



**HAL**  
open science

## Black holes and higher depth mock modular forms

Sergey Alexandrov, Boris Pioline

► **To cite this version:**

Sergey Alexandrov, Boris Pioline. Black holes and higher depth mock modular forms. *Communications in Mathematical Physics*, 2019, 374, pp.549-625. 10.1007/s00220-019-03609-y . hal-01852413v2

**HAL Id: hal-01852413**

**<https://hal.science/hal-01852413v2>**

Submitted on 28 Aug 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Black holes and higher depth mock modular forms

---

**Sergei Alexandrov<sup>1,2</sup> and Boris Pioline<sup>3</sup>**

<sup>1</sup> *Laboratoire Charles Coulomb (L2C), Université de Montpellier, CNRS, F-34095, Montpellier, France*

<sup>2</sup> *Department of High Energy and Elementary Particle Physics, Saint Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg 199034, Russia*

<sup>3</sup> *Laboratoire de Physique Théorique et Hautes Energies (LPTHE), UMR 7589 CNRS-Sorbonne Université, Campus Pierre et Marie Curie, 4 place Jussieu, F-75005 Paris, France*

e-mail: [sergey.alexandrov@umontpellier.fr](mailto:sergey.alexandrov@umontpellier.fr), [pioline@lpthe.jussieu.fr](mailto:pioline@lpthe.jussieu.fr)

**ABSTRACT:** By enforcing invariance under S-duality in type IIB string theory compactified on a Calabi-Yau threefold, we derive modular properties of the generating function of BPS degeneracies of D4-D2-D0 black holes in type IIA string theory compactified on the same space. Mathematically, these BPS degeneracies are the generalized Donaldson-Thomas invariants counting coherent sheaves with support on a divisor  $\mathcal{D}$ , at the large volume attractor point. For  $\mathcal{D}$  irreducible, this function is closely related to the elliptic genus of the superconformal field theory obtained by wrapping M5-brane on  $\mathcal{D}$  and is therefore known to be modular. Instead, when  $\mathcal{D}$  is the sum of  $n$  irreducible divisors  $\mathcal{D}_i$ , we show that the generating function acquires a modular anomaly. We characterize this anomaly for arbitrary  $n$  by providing an explicit expression for a non-holomorphic modular completion in terms of generalized error functions. As a result, the generating function turns out to be a (mixed) mock modular form of depth  $n - 1$ .

---

## Contents

<b>1. Introduction and summary</b>	<b>2</b>
<b>2. BPS indices and wall-crossing</b>	<b>7</b>
2.1 Wall-crossing and attractor flows	7
2.2 Partial tree index	10
2.3 D4-D2-D0 black holes and BPS indices	11
<b>3. D3-instantons and contact potential</b>	<b>14</b>
3.1 $\mathcal{M}_H$ and twistorial description of instantons	14
3.2 D3-instantons in the large volume limit	16
3.3 From the contact potential to the instanton generating function	19
<b>4. Theta series decomposition and modularity</b>	<b>21</b>
4.1 Factorisation	21
4.2 Modularity and Vignéras' equation	22
4.3 Discontinuities and the anomaly	24
<b>5. Modular completion</b>	<b>25</b>
5.1 Completion of the generating function	25
5.2 Generalized error functions and the completion	29
5.3 The structure of the completion	34
5.4 Alternative representations	35
<b>6. Discussion</b>	<b>39</b>
<b>A. Proof of the Theorem</b>	<b>41</b>
<b>B. D3-instanton contribution to the contact potential</b>	<b>42</b>
<b>C. Smoothness of the instanton generating function</b>	<b>43</b>
<b>D. Indefinite theta series and generalized error functions</b>	<b>45</b>
D.1 Vignéras' theorem	45
D.2 Generalized error functions	46
<b>E. Twistorial integrals and generalized error functions</b>	<b>48</b>
<b>F. Proofs of propositions</b>	<b>54</b>
<b>G. Explicit results up to 4th order</b>	<b>64</b>

# 1. Introduction and summary

The degeneracies of BPS black holes in string vacua with extended supersymmetry possess remarkable modular properties, which have been instrumental in recent progress on explaining the statistical origin of the Bekenstein-Hawking entropy in [1] and many subsequent works. Namely, the indices  $\Omega(\gamma)$  counting — with sign — microstates of BPS black holes with electromagnetic charge  $\gamma$  may often be collected into a suitable generating function which exhibits modular invariance, providing powerful constraints on its Fourier coefficients and enabling direct access to their asymptotic growth. When the black holes can be realized as black strings wrapped on a circle, a natural candidate for such a generating function is the elliptic genus of the superconformal field theory supported by the black string, which is modular invariant by construction [2, 3, 4]. Equivalently, one may consider the partition function of the effective three-dimensional gravity living on the near-horizon geometry of the black string [5].

In most cases however, the BPS indices depend not only on the charge  $\gamma$  but also on the moduli  $z^a$  at spatial infinity, due to the wall-crossing phenomenon: some of the BPS bound states with total charge  $\gamma$  only exist in a certain chamber in moduli space, and decay as the moduli are varied across ‘walls of marginal stability’ which delimit this chamber. At strong coupling where the black hole description is accurate, this phenomenon has a transparent interpretation in terms of the (dis)appearance of multi-centered black hole configurations, which can be used to derive a universal wall-crossing formula [6, 7, 8, 9].

In the case of four-dimensional string vacua with  $\mathcal{N} = 4$  supersymmetry, where the BPS index is sensitive only to single-centered 1/4-BPS black holes and to bound states of two 1/2-BPS black holes, the resulting moduli dependence is reflected in poles in the generating function, requiring a suitable choice of contour for extracting the Fourier coefficients in a given chamber [10, 11, 12]. Upon subtracting contributions of two-centered bound states, the generating function of single-centered indices is no longer modular in the usual sense but it transforms as a *mock* Jacobi form with specific ‘shadow’ — a property which is almost as constraining as standard modular invariance [13].

In four-dimensional string vacua with  $\mathcal{N} = 2$  supersymmetry, the situation is much more complicated, firstly due to the fact that the moduli space of scalars receives quantum corrections, and secondly due to BPS bound states potentially involving an arbitrary number of constituents, resulting in an extremely intricate pattern of walls of marginal stability. Thus, it does not seem plausible that a single generating function may capture the BPS indices  $\Omega(\gamma, z^a)$  — which are known in this context as generalized Donaldson-Thomas (DT) invariants — in all chambers. Nevertheless, modular invariance is still expected to constrain them. In particular, D4-D2-D0 black holes in type IIA string theory compactified on a generic compact Calabi-Yau (CY) threefold  $\mathfrak{Y}$  can be lifted to an M5-brane wrapped on a divisor  $\mathcal{D} \subset \mathfrak{Y}$  [2]. If the divisor  $\mathcal{D}$  labelled by the D4-brane charge  $p^a$  is *irreducible*, the indices  $\Omega(\gamma, z^a)$  are independent of the moduli of  $\mathfrak{Y}$ , at least in the limit where the volume of  $\mathfrak{Y}$  is scaled to infinity, and their generating function is known to be a holomorphic (vector valued) modular form of weight  $-\frac{1}{2}b_2(\mathfrak{Y}) - 1$  [4, 3, 6]. But if the divisor  $\mathcal{D}$  is a sum of  $n$  effective divisors  $\mathcal{D}_i$ , the indices  $\Omega(\gamma, z^a)$  do depend on the Kähler moduli  $z^a$ , even in the large volume limit. In general however, the black string SCFT is supposed to capture the states associated

to a single  $AdS_3$  throat, while for generic values of the moduli multiple  $AdS_3$  throats can usually contribute [14, 15]. It is thus natural to consider the modular properties of the DT invariants  $\Omega(\gamma, z^a)$  at the large volume attractor point<sup>1</sup>

$$z_\infty^a(\gamma) = \lim_{\lambda \rightarrow +\infty} (-q^a + i\lambda p^a), \quad (1.1)$$

where only a single  $AdS_3$  throat is allowed [16]. Following [17] we denote these invariants by  $\Omega^{\text{MSW}}(\gamma) = \Omega(\gamma, z_\infty^a(\gamma))$  and call them Maldacena-Strominger-Witten (MSW) invariants. As we discuss in Section 2, the DT invariants  $\Omega(\gamma, z^a)$  can be recovered from the MSW invariants  $\Omega^{\text{MSW}}(\gamma)$  by using a version of the split attractor flow conjecture [18, 6] developed in [19], which we call the flow tree formula.

The case where  $\mathcal{D}$  is the sum of two irreducible divisors was first studied in [16], and more recently in [20, 21], where it was shown that the generating function of MSW invariants is a mock modular form, with a specific non-holomorphic completion obtained by smoothing out the sign functions entering in the bound state contributions, recovering the prescription of [22, 16]. The goal of this paper is to extend this result to the general case where  $\mathcal{D}$  is the sum of  $n$  irreducible divisors  $\mathcal{D}_i$ , where  $n$  can be arbitrarily large. In such generic situation, we characterize the modular properties of the generating function  $h_{p,\mu}$  of MSW invariants and find an explicit expression for its non-holomorphic completion  $\widehat{h}_{p,\mu}$  in terms of the generating functions  $h_{p_i,\mu_i}$  associated to the  $n$  constituents, multiplied by certain iterated integrals introduced in [23, 24], which generalize the usual error function appearing when  $n = 2$ . This result implies that in this case  $h_{p,\mu}$  is a (mixed, vector valued) mock modular form of depth  $n - 1$ , in the sense that its shadow is itself a mock modular form of lower depth (with the depth 1 case reducing to the standard mock modular forms introduced in [22, 25], and the depth 0 case to usual weakly holomorphic modular forms).

In order to establish this result, we follow the same strategy as in our earlier works [20, 21] and analyse D3-D1-D(-1) instanton corrections to the metric on the hypermultiplet moduli space  $\mathcal{M}_H$  in type IIB string theory compactified on  $\mathfrak{Y}$ , at arbitrary order in the instanton expansion. After reducing on a circle and T-dualizing, this moduli space is identical to the vector multiplet moduli space in type IIA string theory compactified on  $\mathfrak{Y} \times S_1$ , where it receives instanton corrections from D4-D2-D0 black holes winding around the circle. In either case, each instanton contribution is weighted by the same generalized DT invariant  $\Omega(\gamma)$  which counts the number of BPS black hole microstates with electromagnetic charge  $\gamma$ . The modular properties of the generalized DT invariants are fixed by requiring that the quaternion-Kähler (QK) metric on  $\mathcal{M}_H$  admits an isometric action of  $SL(2, \mathbb{Z})$ , which comes from S-duality in type IIB, or equivalently from large diffeomorphisms of the torus appearing when viewing type IIA/ $\mathfrak{Y} \times S_1$  as M-theory on  $\mathfrak{Y} \times T^2$  [26]. This QK metric is most efficiently encoded in the complex contact structure on the associated twistor space, a  $\mathbb{C}P^1$ -bundle over  $\mathcal{M}_H$  [27, 28]. The latter is specified by a set of gluing conditions determined by the DT invariants [26, 29], which can in turn be expressed in terms of the MSW invariants using the tree flow formula.

---

<sup>1</sup>Here  $p^a$  and  $q_a$  are D4 and D2-brane charges, respectively, and the index of  $q$  is raised with help of the inverse of the metric  $\kappa_{ab} = \kappa_{abc}p^a$  with  $\kappa_{abc}$  being the triple intersection numbers of  $\mathfrak{Y}$ .

An important quantity appearing in this twistorial formulation<sup>2</sup> is the so called *contact potential*  $e^\phi$ , a real function on  $\mathcal{M}_H$  related to the Kähler potential on the twistor space, and afforded by the existence of a continuous isometry unbroken by D-instantons. On the type IIB side,  $e^\phi$  can be identified with the four-dimensional dilaton  $1/g_4^2$ . When expressed in terms of the ten-dimensional type IIB axio-dilaton  $\tau = c_0 + i/g_s$  (or in terms of the modulus of the 2-torus on the M-theory side), it becomes a complicated function having a classical contribution, a one-loop correction, and a series of instanton corrections expressed as contour integrals on the  $CP^1$  fiber. The importance of the contact potential stems from the fact that it must be a non-holomorphic modular form of weight  $(-\frac{1}{2}, -\frac{1}{2})$  in the variable  $\tau$  [26] in order for  $\mathcal{M}_H$  to admit an isometric action of  $SL(2, \mathbb{Z})$ . This requirement imposes very non-trivial constraints on the instanton contributions to  $e^\phi$ , which can be used to deduce the modular properties of generating functions of DT invariants, order by order in the instanton expansion. This strategy was used in [20] at two-instanton order to characterize the modular behavior of the generating function of MSW invariants in the case of a divisor equal to the sum of  $n = 2$  irreducible divisors. In this paper we generalize this result to all  $n$ , by analyzing the instanton expansion to all orders. Below we summarize the main steps in our analysis and our main results.

1. First, we show that the contact potential  $e^\phi$  in the large volume limit, where the Kähler parameters of the CY are sent to infinity, can be expressed (see (3.24)) through another (complex valued) function  $\mathcal{G}$  (3.22) on the moduli space, which we call the instanton generating function. The expansion of  $\mathcal{G}$  in powers of DT invariants is governed by a sum over unrooted trees decorated by charges  $\gamma_i$  (see (3.25) and (3.26)). The modularity of the contact potential requires  $\mathcal{G}$  to be a modular form of weight  $(-\frac{3}{2}, \frac{1}{2})$ .
2. After expressing the DT invariants  $\Omega(\gamma, z^a)$  through the moduli independent MSW invariants  $\Omega^{\text{MSW}}(\gamma)$  using the tree flow formula of [19], and expanding  $\mathcal{G}$  in powers of  $\Omega^{\text{MSW}}(\gamma)$ , each order in this expansion can be decomposed into a sum of products of certain indefinite theta series and of holomorphic generating functions  $h_{p_i, \mu_i}$  of the invariants  $\Omega^{\text{MSW}}(\gamma)$  (see (4.2)), similarly to the usual decomposition of standard Jacobi forms. Thus, the modular properties of the indefinite theta series  $\vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi_n^{\text{tot}}, n - 2)$  are tied with the modular properties of the generating functions  $h_{p_i, \mu_i}$ , in order for  $\mathcal{G}$  to be modular.
3. In order for an indefinite theta series  $\vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi, \lambda)$  to be modular, its kernel  $\Phi$  must satisfy a certain differential equation (D.3), which we call Vignéras' equation [32]. By construction, the kernels  $\Phi_n^{\text{tot}}$  appearing in our problem are given by iterated contour integrals along the  $CP^1$  fiber of the twistor space, multiplied by so-called 'tree indices'  $g_{\text{tr}, n}$  coming from the expression of  $\Omega(\gamma, z^a)$  in terms of  $\Omega^{\text{MSW}}(\gamma)$ . We evaluate the twistorial

---

<sup>2</sup>A familiarity with the twistorial formulation is not required for the present work. Here we use only two equations relevant for the twistorial description of D-instantons: the integral equation (3.2) for certain Darboux coordinates, which appears also in the study of four-dimensional  $N = 2$  gauge theory on a circle [30]; and the expression for the contact potential (3.20) in terms of these Darboux coordinates. Note that the contact potential was recently extended to the context of gauge theories and interpreted there as a supersymmetric index [31]. These two equations lead to (3.15) and (3.24), respectively, which can be taken as the starting point of our analysis.

integrals in terms of the generalized error functions introduced in [23, 24], and show that the resulting kernels satisfy Vignéras' equation away from certain loci where they have discontinuities. Furthermore, we prove that the discontinuities corresponding to walls of marginal stability cancel between the integrals and the tree indices. But there are additional discontinuities coming from certain moduli independent contributions to the tree index. They spoil Vignéras' equation at the multi-instanton level so that the theta series  $\vartheta_{\mathbf{p},\boldsymbol{\mu}}(\Phi_n^{\text{tot}}, n-2)$  are *not* modular. In turn, this implies that the holomorphic generating functions  $h_{p_i,\mu_i}$  are not modular either.

4. However, we show that one can complete  $h_{p,\boldsymbol{\mu}}$  into a non-holomorphic modular form  $\widehat{h}_{p,\boldsymbol{\mu}}(\tau)$ , by adding to it a series of corrections proportional to products of  $h_{p_i,\mu_i}$  with the same total D4-brane charge  $p = \sum p_i$  (see (5.1)). The non-holomorphic functions  $R_n$ , fixing the completion, are determined by the condition that the expansion of  $\mathcal{G}$  rewritten in terms of  $\widehat{h}_{p,\boldsymbol{\mu}}$  gives rise to a non-anomalous, modular theta series. Equivalently, one can work with functions  $\widehat{g}_n$  appearing in the expansion (5.2) of the holomorphic generating function of DT invariants  $h_{p,q}^{\text{DT}}$  in powers of the non-holomorphic functions  $\widehat{h}_{p_i,\mu_i}$ .
5. Imposing the conditions for modularity, we find that  $\widehat{g}_n$  can be represented in an iterative form (5.8), or more explicitly as a sum (5.32) over rooted trees with valency  $\geq 3$  at each vertex (known as Schröder trees), where  $g_n^{(0)}$  are certain locally polynomial functions defined in (5.26), while  $\mathcal{E}_n$  are smooth solutions of Vignéras' equation constructed in terms of generalized error functions. The functions  $R_n$  are similarly given by a sum over Schröder trees (5.33) in terms of exponentially decreasing and non-decreasing parts of  $\mathcal{E}_n$ . These equations represent the main technical result of this work.
6. For  $n \leq 5$  our general formulas can be drastically simplified. In particular, the main building blocks, the functions  $g_n^{(0)}$  and  $\mathcal{E}_n$ , can be written as a sum over a suitable subset of flow trees, as in (5.38). In addition, we show that  $g_n^{(0)}$  has a natural extension including the refinement parameter conjugate to the spin  $J_3$  of the black hole.

At the end of this lengthy analysis, we thus find a modular completion of the generating functions  $h_{p,\boldsymbol{\mu}}$  of MSW invariants for an arbitrary divisor, i.e. decomposable into a sum of *any* number of irreducible divisors. The result is expressed through sums of products of generalized error functions labelled by trees of various types. For the reader's convenience, in Fig. 1 we list various trees which appear in our construction, up to  $n = 4$ . Since generalized error functions are known to be related to iterated Eichler integrals [20, 33], which occur in the non-holomorphic completion of mock modular forms, we loosely refer to the generating functions of MSW invariants as higher depth mock modular forms, although we do not spell out the precise meaning of this notion.

An unexpected byproduct of our analysis is an interesting combinatorial identity, relating rooted trees to the binomial coefficients, which plays a role in our derivation of the modular completion. Since we are not aware of such an identity in the mathematical literature,<sup>3</sup> we state it here as a theorem which proof can be found in appendix A.

---

<sup>3</sup>We were informed by Karen Yeats that the special case  $m = n - 1$  appears in [34, (124)]. The denominator in (1.2) is sometimes known as the tree factorial  $T'!$ , see (5.19).

	$\mathbb{T}_n = \mathbb{T}_{n,0}$	$\mathbb{T}_{n-3,1}$	$\mathbb{T}_n^r$	$\mathbb{T}_n^{\text{af}}$	$\mathbb{T}_n^{\text{S}}$
$n = 2$					
$n = 3$					
$n = 4$	 				

**Figure 1:** Various types of trees arising in this work.  $\mathbb{T}_{n,m}$  denotes the set of unrooted trees with  $n$  vertices and  $m$  marks distributed between the vertices.  $\mathbb{T}_n \equiv \mathbb{T}_{n,0}$  comprises the trees without marks.  $\mathbb{T}_n^r$  is the set of rooted trees with  $n$  vertices.  $\mathbb{T}_n^{\text{af}}$  is the set of attractor flow trees with  $n$  leaves (of which we only draw the different topologies). And  $\mathbb{T}_n^{\text{S}}$  denotes the set of Schröder trees with  $n$  leaves. In addition, an important role is played by the sets  $\mathbb{T}_{n,m}^\ell$  and  $\mathbb{T}_n^\ell = \mathbb{T}_{n,0}^\ell$  of unrooted, labelled, marked trees which are obtained from  $\mathbb{T}_{n,m}$  and  $\mathbb{T}_n$  by assigning different labels and markings to the vertices in such way that a vertex with  $m_v$  marks is decorated by  $2m_v + 1$  labels.

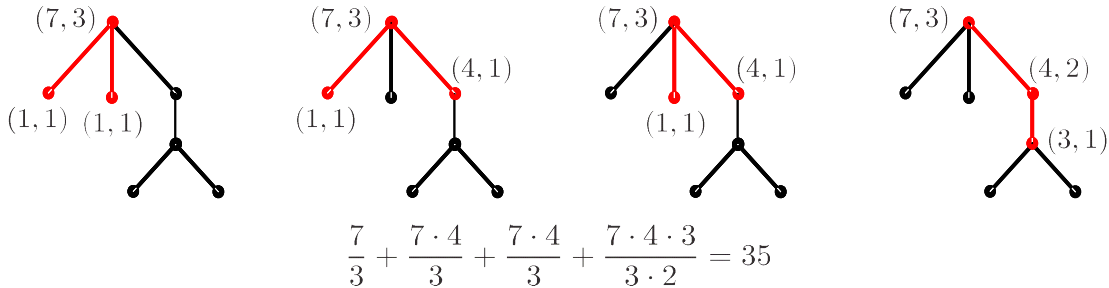
**Theorem 1.** *Let  $V_T$  be the set of vertices of a rooted ordered tree  $T$  and  $n_v(T)$  is the number of descendants of vertex  $v$  inside  $T$  plus 1 (alternatively, the number of vertices of the subtree in  $T$  with the root being the vertex  $v$  and the leaves being the leaves of  $T$ ). Then for a rooted tree  $T$  with  $n$  vertices and  $m < n$  one has*

$$\sum_{T' \subset T} \prod_{v \in V_{T'}} \frac{n_v(T)}{n_v(T')} = \frac{n!}{m!(n-m)!}, \quad (1.2)$$

where the sum goes over all subtrees with  $m$  vertices, having the same root as  $T$  (see Fig. 2).

The organization of the paper is as follows. In section 2 we review the known results about DT invariants, their expression in terms of MSW invariants, and specialize them to the case of D4-D2-D0 black holes in type IIA string theory on a Calabi-Yau threefold. In section 3 we present the twistorial description of the D-instanton corrected hypermultiplet moduli space in the dual type IIB string theory, evaluate the contact potential in the large volume approximation expressing it through a function  $\mathcal{G}$ , and provide an instanton expansion for this function via unrooted labelled trees. Then in section 4 we obtain a theta series decomposition for each order of the expansion of  $\mathcal{G}$  in MSW invariants and analyze the modular properties of the resulting theta series, which leads us to conclude that they possess a modular anomaly thereby implying a modular anomaly for the generating functions  $h_{p,\mu}$ . In section 5 we construct the non-holomorphic completion  $\widehat{h}_{p,\mu}$ , for which the anomaly is cancelled, and determine its explicit form. Section 6 is devoted to discussion of the obtained results. Finally, several appendices contain details of our calculations and proofs of various propositions. Besides, in the last appendix we provide explicit results for various functions appearing in our construction up to order  $n = 4$ .





**Figure 2:** An example illustrating the statement of the Theorem for a tree with  $n = 7$  vertices and subtrees with  $m = 3$  vertices. The subtrees are distinguished by red color. Near each vertex of the subtrees we indicated the pair of numbers  $(n_v(T), n_v(T'))$ .

## 2. BPS indices and wall-crossing

In this section, we review some aspects of BPS indices in theories with  $\mathcal{N} = 2$  supersymmetry, including the tree flow formula relating the moduli-dependent index  $\Omega(\gamma, z^a)$  to the attractor index  $\Omega_\star(\gamma)$ . We then apply this formalism in the context of Calabi-Yau string vacua, and express the generalized DT invariants  $\Omega(\gamma, z^a)$  in terms of their counterparts evaluated at the large volume attractor point (1.1), known as MSW invariants.

### 2.1 Wall-crossing and attractor flows

The BPS index  $\Omega(\gamma, z^a)$  counts (with sign) micro-states of BPS black holes with total electromagnetic charge  $\gamma = (p^\Lambda, q_\Lambda)$ , for a given value  $z^a$  of the moduli at spatial infinity. While  $\Omega(\gamma, z^a)$  is a locally constant function over the moduli space, it can jump across real codimension-one loci where certain bound states, represented by multi-centered black hole solutions of  $N = 2$  supergravity, become unstable. The positions of these loci, known as *walls of marginal stability*, are determined by the central charge  $Z_\gamma(z^a)$ , a complex-valued linear function of  $\gamma$  whose modulus gives the mass of a BPS state of charge  $\gamma$ , while the phase determines the supersymmetry subalgebra preserved by the state. Since a bound state can only decay when its mass becomes equal to the sum of masses of its constituents, it is apparent that the walls correspond to hypersurfaces where the phases of two central charges, say  $Z_{\gamma_L}(z^a)$  and  $Z_{\gamma_R}(z^a)$ , become aligned. The bound states which may become unstable are then those whose constituents have charges in the positive cone spanned by  $\gamma_L$  and  $\gamma_R$ . We shall assume that the charges  $\gamma_L, \gamma_R$  have non-zero Dirac-Schwinger-Zwanziger (DSZ) pairing  $\langle \gamma_L, \gamma_R \rangle \neq 0$ , since otherwise marginal bound states may form, whose stability is hard to control.

The general relation between the values of  $\Omega(\gamma, z^a)$  on the two sides of a wall has been found in the mathematics literature by Kontsevich–Soibelman [35] and Joyce–Song [36, 37], and justified physically in a series of works [6, 7, 8, 9]. However, in this work we require a somewhat different result: an expression of  $\Omega(\gamma, z^a)$  in terms of moduli-independent indices. One such representation, known as the Coulomb branch formula, was developed in a series of papers [38, 39, 40] (see [41] for a review) where the moduli-independent index is the so-called ‘single-centered invariant’ counting single-centered, spherically symmetric BPS black holes. Unfortunately, this representation (and its inverse) is quite involved, as it requires disentangling genuine single-centered solutions from so-called scaling solutions, i.e. multi-

centered solutions with  $n \geq 3$  constituents which can become arbitrarily close to each other [6, 42].

A simpler alternative is to consider the *attractor index*, i.e. the value of the BPS index in the attractor chamber  $\Omega_*(\gamma) \equiv \Omega(\gamma, z_*^a(\gamma))$ , where  $z_*^a(\gamma)$  is fixed in terms of the charge  $\gamma$  via the attractor mechanism [43] (recall that for a spherically symmetric BPS black hole with charge  $\gamma$ , the scalars in the vector multiplets have fixed value  $z_*^a(\gamma)$  at the horizon independently of their value  $z^a$  at spatial infinity.). By definition, the attractor indices are of course moduli independent. The problem of expressing  $\Omega(\gamma, z^a)$  in terms of the attractor indices was addressed recently in [19], extending earlier work in [16, 44]. Relying on the split attractor flow conjecture [18, 6], it was argued that the rational BPS index

$$\bar{\Omega}(\gamma, z^a) = \sum_{d|\gamma} \frac{1}{d^2} \Omega(\gamma/d, z^a) \quad (2.1)$$

can be expanded in powers of  $\bar{\Omega}_*(\gamma_i)$ ,

$$\bar{\Omega}(\gamma, z^a) = \sum_{\sum_{i=1}^n \gamma_i = \gamma} g_{\text{tr},n}(\{\gamma_i\}, z^a) \prod_{i=1}^n \bar{\Omega}_*(\gamma_i), \quad (2.2)$$

where the sum runs over ordered<sup>4</sup> decompositions of  $\gamma$  into sums of vectors  $\gamma_i \in \Gamma_+$ , with  $\Gamma_+$  being the set of all vectors  $\gamma$  whose central charge  $Z_\gamma(z_\infty^a)$  lies in a fixed half-space defining the splitting between BPS particles ( $\gamma \in \Gamma_+$ ) and anti-BPS particles ( $\gamma \in -\Gamma_+$ ). Such decompositions correspond to contributions of multi-centered black hole solutions with constituents carrying charges  $\gamma_i$ .

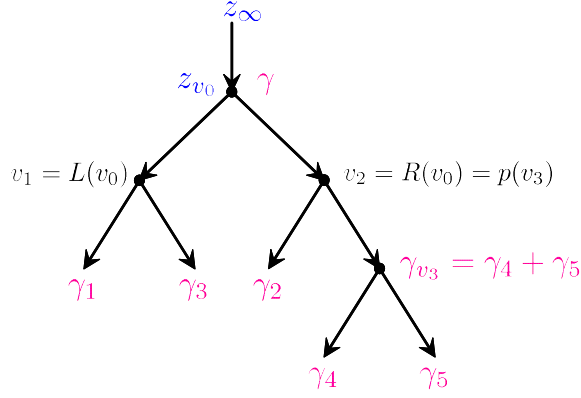
The coefficient  $g_{\text{tr},n}$ , called the *tree index*, is defined as

$$g_{\text{tr},n}(\{\gamma_i\}, z^a) = \frac{1}{n!} \sum_{T \in \mathbb{T}_n^{\text{af}}} \Delta(T) \kappa(T), \quad (2.3)$$

where the sum goes over *attractor flow trees*. These are (unordered, full) rooted binary trees<sup>5</sup>  $T$  with vertices decorated by electromagnetic charges  $\gamma_v$ , such that the leaves of the tree carry the constituent charges  $\gamma_i$ , and the charges propagate along the tree according to  $\gamma_v = \gamma_{L(v)} + \gamma_{R(v)}$  at each vertex, where  $L(v)$ ,  $R(v)$  are the two children of vertex  $v$ . The charge carried by the root of the tree is then the total charge  $\gamma = \sum \gamma_i$ . Note that when all the charges  $\gamma_i$  are distinct, flow trees are in one-to-one correspondence with 2-bracketings of the unordered set  $\{1, \dots, n\}$  (see Fig. 3). The idea of the split attractor flow conjecture is that each tree represents a nested sequence of two-centered bound states describing a multi-centered solution built out of constituents with charges  $\gamma_i$ . With this interpretation, the edges of the graph represent the evolution of the moduli under attractor flow, so that one starts from the moduli at spatial infinity  $z_\infty^a \equiv z^a$  and assigns to the root  $v_0$  the point in the moduli space  $z_{v_0}^a$  where the attractor flow with charge  $\gamma$  crosses the wall of marginal stability where

<sup>4</sup>In [19] this formula was written as a sum over *unordered* decompositions. Thus, the definition of  $g_{\text{tr},n}$  here differs from the one in [19] by the factor of  $1/n!$  which appears explicitly in (2.3).

<sup>5</sup>The number of such trees with  $n$  leaves is  $|\mathbb{T}_n^{\text{af}}| = (2n - 3)!! = (2n - 3)!/[2^{n-2}(n - 2)!] = \{1, 1, 3, 15, 105, 945, \dots\}$  for  $n \geq 1$ .



**Figure 3:** An example of attractor flow tree corresponding to the bracketing  $((13)(2(45)))$ .

$\text{Im} [Z_{\gamma_{L(v_0)}} \bar{Z}_{\gamma_{R(v_0)}}(z_{v_0}^a)] = 0$ . Then one repeats this procedure for every edge, obtaining a set of charges and moduli  $(\gamma_v, z_v^a)$  assigned to each vertex, with the bound state constituents and their attractor moduli  $(\gamma_i, z_{\gamma_i}^a)$  assigned to the leaves. Given these data, the factor  $\Delta(T)$  in (2.3) is given by

$$\Delta(T) = \prod_{v \in V_T} \Delta_{\gamma_{L(v)} \gamma_{R(v)}}^{z_{p(v)}}, \quad \Delta_{\gamma_L \gamma_R}^z = \frac{1}{2} \left[ \text{sgn} \text{Im} [Z_{\gamma_L} \bar{Z}_{\gamma_R}(z^a)] + \text{sgn}(\gamma_{LR}) \right], \quad (2.4)$$

where  $V_T$  denotes the set of vertices of  $T$  excluding the leaves,  $p(v)$  is the parent of vertex  $v$ , and  $\gamma_{LR} = \langle \gamma_L, \gamma_R \rangle$ . This factor vanishes unless the stability condition<sup>6</sup>

$$\gamma_{L(v)R(v)} \text{Im} [Z_{\gamma_{L(v)}} \bar{Z}_{\gamma_{R(v)}}(z_{p(v)}^a)] > 0 \quad (2.5)$$

is satisfied for all  $v \in V_T$ , which ensures admissibility of the flow tree  $T$ , i.e. the existence of the corresponding nested bound state. Importantly, the sign of  $\text{Im} [Z_{\gamma_{L(v)}} \bar{Z}_{\gamma_{R(v)}}(z_{p(v)}^a)]$  entering in (2.4) can be computed recursively in terms of the data at infinity, without evaluating the attractor flow along the edges [19]. The second factor in (2.3),

$$\kappa(T) \equiv (-1)^{n-1} \prod_{v \in V_T} \kappa(\gamma_{L(v)R(v)}), \quad \kappa(x) = (-1)^x x. \quad (2.6)$$

is simply the product of the BPS indices of the nested two-centered solutions associated with the tree  $T$ . Note that the signs of  $\Delta(T)$  and  $\kappa(T)$  are separately ambiguous, due to the choice of ordering  $(\gamma_{L(v)}, \gamma_{R(v)})$  at each vertex, but the sign of their product is unambiguous.

Sometimes, it is useful to consider the refined BPS index  $\Omega(\gamma, z^a, y)$  which carries additional dependence on the fugacity  $y$  conjugate to the spin  $J_3$ . All the above equations remain valid in this case as well (except for the definition of the rational invariant (2.1), which must be slightly modified, see e.g. [8, (1.3)]), but now the function  $\kappa(x)$  appearing in (2.6) becomes a Laurent polynomial in  $y$

$$\kappa(x) = (-1)^x \frac{y^x - y^{-x}}{y - y^{-1}}, \quad (2.7)$$

reducing to  $(-1)^x x$  in the unrefined limit  $y \rightarrow 1$ .

<sup>6</sup>In fact, the admissibility also requires  $\text{Re} [Z_{\gamma_{L(v)}} \bar{Z}_{\gamma_{R(v)}}(z_v^a)] > 0$  at each vertex. This condition will hold automatically for the case of our interest, namely, D4-D2-D0 black holes in the large volume limit, so we do not impose it explicitly.

It is easy to see that the ‘flow tree formula’ (2.2) is consistent with the primitive wall-crossing formula [6, 7]

$$\Delta\bar{\Omega}(\gamma_L + \gamma_R) = -\text{sgn}(\gamma_{LR}) \kappa(\gamma_{LR}) \bar{\Omega}(\gamma_L, z^a) \bar{\Omega}(\gamma_R, z^a), \quad (2.8)$$

which gives the jump of the BPS index due to the decay of bound states after crossing the wall defined by a pair of primitive<sup>7</sup> charges  $\gamma_L$  and  $\gamma_R$ . To this end, it is enough to consider all flow trees which start with the splitting  $\gamma \rightarrow \gamma_L + \gamma_R$  at the root of the tree. It is also useful to note that, assuming that all charges  $\gamma_i$  are distinct, the sum over splittings and flow trees in (2.2) can be generated by iterating the quadratic equation [19]

$$\bar{\Omega}(\gamma, z^a) = \bar{\Omega}_*(\gamma) - \frac{1}{2} \sum_{\substack{\gamma=\gamma_L+\gamma_R \\ \langle\gamma_L, \gamma_R\rangle \neq 0}} \Delta_{\gamma_L\gamma_R}^z \kappa(\gamma_{LR}) \bar{\Omega}(\gamma_L, z_{LR}^a) \bar{\Omega}(\gamma_R, z_{LR}^a), \quad (2.9)$$

where  $z_{LR}^a$  is the point where the attractor flow of charge  $\gamma$  crosses the wall of marginal stability  $\text{Im} [Z_{\gamma_L} \bar{Z}_{\gamma_R}(z_{LR}^a)] = 0$ .

## 2.2 Partial tree index

While the representation of the BPS index based on attractor flows is useful for many purposes, it produces a sum of products of sign functions depending on non-linear combinations of central charges and DSZ products, which are very difficult to work with. A solution to overcome this problem was found in [19]. The key idea is to introduce a refined index with a fugacity  $y$  conjugate to angular momentum, and represent it as

$$g_{\text{tr},n}(\{\gamma_i\}, z^a, y) = \frac{(-1)^{n-1+\sum_{i<j} \gamma_{ij}}}{(y - y^{-1})^{n-1}} \text{Sym} \left\{ F_{\text{tr},n}(\{\gamma_i\}, z^a) y^{\sum_{i<j} \gamma_{ij}} \right\}, \quad (2.10)$$

where  $\text{Sym}$  denotes symmetrization (with weight  $1/n!$ ) with respect to the charges  $\gamma_i$ , and  $F_{\text{tr},n}$  is the ‘partial tree index’ defined by

$$F_{\text{tr},n}(\{\gamma_i\}, z^a) \equiv \sum_{T \in \mathbb{T}_n^{\text{af-pl}}} \Delta(T), \quad (2.11)$$

where the sum runs over the set  $\mathbb{T}_n^{\text{af-pl}}$  of *planar* flow trees with  $n$  leaves<sup>8</sup> carrying ordered charges  $\gamma_1, \dots, \gamma_n$ . Although this is not manifest, the refined tree index (2.10) is regular at  $y = 1$ , and its value (computed e.g. using l’Hôpital rule) reduces to the tree index (2.3). The advantage of the representation (2.10) is that the partial tree index  $F_{\text{tr},n}$  does not involve the  $\kappa$ -factors (2.6) and is independent of the refinement parameter.

The partial index  $F_{\text{tr},n}$  satisfies two important recursions. To formulate them, let us introduce a few convenient notations:

$$c_i(z^a) = \text{Im} [Z_{\gamma_i} \bar{Z}_{\gamma}(z^a)], \quad S_k = \sum_{i=1}^k c_i(z^a), \quad \beta_{kl} = \sum_{i=1}^k \gamma_{il}, \quad \Gamma_{kl} = \sum_{i=1}^k \sum_{j=1}^{\ell} \gamma_{ij}. \quad (2.12)$$

<sup>7</sup>Here, by primitive we mean that all charges with non-zero index in the two-dimensional lattice spanned by  $\gamma_L$  and  $\gamma_R$  are linear combinations  $N_L\gamma_L + N_R\gamma_R$  with coefficients  $N_L, N_R$  of the same sign.

<sup>8</sup>The number of such trees is the  $n - 1$ -th Catalan number  $|\mathbb{T}_n^{\text{af-pl}}| = \frac{(2n-2)!}{n[(n-1)!]^2} = \{1, 1, 2, 5, 14, 42, 132, \dots\}$  for  $n \geq 1$ .

