.

Im Teil 2 werden topologische Defekte als Hilfsmittel benutzt, um nicht-lokale erhaltene Ladungen in konformen Feldtheorien zu konstruieren. Dazu werden die mathematischen Eigenschaften von gestörten topologischen Defektoperatoren in rationalen konformen Feldtheorien untersucht, wobei der existierende Formalismus sowie damit erzielte Resultate auf den Fall getwisteter Randbedingungen verallgemeinert werden. Schliesslich wird eine spezielle Klasse von getwisteten gestörten topologischen Defektoperatoren (T-Operatoren) in der konformen Feldtheorie des kompaktifizierten freien Bosons untersucht. Es wird gezeigt, dass sie die trunkierten getwisteten T-system-Funktionalrelationen erfüllen, welche Bazhanov, Lukyanov und Zamolodchikov durch einen anderen Zugang hergeleitet haben. Damit wird das Spektrum der T-operatoren durch ein Y-system vom in Teil I untersuchten Typ beschrieben. Die Arbeit schliesst mit einer numerischen Untersuchung der Lösungen für einige tief liegende Zustände im Spektrum, wobei insbesondere ausgenutzt wird, dass für das freie Boson ein Kontinuum an getwisteten Randbedingungen existiert.