Note that due to  $\gamma = \sum_{i=1}^n \gamma_i$ , the parameters  $c_i$  satisfy  $\sum_{i=1}^n c_i = 0$ . In terms of these notations, the partial index satisfies the iterative equation [19, (2.59)],

$$F_{\text{tr},n}(\{\gamma_i\}, z^a) = \frac{1}{2} \sum_{\ell=1}^{n-1} (\text{sgn}(S_\ell) - \text{sgn}(\Gamma_{n\ell})) F_{\text{tr},\ell}(\{\gamma_i\}_{i=1}^\ell, z_\ell^a) F_{\text{tr},n-\ell}(\{\gamma_i\}_{i=\ell+1}^n, z_\ell^a), \quad (2.13)$$

where  $z_\ell^a$  is the value of the moduli where the attractor flow crosses the wall for the decay  $\gamma \rightarrow (\gamma_1 + \dots + \gamma_\ell, \gamma_{\ell+1} + \dots + \gamma_n)$  and it can be shown that for this value

$$c_i(z_\ell^a) = c_i^{(\ell)} \equiv c_i(z^a) - \frac{\beta_{ni}}{\Gamma_{n\ell}} S_\ell. \quad (2.14)$$

This iterative equation is in some sense a realization of the quadratic equation (2.9). A similar equation can be written for the tree index  $g_{\text{tr},n}$  itself, but it will involve the  $\kappa$ -factors and symmetrization.

According to [19, Prop. 2], the partial tree index satisfies another recursion

$$F_{\text{tr},n}(\{\gamma_i\}, z^a) = F_n^{(0)}(\{\gamma_i, c_i\}) - \sum_{\substack{n_1+\dots+n_m=n \\ n_k \geq 1, m < n}} F_{\text{tr},m}(\{\gamma'_k\}, z^a) \prod_{k=1}^m F_{n_k}^{(*)}(\gamma_{j_{k-1}+1}, \dots, \gamma_{j_k}), \quad (2.15)$$

where the sum runs over ordered partitions of  $n$ ,  $m$  is the number of parts, and for  $k = 1, \dots, m$  we defined

$$j_0 = 0, \quad j_k = n_1 + \dots + n_k, \quad \gamma'_k = \gamma_{j_{k-1}+1} + \dots + \gamma_{j_k}. \quad (2.16)$$

The two functions appearing in (2.15) are simply products of signs

$$F_n^{(0)}(\{\gamma_i, c_i\}) = \frac{1}{2^{n-1}} \prod_{i=1}^{n-1} \text{sgn}(S_i), \quad F_n^{(*)}(\{\gamma_i\}) = \frac{1}{2^{n-1}} \prod_{i=1}^{n-1} \text{sgn}(\Gamma_{ni}). \quad (2.17)$$

This new recursive relation allows to express the partial index in a way which does not involve sign functions depending on non-linear combinations of parameters, in contrast to the previous relation (2.13) where such sign functions arise due to the attractor flow given by (2.14).

### 2.3 D4-D2-D0 black holes and BPS indices

The results presented above are applicable in any theory with  $N = 2$  supersymmetry. Let us now specialize to the BPS black holes obtained as bound states of D4-D2-D0-branes in type IIA string theory compactified on a CY threefold  $\mathfrak{Y}$ . In this case the moduli  $z^a = b^a + it^a$  (with  $a = 1, \dots, b_2(\mathfrak{Y})$  running over a basis of  $H_2(\mathfrak{Y}, \mathbb{Z})$ ), where  $t^a$  are Kähler moduli and  $b^a$  are components of the Kalb-Ramond field, parametrize the complexified Kähler moduli space  $\mathcal{M}_\mathcal{K}$  of  $\mathfrak{Y}$ . The charge vectors in a fixed basis of  $H_0 \oplus H_2 \oplus H_4 \oplus H_6$  have the form  $\gamma = (0, p^a, q_a, q_0)$  where zero corresponds to the vanishing D6-brane charge, whereas other components satisfy the following quantization conditions [45]

$$p^a \in \mathbb{Z}, \quad q_a \in \mathbb{Z} + \frac{1}{2} (p^2)_a, \quad q_0 \in \mathbb{Z} - \frac{1}{24} p^a c_{2,a}. \quad (2.18)$$

Here  $c_{2,a}$  are components of the second Chern class of  $\mathfrak{Y}$ , and to write the second relation we introduced notations  $(kp)_a = \kappa_{abc} k^b p^c$  and  $(lkp) = \kappa_{abc} l^a k^b p^c$  (recall that  $\kappa_{abc}$  are intersection

numbers of  $\mathfrak{Y}$ ) which will be extensively used below. The lattice of charges satisfying (2.18) will be denoted by  $\Gamma$ . A smaller lattice  $\Gamma_+$  is obtained by further restriction that the D4-brane charge  $p^a$  corresponds to an effective divisor in  $\mathfrak{Y}$  and belongs to the Kähler cone, i.e.

$$p^3 > 0, \quad (rp^2) > 0, \quad k_a p^a > 0, \quad (2.19)$$

for all effective divisors  $r^a \gamma_a \in H_4^+(\mathfrak{Y}, \mathbb{Z})$  and effective curves  $k_a \gamma^a \in H_2^+(\mathfrak{Y}, \mathbb{Z})$ , where  $\gamma_a$  denotes irreducible divisors giving an integer basis of  $\Lambda = H_4(\mathfrak{Y}, \mathbb{Z})$ , whereas  $\gamma^a$  an integer basis of  $\Lambda^* = H_2(\mathfrak{Y}, \mathbb{Z})$ . Such charge  $p^a$  gives rise to the quadratic form  $\kappa_{ab} = \kappa_{abc} p^c$  on  $\Lambda \otimes \mathbb{R} \simeq \mathbb{R}^{b_2}$  of signature  $(1, b_2 - 1)$ . It allows to embed  $\Lambda$  into  $\Lambda^*$ , but since in general the map  $\epsilon^a \mapsto \kappa_{ab} \epsilon^b$  is not surjective, the quotient  $\Lambda^*/\Lambda$  is a finite group of order  $|\det \kappa_{ab}|$ .

The holomorphic central charge for this system is given by

$$Z_\gamma(z^a) = q_\Lambda X^\Lambda(z^a) - p^\Lambda F_\Lambda(z^a), \quad (2.20)$$

where we used the special coordinates  $X^\Lambda(z^a) = (1, z^a)$  and  $F_\Lambda = \partial_{X^\Lambda} F(X)$  is the derivative of the holomorphic prepotential  $F(X)$  on  $\mathcal{M}_\mathcal{K}$ . In the large volume limit  $t^a \rightarrow \infty$ , the prepotential reduces to the classical cubic contribution

$$F(X) \approx F^{\text{cl}}(X) = -\kappa_{abc} \frac{X^a X^b X^c}{6X^0}, \quad (2.21)$$

and the central charge can be approximated as

$$Z_\gamma \approx -\frac{1}{2}(pt^2) + i(q_a t^a + (pb)t) + q_0 + q_a b^a + \frac{1}{2}(pb^2). \quad (2.22)$$

Note, in particular, that it always has a large negative real part. Another useful observation is that both quantities appearing in the definition of  $\Delta_{\gamma_L \gamma_R}^z$  (2.4) are independent of the  $q_0$  component of the charge. Indeed,

$$\begin{aligned} \langle \gamma, \gamma' \rangle &= q_a p'^a - q'_a p^a, \\ \text{Im} [Z_\gamma \bar{Z}_{\gamma'}] &= -\frac{1}{2} \left( (p't^2)(q_a + (pb)_a)t^a - (pt^2)(q'_a + (p'b)_a)t^a \right). \end{aligned} \quad (2.23)$$

The BPS index  $\bar{\Omega}(\gamma, z^a)$  counting D4-D2-D0 black holes is given mathematically by the generalized Donaldson-Thomas invariant, which counts semi-stable coherent sheaves supported on a divisor  $\mathcal{D}$  in the homology class  $p^a \gamma_a$ , with first and second Chern numbers  $(q_a, q_0)$ . An important property of these DT invariants is that they are invariant under the combination of the shift of the Kalb-Ramond field,  $b^a \mapsto b^a + \epsilon^a$ , and the spectral flow transformation acting on the D2 and D0 charges

$$q_a \mapsto q_a - \kappa_{abc} p^b \epsilon^c, \quad q_0 \mapsto q_0 - \epsilon^a q_a + \frac{1}{2}(p\epsilon\epsilon). \quad (2.24)$$

The shift of  $b^a$  cannot be ignored since the DT invariants are only piecewise constant as functions of the complexified Kähler moduli  $z^a = b^a + it^a$  due to wall-crossing.

On the other hand, the MSW invariants  $\bar{\Omega}^{\text{MSW}}(\gamma)$ , defined as the generalized DT invariants  $\bar{\Omega}(\gamma, z^a)$  evaluated at their respective large volume attractor point (1.1), are by construction independent of the moduli, and therefore invariant under the spectral flow (2.24). As

a result, they only depend on  $p^a, \mu_a$  and  $\hat{q}_0$ , where we traded the electric charges  $(q_a, q_0)$  for  $(\epsilon^a, \mu_a, \hat{q}_0)$ . The latter comprise the spectral flow parameter  $\epsilon^a$ , the residue class  $\mu_a \in \Lambda^*/\Lambda$  defined by the decomposition

$$q_a = \mu_a + \frac{1}{2} \kappa_{abc} p^b p^c + \kappa_{abc} p^b \epsilon^c, \quad \epsilon^a \in \Lambda, \quad (2.25)$$

and the invariant charge ( $\kappa^{ab}$  is the inverse of  $\kappa_{ab}$ )

$$\hat{q}_0 \equiv q_0 - \frac{1}{2} \kappa^{ab} q_a q_b, \quad (2.26)$$

which is invariant under (2.24). This allows to write  $\bar{\Omega}^{\text{MSW}}(\gamma) = \bar{\Omega}_{p,\mu}(\hat{q}_0)$ .

An important fact is that the invariant charge  $\hat{q}_0$  is bounded from above by  $\hat{q}_0^{\text{max}} = \frac{1}{24}((p^3) + c_{2,a} p^a)$ . This allows to define two generating functions

$$h_{p,q}^{\text{DT}}(\tau, z^a) = \sum_{\hat{q}_0 \leq \hat{q}_0^{\text{max}}} \bar{\Omega}(\gamma, z^a) \mathbf{e}(-\hat{q}_0 \tau), \quad (2.27)$$

$$h_{p,\mu}(\tau) = \sum_{\hat{q}_0 \leq \hat{q}_0^{\text{max}}} \bar{\Omega}_{p,\mu}(\hat{q}_0) \mathbf{e}(-\hat{q}_0 \tau), \quad (2.28)$$

where we used notation  $\mathbf{e}(x) = e^{2\pi i x}$ . Whereas the generating function of DT invariants  $h_{p,q}^{\text{DT}}$  depends on the full electric charge  $q_a$  and also carries a piecewise constant dependence on the moduli  $z^a$ , the generating function of MSW invariants  $h_{p,\mu}(\tau)$ , due to the spectral flow symmetry, depends only on the residue class  $\mu_a$ . This generating function will be the central object of interest in this paper, and our main goal will be to understand its behavior under modular transformations of  $\tau$ .

In general, the MSW invariants  $\bar{\Omega}^{\text{MSW}}(\gamma)$  are distinct from the attractor moduli  $\bar{\Omega}_*(\gamma)$ , since the latter coincide with the generalized DT invariants  $\bar{\Omega}(\gamma, z^a)$  evaluated at the true attractor point  $z_*^a(\gamma)$  for the charge  $\gamma$ , while the former are the generalized DT invariants evaluated at the large volume attractor point  $z_\infty^a(\gamma)$  defined in (1.1). Nevertheless, we claim that in the large volume limit  $t^a \rightarrow \infty$ , the tree flow formula reviewed in the previous subsections still allows to express  $\bar{\Omega}(\gamma, z^a)$  in terms of the MSW invariants, namely

$$\bar{\Omega}(\gamma, z^a) = \sum_{\sum_{i=1}^n \gamma_i = \gamma} g_{\text{tr},n}(\{\gamma_i\}, z^a) \prod_{i=1}^n \bar{\Omega}^{\text{MSW}}(\gamma_i) \quad (\text{Im } z^a \rightarrow \infty). \quad (2.29)$$

The point is that the only walls of marginal stability which extend to infinite volume are those where the constituents carry only D4-brane charge, and that non-trivial bound states involving constituents with D4-brane charge are ruled out at the large volume attractor point, similarly to the usual attractor chamber. Since the r.h.s. of (2.29) is consistent with wall-crossing in the infinite volume limit and agrees with the left-hand side at  $z^a = z_\infty^a(\gamma)$ , it must therefore hold everywhere at large volume. Of course, some of the states contributing to  $\bar{\Omega}^{\text{MSW}}(\gamma_i)$  may have some substructure, e.g. be realized as D6- $\overline{\text{D6}}$  bound states, but this structure cannot be probed in the large volume limit. Importantly, since the quantities (2.23) entering in the definition of the tree index are independent of the D0-brane charge  $q_0$ , the flow tree formula (2.29) may be rewritten as a relation between the generating functions,

$$h_{p,q}^{\text{DT}}(\tau, z^a) = \sum_{\sum_{i=1}^n \check{\gamma}_i = \check{\gamma}} g_{\text{tr},n}(\{\check{\gamma}_i\}, z^a) e^{\pi i \tau Q_n(\{\check{\gamma}_i\})} \prod_{i=1}^n h_{p_i, \mu_i}(\tau), \quad (2.30)$$

where  $\tilde{\gamma} = (p^a, q_a)$  denotes the projection of the charge vector on  $H_4 \oplus H_2$ , and

$$Q_n(\{\tilde{\gamma}_i\}) = \kappa^{ab} q_a q_b - \sum_{i=1}^n \kappa_i^{ab} q_{i,a} q_{i,b} \quad (2.31)$$

appears due to the quadratic term in the definition (2.26) of the invariant charge  $\hat{q}_0$ .

### 3. D3-instantons and contact potential

In this section, we switch to the dual setup<sup>9</sup> of type IIB string theory compactified on the same CY manifold  $\mathfrak{Y}$ . The DT invariants, describing the BPS degeneracies of D4-D2-D0 black holes in type IIA, now appear as coefficients in front of the D3-D1-D(-1) instanton effects affecting the metric on the hypermultiplet moduli space  $\mathcal{M}_H$ . The main idea of our approach is that these instanton effects are strongly constrained by demanding that  $\mathcal{M}_H$  admits an isometric action of the type IIB S-duality  $SL(2, \mathbb{Z})$ . This constraint uniquely fixes the modular behavior of the generating functions introduced in the previous section. Here we recall the twistorial construction of D-instanton corrections to the hypermultiplet metric, describe the action of S-duality, and analyze the instanton expansion of a particular function on  $\mathcal{M}_H$  known as contact potential.

#### 3.1 $\mathcal{M}_H$ and twistorial description of instantons

The moduli space of four-dimensional  $N = 2$  supergravity is a direct product of vector and hypermultiplet moduli spaces,  $\mathcal{M}_V \times \mathcal{M}_H$ . The former is a (projective) special Kähler manifold, whereas the latter is a quaternion-Kähler (QK) manifold. In type IIB string theory compactified on a CY threefold  $\mathfrak{Y}$ ,  $\mathcal{M}_H$  is a space of real dimension  $4b_2(\mathfrak{Y}) + 4$ , which is fibered over the complexified Kähler moduli space  $\mathcal{M}_{\mathcal{K}}(\mathfrak{Y})$  of dimension  $2b_2(\mathfrak{Y})$ . In addition to the Kähler moduli  $z^a = b^a + it^a$ , it describes the dynamics of the ten-dimensional axio-dilaton  $\tau = c^0 + i/g_s$ , the Ramond-Ramond (RR) scalars  $c^a, \tilde{c}_a, \tilde{c}_0$ , corresponding to periods of the RR 2-form, 4-form and 6-form on a basis of  $H^{\text{even}}(\mathfrak{Y}, \mathbb{Z})$ , and finally, the NS-axion  $\psi$ , dual to the Kalb-Ramond two-form  $B$  in four dimensions.

At tree-level, the QK metric on  $\mathcal{M}_H$  is obtained from the moduli space  $\mathcal{M}_{\mathcal{K}}$  via the  $c$ -map construction [46, 47] and thus is completely determined by the holomorphic prepotential  $F(X)$ . But this metric receives  $g_s$ -corrections, both perturbative and non-perturbative. The latter can be of two types: either from Euclidean D-branes wrapping even dimensional cycles on  $\mathfrak{Y}$ , or from NS5-branes wrapped around the whole  $\mathfrak{Y}$ . In this paper we shall be interested only in the effects of D3-D1-D(-1) instantons, and ignore the effects of NS5 and D5-instantons, which are subleading in the large volume limit. Since NS5-instantons only mix with D5-instantons under S-duality, this truncation does not spoil modular invariance.

The most concise way to describe the D-instanton corrections is to consider type IIA string theory compactified on the mirror CY and use the twistor formalism for quaternionic

---

<sup>9</sup>Our preference for the type IIB set-up is merely for consistency with our earlier works on hypermultiplet moduli spaces in  $d = 4$  Calabi-Yau vacua. The same considerations apply verbatim, with minor changes of wording, to the vector multiplet moduli space in type IIA string theory compactified on  $\mathfrak{Y} \times S^1$ , which is more directly related to the counting of D4-D2-D0 black holes in four dimensions.



geometries [27, 28]. In this approach the metric is encoded in the complex contact structure on the twistor space, a  $\mathbb{C}P^1$ -bundle over  $\mathcal{M}_H$ . The D-instanton corrected contact structure has been constructed to all orders in the instanton expansion in [26, 29], and an explicit expression for the metric has been derived recently in [48, 49]. Here we will present only those elements of the construction which are relevant for the subsequent analysis, and refer to reviews [50, 51] for more details.

The crucial point is that, locally, the contact structure is determined by a set of holomorphic Darboux coordinates  $(\xi^\Lambda, \tilde{\xi}_\Lambda, \alpha)$  on the twistor space, considered as functions of coordinates on  $\mathcal{M}_H$  and of the stereographic coordinate  $t$  on the  $\mathbb{C}P^1$  fiber, so that the contact one-form takes the canonical form  $d\alpha + \tilde{\xi}_\Lambda d\xi^\Lambda$ . Although all Darboux coordinates are important for recovering the metric, for the purposes of this paper the coordinate  $\alpha$  is irrelevant. Therefore, we consider only  $\xi^\Lambda$  and  $\tilde{\xi}_\Lambda$  which can be conveniently packaged into holomorphic Fourier modes  $\mathcal{X}_\gamma = \mathbf{e}\left(p^\Lambda \tilde{\xi}_\Lambda - q_\Lambda \xi^\Lambda\right)$  labelled by a charge vector  $\gamma = (p^\Lambda, q_\Lambda)$ .

At tree level, the Darboux coordinates (multiplied by  $t$ ) are known to be simple quadratic polynomials in  $t$  so that  $\mathcal{X}_\gamma$  take the form<sup>10</sup>

$$\mathcal{X}_\gamma^{\text{sf}}(t) = \mathbf{e}\left(\frac{\tau_2}{2} \left(\bar{Z}_\gamma(\bar{u}^a) t - \frac{Z_\gamma(u^a)}{t}\right) + p^\Lambda \tilde{\xi}_\Lambda - q_\Lambda \xi^\Lambda\right), \quad (3.1)$$

where  $Z_\gamma(u^a)$  is the central charge (2.20), now expressed in terms of the complex structure moduli  $u^a$  of the mirror CY threefold,  $\zeta^\Lambda$  and  $\tilde{\zeta}_\Lambda$  are periods of the RR 3-form of the type IIA formulation, and  $\tau_2 = g_s^{-1}$  is the ten-dimensional string coupling. At non-perturbative level, this expression gets modified and the Darboux coordinates are determined by the integral equation

$$\mathcal{X}_\gamma(t) = \mathcal{X}_\gamma^{\text{sf}}(t) \mathbf{e}\left(\frac{1}{8\pi^2} \sum_{\gamma'} \sigma_{\gamma'} \bar{\Omega}(\gamma') \langle \gamma, \gamma' \rangle \int_{\ell_{\gamma'}} \frac{dt'}{t'} \frac{t+t'}{t-t'} \mathcal{X}_{\gamma'}(t')\right), \quad (3.2)$$

where the sum goes over all charges labelling cycles wrapped by D-branes,  $\bar{\Omega}(\gamma') = \bar{\Omega}(\gamma', z^a)$  is the corresponding rational Donaldson-Thomas invariant,

$$\ell_\gamma = \{t \in \mathbb{C}P^1 : Z_\gamma/t \in i\mathbb{R}^-\} \quad (3.3)$$

is the so called BPS ray, a contour on  $\mathbb{C}P^1$  extending from  $t = 0$  to  $t = \infty$  along the direction fixed by the central charge, and  $\sigma_\gamma$  is a quadratic refinement of the DSZ product on the charge lattice  $\Gamma$ , i.e. a sign factor satisfying the defining relation

$$\sigma_{\gamma_1} \sigma_{\gamma_2} = (-1)^{\langle \gamma_1, \gamma_2 \rangle} \sigma_{\gamma_1 + \gamma_2}, \quad \forall \gamma_1, \gamma_2 \in \Gamma. \quad (3.4)$$

The system of integral equations (3.2) can be solved iteratively by first substituting  $\mathcal{X}_{\gamma'}(t')$  on the r.h.s. with its zero-th order value  $\mathcal{X}_{\gamma'}^{\text{sf}}(t')$  in the weak coupling limit  $\tau_2 \rightarrow \infty$ , computing the leading correction from the integral and iterating this process. This produces an asymptotic series at weak coupling, in powers of the DT invariants  $\bar{\Omega}(\gamma)$ . Using the saddle point method, it is easy to check that the coefficient of each monomial  $\prod_i \bar{\Omega}(\gamma_i)$  is suppressed by a factor  $e^{-\pi\tau_2 \sum_i |Z_{\gamma_i}|}$ , corresponding to an  $n$ -instanton effect [30, 52]. Note that multi-instanton effects become of the same order as one-instanton effects on walls of marginal stability where the phases of  $Z_{\gamma_i}$  become aligned.

<sup>10</sup>The superscript ‘sf’ stands for ‘semi-flat’, which refers to the flatness of the classical geometry in the directions along the torus fibers parametrized by  $\zeta^\Lambda, \tilde{\zeta}_\Lambda$ .

### 3.2 D3-instantons in the large volume limit

The above construction of D-instantons is adapted to the type IIA formulation because the equation (3.2) defines the Darboux coordinates in terms of the type IIA fields appearing explicitly in the tree level expression (3.1). To pass to the mirror dual type IIB formulation, one should apply the mirror map, a coordinate transformation from the type IIA to the type IIB physical fields. This transformation was determined in the classical limit in [53], but it also receives instanton corrections. In order to fix the form of these corrections, we require that the metric on  $\mathcal{M}_H$  carries an isometric action of S-duality group  $SL(2, \mathbb{Z})$  of type IIB string theory, which acts on the type IIB fields by an element  $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  in the following way

$$\begin{aligned} \tau &\mapsto \frac{a\tau + b}{c\tau + d}, & t^a &\mapsto |c\tau + d| t^a, & \begin{pmatrix} c^a \\ b^a \end{pmatrix} &\mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} c^a \\ b^a \end{pmatrix}, \\ \tilde{c}_a &\mapsto \tilde{c}_a - c_{2,a}\varepsilon(g), & \begin{pmatrix} \tilde{c}_0 \\ \psi \end{pmatrix} &\mapsto \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} \tilde{c}_0 \\ \psi \end{pmatrix}, \end{aligned} \quad (3.5)$$

where  $\varepsilon(g)$  is the logarithm of the multiplier system of the Dedekind eta function [45].

For this purpose, one uses the fact that any isometric action on a quaternion-Kähler manifold (preserving the quaternionic structure) can be lifted to a holomorphic contact transformation on twistor space. In the present case,  $SL(2, \mathbb{Z})$  acts on the fiber coordinate  $t$  by a fractional-linear transformation with  $\tau$ -dependent coefficients. This transformation takes a much simpler form when formulated in terms of another coordinate  $z$  on  $\mathbb{C}P^1$  (not to be confused with the Kähler moduli  $z^a$ ), which is related to  $t$  by a Cayley transformation,

$$z = \frac{t + i}{t - i}. \quad (3.6)$$

Then the action of  $SL(2, \mathbb{Z})$  on the fiber is given by a simple phase rotation<sup>11</sup>

$$z \mapsto \frac{c\bar{\tau} + d}{|c\tau + d|} z. \quad (3.7)$$

Using the non-trivial constraint of holomorphicity for the  $SL(2, \mathbb{Z})$  isometry on the twistor space, quantum corrections to the classical mirror map were computed in [55, 56, 17, 21], in the large volume limit where the Kähler moduli are taken to be large,  $t^a \rightarrow \infty$ . In this limit, one finds

$$\begin{aligned} u^a &= b^a + it^a - \frac{i}{2\tau_2} \sum_{\gamma \in \Gamma_+} p^a \left[ \int_{\ell_\gamma} dz (1-z) H_\gamma + \int_{\ell_{-\gamma}} \frac{dz}{z^3} (1-z) H_{-\gamma} \right] \\ \zeta^0 &= \tau_1, & \zeta^a &= -(c^a - \tau_1 b^a) - 3 \sum_{\gamma \in \Gamma_+} p^a \operatorname{Re} \int_{\ell_\gamma} dz z H_\gamma, \\ \tilde{\zeta}_a &= \tilde{c}_a + \frac{1}{2} \kappa_{abc} b^b (c^c - \tau_1 b^c) + \kappa_{abc} t^b \sum_{\gamma \in \Gamma_+} p^c \operatorname{Im} \int_{\ell_\gamma} dz H_\gamma, \end{aligned} \quad (3.8)$$

---

<sup>11</sup>Actually, this is true only when five-brane instanton corrections are ignored. Otherwise, the lift also gets a non-trivial deformation [54].

where we introduced a convenient notation<sup>12</sup>

$$H_\gamma(t) = \frac{\bar{\Omega}(\gamma)}{(2\pi)^2} \sigma_\gamma \mathcal{X}_\gamma(t). \quad (3.9)$$

Similar results are known for  $\tilde{\zeta}_0$  and the NS-axion dual to the  $B$ -field, but we will not need them in this paper.

Note that the integral contributions to the mirror map are written in terms of the coordinate  $z$  (3.6). The reason for using this variable is that, in the large volume limit, the integrals along BPS rays  $\ell_\gamma$  in (3.2) are dominated by the saddle point [17]

$$z'_\gamma \approx -i \frac{(q_a + (pb)_a) t^a}{(pt^2)}, \quad (3.10)$$

for  $(pt^2) > 0$  and  $z'_{-\gamma} = 1/z'_\gamma$  in the opposite case. This shows that all integrands can be expanded in Fourier series either around  $z = 0$  or  $z = \infty$ , keeping constant  $t^a z$  or  $t^a/z$ , respectively. This allows to extract the leading order in the large volume limit in a simple way.

Let us therefore evaluate the combined limit  $t^a \rightarrow \infty$ ,  $z \rightarrow 0$  of the system of integral equations (3.2), assuming that only D3-D1-D(-1) instantons contribute. As a first step, we rewrite the tree level expression (3.1) in terms of the type IIB fields. To this end, we restrict the charge  $\gamma$  to the lattice  $\Gamma_+$ , take the central charge as in (2.20) with the cubic<sup>13</sup> prepotential (2.21), and substitute the mirror map (3.8). Furthermore, we change the coordinate  $t$  to  $z$  and take the combined limit. In this way one finds

$$\mathcal{X}_\gamma^{\text{sf}}(z) = \mathcal{X}_\gamma^{\text{cl}}(z) \exp \left[ 2\pi \sum_{\gamma' \in \Gamma_+} (tpp') \int_{\ell_{\gamma'}} dz' H_{\gamma'} \right], \quad (3.11)$$

where

$$\mathcal{X}_\gamma^{\text{cl}}(z) = \mathbf{e}(-\hat{q}_0 \tau) \mathcal{X}_{p,q}^{(\theta)}(z) \quad (3.12)$$

is the classical part of the Darboux coordinates which we represented as a product of two factors: exponential of the invariant charge (2.26) and the remaining  $q_0$ -independent exponential

$$\mathcal{X}_{p,q}^{(\theta)}(z) = e^{-S_p^{\text{cl}}} \mathbf{e} \left( -\frac{\tau}{2} (q+b)^2 + c^a (q_a + \frac{1}{2}(pb)_a) + i\tau_2 (pt^2) (z^2 - 2zz_\gamma) \right), \quad (3.13)$$

with  $S_p^{\text{cl}}$  being the leading part of the Euclidean D3-brane action in the large volume limit given by  $S_p^{\text{cl}} = \pi\tau_2(pt^2) - 2\pi ip^a \tilde{c}_a$ . Next, we can approximate

$$\frac{dt'}{t'} \frac{t+t'}{t-t'} = \frac{2dz'}{1-z'^2} \frac{1-zz'}{z-z'} \approx \begin{cases} \frac{2dz'}{z-z'}, & \gamma' \in \Gamma_+, \\ \frac{2(1-zz')dz'}{z'^3}, & \gamma' \in -\Gamma_+. \end{cases} \quad (3.14)$$

---

<sup>12</sup>The functions  $H_\gamma$  have a simple geometric meaning [26, 57]: they generate contact transformations (i.e. preserving the contact structure) relating the Darboux coordinates living in the patches separated by the BPS rays. In fact, these functions together with the contours  $\ell_\gamma$  are the fundamental data fixing the contact structure on the twistor space.

<sup>13</sup>The other contributions to the prepotential, representing perturbative and non-perturbative  $\alpha'$ -corrections, combine with D(-1) and D1-instantons, but are irrelevant for our discussion of instantons with non-vanishing D3-brane charge.

This shows that the contribution of  $\gamma' \in -\Gamma_+$  is suppressed comparing to  $\gamma' \in \Gamma_+$  and therefore can be neglected. As a result, the system of integral equations (3.2) in the large volume limit where only D3-D1-D(-1) instantons contribute reduces to the following system of integral equations for  $H_\gamma$

$$H_\gamma(z) = H_\gamma^{\text{cl}} \exp \left[ \sum_{\gamma' \in \Gamma_+} \int_{\ell_{\gamma'}} dz' K_{\gamma\gamma'}(z, z') H_{\gamma'}(z') \right], \quad (3.15)$$

where the effective kernel is

$$K_{\gamma_1\gamma_2}(z_1, z_2) = 2\pi \left( (tp_1p_2) + \frac{i\langle\gamma_1, \gamma_2\rangle}{z_1 - z_2} \right). \quad (3.16)$$

Below we will need a perturbative solution of the integral equation (3.15). Applying the iterative procedure outlined below (3.4), or equivalently using the Lagrange inversion theorem, such solution can be written as a sum over *rooted trees* [30, §C],

$$H_{\gamma_1}(z_1) = H_{\gamma_1}^{\text{cl}}(z_1) \sum_{n=1}^{\infty} \left( \prod_{i=2}^n \sum_{\gamma_i \in \Gamma_+} \int_{\ell_{\gamma_i}} dz_i H_{\gamma_i}^{\text{cl}}(z_i) \right) \sum_{\mathcal{T} \in \mathbb{T}_n^r} \frac{\mathcal{A}(\mathcal{T})}{|\text{Aut}(\mathcal{T})|}, \quad (3.17)$$

where  $\mathbb{T}_n^r$  is the set of rooted trees with  $n$  vertices and  $\mathcal{A}(\mathcal{T})$  is the ‘amplitude’

$$\mathcal{A}(\mathcal{T}) = \prod_{e \in E_{\mathcal{T}}} K_{\gamma_{s(e)}, \gamma_{t(e)}}(z_{s(e)}, z_{t(e)}). \quad (3.18)$$

A rooted tree<sup>14</sup>  $\mathcal{T}$  consists of  $n$  vertices joined by directed edges so that the root vertex has only outgoing edges, whereas all other vertices have one incoming edge and an arbitrary number of outgoing ones. We label the vertices of  $\mathcal{T}$  by  $\mathbf{v} = 1, \dots, n$  in an arbitrary fashion, except for the root which is labelled by  $\mathbf{v} = 1$ . The symmetry factor  $|\text{Aut}(\mathcal{T})|$  is the order of the symmetry group which permutes the labels  $2, \dots, n$  without changing the topology of the tree. Each vertex is decorated by a charge vector  $\gamma_{\mathbf{v}}$  and a complex variable  $z_{\mathbf{v}} \in \mathbb{C}P^1$ . We denote the set of edges by  $E_{\mathcal{T}}$ , the set of vertices by  $V_{\mathcal{T}}$ , and the source and target vertex of an edge  $e$  by  $s(e)$  and  $t(e)$ , respectively. Unpacking these notations, we get, at the few leading orders,

$$\begin{aligned} H_{\gamma_1} &= H_{\gamma_1}^{\text{cl}} + \sum_{\gamma_2} K_{12} H_{\gamma_1}^{\text{cl}} H_{\gamma_2}^{\text{cl}} + \sum_{\gamma_2, \gamma_3} \left( \frac{1}{2} K_{12} K_{13} + K_{12} K_{23} \right) H_{\gamma_1}^{\text{cl}} H_{\gamma_2}^{\text{cl}} H_{\gamma_3}^{\text{cl}} \\ &+ \sum_{\gamma_2, \gamma_3, \gamma_4} \left( \frac{1}{6} K_{12} K_{13} K_{14} + \frac{1}{2} K_{12} K_{23} K_{24} + K_{12} K_{13} K_{24} + K_{12} K_{23} K_{34} \right) H_{\gamma_1}^{\text{cl}} H_{\gamma_2}^{\text{cl}} H_{\gamma_3}^{\text{cl}} H_{\gamma_4}^{\text{cl}} + \dots \end{aligned} \quad (3.19)$$

where we omitted the integrals and denoted  $K_{ij} = K_{\gamma_i, \gamma_j}(z_i, z_j)$ . The expansion (3.17) is effectively a multi-instanton expansion in powers of the DT invariants  $\bar{\Omega}(\gamma_i)$ , which is asymptotic to the exact result at weak coupling.

---

<sup>14</sup>We will use calligraphic letters  $\mathcal{T}$  for trees where charges  $\gamma_i$  are assigned to vertices to distinguish them from rooted trees  $T$  where the charges are assigned to leaves (hence  $T$  has always more than  $n$  vertices). Similarly, we will use notations  $\mathbf{v}$  and  $v$  for vertices of these two types of trees, respectively. Note also that whereas  $V_{\mathcal{T}}$  denotes the set of all vertices,  $V_T$  does not include the leaves. An example of trees of the latter type are attractor flow trees.

### 3.3 From the contact potential to the instanton generating function

Let us recall that our goal is to derive constraints imposed by S-duality on the DT invariants  $\bar{\Omega}(\gamma)$  appearing as coefficients in the multi-instanton expansion. To achieve this goal, rather than studying the full metric on  $\mathcal{M}_H$ , it suffices to consider a suitable function on this moduli space which has a non-trivial dependence on  $\bar{\Omega}(\gamma)$  and specified transformations under S-duality. Remarkably, there is a natural candidate with the above properties — the so-called contact potential  $e^\phi$ , a real function which is well-defined on any quaternion-Kähler manifold with a continuous isometry [28]. Furthermore, there is a general expression for the contact potential in terms of Penrose-type integrals on the  $\mathbb{C}P^1$  fiber. In the present case, the required isometry is the shift of the NS-axion, which survives all quantum corrections as long as NS5-instantons are switched off. The contact potential is then given by the exact formula [26]

$$e^\phi = \frac{i\tau_2^2}{16} (\bar{u}^\Lambda F_\Lambda - u^\Lambda \bar{F}_\Lambda) - \frac{\chi_{\mathfrak{Y}}}{192\pi} + \frac{i\tau_2}{16} \sum_\gamma \int_{\ell_\gamma} \frac{dt}{t} (t^{-1} Z_\gamma(u^a) - t \bar{Z}_\gamma(\bar{u}^a)) H_\gamma, \quad (3.20)$$

where  $\chi_{\mathfrak{Y}}$  is the Euler characteristic of  $\mathfrak{Y}$ . This formula indeed captures contribution from D-instantons due to the last term proportional to  $H_\gamma$ .

On the other hand, in the classical, large volume limit one finds  $e^\phi = \frac{\tau_2^2}{12}(t^3)$ , which shows that the contact potential can be identified with the four-dimensional dilaton and in this approximation behaves as a modular form of weight  $(-\frac{1}{2}, -\frac{1}{2})$  under S-duality transformations (3.5). In fact, one can show [26] that  $SL(2, \mathbb{Z})$  preserves the contact structure, i.e. it is an isometry of  $\mathcal{M}_H$ , only if the *full non-perturbative* contact potential transforms in this way,

$$e^\phi \mapsto \frac{e^\phi}{|c\tau + d|}. \quad (3.21)$$

Furthermore, since S-duality acts by rescaling on both Kähler moduli  $t^a$  and fiber coordinate  $z$  (see (3.7)), it preserves each order in the expansion around the large volume limit. This implies that the large volume limit of the D3-instanton contribution to  $e^\phi$ , which we denote by  $(e^\phi)_{\text{D3}}$ , must itself transform as (3.21). It is this condition that we shall exploit to derive modularity constraints on the DT invariants.

To make this condition more explicit, let us extract the D3-instanton contribution to the function (3.20). The procedure is the same as the one used to get (3.15), and we relegate the details of the calculation to appendix B. The result can be written in a concise way using the complex function defined by

$$\mathcal{G} = \sum_{\gamma \in \Gamma_+} \int_{\ell_\gamma} dz H_\gamma(z) - \frac{1}{2} \sum_{\gamma_1, \gamma_2 \in \Gamma_+} \int_{\ell_{\gamma_1}} dz_1 \int_{\ell_{\gamma_2}} dz_2 K_{\gamma_1 \gamma_2}(z_1, z_2) H_{\gamma_1}(z_1) H_{\gamma_2}(z_2) \quad (3.22)$$

and the modular covariant derivative operator

$$\mathcal{D}_{\mathfrak{h}} = \frac{1}{2\pi i} \left( \partial_\tau + \frac{\mathfrak{h}}{2i\tau_2} + \frac{it^a}{4\tau_2} \partial_{t^a} \right), \quad (3.23)$$

which maps modular functions of weight  $(\mathfrak{h}, \bar{\mathfrak{h}})$  to modular functions of weight  $(\mathfrak{h}+2, \bar{\mathfrak{h}})$ . Then one has (generalizing [20, (4.5)] to all orders in the instanton expansion)

$$(e^\phi)_{\text{D3}} = \frac{\tau_2}{2} \text{Re } \mathcal{D}_{-\frac{3}{2}} \mathcal{G} + \frac{1}{32\pi^2} \kappa_{abc} t^c \partial_{\bar{c}_a} \mathcal{G} \partial_{\bar{c}_b} \bar{\mathcal{G}}. \quad (3.24)$$

It is immediate to see that  $(e^\phi)_{D_3}$  transforms under S-duality as (3.21) provided the function  $\mathcal{G}$  transforms as a modular form of weight  $(-\frac{3}{2}, \frac{1}{2})$ . In order to derive the implications of this fact, we need to express  $\mathcal{G}$  in terms of the generalized DT invariants.

For this purpose, we substitute the multi-instanton expansion (3.17) into (3.22). We claim that the result takes the simple form

$$\mathcal{G} = \sum_{n=1}^{\infty} \left[ \prod_{i=1}^n \sum_{\gamma_i \in \Gamma_+} \int_{\ell_{\gamma_i}} dz_i H_{\gamma_i}^{\text{cl}}(z_i) \right] \mathcal{G}_n(\{\gamma_i, z_i\}), \quad (3.25)$$

where  $\mathcal{G}_n(\{\gamma_i, z_i\})$  is now a sum over *unrooted* trees with  $n$  vertices,

$$\mathcal{G}_n(\{\gamma_i, z_i\}) = \sum_{\mathcal{T} \in \mathbb{T}_n} \frac{\mathcal{A}(\mathcal{T})}{|\text{Aut}(\mathcal{T})|} = \frac{1}{n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \mathcal{A}(\mathcal{T}), \quad (3.26)$$

and in the second equality we rewrote the result as a sum over *unrooted labelled* trees.<sup>15</sup>

To see why this is the case, observe that under the action of the Euler operator  $\hat{D} = H^{\text{cl}} \partial_{H^{\text{cl}}}$  rescaling all functions  $H_\gamma^{\text{cl}}$ , the function  $\mathcal{G}$  maps to the first term in (3.22), which we denote by  $\mathcal{F}$ . Namely,

$$\hat{D} \cdot \mathcal{G} = \mathcal{F}, \quad \mathcal{F} \equiv \sum_{\gamma \in \Gamma_+} \int_{\ell_\gamma} dz H_\gamma, \quad (3.27)$$

as can be easily verified with the help of the integral equation (3.15). The multi-instanton expansion of  $\mathcal{F}$  follows immediately from (3.17),

$$\mathcal{F} = \sum_{n=1}^{\infty} \left( \prod_{i=1}^n \sum_{\gamma_i \in \Gamma_+} \int_{\ell_{\gamma_i}} dz_i H_{\gamma_i}^{\text{cl}}(z_i) \right) \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \frac{\mathcal{A}(\mathcal{T})}{|\text{Aut}(\mathcal{T})|}. \quad (3.28)$$

Integrating back the action of the derivative operator  $\hat{D}$ , we see that the sum over rooted trees in (3.28) turns into the sum over unrooted trees in (3.26). At the first few orders we get, using the same shorthand notation as in (3.19),

$$\begin{aligned} \mathcal{G} &= \sum_{\gamma} H_{\gamma}^{\text{cl}} + \frac{1}{2} \sum_{\gamma_1, \gamma_2} K_{12} H_{\gamma_1}^{\text{cl}} H_{\gamma_2}^{\text{cl}} + \frac{1}{2} \sum_{\gamma_1, \gamma_2, \gamma_3} K_{12} K_{23} H_{\gamma_1}^{\text{cl}} H_{\gamma_2}^{\text{cl}} H_{\gamma_3}^{\text{cl}} \\ &+ \sum_{\gamma_1, \gamma_2, \gamma_3, \gamma_4} \left( \frac{1}{6} K_{12} K_{13} K_{14} + \frac{1}{2} K_{12} K_{23} K_{34} \right) H_{\gamma_1}^{\text{cl}} H_{\gamma_2}^{\text{cl}} H_{\gamma_3}^{\text{cl}} H_{\gamma_4}^{\text{cl}} + \dots \end{aligned} \quad (3.29)$$

The simplicity of the expansion (3.26), and the relation (3.24) to the contact potential, show that the function  $\mathcal{G}$  is very natural and, in some sense, more fundamental<sup>16</sup> than the naive instanton sum  $\mathcal{F}$ . We shall henceforth refer to  $\mathcal{G}$  as the ‘instanton generating function’. In the following we shall postulate that  $\mathcal{G}$  transforms as a modular form of weight  $(-\frac{3}{2}, \frac{1}{2})$ , and analyze the consequences of this requirement for the DT invariants.

<sup>15</sup>The number of such trees is  $|\mathbb{T}_n^\ell| = n^{n-2} = \{1, 1, 3, 16, 125, 1296, \dots\}$  for  $n \geq 1$ . Such trees also appear in the Joyce-Song wall-crossing formula [36, 37] and are conveniently labelled by their Prüfer code.

<sup>16</sup>In [20], it was noticed that the function  $\mathcal{G}$ , denoted by  $\tilde{\mathcal{F}}$  in that reference and computed at second order in the multi-instanton expansion, could be obtained from the seemingly simpler function  $\mathcal{F}$  by *halving* the coefficient of its second order contribution. Now we see that this *ad hoc* prescription is the consequence of going from rooted to unrooted trees, as a result of adding the second term in (3.22).

## 4. Theta series decomposition and modularity

### 4.1 Factorisation

To derive modularity constraints on the DT invariants, we need to perform a theta series decomposition of the generating function  $\mathcal{G}$  defined in (3.22). To this end, let us make use of the fact noticed in (2.23) that the DSZ products  $\langle \gamma, \gamma' \rangle$  and hence the kernels (3.16) do not depend on the  $q_0$  charge. Choosing then the quadratic refinement  $\sigma_\gamma$  as in (D.5), which is also  $q_0$ -independent, and using the factorization (3.12) of  $\mathcal{X}_\gamma^{\text{cl}}$ , one can rewrite the expansion (3.25) as follows

$$\mathcal{G} = \sum_{n=1}^{\infty} \frac{1}{(2\pi)^{2n}} \left[ \prod_{i=1}^n \sum_{p_i, q_i} \sigma_{p_i, q_i} h_{p_i, q_i}^{\text{DT}} \int_{\ell_{\gamma_i}} dz_i \mathcal{X}_{p_i, q_i}^{(\theta)}(z_i) \right] \mathcal{G}_n(\{\gamma_i, z_i\}), \quad (4.1)$$

where the sum over the invariant charges  $\hat{q}_{i,0}$  gave rise to the generating functions of DT invariants defined in (2.27). This is not yet the desired form because these generating functions depend non-trivially on the remaining electric charges  $q_{i,a}$ . If it were not for this dependence, the sum over  $q_{i,a}$  would produce certain non-Gaussian theta series, and at each order we would have a product of this theta series and  $n$  generating functions. Then the modular properties of the theta series would dictate the modular properties of the generating function.

Such a theta series decomposition can be achieved by expressing the DT invariants in terms of the MSW invariants, for which the dependence on electric charges  $q_{i,a}$  reduces to the dependence on the residue classes  $\mu_{i,a}$  due to the spectral flow symmetry. Substituting the expansion (2.30) of  $h_{p,q}^{\text{DT}}$  in terms of  $h_{p,\mu}$ , the expansion (4.1) of the function  $\mathcal{G}$  can be brought to the following factorized form

$$\mathcal{G} = \sum_{n=1}^{\infty} \frac{2^{-\frac{n}{2}}}{\pi \sqrt{2\tau_2}} \left[ \prod_{i=1}^n \sum_{p_i, \mu_i} \sigma_{p_i} h_{p_i, \mu_i} \right] e^{-S_p^{\text{cl}}} \vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi_n^{\text{tot}}, n-2), \quad (4.2)$$

where  $\vartheta_{\mathbf{p}, \boldsymbol{\mu}}$  is the theta series (D.1) with parameter  $\lambda = n-2$  and the kernel which has the following structure

$$\Phi_n^{\text{tot}}(\mathbf{x}) = \text{Sym} \left\{ \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1}} \Phi_m^f(\mathbf{x}') \prod_{k=1}^m \Phi_{n_k}^g(x_{j_{k-1}+1}, \dots, x_{j_k}) \right\}, \quad (4.3)$$

where the sum runs over ordered partitions of  $n$ , whose number of parts is denoted by  $m$ , and we adopted notations from (2.16) for indices  $j_k$ . The argument  $\mathbf{x}$  of the kernel encodes the electric components of the charges (shifted by the B-field and rescaled by  $\sqrt{2\tau_2}$ ) and lives in a vector space  $(\oplus_{i=1}^n \Lambda_i) \otimes \mathbb{R}$  of dimension  $d = nb_2$ , given by  $b_2$  copies of the lattice  $\Lambda$ , where the  $i$ -th copy  $\Lambda_i$  carries the bilinear form  $\kappa_{i,ab} = \kappa_{abc} p_i^c$  of signature  $(1, b_2 - 1)$ . Therefore,  $\vartheta_{\mathbf{p}, \boldsymbol{\mu}}$  is an indefinite theta series of signature  $(n, n(b_2 - 1))$  with the bilinear form given explicitly in (D.4).

Finally, the kernel (4.3) is constructed from two other functions. The first,  $\Phi_n^f$ , is the multiple integral of the coefficient  $\mathcal{G}_n$  in the expansion (3.25),

$$\Phi_n^f(\mathbf{x}) = \left( \frac{\sqrt{2\tau_2}}{2\pi} \right)^{n-1} \left[ \prod_{i=1}^n \int_{\ell_{\gamma_i}} \frac{dz_i}{2\pi} W_{p_i}(x_i, z_i) \right] \mathcal{G}_n(\{\gamma_i, z_i\}), \quad (4.4)$$

with measure factor

$$W_p(x, z) = e^{-2\pi\tau_2 z^2 (pt^2) - 2\pi i \sqrt{2\tau_2} z (pxt)} \quad (4.5)$$

coming from the  $z$ -dependent part of (3.13). Although this function is written in terms of  $\mathcal{G}_n$  depending on full electromagnetic charge vectors  $\gamma_i$ , it is actually independent of their  $q_0$  components. Indeed, using the result (3.26), it can be rewritten as

$$\Phi_n^f(\mathbf{x}) = \frac{1}{n!} \left[ \prod_{i=1}^n \int_{\ell_{\gamma_i}} \frac{dz_i}{2\pi} W_{p_i}(x_i, z_i) \right] \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \prod_{e \in E_{\mathcal{T}}} \hat{K}_{s(e)t(e)}, \quad (4.6)$$

where we introduced a rescaled version of the kernel (3.16)

$$\hat{K}_{ij}(z_i, z_j) = \left( \left( \sqrt{2\tau_2} t + i \frac{x_i - x_j}{z_i - z_j} \right) p_i p_j \right). \quad (4.7)$$

Note that  $\tau_2$  and  $t^a$  appear only in the modular invariant combination  $\sqrt{2\tau_2} t^a$ . In (4.3) this function appears with the argument  $\mathbf{x}'$  and carries a dependence on  $\mathbf{p}'$  (not indicated explicitly) which are both  $mb_2$ -dimensional vectors with components (cf. (2.16))

$$p_k'^a = \sum_{i=j_{k-1}+1}^{j_k} p_i^a, \quad x_k'^a = \kappa_k'^{ab} \sum_{i=j_{k-1}+1}^{j_k} \kappa_{i,bc} x_i^c, \quad (4.8)$$

where  $\kappa_k'^{ab} = \kappa_{abc} p_k'^c$ .

The second function,  $\Phi_n^g$ , appears due to the expansion of DT invariants in terms of the MSW invariants and is given by a suitably rescaled tree index

$$\Phi_n^g(\mathbf{x}) = \frac{\sigma_\gamma (\sqrt{2\tau_2})^{n-1}}{\prod_{i=1}^n \sigma_{\gamma_i}} g_{\text{tr},n}(\{\gamma_i\}, z^a). \quad (4.9)$$

It is also written in terms of functions depending on the full electromagnetic charge vectors  $\gamma_i$  (with  $\gamma = \gamma_1 + \dots + \gamma_n$ ). However, using (3.4), all quadratic refinements can be expressed through  $(-1)^{\langle \gamma_i, \gamma_j \rangle}$  which cancel the corresponding factors in the tree index (see (2.6)). Furthermore, as was noticed in the end of section 2.3, the tree index is independent of the  $q_0$  components of the charge vectors. Therefore, it can be written as a function of  $p_i^a$ ,  $\mu_{i,a}$  and  $x_i^a = \sqrt{2\tau_2} (\kappa_i^{ab} q_{i,b} + b^a)$ . Then, since after canceling the sign factors  $g_{\text{tr},n}$  is homogeneous of degree  $n-1$  in  $q_{i,a}$ , all factors of  $\sqrt{2\tau_2}$  in (4.9) cancel as well.

## 4.2 Modularity and Vignéras' equation

As explained in appendix D.1, the theta series  $\vartheta_{\mathbf{p},\mu}(\Phi, \lambda)$  is a vector-valued modular form of weight  $(\lambda + d/2, 0)$  provided the kernel  $\Phi$  satisfies Vignéras' equation (D.3) — along with certain growth conditions which we expect to be automatically satisfied for the kernels of interest in this work. In our case  $\lambda = n - 2$  and the dimension of the lattice  $d = nb_2(\mathfrak{Q})$  so that the expected weight of the theta series is  $(2(n-1) + nb_2/2, 0)$ . It is consistent with weight  $(-\frac{3}{2}, \frac{1}{2})$  of  $\mathcal{G}$  given in (4.2) only if  $h_{\mathbf{p},\mu}$  is a vector-valued holomorphic modular form of weight  $(-b_2/2 - 1, 0)$ . However for this to be true, the kernel  $\Phi_n^{\text{tot}}$  ought to satisfy Vignéras' equation. Let us check whether or not this is the case.



To this end, we first consider the kernel  $\Phi_n^f$  (4.6). In appendix E we evaluate explicitly the multiple integrals defining this kernel. To present the final result, let us introduce the following  $d$ -dimensional vectors  $\mathbf{v}_{ij}, \mathbf{u}_{ij}$ :

$$\begin{aligned} (\mathbf{v}_{ij})_k^a &= \delta_{ki} p_j^a - \delta_{kj} p_i^a & \text{such that } \mathbf{v}_{ij} \cdot \mathbf{x} &= (p_i p_j (x_i - x_j)), \\ (\mathbf{u}_{ij})_k^a &= \delta_{ki} (p_j t^2)^a - \delta_{kj} (p_i t^2)^a & \text{such that } \mathbf{u}_{ij} \cdot \mathbf{x} &= (p_j t^2) (p_i x_i t) - (p_i t^2) (p_j x_j t), \end{aligned} \quad (4.10)$$

where  $k$  labels the copy in  $\oplus_{k=1}^n \Lambda_k$ ,  $a = 1, \dots, b_2$ , and the bilinear form is given in (D.4). The first scalar product  $\mathbf{v}_{ij} \cdot \mathbf{x}$  corresponds to the DSZ product  $\langle \gamma_i, \gamma_j \rangle$ , whereas the second  $\mathbf{u}_{ij} \cdot \mathbf{x}$  to  $-2 \operatorname{Im} [Z_{\gamma_i} \bar{Z}_{\gamma_j}]$  (2.23), both rescaled by  $\sqrt{2\tau_2}$  and expressed in terms of  $x_i^a$ . From these vectors we also construct two sets of vectors which are assigned to the edges of an unrooted labelled tree  $\mathcal{T}$ , such as the trees appearing in (3.26) and (4.6). Namely,

$$\mathbf{v}_e = \sum_{i \in V_{\mathcal{T}_e^s}} \sum_{j \in V_{\mathcal{T}_e^t}} \mathbf{v}_{ij}, \quad \mathbf{u}_e = \sum_{i \in V_{\mathcal{T}_e^s}} \sum_{j \in V_{\mathcal{T}_e^t}} \mathbf{u}_{ij}, \quad (4.11)$$

where  $\mathcal{T}_e^s, \mathcal{T}_e^t$  are the two disconnected trees obtained from the tree  $\mathcal{T}$  by removing the edge  $e$ . Then the kernel  $\Phi_n^f$  can be expressed as follows<sup>17</sup>

$$\Phi_n^f(\mathbf{x}) = \frac{\Phi_1^f(x)}{2^{n-1} n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \tilde{\Phi}_{n-1}^M(\{\mathbf{u}_e\}, \{\mathbf{v}_{s(e)t(e)}\}; \mathbf{x}). \quad (4.12)$$

Here the first factor is simply a Gaussian

$$\Phi_1^f(x) = \frac{e^{-\frac{\pi(pxt)^2}{(pt^2)}}}{2\pi\sqrt{2\tau_2}(pt^2)} \quad (4.13)$$

which ensures the suppression along the direction of the total charge in the charge lattice. In the second factor one sums over unrooted labelled trees  $\mathcal{T}$  with  $n$  vertices, with summand given by a function  $\tilde{\Phi}_n^M$  defined as in (D.17) with  $m = n$ , upon replacing  $\Phi_n^E$  by  $\Phi_n^M$  in that expression. Both  $\Phi_n^E$  and  $\Phi_n^M$  are the so-called generalized (complementary) error functions introduced in [23] and further studied in [24], whose definitions are recalled in (D.10), (D.11) and (D.14). The functions  $\tilde{\Phi}_{n-1}^M$  in (4.12) depend on two sets of  $n-1$   $d$ -dimensional vectors: the vectors from the first set are given by  $\mathbf{u}_e$  defined above, whereas  $\mathbf{v}_{s(e)t(e)}$  coincide with  $\mathbf{v}_{ij}$  for  $i$  and  $j$  corresponding to the source and target vertices of edge  $e$  of the labelled tree.

The remarkable property of the generalized error functions  $\Phi_{n-1}^M$  and their upgraded versions  $\tilde{\Phi}_{n-1}^M$  is that, away from certain loci where these functions are discontinuous, they satisfy Vignéras' equation for  $\lambda = 0$  and  $n-1$ , respectively. Given that  $\Phi_1^f$  is also a solution for  $\lambda = -1$ , and the vector  $\mathbf{t} = (t^a, \dots, t^a)$  (such that  $\mathbf{t} \cdot \mathbf{x} = (pxt)$ ) is orthogonal to all vectors  $\mathbf{u}_e$  and  $\mathbf{v}_{s(e)t(e)}$ , the kernel (4.12) solves this equation for  $\lambda = n-2$ .

However, as mentioned above, it fails to do so at the discontinuity loci. These discontinuities arise due to dependence of the integration contours  $\ell_\gamma$  on moduli and electric charges. Of course, since the integrands are meromorphic functions, the integrals do not depend on

---

<sup>17</sup>Both vectors  $\mathbf{u}_e$  and  $\mathbf{v}_{s(e)t(e)}$  depend on the choice of orientation of the edge  $e$ , but this ambiguity is cancelled in the function  $\tilde{\Phi}_{n-1}^M$ .

deformations of the contours provided they do not cross the poles. But this is exactly what happens when two BPS rays, say  $\ell_\gamma$  and  $\ell_{\gamma'}$ , exchange their positions, which in turn takes place when the phases of the corresponding central charges  $Z_\gamma$  and  $Z_{\gamma'}$  align, as follows from (3.3). The loci where such alignment takes place are nothing else but the walls of marginal stability. This point will play an important role in the next subsection since it makes possible to recombine the discontinuities of the generalized error functions with discontinuities of the tree indices.

Next, we evaluate the action of Vignéras' operator on  $\Phi_n^g$ . To this end, it is convenient to use the representation of the tree index as a sum over attractor flow trees (2.3). Let us assign a  $nb_2$ -dimensional vector  $\tilde{\mathbf{v}}_v$  to each vertex  $v$  of a flow tree. Denoting by  $\mathcal{I}_v$  the set of leaves which are descendants of vertex  $v$ , we set

$$\tilde{\mathbf{v}}_v = \sum_{i \in \mathcal{I}_L(v)} \sum_{j \in \mathcal{I}_R(v)} \mathbf{v}_{ij}. \quad (4.14)$$

With these definitions the kernel (4.9) can be written as

$$\Phi_n^g(\mathbf{x}) = (-1)^{n-1} \sum_{T \in \mathbb{T}_n^{\text{af}}} \prod_{v \in V_T} (\tilde{\mathbf{v}}_v, \mathbf{x}) \Delta_{\gamma_{L(v)} \gamma_{R(v)}}^{z_p(v)}. \quad (4.15)$$

The factors  $\Delta_{\gamma_L \gamma_R}^z$  are locally constant and therefore, if we stay away from the loci of their discontinuities, the action of Vignéras operator reduces to its action on the scalar products  $(\tilde{\mathbf{v}}_v, \mathbf{x})$ . For a single such factor one finds

$$V_\lambda(\tilde{\mathbf{v}}_v, \mathbf{x}) = (\tilde{\mathbf{v}}_v, \mathbf{x}) V_{\lambda-1} + 2\tilde{\mathbf{v}}_v \cdot \partial_{\mathbf{x}}. \quad (4.16)$$

The crucial observation is that all vectors  $\tilde{\mathbf{v}}_v$  appearing in the product (4.15) for a single tree are mutually orthogonal  $(\tilde{\mathbf{v}}_v, \tilde{\mathbf{v}}_{v'}) = 0$ , which is clear because  $\langle \gamma_{L(v)}, \gamma_{R(v)} \rangle$  is antisymmetric in charges  $\gamma_{L(v)}, \gamma_{R(v)}$ , whereas the factors associated with vertices which are not descendants of  $v$  either depend on their sum or do not depend on them at all. Therefore, one obtains

$$V_\lambda \Phi_{n_k}^g(\mathbf{x}) = \Phi_{n_k}^g(\mathbf{x}) V_{\lambda-n_k+1} + 2(-1)^{n_k-1} \sum_{T \in \mathbb{T}_{n_k}^{\text{af}}} \Delta(T) \sum_{v \in V_T} \left[ \prod_{v' \in V_T \setminus \{v\}} (\tilde{\mathbf{v}}_{v'}, \mathbf{x}) \right] \tilde{\mathbf{v}}_v \cdot \partial_{\mathbf{x}}. \quad (4.17)$$

Let us now evaluate the action of  $V_{n-2}$  on the full kernel  $\Phi_n^{\text{tot}}$ . Applying the result (4.17), we observe that the second term vanishes on the other factors in (4.3) due to the same reason that they either do not depend ( $\Phi_{n_{k'}}^g, k' \neq k$ ) or depend ( $\Phi_m^f$ ) only on the sum of charges entering  $\Phi_{n_k}^g$ . Therefore, one finds that, *away from discontinuities of generalized error functions and  $\Delta(T)$* , one has  $V_{n-2} \cdot \Phi_n^{\text{tot}} = 0$ .

### 4.3 Discontinuities and the anomaly

Let us now turn to the discontinuities of  $\Phi_n^{\text{tot}}$  which we ignored so far and which spoil Vignéras' equation and hence modularity of the theta series. There are three potential sources of these discontinuities:

1. walls of marginal stability — at these loci  $\Phi_m^f$  are discontinuous due to exchange of integration contours and  $\Phi_{n_k}^g$  jump due to factors  $\Delta_{\gamma_{L(v)} \gamma_{R(v)}}^{z_p(v)}$  assigned to the root vertices of attractor flow trees;

2. ‘fake walls’ — these are loci in the moduli space where  $\text{Im} [Z_{\gamma_{L(v)}} \bar{Z}_{\gamma_{R(v)}}(z_{p(v)}^a)] = 0$ , and hence the corresponding  $\Delta$ -factor jumps, where  $v$  is not a root vertex — they have the meaning of walls of marginal stability for the intermediate bound states appearing in the attractor flow;
3. moduli independent loci where  $(\tilde{\mathbf{v}}_v, \mathbf{x}) = 0$  — at these loci the factors  $\Delta_{\gamma_{L(v)}\gamma_{R(v)}}^{z_{p(v)}}$  and hence  $\Phi_{n_k}^g$  are discontinuous due to the second term in (2.4).

Remarkably, the two effects due to the non-trivial charge and moduli dependence of the DT invariants and the contours cancel each other and the function  $\mathcal{G}$  turns out to be smooth at loci of the first type. This is expected because the whole construction of D-instantons has been designed to make the resulting metric on the moduli space smooth across these loci, which required the cancellation of these discontinuities [30, 29]. Besides, in [31] it was proven that the contact potential is also smooth, which suggests that the function  $\mathcal{G}$  must be smooth as well. In appendix C we present an explicit proof of this fact based on the representation in terms of trees.

Furthermore, in [19] it was shown that the discontinuities across ‘fake walls’ cancel in the sum over flow trees as well. In fact, this cancellation is explicit in the representation of the partial tree index given by the recursive formula (2.15) where the signs generating such ‘fake discontinuities’ do not arise at all. As a result, it remains to consider only the discontinuities of the third type corresponding to the moduli independent loci.

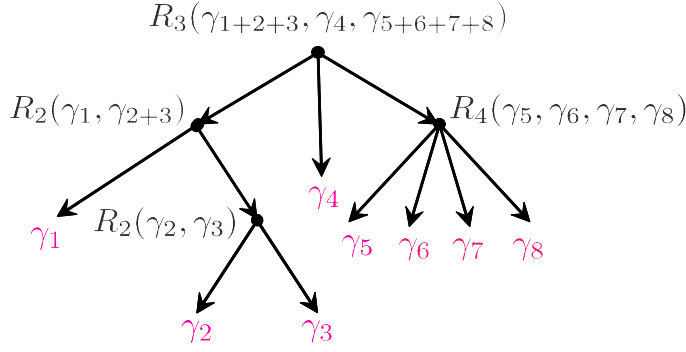
It is straightforward to check that already for  $n = 2$  these discontinuities are present and do spoil modularity of the theta series. For small  $n$  one can explicitly evaluate the anomaly in Vignéras’ equation. It is given by a series of terms proportional to  $\delta(\tilde{\mathbf{v}}_v, \mathbf{x})$ . Note that no derivatives of delta functions or products of two delta functions arise despite the presence of the second derivative in Vignéras’ operator. This is because each  $\text{sgn}(\tilde{\mathbf{v}}_v, \mathbf{x})$  from  $\Delta(T)$  is multiplied by  $(\tilde{\mathbf{v}}_v, \mathbf{x})$  from  $\kappa(T)$  in (2.3) and one gets a non-vanishing result only if the second order derivative operator acts on both factors. In particular, this implies that the anomaly is completely characterized by the action of  $V_\lambda$  on  $\Phi_n^g$ .

## 5. Modular completion

Since the theta series  $\vartheta_{\mathbf{p}, \mu}(\Phi_n^{\text{tot}})$  are not modular for  $n \geq 2$ , the analysis of the previous section implies that the generating function  $h_{p, \mu}$  of the MSW invariants is not modular either whenever the divisor  $\mathcal{D} = p^a \gamma_a$  is reducible. However, its modular anomaly has a definite structure. In particular, in [20] it was shown that for  $n = 2$ ,  $h_{p, \mu}$  must be a vector-valued *mixed mock* modular form, i.e. it has a non-holomorphic completion  $\widehat{h}_{p, \mu}$  constructed in a specific way from a set of holomorphic modular forms and their Eichler integrals [25, 13]. In this section we generalize this result for arbitrary  $n$ , i.e. for any degree of reducibility of the divisor.

### 5.1 Completion of the generating function

Let us recall the notations  $\tilde{\gamma} = (p^a, q_a)$  and  $Q_n$  from (2.31), and decompose the electric



**Figure 4:** An example of Schröder tree contributing to  $W_8$ . Near each vertex we showed the corresponding factor using the shorthand notation  $\gamma_{i+j} = \gamma_i + \gamma_j$ .

component  $q_a$  as in (2.25). Then we define

$$\widehat{h}_{p,\mu}(\tau) = h_{p,\mu}(\tau) - \sum_{n=2}^{\infty} \sum_{\sum_{i=1}^n \check{\gamma}_i = \check{\gamma}} R_n(\{\check{\gamma}_i\}, \tau) e^{\pi i \tau Q_n(\{\check{\gamma}_i\})} \prod_{i=1}^n h_{p_i, \mu_i}(\tau). \quad (5.1)$$

We are looking for *non-holomorphic* functions  $R_n$ , exponentially suppressed for large  $\tau_2$ , such that  $\widehat{h}_{p,\mu}$  transforms as a modular form. The condition for this to be true can be found along the same lines as before: one needs to rewrite the expansion of the function  $\mathcal{G}$  as a series in  $\widehat{h}_{p,\mu}$  and require that at each order the coefficient is given by a modular covariant theta series. For such a theta series decomposition to be possible however, it is important that  $\widehat{h}_{p,\mu}$  be invariant under the spectral flow, which implies that the functions  $R_n$  be *independent* of the spectral flow parameter  $\epsilon^a$  of the total charge  $\check{\gamma}$ . This condition will be an important consistency requirement on our construction.

Rather than inverting (5.1) and substituting the result into (4.2), we can consider the generating function of DT invariants  $h_{p,\mu}^{\text{DT}}$  and, as a first step, rewrite it as a series in  $\widehat{h}_{p,\mu}$ . Denoting the coefficient of such expansion by  $\widehat{g}_n(\{\check{\gamma}_i\}, z^a)$  (with  $\widehat{g}_1 \equiv 1$ ), we get

$$h_{p,q}^{\text{DT}}(\tau, z^a) = \sum_{\sum_{i=1}^n \check{\gamma}_i = \check{\gamma}} \widehat{g}_n(\{\check{\gamma}_i\}, z^a; \tau) e^{\pi i \tau Q_n(\{\check{\gamma}_i\})} \prod_{i=1}^n \widehat{h}_{p_i, \mu_i}(\tau). \quad (5.2)$$

Comparing with (2.30), we see that the coefficients  $\widehat{g}_n$  are a direct analogue of the tree index  $g_{\text{tr},n}$ . The expansion of  $\mathcal{G}$  in terms of  $\widehat{h}_{p,\mu}$  is then simply obtained by replacing  $g_{\text{tr},n}$  by  $\widehat{g}_n$  in (4.9). Our first problem is to express these coefficients in terms of the functions  $R_n$ .

The result can be nicely formulated using the so-called Schröder trees, which are rooted ordered trees such that all vertices (except the leaves) have at least 2 children (see Fig. 4). Their vertices are decorated by charges in such way that the leaves have charges  $\gamma_i$ , whereas the charges at other vertices are defined by the inductive rule  $\gamma_v = \sum_{v' \in \text{Ch}(v)} \gamma_{v'}$  where  $\text{Ch}(v)$  is the set of children of vertex  $v$ . Note that these trees should not be confused with flow trees, since they are not restricted to be binary nor do they carry moduli at the vertices. We denote the set of Schröder trees<sup>18</sup> with  $n$  leaves by  $\mathbb{T}_n^{\text{S}}$ .

<sup>18</sup>The number of Schröder trees with  $n$  leaves is the  $n - 1$ -th super-Catalan number,  $|\mathbb{T}_n^{\text{S}}| = \{1, 1, 3, 11, 45, 197, 903, \dots\}$  for  $n \geq 1$  (sequence A001003 on the Online Encyclopedia of Integer Sequences).

Let us also introduce a convenient notation: for any set of functions  $f_n(\{\check{\gamma}_i\})$  depending on  $n$  charges and a given Schröder tree  $T$ , we set  $f_v \equiv f_{k_v}(\{\check{\gamma}_{v'}\})$  where  $v' \in \text{Ch}(v)$  and  $k_v$  is their number. Using these notations, the expression of  $\widehat{g}_n$  in terms of  $R_n$  can be encoded into a recursive equation provided by the following proposition, whose proof we relegate to appendix F:

**Proposition 1.** *The coefficients  $\widehat{g}_n$  are determined recursively by the following equation*

$$\widehat{g}_n(\{\check{\gamma}_i\}, z^a) = -\frac{1}{2} \text{Sym} \left\{ \sum_{\ell=1}^{n-1} \Delta_{\gamma_L^\ell \gamma_R^\ell}^z \kappa(\gamma_{LR}^\ell) \widehat{g}_\ell(\{\check{\gamma}_i\}_{i=1}^\ell, z_\ell^a) \widehat{g}_{n-\ell}(\{\check{\gamma}_i\}_{i=\ell+1}^n, z_\ell^a) \right\} + W_n(\{\check{\gamma}_i\}), \quad (5.3)$$

where  $\gamma_L^\ell = \sum_{i=1}^\ell \gamma_i$ ,  $\gamma_R^\ell = \sum_{i=\ell+1}^n \gamma_i$ ,  $\gamma_{LR}^\ell = \langle \gamma_L^\ell, \gamma_R^\ell \rangle$ ,  $z_\ell^a$  is the value of the moduli where the attractor flow crosses the wall for the decay  $\gamma \rightarrow \gamma_L^\ell + \gamma_R^\ell$  (cf. (2.13)), and  $W_n$  is a function given by the sum over Schröder trees

$$W_n(\{\check{\gamma}_i\}) = \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^S} \prod_{v \in V_T} R_v \right\}. \quad (5.4)$$

What are the conditions on  $\widehat{g}_n$  for the corresponding theta series to be modular? Let us denote by  $\Phi_n^{\widehat{g}}$  the kernel defined by  $\widehat{g}_n$  analogously to (4.9), namely

$$\Phi_n^{\widehat{g}}(\mathbf{x}) = \frac{\sigma_\gamma (\sqrt{2\tau_2})^{n-1}}{\prod_{i=1}^n \sigma_{\gamma_i}} \widehat{g}_n(\{\gamma_i\}, z^a). \quad (5.5)$$

Then the above analysis implies that the modularity requires  $\Phi_n^{\widehat{g}}$  to satisfy Vignéras equation away from walls of marginal stability, whereas at these walls its discontinuities should follow the same pattern as before in order to cancel the discontinuities from the contour exchange in  $\Phi_m^f$ . Thus, the completion should smoothen out the discontinuities from the moduli independent signs  $\text{sgn}(\tilde{\mathbf{v}}_v, \mathbf{x})$ , but otherwise leave the action of Vignéras' operator unchanged. Formally, this means that  $\Phi_n^{\widehat{g}}$  must satisfy the following equation

$$V_{n-1} \cdot \Phi_n^{\widehat{g}} = \text{Sym} \sum_{\ell=1}^{n-1} \left[ \mathbf{u}_\ell^2 \Delta_{n,\ell}^{\widehat{g}} \delta'(\mathbf{u}_\ell \cdot \mathbf{x}) + 2\mathbf{u}_\ell \cdot \partial_{\mathbf{x}} \Delta_{n,\ell}^{\widehat{g}} \delta(\mathbf{u}_\ell \cdot \mathbf{x}) \right], \quad \Delta_{n,\ell}^{\widehat{g}} = \frac{1}{2} (\mathbf{v}_\ell, \mathbf{x}) \Phi_\ell^{\widehat{g}} \Phi_{n-\ell}^{\widehat{g}}, \quad (5.6)$$

where we introduced two sets of vectors constructed from the vectors (4.10),

$$\mathbf{v}_\ell = \sum_{i=1}^\ell \sum_{j=\ell+1}^n \mathbf{v}_{ij}, \quad \mathbf{u}_\ell = \sum_{i=1}^\ell \sum_{j=\ell+1}^n \mathbf{u}_{ij}. \quad (5.7)$$

Note that  $(\mathbf{v}_\ell, \mathbf{x})$  and  $(\mathbf{u}_\ell, \mathbf{x})$  correspond to the quantities  $-\Gamma_{n\ell}$  and  $-S_\ell$  (2.12), respectively.

To solve the above constraints, let us consider the following iterative ansatz

$$\widehat{g}_n(\{\check{\gamma}_i\}, z^a) = g_n^{(0)}(\{\check{\gamma}_i, c_i\}) - \text{Sym} \left\{ \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1, m < n}} \widehat{g}_m(\{\check{\gamma}'_k\}, z^a) \prod_{k=1}^m \mathcal{E}_{n_k}(\check{\gamma}_{j_{k-1}+1}, \dots, \check{\gamma}_{j_k}) \right\}, \quad (5.8)$$

where the notations for indices and primed charges are the same as in (2.16). This ansatz is motivated by analogy with the iterative equation for the (partial) tree index (2.15). It involves two functions to be determined below,  $g_n^{(0)}$  and  $\mathcal{E}_n$ . The former depend on the moduli through the variables  $c_i$  (2.12), whereas the latter are moduli independent. We set  $g_1^{(0)} = \mathcal{E}_1 = 1$  and also assume that  $g_n^{(0)}$  have discontinuities only at walls of marginal stability, i.e. at  $\sum_{i \in \mathcal{I}} c_i = 0$  for various subsets  $\mathcal{I}$  of indices.

The unknown functions  $g_n^{(0)}$  and  $\mathcal{E}_n$  together with the functions  $R_n$  defining the completion, or their combinations (5.4), are fixed by the conditions (5.3) and (5.6). In appendix F we prove the following result:

**Proposition 2.** *Let us split  $\mathcal{E}_n = \mathcal{E}_n^{(0)} + \mathcal{E}_n^{(+)}$  into  $\mathcal{E}_n^{(+)}$ , which is the part exponentially decreasing for large  $\tau_2$ , and the non-decreasing remainder  $\mathcal{E}_n^{(0)}$ . Then the ansatz (5.8) satisfies the recursive equation (5.6) provided*

1. *the functions  $g_n^{(0)}$  are subject to a similar recursive relation*

$$\frac{1}{4} \text{Sym} \left\{ \sum_{\ell=1}^{n-1} (\text{sgn}(S_\ell) - \text{sgn}(\Gamma_{n\ell})) \kappa(\Gamma_{n\ell}) g_\ell^{(0)}(\{\check{\gamma}_i, c_i^{(\ell)}\}_{i=1}^\ell) g_{n-\ell}^{(0)}(\{\check{\gamma}_i, c_i^{(\ell)}\}_{i=\ell+1}^n) \right\} \quad (5.9)$$

$$= g_n^{(0)}(\{\check{\gamma}_i, c_i\}) - g_n^{(0)}(\{\check{\gamma}_i, \beta_{ni}\}),$$

where  $S_k$ ,  $\beta_{kl}$ ,  $\Gamma_{kl}$  and  $c_i^{(\ell)}$  were defined in (2.12), (2.14);

2. *the non-decreasing part of  $\mathcal{E}_n$  is fixed in terms of  $g_n^{(0)}(\{\check{\gamma}_i, c_i\})$  as*

$$\mathcal{E}_n^{(0)}(\{\check{\gamma}_i\}) = g_n^{(0)}(\{\check{\gamma}_i, \beta_{ni}\}); \quad (5.10)$$

3. *its decreasing part is given by*

$$\mathcal{E}_n^{(+)}(\{\check{\gamma}_i\}) = -\text{Sym} \left\{ \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1, m > 1}} W_m(\{\check{\gamma}'_k\}) \prod_{k=1}^m \mathcal{E}_{n_k}(\check{\gamma}_{j_{k-1}+1}, \dots, \check{\gamma}_{j_k}) \right\}. \quad (5.11)$$

Furthermore, if the functions  $\mathcal{E}_n$  depend on electric charges  $q_{i,a}$  only through the DSZ products  $\gamma_{ij}$  and their kernels  $\Phi_n^\mathcal{E}(\mathbf{x})$  defined as in (4.9) are smooth solutions of Vignéras' equation,

$$V_{n-1} \cdot \Phi_n^\mathcal{E} = 0, \quad (5.12)$$

then the ansatz (5.8) also satisfies the modularity constraint (5.6).

This proposition allows in principle to fix all unknown functions. Indeed, the recursive relation (5.9) determines all  $g_n^{(0)}$ . Then equations (5.10) and (5.11) give the two parts of  $\mathcal{E}_n = \mathcal{E}_n^{(0)} + \mathcal{E}_n^{(+)}$  in terms of  $g_n^{(0)}$  and  $W_n$ . At this point the latter are still undetermined and are defined in terms of the unknown functions  $R_n$ . The additional constraint that  $\mathcal{E}_n$  should satisfy Vignéras' equation links together  $\mathcal{E}_n^{(0)}$  and  $\mathcal{E}_n^{(+)}$  and thus establishes a relation between  $W_n$  and  $g_n^{(0)}$ . Lastly, inverting (5.4) generates a solution for  $R_n$ .

We end this discussion by making an observation which will become relevant in the next subsection: by comparing (5.2) and (2.30), it is clear that  $\widehat{g}_n$  must agree with the tree index

$g_{\text{tr},n}$  in the limit  $\tau_2 \rightarrow \infty$ . In particular,  $g_{\text{tr},n}$  satisfies a similar ansatz as (5.8), upon replacing  $\mathcal{E}_n$  by its non-decaying part:

$$g_{\text{tr},n}(\{\check{\gamma}_i\}, z^a) = g_n^{(0)}(\{\check{\gamma}_i, c_i\}) - \text{Sym} \left\{ \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1, m < n}} g_{\text{tr},m}(\{\check{\gamma}'_k\}, z^a) \prod_{k=1}^m \mathcal{E}_{n_k}^{(0)} \right\}. \quad (5.13)$$

It follows that the function  $g_n^{(0)}$  should be independent of  $\tau_2$ , or at least that any such dependence should cancel in the recursion (5.13).

## 5.2 Generalized error functions and the completion

We see that the first step to be done is to provide an explicit expression for  $g_n^{(0)}$ . Once such an expression is known, all other functions can be determined algebraically. The problem, however, is that the solution of the recursive relation (5.9) is not unique. On the other hand,  $g_n^{(0)}$  determines the non-decaying part  $\mathcal{E}_n^{(0)}$  of  $\mathcal{E}_n$ , which is strongly constrained by the fact that  $\mathcal{E}_n$  must satisfy Vignéras' equation. This restriction turns out to be strong enough to remove the redundancy in the solution of (5.9), but it shows that we have to solve simultaneously two problems: satisfy the recursive relation (5.9) and ensure that it can be promoted to a solution of Vignéras' equation. We will do this by first constructing a solution of Vignéras' equation with the asymptotic part possessing the properties expected from  $g_n^{(0)}$ , and then proving that it does satisfy the recursive relation.

A hint towards a solution of these two problems can be found in the form of the kernel  $\Phi_n^f$  (4.12). The point is that both  $\Phi_n^f$  and  $g_n^{(0)}$  have discontinuities at walls of marginal stability which, as we know, must cancel each other. Furthermore, they should recombine into a smooth solution of Vignéras' equation. Thus,  $\Phi_n^f$  is expected to encode at least part of the completion of  $g_n^{(0)}$ . In addition, as explained in appendix D.2, the function  $\tilde{\Phi}_{n-1}^M$ , from which  $\Phi_n^f$  is constructed, appears as the term with fastest decay in the expansion of the function  $\tilde{\Phi}_{n-1}^E$  defined in (D.17) with the arguments  $\mathcal{V} = \{\mathbf{u}_e\}$ ,  $\check{\mathcal{V}} = \{\mathbf{v}_{s(e)t(e)}\}$ , which is a smooth solution of Vignéras' equation. This motivates us to consider the following function<sup>19</sup>

$$G_n(\{\check{\gamma}_i, c_i\}; \tau_2) = \frac{(-1)^{\sum_{i < j} \gamma_{ij}}}{(\sqrt{2\tau_2})^{n-1}} \tilde{\Phi}_n^{(0)}(\mathbf{x}), \quad (5.14)$$

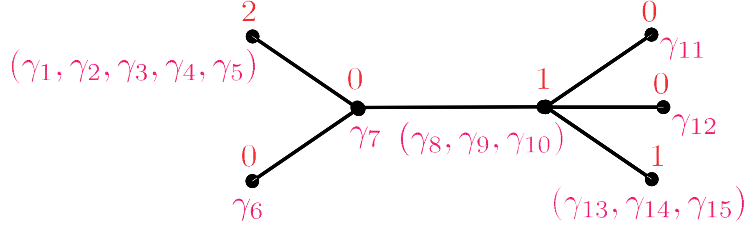
where  $\tilde{\Phi}_n^{(0)}(\mathbf{x})$  denotes the large  $\mathbf{x}$  limit of the function

$$\tilde{\Phi}_n(\mathbf{x}) = \frac{1}{2^{n-1} n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \tilde{\Phi}_{n-1}^E(\{\mathbf{u}_e\}, \{\mathbf{v}_{s(e)t(e)}\}; \mathbf{x}). \quad (5.15)$$

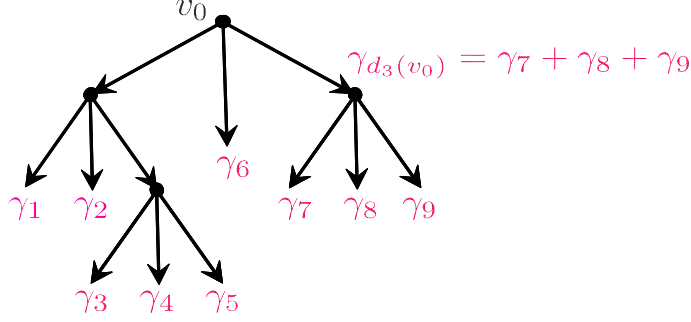
Our first goal will be to evaluate this limit explicitly.

To express the result, we need to introduce two new types of trees, beyond those already encountered. First, we denote by  $\mathbb{T}_{n,m}^\ell$  the set of *marked* unrooted labelled trees with  $n$  vertices and  $m$  marks assigned to vertices (see Fig. 5). Let  $m_{\mathbf{v}} \in \{0, \dots, m\}$  be the number of marks carried by the vertex  $\mathbf{v}$ , so that  $\sum_{\mathbf{v}} m_{\mathbf{v}} = m$ . Furthermore, the vertices are decorated by charges from the set  $\{\gamma_1, \dots, \gamma_{n+2m}\}$  such that a vertex  $\mathbf{v}$  with  $m_{\mathbf{v}}$  marks carries  $1 + 2m_{\mathbf{v}}$

<sup>19</sup>Note that the sign factor  $(-1)^{\sum_{i < j} \gamma_{ij}}$  is equal to the ratio of quadratic refinements appearing in (4.9).



**Figure 5:** An example of marked unrooted tree belonging to  $\mathbb{T}_{7,4}^\ell$  where near each vertex we showed in red an integer counting marks.



**Figure 6:** An example of rooted ternary tree belonging to  $\mathbb{T}_9^{(3)}$ .

charges  $\gamma_{v,s}$ ,  $s = 1, \dots, 1 + 2m_v$  and we set  $\gamma_v = \sum_{s=1}^{1+2m_v} \gamma_{v,s}$ . Second, we define  $\mathbb{T}_n^{(3)}$  to be the set of (unordered, full) rooted ternary trees with  $n$  leaves decorated by charges  $\gamma_i$  and other vertices carrying charges defined by the inductive rule  $\gamma_v = \gamma_{d_1(v)} + \gamma_{d_2(v)} + \gamma_{d_3(v)}$  where  $d_r(v)$  are the three children of vertex  $v$  (see Fig. 6). As usual for such trees,  $V_T$  denotes the set of vertices, with cardinality  $|V_T| = \frac{1}{2}(n - 1)$  (not counting the leaves).

In terms of these notations, we then have the following result proven in appendix F:

**Proposition 3.** *The function (5.14) is given by*

$$G_n(\{\tilde{\gamma}_i, c_i\}; \tau_2) = \frac{(-1)^{n-1+\sum_{i<j} \gamma_{ij}}}{2^{n-1}n!} \sum_{m=0}^{[(n-1)/2]} \frac{(-1)^m}{(4\pi\tau_2)^m} \sum_{T \in \mathbb{T}_{n-2m,m}^\ell} \prod_{v \in V_T} \mathcal{P}_{m_v}(\{p_{v,s}\}) \prod_{e \in E_T} \gamma_{s(e)t(e)} \text{sgn}(S_e), \quad (5.16)$$

where, for each edge  $e$  joining the subtrees  $\mathcal{T}_e^s$  and  $\mathcal{T}_e^t$ ,

$$S_e = \sum_{i \in V_{\mathcal{T}_e^s}} c_i(z^a) = \sum_{i \in V_{\mathcal{T}_e^s}} \sum_{j \in V_{\mathcal{T}_e^t}} \text{Im} [Z_{\gamma_i} \bar{Z}_{\gamma_j}], \quad (5.17)$$

and  $\mathcal{P}_m$  is the weight corresponding to the marks, given by

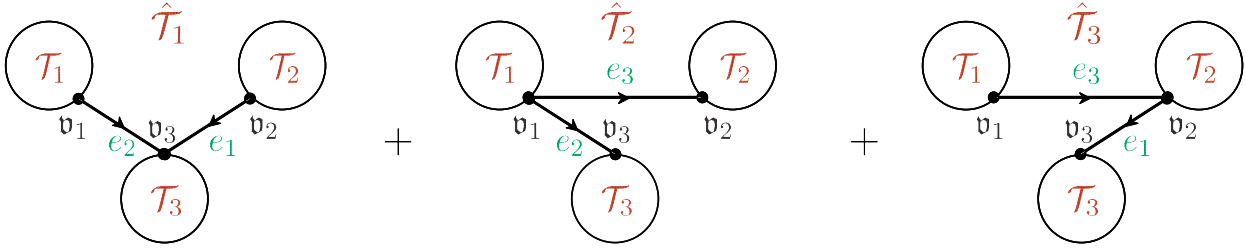
$$\mathcal{P}_m(\{p_s\}) = \sum_{T \in \mathbb{T}_{2m+1}^{(3)}} \frac{1}{T!} \prod_{v \in V_T} (p_{d_1(v)} p_{d_2(v)} p_{d_3(v)}). \quad (5.18)$$

Here  $T!$  is the tree factorial

$$T! = \prod_{v \in V_T} n_v(T), \quad (5.19)$$

where  $n_v(T)$ , as in Theorem 1, denotes the number of vertices (excluding the leaves) of the subtree of  $T$  rooted at  $v$ .





**Figure 7:** Three unrooted trees constructed from the same three subtrees.

To demystify the origin of these structures, note that the sum over  $m$  in (5.16) arises because of the contributions coming from the mutual action of covariant derivative operators  $\mathcal{D}(\mathbf{v}_{s(e)t(e)})$  (D.16) in the definition of the function  $\tilde{\Phi}_{n-1}^E$  (D.17). Such action is non-vanishing for any pair of intersecting edges  $(e_1, e_2)$  and is proportional to  $(p_{\mathbf{v}_1} p_{\mathbf{v}_2} p_{\mathbf{v}_3})$  where  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  are the three vertices belonging to the edges. While in appendix E it is shown that such contributions cancel in the sum over trees defining  $\Phi_n^J$  (4.12), this is not so for  $\tilde{\Phi}_n^{(0)}(\mathbf{x})$ : instead of the identity (E.5), one has to apply the sign identity (F.20) which produces a constant term. As a result, instead of the standard sign factors assigned to  $(e_1, e_2)$ , one finds the weight factor  $(p_{\mathbf{v}_1} p_{\mathbf{v}_2} p_{\mathbf{v}_3})$ . Furthermore, the sum over trees ensures that all other factors except for this weight depend on the charges  $\gamma_{\mathbf{v}_1}, \gamma_{\mathbf{v}_2}, \gamma_{\mathbf{v}_3}$  only through their sum. One can keep track of such contributions by collapsing the corresponding pairs of edges on the tree and marking the remaining vertices with weights  $m_{\mathbf{v}}$ . Essentially, the only non-trivial point here is to understand the form of the weight factor  $\mathcal{P}_m$  for  $m > 1$ . This case corresponds to the situation when more than one pair of joint edges collapse. The representation (5.18) in terms of rooted ternary trees is obtained by collapsing  $m$  pairs of edges successively, where the coefficient  $1/T!$  takes into account that such procedure leads to an overcounting of different assignments of labels.

Unfortunately, the function (5.16) cannot yet be taken as an ansatz for  $g_n^{(0)}$  because it depends non-trivially on the modulus  $\tau_2$ , which is in tension with the observation made at the end of the previous subsection. Therefore, we shall modify the function (5.15) into a function which is still a smooth solution of Vignéras' equation, but whose large  $\mathbf{x}$  limit is independent of  $\tau_2$ . Taking cue from the structure of (5.16), we define<sup>20</sup>

$$\tilde{\Phi}_n^{\mathcal{E}}(\mathbf{x}) = \frac{1}{2^{n-1} n!} \sum_{m=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2m, m}^{\ell}} \left[ \prod_{\mathbf{v} \in V_{\mathcal{T}}} \mathcal{D}_{m_{\mathbf{v}}}(\{\tilde{\gamma}_{\mathbf{v}, s}\}) \right] \tilde{\Phi}_{n-2m-1}^E(\{\mathbf{u}_e\}, \{\mathbf{v}_{s(e)t(e)}\}; \mathbf{x}), \quad (5.20)$$

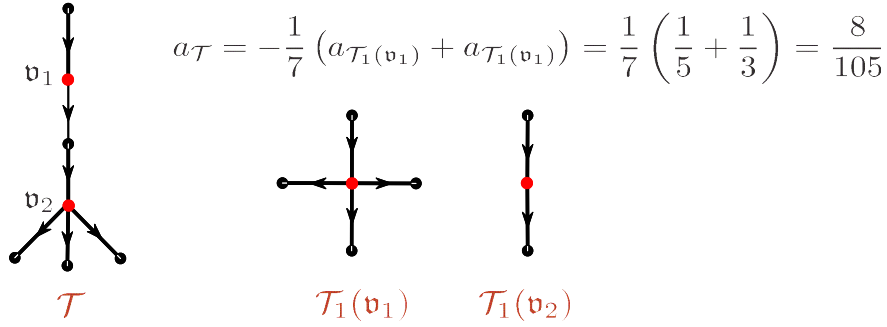
where

$$\mathcal{D}_m(\{\tilde{\gamma}_s\}) = \sum_{\mathcal{T}' \in \mathbb{T}_{2m+1}^{\ell}} a_{\mathcal{T}'} \prod_{e \in E_{\mathcal{T}'}} \mathcal{D}(\mathbf{v}_{s(e)t(v)}). \quad (5.21)$$

The coefficients  $a_{\mathcal{T}}$  are rational numbers which depend (up to a sign determined by the orientation of edges) only on the topology of the tree. We fix them by requiring that they satisfy the following system of equations

$$a_{\hat{\mathcal{T}}_1} + a_{\hat{\mathcal{T}}_2} - a_{\hat{\mathcal{T}}_3} = a_{\mathcal{T}_1} a_{\mathcal{T}_2} a_{\mathcal{T}_3}, \quad (5.22)$$

<sup>20</sup>The term  $m = 0$  in (5.20) reduces to the original function (5.15), while the terms with  $m > 0$  are the afore-mentioned modification.



**Figure 8:** An example of calculation of the coefficient  $a_{\mathcal{T}}$  for a tree with 7 vertices. We indicated in red the only vertices which produce non-vanishing contributions to the sum over vertices. For other vertices at least one of the trees  $\mathcal{T}_s(\mathbf{v})$  has even number of vertices and hence vanishing coefficient.

where  $\mathcal{T}_r$  ( $r = 1, 2, 3$ ) are arbitrary unrooted trees with marked vertex  $\mathbf{v}_r$  and  $\hat{\mathcal{T}}_r$  are the three trees constructed from  $\mathcal{T}_r$  as shown in Fig. 7.<sup>21</sup> This system of equations is imposed in order to ensure certain properties of the operators (5.21) which are crucial for the cancellation of  $\tau_2$ -dependent terms in the asymptotics of  $\tilde{\Phi}_n^{\mathcal{E}}$  (see (F.29)). It is readily seen that the system (5.22) allows to find all coefficients  $a_{\mathcal{T}}$  recursively starting from  $a_{\bullet} = 1$ ,  $a_{\bullet\bullet} = 0$  and going to trees with  $n \geq 3$  vertices. (Starting from  $n = 7$ , the system is overcomplete, but it can be checked that it does have a unique solution, with  $a_{\mathcal{T}} = 0$  for trees with even number of vertices.) For a generic tree, in appendix F we prove the following

**Proposition 4.** *For a tree  $\mathcal{T}$  with  $n$  vertices the coefficient  $a_{\mathcal{T}}$  is given recursively by*

$$a_{\mathcal{T}} = \frac{1}{n} \sum_{\mathbf{v} \in V_{\mathcal{T}}} \epsilon_{\mathbf{v}} \prod_{s=1}^{n_{\mathbf{v}}} a_{\mathcal{T}_s(\mathbf{v})}, \quad (5.23)$$

where  $n_{\mathbf{v}}$  is the valency of the vertex  $\mathbf{v}$ ,  $\mathcal{T}_s(\mathbf{v})$  are the trees obtained from  $\mathcal{T}$  by removing this vertex, and  $\epsilon_{\mathbf{v}}$  is the sign determined by the choice of orientation of edges,  $\epsilon_{\mathbf{v}} = (-1)^{n_{\mathbf{v}}^+}$  with  $n_{\mathbf{v}}^+$  being the number of incoming edges at the vertex. (See Fig. 8 for an example.)

Let us now return to the function (5.20), which is now fully specified given the prescription (5.23) for the coefficients  $a_{\mathcal{T}}$ . Similarly to (5.15) (which coincides with the  $m = 0$  contribution in (5.20)), it is a solution of Vignéras' equation for  $\lambda = n - 1$ . We claim that in the large  $\mathbf{x}$  asymptotics of this function, all  $\tau_2$ -dependent contributions, coming from the mutual action of derivative operators  $\mathcal{D}$ , cancel. More precisely, the asymptotics is given by the following

**Proposition 5.**

$$\lim_{\mathbf{x} \rightarrow \infty} \tilde{\Phi}_n^{\mathcal{E}}(\mathbf{x}) = \frac{1}{2^{n-1} n!} \sum_{m=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2m,m}^{\ell}} \prod_{\mathbf{v} \in V_{\mathcal{T}}} \mathcal{V}_{m_{\mathbf{v}}}(\{\tilde{\gamma}_{\mathbf{v},s}\}) \prod_{e \in E_{\mathcal{T}}} (\mathbf{v}_{s(e)t(e)}, \mathbf{x}) \operatorname{sgn}(\mathbf{u}_e, \mathbf{x}), \quad (5.24)$$

where

$$\mathcal{V}_m(\{\tilde{\gamma}_s\}) = \sum_{\mathcal{T} \in \mathbb{T}_{2m+1}^{\ell}} a_{\mathcal{T}} \prod_{e \in E_{\mathcal{T}}} (\mathbf{v}_{s(e)t(v)}, \mathbf{x}). \quad (5.25)$$

<sup>21</sup>For these equations, it is important to take into account the orientation of the edges: a change of orientation of an edge flips the sign of  $a_{\mathcal{T}}$ . The equations (5.22) are written assuming the orientation shown in Fig. 7, namely  $e_1 = (\mathbf{v}_2, \mathbf{v}_3)$ ,  $e_2 = (\mathbf{v}_1, \mathbf{v}_3)$ ,  $e_3 = (\mathbf{v}_1, \mathbf{v}_2)$ .

Importantly the function (5.24) is locally a homogeneous polynomial of degree  $n - 1$ . Therefore, when written in terms of charges, the  $\tau_2$ -dependence factorizes and is cancelled after the same rescaling as in (5.14). This naturally leads us to the following ansatz for the function  $g_n^{(0)}$ , which we prove in appendix F:

**Proposition 6.** *The function*

$$g_n^{(0)}(\{\check{\gamma}_i, c_i\}) = \frac{(-1)^{n-1+\sum_{i<j}\gamma_{ij}}}{2^{n-1}n!} \sum_{m=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2m,m}^\ell} \prod_{\mathbf{v} \in V_{\mathcal{T}}} \check{\mathcal{V}}_{m_{\mathbf{v}}}(\{\check{\gamma}_{\mathbf{v},s}\}) \prod_{e \in E_{\mathcal{T}}} \gamma_{s(e)t(e)} \operatorname{sgn}(S_e), \quad (5.26)$$

where

$$\check{\mathcal{V}}_m(\{\check{\gamma}_s\}) = \sum_{\mathcal{T}' \in \mathbb{T}_{2m+1}^\ell} a_{\mathcal{T}'} \prod_{e \in E_{\mathcal{T}'}} \gamma_{s(e)t(e)}, \quad (5.27)$$

satisfies the recursive equation (5.9).

Note that the same recursive equation (5.9) is also obeyed by the contribution of unmarked trees (i.e.  $m = 0$ ) to  $g_n^{(0)}$ , which we denote by  $g_n^*$  (see (F.34)). However, the latter cannot be obtained as the large  $\mathbf{x}$  limit of a solution of Vignéras' equation, which is why we have to consider the more complicated function (5.26).

Given the result for  $g_n^{(0)}$  and the relation (5.10), one obtains an explicit expression for the non-decreasing part of  $\mathcal{E}_n$ ,

$$\mathcal{E}_n^{(0)}(\{\check{\gamma}_i\}) = \frac{(-1)^{\sum_{i<j}\gamma_{ij}}}{2^{n-1}n!} \sum_{m=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2m,m}^\ell} \prod_{\mathbf{v} \in V_{\mathcal{T}}} \check{\mathcal{V}}_{m_{\mathbf{v}}}(\{\check{\gamma}_{\mathbf{v},s}\}) \prod_{e \in E_{\mathcal{T}}} \gamma_{s(e)t(e)} \operatorname{sgn}(\Gamma_e), \quad (5.28)$$

where

$$\Gamma_e = \sum_{i \in V_{\mathcal{T}_e^s}} \sum_{j \in V_{\mathcal{T}_e^t}} \gamma_{ij}. \quad (5.29)$$

From Proposition 2 we know that the kernel  $\Phi_n^{\mathcal{E}}$  corresponding to the full function  $\mathcal{E}_n$  must satisfy Vignéras' equation and have the asymptotics captured by (5.28). But we already know a solution of Vignéras' equation with a very similar asymptotics: this is the function (5.20) for which the asymptotics differs only by that the vectors  $\mathbf{v}_e$  are replaced by  $\mathbf{u}_e$ , which does not affect the proof of Vignéras' equation. Thus, we immediately arrive at the following result:

$$\mathcal{E}_n(\{\check{\gamma}_i\}; \tau_2) = \frac{(-1)^{\sum_{i<j}\gamma_{ij}}}{(\sqrt{2\tau_2})^{n-1}} \Phi_n^{\mathcal{E}}(\mathbf{x}), \quad (5.30)$$

where

$$\Phi_n^{\mathcal{E}}(\mathbf{x}) = \frac{1}{2^{n-1}n!} \sum_{m=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2m,m}^\ell} \left[ \prod_{\mathbf{v} \in V_{\mathcal{T}}} \mathcal{D}_{m_{\mathbf{v}}}(\{\check{\gamma}_{\mathbf{v},s}\}) \right] \check{\Phi}_{n-2m-1}^E(\{\mathbf{v}_e\}, \{\mathbf{v}_{s(e)t(e)}\}; \mathbf{x}). \quad (5.31)$$

In words, the function  $\mathcal{E}_n$  is a sum over marked unrooted labelled trees of solutions of Vignéras' equation with  $\lambda = n - 1$ , obtained from the standard generalized error functions  $\Phi_{n-2m-1}^E$  by acting with  $n - 1$  raising operators.

### 5.3 The structure of the completion

After substituting into the iterative ansatz (5.8), the two results (5.26) and (5.31) completely specify the coefficients of the expansion (5.2) of the generating function of DT invariants in terms of  $\widehat{h}_{p,\mu}$ . The result of the iteration can in fact be written explicitly as a sum over Schröder trees. Adopting the same notations as in Proposition 1, one has

**Proposition 7.** *The function  $\widehat{g}_n(\{\check{\gamma}_i, c_i\})$  is given by the sum over Schröder trees with  $n$  leaves,*

$$\widehat{g}_n = \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^S} (-1)^{n_T-1} (g_{v_0}^{(0)} - \mathcal{E}_{v_0}) \prod_{v \in V_T \setminus \{v_0\}} \mathcal{E}_v \right\}, \quad (5.32)$$

where  $v_0$  is the root of the tree and  $n_T$  is the number of its vertices (excluding the leaves).

We conclude that the properties of  $g_n^{(0)}$  and  $\mathcal{E}_n$  ensure that the theta series appearing in the corresponding decomposition of  $\mathcal{G}$  is modular so that the modularity of  $\mathcal{G}$  requires the function  $\widehat{h}_{p,\mu}$  to be a vector valued (non-holomorphic) modular form of weight  $(-b_2/2 - 1, 0)$ . The functions  $R_n$  entering the non-holomorphic completion  $\widehat{h}_{p,\mu}$ , are then given by the following proposition, whose proof can again be found in appendix F:

**Proposition 8.** *Inverting the relations (5.4) and (5.11), one obtains*

$$R_n = \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^S} (-1)^{n_T} \mathcal{E}_{v_0}^{(+)} \prod_{v \in V_T \setminus \{v_0\}} \mathcal{E}_v^{(0)} \right\}. \quad (5.33)$$

It is important to check that the resulting functions  $R_n$  are invariant under the spectral flow of the total charge. This is in fact a simple consequence of the fact that the functions  $\mathcal{E}_n$  entering their definition depend on the electric charges only through the DSZ products  $\gamma_{ij}$ . As follows from (2.23), all such products are invariant under the spectral flow of the total charge, hence  $R_n$  are invariant as well.

In appendix G we present explicit expressions for various elements of our construction up to order  $n = 4$ . Based on these results, we conjecture a general formula for the kernel  $\widehat{\Phi}_n^{\text{tot}}$  of the theta series appearing in the expansion of  $\mathcal{G}$  in terms of  $\widehat{h}_{p,\mu}$ :

**Conjecture 1.** *The generating function of multi-instantons has the theta series decomposition*

$$\mathcal{G} = \sum_{n=1}^{\infty} \frac{2^{-\frac{n}{2}}}{\pi \sqrt{2\tau_2}} \left[ \prod_{i=1}^n \sum_{p_i, \mu_i} \sigma_{p_i} \widehat{h}_{p_i, \mu_i} \right] e^{-S_p^{\text{cl}}} \vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\widehat{\Phi}_n^{\text{tot}}, n-2), \quad (5.34)$$

where the kernels of the theta series are given by

$$\widehat{\Phi}_n^{\text{tot}} = \Phi_1^f \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^S} (-1)^{n_T-1} \left( \widetilde{\Phi}_{v_0}^{\mathcal{E}} - \Phi_{v_0}^{\mathcal{E}} \right) \prod_{v \in V_T \setminus \{v_0\}} \Phi_v^{\mathcal{E}} \right\}. \quad (5.35)$$

Note that this result automatically implies that the theta series  $\vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\widehat{\Phi}_n^{\text{tot}}, n-2)$  is modular, since its kernel is a solution of Vignéras' equation.

We end the discussion of these results by observing that the formula (5.32) allows to obtain a new representation of the tree index  $g_{\text{tr},n}$ . Indeed, as observed at the end of Subsection 5.1,  $g_{\text{tr},n}$  agrees with  $\widehat{g}_n$  in the limit  $\tau_2 \rightarrow \infty$ , which amounts to replacing  $\mathcal{E}_v$  by its non-decaying part  $\mathcal{E}_v^{(0)}$  in (5.32). In fact, as shown in appendix F, all contributions due to trees with non-zero number of marks in (5.26) and (5.28) cancel in the resulting expression, leaving only the contributions coming from the terms with  $m = 0$ . Thus, one arrives at the following representation:

**Proposition 9.** *The tree index  $g_{\text{tr},n}$  defined in (2.3) can be expressed as*

$$g_{\text{tr},n} = \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^{\text{S}}} (-1)^{n_T-1} (g_{v_0}^* - \mathcal{E}_{v_0}^*) \prod_{v \in V_T \setminus \{v_0\}} \mathcal{E}_v^* \right\}, \quad (5.36)$$

where  $g_n^*$  is defined in (F.34) and  $\mathcal{E}_n^* = g_n^*(\{\check{\gamma}_i, \beta_{ni}\})$ .

Note that this representation is more explicit than the one given in section 2.2 since it does not require taking the limit  $y \rightarrow 1$ .

In fact, the mechanism by which contributions of marked trees cancel in the sum over Schröder trees, explained in the proof of Proposition 9, is very general and applies just as well to all the above equations (5.32)-(5.35). As a result, all of them remain valid if we replace  $\widetilde{\Phi}_n^{\mathcal{E}}$  and  $\Phi_n^{\mathcal{E}}$ , respectively, by  $\widetilde{\Phi}_n$  (5.15) and by its cousin with the vectors  $\{\mathbf{u}_e\}$  replaced by  $\{\mathbf{v}_e\}$ ,

$$\Phi_n(\mathbf{x}) = \frac{1}{2^{n-1}n!} \sum_{\mathcal{T} \in \mathbb{T}_n^{\mathcal{E}}} \widetilde{\Phi}_{n-1}^{\mathcal{E}}(\{\mathbf{v}_e\}, \{\mathbf{v}_{s(e)t(e)}\}; \mathbf{x}). \quad (5.37)$$

Correspondingly,  $g_n^{(0)}$  and  $\mathcal{E}_n^{(0)}$  should then be replaced by their asymptotics at large  $\tau_2$ , given by  $G_n(\{\check{\gamma}_i, c_i\})$  and  $G_n(\{\check{\gamma}_i, \beta_{ni}\})$ . The proof that such replacement is possible is completely analogous to the proof of Proposition 9, and we refrain from presenting it. This shows that from the very beginning we could take the function  $G_n$  as the ansatz for  $g_n^{(0)}$  since its  $\tau_2$ -dependence cancels at any rate in the formula for the tree index. We could then avoid most of complications of section 5.2. However, we cannot avoid the introduction of marked trees since they appear in the expression for  $G_n$  (5.16) anyway. Moreover, we prefer to stick to the definition (5.26) because of two reasons. Firstly, due to its  $\tau_2$ -independence,  $g_n^{(0)}$  might have its own interpretation as an index. Secondly, as we discuss in the next subsection, it is related to other potentially interesting representations, which might be at the basis of such interpretation.

## 5.4 Alternative representations

The construction presented above provides an explicit expression for the completion  $\widehat{h}_{p,\mu}$  and all other relevant quantities. Roughly, it can be split into three levels:

1. sum over Schröder trees (5.33);
2. sum over marked unrooted labelled trees (5.31);
3. sum over unrooted labelled trees defining the weights associated to marks (5.21).

The complication due to the appearance of marks and hence the last level arises due to the non-orthogonality of the vectors  $\mathbf{v}_{s(e)t(e)}$  appearing as arguments of the generalized error functions lifted to solutions of Vignéras' equation for  $\lambda = n - 1$ . However, the two sums over unrooted trees are organized in such a way that all additional contributions due to this non-orthogonality cancel. This suggests that there should exist a simpler representation where the vectors appearing in the second argument of functions  $\tilde{\Phi}_{n-1}^E(\mathcal{V}, \tilde{\mathcal{V}}; \mathbf{x})$  are mutually orthogonal from the very beginning. Below we present some results showing that such a representation does exist at least for low values of  $n$ . We then move on to present yet another representation of the functions  $g_n^{(0)}$  which is significantly simpler than (5.26), although its equivalence with the latter remains conjectural.

#### 5.4.1 Simplified representation via flow trees

We have shown, analytically for  $n = 2, 3, 4$  (see appendix G) and numerically for  $n = 5$ , that the function (5.26) can be rewritten as

$$g_n^{(0)}(\{\tilde{\gamma}_i, c_i\}) = \text{Sym} \left\{ \frac{c_n}{2^{n-1}} \sum_{T \in \mathbb{P}_n} \kappa(T) \prod_{k=1}^{n-1} \text{sgn}(S_k) \right\} \quad (n \leq 5), \quad (5.38)$$

where  $\kappa(T)$  is the factor defined in (2.6), whereas  $c_n$  and  $\mathbb{P}_n \subset \mathcal{T}_n^{\text{af}}$  are numerical coefficients and subsets of flow trees with  $n$  leaves, respectively, for which we found the following explicit results:

$$\begin{aligned} c_1 &= 1, & c_2 &= \frac{1}{2}, & c_3 &= \frac{1}{6}, & c_4 &= \frac{1}{12}, & c_5 &= \frac{1}{30}, \\ \mathbb{P}_1 &= \{(1)\}, & \mathbb{P}_2 &= \{(12)\}, & \mathbb{P}_3 &= \{((12)3), (1(23))\}, \\ \mathbb{P}_4 &= \{(((1(23))4), ((12)(34)), (1((23)4)))\}, \\ \mathbb{P}_5 &= \{((((31)4)2)5), (((12)(34))5), ((1((23)4))5), (((12)3)(45)), \\ & \quad ((12)(3(45))), (1((2(34))5)), (1((23)(45))), (1(4(2(53))))\}, \end{aligned} \quad (5.39)$$

where we labelled flow trees using the bracketings as explained below (2.3).

Unfortunately, the simple structure (5.38) fails beyond the fifth order. For instance, for  $n = 6$  one should add a term given by a product of  $n - 3 = 3$  sign functions. Nevertheless, these results suggest the following

**Conjecture 2.** *The function (5.26) can be rewritten as*

$$g_n^{(0)}(\{\tilde{\gamma}_i, c_i\}) = \text{Sym} \left\{ \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1, n_k \text{ odd}}} P_{n,m} \prod_{k=1}^{m-1} \text{sgn}(S_{j_k}) \right\}, \quad (5.40)$$

where as in (2.16)  $j_k = n_1 + \dots + n_k$  and  $P_{n,m}$  are polynomials in  $\kappa(\gamma_{ij})$  homogeneous of degree  $n - 1$  which can be represented as sums over subsets  $\mathbb{P}_{n,m}$  of flow trees with  $n$  leaves

$$P_{n,m} = \frac{1}{2^{n-1}} \sum_{T \in \mathbb{P}_{n,m}} c_T \kappa(T) \quad (5.41)$$

with some numerical coefficients  $c_T$ .

In fact, such a representation (if it exists) is highly ambiguous since there are many linear relations between polynomials  $\kappa(T)$  induced by the identity

$$\gamma_{12}\gamma_{1+2,3} + \gamma_{23}\gamma_{2+3,1} + \gamma_{31}\gamma_{1+3,2} = 0, \quad (5.42)$$

and even more relations after multiplication by sign functions and symmetrization. Eq. (5.38) suggests that there should exist a choice of subsets  $\mathbb{P}_{n,m}$  which leads to simple values of the coefficients  $c_T$ . However, we do not know of a procedure which would allow to determine it systematically beyond  $n = 5$ .

The main advantage of the representation (5.40) is that the factors  $\kappa(T)$  multiplying the products of sign functions are proportional to  $\prod_{v \in V_T} (\tilde{\mathbf{v}}_v, \mathbf{x})$  where the vectors  $\tilde{\mathbf{v}}_v$  defined in (4.14) are all mutually orthogonal. This property allows to immediately promote the kernel corresponding to the function (5.40), or to its value at the attractor point

$$\mathcal{E}_n^{(0)}(\{\tilde{\gamma}_i\}) = g_n^{(0)}(\{\tilde{\gamma}_i, \beta_{ni}\}) = \text{Sym} \left\{ \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1, n_k - \text{odd}}} P_{n,m} \prod_{k=1}^{m-1} \text{sgn}(\Gamma_{nj_k}) \right\} \quad (5.43)$$

to a solution of Vignéras' equation, without the complications that we had to deal with in subsection 5.2. Indeed, both such kernels are linear combinations of terms, labelled by flow trees, of the type (D.18) where the vectors  $\tilde{\mathbf{v}}_i$  coincide with the vectors  $\tilde{\mathbf{v}}_v$  for a given flow tree. Due to their mutual orthogonality, the solutions of Vignéras' equation with  $\lambda = n - 1$  with such asymptotics are given by

$$\Phi_n^{\mathcal{E}}(\mathbf{x}) = \frac{1}{2^{n-1}} \text{Sym} \left\{ \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1, n_k - \text{odd}}} \sum_{T \in \mathbb{P}_{n,m}} c_T \tilde{\Phi}_{m-1, n-1}^E(\{\mathbf{v}_{j_k}\}, \{\tilde{\mathbf{v}}_v\}; \mathbf{x}) \right\}, \quad (5.44)$$

and similarly for  $\tilde{\Phi}_n^{\mathcal{E}}$  with exchange of  $\mathbf{v}_{j_k}$  by  $\mathbf{u}_{j_k}$ . The vectors  $\mathbf{v}_\ell, \mathbf{u}_\ell$  defined in (5.7) can be seen as vectors  $\mathbf{v}_e, \mathbf{u}_e$  for the unrooted tree of trivial topology  $\bullet - \bullet - \dots - \bullet - \bullet$ . Thus, instead of a sum over marked unrooted trees with factors themselves given by another sum over unrooted trees as in (5.31), we arrive at a representation involving only a sum over a suitable subset of flow trees and a single unrooted tree.<sup>22</sup> It becomes particularly simple in the case  $n \leq 5$  where, as follows from (5.38), one can drop the sum over partitions, take  $m = n$  and equate all coefficients  $c_T$  to  $c_n$ .

### 5.4.2 Refinement

Finally, we note that there is yet an alternative way of obtaining the functions  $g_n^{(0)}$ , which also sheds light on the origin of the representation (5.40). This representation is inspired by the solution for the tree index presented in section 2.2. As in that case, the idea is to introduce a refinement with a parameter  $y$ , perform manipulations keeping  $y \neq 1$  and take the limit  $y \rightarrow 1$  in the end.

---

<sup>22</sup>Of course, the function (5.44) must be equal to (5.31), which is guaranteed by the fact that they are solutions of Vignéras' equation with the same asymptotics.

Let us introduce a ‘refined’ analogue of  $g_n^{(0)}$ , which we call  $g_n^{(\text{ref})}$  and which satisfies the refined version of the recursive relation (5.9) where the  $\kappa$ -factor is replaced by its  $y$ -dependent version (2.7). This refined equation can be solved by the following ansatz (cf. (2.10))

$$g_n^{(\text{ref})}(\{\check{\gamma}_i, c_i\}, y) = \frac{(-1)^{n-1+\sum_{i<j} \gamma_{ij}}}{(y-y^{-1})^{n-1}} \text{Sym} \left\{ F_n^{(\text{ref})}(\{\check{\gamma}_i, c_i\}) y^{\sum_{i<j} \gamma_{ij}} \right\}. \quad (5.45)$$

Indeed, it is easy to see that  $g_n^{(\text{ref})}$  satisfy (5.9) with the  $y$ -dependent  $\kappa$  provided  $F_n^{(\text{ref})}$  satisfy their own recursive relation

$$\begin{aligned} \frac{1}{2} \sum_{\ell=1}^{n-1} (\text{sgn}(S_\ell) - \text{sgn}(\Gamma_{n\ell})) F_\ell^{(\text{ref})}(\{\gamma_i, c_i^{(\ell)}\}_{i=1}^\ell) F_{n-\ell}^{(\text{ref})}(\{\gamma_i, c_i^{(\ell)}\}_{i=\ell+1}^n) \\ = F_n^{(\text{ref})}(\{\gamma_i, c_i\}) - F_n^{(\text{ref})}(\{\gamma_i, \beta_{ni}\}). \end{aligned} \quad (5.46)$$

This equation is much simpler because it is  $y$ -independent and does not involve any  $\kappa$ -factors. Furthermore, we already have one solution at our disposal — the functions  $F_n^{(0)}$  (2.17) which can be shown to satisfy (5.46) with help of the sign identity (F.37). However, this solution is not yet suitable because, starting from  $n = 3$ , its substitution into the r.h.s. of (5.45) does not produce a Laurent polynomial in  $y$ , but rather a rational function with a pole at  $y \rightarrow 1$ .

The solution can be promoted to a Laurent polynomial by noting that the recursive equation (5.46) determines  $F_n^{(\text{ref})}$  in terms of  $F_k^{(\text{ref})}$ ,  $k < n$  only up to a constant  $b_n$ . Thus, at each order one can adjust this constant so that to ensure that  $g_n^{(\text{ref})}$  given by (5.45) is regular at  $y = 1$ . It turns out that all  $b_n$  with  $n$  even can be taken to be zero.<sup>23</sup> Then we arrive at the following solution

$$F_n^{(\text{ref})}(\{\check{\gamma}_i, c_i\}) = \frac{1}{2^{n-1}} \sum_{\substack{n_1+\dots+n_m=n \\ n_k \geq 1, n_k \text{-odd}}} \prod_{k=1}^m b_{n_k} \prod_{k=1}^{m-1} \text{sgn}(S_{j_k}), \quad (5.47)$$

where as usual  $j_k = n_1 + \dots + n_k$ . It satisfies the recursion (5.46) for arbitrary coefficients  $b_n$ , as can be easily checked by extracting contributions with the same sets of  $\{n_k\}$  and applying the sign identity (F.37). The coefficients are fixed uniquely by requiring that the numerator of (5.45) vanishes at  $y = 1$ ,

$$\text{Sym} F_n^{(\text{ref})}(\{\check{\gamma}_i, c_i\}) = 0. \quad (5.48)$$

Note that under the permutation  $i \leftrightarrow n - i + 1$  one has  $S_i \rightarrow -S_{n-i}$  and therefore for  $n$  even  $F_n^{(\text{ref})}$  flips the sign. This automatically ensures (5.48) and is consistent with vanishing of  $b_{2i}$ . For  $n$  odd, to find  $b_n$  it is enough to consider only one particular configuration of the moduli, say  $c_i = 1$  for  $i = 1, \dots, n - 1$ . Then numerically one computes

$$b_1 = 1, \quad b_3 = -\frac{1}{3}, \quad b_5 = \frac{2}{15}, \quad b_7 = -\frac{17}{315}, \quad b_9 = \frac{62}{2835}, \quad b_{11} = -\frac{1382}{155925}, \quad \dots \quad (5.49)$$

---

<sup>23</sup>The reason for this is that the terms which are products of even and odd number of signs cannot talk to each other. On the other hand, there are sign identities like (D.12) which can decrease the number of signs in a product by even number. Thus, a constant can appear in  $F_n^{(\text{ref})}$  only for  $n$  odd. This is analogous to vanishing of the coefficients  $a_{\mathcal{T}}$  for trees with even number of vertices.



One observes that these coefficients coincide with the first terms in the series of Taylor coefficients of  $\tanh(x)$  for which the general formula reads

$$b_{n-1} = \frac{2^n(2^n - 1)}{n!} B_n, \quad (5.50)$$

where  $B_n$  are Bernoulli numbers. We conjecture that this formula provides all coefficients.

Using the iterative equation (5.9), the constraint (5.48) and proceeding by induction, it is easy to see that  $g_n^{(\text{ref})}$  defined by (5.45) and its first  $n - 2$  derivatives with respect to  $y$  vanish at  $y = 1$ . This ensures that  $g_n^{(\text{ref})}$  is smooth and its limit  $y \rightarrow 1$  is well-defined.

**Conjecture 3.** *The function  $g_n^{(\text{ref})}$ , defined by (5.45), (5.47) and (5.50), after removing refinement reproduces the function  $g_n^{(0)}$  (5.26),*

$$\lim_{y \rightarrow 1} g_n^{(\text{ref})}(\{\tilde{\gamma}_i, c_i\}, y) = g_n^{(0)}(\{\tilde{\gamma}_i, c_i\}). \quad (5.51)$$

We have checked this conjecture numerically up to  $n = 6$  evaluating the limit using the l'Hôpital rule. Unfortunately, the new representation of  $g_n^{(0)}$  obtained in this way is not really suitable for the purposes of this work because it does not lead to any new representation for the completion. Nevertheless, the comparison of (5.47) and (5.40) shows that the sum over partitions and the simple form of products of sign functions characterizing the representation discussed in the previous subsection find their origin in the formula for  $F_n^{(\text{ref})}$ . Furthermore, this construction strongly suggests that refinement may still preserve modularity and that a suitably defined generating function of refined DT invariants may also possess interesting modular properties, which can even be simpler than those of the usual DT invariants.

## 6. Discussion

In this paper we studied the modular properties of the generating function  $h_{p,\mu}(\tau)$  of MSW invariants encoding BPS degeneracies of D4-D2-D0 black holes in Type IIA string theory on a Calabi-Yau threefold, with fixed D4-charge  $p^a$ , D2-brane charge (up to spectral flow)  $\mu_a$  and invariant D0-brane charge  $\hat{q}_0$  (defined in (2.26)) conjugate to the modular parameter  $\tau$ . These properties follow from demanding that the vector multiplet moduli space in  $D = 3$  (or the hypermultiplet moduli space in the dual type IIB picture) admits an isometric action of  $SL(2, \mathbb{Z})$ . Our main result is an explicit formula for the non-holomorphic modular completion  $\hat{h}_{p,\mu}$  (5.1), where the functions  $R_n(\{\gamma_i\}, \tau)$  are given by Eq. (5.33), with  $\mathcal{E}_v = \mathcal{E}_v^{(0)} + \mathcal{E}_v^{(+)}$  specified by (5.28) and (5.31). This result applies for D4-branes wrapping a general effective divisor which may be the sum of an arbitrary number  $n$  of irreducible divisors.

The existence of such completion was not guaranteed. We arrived at this result by canceling the modular anomaly of an indefinite theta series of signature  $(n, nb_2(\mathfrak{Q}) - n)$  with a complicated kernel (4.3). In particular, Vignéras' equation encoding this anomaly involves Kähler moduli in a non-trivial way through the walls of marginal stability, whereas the functions  $R_n$  in our ansatz (5.1) can only depend (non-holomorphically) on the modular parameter  $\tau$  and on the charges  $\gamma_i$ .

The structure of the completion suggests that, for a divisor decomposable into a sum of  $n$  effective divisors, the holomorphic generating function  $h_{p,\mu}$  is a mixed vector valued mock

modular form of higher depth, equal to  $n - 1$ . Such objects have recently appeared in the mathematical literature [33, 58, 59, 60, 61] and correspond to holomorphic functions whose completion is constructed from period (or Eichler) integrals of mock modular forms but smaller depth (with depth 0 mock modular forms being synonymous with ordinary modular forms). For instance, in the case of standard mock modular forms of depth one, the completion is given by an Eichler integral of a modular function [22]. In our case this structure follows from the fact that the completion  $\widehat{h}_{p,\mu}$  is built from the generalized error functions which have a representation in terms of iterated period integrals [23, 33]. Note however that in order to make manifest the mixed mock modular nature of  $h_{p,\mu}$ , one should represent its completion as a sum of products of holomorphic modular forms times iterated period integral of anti-holomorphic modular forms. This is a non-trivial task which involves the technique of lattice factorization for indefinite theta series, which we leave for future research.

An important caveat in our construction is that we did not establish the convergence of the various generating functions and indefinite theta series. Technically, we constructed a theta series  $\vartheta_{p,\mu}(\widehat{\Phi}_n^{\text{tot}}, n - 2)$  whose kernel satisfies Vignéras' equation (D.3), but we did not demonstrate that it decays as required by Vignéras' criterium. (For  $n = 2$  this is easy to show [20] since the kernel is a product of an exponentially suppressed factor and a difference of two error functions defined by two positive vectors.) In fact, convergence issues already arise when expressing the generating function of DT invariants in terms of MSW invariants in (2.30), and are related to the problem of convergence of the BPS black hole partition function. The latter was proven to converge in the large volume limit for the case involving up to  $n \leq 3$  centers in [44]. We believe that the results for the tree index obtained in [19] will allow to extend this result to arbitrary  $n$ .

We note that many of the complications in our construction originate from the occurrence of polynomial factors in the kernel of theta series, which in turn can be traced to the factors of DSZ products  $\gamma_{ij}$  in the wall-crossing formula. Presumably, most of these complications would disappear if one could find a generalization of the contact potential involving the refined DT invariants, so that all these prefactors can be traded for explicit powers of the fugacity  $y$ , as in section 5.4.2. It is possible that twistorial topological strings [62] may provide the appropriate framework for finding such a refinement, at least in the context of string vacua on non-compact Calabi-Yau threefolds.

On the physics side, we expect that our results will have important implications for the physics of BPS black holes. Indeed, the mock modular property of the generating functions  $h_{p,\mu}$  will affect the growth of its Fourier coefficients [63], and should be taken into account when performing precision tests of the microscopic origin of the Bekenstein-Hawking entropy. Another outstanding question is to understand the physical significance of the instanton generating function  $\mathcal{G}$  defined in (3.22). Its modular properties are exactly those expected from the elliptic genus of a superconformal field theory, except for the fact that it is not holomorphic in  $\tau$ . It is natural to expect that this non-holomorphy can be traced to the existence of a continuum of states with a non-trivial spectral asymmetry between bosons and fermions. It would be very interesting to understand the origin of this continuum in terms of the worldvolume theory of an M5-brane wrapped on a reducible divisor. Moreover, for special  $\mathcal{N} = 2$  theories in  $D = 4$  obtained by circle compactification of an  $\mathcal{N} = 1$  theory in  $D = 5$ ,

$\mathcal{G}$  is closely related to the index considered in [31]. It would be interesting to understand the physical origin of its modular invariance in this context.

Finally, we expect that the structure found in the context of generalized DT-invariants on Calabi-Yau threefolds will also arise in the study of other types of BPS invariants where higher depth mock modular forms are expected to occur, such as Vafa-Witten invariants and Donaldson invariants of four-folds with  $b_2^+ = 1$  [23, 58]. In particular, it would be interesting to determine the modular completion of the generating function of Vafa-Witten invariants on the complex projective plane computed for all ranks in [64], generalizing the rank 3 case studied in [58], and compare with the completion found in the present paper.

## Acknowledgments

The authors are grateful to Sibasish Banerjee and Jan Manschot for useful discussions and collaboration on [20, 23, 21] which paved the way for the present work, and to Karen Yeats for useful communication about the combinatorics of rooted trees. The hospitality and financial support of the Theoretical Physics Department of CERN, where this work was initiated, is also gratefully acknowledged. The research of BP is supported in part by French state funds managed by the Agence Nationale de la Recherche (ANR) in the context of the LABEX ILP (ANR-11-IDEX-0004-02, ANR-10-LABX-63).

## A. Proof of the Theorem

In this appendix we prove Theorem 1 presented in the Introduction. The crucial ingredient is provided by the following Lemma<sup>24</sup>

**Lemma 1.** *The number of ways of labelling the vertices of a rooted ordered tree  $T$  with increasing labels is given by*

$$N_T = n_T! \prod_{v \in V_T} \frac{1}{n_v(T)}, \quad (\text{A.1})$$

where  $n_T$  is the number of vertices of the tree  $T$ .

*Proof.* First, let us find a recursive relation between the numbers  $N_T$ . Namely, consider a vertex  $v$  and the set of its children  $\text{Ch}(v)$ . Denote by  $T_v$  the tree rooted at  $v$  and with leaves coinciding with leaves of  $T$ , for which one has  $n_{T_v} = n_v$ . Then it is clear that

$$N_{T_v} = \frac{\left(\sum_{v' \in \text{Ch}(v)} n_{v'}\right)!}{\prod_{v' \in \text{Ch}(v)} n_{v'}!} \prod_{v' \in \text{Ch}(v)} N_{T_{v'}}. \quad (\text{A.2})$$

From this relation, one easily shows by recursion that (A.1) follows. Indeed, assuming that it holds for the subtrees  $T_{v'}$  and taking into account that  $\sum_{v' \in \text{Ch}(v)} n_{v'} = n_v - 1$ , the r.h.s. can be rewritten as

$$(n_v - 1)! \prod_{v' \in \text{Ch}(v)} \prod_{v'' \in V_{T_{v'}}} \frac{1}{n_{v''}} = n_v! \prod_{v' \in V_{T_v}} \frac{1}{n_{v'}}, \quad (\text{A.3})$$

---

<sup>24</sup>We were informed that this lemma appears as an exercise in [65, p. 70].

which coincides with (A.1) when  $v$  is the root of  $T$ .

□

Given this Lemma, we can now rewrite the l.h.s. of (1.2) as

$$\sum_{T' \subset T} \frac{N_{T'}}{m!} \prod_{v \in V_{T'}} n_v(T) = \frac{n!}{N_T} \sum_{T' \subset T} \frac{N_{T'}}{m!} \prod_{T_r \subset T \setminus T'} \frac{N_{T_r}}{n_{T_r}!}, \quad (\text{A.4})$$

where the last product goes over the subtrees which complement  $T'$  to the full tree  $T$ . On the other hand, it is easy to check that the relation (A.2) is the special case  $m = 1$  of a more general relation

$$N_T = \frac{\left(\sum_{T_r \subset T \setminus T'} n_{T_r}\right)!}{\prod_{T_r \subset T \setminus T'} n_{T_r}!} \sum_{T' \subset T} N_{T'} \prod_{T_r \subset T \setminus T'} N_{T_r}. \quad (\text{A.5})$$

where the sum runs over subtrees  $T'$  with  $m$  vertices. Taking into account that  $\sum_{T_r \subset T \setminus T'} n_{T_r} = n - m$  and substituting the resulting expression into (A.4), one recovers the binomial coefficient as stated in the Theorem.

## B. D3-instanton contribution to the contact potential

To evaluate  $(e^\phi)_{\text{D3}}$ , we first replace in (3.20) the full prepotential  $F$  by its classical, cubic part  $F^{\text{cl}}$  (2.21) and take into account that the sum over  $\gamma \in -\Gamma_+$  is complex conjugate to the sum over  $\gamma \in \Gamma_+$ . This gives

$$e^\phi \approx \frac{\tau_2^2}{12} ((\text{Im } u)^3) + \frac{i\tau_2}{8} \sum_{\gamma \in \Gamma_+} \text{Re} \int_{\ell_\gamma} \frac{dt}{t} (t^{-1} Z_\gamma(u^a) - t \bar{Z}_\gamma(\bar{u}^a)) H_\gamma. \quad (\text{B.1})$$

Next, we substitute the quantum corrected mirror map (3.8), change the integration variable  $t$  to  $z$ , and take the combined limit  $t^a \rightarrow \infty$ ,  $z \rightarrow 0$ . Keeping only the leading contributions, one obtains

$$\begin{aligned} (e^\phi)_{\text{D3}} &= -\frac{\tau_2}{2} \sum_{\gamma \in \Gamma_+} \text{Re} \int_{\ell_\gamma} dz \left[ \hat{q}_0 + \frac{1}{2} (q + b)^2 + 2(pt^2) z z_\gamma - \frac{3}{2} z^2 (pt^2) \right] H_\gamma \\ &\quad - \frac{1}{4} \sum_{\gamma_1, \gamma_2 \in \Gamma_+} (tp_1 p_2) \left( \text{Re} \int_{\ell_{\gamma_1}} dz_1 H_{\gamma_1} \right) \left( \text{Re} \int_{\ell_{\gamma_2}} dz_2 H_{\gamma_2} \right). \end{aligned} \quad (\text{B.2})$$

To further simplify this expression, we note that

$$\begin{aligned} 0 &= \frac{1}{4\pi} \sum_{\gamma \in \Gamma_+} \int_{\ell_\gamma} dz \partial_z (z H_\gamma) \\ &= \sum_{\gamma \in \Gamma_+} \int_{\ell_\gamma} dz \left[ \frac{1}{4\pi} + \tau_2 (pt^2) (z z_\gamma - z^2) - \frac{i}{4} \sum_{\gamma' \in \Gamma_+} \langle \gamma, \gamma' \rangle \int_{\ell_{\gamma'}} \frac{dz'}{z - z'} H_{\gamma'} \right] H_\gamma, \end{aligned} \quad (\text{B.3})$$

where we used the integral equation (3.15). Multiplying this identity by 3/4 and adding its real part to (B.2), one finds

$$\begin{aligned}
(e^\phi)_{\text{D3}} &= \frac{\tau_2}{2} \sum_{\gamma \in \Gamma_+} \text{Re} \int_{\ell_\gamma} dz a_{\gamma, -\frac{3}{2}}(z) H_\gamma \\
&\quad - \frac{1}{4} \sum_{\gamma_1, \gamma_2 \in \Gamma_+} \left[ (tp_1 p_2) \left( \text{Re} \int_{\ell_{\gamma_1}} dz_1 H_{\gamma_1} \right) \left( \text{Re} \int_{\ell_{\gamma_2}} dz_2 H_{\gamma_2} \right) \right. \\
&\quad \left. + \frac{3}{4} \text{Re} \left( \int_{\ell_{\gamma_1}} dz_1 H_{\gamma_1} \int_{\ell_{\gamma_2}} dz_2 H_{\gamma_2} \frac{i\langle \gamma, \gamma' \rangle}{z - z'} \right) \right], \tag{B.4}
\end{aligned}$$

where we introduced

$$a_{\gamma, \mathfrak{h}}(z) = - \left( \hat{q}_0 + \frac{1}{2} (q + b)^2 + \frac{1}{2} (pt^2) z z_\gamma + \frac{\mathfrak{h}}{4\pi\tau_2} \right). \tag{B.5}$$

The meaning of this function is actually very simple: it gives the action of the modular covariant derivative operator  $\mathcal{D}_\mathfrak{h}$  (3.23) on the classical part of the Darboux coordinate (3.12),

$$\mathcal{D}_\mathfrak{h} \mathcal{X}_\gamma^{\text{cl}} = a_{\gamma, \mathfrak{h}}(z) \mathcal{X}_\gamma^{\text{cl}}. \tag{B.6}$$

Combining this fact with equation (3.15), it is easy to check that the function (3.22) satisfies

$$\begin{aligned}
\mathcal{D}_\mathfrak{h} \mathcal{G} &= \sum_{\gamma \in \Gamma_+} \int_{\ell_\gamma} dz a_{\gamma, \mathfrak{h}}(z) H_\gamma \\
&\quad + \frac{1}{8\tau_2} \sum_{\gamma_1, \gamma_2 \in \Gamma_+} \int_{\ell_{\gamma_1}} dz_1 \int_{\ell_{\gamma_2}} dz_2 \left[ (tp_1 p_2) + 2\mathfrak{h} \left( (tp_1 p_2) + \frac{i\langle \gamma_1, \gamma_2 \rangle}{z_1 - z_2} \right) \right] H_{\gamma_1}(z_1) H_{\gamma_2}(z_2), \tag{B.7}
\end{aligned}$$

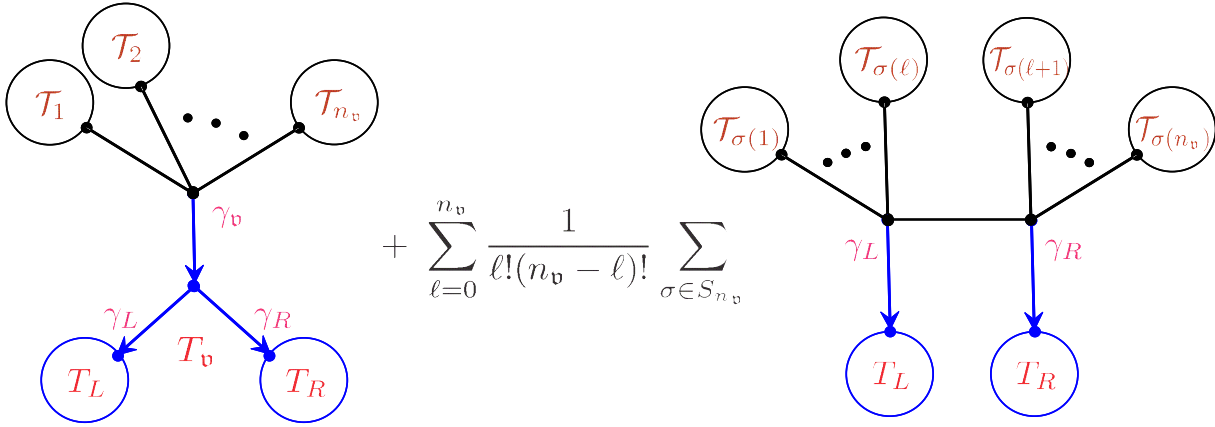
$$\partial_{\tilde{c}_a} \mathcal{G} = 2\pi i \sum_{\gamma \in \Gamma_+} p^a \int_{\ell_\gamma} dz H_\gamma, \tag{B.8}$$

which allows to rewrite (B.4) exactly as in (3.24).

## C. Smoothness of the instanton generating function

In this appendix we prove the smoothness of the function  $\mathcal{G}$  across walls of marginal stability. The starting point is the representation (4.2) where the potential discontinuities are hidden in the kernel of the theta series (4.3). Given that  $\Phi_m^f$  can be represented as a sum over unrooted labelled trees (see (4.6)), whereas  $\Phi_{n_k}^g$  appear as sums over attractor flow trees (see (4.15)), we can view the kernel  $\Phi_n^{\text{tot}}$  as a sum over ‘blooming trees’ which are unrooted trees with a flow tree (the ‘flower’) growing from each vertex. Then the idea is that the discontinuities due to flow trees (i.e. due to DT invariants) of a blooming tree with  $m$  vertices are cancelled by the discontinuities due to exchange of integration contours in  $\Phi_{m+1}^f$  corresponding to blooming trees with  $m + 1$  vertices.

Following this idea, let us consider a tree which has an attractor flow tree  $T_\mathfrak{v}$  growing from vertex  $\mathfrak{v}$  (see Fig. 9). We denote by  $\mathcal{T}_i$ ,  $i = 1, \dots, n_\mathfrak{v}$ , the blooming subtrees connected to this vertex and by  $T_L$  and  $T_R$  the two parts (which may be trivial) of  $T_\mathfrak{v}$  with the total



**Figure 9:** Combination of trees showing cancellation of discontinuities across the wall of marginal stability corresponding to the decay  $\gamma_{\mathbf{v}} \rightarrow \gamma_L + \gamma_R$ . The parts corresponding to attractor flow trees are drawn in blue.

charges  $\gamma_L$  and  $\gamma_R$  so that  $\gamma_{\mathbf{v}} = \gamma_L + \gamma_R$ . Together with the contribution of this tree, we consider the contributions of the trees obtained by splitting the vertex  $\mathbf{v}$  into two vertices  $\mathbf{v}_L$  and  $\mathbf{v}_R$  connected by an edge, carrying charges  $\gamma_L$  and  $\gamma_R$  and all possible allocations of the subtrees  $\mathcal{T}_i$  to these two vertices. Different allocations are accounted for by the sum over permutations, where the weight  $\frac{1}{\ell!(n_{\mathbf{v}}-\ell)!}$  takes into account the fact that permutations between subtrees connected to one vertex are redundant. The attractor flow trees  $T_L$  and  $T_R$  are then connected to  $\mathbf{v}_L$  and  $\mathbf{v}_R$ , respectively, as shown in the picture.

The contribution corresponding to the first blooming tree has a discontinuity at the wall of marginal stability for the bound state  $\gamma_L + \gamma_R$  and originating from the factor  $\Delta_{\gamma_L \gamma_R}^z$  assigned to the root vertex of  $T_{\mathbf{v}}$ . The other contributions have discontinuities at the same wall due to the exchange of the contours  $\ell_{\gamma_L}$  and  $\ell_{\gamma_R}$  for the integrals assigned to  $\mathbf{v}_L$  and  $\mathbf{v}_R$ , respectively. They are given by the residues at the pole of the integration kernel  $K_{\gamma_L \gamma_R}$ . It is clear that the structure of all jumps is very similar since different subtrees produce essentially the same weights. Let us analyze which differences they may have.

- First, the contributions of the flow trees  $T_L$  and  $T_R$  could differ in the two cases because they have different starting points for the attractor flows: for the trees on the right side of Fig. 9 this is  $z^a \in \mathcal{M}_K$ , whereas for the tree on the left this is the point on the wall for  $\gamma_L + \gamma_R$  reached by the flow from  $z^a$ . But we are evaluating the discontinuity exactly on the wall where the two points coincide. Thus, the contributions are the same.
- Although the subtrees  $\mathcal{T}_i$  give rise to the same contributions for all blooming trees shown in Fig. 9, the contributions of the edges connecting them to either  $\mathbf{v}$  or  $\mathbf{v}_L$ ,  $\mathbf{v}_R$  are not exactly the same. Each of them contributes the factor given by the kernel (3.16). After taking the residue, the  $z$ -dependence of these kernels for the trees shown on the left and the right sides of the picture is identical. However, their charge dependence is different: for the tree on the left they depend on  $\gamma_{\mathbf{v}}$ , whereas for the trees on the right they depend on  $\gamma_L$  or  $\gamma_R$ , depending on which vertex they are connected to. But it is easy to see that the sums over  $\ell$  and permutations produce the standard binomial expansion of a

single product of kernels which all depend on  $\gamma_L + \gamma_R = \gamma_{\mathbf{v}}$  and thus coinciding with the contribution of the tree on the left.

- Finally, one should take into account that the discontinuity of  $\Delta_{\gamma_L \gamma_R}^z$  in the first contribution gives the factor  $\langle \gamma_L, \gamma_R \rangle$ . But exactly the same factor arises as the residue of  $K_{\gamma_L \gamma_R}$  corresponding to the additional edge.

Thus, it remains only to check that all numerical factors work out correctly. Except the factors which are common to both contributions, we have

$$-\frac{(-1)^{\langle \gamma_L, \gamma_R \rangle}}{2} \frac{\sigma_{\gamma_{\mathbf{v}}}}{(2\pi)^2} \frac{2m}{m!} + (2\pi i)(-2\pi i) \frac{\sigma_{\gamma_L} \sigma_{\gamma_R}}{(2\pi)^4} \frac{m(m+1)}{(m+1)!}. \quad (\text{C.1})$$

Here  $-\frac{(-1)^{\langle \gamma_L, \gamma_R \rangle}}{2}$  comes from the factor  $-\Delta_{\gamma_L \gamma_R}^z$  in  $T_{\mathbf{v}}$ , the factors with quadratic refinement are due to functions  $H_{\gamma}$  (3.9) assigned to  $\mathbf{v}$  or  $\mathbf{v}_L, \mathbf{v}_R$ ,  $(2\pi i)$  is the standard weight of the residue,  $(-2\pi i)$  is the residue of  $K_{\gamma_L \gamma_R}$  (3.16), and factorials are the weights of the trees in the expansion (3.26). Finally, the factors  $2m$  and  $m(m+1)$  arise due to the freedom to relabel charges assigned to the marked vertices: on the left these are vertex  $\mathbf{v}$  and the two children of the root in  $T_{\mathbf{v}}$ , whereas on the right these are  $\mathbf{v}_L$  and  $\mathbf{v}_R$ . It is immediate to check that all these numerical weights cancel, which ensures that the function  $\mathcal{G}$  is continuous across walls of marginal stability. Moreover, in this cancellation the condition that we sit on the wall was used only in locally constant factors. Therefore, it proves not only that  $\mathcal{G}$  is continuous, but that it is actually smooth around these loci.

## D. Indefinite theta series and generalized error functions

### D.1 Vignéras' theorem

Let  $\Lambda$  be a  $d$ -dimensional lattice equipped with a bilinear form  $(\mathbf{x}, \mathbf{y}) \equiv \mathbf{x} \cdot \mathbf{y}$ , where  $\mathbf{x}, \mathbf{y} \in \Lambda \otimes \mathbb{R}$ , such that its associated quadratic form has signature  $(n, d-n)$  and is integer valued, i.e.  $\mathbf{k}^2 \equiv \mathbf{k} \cdot \mathbf{k} \in \mathbb{Z}$  for  $\mathbf{k} \in \Lambda$ . Furthermore, let  $\mathbf{p} \in \Lambda$  be a characteristic vector (such that  $\mathbf{k} \cdot (\mathbf{k} + \mathbf{p}) \in 2\mathbb{Z}, \forall \mathbf{k} \in \Lambda$ ),  $\boldsymbol{\mu} \in \Lambda^*/\Lambda$  a glue vector, and  $\lambda$  an arbitrary integer. We consider the following family of theta series

$$\vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi, \lambda; \tau, \mathbf{b}, \mathbf{c}) = \tau_2^{-\lambda/2} \sum_{\mathbf{k} \in \Lambda + \boldsymbol{\mu} + \frac{1}{2}\mathbf{p}} (-1)^{\mathbf{k} \cdot \mathbf{p}} \Phi(\sqrt{2\tau_2}(\mathbf{k} + \mathbf{b})) \mathbf{e}\left(-\frac{\tau}{2}(\mathbf{k} + \mathbf{b})^2 + \mathbf{c} \cdot (\mathbf{k} + \frac{1}{2}\mathbf{b})\right) \quad (\text{D.1})$$

defined by a kernel  $\Phi(\mathbf{x})$  such that the function  $f(\mathbf{x}) \equiv \Phi(\mathbf{x}) e^{\frac{\pi}{2}\mathbf{x}^2} \in L^1(\Lambda \otimes \mathbb{R})$  so that the sum is absolutely convergent. Irrespective of the choice of this kernel and of the parameter  $\lambda$ , any such theta series satisfies the following elliptic properties

$$\begin{aligned} \vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi, \lambda; \tau, \mathbf{b} + \mathbf{k}, \mathbf{c}) &= (-1)^{\mathbf{k} \cdot \mathbf{p}} \mathbf{e}\left(-\frac{1}{2}\mathbf{c} \cdot \mathbf{k}\right) \vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi, \lambda; \tau, \mathbf{b}, \mathbf{c}), \\ \vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi, \lambda; \tau, \mathbf{b}, \mathbf{c} + \mathbf{k}) &= (-1)^{\mathbf{k} \cdot \mathbf{p}} \mathbf{e}\left(\frac{1}{2}\mathbf{b} \cdot \mathbf{k}\right) \vartheta_{\mathbf{p}, \boldsymbol{\mu}}(\Phi, \lambda; \tau, \mathbf{b}, \mathbf{c}). \end{aligned} \quad (\text{D.2})$$

Now let us require that in addition the kernel satisfies the following two conditions:

1. Let  $D(\mathbf{x})$  be any differential operator of order  $\leq 2$ , and  $R(\mathbf{x})$  any polynomial of degree  $\leq 2$ . Then  $f(\mathbf{x})$  defined above must be such that  $f(\mathbf{x})$ ,  $D(\mathbf{x})f(\mathbf{x})$  and  $R(\mathbf{x})f(\mathbf{x}) \in L^2(\Lambda \otimes \mathbb{R}) \cap L^1(\Lambda \otimes \mathbb{R})$ .

2.  $\Phi(\mathbf{x})$  must satisfy

$$V_\lambda \cdot \Phi(\mathbf{x}) = 0, \quad V_\lambda = \partial_{\mathbf{x}}^2 + 2\pi(\mathbf{x} \cdot \partial_{\mathbf{x}} - \lambda). \quad (\text{D.3})$$

Then in [32] it was proven that the theta series (D.1) transforms as a vector-valued modular form of weight  $(\lambda + d/2, 0)$  (see Theorem 2.1 in [23] for the detailed transformation under  $\tau \rightarrow -1/\tau$ ). We refer to  $V_\lambda$  as Vignéras' operator. The simplest example is the Siegel theta series for which the kernel is  $\Phi(\mathbf{x}) = e^{-\pi \mathbf{x}_+^2}$  where  $\mathbf{x}_+$  is the projection of  $\mathbf{x}$  on a fixed positive plane of dimension  $n$ . This kernel is annihilated by  $V_{-n}$ .

In this paper we apply the Vignéras' theorem to the case of  $\Lambda = \bigoplus_{i=1}^n \Lambda_i$ . Thus, the charges appearing in the description of the theta series (D.1) are of the type  $\mathbf{k} = (k_1^a, \dots, k_n^a)$ , whereas the vectors  $\mathbf{b}$  and  $\mathbf{c}$  are taken with  $i$ -independent components, namely,  $\mathbf{b}_i^a = b^a$ ,  $\mathbf{c}_i^a = c^a$  for  $i = 1, \dots, n$ . The lattices  $\Lambda_i$  carry the bilinear forms  $\kappa_{i,ab} = \kappa_{abc} p_i^c$  which are all of signature  $(1, b_2 - 1)$ . This induces a natural bilinear form on  $\Lambda$ :

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n (p_i x_i y_i). \quad (\text{D.4})$$

Note also that the sign factor  $(-1)^{\mathbf{k} \cdot \mathbf{p}}$  in (D.1) can be traced back to the quadratic refinement provided we choose it as

$$\sigma_\gamma = \sigma_{p,q} \equiv \mathbf{e} \left( \frac{1}{2} p^a q_a \right) \sigma_p, \quad \sigma_p = \mathbf{e} \left( \frac{1}{2} A_{ab} p^a p^b \right). \quad (\text{D.5})$$

It involves the matrix  $A_{ab}$  satisfying

$$A_{ab} p^p - \frac{1}{2} \kappa_{abc} p^b p^c \in \mathbb{Z} \quad \text{for } \forall p^a \in \mathbb{Z}. \quad (\text{D.6})$$

This matrix appears due to the non-trivial quantization of charges on the type IIB side (2.18) and can be used to perform a symplectic rotation to identify them with mirror dual integer charges on the type IIA side [45]. It is easy to check that the quadratic refinement (D.5) satisfies (3.4).

## D.2 Generalized error functions

An important class of solutions of Vignéras' equation is given by the error function and its generalizations constructed in [23] and further elaborated in [24]. Let us take

$$M_1(u) = -\text{sgn}(u) \text{Erfc}(|u|\sqrt{\pi}) = \frac{i}{\pi} \int_{\ell} \frac{dz}{z} e^{-\pi z^2 - 2\pi i z u}, \quad (\text{D.7})$$

$$\begin{aligned} E_1(u) &= \text{sgn}(u) + M_1(u) \\ &= \text{Erf}(u\sqrt{\pi}) = \int_{\mathbb{R}} du' e^{-\pi(u-u')^2} \text{sgn}(u'), \end{aligned} \quad (\text{D.8})$$



where the contour  $\ell = \mathbb{R} - iu$  runs parallel to the real axis through the saddle point at  $z = -iu$ . Then, given a vector with a positive norm  $\mathbf{v}^2 > 0$  so that  $|\mathbf{v}| = \sqrt{\mathbf{v}^2}$ , we define

$$\Phi_1^E(\mathbf{v}; \mathbf{x}) = E_1\left(\frac{\mathbf{v} \cdot \mathbf{x}}{|\mathbf{v}|}\right), \quad \Phi_1^M(\mathbf{v}; \mathbf{x}) = M_1\left(\frac{\mathbf{v} \cdot \mathbf{x}}{|\mathbf{v}|}\right). \quad (\text{D.9})$$

It is easy to check that the first function is a smooth solution of (D.3) with  $\lambda = 0$ , whereas the second is exponentially suppressed at large  $\mathbf{x}$  and also solves the same equation, but only away from the locus  $\mathbf{v} \cdot \mathbf{x} = 0$  where it has a discontinuity.

Generalizing the integral representations (D.7) and (D.8), we define

$$M_n(\mathcal{M}; \mathbf{u}) = \left(\frac{i}{\pi}\right)^n |\det \mathcal{M}|^{-1} \int_{\mathbb{R}^n - i\mathbf{u}} d^n z \frac{e^{-\pi \mathbf{z}^{\text{tr}} \mathbf{z} - 2\pi i \mathbf{z}^{\text{tr}} \mathbf{u}}}{\prod(\mathcal{M}^{-1} \mathbf{z})}, \quad (\text{D.10})$$

$$E_n(\mathcal{M}; \mathbf{u}) = \int_{\mathbb{R}^n} d\mathbf{u}' e^{-\pi(\mathbf{u}-\mathbf{u}')^{\text{tr}}(\mathbf{u}-\mathbf{u}')} \text{sgn}(\mathcal{M}^{\text{tr}} \mathbf{u}'), \quad (\text{D.11})$$

where  $\mathbf{z} = (z_1, \dots, z_n)$  and  $\mathbf{u} = (u_1, \dots, u_n)$  are  $n$ -dimensional vectors,  $\mathcal{M}$  is  $n \times n$  matrix of parameters, and we used the shorthand notations  $\prod \mathbf{z} = \prod_{i=1}^n z_i$  and  $\text{sgn}(\mathbf{u}) = \prod_{i=1}^n \text{sgn}(u_i)$ . The detailed properties of these functions can be found in [24]. Here we mention only a few:

- $M_n$  are exponentially suppressed for large  $\mathbf{u}$  as  $M_n \sim \frac{(-1)^n}{\pi^n} |\det \mathcal{M}|^{-1} \frac{e^{-\pi \mathbf{u}^{\text{tr}} \mathbf{u}}}{\prod(\mathcal{M}^{-1} \mathbf{u})}$ , whereas  $E_n$  are locally constant for large  $\mathbf{u}$  as  $E_n \sim \text{sgn}(\mathcal{M}^{\text{tr}} \mathbf{u})$ .
- More generally,  $E_n$  can be expressed as a linear combination of  $M_k$ ,  $k = 0, \dots, n$ , multiplied by  $n - k$  sign functions, generalizing the first relation in (D.8) (see Eq. (D.15) below for a precise statement).
- From (D.11) it follows that every identity between products of sign functions implies an identity between generalized error functions  $E_n$ . Moreover, expanding the  $E_n$  functions in terms of  $M_k$ 's and sign functions, one obtains similar identities for functions  $M_n$ . For instance, the identity

$$(\text{sgn}(x_1) + \text{sgn}(x_2)) \text{sgn}(x_1 + x_2) = 1 + \text{sgn}(x_1) \text{sgn}(x_2) \quad (\text{D.12})$$

implies

$$\begin{aligned} E_2((\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2); \mathbf{u}) + E_2((\mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2); \mathbf{u}) &= 1 + E_2((\mathbf{v}_1, \mathbf{v}_2); \mathbf{u}), \\ M_2((\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2); \mathbf{u}) + M_2((\mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2); \mathbf{u}) &= M_2((\mathbf{v}_1, \mathbf{v}_2); \mathbf{u}), \end{aligned} \quad (\text{D.13})$$

where  $\mathbf{v}_1, \mathbf{v}_2$  are two-dimensional vectors used to encode the  $2 \times 2$  matrix of parameters.

The main reason to introduce these functions is that, similarly to the usual error and complementary error functions, they can be used to produce solutions of Vignéras' equation on  $\mathbb{R}^{n, d-n}$ . To write them down, let us consider  $d \times n$  matrix  $\mathcal{V}$  which can be viewed as a collection of  $n$  vectors,  $\mathcal{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ . We assume that these vectors span a positive definite subspace, i.e.  $\mathcal{V}^{\text{tr}} \cdot \mathcal{V}$  is a positive definite matrix. Let  $\mathcal{B}$  be  $n \times d$  matrix whose rows define an orthonormal basis for this subspace. Then we define the *boosted* generalized error functions

$$\Phi_n^M(\mathcal{V}; \mathbf{x}) = M_n(\mathcal{B} \cdot \mathcal{V}; \mathcal{B} \cdot \mathbf{x}), \quad \Phi_n^E(\mathcal{V}; \mathbf{x}) = E_n(\mathcal{B} \cdot \mathcal{V}; \mathcal{B} \cdot \mathbf{x}). \quad (\text{D.14})$$

It can be shown that both these functions satisfy Vignéras' equation (for  $\Phi_n^M$  one should stay away from its discontinuities, i.e. loci where  $\text{sgn}((\mathcal{B} \cdot \mathcal{V})^{-1} \mathcal{B} \cdot \mathbf{x}) = 0$ ). Since at large  $\mathbf{x}$  one has  $\Phi_n^E \sim \text{sgn}(\mathcal{V}^{\text{tr}} \cdot \mathbf{x}) = \prod_{i=1}^n \text{sgn}(\mathbf{v}_i \cdot \mathbf{x})$ , one can think about this function as providing the modular completion for (indefinite) theta series with kernel given by a product of signs.

The relation between functions  $E_n$  and  $M_n$  mentioned above implies a similar relation between the functions (D.14). For generic  $n$ , it takes the following form<sup>25</sup>

$$\Phi_n^E(\mathcal{V}; \mathbf{x}) = \sum_{\mathcal{I} \subseteq \mathcal{Z}_n} \Phi_{|\mathcal{I}|}^M(\{\mathbf{v}_i\}_{i \in \mathcal{I}}; \mathbf{x}) \prod_{j \in \mathcal{Z}_n \setminus \mathcal{I}} \text{sgn}(\mathbf{v}_{j \perp \mathcal{I}}, \mathbf{x}), \quad (\text{D.15})$$

where the sum goes over all possible subsets (including the empty set) of the set  $\mathcal{Z}_n = \{1, \dots, n\}$ ,  $|\mathcal{I}|$  is the cardinality of  $\mathcal{I}$ , and  $\mathbf{v}_{j \perp \mathcal{I}}$  denotes the projection of  $\mathbf{v}_j$  orthogonal to the subspace spanned by  $\{\mathbf{v}_i\}_{i \in \mathcal{I}}$ . The cardinality  $|\mathcal{I}|$  can also be interpreted as the number of directions in  $\mathbb{R}^d$  along which the corresponding contribution has an exponential fall off.

Finally, note that a solution of Vignéras' equation with a given  $\lambda$  can be upgraded to a solution with  $\lambda + 1$  by the action of the differential operator

$$\mathcal{D}(\mathbf{v}) = \mathbf{v} \cdot \left( \mathbf{x} + \frac{1}{2\pi} \partial_{\mathbf{x}} \right), \quad (\text{D.16})$$

which realizes the action of the covariant derivative raising the holomorphic weight by 1. In particular, we can construct solutions with  $\lambda = m$  which behave for large  $\mathbf{x}$  as products of  $n$  sign functions. To this end, it is enough to act on  $\Phi_n^E$  by this operator  $m$  times. Thus, we define

$$\tilde{\Phi}_{n,m}^E(\mathcal{V}, \tilde{\mathcal{V}}; \mathbf{x}) = \left[ \prod_{i=1}^m \mathcal{D}(\tilde{\mathbf{v}}_i) \right] \Phi_n^E(\mathcal{V}; \mathbf{x}), \quad (\text{D.17})$$

where  $\tilde{\mathcal{V}} = (\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_m)$  encodes the vectors contracted with the covariant derivatives. In the case where all  $\tilde{\mathbf{v}}_i$  are mutually orthogonal, one finds the following asymptotics at large  $\mathbf{x}$

$$\lim_{\mathbf{x} \rightarrow \infty} \tilde{\Phi}_{n,m}^E(\mathcal{V}, \tilde{\mathcal{V}}; \mathbf{x}) = \prod_{i=1}^m (\tilde{\mathbf{v}}_i, \mathbf{x}) \prod_{j=1}^n \text{sgn}(\mathbf{v}_j, \mathbf{x}). \quad (\text{D.18})$$

Note that the derivative  $\partial_{\mathbf{x}}$  in (D.17) does not act on sign functions since  $\Phi_n^E$  is smooth and all discontinuities due to signs are guaranteed to cancel. Similarly to (D.17), we can also define  $\tilde{\Phi}_{n,m}^M(\mathcal{V}, \tilde{\mathcal{V}}; \mathbf{x})$  where the action of derivatives on the discontinuities of  $\Phi_n^M$  is ignored as well. In the particular case  $m = n$ , we will omit the second label and simply write  $\tilde{\Phi}_n^E$  or  $\tilde{\Phi}_n^M$ .

## E. Twistorial integrals and generalized error functions

In this appendix we evaluate the kernels  $\Phi_n^f$  (4.6) and show that they can be expressed through the generalized error functions introduced in appendix D.2. To this end, let us note the following identity

$$i \mathcal{D}(\mathbf{v}_{ij}) \frac{W_{p_i}(x_i, z_i) W_{p_j}(x_j, z_j)}{z_i - z_j} = \hat{K}_{ij} W_{p_i}(x_i, z_i) W_{p_j}(x_j, z_j). \quad (\text{E.1})$$

<sup>25</sup>This relation reduces to Eq. (3.53) and (6.18) in [23] for  $n = 2, 3$ , and follows from Eq. (63) in [24] for any  $n$ .

Due to it, the kernel can be represented as

$$\Phi_n^J(\mathbf{x}) = \frac{1}{n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \left[ \prod_{e \in E_{\mathcal{T}}} \mathcal{D}(\mathbf{v}_{s(e)t(e)}) \right] \Phi_{\mathcal{T}}(\mathbf{x}), \quad (\text{E.2})$$

where

$$\Phi_{\mathcal{T}}(\mathbf{x}) = \frac{i^{n-1}}{(2\pi)^n} \left[ \prod_{i=1}^n \int_{\ell_{\gamma_i}} dz_i W_{p_i}(x_i, z_i) \right] \frac{1}{\prod_{e \in E_{\mathcal{T}}} (z_{s(e)} - z_{t(e)})}. \quad (\text{E.3})$$

One may think that the representation (E.2) is not complete because it misses contributions from the covariant derivatives acting on each other. Such contributions are proportional to the scalar products of two vectors  $\mathbf{v}_{s(e)t(e)}$  and are non-vanishing provided the two edges have a common vertex. If this is the case, the two edges generate the following factor

$$\frac{(p_{\mathbf{v}_1} p_{\mathbf{v}_2} p_{\mathbf{v}_3})}{(z_{\mathbf{v}_1} - z_{\mathbf{v}_3})(z_{\mathbf{v}_2} - z_{\mathbf{v}_3})}, \quad (\text{E.4})$$

where  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  are the tree vertices joint by the edges and  $\mathbf{v}_3 = e_1 \cap e_2$ . The crucial observation is that if we pick up 3 subtrees, each with a marked vertex, there are exactly 3 ways to form a labelled tree out of them by joining the marked vertices as shown in Fig. 7 on page 31. Each subtree contributes the same factor in all 3 cases, whereas the joining edges and the sum over trees give rise to the vanishing factor

$$\frac{(p_1 p_2 p_3)}{(z_1 - z_3)(z_2 - z_3)} + \frac{(p_1 p_2 p_3)}{(z_1 - z_2)(z_1 - z_3)} - \frac{(p_1 p_2 p_3)}{(z_2 - z_3)(z_1 - z_2)} = 0. \quad (\text{E.5})$$

This ensures that no additional contributions arise and thereby proves (E.2). Note that for this proof it was crucial that inequivalent labelled trees enter the sum with the same weight.

Next, note that the contours  $\ell_{\gamma_i}$  in the  $z$ -plane go along the arcs running from  $-1$  to  $1$  and passing through the saddle points  $z_{\gamma_i}$  (3.10). However, in the large volume limit one can deform the contours into the straight lines  $z_{\gamma_i} + \mathbb{R}$  since this changes the integrals by exponentially small contributions.

Then let us do the change of variables

$$z_i = z' - \frac{i(pxt)}{\sqrt{2\tau_2}(pt^2)} + \sum_{\alpha=1}^{n-1} e_i^\alpha z'_\alpha, \quad (\text{E.6})$$

where  $x^a = \kappa^{ab} \sum \kappa_{i,bc} x_i^c$  (cf. (4.8)), and  $e_i^\alpha$  are such that

$$\sum_{i=1}^n \mathbf{p}_i e_i^\alpha = 0, \quad \sum_{i=1}^n \mathbf{p}_i e_i^\alpha e_i^\beta = 0, \quad \alpha \neq \beta \quad (\text{E.7})$$

and we introduced the convenient notation  $\mathbf{p}_i = (p_i t^2)$ . Labeling the  $n - 1$  edges of the tree by the same index  $\alpha$ , one can rewrite the function (E.3) in the new variables as

$$\Phi_{\mathcal{T}}(\mathbf{x}) = \frac{i^{n-1} \sqrt{\Delta}}{(2\pi)^n} e^{-\frac{\pi(pxt)^2}{(pt^2)}} \int dz' e^{-2\pi\tau_2(pt^2)z'^2} \prod_{\alpha=1}^{n-1} \int \frac{dz'_\alpha e^{-2\pi\tau_2 \Delta_\alpha (z'_\alpha)^2 - 2\pi i \sqrt{2\tau_2} w^\alpha z'_\alpha}}{\sum_{\beta=1}^{n-1} \left( e_{s(\alpha)}^\beta - e_{t(\alpha)}^\beta \right) z'_\beta}, \quad (\text{E.8})$$

where

$$w^\alpha = \sum_{i=1}^n (p_i x_i t) e_i^\alpha, \quad \Delta_\alpha = \sum_{i=1}^n \mathbf{p}_i (e_i^\alpha)^2, \quad \Delta = \frac{\mathbf{p} \prod_{\alpha=1}^{n-1} \Delta_\alpha}{\prod_{i=1}^n \mathbf{p}_i}. \quad (\text{E.9})$$

The integral over  $z'$  is Gaussian and is easily evaluated. In the remaining integrals, rescaling the integration variables by  $\sqrt{2\tau_2 \Delta_\alpha}$ , one recognizes the generalized error functions (D.10). Thus, one obtains

$$\Phi_{\mathcal{T}}(\mathbf{x}) = \frac{1}{2^{n-1}} \frac{\sqrt{|\Delta|} \det \mathcal{M}}{\prod_{\alpha=1}^{n-1} \Delta_\alpha} \Phi_1^f(x) M_{n-1} \left( \mathcal{M}; \left\{ \frac{w^\alpha}{\sqrt{\Delta_\alpha}} \right\} \right). \quad (\text{E.10})$$

where we used the function  $\Phi_1^f$  (4.13) and introduced the matrix  $\mathcal{M}$  such that

$$\mathcal{M}_{\alpha\beta}^{-1} = (\Delta_\alpha \Delta_\beta)^{-1/2} \left( e_{s(\alpha)}^\beta - e_{t(\alpha)}^\beta \right). \quad (\text{E.11})$$

There is a simple solution to the conditions (E.7). Let  $T$  be a rooted ordered binary tree with  $n$  leaves decorated by  $\gamma_i$ ,  $i = 1 \dots n$ . As usual for such trees, other vertices  $v$  carry charges given by the sum of charges of their children, i.e.  $\gamma_v = \sum_{i \in \mathcal{I}_v} \gamma_i$  where  $\mathcal{I}_v$  is the set of leaves which are descendants of  $v$ . There are  $n - 1$  such vertices which we label by index  $\alpha$ . Then we can choose

$$e_i^\alpha = \sum_{j \in \mathcal{I}_L(v_\alpha)} \sum_{k \in \mathcal{I}_R(v_\alpha)} (\delta_{ij} \mathbf{p}_k - \delta_{ik} \mathbf{p}_j), \quad (\text{E.12})$$

which satisfy (E.7) as can be easily checked. For this choice (cf. (4.14))

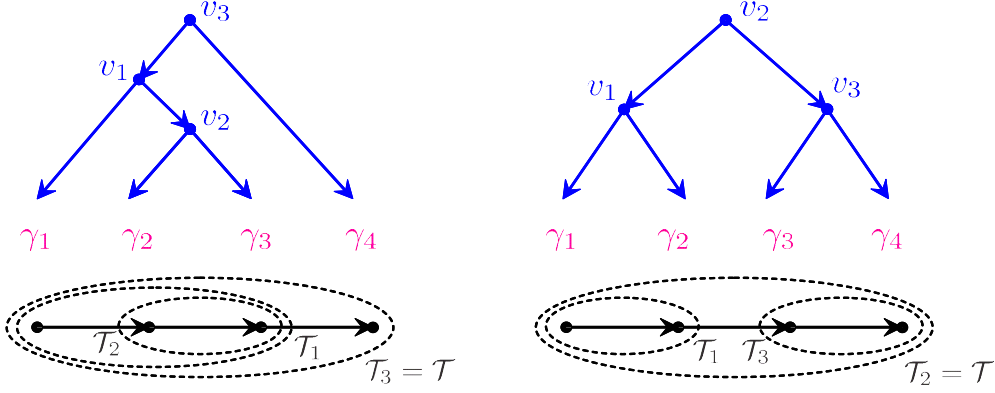
$$w^\alpha = (\tilde{\mathbf{u}}_\alpha, \mathbf{x}), \quad \text{where} \quad \tilde{\mathbf{u}}_\alpha = \sum_{i \in \mathcal{I}_L(v_\alpha)} \sum_{j \in \mathcal{I}_R(v_\alpha)} \mathbf{u}_{ij}, \quad (\text{E.13})$$

$$\Delta_\alpha = \tilde{\mathbf{u}}_\alpha^2 = \mathbf{p}_{v_\alpha} \mathbf{p}_{L(v_\alpha)} \mathbf{p}_{R(v_\alpha)}.$$

Note that the vectors  $\tilde{\mathbf{u}}_\alpha$  are mutually orthogonal.

In principle, any rooted binary tree  $T$  is suitable for the above construction. However, given the unrooted tree  $\mathcal{T}$ , there is a simple (but non-unique) choice of  $T$  which simplifies the resulting matrix  $\mathcal{M}$ . To define it, let us construct a partially increasing family of subtrees of  $\mathcal{T}$ , such that two members in this family are either disjoint, or contained in one another. Moreover, we require that the largest subtree is  $\mathcal{T}$  itself, while each subtree containing more than one vertex is obtained by joining two smaller subtrees along an edge of  $\mathcal{T}$ . Any such family contains  $2n - 1$  subtrees  $\mathcal{T}_{\hat{\alpha}}$  labelled by  $\hat{\alpha} = 1, \dots, 2n - 1$ . Among them,  $n - 1$  subtrees, which we label by  $\alpha = \hat{\alpha} = 1, \dots, n - 1$ , contain several vertices, while the remaining  $n$  subtrees with label  $\hat{\alpha} = n, \dots, 2n - 1$  have only one vertex. For each subtree  $T_\alpha$ , we denote by  $e_\alpha$  the edge of  $\mathcal{T}$  which is used to reconstruct  $T_\alpha$  from two smaller subtrees  $\mathcal{T}_{\alpha_L}, \mathcal{T}_{\alpha_R}$ . From this data, we construct a rooted binary tree  $T$  with  $n - 1$  vertices in one-to-one correspondence with the subtrees  $\mathcal{T}_\alpha$  and  $n$  leaves in one-to-one correspondence with the one-vertex subtrees  $\mathcal{T}_{\hat{\alpha}}$  with  $\hat{\alpha} \geq n$ . In this correspondence, the two children of a vertex associated to  $\mathcal{T}_\alpha$  are the vertices associated to the two subtrees  $\mathcal{T}_{\alpha_L}, \mathcal{T}_{\alpha_R}$ . The ordering at each vertex is defined to be such that the subtree containing the source/target vertex of the corresponding edge  $e_\alpha$  is on the left/right.<sup>26</sup> Of course, this construction is not unique since there are many ways to decompose  $\mathcal{T}$  into such a set of subtrees (see Fig. 10).

<sup>26</sup>The orientation of the edges of  $\mathcal{T}$  is fixed already in (E.2), but the full kernel  $\Phi_n^f$  does not depend on its choice.



**Figure 10:** An example of an unrooted labelled tree with 4 vertices and two choices of decompositions into subtrees with the corresponding rooted binary trees. The edges are labelled  $e_1, e_2, e_3$  from left to right along  $\mathcal{T}$ .

Applying the above construction to this particular choice of rooted tree, one finds

$$\sqrt{\Delta_\alpha \Delta_\beta} \mathcal{M}_{\alpha\beta}^{-1} = \begin{cases} \mathbf{p}_{v_\alpha}, & \alpha = \beta, \\ \epsilon_{\alpha\beta} \mathbf{p}_{L(v_\beta)}, & e_\alpha \cap \mathcal{T}_{\beta_R} \neq \emptyset, e_\alpha \not\subseteq \mathcal{T}_\beta, \\ \epsilon_{\alpha\beta} \mathbf{p}_{R(v_\beta)}, & e_\alpha \cap \mathcal{T}_{\beta_L} \neq \emptyset, e_\alpha \not\subseteq \mathcal{T}_\beta, \\ 0, & e_\alpha \cap \mathcal{T}_\beta = \emptyset \text{ or } e_\alpha \subset \mathcal{T}_\beta, \end{cases} \quad (\text{E.14})$$

where  $\epsilon_{\alpha\beta} = -1$  if the orientations of  $e_\alpha$  and  $e_\beta$  on the path joining them are the same and  $+1$  otherwise. This result shows that the matrix  $\mathcal{M}^{-1}$  turns out to be triangular which makes it much simpler to find its inverse. On the basis of (E.14), below we will prove the following

**Lemma 2.** *One has  $\mathcal{B} \cdot \mathcal{V} = \mathcal{M}$  and  $\mathcal{B} \cdot \mathbf{x} = \left\{ \frac{w^\alpha}{\sqrt{\Delta_\alpha}} \right\}$  provided*

$$\begin{aligned} \mathcal{B} &= \left( \frac{\tilde{\mathbf{u}}_1}{\sqrt{\Delta_1}}, \dots, \frac{\tilde{\mathbf{u}}_{n-1}}{\sqrt{\Delta_{n-1}}} \right)^{\text{tr}}, \\ \mathcal{V} &= \mathbf{p}^{-1} \left( \sqrt{\Delta_1} \mathbf{u}_1, \dots, \sqrt{\Delta_{n-1}} \mathbf{u}_{n-1} \right), \end{aligned} \quad (\text{E.15})$$

where the vectors  $\mathbf{u}_\alpha$  are defined in (4.11). Moreover, the vectors  $\tilde{\mathbf{u}}_\alpha$  form an orthogonal basis in the subspace spanned by  $\mathbf{u}_\alpha$ .

This lemma allows to reexpress the kernel  $\Phi_{\mathcal{T}}$  (E.10) in terms of the boosted generalized error function  $\Phi_{n-1}^M$  (D.14). It is important that its argument  $\mathcal{V}$  does not depend on the choice of the binary tree  $T$ , but only on  $\mathcal{T}$ . In addition, the function actually does not depend on the normalization of the vectors composing  $\mathcal{V}$ . Given also that the determinant of  $\mathcal{M}$  is found to be

$$|\det \mathcal{M}| = \prod_{\alpha=1}^{n-1} \frac{\Delta_\alpha}{\mathbf{p}_{v_\alpha}} = \prod_{i=1}^n \mathbf{p}_i \prod_{v \in V_T \setminus \{v_0\}} \mathbf{p}_v = \sqrt{\mathbf{p}^{-1} \prod_{i=1}^n \mathbf{p}_i \prod_{\alpha=1}^{n-1} \Delta_\alpha}, \quad (\text{E.16})$$

so that the prefactor in (E.10) cancels, one arrives at

$$\Phi_T(\mathbf{x}) = \frac{1}{2^{n-1}} \Phi_1^f(x) \Phi_{n-1}^M(\{\mathbf{u}_e\}; \mathbf{x}). \quad (\text{E.17})$$

Finally, since the differential operator in (E.2) commutes with functions of  $x$  due to orthogonality of  $\mathbf{v}_{ij}$  and  $\mathbf{t}$ , one can write the kernel  $\Phi_n^f$  as

$$\Phi_n^f(\mathbf{x}) = \frac{\Phi_1^f(x)}{2^{n-1}n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \left[ \prod_{e \in E_{\mathcal{T}}} \mathcal{D}(\mathbf{v}_{s(e)t(e)}) \right] \Phi_{n-1}^M(\{\mathbf{u}_\alpha\}; \mathbf{x}). \quad (\text{E.18})$$

which is the same as (4.12).

### Proof of Lemma 2

We start by proving that the vectors  $\tilde{\mathbf{u}}_\alpha$  form an orthogonal basis in the subspace spanned by  $\mathbf{u}_\alpha$ . Since the orthogonality is ensured by construction based on a rooted binary tree, it remains to show that any vector  $\mathbf{u}_\alpha$  can be decomposed as a linear combination of  $\tilde{\mathbf{u}}_\alpha$ . To this end, we show that the determinant of the Gram matrix constructed from the set of vectors  $\{\tilde{\mathbf{u}}_\alpha\}_{\alpha=1}^{n-1} \cup \{\mathbf{u}_\beta\}$  vanishes. This requires to calculate the scalar product  $(\tilde{\mathbf{u}}_\alpha, \mathbf{u}_\beta)$  which can be done using

$$(\mathbf{u}_{ij}, \mathbf{u}_{kl}) = \begin{cases} \mathfrak{p}_i \mathfrak{p}_j \mathfrak{p}_{i+j}, & i = k, j = l, \\ \mathfrak{p}_i \mathfrak{p}_j \mathfrak{p}_l, & i = k, j \neq l, \\ 0, & i, j \neq k, l. \end{cases} \quad (\text{E.19})$$

Summing  $i, j, k, l$  over appropriate subsets, it is immediate to see that  $(\tilde{\mathbf{u}}_\alpha, \mathbf{u}_\beta) = 0$  if  $\mathcal{T}_\alpha \subset \mathcal{T}_\beta^s$  or  $\mathcal{T}_\beta^t$ , which are the two trees obtained by dividing the tree  $\mathcal{T}$  into two parts by cutting the edge  $e_\alpha$ . In other words, it is non-vanishing only if  $e_\beta \subseteq \mathcal{T}_\alpha$ . Then there are two cases which give

$$(\tilde{\mathbf{u}}_\alpha, \mathbf{u}_\beta) = \begin{cases} \mathfrak{p}_\alpha^L \mathfrak{p}_\alpha^R \mathfrak{p}, & \alpha = \beta, \\ \mathfrak{p}_{\alpha\beta}^{Ls} \mathfrak{p}_\alpha^R \mathfrak{p}_\beta^t - \mathfrak{p}_{\alpha\beta}^{Lt} \mathfrak{p}_\alpha^R \mathfrak{p}_\beta^s + \mathfrak{p}_\alpha^L \mathfrak{p}_\alpha^R \mathfrak{p}_\beta^s = \mathfrak{p}_{\alpha\beta}^{Ls} \mathfrak{p}_\alpha^R \mathfrak{p}, & e_\beta \subset \mathcal{T}_\alpha, \end{cases} \quad (\text{E.20})$$

where we introduced

$$\mathfrak{p}_\alpha^L = \sum_{i \in \mathcal{I}_{L(v_\alpha)}} \mathfrak{p}_i, \quad \mathfrak{p}_\alpha^s = \sum_{i \in V_{\mathcal{T}_\alpha^s}} \mathfrak{p}_i, \quad \mathfrak{p}_{\alpha\beta}^{Ls} = \sum_{i \in \mathcal{I}_{L(v_\alpha)} \cap V_{\mathcal{T}_\beta^s}} \mathfrak{p}_i, \quad \mathfrak{p}_{\alpha\beta}^{st} = \sum_{i \in V_{\mathcal{T}_\alpha^s} \cap V_{\mathcal{T}_\beta^t}} \mathfrak{p}_i, \quad (\text{E.21})$$

and similarly for variables with labels  $R$  and  $t$ . In (E.20) in the second case we assumed that the orientation of edges is such that  $e_\beta \subset \mathcal{T}_\alpha^s$  and  $e_\alpha \subset \mathcal{T}_\beta^t$ . If this is not the case, one should replace  $s$  by  $t$ ,  $L$  by  $R$  and flip the sign for each change of orientation. Below we use the same assumption, but the computation can easily be generalized to a more general situation.

Given the result (E.20),  $(\tilde{\mathbf{u}}_\alpha, \tilde{\mathbf{u}}_\beta) = \Delta_\alpha \delta_{\alpha\beta}$  and  $(\mathbf{u}_\beta, \mathbf{u}_\beta) = \mathfrak{p}_\beta^s \mathfrak{p}_\beta^t \mathfrak{p}$ , the determinant of the Gram matrix is easily found to be

$$\det \text{Gram}(\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_{n-1}, \mathbf{u}_\beta) = \mathfrak{p} \prod_{\alpha=1}^{n-1} \Delta_\alpha \left[ \mathfrak{p}_\beta^s \mathfrak{p}_\beta^t - \mathfrak{p} \sum_{\mathcal{T}_\alpha \supseteq e_\beta} \frac{(\mathfrak{p}_{\alpha\beta}^{Ls})^2 \mathfrak{p}_\alpha^R}{\mathfrak{p}_{v_\alpha} \mathfrak{p}_\alpha^L} \right]. \quad (\text{E.22})$$

Note that the subtrees  $\mathcal{T}_\alpha$  containing the edge  $e_\beta$  form an ordered set so that the sum in the square brackets goes over  $\alpha_\ell$ ,  $\ell = 1, \dots, m$ , such that  $\mathcal{T}_{\alpha_\ell} \subset \mathcal{T}_{\alpha_{\ell+1}}$ . The first element of this set  $\alpha_1 = \beta$ , whereas the last corresponds to the total tree,  $\mathcal{T}_{\alpha_m} = \mathcal{T}$ . Due to  $\mathfrak{p}_{\alpha_m}^L = \mathfrak{p}_{\alpha_m}^s$ ,  $\mathfrak{p}_{\alpha_m}^R = \mathfrak{p}_{\alpha_m}^t$  and  $\mathfrak{p}_{\alpha_m\beta}^{Ls} = \mathfrak{p}_\beta^s$ , the first term in the square brackets together with the term in the sum corresponding to  $\alpha_m$  gives

$$\mathfrak{p}_\beta^s \left( \mathfrak{p}_\beta^t - \frac{\mathfrak{p}_\beta^s \mathfrak{p}_{\alpha_m}^t}{\mathfrak{p}_{\alpha_m}^s} \right) = \frac{\mathfrak{p} \mathfrak{p}_\beta^s \mathfrak{p}_{\alpha_m\beta}^{st}}{\mathfrak{p}_{\alpha_m}^s}, \quad (\text{E.23})$$

where we have used that  $\mathbf{p}_\beta^t = \mathbf{p}_{\alpha_m\beta}^{st} + \mathbf{p}_{\alpha_m}^t$  and  $\mathbf{p}_{\alpha_m}^s = \mathbf{p}_{\alpha_m\beta}^{st} + \mathbf{p}_\beta^s$ . Thus, the expression in the square brackets in (E.22) becomes

$$\frac{\mathbf{p}}{\mathbf{p}_{\alpha_m}^s} \left[ \mathbf{p}_\beta^s \mathbf{p}_{\alpha_m\beta}^{st} - \mathbf{p}_{\alpha_m}^s \sum_{\ell=1}^{m-1} \frac{(\mathbf{p}_{\alpha_\ell\beta}^{Ls})^2 \mathbf{p}_{\alpha_\ell}^R}{\mathbf{p}_{v_{\alpha_\ell}}^L \mathbf{p}_{\alpha_\ell}^L} \right]. \quad (\text{E.24})$$

The new expression in the square brackets is exactly the same as in (E.22) where tree  $\mathcal{T}$  was replaced by subtree  $\mathcal{T}_{\alpha_m}^s = \mathcal{T}_{\alpha_{m-1}}$ . Thus, one can repeat the above manipulation until one exhausts all terms in the sum. As a result, the determinant of the Gram matrix turns out to be proportional to  $\mathbf{p}_{\alpha_1\beta}^{st}$ . But since  $\alpha_1 = \beta$ , this quantity, and hence the whole determinant, trivially vanish.

Next, we prove that (E.15) is consistent with  $\mathcal{B} \cdot \mathbf{x} = \left\{ \frac{w_\alpha}{\sqrt{\Delta_\alpha}} \right\}$  and  $\mathcal{B} \cdot \mathcal{V} = \mathcal{M}$ . The first relation is a direct consequence of (E.13). The second relation requires to show that  $(\tilde{\mathbf{u}}_\alpha, \mathbf{u}_\beta) = \mathbf{p} \Delta_\alpha^{1/2} \Delta_\beta^{-1/2} \mathcal{M}_{\alpha\beta}$  or equivalently  $\sum_\gamma (\tilde{\mathbf{u}}_\alpha, \mathbf{u}_\gamma) \sqrt{\Delta_\gamma \Delta_\beta} \mathcal{M}_{\gamma\beta}^{-1} = \mathbf{p} \Delta_\alpha \delta_{\alpha\beta}$ . Using the result (E.14) for the matrix  $\mathcal{M}_{\gamma\beta}^{-1}$ , this relation can be written explicitly as

$$\mathbf{p}_{v_\beta} (\tilde{\mathbf{u}}_\alpha, \mathbf{u}_\beta) + \left( \sum_{e_\gamma \in E_\beta^R} \epsilon_{\gamma\beta} \mathbf{p}_{L(v_\beta)} + \sum_{e_\gamma \in E_\beta^L} \epsilon_{\gamma\beta} \mathbf{p}_{R(v_\beta)} \right) (\tilde{\mathbf{u}}_\alpha, \mathbf{u}_\gamma) = \mathbf{p} \Delta_\alpha \delta_{\alpha\beta}, \quad (\text{E.25})$$

where  $E_\beta^L = \{e \in E_\mathcal{T} : e \cap \mathcal{T}_{\beta L} \neq \emptyset, e \not\subseteq \mathcal{T}_\beta\}$  and similarly for  $E_\beta^R$ .

Consider first the case  $\alpha = \beta$ . From (E.20), it immediately follows that the first term gives  $\mathbf{p} \Delta_\alpha$ . On the other hand, the second contribution sums over edges for which  $\mathcal{T}_\alpha \subseteq \mathcal{T}_\gamma^s$  or  $\mathcal{T}_\gamma^t$ , and as noted above (E.20) this leads to vanishing of the scalar product. Thus, in this case the relation (E.25) indeed holds.

Let us now show that it holds as well for  $\alpha \neq \beta$ . To this end, one should consider several options. If  $\mathcal{T}_\alpha \cap \mathcal{T}_\beta = \emptyset$ , then  $\mathcal{T}_\alpha \subset \mathcal{T}_\beta^s$  or  $\mathcal{T}_\beta^t$  which implies vanishing of the first term. But the second term vanishes as well because the conditions  $e_\gamma \subseteq \mathcal{T}_\alpha$  and  $e_\gamma \cap \mathcal{T}_\beta \neq \emptyset$  are inconsistent with  $\mathcal{T}_\alpha \cap \mathcal{T}_\beta = \emptyset$ .

Similarly, if  $\mathcal{T}_\alpha \subset \mathcal{T}_\beta$ , one has  $\mathcal{T}_\alpha \subset \mathcal{T}_\beta^s$  or  $\mathcal{T}_\beta^t$  which again leads to the vanishing of the first term, whereas the vanishing of the second is a consequence of that  $e_\gamma \subseteq \mathcal{T}_\alpha$  implies  $e_\gamma \subset \mathcal{T}_\beta$  so that the sum over  $e_\gamma$  is empty.

It remains to consider the case  $\mathcal{T}_\beta \subset \mathcal{T}_\alpha$ . It is clear that  $\mathcal{T}_\beta \subset \mathcal{T}_\alpha^s$  or  $\mathcal{T}_\alpha^t$ . Without loss of generality, let us assume that  $\mathcal{T}_\beta \subset \mathcal{T}_\alpha^s$  and  $e_\alpha \subset \mathcal{T}_\beta^t$ . Then according to (E.20), the first term gives  $\mathbf{p}_{v_\beta} \mathbf{p}_{\alpha\beta}^{Ls} \mathbf{p}_\alpha^R \mathbf{p}$ . If instead we have chosen the orientation such that  $e_\alpha \subset \mathcal{T}_\beta^s$ , then we would find  $-\mathbf{p}_{v_\beta} \mathbf{p}_{\alpha\beta}^{Lt} \mathbf{p}_\alpha^R \mathbf{p}$ . Similar results are obtained for each term in the sum of the second contribution. Again without loss of generality we assume that for all relevant edges  $e_\gamma$  one has  $e_\alpha \subset \mathcal{T}_\gamma^t$ , otherwise one flips their orientation. Then the l.h.s. of (E.25) is proportional to

$$\mathbf{p}_{v_\beta} \mathbf{p}_{\alpha\beta}^{Ls} + \sum_{e_\gamma \in E_\beta^R} \epsilon_{\gamma\beta} \mathbf{p}_\beta^L \mathbf{p}_{\alpha\gamma}^{Ls} + \sum_{e_\gamma \in E_\beta^L} \epsilon_{\gamma\beta} \mathbf{p}_\beta^R \mathbf{p}_{\alpha\gamma}^{Ls}. \quad (\text{E.26})$$

Let  $e_\alpha^* \in E_\beta^R$  is the edge belonging to the path from  $\mathcal{T}_\beta$  to  $e_\alpha$  (which may coincide with  $e_\alpha$ ). Then our choice of orientation implies that  $\epsilon_{\gamma\beta} = -1$  for  $e_\gamma \in E_\beta^L \cup \{e_\alpha^*\}$  and  $\epsilon_{\gamma\beta} = 1$  for  $e_\gamma \in E_\beta^R \setminus \{e_\alpha^*\}$ . Furthermore, one has

$$\mathbf{p}_{\alpha\gamma_\alpha^*}^{Ls} = \mathbf{p}_{v_\beta} + \sum_{e_\gamma \in E_\beta^L \cup E_\beta^R \setminus \{e_\alpha^*\}} \mathbf{p}_{\alpha\gamma}^{Ls}, \quad \mathbf{p}_{\alpha\beta}^{Ls} = \mathbf{p}_\beta^L + \sum_{e_\gamma \in E_\beta^L} \mathbf{p}_{\alpha\gamma}^{Ls}. \quad (\text{E.27})$$

As a result, one finds

$$(\mathfrak{p}_\beta^L + \mathfrak{p}_\beta^R) \left( \mathfrak{p}_\beta^L + \sum_{e_\gamma \in E_\beta^L} \mathfrak{p}_{\alpha\gamma}^{Ls} \right) - \mathfrak{p}_\beta^L \mathfrak{p}_{\alpha\gamma_\alpha^*}^{Ls} + \sum_{e_\gamma \in E_\beta^R \setminus \{e_\alpha^*\}} \mathfrak{p}_\beta^L \mathfrak{p}_{\alpha\gamma}^{Ls} - \sum_{e_\gamma \in E_\beta^L} \mathfrak{p}_\beta^R \mathfrak{p}_{\alpha\gamma}^{Ls} = 0. \quad (\text{E.28})$$

This completes the proof of the required statement.

## F. Proofs of propositions

In this appendix we fill the gap and prove several propositions which we stated in the main text.

### Proposition 1

To prove the recursive equation (5.3), let us note that the tree index  $g_{\text{tr},n}$  satisfies a very similar equation (cf. (2.9) or (2.15)) which can be seen as the origin of its representation (2.3) in terms of attractor flow trees. The only difference is the absence of the last term in (5.3). Therefore, it is easy to see that this equation implies a simple relation between  $\widehat{g}_n$  and  $g_{\text{tr},n}$

$$\widehat{g}_n(\{\check{\gamma}_i\}, z^a) = \sum_{\substack{n_1 + \dots + n_m = n \\ n_k \geq 1}} g_{\text{tr},m}(\{\check{\gamma}'_k\}, z^a) \prod_{k=1}^m W_{n_k}(\check{\gamma}_{j_{k-1}+1}, \dots, \check{\gamma}_{j_k}), \quad (\text{F.1})$$

where as usual we use notations from (2.16).

Next, we substitute this relation into the expansion (5.2). The result can be represented in the following form

$$h_{p,q}^{\text{DT}} = \sum_{\sum_{i=1}^n \check{\gamma}_i = \check{\gamma}} g_{\text{tr},n}(\{\check{\gamma}_i\}, z^a) e^{\pi i \tau Q_n(\{\check{\gamma}_i\})} \prod_{i=1}^n h_{p_i, \mu_i}^R(\tau), \quad (\text{F.2})$$

where we introduced

$$\begin{aligned} h_{p,\mu}^R &= \sum_{\sum_{i=1}^n \check{\gamma}_i = \check{\gamma}} W_n(\{\check{\gamma}_i\}) e^{\pi i \tau Q_n(\{\check{\gamma}_i\})} \prod_{i=1}^n \widehat{h}_{p_i, \mu_i}(\tau) \\ &= \sum_{\sum_{i=1}^n \check{\gamma}_i = \check{\gamma}} \left[ \sum_{T \in \mathbb{T}_n^S} \prod_{v \in V_T} R_v \right] e^{\pi i \tau Q_n(\{\check{\gamma}_i\})} \\ &\quad \times \prod_{i=1}^n \left( h_{p_i, \mu_i} - \sum_{n_i=2}^{\infty} \sum_{\sum_{j=1}^{n_i} \check{\gamma}'_j = \check{\gamma}} R_{n_i}(\{\check{\gamma}'_j\}) e^{\pi i \tau Q_{n_i}(\{\check{\gamma}'_j\})} \prod_{j=1}^{n_i} h_{p'_{j_i}, \mu'_{j_i}} \right) \end{aligned} \quad (\text{F.3})$$

and in the last relation we used the definition of  $W_n$  (5.4) and the expansion of  $\widehat{h}_{p,\mu}$  (5.1). The crucial observation is that if one picks up a factor  $R_{n_i}$  from the second line of (F.3), appearing due to the expansion of  $\widehat{h}_{p_i, \mu_i}$ , and combines it with the contribution of a tree  $T$  from the first line, one finds the opposite of the contribution of another tree obtained from



$T$  by adding  $n_i$  children to its  $i$ th leaf. As a result, all such contributions cancel and the function (F.3) reduces to the trivial term

$$h_{p,\mu}^R = h_{p,\mu}. \quad (\text{F.4})$$

Substituting this into (F.2), it gives back the original expansion (2.30) of the generating function of DT invariants, which proves the recursive equation (5.3).

## Proposition 2

We will prove proposition 2 by induction. For  $n = 2$  the recursive relation (5.9) reads

$$g_2^{(0)}(\check{\gamma}_1, \check{\gamma}_2; c_1) - g_2^{(0)}(\check{\gamma}_1, \check{\gamma}_2; \beta_{21}) = -\frac{1}{4} (\text{sgn}(c_1) - \text{sgn}(\beta_{21})) \kappa(\gamma_{12}), \quad (\text{F.5})$$

where we took into account that  $S_1 = c_1$  and  $\Gamma_{21} = \beta_{21} = -\gamma_{12}$ . Since  $g_2^{(0)}$  is supposed to have discontinuities only at walls of marginal stability, it must not involve signs of DSZ products. Therefore, we are led to take

$$g_2^{(0)}(\check{\gamma}_1, \check{\gamma}_2; c_1) = -\frac{1}{4} \text{sgn}(c_1) \kappa(\gamma_{12}). \quad (\text{F.6})$$

Then (5.10) and (5.11) imply

$$\mathcal{E}_2 = \frac{1}{4} \text{sgn}(\gamma_{12}) \kappa(\gamma_{12}) - R_2, \quad (\text{F.7})$$

so that the ansatz (5.8) reads

$$\widehat{g}_2(\check{\gamma}_1, \check{\gamma}_2; z) = g_2^{(0)}(\check{\gamma}_1, \check{\gamma}_2; c_1) - \mathcal{E}_2(\check{\gamma}_1, \check{\gamma}_2) = -\frac{1}{4} \left[ \text{sgn}(c_1) + \text{sgn}(\gamma_{12}) \right] \kappa(\gamma_{12}) + R_2(\check{\gamma}_1, \check{\gamma}_2), \quad (\text{F.8})$$

which reproduces the recursive equation (5.3). Furthermore, in appendix D.2 it is shown that there is a smooth solution of Vignéras' equation which asymptotes the function  $(\mathbf{v}, \mathbf{x}) \text{sgn}(\mathbf{v}, \mathbf{x})$ , coinciding with the (rescaled) first term in  $\mathcal{E}_2$  for  $\mathbf{v} = \mathbf{v}_{12}$  (4.10). It is given by

$$\widetilde{\Phi}_1^E(\mathbf{v}, \mathbf{v}; \mathbf{x}) = \mathbf{v} \cdot \left( \mathbf{x} + \frac{1}{2\pi} \partial_{\mathbf{x}} \right) \text{Erf} \left( \frac{\sqrt{\pi} \mathbf{v} \cdot \mathbf{x}}{|\mathbf{v}|} \right), \quad (\text{F.9})$$

which corresponds to the following choice of  $R_2$  [20]

$$R_2 = \frac{(-1)^{1+\gamma_{12}}}{8\pi} |\gamma_{12}| \beta_{\frac{3}{2}} \left( \frac{2\tau_2 \gamma_{12}^2}{(pp_1 p_2)} \right), \quad (\text{F.10})$$

where  $\beta_{\frac{3}{2}}(x^2) = 2|x|^{-1} e^{-\pi x^2} - 2\pi \text{Erfc}(\sqrt{\pi}|x|)$ . Note that the resulting  $\mathcal{E}_2$  depends on the electric charges only through the DSZ product  $\gamma_{12}$ . Finally, it is immediate to see that the kernel  $\Phi_2^{\widehat{g}}$  (5.5) satisfies (5.6).

Now we assume that (5.3) is consistent with the ansatz (5.8) for all orders up to  $n-1$  and check it at order  $n$ . Denoting the second term in (5.8) by  $\widehat{g}_n^{(+)}$  and substituting this ansatz into the r.h.s. of (5.3), one finds

$$\text{Sym} \left\{ \sum_{\ell=1}^{n-1} g_2^\ell \left[ g_\ell^{(0)} g_{n-\ell}^{(0)} - \widehat{g}_\ell^{(+)} \widehat{g}_{n-\ell} - \widehat{g}_\ell \widehat{g}_{n-\ell}^{(+)} - \widehat{g}_\ell^{(+)} \widehat{g}_{n-\ell}^{(+)} \right]_{z^a \rightarrow z_\ell^a} \right\} + W_n, \quad (\text{F.11})$$

where

$$g_2^\ell = -\frac{1}{2} \Delta_{\gamma_L^\ell \gamma_R^\ell}^z \kappa(\gamma_{LR}^\ell) = \frac{1}{4} (\text{sgn}(S_\ell) - \text{sgn}(\Gamma_{n\ell})) \kappa(\Gamma_{n\ell}). \quad (\text{F.12})$$

In the first term proportional to  $g_\ell^{(0)} g_{n-\ell}^{(0)}$ , one can apply the relation (5.9), which together with (5.10) gives  $g_n^{(0)} - \mathcal{E}_n^{(0)}$ . The other terms in the sum over  $\ell$  can be reorganized as follows

$$\begin{aligned} & - \text{Sym} \left\{ \sum_{\ell=1}^{n-1} g_2^\ell \sum_{\substack{n_1+\dots+n_m=n \\ n_k \geq 1, m < n, \ell \in \{j_k\}}} \widehat{g}_{k_0}(z_\ell^a) \widehat{g}_{m-k_0}(z_\ell^a) \prod_{k=1}^m \mathcal{E}_{n_k} \right\} \\ & = - \text{Sym} \left\{ \sum_{\substack{n_1+\dots+n_m=n \\ n_k \geq 1, 1 < m < n}} \left[ \sum_{k_0=1}^{m-1} g_2^{j_{k_0}} \widehat{g}_{k_0}(z_\ell^a) \widehat{g}_{m-k_0}(z_\ell^a) \right] \prod_{k=1}^m \mathcal{E}_{n_k} \right\}. \end{aligned} \quad (\text{F.13})$$

Here we first combined three contributions into one sum over splittings by adding the condition  $\ell \in \{j_k\}$ , with  $k_0$  being the index for which  $j_{k_0} = \ell$ , and then interchanged the two sums which allows to drop the condition  $\ell \in \{j_k\}$ , but adds the requirement  $m > 1$  (following from  $\ell \in \{j_k\}$  in the previous representation). In square brackets one recognizes the first contribution from the r.h.s. of (5.3) with  $n$  replaced by  $m < n$ . Hence, it is subject to the induction hypothesis which allows to replace this expression by  $\widehat{g}_m(\{\gamma'_k\}, z^a) - W_m(\{\gamma'_k\})$ . Combining all contributions together, one concludes that (F.11) is equal to

$$g_n^{(0)} - \mathcal{E}_n^{(0)} - \text{Sym} \left\{ \sum_{\substack{n_1+\dots+n_m=n \\ n_k \geq 1, 1 < m < n}} (\widehat{g}_m - W_m) \prod_{k=1}^m \mathcal{E}_{n_k} \right\} + W_n. \quad (\text{F.14})$$

The contributions involving  $W$ 's can be combined into one sum by dropping the condition  $m < n$ . The resulting sum coincides with the r.h.s. of (5.11) so that these contributions can be replaced by  $-\mathcal{E}_n^{(+)}$ . Combined with  $-\mathcal{E}_n^{(0)}$ , this gives  $-\mathcal{E}_n$  and allows to drop the condition  $m > 1$  in the remaining sum with  $\widehat{g}_m$ . As a result, (F.14) becomes equivalent to the r.h.s. of (5.8), which proves the consistency of this ansatz with the recursive equation.

Finally, let us show that the ansatz satisfies the modularity constraint (5.6). The crucial observation is that the vectors  $\mathbf{v}_{ij}$  and  $\mathbf{u}_{ij}$  (4.10) satisfy

$$(\mathbf{v}_{i+j,k}, \mathbf{v}_{ij}) = (\mathbf{u}_{i+j,k}, \mathbf{v}_{ij}) = 0, \quad (\text{F.15})$$

where we abused notation and denoted  $\mathbf{v}_{i+j,k} = \mathbf{v}_{ik} + \mathbf{v}_{jk}$ , etc. These orthogonality relations, together with the assumption that  $\mathcal{E}_n$  depend on electric charges only through the DSZ products  $\gamma_{ij} \sim (\mathbf{v}_{ij}, \mathbf{x})$ , imply factorization of the action of Vignéras' operator on the kernel corresponding to the second term in (5.8). Indeed, all contributions of  $\partial_{\mathbf{x}}^2$  where two derivatives act on different factors vanish and the action reduces to the sum of terms where Vignéras' operator acts on one of the factors. But since it is supposed to vanish on  $\Phi_n^\mathcal{E}$ , one obtains the simple result

$$V_{n-1} \cdot \Phi_n^{\widehat{g}} = V_{n-1} \cdot \Phi_n^{g^{(0)}} - \text{Sym} \left\{ \sum_{\substack{n_1+\dots+n_m=n \\ n_k \geq 1, m < n}} (V_{m-1} \cdot \Phi_m^{\widehat{g}}) \prod_{k=1}^m \Phi_{n_k}^\mathcal{E} \right\}. \quad (\text{F.16})$$

Since for  $n = 2$  the constraint was already shown to hold, one can proceed by induction. Then in the second term one can substitute the r.h.s. of (5.6), whereas the first term can be evaluated using the recursive relation (5.9). First of all, by the same reasoning as above, away from discontinuities, the action of Vignéras' operator is factorized and actually vanishes. Furthermore, since  $g_n^{(0)}$  have discontinuities only at walls of marginal stability, to obtain the complete action, it is enough to consider it only on  $\text{sgn}(S_\ell)$ . Since at  $S_\ell = 0$  one has  $c_i^{(\ell)} = c_i$  (see (2.14)), one finds that  $\Phi_n^{g^{(0)}}$  satisfy exactly the same constraint as (5.6). Thus, one can rewrite (F.16) as

$$V_{n-1} \cdot \Phi_n^{\hat{g}} = \text{Sym} \sum_{\ell=1}^{n-1} \left( \mathbf{u}_\ell^2 \Delta_{n,\ell}^{g^{(0)}} \delta'(\mathbf{u}_\ell \cdot \mathbf{x}) + 2\mathbf{u}_\ell \cdot \partial_{\mathbf{x}} \Delta_{n,\ell}^{g^{(0)}} \delta(\mathbf{u}_\ell \cdot \mathbf{x}) \right) \quad (\text{F.17})$$

$$- \text{Sym} \left\{ \sum_{\substack{n_1+\dots+n_m=n \\ n_k \geq 1, 1 < m < n}} \left[ \sum_{\ell=1}^{n-1} \left( \mathbf{u}_\ell^2 \Delta_{m,\ell}^{\hat{g}} \delta'(\mathbf{u}_\ell \cdot \mathbf{x}) + 2\mathbf{u}_\ell \cdot \partial_{\mathbf{x}} \Delta_{m,\ell}^{\hat{g}} \delta(\mathbf{u}_\ell \cdot \mathbf{x}) \right) \right] \prod_{k=1}^m \Phi_{n_k}^{\mathcal{E}} \right\}.$$

Note that the orthogonality relation allows to include  $\Phi_{n_k}^{\mathcal{E}}$  under the derivative in the last term. Then one can perform the same manipulations with the sum over splittings as in (F.13) but in the inverse direction, which directly leads to the constraint (5.6).

### Proposition 3

Our goal is to find an explicit expression for  $\tilde{\Phi}_n^{(0)}$ . Using the fact that the limit of large  $\mathbf{x}$  of the generalized error function  $\Phi_n^E$  is the product of  $n$  sign functions, one immediately obtains

$$\tilde{\Phi}_n^{(0)}(\mathbf{x}) = \frac{1}{2^{n-1} n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \left[ \prod_{e \in E_{\mathcal{T}}} \mathcal{D}(\mathbf{v}_{s(e)t(e)}) \right] \left[ \prod_{e \in E_{\mathcal{T}}} \text{sgn}(\mathbf{u}_e, \mathbf{x}) \right], \quad (\text{F.18})$$

where  $\mathcal{D}(\mathbf{v})$  are the covariant derivative operators (D.16). The action of the derivatives  $\partial_{\mathbf{x}}$  on the sign functions can be ignored (since the original function is smooth), however there are additional contributions due to the mutual action of the operators  $\mathcal{D}$ . Similar contributions were discussed in a similar context in appendix E, where they were shown to cancel, but here they turn out to leave a finite remainder. The contribution generated by the mutual action of two operators  $\mathcal{D}(\mathbf{v}_{s(e)t(e)})$  is non-vanishing only if the two edges  $e_1, e_2$  have a common vertex. In this case it contributes the factor

$$\frac{(p_{\mathbf{v}_1} p_{\mathbf{v}_2} p_{\mathbf{v}_3})}{2\pi} \text{sgn}(\mathbf{u}_{e_1}, \mathbf{x}) \text{sgn}(\mathbf{u}_{e_2}, \mathbf{x}), \quad (\text{F.19})$$

where  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  are the tree vertices joint by the edges  $e_1 = (\mathbf{v}_2, \mathbf{v}_3)$ ,  $e_2 = (\mathbf{v}_1, \mathbf{v}_3)$ . Again, considering the three trees shown in Fig. 7, one can note that the vectors  $\mathbf{u}_e$  defined by these trees satisfy the following relations:  $\mathbf{u}_{e_1}^{(1)} = -\mathbf{u}_{e_3}^{(2)}$ ,  $\mathbf{u}_{e_2}^{(1)} = \mathbf{u}_{e_3}^{(3)}$ ,  $\mathbf{u}_{e_2}^{(2)} = \mathbf{u}_{e_1}^{(3)}$  and  $\mathbf{u}_{e_2}^{(2)} = \mathbf{u}_{e_1}^{(1)} + \mathbf{u}_{e_2}^{(1)}$ , where  $e_3 = (\mathbf{v}_1, \mathbf{v}_2)$  and we indicated by upper index the tree with respect to which the vector is defined. Therefore, the contributions generated by these trees combine into

$$\frac{(p_{\mathbf{v}_1} p_{\mathbf{v}_2} p_{\mathbf{v}_3})}{2\pi} \left[ \text{sgn}(\mathbf{u}_{e_1}^{(1)}, \mathbf{x}) \text{sgn}(\mathbf{u}_{e_2}^{(1)}, \mathbf{x}) + \text{sgn}(\mathbf{u}_{e_2}^{(2)}, \mathbf{x}) \text{sgn}(\mathbf{u}_{e_3}^{(2)}, \mathbf{x}) - \text{sgn}(\mathbf{u}_{e_1}^{(3)}, \mathbf{x}) \text{sgn}(\mathbf{u}_{e_3}^{(3)}, \mathbf{x}) \right]$$

$$= -\frac{(p_{\mathbf{v}_1} p_{\mathbf{v}_2} p_{\mathbf{v}_3})}{2\pi}, \quad (\text{F.20})$$

where we used the above relations between the vectors and the sign identity (D.12) for  $x_s = (\mathbf{u}_{e_s}^{(1)}, \mathbf{x})$ ,  $s = 1, 2$ . Thus, unlike in (E.5), we now get a non-vanishing result, due to the mutual action of derivative operators.

Each pair of intersecting edges leads to a contribution which recombines the contributions of the edges from the three trees into a single factor (F.20). Furthermore, the sum over trees implies that we have to sum over all possible subtrees  $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$ , in particular, over all possible allocations of different subtrees to the vertices  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ . The sign factors  $\text{sgn}(\mathbf{u}_e, \mathbf{x})$  for edges of these subtrees do not depend on this allocation, but the factors  $\mathcal{D}(\mathbf{v}_{s(e)t(e)})$  do depend for the edges connecting to one of these vertices. It is easy to see that the sum over allocations effectively replaces the three vertices by a single one labelled by the total charge  $p_{\mathbf{v}_1} + p_{\mathbf{v}_2} + p_{\mathbf{v}_3}$ . As a result, one obtains that the function (F.18) can be represented as a sum over marked trees where a mark corresponds to a collapse of a pair of intersecting edges and contributes the factor (F.20). More precisely, one has

$$\tilde{\Phi}_n^{(0)}(\mathbf{x}) = \frac{1}{2^{n-1}n!} \sum_{m=0}^{\lfloor (n-1)/2 \rfloor} \frac{(-1)^m}{(2\pi)^m} \sum_{\mathcal{T} \in \mathbb{T}_{n-2m, m}^\ell} \left[ \prod_{\mathbf{v} \in V_{\mathcal{T}}} \mathcal{P}_{m_{\mathbf{v}}}(\{p_{\mathbf{v}, s}\}) \right] \prod_{e \in E_{\mathcal{T}}} (\mathbf{v}_{s(e)t(e)}, \mathbf{x}) \text{sgn}(\mathbf{u}_e, \mathbf{x}), \quad (\text{F.21})$$

where

$$\mathcal{P}_m(\{p_s\}) = \sum_{\mathcal{I}_1 \cup \dots \cup \mathcal{I}_m = \mathcal{L}_{2m+1}} \prod_{j=1}^m (p_{j_1} p_{j_2} p_{j_3}). \quad (\text{F.22})$$

This factor collects the weights (F.20) assigned to a vertex due to collapse of  $m$  pairs of edges. It is represented as a sum over all possible splittings of the set  $\mathcal{L}_{2m+1} = \{1, \dots, 2m+1\}$  into union of triples  $\mathcal{I}_j = \{j_1, j_2, j_3\}$  such that the labels in one triple  $\mathcal{I}_j$  are all different, two different triples can have at most one common label, and there are no closed cycles in the sense that there are no subsets  $\{j_k\}_{k=1}^r$  such that  $\mathcal{I}_{j_k} \cap \mathcal{I}_{j_{k+1}} \neq \emptyset$  where  $j_{r+1} \equiv j_1$ . This sum simply counts all possible splittings of a collection of  $2m$  joint edges into  $m$  intersecting pairs, suppressing for each pair the distinction between the three configurations of Fig. 7.

However, the representation (F.22) is not very convenient for our purposes. An alternative representation can be obtained by noting that, instead of collapsing all edges at once, one can collapse first one pair, sum over all configurations (i.e. allocations of subtrees), then collapse another pair, and so on. In this approach at each step the sum over different configurations ensures that all factors assigned to the elements of the surviving tree depend only on the sum of charges being collapsed. As a result, one obtains a hierarchical structure described by a rooted ternary tree  $T$ , where the leaves correspond to the vertices of the original unrooted tree which have collapsed into one vertex with  $m_{\mathbf{v}}$  marks corresponding to the root of  $T$ . The other vertices of  $T$  are then in one-to-one correspondence with marked vertices appearing at intermediate stages of the above process.

This procedure gives rise to the representation (5.18) of the weight factor  $\mathcal{P}_m$  in terms of a sum over rooted ternary trees. A non-trivial point which must be taken into account is that the procedure leading to this representation overcounts different configurations. As a result, each tree is weighted by the rational coefficient  $N_{\hat{T}}/m!$  where  $\hat{T}$  is the rooted tree obtained from  $T$  by dropping all leaves and  $N_{\hat{T}}$  is the number of ways of labelling the vertices of  $\hat{T}$  with increasing labels, which already appeared in appendix A. Here the numerator

takes into account that the tree  $T$  is generated  $N_{\hat{T}}$  times in the collapsing process, whereas the denominator removes the overcounting produced by specifying the order of collapsing  $m$  pairs of edges. Finally, we apply Lemma 1 from appendix A. Since  $n_{\hat{T}} = m$ , the coefficients coincides with the inverse of the tree factorial (5.19). Taking into account that substitution of (F.21) into (5.14) gives exactly (5.16), this completes the proof of the proposition.

### Proposition 4

Our aim is to show that the recursive formula (5.23) solves the equations (5.22). We will proceed by induction, starting with the case  $n = 3$  where there is a single unrooted tree and the system (5.22) contains a single equation corresponding to the trivial trees  $\mathcal{T}_r$  consisting of one vertex. Thus, it is solved by  $a_{\bullet\bullet\bullet} = \frac{1}{3}$  consistently with (5.23).

Let us now consider trees with  $n$  vertices assuming that for all trees with less number of vertices (5.23) holds. We start by noting that for every vertex  $\mathbf{v} \in V_{\mathcal{T}_r}$ , among the subtrees  $\mathcal{T}_{r,s}(\mathbf{v}) \subset \mathcal{T}_r$  obtained by removing the vertex  $\mathbf{v}$ , there is one which contains  $\mathbf{v}_r$ , which we denote by  $\mathcal{T}_{r,s_0}(\mathbf{v})$ . (If  $\mathbf{v} = \mathbf{v}_r$ , we take  $\mathcal{T}_{r,s_0}(\mathbf{v}) = \emptyset$ .) Since every tree  $\hat{\mathcal{T}}_r$  is a union of the three trees  $\mathcal{T}_r$ , substituting (5.23) into the l.h.s. of (5.22), one obtains that the sum over vertices of  $\hat{\mathcal{T}}_r$  can be represented as

$$\frac{1}{n} \sum_{r=1}^3 \sum_{\mathbf{v} \in V_{\mathcal{T}_r}} \epsilon_{\mathbf{v}} \left( a_{\hat{\mathcal{T}}_{1,r}(\mathbf{v})} + a_{\hat{\mathcal{T}}_{2,r}(\mathbf{v})} - a_{\hat{\mathcal{T}}_{3,r}(\mathbf{v})} \right) \prod_{\substack{s=1 \\ s \neq s_0}}^{n_{\mathbf{v}}} a_{\mathcal{T}_{r,s}(\mathbf{v})}, \quad (\text{F.23})$$

where  $\hat{\mathcal{T}}_{r',r}(\mathbf{v})$  is obtained from  $\hat{\mathcal{T}}_{r'}$  by replacing  $\mathcal{T}_r$  by  $\mathcal{T}_{r,s_0}(\mathbf{v})$ . Applying the equations (5.22), this expression gives

$$\frac{1}{n} \left[ a_{\mathcal{T}_2} a_{\mathcal{T}_3} \sum_{\mathbf{v} \in V_{\mathcal{T}_1}} \epsilon_{\mathbf{v}} \prod_{s=1}^{n_{\mathbf{v}}} a_{\mathcal{T}_{1,s}(\mathbf{v})} + a_{\mathcal{T}_1} a_{\mathcal{T}_3} \sum_{\mathbf{v} \in V_{\mathcal{T}_2}} \epsilon_{\mathbf{v}} \prod_{s=1}^{n_{\mathbf{v}}} a_{\mathcal{T}_{2,s}(\mathbf{v})} + a_{\mathcal{T}_1} a_{\mathcal{T}_2} \sum_{\mathbf{v} \in V_{\mathcal{T}_3}} \epsilon_{\mathbf{v}} \prod_{s=1}^{n_{\mathbf{v}}} a_{\mathcal{T}_{3,s}(\mathbf{v})} \right]. \quad (\text{F.24})$$

Since the trees  $\mathcal{T}_r$  have less than  $n$  vertices, they are subject to the induction hypothesis which allows to replace the sums over vertices by  $n_r a_{\mathcal{T}_r}$  where  $n_r$  is the number of vertices in  $\mathcal{T}_r$ . Taking into account that  $n_1 + n_2 + n_3 = n$ , the expression (F.24) reduces to  $a_{\mathcal{T}_1} a_{\mathcal{T}_2} a_{\mathcal{T}_3}$ , which proves that the equations (5.22) are indeed satisfied.

### Proposition 5

The evaluation of the large  $\mathbf{x}$  limit of the function  $\tilde{\Phi}_n^{\mathcal{E}}(\mathbf{x})$  is very similar to the calculation done in the proof of Proposition 3. First, using the asymptotics of the generalized error functions, we can write

$$\lim_{\mathbf{x} \rightarrow \infty} \tilde{\Phi}_n^{\mathcal{E}}(\mathbf{x}) = \frac{1}{2^{n-1} n!} \sum_{m=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2m,m}^{\ell}} \left[ \prod_{\mathbf{v} \in V_{\mathcal{T}}} \mathcal{D}_{m_{\mathbf{v}}}(\{\check{\gamma}_{\mathbf{v},s}\}) \right] \left[ \prod_{e \in E_{\mathcal{T}}} \mathcal{D}(\mathbf{v}_{s(e)t(e)}) \right] \left[ \prod_{e \in E_{\mathcal{T}}} \text{sgn}(\mathbf{u}_e, \mathbf{x}) \right]. \quad (\text{F.25})$$

Then note that the two last factors depend only on sums of charges appearing in the operators  $\mathcal{D}_{m_{\mathbf{v}}}$  in the first factor. This implies that the vectors on which these factors depend are

orthogonal to the vectors determining the operators in (5.21). Therefore, these operators effectively act on a constant and can be expanded as

$$\mathcal{D}_m \cdot 1 = \sum_{k=0}^m \mathcal{V}_{m,k}, \quad (\text{F.26})$$

where  $\mathcal{V}_{m,k}$  are homogeneous polynomials in  $\mathbf{x}$  of degree  $2(m-k)$ . In particular, the highest degree term coincides with the function defined in (5.25),  $\mathcal{V}_{m,0}(\{\tilde{\gamma}_s\}) = \mathcal{V}_m(\{\tilde{\gamma}_s\})$ .

Next, the mutual action of the derivative operators from the product over edges in the second factor generates contributions described by trees with pairs of collapsed edges replaced by marks. The difference here is that the original trees were also marked. This fact does not change the structure of the result, which is again given by a sum over marked trees, but it affects the weight associated with marks. Denoting this weight for a vertex with total  $m$  marks (old and new) by  $\mathcal{V}_m^{\text{tot}}$ , we recover the equation (5.24) in the statement of the Proposition, provided that  $\mathcal{V}_m^{\text{tot}} = \mathcal{V}_m$ . We now proceed to prove the latter identity.

The weight factor  $\mathcal{V}_m^{\text{tot}}$  coming from the above procedure is given by

$$\mathcal{V}_m^{\text{tot}}(\{\tilde{\gamma}_s\}) = \text{Sym} \left\{ \sum_{m_0=0}^m \frac{(-1)^{m_0}}{(2\pi)^{m_0}} \sum_{\substack{\sum_{r=1}^{2m_0+1} m_r = m - m_0}} C(\{m_r\}) \mathcal{P}_{m_0}(\{p'_r\}) \prod_{r=1}^{2m_0+1} \sum_{k_r=0}^{m_r} \mathcal{V}_{m_r, k_r}(\tilde{\gamma}_{j_{r-1}+1}, \dots, \tilde{\gamma}_{j_r}) \right\}, \quad (\text{F.27})$$

where the second sum goes over all ordered decompositions of  $m - m_0$  into non-negative integers,  $C(\{m_r\}) = \frac{(2m+1)!}{\prod_r (2m_r+1)!}$ , and we used notations similar to (2.16),

$$j_0 = 0, \quad j_r = m_1 + \dots + m_r, \quad \gamma'_r = \gamma_{j_{r-1}+1} + \dots + \gamma_{j_r}. \quad (\text{F.28})$$

Here the first factor  $\mathcal{P}_{m_0}$  arises due to collapse of  $m_0$  pairs of edges in the marked trees one sums over in (F.25), whereas the factors given by the sums over  $k_r$  are the ones corresponding to the ‘‘old’’ marks assigned to that trees.

Now let us use the equations (5.22) determining the coefficients  $a_{\mathcal{T}}$ . It is easy to see that they imply the following constraint on the next-to-highest degree term in the expansion (F.26)

$$\mathcal{V}_{m,1}(\{\tilde{\gamma}_s\}) = \frac{1}{6 \cdot 2\pi} \sum_{\substack{m_1+m_2+m_3=m-1 \\ m_r \geq 0}} C(\{m_r\}) \text{Sym} \left\{ (p'_1 p'_2 p'_3) \prod_{r=1}^3 \mathcal{V}_{m_r}(\{\tilde{\gamma}_i\}_{i=j_{r-1}+1}^{j_r}) \right\}, \quad (\text{F.29})$$

where  $j_r = \sum_{s=1}^r (2m_s + 1)$  and  $p'_r = \sum_{i=1}^{2m_r+1} p_{j_{r-1}+i}$ . Applying this constraint recursively, one can express all  $\mathcal{V}_{m,k}$  for  $k > 0$  through  $\mathcal{V}_m$ . The idea is to replace the factors  $\mathcal{V}_{m_r}$  by the operators  $\mathcal{D}_{m_r}$ . Then one can realize that, extracting from the resulting function the terms homogeneous in  $\mathbf{x}$  of order  $2(m-2)$  (which have two factors of  $(p^3)$ ), one obtains the result for  $2\mathcal{V}_{m,2}$ . Proceeding in this way, one arrives at the representation very similar to (F.27),

$$\mathcal{V}_{m,k}(\{\tilde{\gamma}_s\}) = \frac{1}{(2\pi)^k} \text{Sym} \left\{ \sum_{\substack{\sum_{r=1}^{2k+1} m_r = m - k}} C(\{m_r\}) \mathcal{P}_k(\{p'_r\}) \prod_{r=1}^{2k+1} \mathcal{V}_{m_r}(\tilde{\gamma}_{j_{r-1}+1}, \dots, \tilde{\gamma}_{j_r}) \right\}. \quad (\text{F.30})$$

Substituting it into (F.27) and using the expression (5.18) for the factors  $\mathcal{P}_m$  through the sum over rooted ternary trees, one can recombine all these sums in the following way

$$\begin{aligned} \mathcal{V}_m^{\text{tot}} = & \text{Sym} \left\{ \sum_{k=0}^m \sum_{\substack{2k+1 \\ \sum_{r=1} m_r = m-k}} C(\{m_r\}) \sum_{T \in \mathbb{T}_{2k+1}^{(3)}(\{\check{\gamma}_r\})} \frac{1}{T!} \sum_{T' \subseteq T} (-1)^{n_{T'}} \prod_{v \in V_{T'}} \frac{n_v(T)}{n_v(T')} \right. \\ & \left. \times \prod_{v \in V_T} \frac{(p_{d_1(v)} p_{d_2(v)} p_{d_3(v)})}{2\pi} \prod_{r=1}^{2k+1} \mathcal{V}_{m_r}(\check{\gamma}_{j_{r-1}+1}, \dots, \check{\gamma}_{j_r}) \right\}, \end{aligned} \quad (\text{F.31})$$

where the sum over  $T'$  is the sum over subtrees of  $T$  having the same root. In terms of the variables appearing in (F.27), one can identify  $k = m_0 + \sum_r k_r$  and  $n_{T'} = m_0$ . Thus,  $T'$  is the tree appearing in the decomposition of  $\mathcal{P}_{m_0}$ , whereas  $T$  is its union with  $2m_0 + 1$  trees  $T_r$  appearing in the decomposition of  $\mathcal{P}_{k_r}$ , i.e.  $T = T' \cup (\cup_r T_r)$  where  $T_r$  are the trees rooted at leaves of  $T'$ . Finally, we took into account that for such trees one has

$$T! = T'! \prod_r T_r! \prod_{v \in V_{T'}} \frac{n_v(T)}{n_v(T')}. \quad (\text{F.32})$$

Remarkably, the sum over subtrees in (F.31) factorizes and for a fixed number of vertices  $n_{T'} = m_0$  is subject to Theorem 1 where the role of the trees is played by rooted ternary trees  $T$  and  $T'$  after stripping out their leaves. As a result, one finds for  $k > 0$

$$\sum_{T' \subseteq T} (-1)^{n_{T'}} \prod_{v \in V_{T'}} \frac{n_v(T)}{n_v(T')} = \sum_{m_0=0}^k \frac{(-1)^{m_0} k!}{m_0! (k - m_0)!} = (1 - 1)^k = 0. \quad (\text{F.33})$$

Thus, the only non-vanishing contribution is the one with  $k = 0$  which coincides with  $\mathcal{V}_m$ . This is what we had to show and therefore completes the proof of the Proposition.

### Proposition 6

For clarity, let us first show that the recursive equation (5.9) is satisfied by the contribution to  $g_n^{(0)}(\{\check{\gamma}_i, c_i\})$  given by the trees without any marks, i.e. by the function

$$g_n^*(\{\check{\gamma}_i, c_i\}) = \frac{(-1)^{n-1+\sum_{i<j} \gamma_{ij}}}{2^{n-1} n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \prod_{e \in E_{\mathcal{T}}} \gamma_{s(e)t(e)} \text{sgn}(S_e). \quad (\text{F.34})$$

Then the inclusion of marks will be straightforward because the corresponding contributions can be dealt with essentially in the same way as the contribution (F.34).

To start with, we substitute  $g_n^*$  into the l.h.s. of the recursive equation and decompose  $\Gamma_{n\ell} = -\sum_{i=1}^{\ell} \sum_{j=\ell+1}^n \gamma_{ij}$ . Then the crucial observation is that this double sum, the sum over  $\ell$  and the two sums over trees (over  $\mathbb{T}_\ell^\ell$  and  $\mathbb{T}_{n-\ell}^\ell$ ) are all equivalent to a single sum over trees with  $n$  vertices, i.e. over  $\mathbb{T}_n^\ell$ , supplemented by the sum over edges. Namely, one can do the following replacement

$$\frac{1}{2} \text{Sym} \sum_{\ell=1}^{n-1} \frac{1}{\ell!(n-\ell)!} \sum_{\mathcal{T}_L \in \mathbb{T}_\ell^\ell} \sum_{\mathcal{T}_R \in \mathbb{T}_{n-\ell}^\ell} \sum_{i=1}^{\ell} \sum_{j=\ell+1}^n = \frac{1}{n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \sum_{e \in E_{\mathcal{T}}}. \quad (\text{F.35})$$

The idea is that on the l.h.s. one sums over all possible splittings of unrooted labelled trees with  $n$  vertices into two trees with  $\ell$  and  $n - \ell$  vertices. Such splitting can be done by cutting an edge and then  $i, j$  correspond to the labels of the vertices joined by the cutting edge. The relative coefficient  $\frac{n!}{\ell!(n-\ell)!}$  takes into account that after splitting the vertices of the first tree can have arbitrary labels from the set  $\{1, \dots, n\}$  and not necessarily  $\{1, \dots, \ell\}$ , whereas  $\frac{1}{2}$  avoids doubling due to the symmetry between  $\mathcal{T}_L$  and  $\mathcal{T}_R$ .

It is easy to check that all factors in (5.9) fit this interpretation and the l.h.s. takes the following form

$$\frac{(-1)^{n-1+\sum_{i<j}\gamma_{ij}}}{2^{n-1}n!} \sum_{\mathcal{T} \in \mathbb{T}_n^\ell} \prod_{e \in E_{\mathcal{T}}} \gamma_{s(e)t(e)} \sum_{e \in E_{\mathcal{T}}} (\text{sgn}(S_e) - \text{sgn}(\Gamma_e)) \prod_{e' \in E_{\mathcal{T}} \setminus \{e\}} \text{sgn} \left( S_{e'} - \frac{\Gamma_{e'}}{\Gamma_e} S_e \right), \quad (\text{F.36})$$

where  $\Gamma_e$  was defined in (5.29). Finally, we apply the following sign identity, established in [19, Eq.(A.7)],

$$\sum_{\beta=1}^m (\text{sgn}(x_\beta) - 1) \prod_{\substack{\alpha=1 \\ \alpha \neq \beta}}^m \text{sgn}(x_\alpha - x_\beta) = \prod_{\alpha=1}^m \text{sgn}(x_\alpha) - 1, \quad (\text{F.37})$$

where one should take the label  $\alpha$  to run over  $m = n - 1$  edges of a tree  $\mathcal{T}$ , identify  $x_\alpha = S_\alpha/\Gamma_\alpha$ , and multiply it by  $\prod_{\alpha=1}^{n-1} \text{sgn}(\Gamma_\alpha)$ . Then the expression (F.36) reduces to  $g_n^*(\{\tilde{\gamma}_i, c_i\}) - g_n^*(\{\tilde{\gamma}_i, \beta_{ni}\})$ , i.e. the r.h.s. of (5.9) evaluated for function (F.34). This proves that this function solves the recursive equation.

The generalization of this proof to the full ansatz (5.26) is elementary. Instead of the relation (F.35), one now has

$$\frac{1}{2} \text{Sym} \sum_{\ell=1}^{n-1} \frac{1}{\ell!(n-\ell)!} \sum_{k_L=0}^{[(\ell-1)/2]} \sum_{\mathcal{T}_L \in \mathbb{T}_{\ell-2k_L, k_L}^\ell} \sum_{k_R=0}^{[(n-\ell-1)/2]} \sum_{\mathcal{T}_R \in \mathbb{T}_{n-\ell-2k_R, k_R}^\ell} \sum_{i=1}^{\ell} \sum_{j=\ell+1}^n = \frac{1}{n!} \sum_{k=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2k, k}^\ell} \sum_{e \in E_{\mathcal{T}}}, \quad (\text{F.38})$$

which again reflects the fact the sum over marked unrooted labelled trees can be represented as a sum over all possible splittings into two such trees by cutting them along an edge. If the cutting edge joins a marked vertex, it counts  $1 + 2m_v$  times, producing the right factor  $\gamma_{v_L v_R}$ , which is reflected by the fact that the sums over  $i, j$  run over  $\ell$  and  $n - \ell$  values, respectively. Other numerical factors work in the same way as before. Thus, the l.h.s. of (5.9) can again be rewritten as in (F.36) with the only difference that  $\sum_{\mathcal{T} \in \mathbb{T}_n^\ell}$  should now be replaced by  $\sum_{k=0}^{[(n-1)/2]} \sum_{\mathcal{T} \in \mathbb{T}_{n-2k, k}^\ell} \prod_{v \in V_{\mathcal{T}}} \tilde{\mathcal{V}}_v$ . Applying the same sign identity (F.37) for  $m = n - 1 - 2k$  and the same identification for  $x_\alpha$ , one recovers the r.h.s. of (5.9). This completes the proof of the proposition.

### Proposition 7

This Proposition trivially follows from (5.8).

### Proposition 8

The easiest way to prove (5.33) is to substitute it into (5.4) and then check that the result is consistent with the constraint (5.11). The substitution generates a sum over trees which



resemble the blooming trees of appendix C: these are trees with vertices from which other trees grows. But now the two types of trees, representing the ‘base’ and the ‘flowers’, are actually the same — both of them are Schröder trees. The only difference between them is that vertices of the ‘base’ carry weights  $\mathcal{E}_v^{(+)}$ , whereas the vertices of ‘flowers’ have weights  $\mathcal{E}_v^{(0)}$ . Equivalently, this can be written as

$$W_n = \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^S} (-1)^{n_T} \sum_{T' \subseteq T} \prod_{v' \in V_{T'}} \mathcal{E}_{v'}^{(+)} \prod_{v \in V_T \setminus V_{T'}} \mathcal{E}_v^{(0)} \right\}, \quad (\text{F.39})$$

where the second sum goes over all subtrees of  $T$  containing its root. It is clear that the two contributions differing only by whether a vertex  $v$  is included into  $T'$  or not, combine into the factor  $\mathcal{E}_v$  assigned to this vertex. Thus, the sum over subtrees can be evaluated explicitly and gives

$$W_n = \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^S} (-1)^{n_T} \mathcal{E}_{v_0}^{(+)} \prod_{v \in V_T \setminus \{v_0\}} \mathcal{E}_v \right\}. \quad (\text{F.40})$$

Given this result, the proof of the constraint (5.11) is analogous to the proof of the relation (F.4): the contribution of each tree  $T$  (from the sum in (F.40)) and a splitting with  $n_k > 1$  (from the sum in (5.11)) is cancelled by the contribution of another tree obtained from  $T$  by adding  $n_k$  children to its  $k$ th leaf and the same splitting but with  $n_k$  replaced by  $1 + \dots + 1$  ( $n_k$  times). The only contribution which survives is the one generated by the tree with one vertex and  $n$  leaves and the splitting with all  $n_k = 1$ . It is given by  $\mathcal{E}_n^{(+)}$ , which verifies the constraint and proves the proposition.

### Proposition 9

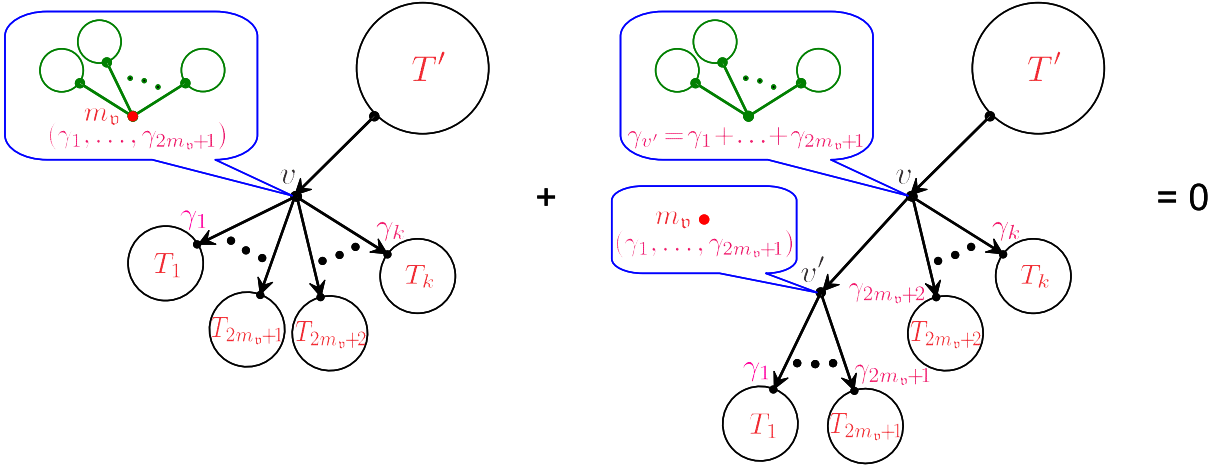
Our goal is to prove that in the expression

$$g_{\text{tr},n} = \text{Sym} \left\{ \sum_{T \in \mathbb{T}_n^S} (-1)^{n_T-1} (g_{v_0}^{(0)} - \mathcal{E}_{v_0}^{(0)}) \prod_{v \in V_T \setminus \{v_0\}} \mathcal{E}_v^{(0)} \right\}, \quad (\text{F.41})$$

following from (5.32), all contributions due to marked trees cancel leaving only contributions of trees without marks, i.e with  $m = 0$ . There are several possible situations which we need to analyze.

First, let us consider the contributions generated by non-trivial marked trees, i.e. trees having more than one vertex and at least one mark. Let us focus on the contribution corresponding to a vertex  $\mathbf{v}$  with  $m_{\mathbf{v}} > 0$  marks of a tree  $\mathcal{T}$ , which appears in the sum over marked unrooted trees living at a vertex  $v$  of a Schröder tree  $T$ . Let  $k = n_{\mathbf{v}}$  be the number of children of vertex  $v$  and  $\gamma_i$  ( $i = 1, \dots, k$ ) their charges so that  $\gamma_s$  ( $s = 1, \dots, 2m_{\mathbf{v}} + 1$ ) are the charges labelling the marked vertex  $\mathbf{v}$ . Note that  $k \geq 2m_{\mathbf{v}} + 2$  because the tree  $\mathcal{T}$  has at least one additional vertex except  $\mathbf{v}$ . Then the contribution we described is cancelled by the contribution coming from another Schröder tree, which is obtained from  $T$  by adding an edge connecting the vertex  $v$  to a new vertex  $v'$ , whose children are the  $2m_{\mathbf{v}} + 1$  children of  $v$  in  $T$  carrying charges  $\gamma_s$  (see Fig. 11).<sup>27</sup> Indeed, choosing in the sum over marked trees at vertex

<sup>27</sup>The new tree is of Schröder type because its vertex  $v$  has  $k - 2m_{\mathbf{v}} \geq 2$  children and vertex  $v'$  has  $2m_{\mathbf{v}} + 1 \geq 3$  children.



**Figure 11:** Combination of two Schröder trees ensuring the cancellation of contributions generated by marked trees.

$v$  the same tree  $\mathcal{T}$  as before, but now with  $m_v = 0$ , and in the sum at vertex  $v'$  the trivial tree having one vertex and  $m_v$  marks, one gets exactly the same contribution as before, but now with an opposite sign due to the presence of an additional vertex in the Schröder tree. Thus, all contributions from non-trivial marked trees are cancelled.

As a result, we remain only with the contributions generated by trivial marked trees, i.e. having only one vertex and  $m_v$  marks. One has to distinguish two cases: either the corresponding vertex  $v$  of the Schröder tree is the root or not. In the former case, this contribution is trivially cancelled in the difference  $g_{v_0}^{(0)} - \mathcal{E}_{v_0}^{(0)}$  in (F.41). In the latter case, this is precisely the contribution used above to cancel the contributions from non-trivial marked trees. This exhausts all possibilities and we arrive at the formula (5.36).

## G. Explicit results up to 4th order

In this appendix we provide explicit expressions for various functions appearing in our construction up to the fourth order. To write them down, we will use the shorthand notation  $\gamma_{i+j} = \gamma_i + \gamma_j$ ,  $c_{i+j} = c_i + c_j$ , etc. as well as indicate the arguments of functions through their indices, for instance,  $\mathcal{E}_{i_1 \dots i_n} = \mathcal{E}_n(\check{\gamma}_{i_1}, \dots, \check{\gamma}_{i_n})$ . These expressions are obtained by simple substitutions using the results found in the main text and the sets of trees shown in Fig. 1. For  $n = 2$  they all agree with the results of [20].

The results (5.32) and (5.33) generate the following expansions

$$\begin{aligned}
h_{p,q}^{\text{DT}} &= \widehat{h}_{p,\mu} + \sum_{\check{\gamma}_1 + \check{\gamma}_2 = \check{\gamma}} \left[ g_{12}^{(0)} - \mathcal{E}_{12} \right] e^{\pi i \tau Q_2(\{\check{\gamma}_i\})} \widehat{h}_{p_1, \mu_1} \widehat{h}_{p_2, \mu_2} \\
&+ \sum_{\sum_{i=1}^3 \check{\gamma}_i = \check{\gamma}} \left[ g_{123}^{(0)} - \mathcal{E}_{123} - 2 \left( g_{1+2,3}^{(0)} - \mathcal{E}_{1+2,3} \right) \mathcal{E}_{12} \right] e^{\pi i \tau Q_3(\{\check{\gamma}_i\})} \prod_{i=1}^3 \widehat{h}_{p_i, \mu_i} \\
&+ \sum_{\sum_{i=1}^4 \check{\gamma}_i = \check{\gamma}} \left[ g_{1234}^{(0)} - \mathcal{E}_{1234} - 2 \left( g_{1+2+3,4}^{(0)} - \mathcal{E}_{1+2+3,4} \right) (\mathcal{E}_{123} - 2\mathcal{E}_{1+2,3}\mathcal{E}_{12}) \right. \\
&\quad \left. - 3 \left( g_{1+2,34}^{(0)} - \mathcal{E}_{1+2,34} \right) \mathcal{E}_{12} + \left( g_{1+2,3+4}^{(0)} - \mathcal{E}_{1+2,3+4} \right) \mathcal{E}_{12}\mathcal{E}_{34} \right] e^{\pi i \tau Q_4(\{\check{\gamma}_i\})} \prod_{i=1}^4 \widehat{h}_{p_i, \mu_i} + \dots,
\end{aligned} \tag{G.1}$$

$$\begin{aligned}
\widehat{h}_{p,q} &= h_{p,\mu} + \sum_{\tilde{\gamma}_1 + \tilde{\gamma}_2 = \tilde{\gamma}} \mathcal{E}_{12}^{(+)} e^{\pi i \tau Q_2(\{\tilde{\gamma}_i\})} h_{p_1, \mu_1} h_{p_2, \mu_2} \\
&+ \sum_{\sum_{i=1}^3 \tilde{\gamma}_i = \tilde{\gamma}} \left[ \mathcal{E}_{123}^{(+)} - 2\mathcal{E}_{1+2,3}^{(+)} \mathcal{E}_{12}^{(0)} \right] e^{\pi i \tau Q_3(\{\tilde{\gamma}_i\})} \prod_{i=1}^3 h_{p_i, \mu_i} \\
&+ \sum_{\sum_{i=1}^4 \tilde{\gamma}_i = \tilde{\gamma}} \left[ \mathcal{E}_{1234}^{(+)} - 2\mathcal{E}_{1+2+3,4}^{(+)} \left( \mathcal{E}_{123}^{(0)} - 2\mathcal{E}_{1+2,3}^{(0)} \mathcal{E}_{12}^{(0)} \right) - 3\mathcal{E}_{1+2,34}^{(+)} \mathcal{E}_{12}^{(0)} \right. \\
&\quad \left. + \mathcal{E}_{1+2,3+4}^{(+)} \mathcal{E}_{12}^{(0)} \mathcal{E}_{34}^{(0)} \right] e^{\pi i \tau Q_4(\{\tilde{\gamma}_i\})} \prod_{i=1}^4 h_{p_i, \mu_i} + \dots,
\end{aligned} \tag{G.2}$$

where the functions  $g_n^{(0)}$  and  $\mathcal{E}_n$  can be read off from (5.26), (5.31) and (5.30),

$$\begin{aligned}
g_2^{(0)} &= \frac{(-1)^{1+\gamma_{12}}}{4} \gamma_{12} \operatorname{sgn}(c_1), \\
g_3^{(0)} &= \frac{(-1)^{1+\gamma_{12}+\gamma_{1+2,3}}}{8} \operatorname{Sym} \left\{ \gamma_{12} \gamma_{23} \operatorname{sgn}(c_1) \operatorname{sgn}(c_3) + \frac{1}{3} \gamma_{12} \gamma_{23} \right\}, \\
g_4^{(0)} &= \frac{(-1)^{\gamma_{12}+\gamma_{1+2,3}+\gamma_{1+2+3,4}}}{16} \operatorname{Sym} \left\{ \gamma_{12} \gamma_{23} \gamma_{34} \operatorname{sgn}(c_1) \operatorname{sgn}(c_{1+2}) \operatorname{sgn}(c_4) \right. \\
&\quad \left. - \frac{1}{3} \gamma_{12} \gamma_{23} \gamma_{24} \operatorname{sgn}(c_1) \operatorname{sgn}(c_3) \operatorname{sgn}(c_4) - \frac{1}{3} \gamma_{12} \gamma_{23} \gamma_{1+2+3,4} \operatorname{sgn}(c_4) \right\},
\end{aligned} \tag{G.3}$$

$$\begin{aligned}
\mathcal{E}_{12} &= \frac{(-1)^{\gamma_{12}}}{4\sqrt{2\tau_2}} \tilde{\Phi}_1^E(\mathbf{v}_{12}, \mathbf{v}_{12}) \\
&= \frac{(-1)^{\gamma_{12}}}{4} \left[ \gamma_{12} E_1 \left( \frac{\sqrt{2\tau_2} |\gamma_{12}|}{\sqrt{(pp_1 p_2)}} \right) + \frac{\sqrt{(pp_1 p_2)}}{\pi \sqrt{2\tau}} e^{-\frac{2\pi \tau_2 \gamma_{12}^2}{(pp_1 p_2)}} \right], \\
\mathcal{E}_{123} &= \frac{(-1)^{\gamma_{12}+\gamma_{1+2,3}}}{8} \operatorname{Sym} \left\{ \frac{1}{2\tau_2} \tilde{\Phi}_2^E((\mathbf{v}_{1,2+3}, \mathbf{v}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{23})) \right. \\
&\quad \left. - \frac{1}{3} \left( \gamma_{12} \gamma_{23} - \frac{(p_1 p_2 p_3)}{4\pi \tau_2} \right) \right\}, \\
\mathcal{E}_{1234} &= \frac{(-1)^{\gamma_{12}+\gamma_{1+2,3}+\gamma_{1+2+3,4}}}{16\sqrt{2\tau_2}} \operatorname{Sym} \left\{ \frac{1}{2\tau_2} \left( \tilde{\Phi}_3^E((\mathbf{v}_{1,2+3+4}, \mathbf{v}_{1+2,3+4}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{12}, \mathbf{v}_{23}, \mathbf{v}_{34})) \right) \right. \\
&\quad \left. + \frac{1}{3} \tilde{\Phi}_3^E((\mathbf{v}_{1,2+3+4}, \mathbf{v}_{1+2+4,3}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{12}, \mathbf{v}_{23}, \mathbf{v}_{24})) \right) \\
&\quad \left. - \frac{1}{3} \left( \gamma_{12} \gamma_{23} - \frac{(p_1 p_2 p_3)}{4\pi \tau_2} \right) \tilde{\Phi}_1^E(\mathbf{v}_{1+2+3,4}, \mathbf{v}_{1+2+3,4}) \right\},
\end{aligned} \tag{G.4}$$

where all generalized error functions are evaluated at  $\mathbf{x} = \sqrt{2\tau_2}(\mathbf{q} + \mathbf{b})$ . The last terms appearing in the above quantities for  $n = 3$  and  $n = 4$  correspond to contributions of trees with one mark ( $m = 1$ ). In fact, at these orders these results can be rewritten in a simpler form, which coincides with the representations (5.38) and (5.44) (in the latter formula one should drop the sum over partitions and take  $c_T = c_n$ ):

$$\begin{aligned}
g_3^{(0)} &= \frac{(-1)^{1+\gamma_{12}+\gamma_{1+2,3}}}{12} \operatorname{Sym} \left\{ \gamma_{12} \gamma_{1+2,3} \operatorname{sgn}(c_1) \operatorname{sgn}(c_3) \right\}, \\
g_4^{(0)} &= \frac{(-1)^{\gamma_{12}+\gamma_{1+2,3}+\gamma_{1+2+3,4}}}{96} \operatorname{Sym} \left\{ \left( 2 \gamma_{23} \gamma_{1,2+3} \gamma_{1+2+3,4} + \gamma_{12} \gamma_{34} \gamma_{1+2,3+4} \right) \operatorname{sgn}(c_1) \operatorname{sgn}(c_{1+2}) \operatorname{sgn}(c_4) \right\},
\end{aligned} \tag{G.5}$$

$$\begin{aligned}
\mathcal{E}_{123} &= \frac{(-1)^{\gamma_{12}+\gamma_{1+2,3}}}{24\tau_2} \text{Sym} \left\{ \tilde{\Phi}_2^E((\mathbf{v}_{1,2+3}, \mathbf{v}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{1+2,3})) \right\}, \\
\mathcal{E}_{1234} &= \frac{(-1)^{\gamma_{12}+\gamma_{1+2,3}+\gamma_{1+2+3,4}}}{96(2\tau_2)^{3/2}} \text{Sym} \left\{ 2 \tilde{\Phi}_3^E((\mathbf{v}_{1,2+3+4}, \mathbf{v}_{1+2,3+4}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{23}, \mathbf{v}_{1,2+3}, \mathbf{v}_{1+2+3,4})) \right. \\
&\quad \left. + \tilde{\Phi}_3^E((\mathbf{v}_{1,2+3+4}, \mathbf{v}_{1+2,3+4}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{12}, \mathbf{v}_{34}, \mathbf{v}_{1+2,3+4})) \right\}.
\end{aligned} \tag{G.6}$$

The simplest way to prove the equality of the two representations of  $g_n^{(0)}$  is to expand the DSZ products appearing in (G.5) into elementary  $\gamma_{ij}$ 's and then, using symmetrization, bring all their products to the form appearing in (G.3). These products are multiplied by combinations of sign functions which can be recombined with the help of the identity (D.12). As an example, let us perform these manipulations for  $n = 3$ :

$$\begin{aligned}
&\text{Sym} \left\{ \gamma_{12} \gamma_{1+2,3} \text{sgn}(c_1) \text{sgn}(c_3) \right\} = \text{Sym} \left\{ \gamma_{12} \gamma_{23} (\text{sgn}(c_1) - \text{sgn}(c_2)) \text{sgn}(c_3) \right\} \\
&= \text{Sym} \left\{ \gamma_{12} \gamma_{23} \left( \text{sgn}(c_1) \text{sgn}(c_3) + \frac{1}{2} \text{sgn}(c_{1+3}) (\text{sgn}(c_1) + \text{sgn}(c_3)) \right) \right\} \\
&= \frac{1}{2} \text{Sym} \left\{ \gamma_{12} \gamma_{23} (3 \text{sgn}(c_1) \text{sgn}(c_3) + 1) \right\},
\end{aligned} \tag{G.7}$$

where we used that  $c_2 = -(c_1 + c_3)$ . This identity then shows the equality of the two forms of  $g_3^{(0)}$  given above. For  $g_4^{(0)}$  the manipulations are very similar, but a bit more cumbersome. The equality of the two forms of  $\mathcal{E}_n$  follows from the equality of their asymptotics  $\mathcal{E}_n^{(0)}$ , which is in turn ensured by the same identities as for  $g_n^{(0)}$ .

Finally, we provide expressions for the kernels  $\widehat{\Phi}_n^{\text{tot}}$  of the indefinite theta series appearing in the expansion (5.34) of  $\mathcal{G}$  in powers of  $\widehat{h}_{p,\mu}$ . For the first two orders, one has

$$\begin{aligned}
\widehat{\Phi}_1^{\text{tot}} &= \Phi_1^f, \\
\widehat{\Phi}_2^{\text{tot}} &= \Phi_2^f + \Phi_1^f \Phi_2^{\widehat{g}} = \frac{1}{4} \Phi_1^f \left( \tilde{\Phi}_1^E(\mathbf{u}_{12}, \mathbf{v}_{12}) - \tilde{\Phi}_1^E(\mathbf{v}_{12}, \mathbf{v}_{12}) \right),
\end{aligned} \tag{G.8}$$

where  $\Phi_1^f(x)$  is defined in (4.13). At the next order,

$$\widehat{\Phi}_3^{\text{tot}} = \Phi_3^f + 2 \text{Sym} \left\{ \Phi_2^f(x_{1+2}, x_3) \Phi_2^{\widehat{g}}(x_1, x_2) \right\} + \Phi_1^f \Phi_3^{\widehat{g}}. \tag{G.9}$$

To get an explicit expression in terms of smooth solutions of Vignéras' equation, one should use the relation (D.15). Applying it to the case  $n = 2$  with  $\mathcal{V} = (\mathbf{u}_{1,2+3}, \mathbf{u}_{1+2,3})$ , using the orthogonality properties

$$\mathbf{u}_{(1,2+3)\perp(1+2,3)} = \mathbf{u}_{12}, \quad \mathbf{u}_{(1+2,3)\perp(1,2+3)} = \mathbf{u}_{23}, \tag{G.10}$$

and acting by the operator  $\mathcal{D}(\mathbf{v}_{12})\mathcal{D}(\mathbf{v}_{23})$ , one can show that

$$\begin{aligned}
&\text{Sym} \left\{ \tilde{\Phi}_2^E((\mathbf{u}_{1,2+3}, \mathbf{u}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{23})) \right\} = \text{Sym} \left\{ \tilde{\Phi}_2^M((\mathbf{u}_{1,2+3}, \mathbf{u}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{23})) \right. \\
&\quad + (\mathbf{v}_{12}, \mathbf{x}) \text{sgn}(\mathbf{u}_{12}, \mathbf{x}) \tilde{\Phi}_1^M(\mathbf{u}_{1+2,3}, \mathbf{v}_{1+2,3}) + (\mathbf{v}_{12}, \mathbf{x}) (\mathbf{v}_{23}, \mathbf{x}) \text{sgn}(\mathbf{u}_{1,2+3}, \mathbf{x}) \text{sgn}(\mathbf{u}_{1+2,3}, \mathbf{x}) \\
&\quad \left. - \frac{(p_1 p_2 p_3)}{6\pi} \right\}.
\end{aligned} \tag{G.11}$$

This result allows to obtain the following representation for the kernel

$$\begin{aligned} \widehat{\Phi}_3^{\text{tot}} = \frac{1}{8} \Phi_1^f \text{Sym} \left\{ \widetilde{\Phi}_2^E((\mathbf{u}_{1,2+3}, \mathbf{u}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{23})) - \widetilde{\Phi}_2^E((\mathbf{v}_{1,2+3}, \mathbf{v}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{23})) \right. \\ \left. - \left( \widetilde{\Phi}_1^E(\mathbf{u}_{1+2,3}, \mathbf{v}_{1+2,3}) - \widetilde{\Phi}_1^E(\mathbf{v}_{1+2,3}, \mathbf{v}_{1+2,3}) \right) \widetilde{\Phi}_1^E(\mathbf{v}_{12}, \mathbf{v}_{12}) \right\}. \end{aligned} \quad (\text{G.12})$$

Using the identity between functions  $\widetilde{\Phi}_2^E$  implied by the identity (G.7), the kernel can also be rewritten as

$$\begin{aligned} \widehat{\Phi}_3^{\text{tot}} = \frac{1}{8} \Phi_1^f \text{Sym} \left\{ \frac{2}{3} \left( \widetilde{\Phi}_2^E((\mathbf{u}_{1,2+3}, \mathbf{u}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{1+2,3})) - \widetilde{\Phi}_2^E((\mathbf{v}_{1,2+3}, \mathbf{v}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{1+2,3})) \right) \right. \\ \left. - \left( \widetilde{\Phi}_1^E(\mathbf{u}_{1+2,3}, \mathbf{v}_{1+2,3}) - \widetilde{\Phi}_1^E(\mathbf{v}_{1+2,3}, \mathbf{v}_{1+2,3}) \right) \widetilde{\Phi}_1^E(\mathbf{v}_{12}, \mathbf{v}_{12}) \right\}, \end{aligned} \quad (\text{G.13})$$

so that the vectors appearing in the second argument of  $\widetilde{\Phi}_2^E$  are now mutually orthogonal.

For  $n = 4$ , the kernel is given by

$$\begin{aligned} \widehat{\Phi}_4^{\text{tot}} = \Phi_4^f + \text{Sym} \left\{ 3\Phi_3^f(x_{1+2}, x_3, x_4)\Phi_2^{\widehat{g}}(x_1, x_2) + \Phi_2^f(x_{1+2}, x_{3+4})\Phi_2^{\widehat{g}}(x_1, x_2)\Phi_2^{\widehat{g}}(x_3, x_4) \right. \\ \left. + 2\Phi_2^f(x_{1+2+3}, x_4)\Phi_3^{\widehat{g}}(x_1, x_2, x_3) \right\} + \Phi_1^f\Phi_4^{\widehat{g}}. \end{aligned} \quad (\text{G.14})$$

Proceeding in the same way as for  $n = 3$ , obtaining a generalization of (G.11) to  $n = 4$ , one arrives at

$$\begin{aligned} \widehat{\Phi}_4^{\text{tot}} = \frac{1}{16} \Phi_1^f \text{Sym} \left\{ \widetilde{\Phi}_3^E((\mathbf{u}_{1,2+3+4}, \mathbf{u}_{1+2,3+4}, \mathbf{u}_{1+2+3,4}), (\mathbf{v}_{12}, \mathbf{v}_{23}, \mathbf{v}_{34})) \right. \\ \left. - \widetilde{\Phi}_3^E((\mathbf{v}_{1,2+3+4}, \mathbf{v}_{1+2,3+4}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{12}, \mathbf{v}_{23}, \mathbf{v}_{34})) \right. \\ \left. + \frac{1}{3} \left( \widetilde{\Phi}_3^E((\mathbf{u}_{1,2+3+4}, \mathbf{u}_{1+2+4,3}, \mathbf{u}_{1+2+3,4}), (\mathbf{v}_{12}, \mathbf{v}_{24}, \mathbf{v}_{34})) \right. \right. \\ \left. \left. - \widetilde{\Phi}_3^E((\mathbf{v}_{1,2+3+4}, \mathbf{v}_{1+2+4,3}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{12}, \mathbf{v}_{24}, \mathbf{v}_{34})) \right) \right. \\ \left. - \left( \widetilde{\Phi}_2^E((\mathbf{u}_{1+2,3+4}, \mathbf{u}_{1+2+3,4}), (\mathbf{v}_{1+2,3}, \mathbf{v}_{34})) - \widetilde{\Phi}_2^E((\mathbf{v}_{1+2,3+4}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{1+2,3}, \mathbf{v}_{34})) \right) \right. \\ \left. + \frac{1}{2} \left( \widetilde{\Phi}_2^E((\mathbf{u}_{1+2+4,3}, \mathbf{u}_{1+2+3,4}), (\mathbf{v}_{1+2,3}, \mathbf{v}_{1+2,4})) \right. \right. \\ \left. \left. - \widetilde{\Phi}_2^E((\mathbf{v}_{1+2+4,3}, \mathbf{v}_{1+2+3,4}), (\mathbf{v}_{1+2,3}, \mathbf{v}_{1+2,4})) \right) \right) \widetilde{\Phi}_1^E(\mathbf{v}_{12}, \mathbf{v}_{12}) \\ \left. - \left( \widetilde{\Phi}_1^E(\mathbf{u}_{1+2+3,4}, \mathbf{v}_{1+2+3,4}) - \widetilde{\Phi}_1^E(\mathbf{v}_{1+2+3,4}, \mathbf{v}_{1+2+3,4}) \right) \widetilde{\Phi}_2^E((\mathbf{v}_{1,2+3}, \mathbf{v}_{1+2,3}), (\mathbf{v}_{12}, \mathbf{v}_{23})) \right. \\ \left. + \left( \widetilde{\Phi}_1^E(\mathbf{u}_{1+2+3,4}, \mathbf{v}_{1+2+3,4}) - \widetilde{\Phi}_1^E(\mathbf{v}_{1+2+3,4}, \mathbf{v}_{1+2+3,4}) \right) \widetilde{\Phi}_1^E(\mathbf{v}_{12}, \mathbf{v}_{12}) \widetilde{\Phi}_1^E(\mathbf{v}_{1+2,3}, \mathbf{v}_{1+2,3}) \right. \\ \left. + \frac{1}{4} \left( \widetilde{\Phi}_1^E(\mathbf{u}_{1+2,3+4}, \mathbf{v}_{1+2,3+4}) - \widetilde{\Phi}_1^E(\mathbf{v}_{1+2,3+4}, \mathbf{v}_{1+2,3+4}) \right) \widetilde{\Phi}_1^E(\mathbf{v}_{12}, \mathbf{v}_{12}) \widetilde{\Phi}_1^E(\mathbf{v}_{34}, \mathbf{v}_{34}) \right\}. \end{aligned} \quad (\text{G.15})$$

It is possible also to rewrite this expression in terms of generalized error functions  $\widetilde{\Phi}_n^E(\mathcal{V}, \tilde{\mathcal{V}})$  where the vectors entering the second argument are mutually orthogonal, as in (G.13). The reader can easily guess the result by comparing (G.4) and (G.6).

The explicit results for  $\widehat{\Phi}_n^{\text{tot}}$ ,  $n \leq 4$ , presented above are the basis for the conjectural formula (5.35). Note that all terms in these expressions have the sum of ranks of the generalized error functions equal to  $n - 1$ . This shows that all contributions due to trees with non-zero number of marks cancel in the sum over Schröder trees.

## References

- [1] A. Strominger and C. Vafa, “Microscopic origin of the Bekenstein-Hawking entropy,” *Phys. Lett.* **B379** (1996) 99–104, [hep-th/9601029](#).
- [2] J. M. Maldacena, A. Strominger, and E. Witten, “Black hole entropy in M-theory,” *JHEP* **12** (1997) 002, [hep-th/9711053](#).
- [3] D. Gaiotto, A. Strominger, and X. Yin, “The M5-brane elliptic genus: Modularity and BPS states,” *JHEP* **08** (2007) 070, [hep-th/0607010](#).
- [4] J. de Boer, M. C. N. Cheng, R. Dijkgraaf, J. Manschot, and E. Verlinde, “A farey tail for attractor black holes,” *JHEP* **11** (2006) 024, [hep-th/0608059](#).
- [5] R. Dijkgraaf, J. M. Maldacena, G. W. Moore, and E. P. Verlinde, “A black hole Farey tail,” [hep-th/0005003](#).
- [6] F. Denef and G. W. Moore, “Split states, entropy enigmas, holes and halos,” *JHEP* **1111** (2011) 129, [hep-th/0702146](#).
- [7] E. Diaconescu and G. W. Moore, “Crossing the wall: Branes versus bundles,” *Adv. Theor. Math. Phys.* **14** (2010), no. 6, 1621–1650, [0706.3193](#).
- [8] J. Manschot, B. Pioline, and A. Sen, “Wall Crossing from Boltzmann Black Hole Halos,” *JHEP* **1107** (2011) 059, [1011.1258](#).
- [9] E. Andriyash, F. Denef, D. L. Jafferis, and G. W. Moore, “Wall-crossing from supersymmetric galaxies,” *JHEP* **1201** (2012) 115, [1008.0030](#).
- [10] R. Dijkgraaf, H. L. Verlinde, and E. P. Verlinde, “Counting dyons in  $N = 4$  string theory,” *Nucl. Phys.* **B484** (1997) 543–561, [hep-th/9607026](#).
- [11] M. C. N. Cheng and E. Verlinde, “Dying Dyons Don’t Count,” *JHEP* **09** (2007) 070, [0706.2363](#).
- [12] S. Banerjee, A. Sen, and Y. K. Srivastava, “Genus Two Surface and Quarter BPS Dyons: The Contour Prescription,” *JHEP* **03** (2009) 151, [0808.1746](#).
- [13] A. Dabholkar, S. Murthy, and D. Zagier, “Quantum Black Holes, Wall Crossing, and Mock Modular Forms,” [1208.4074](#).
- [14] J. de Boer, F. Denef, S. El-Showk, I. Messamah, and D. Van den Bleeken, “Black hole bound states in  $AdS_3 \times S^2$ ,” *JHEP* **0811** (2008) 050, [0802.2257](#).
- [15] E. Andriyash and G. W. Moore, “Ample D4-D2-D0 Decay,” [0806.4960](#).
- [16] J. Manschot, “Stability and duality in N=2 supergravity,” *Commun.Math.Phys.* **299** (2010) 651–676, [0906.1767](#).
- [17] S. Alexandrov, J. Manschot, and B. Pioline, “D3-instantons, Mock Theta Series and Twistors,” *JHEP* **1304** (2013) 002, [1207.1109](#).
- [18] F. Denef, B. R. Greene, and M. Raugas, “Split attractor flows and the spectrum of BPS D-branes on the quintic,” *JHEP* **05** (2001) 012, [hep-th/0101135](#).

- [19] S. Alexandrov and B. Pioline, “Attractor flow trees, BPS indices and quivers,” 1804.06928.
- [20] S. Alexandrov, S. Banerjee, J. Manschot, and B. Pioline, “Multiple D3-instantons and mock modular forms I,” *Commun. Math. Phys.* **353** (2017), no. 1, 379–411, 1605.05945.
- [21] S. Alexandrov, S. Banerjee, J. Manschot, and B. Pioline, “Multiple D3-instantons and mock modular forms II,” *Commun. Math. Phys.* **359** (2018), no. 1, 297–346, 1702.05497.
- [22] S. Zwegers, “Mock theta functions.” PhD dissertation, Utrecht University, 2002.
- [23] S. Alexandrov, S. Banerjee, J. Manschot, and B. Pioline, “Indefinite theta series and generalized error functions,” 1606.05495.
- [24] C. Nazaroglu, “ $r$ -Tuple Error Functions and Indefinite Theta Series of Higher-Depth,” 1609.01224.
- [25] D. Zagier, “Ramanujan’s mock theta functions and their applications (after Zwegers and Ono-Bringmann),” *Astérisque* (2009), no. 326, Exp. No. 986, vii–viii, 143–164 (2010). Séminaire Bourbaki. Vol. 2007/2008.
- [26] S. Alexandrov, B. Pioline, F. Saueressig, and S. Vandoren, “D-instantons and twistors,” *JHEP* **03** (2009) 044, 0812.4219.
- [27] C. LeBrun, “Fano manifolds, contact structures, and quaternionic geometry,” *Internat. J. Math.* **6** (1995), no. 3, 419–437, dg-ga/9409001.
- [28] S. Alexandrov, B. Pioline, F. Saueressig, and S. Vandoren, “Linear perturbations of quaternionic metrics,” *Commun. Math. Phys.* **296** (2010) 353–403, 0810.1675.
- [29] S. Alexandrov, “D-instantons and twistors: some exact results,” *J. Phys.* **A42** (2009) 335402, 0902.2761.
- [30] D. Gaiotto, G. W. Moore, and A. Neitzke, “Four-dimensional wall-crossing via three-dimensional field theory,” *Commun. Math. Phys.* **299** (2010) 163–224, 0807.4723.
- [31] S. Alexandrov, G. W. Moore, A. Neitzke, and B. Pioline, “ $\mathbb{R}^3$  Index for Four-Dimensional  $N = 2$  Field Theories,” *Phys. Rev. Lett.* **114** (2015) 121601, 1406.2360.
- [32] M.-F. Vignéras, “Séries thêta des formes quadratiques indéfinies,” *Springer Lecture Notes* **627** (1977) 227 – 239.
- [33] K. Bringmann, J. Kaszian, and A. Milas, “Higher depth quantum modular forms, multiple Eichler integrals, and  $\mathfrak{sl}_3$  false theta functions,” *ArXiv e-prints* (Apr., 2017) 1704.06891.
- [34] D. Kreimer, “Chen’s iterated integral represents the operator product expansion,” hep-th/9901099.
- [35] M. Kontsevich and Y. Soibelman, “Stability structures, motivic Donaldson-Thomas invariants and cluster transformations,” 0811.2435.
- [36] D. Joyce and Y. Song, “A theory of generalized Donaldson-Thomas invariants,” *Memoirs of the Am. Math. Soc.* **217** (2012), no. 1020, 0810.5645.

- [37] D. Joyce, “Generalized Donaldson-Thomas invariants,” *Surveys in differential geometry* **16** (2011), no. 1, 125–160, 0910.0105.
- [38] J. Manschot, B. Pioline, and A. Sen, “A Fixed point formula for the index of multi-centered  $N=2$  black holes,” *JHEP* **1105** (2011) 057, 1103.1887.
- [39] J. Manschot, B. Pioline, and A. Sen, “From Black Holes to Quivers,” *JHEP* **1211** (2012) 023, 1207.2230.
- [40] J. Manschot, B. Pioline, and A. Sen, “On the Coulomb and Higgs branch formulae for multi-centered black holes and quiver invariants,” *JHEP* **05** (2013) 166, 1302.5498.
- [41] J. Manschot, B. Pioline, and A. Sen, “The Coulomb Branch Formula for Quiver Moduli Spaces,” *Confluentes Mathematici* **2** (2017) 49–69, 1404.7154.
- [42] I. Bena, M. Berkooz, J. de Boer, S. El-Showk, and D. Van den Bleeken, “Scaling BPS Solutions and pure-Higgs States,” *JHEP* **1211** (2012) 171, 1205.5023.
- [43] S. Ferrara, R. Kallosh, and A. Strominger, “ $N = 2$  extremal black holes,” *Phys. Rev.* **D52** (1995) 5412–5416, hep-th/9508072.
- [44] J. Manschot, “Wall-crossing of D4-branes using flow trees,” *Adv.Theor.Math.Phys.* **15** (2011) 1–42, 1003.1570.
- [45] S. Alexandrov, D. Persson, and B. Pioline, “Fivebrane instantons, topological wave functions and hypermultiplet moduli spaces,” *JHEP* **1103** (2011) 111, 1010.5792.
- [46] S. Cecotti, S. Ferrara, and L. Girardello, “Geometry of type II superstrings and the moduli of superconformal field theories,” *Int. J. Mod. Phys.* **A4** (1989) 2475.
- [47] S. Ferrara and S. Sabharwal, “Quaternionic manifolds for type II superstring vacua of Calabi-Yau spaces,” *Nucl. Phys.* **B332** (1990) 317.
- [48] S. Alexandrov and S. Banerjee, “Hypermultiplet metric and D-instantons,” *JHEP* **02** (2015) 176, 1412.8182.
- [49] S. Alexandrov, S. Banerjee, and P. Longhi, “Rigid limit for hypermultiplets and five-dimensional gauge theories,” *JHEP* **01** (2018) 156, 1710.10665.
- [50] S. Alexandrov, “Twistor Approach to String Compactifications: a Review,” *Phys.Rept.* **522** (2013) 1–57, 1111.2892.
- [51] S. Alexandrov, J. Manschot, D. Persson, and B. Pioline, “Quantum hypermultiplet moduli spaces in  $N=2$  string vacua: a review,” in *Proceedings, String-Math 2012, Bonn, Germany, July 16-21, 2012*, pp. 181–212. 2013. 1304.0766.
- [52] J. Stoppa, “Joyce-Song wall crossing as an asymptotic expansion,” *Kyoto J. Math.* **54** (2014), no. 1, 103–156, 1112.2174.
- [53] R. Böhm, H. Günther, C. Herrmann, and J. Louis, “Compactification of type IIB string theory on Calabi-Yau threefolds,” *Nucl. Phys.* **B569** (2000) 229–246, hep-th/9908007.
- [54] S. Alexandrov and S. Banerjee, “Modularity, quaternion-Kähler spaces, and mirror symmetry,” *J.Math.Phys.* **54** (2013) 102301, 1306.1837.



- [55] S. Alexandrov and F. Saueressig, “Quantum mirror symmetry and twistors,” *JHEP* **09** (2009) 108, 0906.3743.
- [56] S. Alexandrov and B. Pioline, “S-duality in Twistor Space,” *JHEP* **1208** (2012) 112, 1206.1341.
- [57] S. Alexandrov and S. Banerjee, “Fivebrane instantons in Calabi-Yau compactifications,” *Phys.Rev.* **D90** (2014) 041902, 1403.1265.
- [58] J. Manschot, “Vafa-Witten theory and iterated integrals of modular forms,” 1709.10098.
- [59] K. Bringmann, J. Kaszian, and A. Milas, “Vector-valued higher depth quantum modular forms and higher mordell integrals,” 1803.06261.
- [60] K. Bringmann and C. Nazaroglu, “An exact formula for U(3) Vafa-Witten invariants on  $\mathbb{P}^2$ ,” 1803.09270.
- [61] R. K. Gupta, S. Murthy, and C. Nazaroglu, “Squashed Toric Manifolds and Higher Depth Mock Modular Forms,” 1808.00012.
- [62] S. Cecotti, A. Neitzke, and C. Vafa, “Twistorial topological strings and a tt\* geometry for  $\mathcal{N} = 2$  theories in 4d,” *Adv. Theor. Math. Phys.* **20** (2016) 193–312, 1412.4793.
- [63] K. Bringmann and J. Manschot, “From sheaves on  $\mathbb{P}^2$  to a generalization of the Rademacher expansion,” *Am. J. of Math.* **135** (2013) 1039–1065, 1006.0915.
- [64] J. Manschot, “Sheaves on  $\mathbb{P}^2$  and generalized Appell functions,” 1407.7785.
- [65] D. E. Knuth, *The Art of Computer Programming, Vol. 3: Sorting and Searching*. Addison-Wesley, 1973.